United for Change: Deliberative Coalition Formation to Change the Status Quo

Edith Elkind[1], Davide Grossi[2], Ehud Shapiro[3], Nimrod Talmon[4]
[1]University of Oxford
[2]University of Groningen and University of Amsterdam
[3]Weizmann Institute of Science
[4]Ben-Gurion University

Abstract

We study a setting in which a community wishes to identify a strongly supported proposal from a large space of alternatives, in order to change the status quo. We describe a deliberation process in which agents dynamically form coalitions around proposals that they prefer over the status quo. We formulate conditions on the space of proposals and on the ways in which coalitions are formed that guarantee deliberation to succeed, that is, to terminate by identifying a proposal with the largest possible support. Our results provide theoretical foundations for the analysis of deliberative processes in systems for democratic deliberation support, such as, e.g., LiquidFeedback or Polis.

1 Introduction

Democratic decision-making requires equality in voting on alternatives, and social choice theory provides us with a plethora of useful tools for preference aggregation (Zwicker 2016). However, another important dimension, which has received considerably less attention in the literature, is equality in deciding on which alternatives to vote upon. In practice, agents engaging in group decisions do not just vote on an externally determined set of choices: rather, they make proposals, deliberate over them, and join coalitions to push their proposals through. Understanding these processes is critical for interactive democracy applications (Brill 2018), which provide support for online voting and deliberation to self-governing communities of users that need to make democratic decisions; relevant examples are the LiquidFeedback[1] (Behrens et al. 2014) and the Polis[2] platforms.

In this paper, we aim to provide a plausible model of deliberative processes in self-governed online systems. We abstract away from the communication mechanisms by means of which deliberation may be concretely implemented, and focus instead on the coalitional effects of deliberation, that is, on how coalitions in support of various proposals may be formed or broken in the face of new suggestions. In our model, agents and alternatives are located in a metric space, and there is one distinguished alternative, which we refer to as the *status quo*. We assume that the number of alternatives is large (possibly infinite), so that the agents cannot be expected to rank the alternatives or even list all alternatives they prefer to the status quo. Rather, during the deliberative process, some agents formulate new proposals, and then each agent can decide whether she prefers a given proposal p to the status quo, i.e., whether p is closer to her location than the status quo alternative; if so, she may join a coalition of users supporting p. Importantly, such coalitions are dynamic: a user may move to another coalition if she thinks that its proposal is more likely to attract a large number of supporters; or, two coalitions may merge, possibly leaving some members behind. We assume that agents are driven by a desire to get behind a winning proposal; thus, they may prefer a larger coalition with a less appealing proposal to a smaller coalition with a more appealing proposal, as long as the former proposal is more attractive to them than the status quo. The deliberation succeeds if it identifies a most popular alternative to the status quo.

This process is democratic in that each participant who is capable of formulating a new proposal is welcome to do so; furthermore, participants who do not have the time or sophistication to work out a proposal can still take part in the deliberation by choosing which coalition to join. Also, participants could be assisted by AI-enabled tools, to help them search for a suitable compromise proposal.

To flesh out the broad outline of the deliberative coalition formation process described above, we specify rules that govern the dynamics of coalition formation. We explore several such rules ranging from single-agent moves to more complex transitions where two coalitions merge behind a new proposal, possibly leaving some dissenting members behind. Our aim is to understand whether the agents can succeed at identifying credible alternatives to the status quo if they conduct deliberation in a certain way.

We concentrate on how the properties of the underlying abstract space of proposals and the coalition formation operators available to the agents affect the guaranteed success of the deliberative coalition formation process. We show that, as the complexity of the proposal space increases, more complex forms of coalition formation are required in order to guarantee success. Intuitively, this seems to suggest that complex deliberative spaces require more sophisticated coalition formation abilities on the side of the agents.

[1]https://liquidfeedback.org
[2]https://pol.is/home

Specifically, we study four ways in which agents can form coalitions: (1) by *deviating*, when a single agent moves to another coalition; (2) by *following*, when a coalition joins another coalition in supporting its proposal; (3) by *merging*, when two coalitions join forces behind a new proposal; (4) by *compromising*, when agents belonging to two coalitions form a new larger coalition, possibly leaving dissenting agents behind. We refer to these types of coalition formation operations as *transitions*.

We show that, for each class of transitions, the deliberation process is guaranteed to terminate; in fact, for single-agent, follow and merge transitions the number of steps until convergence is at most polynomial in the number of agents. Furthermore, follow transitions are sufficient for deliberation to succeed if the set of possible proposals is a subset of the 1-dimensional Euclidean space, but this is no longer the case in two or more dimensions; in contrast, compromise transitions are sufficient in sufficiently rich subsets of $\mathbb{R}^d$ for each $d \geq 1$. However, the 'richness' is essential: we provide an example where the space of proposals is a finite subset of $\mathbb{R}^3$, but compromise transitions are not capable of identifying most popular proposals.

We view our work as an important step towards modeling a form of pre-vote deliberation usually not studied within the social choice literature: how voters can identify proposals with large support. Such a theory provides formal foundations for the design and development of practical systems that can support egalitarian deliberative decision-making, for instance by helping agents to identify mutually beneficial compromise positions (di Fenizio and Velikanov 2016).

Related Work Many social choice settings are naturally embedded in a metric space, and there is a large literature that considers preference aggregation and coalition formation in such scenarios (Coombs 1964; Merrill and Grofman 1999; de Vries 1999). We mention in particular the work of Shahaf et al. (2019), which considers an egalitarian framework for voting and proposing, instantiated to several metric spaces, modeling a broad range of social choice settings, and anticipating their use for deliberative decision making.

Group deliberation has been analyzed by several disciplines, from economics to social choice and artificial intelligence. Deliberation via sequential interaction in small groups has been studied as a means to implement large-scale decision-making when alternatives form a median graph (Goel and Lee 2016; Fain et al. 2017). Our work is not based on interaction in small groups, and focuses on how widely supported proposals can be identified in a decentralized way.

In social choice theory and economics, papers have developed axiomatic (List 2011), experimental (List et al. 2013), as well as game-theoretic approaches to deliberation. The latter have focused, for instance, on persuasion (Glazer and Rubinstein 2001, 2004, 2006) or on the way in which deliberation interacts with specific voting rules in determining group decisions (Austen-Smith and Feddersen 2005). The interaction between deliberation and voting has also been recently investigated by Karanikolas, Bisquert, and Kaklamanis (2019) using tools from argumentation theory. We, however, abstract from the concrete interaction mechanism by which agents discuss the proposals themselves, as well as on how deliberation might interact with specific voting rules, and concentrate on the results of such interactions, as they are manifested by changes in the structure of coalitions supporting different proposals.

The fact that we abstract from strategic issues that agents may be confronted with in deciding to join or leave coalitions differentiates our work also from related literature on dynamic coalition formation (Dieckmann and Schwalbe 2002; Chalkiadakis and Boutilier 2008).

Finally, unlike influential opinion dynamics models (e.g., (De Groot 1974)), in our model deliberation is driven by compromise rather than by influence or trust.

2 Formal Model

We view deliberation as a process in which agents aim to find an alternative preferred over the status quo by as many agents as possible. Thus, we assume a (possibly infinite) domain X of *alternatives*, or *proposals*, which includes the *status quo*, or *reality*, $r \in X$. We also assume a set $V = \{v_1, \ldots, v_n\}$ of n *agents*. For each proposal $x \in X$, an agent v is able to articulate whether she (strictly) prefers x over the status quo r (denoted as $x >_v r$); when $x >_v r$ we say that v *approves* x. For each $v \in V$, let $X^v = \{x \in X : x >_v r\}$; the set X^v is the *approval set* of v. Conversely, given a subset of agents $C \subseteq V$ and a proposal $p \in X$, let $C^p := \{v \in C : p >_v r\}$; the agents in C^p are the *approvers* of p in C.

Throughout this paper, we focus on the setting where X and V are contained in a metric space (M, ρ), i.e., (1) $X, V \subseteq M$, (2) for every $x \in X$ and every $v \in V$ we have $x >_v r$ if and only if $\rho(v, x) < \rho(v, r)$, and (3) the mapping $\rho : M \times M \to \mathbb{R}^+ \cup \{0\}$ satisfies (i) $\rho(x, y) = 0$ if and only if $x = y$, (ii) $\rho(x, y) = \rho(y, x)$, and (iii) $\rho(x, y) + \rho(y, z) \leq \rho(x, z)$ for all $x, y, z \in M$. E.g., if $M = \mathbb{R}^2$ and the metric is the usual Euclidean metric in $\mathbb{R}^2$, then the approval set of v consists of all points in X that are located inside the circle with center v and radius $\rho(v, r)$ (see Figure 1), whereas the set of supporters of a proposal p in C consists of all agents $v \in C$ such that $\rho(v, p) < \rho(v, r)$ (geometrically, consider the line ℓ that passes through the midpoint of the segment $[p, r]$ and is orthogonal to it; then C^p consists of all agents $v \in C$ such that v and p lie on the same side of ℓ).

Thus, an instance of our problem can be succinctly encoded by a 4-tuple (X, V, r, ρ); we will refer to such tuples as *deliberation spaces*.

Remark 1. Note that we do not require that $X = M$. By allowing X to be a proper subset of M we can capture the case where the space of proposals is, e.g., a finite subset of $\mathbb{R}^d$ for some $d \geq 1$. Moreover, in our model it need not be the case that $V \subseteq X$, i.e., we do not assume that for each agent there exists a 'perfect' proposal. Furthermore, while for each agent v the quantities $\rho(v, x)$ are well-defined for each $x \in X$, we do not expect the agents to compare different proposals based on distance; rather, the distance only determines which proposals are viewed as acceptable (i.e., preferred to the status quo).

Agents proceed by forming coalitions around proposals.

Thus, at each point in the deliberation, agents can be partitioned into coalitions, so that each coalition C is identified with a proposal p_C and all agents in C support p_C. Agents may then move from one coalition to another as well as select a proposal from X that is not associated with any of the existing coalitions and form a new coalition around it. We consider a variety of permissible moves, ranging from single-agent transitions (when an agent abandons her current coalition and joins a new one), to complex transitions that may involve agents from multiple coalitions and a new proposal. In each case, we assume that agent v is unwilling to join a coalition if this coalition advocates a proposal p such that v prefers the status quo r to p. To reason about the coalition formation process, we introduce additional notation and terminology.

Definition 1 (Deliberative Coalition)**.** A *deliberative coalition* is a pair $\mathbf{d} = (C, p)$, where $C \subseteq V$ is a set of agents, $p \in X$ is a proposal, and either (i) $p = r$ and $x \not>_v r$ for all $v \in C, x \in X \setminus \{r\}$, or (ii) $p \neq r$ and $p >_v r$ for all $v \in C$. We refer to p as the *supported proposal* of $\mathbf{d}$. The set of all deliberative coalitions is denoted by $\mathcal{D}$.

When convenient, we identify a coalition $\mathbf{d} = (C, p)$ with its set of agents C and write $\mathbf{d}^p := C^p$, $|\mathbf{d}| := |C|$.

Remark 2. While we allow coalitions that support the status quo, we require that such coalitions consist of agents who weakly prefer the status quo to all other proposals in X. We discuss a relaxation of this constraint in Section 7.

A partition of the agents into deliberative coalitions is called a *deliberative coalition structure*.

Definition 2 (Deliberative Coalition Structure)**.** A *deliberative coalition structure* (*coalition structure* for short) is a set $\mathbf{D} = \{\mathbf{d}_1, \ldots, \mathbf{d}_m\}$, $m \geq 1$, such that:

- $\mathbf{d}_i = (C_i, p_i) \in \mathcal{D}$ for each $i \in [m]$;
- $\bigcup_{i \in [m]} C_i = V$, $C_i \cap C_j = \emptyset$ for all $i, j \in [m]$ with $i \neq j$.

The set of all deliberative coalition structures over V and X is denoted by $\mathfrak{D}$.

Note that a deliberative coalition structure may contain several coalitions supporting the same proposal; also, for technical reasons we allow empty deliberative coalitions, i.e., coalitions (C, p) with $C = \emptyset$.

Example 1. *Consider a set of agents $V = \{v_1, v_2, v_3\}$ and a space of proposals $X = \{a, b, c, d, r\}$, where r is the status quo. Suppose that $X^{v_1} = \{a, b\}$, $X^{v_2} = \{b, c\}$, and $X^{v_3} = \{b, c, d\}$. Then for $C = \{v_1, v_2\}$ we have $C^a = \{v_1\}$ and $C^b = \{v_1, v_2\}$. Furthermore, let $C_1 = \{v_1, v_2\}$, $C_2 = \{v_3\}$, and let $\mathbf{d}_1 = (C_1, b)$, $\mathbf{d}_2 = (C_2, c)$. Then $\mathbf{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ is a deliberative coalition structure; see Figure 1 for an illustration.*

Remark 3. An important feature of our model is that agents do not need to explicitly list all proposals they approve of, or reason about them. At any time during the deliberation, each agent supports a single proposal in her approval set. Thus, this model can be used even if the agents have limited capability to reason about the available proposals.

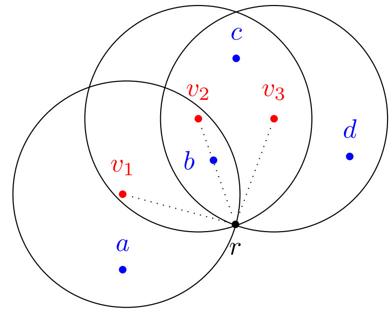

Figure 1: A deliberation space with $X = \{a, b, c, d, r\}$. The circle with center v_i, $i \in [3]$, contains all proposals approved by v_i.

2.1 Deliberative Transition Systems

As suggested above, we model deliberation as a process whereby deliberative coalition structures change as their constituent coalitions evolve. We provide general definitions for deliberative processes, modeled as transition systems (Definitions 3–6). Subsequently, we explore several specific kinds of transitions. Generally, a *transition system* is characterized by a set of *states* S, a subset $S_0 \subseteq S$ of *initial states*, and a set of *transitions* T, where each transition $t \in T$ is represented by a pair of states $(s, s') \in S \times S$; we write $t = (s, s')$ and $s \xrightarrow{t} s'$ interchangeably. We use $s \xrightarrow{T} s'$ to denote that $s \xrightarrow{t} s'$ for some $t \in T$. A *run* of a transition system is a (finite or infinite) sequence $s_0 \xrightarrow{T} s_1 \xrightarrow{T} s_2 \cdots$; such a run is *initialized* if $s_0 \in S_0$. The last state of a finite run is called its *terminal state*.

Definition 3 (Deliberative Transition System)**.** A *deliberative transition system* over a set of proposals X and a set of agents V is a transition system that has $\mathfrak{D}$ as its set of states, a subset of states $\mathfrak{D}_0 \subseteq \mathfrak{D}$ as its set of initial states, and a set of transitions $\mathcal{S}$.

Deliberative transition systems can be used to analyze the dynamics of the deliberation process; we focus on maximal runs over a deliberative transition system, as we are interested in the results of such a process.

Definition 4 (Deliberation)**.** A *deliberation* is a maximal run of a deliberative transition system, that is, a run that does not occur as a prefix of any other run (i.e., cannot be extended).

A *successful* deliberation is one that identifies some of the most popular proposals in $X \setminus \{r\}$. In particular, if there is a majority-supported proposal, a successful deliberation process allows the agent population to identify some such proposal; this can then result in a majority-supported change to the status quo.

Definition 5 (Most-Supported Alternatives)**.** For a set of agents V and a set of proposals X, the set of *most-supported alternatives* is $M^* = \text{argmax}_{x \in X \setminus \{r\}} |V^x|$ and the *maximum support* is $m^* = \max_{x \in X \setminus \{r\}} |V^x|$.

Note that, as long as all approval sets are non-empty, $M^* \neq \emptyset$ and $|V^x| = m^*$ for every $x \in M^*$. We are now ready to define success in deliberation.

Definition 6. A coalition $\mathbf{d} = (C, p)$ is *successful* if $p \in M^*$ and $|C| = m^*$. A coalition structure is *successful* if it contains a successful coalition, and *unsuccessful* otherwise. A deliberation is *successful* if it is finite and its terminal coalition structure is successful.

Intuitively, a coalition is successful if it consists of all approvers of a most-supported proposal. Successful deliberations are maximal finite runs that lead to a successful coalition structure, i.e., to a coalition structure that contains some successful coalition.

Example 2. *Consider a toy example of a deliberative transition system that consists of (1) a single initial coalition structure* $\mathbf{D}_0 = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$*, with* $\mathbf{d}_1 = (C_1, a)$ *where* $C_1 = \{v_1, v_2\}$*,* $\mathbf{d}_2 = (C_2, b)$ *where* $C_2 = \{v_3, v_4\}$*, and* $\mathbf{d}_3 = (C_3, c)$ *where* $C_3 = \{v_5, v_6\}$*; and (2) the set of transitions* $\mathcal{S}$ *only containing a transition* t *from* $\mathbf{D}_0$ *to* $\mathbf{D}_1 = \{\mathbf{d}_4, \mathbf{d}_5\}$*, with* $\mathbf{d}_4 = (C_4, e)$ *where* $C_4 = \{v_1, v_2, v_3, v_4, v_5\}$*,* $\mathbf{d}_5 = (C_5, f)$ *where* $C_5 = \{v_6\}$*. Then, assuming that there is no proposal approved by all agents, we have that* $\mathbf{D}_0 \xrightarrow{t} \mathbf{D}_1$ *is a successful deliberation.*

Intuitively, the abstract transition system described above models the dynamics of deliberation that might occur. As we want to model specific possibilities available to agents participating in such deliberation processes, below we consider specific types of transitions; each can be thought of as a *deliberation operator* that might be available to the agents. For each transition type, we aim to determine if a deliberation that only uses such transitions is guaranteed to terminate, and, if so, whether the final coalition structure is guaranteed to be successful. We show that the answer to this question depends on the underlying metric space: simple transition rules guarantee success in simple metric spaces, but may fail in richer spaces.

3 Single-Agent Transitions

The simplest kind of transition we consider is a deviation by a single agent. As we assume that agents aim to form a successful coalition and are not necessarily able to distinguish among approved proposals, it is natural to focus on transitions where an agent moves from a smaller group to a larger group; of course, this move is only possible if the agent approves the proposal supported by the larger group.

Definition 7 (Single-Agent Transition)**.** A pair of coalition structures $(\mathbf{D}, \mathbf{D}')$ forms a *single-agent transition* if there exist coalitions $\mathbf{d}_1, \mathbf{d}_2 \in \mathbf{D}$ and $\mathbf{d}'_1, \mathbf{d}'_2 \in \mathbf{D}'$ such that $|\mathbf{d}_2| \geq |\mathbf{d}_1|$, $\mathbf{D} \setminus \{\mathbf{d}_1, \mathbf{d}_2\} = \mathbf{D}' \setminus \{\mathbf{d}'_1, \mathbf{d}'_2\}$, and there exists an agent $v \in \mathbf{d}_1$ such that $\mathbf{d}'_1 = \mathbf{d}_1 \setminus \{v\}$, and $\mathbf{d}'_2 = \mathbf{d}_2 \cup \{v\}$.[3] We refer to v as the *deviating agent*.

Since $\mathbf{D}'$ is a deliberative coalition structure, agent v must approve the proposal supported by $\mathbf{d}_2$. As a consequence, no agent can deviate from a coalition that supports r to a coalition that supports some $p \in X \setminus \{r\}$ or vice versa.

Next, we show that a sequence of single-agent transitions necessarily terminates after polynomially many steps.

[3]Note that $\mathbf{d}'_1$ may be empty; we allow such 'trivial' coalitions as it simplifies our definitions.

Proposition 1. *A deliberation that consists of single-agent transitions can have at most* n^2 *transitions.*

Proof. Given a coalition structure $\mathbf{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_m\}$ such that $\mathbf{d}_i = (C_i, p_i)$ for each $i \in [m]$, let $\lambda(\mathbf{D}) = \sum_{i \in [m]} |C_i|^2$; we will refer to $\lambda(\mathbf{D})$ as the *potential* of $\mathbf{D}$. Consider a single-agent transition where an agent moves from a coalition of size x to a coalition of size y; note that $1 \leq x \leq y$. This move changes the potential by $(y+1)^2 + (x-1)^2 - y^2 - x^2 = 2 + 2(y - x) \geq 2$. Now, we claim that for every deliberative coalition structure $\mathbf{D}$ over n agents we have $\lambda(\mathbf{D}) \leq n^2$. Indeed, for the coalition structure $\mathbf{D}_0$ where all agents are in one coalition we have $\lambda(\mathbf{D}_0) = n^2$. On the other hand, if a coalition structure contains two non-empty coalitions $\mathbf{d}_1, \mathbf{d}_2$ with $|\mathbf{d}_1| \leq |\mathbf{d}_2|$, then the calculation above shows that we can increase the potential by moving one agent from $\mathbf{d}_1$ to $\mathbf{d}_2$. Thus, $\lambda(\mathbf{D}_0) \geq \lambda(\mathbf{D})$ for each $\mathbf{D} \in \mathfrak{D}$. As every single-agent transition increases the potential by at least 2, and the potential takes values in $\{1, \dots, n^2\}$, the claim follows. $\square$

However, a deliberation consisting of single-agent transitions is not necessarily successful, even for very simple metric spaces. The next example shows that such a deliberation may fail to identify a majority-supported proposal even if the associated metric space is the 1D Euclidean space.

Example 3. *Suppose that* $X, V \subseteq \mathbb{R}$*, and* $X = \{r, a, b, c\}$ *with* $r = 0$*,* $a = 1$*,* $b = 5$*,* $c = -1$*. There are three agents* v_1, v_2, v_3 *located at* a*, four agents* v_4, v_5, v_6, v_7 *located at* b*, and three agents* v_8, v_9, v_{10} *located at* c*. Observe that a majority of the agents prefer* a *to* r*. Consider the deliberative coalition structure* $\{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ *with* $\mathbf{d}_1 = (C_1, a)$*,* $\mathbf{d}_2 = (C_2, b)$*,* $\mathbf{d}_3 = (C_3, c)$*, and such that* $C_1 = \{v_1, v_2, v_3\}$*,* $C_2 = \{v_4, v_5, v_6, v_7\}$*, and* $C_3 = \{v_8, v_9, v_{10}\}$*. There are no single-agent deviations from this coalition structure: in particular, the agents in* $\mathbf{d}_2$ *do not want to deviate to* $\mathbf{d}_1$ *because* $|\mathbf{d}_1| < |\mathbf{d}_2|$ *and the agents in* $\mathbf{d}_1$ *do not want to deviate to* $\mathbf{d}_2$*, because they do not approve* b*.*

Note that this argument still applies to any proposal space X' *with* $X \subseteq X'$*; e.g., we can take* $X' = \mathbb{R}$*.*

Thus, to ensure success, we need to consider more powerful transitions.

4 Follow Transitions

A straightforward generalization of single-agent transitions are transitions where an entire coalition simply joins another coalition in supporting its current proposal.

Definition 8 (Follow Transition)**.** A pair of coalition structures $(\mathbf{D}, \mathbf{D}')$ forms a *follow transition* if there exist non-empty coalitions $\mathbf{d}_1, \mathbf{d}_2 \in \mathbf{D}$ and $\mathbf{d}'_2 \in \mathbf{D}'$ such that $\mathbf{d}_1 = (C_1, p_1)$, $\mathbf{d}_2 = (C_2, p_2)$, $\mathbf{D} \setminus \{\mathbf{d}_1, \mathbf{d}_2\} = \mathbf{D}' \setminus \{\mathbf{d}'_2\}$, and $\mathbf{d}'_2 = (C_1 \cup C_2, p_2)$.

Note that each follow transition reduces the number of coalitions by one, so a deliberation that consists only of follow transitions converges in at most n steps. Also, $|\mathbf{d}'_2|^2 = (|\mathbf{d}_1| + |\mathbf{d}_2|)^2 > |\mathbf{d}_1|^2 + |\mathbf{d}_2|^2$, i.e., follow transitions increase the potential function defined in the proof of Proposition 1. This implies the following bound.

Proposition 2. *A deliberation that consists of single-agent transitions and follow transitions can have at most n^2 transitions.*

However, in contrast to single-agent transitions, follow transitions are sufficient for successful deliberation in one dimension.

Theorem 1. *Consider a deliberation space (X, V, r, ρ), where $X, V \subseteq \mathbb{R}$ and $\rho(x, y) = |x - y|$. Then any deliberation that consists of follow transitions only, or of a combination of follow transitions and single-agent transitions, is successful.*

Proof. Assume for convenience that $r = 0$. Consider a deliberation that consists of single-agent transitions and follow transitions. By Proposition 2 we know that it is finite; let $\mathbf{D}$ be its terminal state, let p be some proposal in M^*, and assume without loss of generality that $p > 0$.

Suppose that $\mathbf{D}$ contains two deliberative coalitions (C_1, p_1) and (C_2, p_2) with $p_1, p_2 \in \mathbb{R}^+$; assume without loss of generality that $p_1 \leq p_2$. Note that $C_1, C_2 \subseteq \mathbb{R}^+$: all agents in $\mathbb{R}^- \cup \{0\}$ prefer r to p_1, p_2. Furthermore, every agent in C_2 approves p_1: indeed, if $v \in C_2$ does not approve p_1 then $|v - r| \leq |v - p_1|$, i.e., $0 < v \leq p_1 - v \leq p_2 - v$ and hence v does not approve p_2, a contradiction. Hence there is a follow transition in which C_2 joins C_1, a contradiction with $\mathbf{D}$ being a terminal coalition structure. Thus, $\mathbf{D}$ contains at most one coalition, say, (C^+, q^+), that supports a proposal in $\mathbb{R}^+$; by the same argument, it also contains at most one coalition, say, (C^-, q^-), that supports a proposal in $\mathbb{R}^-$ and at most one coalition, say, (C^0, r), that supports r. Since agents in $\mathbb{R}^-$ prefer r to p, we have $C^- \cap V^p = \emptyset$; also, by definition all agents in C^0 weakly prefer r to p. Hence $V^p \subseteq C^+$ and therefore $|C^+| = m^*$. □

If we modify the definition of follow transitions to require $|\mathbf{d}_2| \geq |\mathbf{d}_1|$ (i.e., if we require that the joint proposal of the new coalition is the proposal originally supported by the larger of the two coalitions — a seemingly sensible condition), then the proof of Theorem 1 no longer goes through. In fact, Example 3 illustrates that in this case the transition system may be unable to reach a successful state.

Unfortunately, Theorem 1 does not extend beyond one dimension. The following example shows that in the Euclidean plane a deliberation that only uses single-agent transitions and follow transitions is not necessarily successful.

Example 4. *Consider a space of proposals $\{a, b, p, r\}$ embedded into $\mathbb{R}^2$, where r is located at $(0, 0)$, p is located at $(0, 3)$, a is located at $(-3, 3)$, and b is located at $(3, 3)$. There are four agents v_1, v_2, v_3, v_4 located at $(-3, 3)$, $(-3, 4)$, $(3, 3)$, and $(3, 4)$, respectively. Note that all agents prefer p to r. Consider a deliberative coalition structure $\mathbf{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$, where $\mathbf{d}_1 = (\{v_1, v_2\}, a)$, $\mathbf{d}_2 = (\{v_3, v_4\}, b)$. Then no agent in $\mathbf{d}_1$ approves b, and no agent in $\mathbf{d}_2$ approves a, so there are no follow transitions and no single-agent transitions from $\mathbf{D}$.*

5 Merge Transitions

So far, we focused on transitions that did not introduce new proposals. Example 4, however, shows that new proposals may be necessary to reach success: indeed, none of the proposals supported by existing coalitions in that example had majority support. Thus, next we explore transitions that identify new proposals.

As a first step, it is natural to relax the constraint in the definition of the follow transitions that requires the new coalition to adopt the proposal of one of the two component coalitions, and, instead, allow the agents to identify a new proposal that is universally acceptable.

We do not specify how the compromise proposal p is identified. In practice, we expect that the new proposal would be put forward by one of the agents in $\mathbf{d}_1, \mathbf{d}_2$ or by an external mediator whose goal is to help the agents reach a consensus.

Definition 9 (Merge Transition)**.** A pair of coalition structures $(\mathbf{D}, \mathbf{D}')$ forms a *merge transition* if there exist non-empty coalitions $\mathbf{d}_1, \mathbf{d}_2 \in \mathbf{D}$ and $\mathbf{d}'_2 \in \mathbf{D}'$ such that $\mathbf{d}_1 = (C_1, p_1)$, $\mathbf{d}_2 = (C_2, p_2)$, $\mathbf{D} \setminus \{\mathbf{d}_1, \mathbf{d}_2\} = \mathbf{D}' \setminus \{\mathbf{d}'_2\}$, and $\mathbf{d}'_2 = (C_1 \cup C_2, p)$ for some proposal p.

One can verify that in Example 4 the agents have a merge transition to the deliberative coalition $(\{v_1, v_2, v_3, v_4\}, p)$; indeed, merge transitions are strictly more powerful than follow transitions. Moreover, our potential function argument shows that a deliberation that consists of single-agent transitions and merge transitions can have at most n^2 steps. However, the next example shows that, either on their own or combined with single-agent transitions, merge transitions are not sufficient for successful deliberation.

Example 5. *Consider a modification of Example 4, where we add an agent v_5 at $(-4, 0)$ to $\mathbf{d}_1$ and an agent v_6 at $(4, 0)$ to $\mathbf{d}_2$; let $\mathbf{D}'$ be the resulting coalition structure. Note that v_5 approves a, but does not approve p, thereby preventing a merge transition between $\mathbf{d}_1$ and $\mathbf{d}_2$ at p.*

6 Compromise Transitions

Example 5 illustrates that, to reach a successful outcome, coalitions may need to leave some of their members behind when joining forces. Next we formalize this idea.

Definition 10 (Compromise Transition)**.** A pair of coalition structures $(\mathbf{D}, \mathbf{D}')$ forms a *compromise transition* if there exist coalitions $\mathbf{d}_1, \mathbf{d}_2 \in \mathbf{D}$, $\mathbf{d}'_1, \mathbf{d}'_2, \mathbf{d}' \in \mathbf{D}'$ and a proposal p such that $\mathbf{D} \setminus \{\mathbf{d}_1, \mathbf{d}_2\} = \mathbf{D}' \setminus \{\mathbf{d}'_1, \mathbf{d}'_2, \mathbf{d}'\}$, $\mathbf{d}_1 = (C_1, p_1)$, $\mathbf{d}_2 = (C_2, p_2)$, $\mathbf{d}' = (C_1^p \cup C_2^p, p)$, $\mathbf{d}'_1 = (C_1 \setminus C_1^p, p_1)$, $\mathbf{d}_2 = (C_2 \setminus C_2^p, p_2)$, and $|C_1^p \cup C_2^p| > |C_1|, |C_2|$.

Intuitively, under a compromise transition agents in $\mathbf{d}_1$ and $\mathbf{d}_2$ identify a suitable proposal p, and then those of them who approve p move to form a coalition that supports p, leaving the rest of their old friends behind; a necessary condition for the transition is that the new coalition should be larger than both $\mathbf{d}_1$ and $\mathbf{d}_2$.

Example 6. *Consider again the deliberative coalition structure $\mathbf{D}'$ in Example 5. There is a compromise transition from $\mathbf{D}'$ to the coalition structure $(\mathbf{d}'_1, \mathbf{d}'_2, \mathbf{d}')$, where $\mathbf{d}'_1 = (\{v_5\}, a)$, $\mathbf{d}'_2 = (\{v_6\}, b)$, and $\mathbf{d}' = (\{v_1, v_2, v_3, v_4\}, p)$. As $|\mathbf{d}'| = 4$, the resulting coalition structure is successful.*

Importantly, we assume that all agents in $\mathbf{d}_1$ and $\mathbf{d}_2$ who support p join the compromise coalition; indeed, this is what we expect to happen if the agents myopically optimize the size of their coalition.

An important feature of compromise transitions is that they ensure termination.

Proposition 3. *A deliberation that consists of compromise transitions is necessarily finite.*

Proof. Given a coalition structure $\mathbf{D} = \{\mathbf{d}_1, \ldots, \mathbf{d}_m\}$ such that $\mathbf{d}_i = (C_i, p_i)$ for each $i \in [m]$, $|C_1| \geq \cdots \geq |C_m|$, and there exists an $\ell \in [m]$ such that $|C_i| > 0$ for $i \in [\ell]$, $|C_i| = 0$ for $i = \ell+1, \ldots, m$, let $\gamma(\mathbf{D}) = (|C_1|, \ldots, |C_\ell|)$. Note that $\gamma(\mathbf{D})$ is a non-increasing sequence of positive integers. Given two non-increasing sequences $(a_1, \ldots, a_s)$, $(b_1, \ldots, b_t)$ of positive integers we write $(a_1, \ldots, a_s) <_{\text{lex}} (b_1, \ldots, b_t)$ if either (a) there exists a $j \leq \min\{s, t\}$ such that $a_i = b_i$ for all $i < j$ and $a_j < b_j$, or (b) $s < t$, and $a_i = b_i$ for all $i \in [s]$. Now, observe that if $(\mathbf{D}, \mathbf{D}')$ is a compromise transition, then for the respective coalitions $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}'$ we have $|\mathbf{d}'| > \max\{|\mathbf{d}_1|, |\mathbf{d}_2|\}$, and hence $\gamma(\mathbf{D}) <_{\text{lex}} \gamma(\mathbf{D}')$. Since for every coalition structure $\mathbf{D}$ over n agents we have $\gamma(\mathbf{D}) \leq_{\text{lex}} (n)$, the claim follows. □

In contrast to the case of single-agent transitions and follow/merge transitions, we are unable to show that a deliberation consisting of compromise transitions always terminates after polynomially many steps; it remains an open problem whether this is indeed the case. We note that a compromise transition does not necessarily increase the potential function defined in the proof of Proposition 1; e.g., the transition described in Example 6 does not change the potential.

We say that a deliberation space (X, V, r, ρ) is a *Euclidean deliberation space* if $X = \mathbb{R}^d$, $V \subseteq \mathbb{R}^d$ for some $d \geq 1$, and ρ is the standard Euclidean metric on $\mathbb{R}^d$. The main result of this section is that, in every Euclidean deliberation space, every maximal run of compromise transitions is successful. To prove it, we need two auxiliary lemmas. In what follows, for a coalition structure $\mathbf{D}$, $|\mathbf{D}|$ denotes the number of non-empty coalitions in $\mathbf{D}$ that do not support r.

Lemma 1. *In every deliberation space, a deliberation that consists of compromise transitions and has a coalition structure $\mathbf{D}$ with $|\mathbf{D}| = 2$ as its terminal state is successful.*

Proof. Consider a coalition $\mathbf{D}$ with $|\mathbf{D}| = 2$ that is not successful. Let $\mathbf{d}_1$, $\mathbf{d}_2$ be the two coalitions in $\mathbf{D}$ that do not support r; we have $|\mathbf{d}_1| < m^*$, $|\mathbf{d}_2| < m^*$. For each $p \in M^*$ we have $V^p = \mathbf{d}_1^p \cup \mathbf{d}_2^p$ and hence $|\mathbf{d}_1^p \cup \mathbf{d}_2^p| = m^* > |\mathbf{d}_1|, |\mathbf{d}_2|$. Thus, there exists a compromise transition from $\mathbf{D}$ in which agents in $\mathbf{d}_1^p \cup \mathbf{d}_2^p$ form a coalition around p. □

Lemma 2. *In every Euclidean deliberation space, a deliberation that consists of compromise transitions and has a coalition structure $\mathbf{D}$ with $|\mathbf{D}| \geq 3$ as its terminal state is successful.*

Proof. For the case $d = 1$ our claim follows from the proof of Theorem 1. We will now provide a proof for $d = 2$; it generalizes straightforwardly to $d > 2$.

Fix a coalition structure $\mathbf{D}$ that contains at least three coalitions, all of which are not successful and none of which supports r; we will show that $\mathbf{D}$ is not terminal. Let $\mathbf{d}^a =$

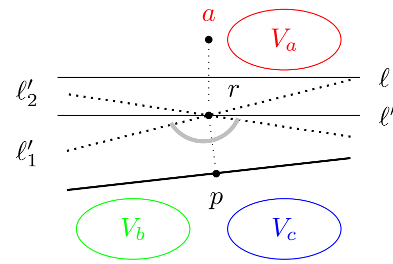

Figure 2: Proof of Lemma 2.

(C_a, a) be a maximum-size coalition in $\mathbf{D}$ that does not support r. Let ℓ be the line that passes through the middle of the a–r segment and is orthogonal to it; this line separates $\mathbb{R}^2$ into two open half-planes so that r lies in one of these half-planes while all points in $\mathbf{d}^a$ lie in the other half-plane (see Figure 2). Let ℓ' be the line that passes through r and is parallel to ℓ. For a positive α, let ℓ_1' be the line obtained by rotating ℓ' about r clockwise by α, and let ℓ_2' be the line obtained by rotating ℓ' counterclockwise by α. The line ℓ_1' (resp. ℓ_2') partitions $\mathbb{R}^2$ into open half-planes H_1 and H_1' (resp. H_2 and H_2'). We can choose α to be small enough that $\mathbf{d}^a \subset H_1, \mathbf{d}^a \subset H_2$ and so that no agent lies on ℓ_1' or on ℓ_2'. Now, if there exists a coalition $\mathbf{d}^b = (C_b, b) \in \mathbf{D}, \mathbf{d}^b \neq \mathbf{d}^a$, $b \neq r$, such that $v \in H_1$ or $v \in H_2$ for some $v \in C_b$, then r is not in the convex hull of C_a and v, and hence there is a line that separates $C_a \cup \{v\}$ from r; by projecting r onto this line, we obtain a proposal r' that is approved by v and all agents in C_a. Thus, there is a compromise transition in which a non-empty subset of agents in $C_b^{r'}$ joins C_a to form a deliberative coalition around r'.

Otherwise, $\mathbf{d}^x \subseteq H_1' \cap H_2'$ for all $\mathbf{d}^x \in S \setminus \{\mathbf{d}^a\}$. Consider two distinct coalitions $\mathbf{d}^b, \mathbf{d}^c \in \mathbf{D}$ with $b, c \neq r$. As $H_1' \cap H_2'$ is bounded by two rays that start from r, and the angle between these rays is $2\pi - 2\alpha < 2\pi$, there is a line ℓ^* that divides $\mathbb{R}^2$ into two open half-planes so that r is in one half-plane and $\mathbf{d}^b \cup \mathbf{d}^c$ are in the other half-plane; thus, all agents in $\mathbf{d}^b \cup \mathbf{d}^c$ approve the proposal p obtained by projecting r onto ℓ^*, and hence there is a merge transition. □

Combining Lemmas 1 and 2, we obtain the main result of this section.

Theorem 2. *In every Euclidean deliberation space, every maximal run of compromise transitions is successful.*

Proof. Consider a maximal run of compromise transitions, and let $\mathbf{D}$ be its terminal state. Suppose for the sake of contradiction that $\mathbf{D}$ is not successful. Note that this implies that $|\mathbf{D}| > 1$. If $|\mathbf{D}| = 2$ there is a transition from $\mathbf{D}$ by Lemma 1 and if $|\mathbf{D}| \geq 3$, then there is a transition from $\mathbf{D}$ by Lemma 2. □

For the proof of Lemma 2 to go through, the underlying metric space should be sufficiently rich: to obtain the proposal approved by the new coalition, we project the status quo r on a certain line. The argument goes through if we replace $\mathbb{R}^d$ with $\mathbb{Q}^d$; however, it does not extend to the case where X is an arbitrary finite subset of $\mathbb{R}^d$. Intuitively, for

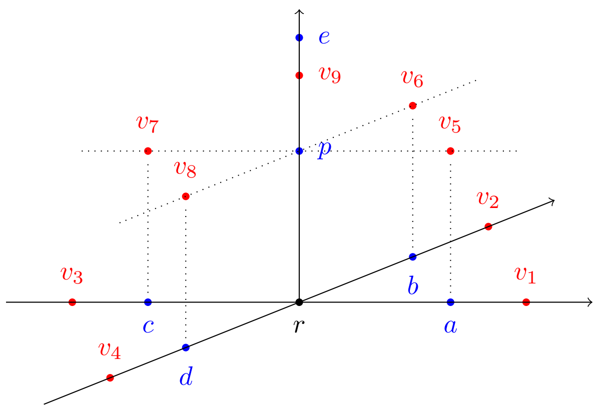

Figure 3: The deliberative space in Example 7.

deliberation to converge, at least some agents should be able to spell out nuanced compromise proposals.

Observe that the compromise transitions in Lemma 2 have a special form: when two coalitions join forces, at least one of them is fully behind the new proposal. This motivates the following definition.

Definition 11 (Subsume Transition)**.** A pair of coalition structures $(\mathbf{D}, \mathbf{D}')$ forms a *subsume transition* if there exist coalitions $\mathbf{d}_1, \mathbf{d}_2 \in \mathbf{D}$, $\mathbf{d}'_1, \mathbf{d}' \in \mathbf{D}'$ and a proposal p such that $\mathbf{D} \setminus \{\mathbf{d}_1, \mathbf{d}_2\} = \mathbf{D}' \setminus \{\mathbf{d}'_1, \mathbf{d}'\}$, $\mathbf{d}_1 = (C_1, p_1)$, $\mathbf{d}_2 = (C_2, p_2)$, $C_2^p = C_2$, $C_1^p \neq \emptyset$, $\mathbf{d}' = (C_1^p \cup C_2, p)$, $\mathbf{d}'_1 = (C_1 \setminus C_1^p, p_1)$, and $|C_1^p \cup C_2| > |C_1|$.

By construction, every subsume transition is a compromise transition, and Lemma 2 holds for subsume transitions.

While subsume transitions by themselves are not sufficient for successful deliberation in $\mathbb{R}^d$ (as the transition in Lemma 1 is not necessarily a subsume transition), they are nearly sufficient: the proof of Theorem 2 shows that we need at most one general compromise transition. Noting that every subsume transition increases the potential by $2(|C_2|-|C_1|+|C_1^p|) \geq 2$, we obtain the following corollary.

Corollary 1. *For every Euclidean deliberation space there exists a successful deliberation that consists of compromise transitions and has length at most $n^2 + 1$.*

Proof. Suppose that agents perform subsume transitions until no such transitions are available; as every subsume transition increases the potential, this process ends after at most n^2 steps. By Lemma 2, if the resulting coalition structure $\mathbf{D}$ is not successful, then $|\mathbf{D}| = 2$, in which case, by Lemma 1, there is a compromise transition from $\mathbf{D}$ to a successful coalition structure. □

Given our positive results for Euclidean deliberation spaces, it is natural to ask whether compromise transitions are sufficient for convergence in other metric spaces. The following example, however, illustrates that this is not the case even if X is a finite subset of $\mathbb{R}^d$.

Example 7. *Figure 3 depicts a deliberation space that is embedded in $\mathbb{R}^3$, so that ρ is the usual Euclidean metric in $\mathbb{R}^3$, and $X = \{r, a, b, c, d, p, e\}$, $V = \{v_1, \ldots, v_9\}$ with*
$r = (0, 0, 0)$,
$a = (2, 0, 0), b = (0, 2, 0), c = (-2, 0, 0), d = (0, -2, 0)$,
$p = (0, 0, 2), e = (0, 0, 3.5)$,
$v_1 = (3, 0, 0), v_2 = (0, 3, 0), v_3 = (-3, 0, 0), v_4 = (0, -3, 0)$,
$v_5 = (2, 0, 2), v_6 = (0, 2, 2), v_7 = (-2, 0, 2), v_8 = (0, -2, 2)$,
$v_9 = (0, 0, 3)$.
Then $\mathbf{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4, \mathbf{d}_5\}$, where $\mathbf{d}_1 = (\{v_1, v_5\}, a)$, $\mathbf{d}_2 = (\{v_2, v_6\}, b)$, $\mathbf{d}_3 = (\{v_3, v_7\}, c)$, $\mathbf{d}_4 = (\{v_4, v_8\}, d)$, $\mathbf{d}_5 = (\{v_9\}, e)$ is a deliberative coalition structure, and there are no compromise transitions from $\mathbf{D}$. However, $\mathbf{D}$ is not successful, as agents $v_5, \ldots, v_9$ all support p.

Thus, for general metric spaces one needs an even richer class of transitions to identify credible alternatives to the status quo.

7 Conclusions and Future Work

We proposed a formal model of deliberation for agent populations forming coalitions around proposals in order to change the status quo. We identified several natural types of deliberation transitions, together with metric spaces where deliberations that consist of such transitions are always successful. We intend our model as a foundation for a system allowing an online community to self govern.

Future work includes considering other transition types and identifying further deliberative spaces that guarantee success. One can also ask whether our positive results are preserved if we impose additional conditions on the structure of the deliberative process, e.g., require new proposals to be close to the original proposals. We may also revisit our approach to modeling coalitions that support the status quo: we now assume that each agent v with $X^v \neq \emptyset$ is capable of identifying some proposal in X^v, and this assumption may be too strong for many deliberation scenarios. It is perhaps more realistic to assume that some agents start out by supporting the status quo, and then learn about a new proposal p that they prefer to the status quo by observing a coalition that supports p, and then move to join this coalition. Investigating a model that permits such transitions is another direction for future work.

Further afield, it would be interesting to consider a stochastic variant of our model, in which each transition from a state is assigned a certain probability. A Markovian analysis of such systems might shed further light on additional properties of deliberation. Also, in our current model agents are not strategic: they truthfully reveal whether they support a given proposal. Yet, revealed support for proposals may well be object of manipulation. Such a game-theoretic extension is a natural direction for future research. Another ambitious direction for future work is to design a practical tool for deliberation and self-government of an online community, building on top of our model and analysis. Such a tool could take the form of an AI bot over existing on-line deliberation platforms such as the LiquidFeedback and Polis platforms mentioned earlier. The bot would suggest proposals to agents in order to support compromise, hopefully fostering successful deliberations.

Acknowledgements

We thank Schloss Dagstuhl—Leibniz Center for Informatics. The paper develops ideas discussed by the authors and other participants during the Dagstuhl Seminar 19381 (Application-Oriented Computational Social Choice), summer 2019. We thank the generous support of the Braginsky Center for the Interface between Science and the Humanities. Edith Elkind was supported by the ERC Starting Grant ACCORD (GA 639945). Nimrod Talmon was supported by the Israel Science Foundation (ISF; Grant No. 630/19).

References

Austen-Smith, D.; and Feddersen, T. J. 2005. Deliberation and Voting Rules. In *Social Choice and Strategic Decisions*, Studies in Social Choice and Welfare. Springer.

Behrens, J.; Kistner, A.; Nitsche, A.; and Swierczek, B. 2014. *Principles of Liquid Feedback*. Interaktive Demokratie.

Brill, M. 2018. Interactive Democracy. In *Proceedings of AAMAS '18*, 1183–1187.

Chalkiadakis, G.; and Boutilier, C. 2008. Sequential Decision Making in Repeated Coalition Formation under Uncertainty. In *Proceedings of AAMAS '08*, 347–354.

Coombs, C. H. 1964. *A theory of data*. Wiley.

De Groot, M. 1974. Reaching a Consensus. *Journal of the American Statistical Association* 69(345): 118–121.

de Vries, M. W. M. 1999. *Governing with your closest neighbour : an assessment of spatial coalition formation theories*. Radboud University Nijmegen.

di Fenizio, P. S.; and Velikanov, C. 2016. System-Generated Requests for Rewriting Proposals. *CoRR* abs/1611.10095. URL http://arxiv.org/abs/1611.10095.

Dieckmann, T.; and Schwalbe, U. 2002. Dynamic coalition formation and the core. *Journal of Economic Behavior & Organization* 49(3): 363–380.

Fain, B.; Goel, A.; Munagala, K.; and Sakshuwong, S. 2017. Sequential deliberation for social choice. In *Proceedings of WINE '17*, 177–190.

Glazer, J.; and Rubinstein, A. 2001. Debates and Decisions: On a Rationale of Argumentation Rules. *Games and Economic Behavior* 36(2): 158–173.

Glazer, J.; and Rubinstein, A. 2004. On Optimal Rules of Persuasion. *Econometrica* 1715–1736.

Glazer, J.; and Rubinstein, A. 2006. A Study in the Pragmatics of Persuasion: A Game Theoretical Approach. *Theoretical Economics* 1: 395–410.

Goel, A.; and Lee, D. T. 2016. Towards large-scale deliberative decision-making: Small groups and the importance of triads. In *Proceedings of EC '16*, 287–303.

Karanikolas, N.; Bisquert, P.; and Kaklamanis, C. 2019. A Voting argumentation framework: Considering the reasoning behind preferences. In *Proceedings of ICAART '19*, 42–53.

List, C. 2011. Group Communication and the Transformation of Judgments: An Impossibility Result. *The Journal of Political Philosophy* 19(1): 1–27.

List, C.; Luskin, R.; Fishkin, J.; and McLean, I. 2013. Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls. *The Journal of Politics* 75(1): 80–95.

Merrill, S.; and Grofman, B. 1999. *A unified theory of voting: Directional and proximity spatial models*. Cambridge University Press.

Shahaf, G.; Shapiro, E.; and Talmon, N. 2019. Aggregation over metric spaces: Proposing and voting in elections, budgeting, and legislation. In *Proceedings of ADT '19*.

Zwicker, W. S. 2016. Introduction to the Theory of Voting. In *Handbook of Computational Social Choice*, 23–56. Cambridge University Press.