Sequential Deliberation for Social Choice

Brandon Fain[1]*, Ashish Goel[2]**, Kamesh Munagala[1]***, and Sukolsak Sakshuwong[2].

[1] Computer Science Dept., Duke University, Durham NC
{btfain, kamesh}@cs.duke.edu
[2] Management Science and Engineering Dept., Stanford University, Stanford CA
{ashishg, sukolsak}@stanford.edu

Abstract. Social choice is a normative study of designing protocols for collective decision making. However, in instances where the underlying decision space is too large or complex for ordinal voting, standard voting methods may be impractical. How then can we design a protocol - preferably decentralized, simple, scalable, and not requiring any special knowledge of the decision space - to reach consensus? We propose sequential deliberation as a natural solution to this problem. In this iterative method, successive pairs of agents bargain over the decision space using the previous decision as a disagreement alternative. We show that sequential deliberation finds a 1.208-approximation to the optimal social cost when the space of preferences define a median graph, coming very close to this value with only a small constant number of agents sampled from the population. We also give lower bounds on simpler classes of mechanisms to justify our design choices. We further show that sequential deliberation is ex-post Pareto efficient and has truthful reporting as an equilibrium of the induced extensive form game. Finally, we prove that for general metric spaces, the first and second moment of of the distribution of social cost of the outcomes produced by sequential deliberation are also bounded by constants.

1 Introduction

Suppose a university administrator plans to spend millions of dollars to update her campus, and she wants to elicit the input of students, staff, and faculty. In a typical social choice setting, she could first elicit the bliss points of the students, say "new gym," "new library," and "new student center." However, voting on these options need not find the social optimum, because it is not clear that the social optimum is even on the ballot. In such a setting, *deliberation* between individuals would find entirely new alternatives, for example "replace

* Supported by NSF grants CCF-1637397 and IIS-1447554.
** Supported by the Army Research Office Grant No. 116388, the Office of Naval Research Grant No. 11904718, by NSF grant CCF-1637418, and by the Stanford Cyber Initiative.
*** Supported by NSF grants CCF-1408784, CCF-1637397, and IIS-1447554.

gym equipment plus remodeling campus dining plus money for scholarship." This leads to finding a social optimum over a wider space of semi-structured outcomes that the system/mechanism designer was not originally aware of, and the participants had not initially articulated.

We therefore start with the following premise: The mechanism designer may not be able to enumerate the outcomes in the decision space or know their structure, and this decision space may be too big for most ordinal voting schemes. (For instance, ordinal voting is difficult to implement in complex combinatorial spaces [25] or in continuous spaces [15].) However, we assume that agents can still reason about their preferences and small groups of agents can negotiate over this space and collaboratively propose outcomes that appeal to all of them. Our goal is to design protocols based on such a primitive by which small group negotiation can lead to an aggregation of societal preferences without a need to formally articulate the entire decision space and without every agent having to report ordinal rankings over this space.

The need for small groups is motivated by a practical consideration as well as a theoretical one. On the practical side, there is no online platform, to the best of our knowledge, that has a successful history of large scale deliberation and decision making on complex issues; in fact, large online forums typically degenerate into vitriol and name calling when there is substantive disagreement among the participants. Thus, if we are to develop practical tools for decision making at scale, a sequence of small group deliberations appears to be the most plausible path. On the theoretical side, we understand the connections between sequential protocols for deliberation and axiomatic theories of bargaining for small groups, e.g. for pairs [34, 8], but not for large groups, and we seek to bridge this gap.

Summary of Contributions. Our main contributions are two-fold:

- A simple and practical sequential protocol that only requires agents to negotiate in pairs and propose outcomes that appeal to both agents.
- A canonical analytic model in which we can precisely state properties of this protocol in terms of approximation of the social optimum, Pareto-efficiency, and incentive-compatibility, as well as compare it with simpler protocols.

1.1 Background: Bargaining Theory

Before proceeding further, we review bargaining, the classical framework for two-player negotiation in Economics. Two-person bargaining, as framed in [29], is a game wherein there is a disagreement outcome and two agents must cooperate to reach a decision; failure to cooperate results in the adoption of the disagreement outcome. Nash postulated four axioms that the bargaining solution ought to satisfy assuming a convex space of alternatives: Pareto optimality (agents find an outcome that cannot be simultaneously improved for both of them), symmetry between agents, invariance with respect to affine transformations of utility (scalar multiplication or additive translation of any agent's utility should not change the outcome), and independence of irrelevant alternatives (informally that the presence of a feasible outcome that agents do not select does not influence their decision). Nash proved that the solution maximizing the Nash product (that we describe later) is the unique solution satisfying these axioms. To provide some explanation of how two agents might find such a solution, [34] shows that Nash's solution is the subgame perfect equilibrium of a simple repeated game on the two agents, where the agents take turns making offers, and at each round, there is an exogenous probability of the process terminating with no agreement.

The two-person bargaining model is therefore clean and easy to reason about. As a consequence, it has been extensively studied. In fact, there are other models and solutions to two-person bargaining, each with a slightly different axiomatization [21, 22, 28], as well as several experimental studies [33, 30, 7]. In a social choice setting, there are typically many more than two agents, each agent having their own complex preferences. Though bargaining can be generalized to n agents with similar axiomatization and solution structure, such a generalization is considered impractical. This is because in reality it is difficult to get a large number of individuals to negotiate coherently; complexities come with the formation of coalitions and power structures [19, 24]. Any model for simultaneous bargaining, even with three players [6], needs to take these messy aspects into account.

1.2 A Practical Compromise: Sequential Pairwise Deliberation

In this paper, we take a middle path, avoiding both the complexity of explicitly specifying preferences in a large decision space that any individual agent may not even fully know (fully centralized voting), and that of simultaneous n-person bargaining (a fully decentralized cooperative game). We term this approach *sequential deliberation*. We use 2-person bargaining as a basic primitive, and view deliberation as a sequence of pairwise interactions that refine good alternatives into better ones as time goes by.

More formally, there is a decision space $\mathcal{S}$ of feasible alternatives (these may be projects, sets of projects, or continuous allocations) and a set $\mathcal{N}$ of agents. We assume each agent has a hidden cardinal utility for each alternative. We encapsulate deliberation as a sequential process. The framework that we analyze in the rest of the paper is captured in Fig. 1.

Our framework is simple with low cognitive overhead, and is easy to implement and reason about. Though we don't analyze other variants in this paper, we note that the framework is flexible. For instance, the bargaining step can be replaced with any function $\mathcal{B}(u, v, a)$ that corresponds to an interaction between u and v using a as the disagreement outcome; we assume that this function maximizes the Nash product, that is, it corresponds to the Nash bargaining solution. Similarly, the last step of social choice could be implemented by a central planner based on the distribution of outcomes produced.

1. In each round $t = 1, 2, \dots, T$:
   (a) A pair of agents u^t and v^t are chosen independently and uniformly at random with replacement.
   (b) These agents are presented with a disagreement alternative a^t, and perform bargaining, which is encoded as a function $\mathcal{B}(u, v, a)$ as described below.
   (c) Agents u_t and v_t are asked to output a consensus alternative; if they fail to reach a consensus then the alternative a^t is output.
   (d) Let o^t denote the alternative that is output in round t. We set $a^{t+1} = o^t$, where we assume a^1 is the bliss point of an arbitrary agent.
2. The final social choice is a^T. Note that this is equivalent to drawing an outcome at random from the distribution generated by repeating this protocol several times.

Fig. 1. A framework for sequential pairwise deliberation.

1.3 Analytical Model: Median Graphs and Nash Bargaining

The framework in Fig. 1 is well-defined and practical irrespective of an analytical model. However, we provide a simple analytical model for specifying the preferences of the agents in which we can precisely quantify the behavior of this framework as justification.

Median Graphs. We assume that the set $\mathcal{S}$ of alternatives are vertices of a *median graph*. A median graph has the property that for each triplet of vertices u, v, w, there is a unique point that is common to the three sets of shortest paths (since there may be multiple pairwise shortest paths), those between u, v, between v, w, and between u, w. This point is the unique *median* of u, v, w. We assume each agent u has a bliss point $p_u \in \mathcal{S}$, and his disutility for an alternative $a \in \mathcal{S}$ is simply $d(p_u, a)$, where $d(\cdot)$ is the shortest path distance function on the median graph. (Note that this disutility can have an agent-dependent scale factor.) Several natural graphs are median graphs, including trees, points on the line, hypercubes, and grid graphs in arbitrary dimensions [23]. As we discuss in Sect. 1.5, because of their analytic tractability and special properties, median graphs have been extensively studied as structured models for spatial preferences in voting theory. Some of our results generalize to metric spaces beyond median graphs; see Sect. 5.

Nash Bargaining. The model for two-person bargaining is simply the classical *Nash bargaining* solution described before. Given a disagreement alternative a, agents u and v choose that alternative $o \in \mathcal{S}$ that maximizes:

$$\text{Nash product } = (d(p_u, a) - d(p_u, o)) \times (d(p_v, a) - d(p_v, o))$$

subject to individual rationality, that is, $d(p_v, o) \leq d(p_v, a)$ and $d(p_u, o) \leq d(p_u, a)$. The Nash product maximizer need not be unique; in the case of ties we postulate that agents select the outcome that is closest to the disagreement outcome. As mentioned before, the Nash product is a widely studied axiomatic notion of pairwise interactions, and is therefore a natural solution concept in our framework.

Social Cost and Distortion. The *social cost* of an alternative $a \in \mathcal{S}$ is given by $SC(a) = \sum_{u \in \mathcal{N}} d(p_u, a)$. Let $a^* \in \mathcal{S}$ be the minimizer of social cost, *i.e.*, the *generalized median*. We measure the Distortion of outcome a as

$$\text{Distortion}(a) = \frac{SC(a)}{SC(a^*)} \tag{1}$$

where we use the expected social cost if a is the outcome of a randomized algorithm. Note that our model is fairly general in that the bliss points of the agents in $\mathcal{N}$ form an arbitrary subset of $\mathcal{S}$. Assuming that disutility is some metric over the space follows recent literature [2, 3, 9, 10, 17], and our tightest results are for median graphs specifically.

1.4 Our Results

Before presenting our results, we re-emphasize that while we present analytical results for sequential deliberation in specific decision spaces, the framework in Fig. 1 is well defined regardless of the underlying decision space and the mediator's understanding of the space. At a high level, this flexibility and generality in practice are its key advantages.

Bounding Distortion. Our main result is in Sect. 3, and shows that for sequential Nash bargaining on a median graph, the expected Distortion of outcome a^T has an upper bound approaching 1.208 as $T \to \infty$. Surprisingly, we show that in $T = \log_2 \frac{1}{\epsilon} + 2.575$ steps, the expected Distortion is at most $1.208 + \epsilon$, independent of the number of agents, the size of the median space, and the initial disagreement point a^1. For instance, the Distortion falls below 1.22 in at most 9 steps of deliberation, which only requires a random sample of at most 20 agents from the population to implement.

In Sect. 3.2, we ask: *How good is our numerical bound?* We present a sequence of lower bounds for social choice mechanisms that are allowed to use increasingly richer information about the space of alternatives on the median graph. This also leads us to make qualitative statements about our deliberation scheme.

- We show that any social choice mechanism that is restricted to choosing the bliss point of some agent cannot have Distortion better than 2. More generally, it was recently shown [18] that even eliciting the top k alternatives for each agent does not improve the bound of 2 for median graphs unless $k = \Omega(|\mathcal{S}|)$.
- Next consider mechanisms that choose, for some triplet (u, v, w) of agents with bliss points p_u, p_v, p_w, the median outcome $m_{uvw} = \mathcal{B}(u, v, p_w)$. We

show this has Distortion at least 1.316, which means that sequential deliberation is superior to one-shot deliberation that outputs o^1 where a^1 is the bliss point of some agent.
– Finally, for every pair of agents (u, v), consider the set of alternatives on a shortest path between p_u and p_v. This encodes all deliberation schemes where $\mathcal{B}$ finds a Pareto-efficient alternative for some 2 agents at each step. We show that any such mechanisms has Distortion ratio at least $9/8 = 1.125$. This space of mechanisms captures sequential deliberation, and shows that sequential deliberation is close to best possible within this space.

Properties of Sequential Deliberation. We next show that sequential deliberation has several natural desiderata on median graphs in Sect. 4. In particular:

– Under mild assumptions, the limiting distribution over outcomes of sequential deliberation is *unique*.
– For every $T \geq 1$, the outcome o^T of sequential deliberation is *ex-post Pareto-efficient*, meaning that there is no other alternative that has at most that social cost for all agents and strictly better cost for one agent. This is not a priori obvious, since the outcome at any one round only uses inputs from two agents.
– Interpreted as a mechanism, truthful play is a *sub-game perfect Nash equilibrium* of sequential deliberation. This interpretation is made precise in Sect. 4, but at a high level, we seek to address the following concern: In a sequential setting, would any agent have incentive to misrepresent their preferences so that they gain an advantage?

Beyond Median Graphs. In Sect. 5, we consider general metric spaces. We show that the Distortion of sequential deliberation is always at most a factor of 3. More surprisingly, we show that sequential deliberation has constant distortion even for the second moment of the distribution of social cost of the outcomes, *i.e.*, the latter is at most a constant factor worse than the optimum squared social cost. The practical implication is that one can look at the distribution of outcomes produced by deliberation and know that the standard deviation in social cost is comparable to its expected value.[3]

1.5 Related Work

While the real world complexities of the model are beyond the analytic confines of this work, deliberation as an important component of collective decision making and democracy is studied in political science. For examples (by no means exhaustive), see [14, 37]. There is ongoing related work on Distortion of voting for simple analytical models like points in $\mathbb{R}$ [13], and in general metric spaces [3,

[3] See also recent work by [38] that considers minimizing the variance of randomized truthful mechanisms.

2, 9, 10, 17]. This work focuses on optimally aggregating ordinal preferences, say the top k preferences of a voter [18]. In contrast, our scheme elicits alternatives as the outcome of bargaining rounds that require agents to reason about cardinal preferences. We essentially show that for median graphs, unless k is very large, such deliberation has provably lower distortion than social choice schemes that elicit purely ordinal rankings.

Median graphs and their ordinal generalization, median spaces, have been extensively studied in social choice. The special cases of trees and grids have been studied as structured models for voter preferences [36, 5]. For general median spaces, the Condorcet winner (an alternative that pairwise beats any other alternative in terms of voter preferences) is related to the generalized median [4, 35, 39] – if the former exists, it coincides with the latter. Nehring and Puppe [31] show that any single-peaked domain which admits a non-dictatorial and neutral strategy-proof social choice function is a median space. Clearwater *et al.* [11] also showed that any set of voters and alternatives on a median graph will have a Condorcet winner. In a sense, these are the largest class of structured preferences where ordinal voting over the entire space of alternatives leads to a "clear winner" (importantly, we assume this is impractical).

Our paper is inspired by the *triadic consensus* results of Goel and Lee [16], where the authors also focus on small group interactions to make a collective decision. The authors show that the Distortion of their protocol approaches 1 on median graphs. However, the protocol crucially assumes individuals know the positions of other individuals, and requires the space of alternatives to coincide with the space of individuals. We make neither of these assumptions – in our case, the space of alternatives can be much larger than the number of agents, and individuals interact with others only via bargaining. This makes our protocol more practical, but restricts our Distortion to be bounded away from 1.

The notion of *democratic equilibrium* [20, 15] considers social choice mechanisms in continuous spaces where individual agents with complex utility functions perform update steps inspired by gradient descent. However, these schemes do not involve deliberation between agents and have little formal analysis of convergence. Several works have considered *iterative voting* where the current alternative is put to vote against one proposed by different random agent chosen each step [1, 26, 32], or other related schemes [27]. In contrast with our work, these protocols are not deliberative and require voting among several agents each step; furthermore, the analysis focuses on convergence to an equilibrium instead of welfare or efficiency guarantees.

2 Median Graphs and Nash Bargaining

In this section we will use the notation $\mathcal{N}$ for a set of agents, $\mathcal{S}$ for the space of feasible alternatives, and $\mathcal{H}$ for a distribution over $\mathcal{S}$. Most of our results are for the analytic model given earlier wherein the set $\mathcal{S}$ of alternatives are vertices of a *median graph*. All proofs are given in the full version of the paper [12].

Definition 1. *A median graph $G(\mathcal{S}, E)$ is an unweighted and undirected graph with the following property: For each triplet of vertices $u, v, w \in \mathcal{S} \times \mathcal{S} \times \mathcal{S}$, there is a unique point that is common to the shortest paths (which need not be unique between a given pair) between u, v, between v, w, and between u, w. This point is the unique* median *of u, v, w.*

In the framework of Fig. 1, we assume that at every step, two agents perform Nash bargaining with a disagreement alternative. The first results characterize Nash bargaining on a median graph. In particular, we show that Nash bargaining at each step will select the median of bliss points of the two agents and the disagreement alternative. After that, we show that we can analyze the Distortion of sequential deliberation on a median graph by looking at the embedding of that graph onto the hypercube.

Lemma 1. *For any median graph $G = (\mathcal{S}, E)$, any two agents u, v with bliss points $p_u, p_v \in \mathcal{S}$, and any disagreement outcome $a \in \mathcal{S}$, let M be the median. Then M maximizes the Nash product of u and v given a, and is the maximizer closest to a.*

Hypercube Embeddings. For any median graph $G = (\mathcal{S}, E)$, there is an isometric embedding $\phi : G \to Q$ of G into a hypercube Q [23]. This embedding maps vertices $\mathcal{S}$ into a subset of vertices of Q so that all pairwise distances between vertices in $\mathcal{S}$ are preserved by the embedding. A simple example of this embedding for a tree is shown in Fig. 2. We use this embedding to show the following result, in order to simplify subsequent analysis.

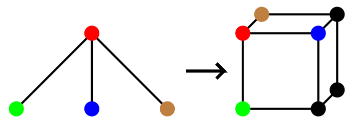

Fig. 2. The hypercube embedding of a 4-vertex star graph

Lemma 2. *Let $G(\mathcal{S}, E)$ be a median graph, and let ϕ be its isometric embedding into hypercube $Q(V, E')$. For any three points $t, u, v \in \mathcal{S}$, let M_G be the median of vertices t, u, v and let M_Q be the median of vertices $\phi(t), \phi(u), \phi(v) \in V$. Then $\phi(M_G) = M_Q$.*

3 The Efficiency of Sequential Deliberation

In this section, we show that the Distortion of sequential deliberation is at most 1.208. We then show that this bound is significant, meaning that mechanisms from simpler classes are necessarily constrained to have higher Distortion values. All proofs are given in the full version of the paper [12].

3.1 Upper Bounding Distortion

Recall the framework for sequential deliberation in Fig. 1 and the definition of Distortion in Equation (1). We first map the problem into a problem on hypercubes using Lemma 2.

Corollary 1. *Let $G = (\mathcal{S}, E)$ be a median graph, let $\phi : G \to Q$ be an isometric embedding of G onto a hypercube $Q(V, E')$, and let $\mathcal{N}$ be a set of agents such that each agent u has a bliss point $p_u \in \mathcal{S}$. Then the Distortion of sequential deliberation on G is at most the Distortion of sequential deliberation on $\phi(G)$ where each agent's bliss point is $\phi(p_u)$.*

Our main result in this section shows that as $t \to \infty$, the Distortion of sequential deliberation approaches 1.208, with the convergence rate being exponentially fast in t and independent of the number of agents $|\mathcal{N}|$, the size of the median space $|\mathcal{S}|$, and the initial disagreement point a^1. In particular, the Distortion is at most 1.22 in at most 9 steps of deliberation, which is indeed a very small number of steps.

Theorem 1. *Sequential deliberation among a set $\mathcal{N}$ of agents, where the decision space $\mathcal{S}$ is a median graph, yields $\mathbb{E}[Distortion(a^t)] \leq 1.208 + \frac{6}{2^t}$.*

3.2 Lower Bounds on Distortion

We will now show that the Distortion bounds of sequential deliberation are significant, meaning that mechanisms from simpler classes are constrained to have higher Distortion values. We present a sequence of lower bounds for social choice mechanisms that use increasingly rich information about the space of alternatives on a median graph $G = (\mathcal{S}, E)$ with a set of agents $\mathcal{N}$ with bliss points $V_{\mathcal{N}} \subseteq \mathcal{S}$. We first consider mechanisms that are constrained to choose outcomes in $V_{\mathcal{N}}$. For instance, this captures the Random Dictatorship algorithm that chooses the bliss point of a random agent. It shows that the compromise alternatives found by deliberation do play a role in reducing Distortion.

Lemma 3. *Any mechanism constrained to choose outcomes in $V_{\mathcal{N}}$ has Distortion at least 2.*

We next consider mechanisms that are restricted to choosing the median of the bliss points of some three agents in $\mathcal{N}$. This captures sequential deliberation run for $T = 1$ steps, as well as mechanisms that generalize dictatorship to an oligarchy composed of at most 3 agents. This shows that iteratively refining the bargaining outcome has better Distortion than performing only one iteration.

Lemma 4. *Any mechanism constrained to choose outcomes in $V_{\mathcal{N}}$ or a median of three points in $V_{\mathcal{N}}$ must have Distortion at least 1.316.*

We finally consider a class of mechanisms that includes sequential deliberation as a special case. We show that any mechanism in this class cannot have Distortion arbitrarily close to 1. This also shows that sequential deliberation is close to best possible in this class.

Lemma 5. *Any mechanism constrained to choose outcomes on shortest paths between pairs of outcomes in $V_{\mathcal{N}}$ must have Distortion at least* $9/8 = 1.125$.

The significance of the lower bound in Lemma 5 should be emphasized: though there is always a Condorcet winner in median graphs, it need not be any agent's bliss point, nor does it need to be Pareto optimal for any pair of agents. The somewhat surprising implication is that any local mechanism (in the sense that the mechanism chooses locally Pareto optimal points) is constrained away from finding the Condorcet winner.

4 Properties of Sequential Deliberation

In this section, we study some natural desirable properties for our mechanism: uniqueness of the stationary distribution of the Markov chain, ex-post Pareto-efficiency of the final outcome, and subgame perfect Nash equilibrium. All proofs are given in the full version of the paper [12].

Uniqueness of the Stationary Distribution. We first show that the Markov chain corresponding to sequential deliberation converges to a unique stationary distribution on the actual median graph, rather than just showing that the marginals and thus the expected distances converge.

Theorem 2. *The Markov chain defined in Theorem 1 has a unique stationary distribution.*

Pareto-Efficiency. The outcome of sequential deliberation is ex-post Pareto-efficient on a median graph. In other words, in any realization of the random process, suppose the final outcome is o; then there is no other alternative a such that $d(a, v) \leq d(o, v)$ for every $v \in \mathcal{N}$, with at least one inequality being strict. This is a weak notion of efficiency, but it is not trivial to show; while it is easy to see that a one shot bargaining mechanism using only bliss points is Pareto efficient by virtue of the Pareto efficiency of bargaining, sequential deliberation defines a potentially complicated Markov chain for which many of the outcomes need not be bliss points themselves.

Theorem 3. *Sequential deliberation among a set $\mathcal{N}$ of agents, where the decision space $\mathcal{S}$ is a median graph and the initial disagreement point a^1 is the bliss point of some agent, yields an ex-post Pareto Efficient alternative.*

Truthfulness of Extensive Forms. Finally, we show that sequential deliberation has truth-telling as a sub-game perfect Nash equilibrium in its induced extensive form game. Towards this end, we formalize a given round of bargaining as a 2-person non-cooperative game between two players who can choose as a strategy to report any point v on a median graph; the resulting outcome is the median of the two strategy points chosen by the players and the disagreement alternative presented. The payoffs to the players are just the utilities already defined; i.e., the player wishes to minimize the distance from their true bliss point to the outcome point. Call this game the non-cooperative bargaining game (NCBG).

The extensive form game tree defined by non-cooperative bargaining consists of $2T$ alternating levels: Nature draws two agents at random, then the two agents play NCBG and the outcome becomes the disagreement alternative for the next NCBG. The leaves of the tree are a set of points in the median graph; agents want to minimize their expected distance to the final outcome.

Theorem 4. *Sequential NCBG on a median graph has a sub-game perfect Nash equilibrium where every agent truthfully reports their bliss point at all rounds of bargaining.*

5 General Metric Spaces

We now work in the very general setting that the set $\mathcal{S}$ of alternatives are points in a finite metric space equipped with a distance function $d(\cdot)$ that is a metric. As before, we assume each agent $u \in \mathcal{N}$ has a bliss point $p_u \in \mathcal{S}$. An agent's disutility for an alternative $a \in \mathcal{S}$ is simply $d(p_u, a)$. We first present results for the Distortion, and subsequently define the second moment, or Squared-Distortion. For both measures, we show that the upper bound for sequential deliberation is at most a constant regardless of the metric space. All proofs are given in the full version of the paper [12].

Theorem 5. *The Distortion of sequential deliberation is at most 3 when the space of alternatives and bliss points lies in some metric, and this bound is tight.*

The bound of 3 above is quite pessimistic. The metric space employed in the lower bound is contrived in the following sense: Every pair of agents has some unique (to that pair) alternative they very slightly prefer to the social optimum. For structured spaces, we expect the bound to be much better. We have already shown this for median spaces. In the full version of this paper [12], we provide more evidence in this direction by considering a structured space motivated by budgeting applications that is not a median graph. For this space, we show that sequential deliberation has Distortion at most 4/3.

5.1 Second Moment of Social Cost

We now show that for any metric space, sequential deliberation has a crucial advantage in terms of the distribution of outcomes it produces. For this, we consider the second moment, or the expected squared social cost. Recall that the *social cost* of an alternative $a \in \mathcal{S}$ is given by $SC(a) = \sum_{u \in \mathcal{N}} d(p_u, a)$. Let $a^* \in \mathcal{S}$ be the minimizer of social cost, *i.e.*, the *generalized median.* Then define:

$$\text{Squared-Distortion} = \frac{\mathbb{E}[(SC(a))^2]}{(SC(a^*))^2}$$

where the expectation is over the set of outcomes a produced by Sequential Deliberation.[4] We show that sequential deliberation has Squared-Distortion upper bounded by a constant. This means the standard deviation in social cost of the distribution of outcomes is comparable to the optimal social cost. This has a practical implication: A policy designer can run sequential deliberation for a few steps, and be sure that the probability of observing an outcome that has γ times the optimal social cost is at most $O(1/\gamma^2)$. In contrast, Random Dictatorship (choosing an agent uniformly at random and using her bliss point as the solution) has unbounded Squared-Distortion, which means its standard deviation in social cost cannot be bounded. In other words, deliberation between agents eliminates the outlier agent, and concentrates probability mass on central outcomes.

Theorem 6. *The Squared-Distortion of sequential deliberation for $T \geq 1$ is at most 41 when the space of alternatives and bliss points lies in some metric. Furthermore, the Squared-Distortion of random dictatorship is unbounded.*

6 Open Questions

Our work is the first step to developing a theory around practical deliberation schemes. We suggest several future directions. First, we do not have a general characterization of the Distortion of sequential deliberation for metric spaces. We have shown that for general metric spaces there is a small but pessimistic bound on the Distortion of 3, but that for specific metric spaces the Distortion may be much lower. We do not have a complete characterization of what separates these good and bad regimes.

More broadly, an interesting question is extending our work to take opinion dynamics into account, *i.e.*, proving stronger guarantees if we assume that when two agents deliberate, each agent's opinion moves slightly towards the other agent's opinion and the outside alternative. Furthermore, though we have shown that all agents deliberating at the same time does not improve on dictatorship, it is not clear how to extend our results to more than two agents negotiating at the same time. This runs into the challenges in understanding and modeling multiplayer bargaining [19, 24, 6].

Finally, it would be interesting to conduct experiments to measure the efficacy of our framework on complex, real world social choice scenarios. There are several practical hurdles that need to be overcome before such a system can be feasibly deployed. In a related sense, it would be interesting to develop an axiomatic theory for deliberation, much like that for bargaining [29], and show that sequential deliberation arises naturally from a set of axioms.

[4] The motivation for considering Squared-Distortion instead of the standard deviation is that the latter might prefer a more deterministic mechanism with a worse social cost, a problem that the Squared-Distortion avoids.

References

1. Airiau, S., Endriss, U.: Iterated majority voting. In: International Conference on Algorithmic Decision Theory. pp. 38–49. Springer (2009)
2. Anshelevich, E., Bhardwaj, O., Postl, J.: Approximating optimal social choice under metric preferences. In: AAAI. vol. 15, pp. 777–783 (2015)
3. Anshelevich, E., Postl, J.: Randomized social choice functions under metric preferences. 25th International Joint Conference on Artificial Intelligence (2016)
4. Bandelt, H.J., Barthelemy, J.P.: Medians in median graphs. Discrete Applied Mathematics 8(2), 131 – 142 (1984)
5. Barbera, S., Gul, F., Stacchetti, E.: Generalized median voter schemes and committees. Journal of Economic Theory 61(2), 262 – 289 (1993)
6. Bennett, E., Houba, H.: Odd man out: Bargaining among three players. Working Papers 662, UCLA Department of Economics (May 1992), http://www.econ.ucla.edu/workingpapers/wp662.pdf
7. Binmore, K., Shaked, A., Sutton, J.: Testing noncooperative bargaining theory: A preliminary study. The American Economic Review 75(5), 1178–1180 (1985)
8. Binmore, K., Rubinstein, A., Wolinsky, A.: The nash bargaining solution in economic modelling. The RAND Journal of Economics 17(2), 176–188 (1986)
9. Boutilier, C., Caragiannis, I., Haber, S., Lu, T., Procaccia, A.D., Sheffet, O.: Optimal social choice functions: A utilitarian view. Artificial Intelligence 227, 190–213 (2015)
10. Cheng, Y., Dughmi, S., Kempe, D.: Of the people: Voting is more effective with representative candidates. In: EC (2017)
11. Clearwater, A., Puppe, C., Slinko, A.: Generalizing the single-crossing property on lines and trees to intermediate preferences on median graphs. In: Proceedings of the 24th International Conference on Artificial Intelligence. pp. 32–38. IJCAI'15 (2015)
12. Fain, B., Goel, A., Munagala, K., Shaksuwong, S.: Sequential deliberation for social choice (2017), https://arxiv.org/abs/1710.00771
13. Feldman, M., Fiat, A., Golomb, I.: On voting and facility location. In: Proceedings of the 2016 ACM Conference on Economics and Computation. pp. 269–286. EC '16, ACM, New York, NY, USA (2016)
14. Fishkin, J., Luskin, R.: Experimenting with a democratic ideal: Deliberative polling and public opinion. Acta Politica 40(3), 284298 (2005)
15. Garg, N., Kamble, V., Goel, A., Marn, D., Munagala, K.: Collaborative optimization for collective decision-making in continuous spaces. In: Proc. World Wide Web (WWW) Conference (2017)
16. Goel, A., Lee, J.: Towards large-scale deliberative decision-making: Small groups and the importance of triads. In: ACM EC (2016)
17. Goel, A., Krishnaswamy, A.K., Munagala, K.: Metric distortion of social choice rules. In: EC (2017)
18. Gross, S., Anshelevich, E., Xia, L.: Vote until two of you agree: Mechanisms with small distortion and sample complexity. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA. pp. 544–550 (2017)
19. Harsanyi, J.C.: A simplified bargaining model for the n-person cooperative game. International Economic Review 4(2), 194–220 (1963)
20. Hylland, A., Zenkhauser, R.: A mechanism for selecting public goods when preferences must be elicited. Working Paper (1980)

21. Kalai, E.: Proportional solutions to bargaining situations: Interpersonal utility comparisons. Econometrica 45(7), 1623–1630 (1977)
22. Kalai, E., Smorodinsky, M.: Other solutions to nash's bargaining problem. Econometrica 43(3), 513–518 (1975)
23. Knuth, D.E.: The Art of Computer Programming: Combinatorial Algorithms, Part 1. Addison-Wesley Professional, 1st edn. (2011)
24. Krishna, V., Serrano, R.: Multilateral bargaining. The Review of Economic Studies 63(1), 61–80 (1996)
25. Lang, J., Xia, L.: Voting in combinatorial domains. In: Brandt, F., Conitzer, V., Endriss, U., Lang, J., Procaccia, A.D. (eds.) Handbook of Computational Social Choice. Cambridge University Press (2016)
26. Lev, O., Rosenschein, J.S.: Convergence of iterative voting. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2. pp. 611–618. International Foundation for Autonomous Agents and Multiagent Systems (2012)
27. Meir, R., Polukarov, M., Rosenschein, J.S., Jennings, N.R.: Convergence to equilibria in plurality voting. In: Proc. of 24th Conference on Artificial Intelligence (AAAI-10). pp. 823–828 (2010)
28. Myerson, R.B.: Two-person bargaining problems and comparable utility. Econometrica 45(7), 1631–1637 (1977)
29. Nash, J.F.: The bargaining problem. Econometrica 18(2), 155–162 (1950), http://www.jstor.org/stable/1907266
30. Neelin, J., Sonnenschein, H., Spiegel, M.: An Experimental Test of Rubinstein's Theory of Bargaining. Working Papers 587, Princeton University, Department of Economics, Industrial Relations Section (May 1986)
31. Nehring, K., Puppe, C.: The structure of strategy-proof social choice. part i: General characterization and possibility results on median spaces. Journal of Economic Theory 135(1), 269 – 305 (2007)
32. Rabinovich, Z., Obraztsova, S., Lev, O., Markakis, E., Rosenschein, J.S.: Analysis of equilibria in iterative voting schemes. In: AAAI. vol. 15, pp. 1007–1013. Citeseer (2015)
33. Roth, A.E.: Bargaining phenomena and bargaining theory. In: Roth, A.E. (ed.) Laboratory Experiments in Economics: Six Points of View, pp. 14–41. Cambridge University Press (1987)
34. Rubinstein, A.: Perfect equilibrium in a bargaining model. Econometrica 50(1), 97–109 (1982), http://www.jstor.org/stable/1912531
35. Saban, D., Stier-Moses, N.: The Competitive Facility Location Problem in a Duopoly: Connections to the 1-Median Problem, pp. 539–545 (2012)
36. Schummer, J., Vohra, R.V.: Strategy-proof location on a network. Journal of Economic Theory 104(2), 405 – 428 (2002)
37. Thompson, D.F.: Deliberative democratic theory and empirical political science. Annual Review of Political Science 11(1), 497–520 (2008)
38. Wajc, D., Procaccia, A., Zhang, H.: Approximation-variance tradeoffs in mechanism design (January 2017)
39. Wendell, R.E., McKelvey, R.D.: New perspectives in competitive location theory. European Journal of Operational Research 6(2), 174–182 (1981)