Complexity of Deliberative Coalition Formation

Edith Elkind, Abheek Ghosh, Paul Goldberg

Department of Computer Science, University of Oxford, UK.
{edith.elkind,abheek.ghosh,paul.goldberg}@cs.ox.ac.uk

Abstract

Elkind et al. (2021, 2020) introduced a model for deliberative coalition formation, where a community wishes to identify a strongly supported proposal from a space of alternatives, in order to change the status quo. In their model, agents and proposals are points in a metric space, agents' preferences are determined by distances, and agents deliberate by dynamically forming coalitions around proposals that they prefer over the status quo. The deliberation process operates via k-*compromise* transitions, where agents from k (current) coalitions come together to form a larger coalition in order to support a (perhaps new) proposal, possibly leaving behind some of the dissenting agents from their old coalitions. A deliberation succeeds if it terminates by identifying a proposal with the largest possible support. For deliberation in d dimensions, Elkind et al. consider two variants of their model: in the *Euclidean model*, proposals and agent locations are points in $\mathbb{R}^d$ and the distance is measured according to $||\cdot||_2$; and in the *hypercube model*, proposals and agent locations are vertices of the d-dimensional hypercube and the metric is the Hamming distance. They show that in the continuous model 2-compromises are guaranteed to succeed, but in the discrete model for deliberation to succeed it may be necessary to use k-compromises with $k \geq d$. We complement their analysis by (1) proving that in both models it is hard to find a proposal with a high degree of support, and even a 2-compromise transition may be hard to compute; (2) showing that a sequence of 2-compromise transitions may be exponentially long; (3) strengthening the lower bound on the size of the compromise for the d-hypercube model from d to $2^{\Omega(d)}$.

1 Introduction

Imagine a nominally democratic country where the current ruling party has been in power for many years, through a combination of clever political strategizing and a range of more or less non-democratic means (e.g., gerrymandering, vote suppression, media control, etc.). Even if, initially, political parties and alliances other than the ruling party may differ in positions they take on various issues facing the society, after many years of being out of power, they may choose to focus on what unites them rather than on what differentiates them, and seek to identify a common platform that

would enjoy broad popular support and enable them to overturn the status quo. In doing so, they may pursue a variety of goals: to increase their chances of winning seats in an election, to place themselves in a better position in post-electoral bargaining, or—in the most extreme cases—to initiate a revolt against the current regime. This common platform may be quite different from the party's true position—what matters is that it attracts a large number of supporters and that it offers an improvement over the status quo, i.e., is preferable to the policies of the current governing party.

In a recent paper, Elkind et al. (2021, 2020) propose a simple model that aims to capture the essential features of such scenarios[1]. In this model, both voters and proposals are associated with points in a metric space, with voters' preference being driven by distances: voters prefer proposals that are close to them to ones that are further away. The number of voters is finite, but the set of feasible proposals may be any (potentially infinite) subset of the metric space. There is also a distinguished point, referred to as the *status quo* and denoted by r. A voter v approves a proposal p if her distance to p is strictly less than her distance to r. Voters deliberate in order to identify a proposal that is supported by as many voters as possible. At each point, each voter selects some approved proposal to support, with voters who support a given proposal forming a *deliberative coalition* around it. This collection of deliberative coalitions—a deliberative coalitions structure—evolves based on *transition rules*: for instance, one transition rule allows two coalitions to identify a new proposal supported by all members of both coalitions and to form a new joint coalition around it. The transition rules aim to capture the behavior of agents who are consensus-driven—i.e., they desire to form a large coalition to overturn the status quo—and myopic, in the sense that they make a decision whether to participate in a transition based on the outcome of that transition only and not the entire deliberative process.

In their work, Elkind et al. (2020) primarily focus on the power of various transition rules to enable the identification

---

[1] The conference version of their paper (Elkind et al. 2021) was published in AAAI'21, and an extended version was presented at COMSOC'21 and is available from arXiv (Elkind et al. 2020). One of the models we consider, namely, the hypercube model, is only described in the arXiv version, so in what follows we refer to the arXiv version only.

of popular proposals, in a range of metric spaces. They do not investigate the complexity of the associated algorithmic challenges, and only offer very crude bounds on the number of steps it may take the deliberation process to converge.

1.1 Our Contribution

Our goal is to complement the analysis of Elkind et al. (2020), by exploring the complexity of the deliberation process in their model. We focus on two deliberation spaces: in the *hypercube* model, voters and proposals are vertices of the d-dimensional hypercube, with distances measured according to the Hamming distance, and in the *Euclidean* model voters and proposals are elements of the d-dimensional Euclidean space, with $||\cdot||_2$ being the underlying distance measure. In both models, each point of the underlying metric space is considered to be a feasible proposal. We consider three types of questions:

- What is the computational complexity of identifying a proposal approved by as many voters as possible, both from a centralized perspective (i.e., how can we compute a popular proposal given the positions of all voters), and from a decentralized perspective (i.e., how can a group of voters identify the next step in the deliberative process)? We consider both the worst-case complexity of this problem and its parameterized complexity, with two natural parameters being the number of voters and the dimension of the space.

- How many transitions may be necessary for a deliberation to converge? Elkind et al. (2020) show that deliberation always converges after at most n^n steps, where n is the number of voters; we improve this upper bound to 2^n and derive exponential lower bounds for both of the models we consider.

- How many coalitions need to be involved in each deliberation step to ensure that a most approved proposal is identified? Elkind et al. (2020) prove that in Euclidean deliberation spaces 2-coalition deals are sufficient irrespective of dimension, and in a d-dimensional hypercube we may need transitions involving at least d coalitions; we improve this lower bound from d to $2^{\Theta(d)}$.

The work of Elkind et al. (2020) is an important step towards modeling coalition formation in the presence of a status quo option. Such a theory provides formal foundations for the design and development of practical systems that can support successful deliberation, for instance, by helping agents to identify mutually beneficial compromise positions (cf. (di Fenizio and Velikanov 2016)). Our paper supplements the analysis of Elkind et al. (2020) by resolving several challenging open questions posed by their work.

1.2 Related Work

Our work builds directly on the work of Elkind et al. (2020). In turn, their paper belongs to a rich tradition in political science that studies *spatial coalition formation* (Coombs 1964; Enelow and Hinich 1984; Merrill III, Merrill, and Grofman 1999; Vries 1999). An important feature of their model that sets it somewhat apart from prior work is the presence of a special reference point, i.e., the status quo. The agenda of social choice in the presence of status quo has recently been pursued by Shapiro, Talmon, and co-authors in a series of papers (Shapiro and Talmon 2018; Shahaf, Shapiro, and Talmon 2019; Bulteau et al. 2021; Abramowitz, Shapiro, and Talmon 2020).

The study of group deliberation is a broad and interdisciplinary area, see, e.g., (Austen-Smith and Feddersen 2005; Hafer and Landa 2007; Patty 2008; List 2011; List et al. 2013; Rad and Roy 2021; Perote-Peña and Piggins 2015; Karanikolas, Bisquert, and Kaklamanis 2019); we refer the reader to the work of Elkind et al. (2020) for further discussion. In particular, an important consideration is whether simple deliberation protocols that only involve a small number of participants can achieve a desirable outcome (Goel and Lee 2016; Fain et al. 2017); this question is similar to the one considered in Section 5 of our paper.

The process of deliberation that we consider can be viewed as one of dynamic coalition formation (Konishi and Ray 2003; Arnold and Schwalbe 2002; Chalkiadakis and Boutilier 2012). In particular, the transitions from one deliberative coalition structure to the next one can be viewed as a form of better response dynamics, where agents prefer deliberative coalitions whose proposal they approve to those whose proposal they do not approve, and, among coalitions whose proposal they approve, they favor larger coalitions. Potential functions are commonly used in the literature to establish upper bounds on convergence time of better response dynamics (see, e.g., (Tardos and Wexler 2007)); however, our use of a potential function to establish a lower bound (Section 4) is somewhat unconventional.

2 Preliminaries

Following Elkind et al. (2020), we define a *deliberation space* as a 4-tuple $(\mathcal{X}, \mathcal{V}, r, \rho)$, where $\mathcal{X}$ is a (possibly infinite) set of *proposals*, $\mathcal{V} = \{v_1, \ldots, v_n\}$ is a set of of n *agents*, $r \in \mathcal{X}$ is a special element of $\mathcal{X}$, which we will refer to as the *status quo*, and ρ is a metric on $\mathcal{X} \cup \mathcal{V}$. An agent v_i is said to *approve* a proposal $x \in \mathcal{X} \setminus \{r\}$ if $\rho(v_i, x) < \rho(v_i, r)$: intuitively, v_i approves x if x is more representative of her position than r is. It will be convenient to require that each agent approves at least one proposal, i.e., to exclude agents that are happy with the status quo.

We will consider two families of d-dimensional deliberation spaces, where d is a positive integer: Euclidean deliberation spaces and hypercube deliberation spaces. We refer to Elkind et al. (2020) for a discussion of some other deliberation spaces.

- d-**Euclidean** The space of proposals $\mathcal{X}$ is $\mathbb{R}^d$, $r = (0, 0, \ldots, 0)$, and $v_i \in \mathbb{R}^d$ for each $i \in [n]$. The metric ρ is the usual Euclidean norm: $\rho(x, y) = ||x - y||_2$.

- d-**Hypercube** The space of proposals $\mathcal{X}$ is $\{0, 1\}^d$, $r = (0, 0, \ldots, 0)$, and $v_i \in \{0, 1\}^d$ for each $i \in [n]$. The metric ρ is the Hamming distance: $\rho(x, y) = ||x - y||_1$.

At any point in the deliberation process, the set of agents is split into deliberative coalitions. A *deliberative coalition* is a pair $\mathbf{d} = (C, x)$, where C is a non-empty subset of

$\mathcal{V}$, $x \in \mathcal{X} \setminus \{r\}$, and each agent in C approves x. When convenient, we identify a deliberative coalition $\mathbf{d} = (C, x)$ with its set of agents C, and say that agents in C *support* x.

A *deliberative coalition structure* is a set $\mathbf{D} = \{\mathbf{d}_1, \ldots, \mathbf{d}_m\}$, $m \geq 1$, such that:

- $\mathbf{d}_j = (C_j, x_j)$ is a deliberative coalition for each $j \in [m]$;
- $\cup_{j \in [m]} C_j = \mathcal{V}$ and $C_j \cap C_\ell = \varnothing$ for all $j, \ell \in [m]$ such that $j \neq \ell$.

The agents start out in some deliberative coalition structure, and then this structure evolves according to *transition rules*. Elkind et al. (2020) define several such rules; in this paper, we focus on *k-compromise transitions*. The intuition behind all these rules is that agents seek to form large coalitions and act myopically, i.e., an agent in a deliberative coalition (C, x) is willing to deviate so as to end up in a coalition (C', x') if (i) she approves x' and (ii) $|C'| > |C|$. Specifically, under a k-compromise transition, agents from $\ell \leq k$ existing deliberative coalitions $\mathbf{d}_1, \ldots, \mathbf{d}_\ell$, where $\mathbf{d}_j = (C_j, x_j)$ for $j \in [\ell]$, get together and identify a proposal x such that t of them approve x and $t > |C_j|$ for each $j \in [\ell]$. Then all agents who approve x form a deliberative coalition that supports proposal x. The agents in C_j who do not approve x stay put, i.e., they end up in deliberative coalition (C'_j, x_j) with $C'_j = C_j \setminus C$ (note that C'_j may be empty). This notion is formalized as follows.

Definition 2.1 (k-Compromise Transitions). A pair of coalition structures $(\mathbf{D}, \mathbf{D}')$ forms a *k-compromise transition* if there exist ℓ coalitions $\mathbf{d}_1, \ldots, \mathbf{d}_\ell \in \mathbf{D}$, $1 \leq \ell \leq k$, where $\mathbf{d}_j = (C_j, x_j)$ for $j \in [\ell]$, such that $\mathbf{D}'$ is obtained from $\mathbf{D}$ by (1) removing $\mathbf{d}_1, \ldots, \mathbf{d}_\ell$, (2) adding a deliberative coalition (C, x) such that $C \subseteq \cup_{j \in [\ell]} C_j$ and for each $j \in [\ell]$ it holds that $|C| > |C_j|$ and $C \cap C_j$ consists of all agents in C_j that approve x; (3) for each $j \in [\ell]$ such that $C_j \nsubseteq C$ adding a deliberative coalition $(C_j \setminus C, x_j)$.

We say that a deliberative coalition structure $\mathbf{D}$ is *k-terminal* if there does not exist a k-compromise transition of the form $(\mathbf{D}, \mathbf{D}')$. A *$k$-deliberation* is a sequence of k-compromise transitions such that for the last transition $(\mathbf{D}, \mathbf{D}')$ in this sequence it holds that $\mathbf{D}'$ is terminal. We will refer to $\mathbf{D}'$ as the *outcome* of the respective k-deliberation.

We define the *score* of a proposal $x \in \mathcal{X} \setminus \{r\}$ as the number of agents in $\mathcal{V}$ who approve x. We say that a proposal $x \in \mathcal{X} \setminus \{r\}$ is *popular* if its score is at least as high as that of any other proposal in $\mathcal{X} \setminus \{r\}$. A deliberative coalition structure $\mathbf{D}$ is *successful* if it contains a deliberative coalition (C, x) such that x is a popular proposal and C consists of all agents who approve x. A k-deliberation is *successful* if its outcome is successful.

Note that if there is a majority-approved proposal, a successful k-deliberation identifies some such proposal, enabling a majority-supported change to the status quo.

An important result of Elkind et al. (2020) is that a deliberation process with k-compromise transitions always terminates.

Theorem 1. *(Elkind et al. 2020) For each integer k with $2 \leq k \leq n$ a k-deliberation can have at most n^n transitions.*

This result holds for any deliberation space, so in particular both for the d-Euclidean space and the d-hypercube for any $d \geq 1$.

3 Complexity of Finding Popular Proposals

In this section, we focus on the complexity-theoretic challenges presented by the deliberation process. We first consider two computational problems: SCORE and PERFECT SCORE. For SCORE, the input is a deliberation space and a positive integer η, and we ask if there is a proposal that is approved by at least η agents in $\mathcal{V}$; in PERFECT SCORE, we are given a deliberation space and ask if there is a proposal that is approved by all agents in $\mathcal{V}$. These problems model the challenge of finding a good proposal in a centralized way.

As PERFECT SCORE is a special case of SCORE, an NP-hardness result for PERFECT SCORE implies that SCORE, too, is NP-hard, whereas a polynomial-time algorithm for SCORE can be used to solve PERFECT SCORE in polynomial time.

We will also consider another computational problem, which captures the complexity of the decentralized deliberation process: in k-COMPROMISE, we are given a deliberative space and a deliberative coalition structure, and the goal is to find a k-compromise transition from this coalition structure if one exists; note that k-COMPROMISE is a function problem and not a decision problem.

For all three problems, we use prefixes *Euc* and *Hyp* to indicate whether we consider the variant of the problem for the Euclidean space or for the hypercube.

Given a deliberation space $I = (\mathcal{S}, \mathcal{V}, r, \rho)$, we denote by $|I|$ the number of bits in the description of I. For d-hypercube deliberation spaces, this is simply $O(nd)$ (we need to specify d bits per agent). For Euclidean deliberation spaces, we assume that the locations of all agents are vectors of rational numbers given in binary; similarly, when we consider k-COMPROMISE, we assume that the proposals supported by the coalitions in the initial deliberative coalition structure are vectors of rational numbers given in binary.

We will first consider hypercubes, and then Euclidean spaces.

3.1 Hypercube Deliberation Spaces

In hypercube deliberation spaces, even deciding whether there is a unanimously approved proposal is computationally difficult.

Theorem 2. *Hyp-PERFECT SCORE is NP-complete.*

We prove Theorem 2 by a reduction from INDEPENDENT SET (Garey and Johnson 1979); proof provided in the appendix.

Since SCORE is at least as hard as PERFECT SCORE, and it is clearly in NP, we obtain the following corollary.

Corollary 3. *Hyp-SCORE is NP-complete.*

Moreover, for hypercube deliberation spaces, we can show that computing a k-compromise is at least as hard as finding a proposal with a perfect score.

Corollary 4. *For each $k \geq 2$, there does not exist a polynomial-time algorithm for Hyp-k-COMPROMISE unless P=NP.*

Proof. Suppose we have a polynomial-time algorithm for *Hyp-k*-COMPROMISE for some $k \geq 2$. We will explain how it can be used to solve *Hyp*-PERFECT SCORE in polynomial time.

Given an instance I of *Hyp*-PERFECT SCORE, we proceed inductively as follows. For each $i \in [n]$, let I_i be our instance of *Hyp*-PERFECT SCORE restricted to the first i agents. We will explain how to find a proposal x_i approved by each of the agents $v_1 \ldots, v_i$ if it exists (and output 'no' if it does not). For I_1, we can set $x_1 = v_1$. To solve I_i, we first solve I_{i-1}. If the answer for I_{i-1} is 'no', then we output 'no' as well. Otherwise, we form an instance of *Hyp-k*-COMPROMISE that contains the first i agents; the initial deliberative coalition structure consists of $(\{v_1, \ldots, v_{i-1}\}, x_{i-1})$ and the singleton deliberative coalition $(\{v_i\}, v_i)$. We can then run the algorithm for *Hyp-k*-COMPROMISE, $k \geq 2$, on this instance; it returns a proposal if and only if I_i is a yes-instance of PERFECT SCORE. $\square$

On the positive side, the problem of finding a popular proposal becomes easy if the number of agents or the number of dimensions is small: indeed, we can obtain an FPT algorithm with respect to each of these parameters.

Proposition 5. *Given a d-hypercube deliberation space, we can compute a popular proposal in time $O(2^d dn)$.*

Proof. We can go through the proposals one by one; it takes $O(d)$ time to decide whether a given agent prefers a given proposal to the status quo, so we can determine the score of each proposal in time $O(dn)$. As there are $2^d - 1$ proposals other than the status quo, the bound on the running time follows. $\square$

Proposition 6. *Given a d-hypercube deliberation space, we can compute a popular proposal in time $poly(2^{n2^n}, \log d)$.*

Proof. Given an n-bit vector $\mathbf{b} = (b_1, \ldots, b_n) \in \{0,1\}^n$, we say that a dimension $j \in [d]$ is of *type* $\mathbf{b}$ if for each $i \in [n]$ the j-th coordinate of v_i is equal to b_i. Thus, each dimension belongs to one of the 2^n possible types. For a type $\mathbf{b}$, let $n_{\mathbf{b}}$ denote the number of dimensions of type $\mathbf{b}$. We can then represent a proposal x by a sequence of numbers $(x_{\mathbf{b}})_{\mathbf{b} \in \{0,1\}^n}$, where $0 \leq x_{\mathbf{b}} \leq n_{\mathbf{b}}$: the value $x_{\mathbf{b}}$ indicates the number of dimensions of type $\mathbf{b}$ that are set to 1 in x. Now, for each subset of agents $S \subseteq \mathcal{V}$, we formulate a set of constraints on the values $(x_{\mathbf{b}})_{\mathbf{b} \in \{0,1\}^n}$ which ensure that x is approved exactly by the agents in S.

1. For each agent i in S,

$$\sum_{\mathbf{b}:b_i=0} x_{\mathbf{b}} + \sum_{\mathbf{b}:b_i=1} (n_{\mathbf{b}} - x_{\mathbf{b}}) < \sum_{\mathbf{b}:b_i=1} n_{\mathbf{b}}$$
$$\iff \sum_{\mathbf{b}:b_i=0} x_{\mathbf{b}} - \sum_{\mathbf{b}:b_i=1} x_{\mathbf{b}} \leq -1.$$

In the constraint above, $\sum_{\mathbf{b}:b_i=0} x_{\mathbf{b}}$ counts the number of dimensions where the agent has value 0, but the proposal has value 1, while $\sum_{\mathbf{b}:b_i=1} (n_{\mathbf{b}} - x_{\mathbf{b}})$ counts the number of dimensions where the agent has value 1 but the proposal has value 0; overall, it measures the distance

of the agent from the proposal. $\sum_{\mathbf{b}:b_i=1} n_{\mathbf{b}}$ measures the distance of the agent from the status quo.

2. For each agent i not in S,

$$\sum_{\mathbf{b}:b_i=0} x_{\mathbf{b}} + \sum_{\mathbf{b}:b_i=1} (n_{\mathbf{b}} - x_{\mathbf{b}}) \geq \sum_{\mathbf{b}:b_i=1} n_{\mathbf{b}}$$
$$\iff \sum_{\mathbf{b}:b_i=0} x_{\mathbf{b}} - \sum_{\mathbf{b}:b_i=1} x_{\mathbf{b}} \geq 0.$$

3. Feasibility constraints: for each $\mathbf{b} \in \{0,1\}^n$, $0 \leq x_{\mathbf{b}} \leq n_{\mathbf{b}}$.

The constraints above form an ILP with 2^n variables and $n + 2^n$ constraints. Lenstra's algorithm (and subsequent improvements of it) can solve an ILP in time exponential in the number of variables, but linear in the number of bits required to represent the problem (constraints). As $n_{\mathbf{b}} \leq d$ and can be represented in $O(\log d)$ bits for each $\mathbf{b} \in \{0,1\}^n$, we obtain a time complexity of $poly(2^{n2^n}, \log d)$. We can find the optimal proposal by searching over all 2^n subsets of agents $S \subseteq \mathcal{V}$. The overall running time is then $2^n \cdot poly(2^{n2^n}, \log d) = poly(2^{n2^n}, \log d)$. $\square$

3.2 Euclidean Deliberation Spaces

Recall that, in a Euclidean space, an agent v approves approves a proposal p if and only if

$$\rho(v, p) < \rho(v, 0) \iff ||p||^2 < 2\langle v, p \rangle. \tag{1}$$

That is, a given proposal p splits $\mathbb{R}^d$ into two half-spaces; the half-spaces are divided by the hyper-plane orthogonal to and bisecting the line segment joining p to the origin. We will use this correspondence between proposals and half-spaces throughout this section.

We first observe that, in contrast to hypercube spaces, for Euclidean deliberation spaces the PERFECT SCORE problem is easy.

Proposition 7. *Euc-*PERFECT SCORE *is polynomial-time solvable.*

Proof. It suffices to check whether there exists a hyperplane $\mathcal{H}$ passing through r such that the entire set $\mathcal{V}$ lies on the same side of this hyperplane. This is equivalent to checking whether there exists an $x \in \mathbb{R}^d$ that satisfies the following linear program: $\langle v, x \rangle > 0, \forall v \in \mathcal{V}$. Once we find such an x, we can choose the proposal to be $p = \epsilon x / ||x||$ for sufficiently small $\epsilon > 0$. $\square$

However, finding a proposal that enjoys a specific degree of support turns out to be computationally challenging,

Theorem 8. *Euc-*SCORE *is NP-complete.*

We prove Theorem 8 by a reduction from 3-SAT; proof provided in the appendix.

Just like for hypercubes, *Euc*-SCORE is fixed-parameter tractable with respect to the number of agents.

Proposition 9. *Euc-*SCORE *can be solved in time $2^n \cdot poly(|I|)$.*

Proof. We reuse the argument in the proof of Proposition 7, but apply it to every subset of agents. That is, for every subset $S \subseteq \mathcal{V}$ of agents, we check whether there exists a hyper-plane passing through the origin that has all the agents in S strictly on one side of the hyper-plane, by solving a linear program. Altogether, we have to solve 2^n linear programs, so the bound on the running time holds. $\square$

However, for the Euclidean case, we were unable to obtain an FPT algorithm with respect to the number of dimensions. On the positive side, we can place our problem in the complexity class XP with respect to this parameter, i.e., show that it can be solved in polynomial time for the practically important case where the dimension of the underlying Euclidean space is small.

Proposition 10. *Euc-*SCORE *can be solved in time polynomial in* n^{d+1} *and* $|I|$.

Proof. As we observed earlier, a proposal divides $\mathbb{R}^d$ into two half-spaces. The set of half-spaces in $\mathbb{R}^d$ has a VC-dimension of $d+1$ (Kearns and Vazirani 1994), and therefore, the set of proposals also has a VC-dimension of at most $d+1$. Given a set $\mathcal{V}$ of n agents, let $\mathcal{S}$ be the following set of subsets of agents:

$$\mathcal{S} = \{S \subseteq \mathcal{V} \mid \exists x \in \mathcal{X} \text{ s.t. } ||x||^2 < 2\langle v, x\rangle, \forall v \in S,$$
$$\text{and } ||x||^2 \geq 2\langle v, x\rangle, \forall v \notin S\}.$$

In other words, for every set $S \in \mathcal{S}$, there exists a proposal that is supported by all the agents in S and none of the agents not in S. As the set of proposals has a VC-dimension of $d+1$, the set $\mathcal{S}$ is of size $O(n^{d+1})$ and the elements of $\mathcal{S}$ can be enumerated in time polynomial in n when d is fixed.[2] The elements of $\mathcal{S}$ can be computed inductively, see the proof of Lemma 3.1 in Kearns and Vazirani (1994). We are also using the fact that given an arbitrary set of agents S, we can efficiently compute a proposal that is supported by exactly the agents in S by solving an LP, as shown in Proposition 7. So, we can find the largest S in $\mathcal{S}$ and the corresponding proposal efficiently. $\square$

To conclude this section, we consider the complexity of 2-COMPROMISE in d-Euclidean deliberation spaces. Recall that Elkind et al. (2020) prove (Theorem 3) that a sequence of 2-compromises is guaranteed to converge to a successful deliberative coalition structure. Their argument proceeds by showing that whenever a deliberative coalition structure is not successful, then one of the following conditions holds: (i) two existing coalitions can merge; (ii) there exists an agent that can join a maximum-size coalition (the new coalition may need to choose a proposal that differs from the proposal originally supported by the maximum-size coalition), or (iii) the deliberative coalition structure consists of two coalitions (and hence there is a 2-compromise transition). Note that the transitions in (ii) and (iii) increase the size of the largest coalition and (i) does not decrease it, whereas (i) decreases the number of coalitions. Consequently, one

---

[2]See Lemma 3.1 and 3.2 of Chapter 3 of the book by Kearns and Vazirani (1994)

can reach a successful outcome in $O(n^2)$ steps by verifying conditions (i)–(iii) and performing the respective transition when one of them holds. Now, one can efficiently verify whether condition (i) or (ii) holds, and compute the outcome of the respective transition; if this was also true for (iii), we would be able to solve *Euc-*SCORE in polynomial time, in contradiction to Theorem 8 (assuming P$\neq$NP). We obtain the following corollary.

Corollary 11. *For each* $k \geq 2$, *there is no polynomial-time algorithm for Euc-*k*-*COMPROMISE *unless P$\neq$NP.*

4 Number of Transitions

In this section, we go back to viewing deliberation as a decentralized process, and ask whether all sequences of k-compromise transitions terminate after a number of steps that is polynomial in n. Recall that, in Euclidean spaces, if the agents are told to perform 2-compromise transitions in a specific easy-to-compute order, then deliberation terminates in n^2+1 steps (Elkind et al. 2020) (see also the exposition at the end of Section 3). However, if the agents can choose the order of transitions arbitrarily, the only known upper bound (which applies to all deliberation spaces) is n^n, proved using a lexicographic potential function. In the next theorem, we put forward a different potential function, which enables us to improve the upper bound to 2^n.

Theorem 12. *The number of transitions in a* 2*-deliberation is at most* 2^n. *This can be shown using the following potential function:*

$$\phi(\mathbf{D}) = -|\mathbf{D}| + \sum_{(C,x)\in\mathbf{D}} 2^{|C|}. \qquad (2)$$

Proof. Consider a deliberative coalition structure $\mathbf{D}$. Note that $\bigcup_{(C,x)\in\mathbf{D}} C = \mathcal{V}$. From now on, to simplify notation, we will identify a deliberative coalition (C,x) with its set of agents C, i.e., we will speak of a coalition C in $\mathbf{D}$.

Suppose a 2-compromise transition occurs, where two coalitions A and B of sizes a and b in $\mathbf{D}$ compromise to form coalitions C, D, and E of sizes c, d and e in $\mathbf{D}'$, where $D \subsetneq A$, $E \subsetneq B$. We can assume without loss of generality that $a \leq b < c$. Note that D and E may be empty, in which case they do not appear in the new coalition structure $\mathbf{D}'$. We have $a + b = c + d + e$, and the value of $|\mathbf{D}'| - |\mathbf{D}|$ may be either -1, 0, or 1, depending upon whether D and/or E are empty. The change in potential is

$$\phi(\mathbf{D}') - \phi(\mathbf{D}) = 2^c - 2^a - 2^b + 1 + (2^d - 1) + (2^e - 1).$$

Indeed, if D is non-empty then it contributes 2^d to $\sum_{C\in\mathbf{D}'} 2^{|C|}$ and -1 to the $-|\mathbf{D}'|$ portion of $\phi(\mathbf{D}')$. On the other hand, if D is empty, then it makes neither of these contributions to $\phi(\mathbf{D}')$, but also $2^d - 1 = 0$. The same argument applies to E. As $2^d - 1$ and $2^e - 1$ are always non-negative, we obtain

$$\phi(\mathbf{D}') - \phi(\mathbf{D}) \geq 2^c - 2^a - 2^b + 1 \geq 2\cdot 2^b - 2^b - 2^b + 1 = 1,$$

as $c > b \geq a$. Hence, any 2-compromise transition increases the potential by at least 1. On the other hand, we have $n \leq \phi(\mathbf{D}) \leq 2^n - 1$ for any $\mathbf{D}$. $\square$

Observe that Theorem 12 is independent of the deliberation space and is a property of 2-compromise transitions. A similar result can be proven for k-compromise transitions: they converge in at most k^n steps.

We now focus on proving a lower bound on the convergence of 2-compromise transitions. We first prove a lemma that applies to any deliberation space that satisfies a certain property. We then use this lemma to construct a family of examples for hypercube deliberation spaces, and for Euclidean deliberation spaces.

Lemma 13. *Suppose a deliberation space with the set of proposals $\mathcal{X}$ and the set of agents $\mathcal{V}$ satisfies the following property: for every $C \subseteq \mathcal{V}$ there exists a proposal $p \in \mathcal{X}$ s.t. all agents in C approve p and none of the agents not in C approve p. Then, a deliberation may take $\Omega(2^{\sqrt{n}/2})$ 2-compromise transitions.*

Proof. Fix a coalition structure $\mathbf{D}$. The property of the deliberation space formulated in the theorem statement implies that for every pair of deliberative coalitions $(A, x), (B, y) \in \mathbf{D}$ and every $C \subseteq A \cup B$ we can find a proposal z approved by agents in C (and no other agents). Thus, in what follows, we can reason in terms of sets of agents rather than deliberative coalitions. Consequently, to simplify notation, we will write $C \in \mathbf{D}$ instead of $(C, x) \in \mathbf{D}$.

We now prove that if the agents end up executing the following types of 2-compromise transitions, then they will take an exponential time to converge, starting from the coalition structure where all the agents are in singleton coalitions.

1. **Type 1.** If there is a pair $C, C' \in \mathbf{D}$ such that $|C| = |C'|$, then make the following transition (expressed in terms of coalition sizes):

$$(a) + (a) \longrightarrow (a+1) + \left\lfloor \frac{a-1}{2} \right\rfloor + \left\lceil \frac{a-1}{2} \right\rceil,$$

where $a = |C| = |C'|$. The change in the potential function (as defined in (2)) for such a transition is

$$\Delta\phi \leq 2^{a+1} + 2^{\lfloor \frac{a-1}{2} \rfloor} + 2^{\lceil \frac{a-1}{2} \rceil} - 2 \cdot 2^a$$
$$= 2^{\lfloor \frac{a-1}{2} \rfloor} + 2^{\lceil \frac{a-1}{2} \rceil} \leq \frac{3}{2} \cdot 2^{a/2}. \quad (3)$$

(Note that $\lfloor \frac{a-1}{2} \rfloor$ and $\lceil \frac{a-1}{2} \rceil$ may be zero, which implies that the number of coalitions decreases, and therefore, may contribute 1 to $\Delta\phi$ due to change in $|\mathbf{D}|$. But, in that case, $2^{\lfloor \frac{a-1}{2} \rfloor}$ would not be in $\Delta\phi$; compensates.)

2. **Type 2.** If there are no Type 1 transitions available, then select two smallest coalitions $C, C' \in \mathbf{D}$ and make the following transition:

$$(a) + (b) \longrightarrow (b+1) + \left\lfloor \frac{a-1}{2} \right\rfloor + \left\lceil \frac{a-1}{2} \right\rceil, \quad (4)$$

where $a = |C|$, $b = |C'|$ and $a \leq b$.

Now, if $\max_{C \in \mathbf{D}} |C| \leq \sqrt{n}$, then there must be at least one pair of coalitions of the same size. If not, then

$$\sum_{i=1}^{\sqrt{n}} i = \sqrt{n}(\sqrt{n}+1)/2 = n/2 + \sqrt{n}/2 < n,$$

for large enough n, we get a contradiction. So, only Type 1 transitions are made until $\max_{C \in \mathbf{D}} |C|$ exceeds $\sqrt{n}$.

From (3), until $\max_{C \in \mathbf{D}} |C| \leq \sqrt{n}$ holds, we know that the change in potential for Type 1 transitions is bounded by

$$\Delta\phi \leq \frac{3}{2} \cdot 2^{a/2} \leq \frac{3}{2} \cdot 2^{\sqrt{n}/2}.$$

We also know that if $\max_{C \in \mathbf{D}} |C|$ goes above $\sqrt{n}$ then the value of ϕ must exceed $2^{\sqrt{n}} - n$, and the initial value of ϕ is n. Hence, the number of transitions must be at least

$$\frac{2}{3} \cdot \frac{2^{\sqrt{n}} - 2n}{2^{\sqrt{n}/2}} = \frac{2}{3} \cdot \left(2^{\sqrt{n}/2} - \frac{2n}{2^{\sqrt{n}/2}} \right).$$

$\square$

4.1 Hypercube Deliberation Spaces

In the next theorem, we apply Lemma 13 to hypercube deliberation spaces by constructing a family of the deliberation spaces that satisfies the required conditions for the lemma.

Theorem 14. *There exists a family of hypercube deliberation spaces where a 2-deliberation may take $\Omega(2^{\sqrt{n}/2})$ 2-compromise transitions.*

Proof. For every even $d \geq 2$, we will construct an instance with $n = d/2$ agents. All these agents have 0s in their first $d/2$ dimensions and 1s in all but one (which is different for each agent) of the last $d/2$ dimensions. In particular, the i-th agent $v_i = (v_{i,j})_{j \in [d]}$ has

$$v_{i,j} = \begin{cases} 0, & \text{if } j \leq d/2 \text{ or } j - 1 - d/2 = i \\ 1, & \text{otherwise} \end{cases}.$$

Consider a set S of $m < d/2$ agents. They agree on $d/2 - m$ 1s. Now consider a proposal that has $d/2 - m - 1$ 1s somewhere in the first $d/2$ dimensions and $d/2 - m$ 1s in the dimensions where the agents in S agree on, and 0s everywhere else. Each agent in S is at a distance of $(d/2-m-1)+((d/2-1)-(d/2-m)) = d/2-2$ from this proposal, and at a distance of $d/2-1$ from the origin, so they approve the proposal. However, an agent not in S is at a distance of at least $(d/2-m-1)+((d/2-1)-(d/2-m-2)) = d/2$ from the proposal, so they do not approve the proposal.

So, for any subset of agents S, there is a proposal that is approved exactly by the agents in S. Applying Lemma 13, we get the desired result. $\square$

4.2 Euclidean Deliberation Spaces

As we did for hypercube deliberation spaces, we construct a family of Euclidean deliberation spaces that satisfies the required conditions for Lemma 13.

Theorem 15. *There exists a family of Euclidean deliberation spaces where a 2-deliberation may take $\Omega(2^{\sqrt{n}/2})$ 2-compromise transitions.*

Proof. For each $d \geq 2$, we will construct an instance with d agents. Let the agent $v_i \in \mathcal{V}$ be positioned on the i-th axis at a distance of 1 from the origin, i.e., agent v_1 is located at $(1, 0, \ldots, 0)$, agent v_2 at $(0, 1, 0, \ldots, 0)$, and so on.

For every $S \subseteq \mathcal{V}$, let the point x^S be defined as

$$x_i^S = \begin{cases} 1/|S|, & \text{if } v_i \in S \\ 0, & \text{otherwise} \end{cases}.$$

Observe that the distance of agent $v_i \in S$ from x^S is $\rho(v_i, x^S) = (1 - 1/|S|)^2 + (1/|S|)^2(|S| - 1) = 1 - 1/|S| < 1$, so agent v_i prefers x^S to the status quo. But, for an agent $v_i \notin S$, the distance of v_i from x^S is $1 + |S|(1/|S|)^2 = 1 + 1/|S| > 1$. So, for any subset of agents S, there is a proposal x^S that is supported exactly by the agents in S. Applying Lemma 13, we obtain the desired result. $\square$

5   Beyond Two-Way Compromises

Elkind et al. (2020) showed that, while in Euclidean deliberation spaces 2-compromise transitions guarantee successful deliberation, in hypercube deliberation spaces this is not the case. Specifically, they proved the following result:

Theorem 16. *(Elkind et al. 2020) There are d-hypercube deliberation spaces where d-compromise transitions are necessary for a successful deliberation; on the other hand, in every d-hypercube deliberation space, $(2^{d-1} + (d+1)/2)$-compromise transitions are sufficient for a successful deliberation.*

Theorem 16 leaves a big gap between the lower bound of d and the upper bound of $2^{d-1} + (d+1)/2$. We tighten this bound by proving a lower bound of $2^{\Theta(d)}$.

Theorem 17. *There are d-hypercube deliberation spaces where $2^{\Theta(d)}$-compromise transitions are necessary for a successful deliberation.*

The proof of Theorem 17 is given in the appendix. To prove that k-compromise transitions are necessary for a successful deliberation, we need to describe a deliberation space and a coalition structure, where:

1. The coalition structure is sub-optimal, i.e., there exists a coalition structure with a larger coalition. One way to prove this is to show that an ℓ-compromise transition, $\ell \geq k$, from this coalition structure leads to a strictly larger coalition. That is, we need to describe a particular proposal and a particular set of agents, and argue that these agents support the new proposal, and the new coalition is larger than the current coalitions of all these agents.

2. Any ℓ-compromise, for $\ell < k$, does not lead to a strictly larger coalition. First, there must be at least k coalitions in the current coalition structure. Second, we need to prove that for any proposal in the deliberation space, any set of agents that supports this proposal and is contained in fewer than k current coalitions is at most as large as the current coalition of one of the members of this set.

If we focus on single-agent transitions, the second point above says that, in the current coalition structure, any agent in a (weakly) smaller coalition should not support the proposal of a (weakly) larger coalition. So, the proposals of all the coalitions should be different and should not be supported by any agent from an equal or smaller coalition. k-compromise transitions include single-agent transitions and

other more complex transitions, and the construction should address all of them. Elkind et al. (2020) constructed such an example for $k = d$, we do this for

$$k = \binom{(d-1)/9}{(d-1)/27} - 1 \geq 3^{(d-1)/27} - 1 = 2^{\Theta(d)}.$$

6   Conclusions and Future Work

We have provided an in-depth investigation of the complexity of deliberation in two models proposed by Elkind et al. (2020), answering several open questions formulated in that paper. Our results are mostly negative: in both models we have considered, identifying a successful proposal is hard even for a centralized algorithm, and agents will find it computationally challenging to discover feasible transitions from the status quo. Moreover, a completely decentralized deliberation procedure, in which groups of agents are free to execute compromise transitions in any order, may take a very long time to converge. Finally, while the Euclidean deliberation spaces have the attractive feature that a successful deliberation is possible even if each transition only involves agents from two coalitions, in hypercube spaces we may need transitions that involve exponentially many coalitions, negating the benefits of a decentralized process.

Nevertheless, we do not feel that these negative results mean that we should give up on this model of deliberation. Rather, it would be interesting to identify 'islands of tractability', i.e., additional conditions that make this model tractable, both in terms of computational complexity and in terms of the length of the deliberation process and the number of coalitions involved in each transition; our FPT and XP results are a step in that direction. It would also be interesting to complement our theoretical findings with empirical work, checking if natural heuristics enable the agents to quickly converge to good (even if not necessarily optimal) outcomes.

There are several questions concerning the complexity of deliberative coalition formation that are left open by our work. For instance, while Theorem 15 shows that convergence may be slow if the number of dimensions scales with the number of agents, it is not clear if this remains true if the number of dimensions is a fixed constant. Further, throughout the paper, we assume that the space of feasible proposals is the entire metric space. A more general approach is to assume that it is a proper subset of the metric space: this subset can be described implicitly by constraints or, in case it is finite, listed explicitly as part of the input. Of course, the computational complexity questions become trivial if the feasible proposals are listed explicitly, but it is not clear if we can bound the length of deliberation as a polynomial function of the number of proposals; on the other hand, the lower bound arguments of Theorems 14 and 15 no longer apply. Finally, k-compromise transitions, as defined by Elkind et al. (2020) can be viewed as better responses in the respective game; it would also be interesting to explore the speed of convergence of *best* response dynamics, where a negotiation among k coalitions always results in the largest possible coalition that can be formed by their members.