

# Approximate Core for Committee Selection via Multilinear Extension and Market Clearing

Kamesh Munagala\*    Yiheng Shen\*    Kangning Wang\*    Zhiyi Wang\*

## Abstract

Motivated by civic problems such as participatory budgeting and multiwinner elections, we consider the problem of *public good* allocation: Given a set of indivisible projects (or candidates) of different sizes, and voters with different monotone utility functions over subsets of these candidates, the goal is to choose a budget-constrained subset of these candidates (or a committee) that provides fair utility to the voters. The notion of fairness we adopt is that of *core stability* from cooperative game theory: No subset of voters should be able to choose another *blocking committee* of proportionally smaller size that provides strictly larger utility to all voters that deviate. The core provides a strong notion of fairness, subsuming other notions that have been widely studied in computational social choice.

It is well-known that an exact core need not exist even when utility functions of the voters are additive across candidates. We therefore relax the problem to allow *approximation*: Voters can only deviate to the blocking committee if after they choose any extra candidate (called an *additament*), their utility still increases by an  $\alpha$  factor. If no blocking committee exists under this definition, we call this an  $\alpha$ -core.

Our main result is that an  $\alpha$ -core, for  $\alpha < 67.37$ , always exists when utilities of the voters are arbitrary monotone submodular functions, and this can be computed in polynomial time. This result improves to  $\alpha < 9.27$  for additive utilities, albeit without the polynomial time guarantee. Our results are a significant improvement over prior work that only shows logarithmic approximations for the case of additive utilities. We complement our results with a lower bound of  $\alpha > 1.015$  for submodular utilities, and a lower bound of any function in the number of voters and candidates for general monotone utilities.

## 1 Introduction

Consider the following scenario: A city is deciding what public projects to fund using its limited budget  $b$ . There is a list of  $m$  candidate projects (forming a set  $C$ ), where each  $j \in C$  is associated with a cost  $s_j$ . The city needs to select a subset  $O$  of these projects whose total cost is at most the budget, that is,  $\sum_{j \in O} s_j \leq b$ . There are  $n$  residents or *voters* (forming a set  $V$ ) in the city, and each of them has preferences on how the city should spend its budget. These preferences may not be perfectly compatible with each other: For instance, families with children may prefer a public school, while others may prefer a park; people living in the east may prefer projects there, and so for those living in the west. It is desirable to have a fair process to decide on the projects to pay for. In participatory budgeting, voters express their preferences through their votes and influence the decision process. The paradigm has been implemented in numerous cities across the world [13, 4, 11, 31].

A similar problem is seen in multiwinner elections [7, 24, 11, 15, 41], where voters select a committee of size  $b$  from  $m$  candidates. Each voter holds her own opinions on the committees, and a fair method to incorporate the preferences of all voters is called for. Mathematically, it is a special case of the participatory budgeting problem, where each candidate has the same “cost”.

In these settings, simple methods to aggregate the preferences may have drawbacks. In majority voting, the utilities of a coherent minority group might be entirely ignored in favor of the majority. Utilitarian ways (maximizing the sum of utilities) may overly focus on a certain group, disregarding the welfare of the vast majority. Indeed, many practically implemented voting rules for choosing parliaments and civic body members, such as the well-known Single Transferable Voting (STV) [42], attempt to address precisely this issue. We naturally ask:

---

\*Department of Computer Science, Duke University, Durham, NC 27708-0129. Emails: kamesh@cs.duke.edu, yiheng.shen@duke.edu, knwang@cs.duke.edu, zhiyi.wang@duke.edu. This work is supported by NSF grants CCF-1637397 and CCF-2113798, ONR award N00014-19-1-2268, and DARPA award FA8650-18-C-7880.

What is a fair solution for participatory budgeting and multiwinner elections, and how do we reach such a solution?

**1.1 The Core and Its Multiplicative Approximation** Recall that there are  $n$  voters forming a set  $V$ , and  $m$  candidates forming a set  $C$ , where candidate  $j$  has size (or cost)  $s_j$ . We need to choose a subset  $O$  of candidates with total size at most  $b$  (that is,  $\sum_{j \in O} s_j \leq b$ ). Following social choice terminology, this subset of candidates is called a *committee*, and the problem of choosing a budget-constrained committee is called the *committee selection* problem. Denote the utility of voter  $i$  for a committee  $T$  by a *utility function*  $u_i(T)$ . We will assume this function is non-negative and monotone, with  $u_i(\emptyset) = 0$ .

Although there are copious notions of fairness for committee selection, the *core* is a classic and influential one among them. This idea has existed for more than a century [21, 41, 34], and serves as a strong notion of proportional representation. Towards defining this concept, imagine we split the size  $b$  among the voters, so that each voter has an *endowment* of  $\frac{b}{n}$  that they can use to “buy” candidates. A candidate of size  $s_j$  requires an endowment of  $s_j$  to “buy”. A committee  $O \subseteq C$  with total size at most  $b$  is said to be in the core, if no subset  $S$  of voters can deviate and purchase another committee  $T \subseteq C$  by pooling their endowments, so that each voter in  $S$  prefers the new committee  $T$  to the original one  $O$ . Note that the total endowment of  $S$  is  $|S|\frac{b}{n}$ , so that this set of voters can buy a committee  $T$  of size at most  $|S|\frac{b}{n}$ . Formally,

**DEFINITION 1.1. (CORE)** *A committee  $O$  is in the core if there is no  $S \subseteq V$  and committee  $T \subseteq C$  with  $\sum_{j \in T} s_j \leq \frac{|S|}{n} \cdot b$ , such that  $u_i(T) > u_i(O)$  for every  $i \in S$ .*

The core has a “fair taxation” interpretation [34, 28]. The quantity  $\frac{b}{n}$  can be thought of as the tax contribution of a voter, and a committee in the core has the property that no sub-group of voters could have spent their share of tax money in a way that *all* of them were better off. As such it subsumes notions of fairness such as Pareto-optimality, proportionality, and various forms of justified representation [5, 27, 6] that have been extensively studied in multiwinner election and fairness literature.

Despite the satisfying properties of the core, its strength is also its limitation: Even in the simple setting of unit sizes, integer budget, and additive utilities (the so-called *approval-set* setting with general utilities), the core can be empty. (See for example, [26].)

A natural approach to circumvent this problem is to show the existence of a committee that *multiplicatively* approximates the core. We define the  $\alpha$ -core as follows.

**DEFINITION 1.2. ( $\alpha$ -CORE)** *A committee  $O$  is in the  $\alpha$ -core if there is no  $S \subseteq V$  and  $T \subseteq C$  with  $\sum_{j \in T} s_j \leq \frac{|S|}{n} \cdot b$ , such that  $u_i(T) > \alpha \cdot u_i(O \cup \{q\})$  for every  $i \in S$  and  $q \in C$ . We call the candidate  $q$  an additament.*

Note that we introduce the additament in the definition, since no multiplicative approximation is possible without it even in the setting with unit candidate sizes and additive utilities. This follows from examples in previous work [26, 19]. The idea of a bicriteria (multiplicative and additive) approximation to utilities was first presented in [26]. The work of [40] presents an almost identical definition as [Definition 1.2](#) except that the additament  $q$  must come from the set  $T$ . This makes their definition more restrictive and the  $\alpha$ -core smaller. They show that when utilities are additive, an  $O\left(\log \frac{u_{\max}}{u_{\min}}\right)$ -core solution not only exists, but can also be computed in polynomial time, where  $u_{\max}$  and  $u_{\min}$  are the largest and smallest non-zero utilities any voter has for any feasible committee.

**1.2 Our Results** Our main result in [Section 3](#) is to show the existence of an  $O(1)$ -core when the utility functions of the voters are monotone and submodular. Submodular utilities capture the notion of “diminishing returns”, and are therefore well-motivated in participatory budgeting and multiwinner election settings where each additional project or candidate provides diminishing marginal utility.

Our result is a significant improvement over the previous work mentioned above [40], which only presents a logarithmic approximation for the restricted case of additive utilities. Our main result is formally the following.

**THEOREM 1.1. (MAIN THEOREM)** *For monotone submodular utilities, a 67.37-core is always non-empty. One such solution can be computed in polynomial time.*

Our second result in [Section 4](#) is the following improved result for the well-studied special case of additive utilities:

**THEOREM 1.2. (ADDITIVE UTILITIES)** *For additive utilities, a 9.27-core is always non-empty.*

Unlike [Theorem 1.1](#), we do not know how to implement the algorithm in [Theorem 1.2](#) in polynomial time. We remark that the previous two results can be combined to show that a 15.2-core solution for additive utilities can be computed in polynomial time, without providing the details.

**Lower Bounds.** The next natural question is whether our results can be extended to arbitrary monotone utilities. We show that this is not possible. We present an example in [Section 5](#) to show that an  $O(1)$ -core (or even any  $f(n, m)$ -core where  $n$  is the number of voters and  $m$  is the number of candidates) may not exist for general monotone utilities.

**THEOREM 1.3. (GENERAL UTILITY LOWER BOUND)** *For general monotone utilities, for any function  $\varphi : \mathbb{Z}^+ \times \mathbb{Z}^+ \rightarrow \mathbb{R}^+$ , a  $\varphi(n, m)$ -core can be empty.*

We also show that the case of submodular utilities does need a multiplicative approximation to the core, and the 1-core can be empty. This justifies the form of [Theorem 1.1](#) that involves both a multiplicative approximation and an additament.

**THEOREM 1.4. (SUBMODULAR LOWER BOUND)** *For monotone submodular utilities, a 1.015-core can be empty.*

**1.3 Technical Contributions** The main technical idea in our paper is to consider *fractional relaxations* of the problem, via considering outcomes where candidates are fractionally allocated. For submodular utilities, we extend the utility to fractional values via the so-called multilinear extension [\[43, 14\]](#). For additive utilities, the extension is straightforward. We then construct an exact or approximate core outcome for the fractional version. We finally round this solution to find an integer core. We now describe these steps in more detail.

**Fractional Core.** We use two entirely different techniques for constructing fractional core outcomes to show [Theorem 1.1](#) and [Theorem 1.2](#). For both solutions, we extend the standard core property and show a stronger *approximation property* ([Theorem 3.1](#) and [Lemma 4.3](#)) on how the size of a coalition of voters relates to the budget of the committee to which they can deviate to and obtain a multiplicative factor larger utility.

The first approach that we use for arbitrary submodular functions is via optimizing the classic Nash Welfare objective [\[37\]](#), that maximizes the product of utilities of the voters. Though the Nash Welfare is a convex optimization problem when utilities are continuous, monotone, and concave, in our case, the utility is now defined as the multilinear extension and this is only concave in positive directions [\[14\]](#). Nevertheless, we show a simple continuous local search procedure for the Nash Welfare objective, such that the local optimum is a 2-approximate fractional core. For our rounding, we need a stronger approximation property that we prove in [Theorem 3.1](#). This property may be of independent interest.

The second approach that we use for additive utilities, finds an *exact* fractional core. This uses a classic market clearing solution is termed the *Lindahl equilibrium* [\[34, 28\]](#). In this equilibrium, candidates can be chosen to any fraction (even greater than one). Each candidate is assigned a per-voter “price”, and the voters are assigned an endowment of  $\frac{b}{n}$  just as in [Definition 1.1](#). At these prices, if the voters choose their utility maximizing allocation subject to spending their endowment, then (1) they all choose the same fractional committee, and they spend their entire endowment; and (2) for each chosen candidate, the total price collected exactly pays for that candidate. The work of [\[28\]](#) shows via a fixed point argument that such an equilibrium always exists when each voter’s utility for fractional committees is any continuous, monotone, and concave function. Further, the resulting equilibrium is a core outcome.

**Rounding.** We then construct the integer solution by rounding of the fractional core. Since the allocation is common to all voters, it is easier to work with rounding processes that are oblivious to the utility functions, and we therefore simply use randomized rounding. If we randomly round the allocation to integer values, the expected utility of any voter is preserved. This can now be combined with Chernoff-type bounds for multilinear extensions [\[16\]](#) to argue that a constant factor of voters see utility that is at least a constant factor of the corresponding utility in the fractional solution. For additive utilities, we use dependent rounding [\[30, 12\]](#) to provide an improved constant here.

This brings up the next hurdle: The remaining unsatisfied set of voters can still deviate to a different coalition. Our second ingredient is the iterative rounding framework of [32]: We recurse on the unsatisfied voters to find a fractional solution with a smaller budget on total size. The key idea is that every voter is satisfied at some level of the recursion. We now use the approximation property of the fractional solution at each level of recursion to argue that the number of voters deviating to any fixed committee that provides them a constant factor more utility is bounded. Summing this argument over all levels completes the proof.

Our overall argument in both cases – submodular and additive utilities – is quite delicate, and crucially requires strong approximation properties of the fractional solutions of Lindahl equilibrium and locally optimal Nash Welfare. This forms our main technical contribution.

## 1.4 Related Work

**Proportionality and the Core.** One classic objective in committee selection is achieving fairness via proportionality, where different demographic slices of voters feel they have been represented fairly. This general idea dates back more than a century [21], and has recently received significant attention [15, 35, 10, 5, 27, 6]. In fact, there are several elections, both at a group level and a national level, that attempt to find committees (or parliaments) that provide approximately proportional representation. For instance, the popular Single Transferable Vote (STV) rule is used in parliamentary elections in Ireland and Australia, and in several municipal elections in the USA. This rule attempts to find a proportional solution.

A long line of recent literature has studied the complexity and axiomatization of voting rules that achieve proportionality; see [7, 24, 11] for recent surveys. Proportionality in committee selection arises in many other applications outside of social choice as well. For example, consider a shared cache for data items in a multi-tenant cloud system, where each data item is used by several long-running applications [33, 29]. Each data item can be treated as a candidate, and each application as a voter whose utility for an item corresponds to the speedup obtained by caching that item. In this context, a desirable caching policy provides proportional speedup to all applications.

The core represents the ultimate form of proportionality: Every demographic of voters feel that they have been fairly represented and do not have the incentive to deviate and choose their own committee of proportionally smaller size which gives all of them higher utility. In the typical setting where these demographic slices are not known upfront, the notion of core attempts to be fair to all subsets of voters. The work of [36] formally argues that in certain multiwinner election settings, the core also approximately optimizes simpler diversity measures of the resulting committee.

**Fisher Markets.** Our fractional solutions are superficially related to the Fisher market equilibrium [9, 37, 3] when divisible items need to be *allocated* to agents, and agents' utilities are additive. For the Fisher market, the optimum Nash Welfare solution finds market clearing prices. However, in a Fisher market, the prices are common to the agents while the allocations are different, while in a Lindahl equilibrium, the prices are per-voter while the allocation (or committee) is common and provides shared utility to all the voters. This is a key difference – the Fisher market has a polynomial time algorithm via convex programming [23], while no polynomial time algorithm is known for the Lindahl equilibrium even when candidates have unit sizes and voters' utilities are additive (or linear till the maximum size of the candidate). Similarly, though the Nash Welfare solution finds market clearing prices for the Fisher market via strong duality, in the case of public goods, there is no obvious way to interpret the dual of the Nash Welfare solution as market clearing prices. Moreover, for submodular utilities and multilinear extensions, the Nash Welfare objective is no longer a convex program, so that strong duality does not apply.

**Approximate Core.** In this paper, we have focused on approximating the utility voters obtain on deviating (see Definition 1.2). As mentioned before, this notion first appeared in [26], and the notion of a single additament in approximation is due to [40]. The latter work present a logarithmic approximation for the special case of additive utilities. The work of [39] shows that the well-known Proportional Approval Voting method [41] achieves a 2-core for the special case where the utilities are additive and candidates are unit size, with each voter having utility either zero or one for each candidate. This algorithm can be viewed as a discrete version of Nash Welfare, and in essence, Theorem 3.1 extends this result to the case of submodular utilities and general costs, showing that it yields a 2-core for the *fractional* case of multilinear extension via a polynomial time local search algorithm. The work of [18] presents a constant approximation for the  $K$ -clustering problem, where the committee is a set of  $K$  centers in a metric space, and the cost of a voter is the distance to the closest center. However, these ideas do not extend to the committee selection problem we consider in this paper.

The work of [32] considers a different notion of approximation: Instead of approximating the utility, they approximate the *endowment* that a voter can use to buy the deviating committee. Building on the work of [19], it shows a different fractional relaxation, to which a 2-approximation always exists. They then iteratively round this fractional solution to an integer solution that is a 32-approximation for *all* monotone utility functions. The problem of approximating utilities is very different; indeed, Theorem 1.3 shows we cannot hope to have a similar constant approximation for all utility functions. Nevertheless, we use the idea of iterative rounding from that work, albeit with an entirely different fractional solution and analysis. In effect, we showcase the power of iterative rounding as a unifying framework for finding approximate core solutions, regardless of the notion of approximation.

**Rounding Techniques.** The notion of multilinear extension and correlation gap has been widely used in stochastic optimization [2], mechanism design [44], and rounding [43, 14, 17]. Typically, it has been used to develop computationally efficient approaches; on the other hand, we demonstrate an application to showing a purely existential result. Similarly, rounding of market clearing solutions have been used to show approximately fair allocations of indivisible goods among agents [8, 20]. The structure of these problems (common prices but different allocations) is very different from ours (common allocations and different prices), and we need different techniques. Again, in contrast with the resource allocation literature, we need the rounding just to show an existence result as opposed to a computational one.

## 2 Model and Preliminaries

Recall that  $C$  is a set of  $m$  candidates and  $V$  is a set of  $n$  voters. Each candidate  $j \in C$  is associated with a size  $s_j > 0$ . For each  $i \in V$  and each  $T \subseteq C$ , we denote by  $u_i(T)$  the utility of voter  $i$  on committee  $T$ . We assume  $u_i$  is a monotone, submodular set function with  $u_i(\emptyset) = 0$ :

- Monotonicity:  $u_i(T) \leq u_i(T') \quad \forall T \subseteq T'$ .
- Submodularity:  $u_i(T \cup \{t\}) - u_i(T) \geq u_i(T' \cup \{t\}) - u_i(T') \quad \forall T \subseteq T', t \notin T'$ .

We have a constraint of  $b$  on the total size of the committee. We call a committee  $O \subseteq C$  feasible (with respect to  $b$ ) if:  $\sum_{j \in O} s_j \leq b$ .

Our goal is to find a committee in the  $\alpha$ -core for  $\alpha = O(1)$ . Recall its definition:

**DEFINITION 1.2. ( $\alpha$ -CORE)** *A committee  $O$  is in the  $\alpha$ -core if there is no  $S \subseteq V$  and  $T \subseteq C$  with  $\sum_{j \in T} s_j \leq \frac{|S|}{n} \cdot b$ , such that  $u_i(T) > \alpha \cdot u_i(O \cup \{q\})$  for every  $i \in S$  and  $q \in C$ . We call the candidate  $q$  an additament.*

**Fractional Allocations and Core.** In the sequel, we will consider continuous extensions of the committee and the utility function. We define a *fractional committee* as a  $m$ -dimensional vector  $\mathbf{x} \geq 0$ . The quantity  $x_j$  denotes the fraction to which candidate  $j$  is allocated.

We denote by  $\text{Cost}(\mathbf{x}) := \sum_{j \in C} x_j \cdot s_j$  as the cost of the allocation  $\mathbf{x}$ . Without loss of generality, we assume that  $\sum_{j \in C} s_j > b$ .

Similarly, we consider continuous utilities  $U_i(\mathbf{x})$  for the voters, whose construction will be specified later. Given this notation, an  $\alpha$ -approximate fractional core, for  $\alpha \geq 1$  is defined as follows:

**DEFINITION 2.1. ( $\alpha$ -APPROXIMATE FRACTIONAL CORE)** *A fractional committee  $\mathbf{x} \geq 0$  with  $\text{Cost}(\mathbf{x}) \leq b$  lies in the  $\alpha$ -approximate fractional core if there is no  $S \subseteq V$  and allocation  $\mathbf{z} \geq 0$  with  $\text{Cost}(\mathbf{z}) \leq \frac{|S|}{|V|} \cdot b$  such that  $U_i(\mathbf{z}) > \alpha U_i(\mathbf{x})$  for all  $i \in S$ .*

When  $\alpha = 1$ , we call the fractional committee  $\mathbf{x}$  as simply lying in the *fractional core*.

## 3 Constant Approximate Core for Submodular Utilities

Our overall algorithm proceeds via constructing approximate fractional core for a well-known continuous relaxation, called the *multilinear relaxation*. We use a continuous time local search procedure for the Nash Welfare objective to find this fractional solution. We show that it is almost in a 2-approximate fractional core, and subsequently round it iteratively to find a solution in the approximate integer core. Our overall algorithm runs in time polynomial in the number of voters and candidates. We first present the entire algorithm as an existence proof, and then delve into the details of the running time in [Section 3.5](#).

**3.1 First Observations** The following steps are without loss of generality, aiming to simplify future derivations in this section. We will assume  $\varepsilon \in (0, 1/20)$  is a small constant that we choose later. In a proof of existence we can set  $\varepsilon = 0$ ; however, we need  $\varepsilon > 0$  to achieve polynomial running time.

- Let  $u_i^{\max} := \max_j \{u_i(\{j\})\}$ , the maximum utility of voter  $i$  with a single item. We scale down every  $u_i(T)$  by a factor of  $u_i^{\max}$  and denote this normalized utility as  $u'_i(T) := \frac{u_i(T)}{u_i^{\max}}$ . (We ignore every voter with  $u_i(C) = 0$ , since they will not deviate in any situation.) By submodularity of the utility functions,  $u'_i(T) \in [0, m]$  for every  $i$  and  $T \subseteq C$ . This step preserves the original problem, since every voter multiplicatively compares utilities of different committees in [Definition 1.2](#) and different voters compare separately. Therefore, without loss of generality, we assume  $\max_j \{u_i(\{j\})\} = 1$  and thus  $u_i(T) \in [0, m], \forall T \subseteq C$ .
- We further assume  $s_j \leq b$  for each  $j \in C$ , as otherwise, the candidate cannot be selected in any feasible committee.
- Now define

$$C_s := \{j \mid s_j \leq \varepsilon b/m\}$$

as a set of “small enough” candidates. We place all these candidates in the solution. Clearly all candidates in  $C_s$  take up a total budget of at most  $\varepsilon b$  – the remaining budget is at least  $(1 - \varepsilon)b$ . Our algorithm then works with candidates in:

$$C_\ell := C \setminus C_s = \{j \mid s_j > \varepsilon b/m\}.$$

Note that each candidate  $j$  in  $C_\ell$  now has size  $s_j > \varepsilon b/m$ .

**3.2 The 2-Approximate Fractional Core** We now present a procedure that finds an approximate fractional core solution for an arbitrary budget  $B \in [\frac{\varepsilon}{5m}b, b]$ , and an arbitrary subset  $W \subseteq V$  of voters. We present the algorithm and its optimality properties below; however, we defer the polynomial running time analysis to [Section 3.5](#).

**3.2.1 Multilinear Extension** Recall that the utility function  $u_i(S)$  of voter  $i$  for  $S \subseteq C$  is non-negative, monotone, and submodular. We first extend these utilities to fractional allocations. A natural way to do that is to use the multilinear extension.

**DEFINITION 3.1.** (MULTILINEAR EXTENSION, [\[14, 43\]](#)) *Given a monotone, non-negative submodular function  $f$ , its multilinear extension  $F$  is defined for any  $\mathbf{x} \in [0, 1]^m$  as:*

$$F(\mathbf{x}) = \sum_{T \subseteq C} f(T) \left( \prod_{j \in T} x_j \right) \left( \prod_{j \notin T} (1 - x_j) \right).$$

We will apply the following property of this function.

**LEMMA 3.1.** (CONCAVITY ALONG POSITIVE DIRECTIONS, [\[14\]](#)) *Given a monotone, non-negative submodular function  $f$ , its multilinear extension  $F$  is concave along positive directions, i.e., for all  $\mathbf{x}_0$  and all  $\mathbf{d} \geq \mathbf{0}$ , we have that  $F_{\mathbf{x}_0, \mathbf{d}}(\lambda) = F(\mathbf{x}_0 + \lambda \mathbf{d})$  is a monotone, non-negative and concave function of  $\lambda$ .*

We denote the multilinear extension of voter  $i$ ’s utility function  $u_i$  for a fractional committee  $\mathbf{x}$  as  $U_i(\mathbf{x})$ . The above lemma implies that  $\frac{\partial U_i}{\partial x_j} \geq 0$  for all  $i \in W, j \in C$ . The next lemma upper bounds the gradient. Here,  $\mathbf{x}_{-j}$  is  $\mathbf{x}$  with the  $j^{\text{th}}$  dimension removed, and  $T \sim \mathbf{x}_{-j}$  means that  $T$  is chosen by including  $\ell \in C \setminus \{j\}$  independently with probability  $x_\ell$ .

**LEMMA 3.2.** (BOUNDED GRADIENT, [\[14\]](#))

$$\frac{\partial U_i(\mathbf{x})}{\partial x_j} = \sum_{T \subseteq C \setminus \{j\}} \Pr_{T \sim \mathbf{x}_{-j}} [T] \cdot (u_i(T \cup \{j\}) - u_i(T)) \leq \max_j u_i(\{j\}) = 1.$$

**3.2.2 Nash Welfare and Local Search** The input of the algorithm consists of three parts: The utility function  $u_i(T)$  for each committee  $T \subseteq C$  and voter  $i \in W$ , the sizes of the candidates  $\{s_j\}$  and the total budget  $B \in [\frac{\varepsilon}{5m}b, b]$  on candidates in  $C_\ell$ . Recall that we assume all candidates in  $C_s$  have already been chosen. We assume w.l.o.g. that  $\sum_{j \in C_\ell} s_j > B$ . In our algorithm, we treat the utility function as an oracle which returns the utility of an agent  $i$  on a given committee in  $O(1)$  time.

In our algorithm, we will lower bound the allocation  $x_j$  by an amount  $\underline{x}_j$  that is defined as follows.

$$\underline{x}_j = \begin{cases} 1 & j \in C_s, \\ \frac{B \cdot \varepsilon}{\sum_{j \in C_\ell} s_j} & j \in C_\ell. \end{cases}$$

**Nash Welfare.** We will find a local optimum to the following *Nash Welfare* program to find the fractional solution.

$$\begin{aligned} & \text{Maximize } \sum_{i \in W} \log U_i(\mathbf{x}) \\ & \sum_{j \in C_\ell} s_j x_j \leq B \\ & x_j \in [\underline{x}_j, 1] \quad \forall j \in C \end{aligned}$$

Let  $\phi_i(\mathbf{x}) = \log U_i(\mathbf{x})$ , and  $\phi(\mathbf{x}) = \sum_{i \in W} \phi_i(\mathbf{x})$ . This is the objective value of the above program.

**Local Search.** We call the local search procedure as  $\text{NW}(C, W, \{U_i\}, B)$ . This procedure has a *step-size* parameter  $\delta = \frac{\varepsilon^7 b}{312m^6}$ . We start with any allocation  $\mathbf{x}_0 \in [\underline{x}_j, 1]^m$  with  $\sum_{j \in C_\ell} s_j x_j = B$ . Note that such an allocation always exists since we assumed  $\sum_{j \in C_\ell} s_j > B$ , and further,  $\sum_{j \in C_\ell} s_j \underline{x}_j = \varepsilon B$ .

Given the current allocation  $\mathbf{x}$ , we find candidates  $j, \ell \in C_\ell$  such that (i)  $x_j \leq 1 - \frac{\delta}{s_j}$ ; (ii)  $x_\ell \geq \underline{x}_\ell + \frac{\delta}{s_\ell}$ ; and (iii) the following condition holds:

$$(3.1) \quad \frac{\partial \phi(\mathbf{x})}{\partial x_j \cdot s_j} > \frac{\partial \phi(\mathbf{x})}{\partial x_\ell \cdot s_\ell} + \frac{\varepsilon}{b}.$$

Recall that  $b$  is the budget for the overall problem, not for the subroutine. While such a pair of candidates  $(j, \ell)$  exists, the algorithm increases  $x_j$  by  $\delta/s_j$ , and decreases  $x_\ell$  by  $\delta/s_\ell$ . Note that this update is feasible for the program, and preserves the cost of the allocation in  $C_\ell$  at exactly  $B$ , and is hence feasible for the above program. The procedure stops when such a pair of candidates  $(j, \ell)$  no longer exists.

**3.2.3 Analysis** In the analysis below, we assume the local optimum can be efficiently computed and present properties of this solution. The running time analysis is presented in [Section 3.5](#).

We show the following result, which at a high level, extends the analysis of the PAV rule for binary additive utilities and unit sizes in [\[39\]](#) to the continuous setting, with submodular utilities and arbitrary sizes. This result will show as a corollary that  $\text{NW}(C, W, \{U_i\}, B)$  finds a 2-approximate fractional core outcome.

**THEOREM 3.1.** *Given a set of voters  $W$  with utilities  $\{u_i\}$  and multilinear extensions  $\{U_i\}$ , and a cost budget  $B \in [\frac{\varepsilon}{5m}b, b]$  on  $C_\ell$ , let  $\mathbf{x}$  denote the solution to  $\text{NW}(C, W, \{U_i\}, B)$ . Suppose a subset  $S \subseteq W$  of voters can choose an allocation  $\mathbf{y} \in [0, 1]^m$  with  $\text{Cost}(\mathbf{y}) \leq b$  such that  $U_i(\mathbf{y}) > \theta U_i(\mathbf{x})$  for all  $i \in S$ , where  $\theta \geq 1$ . Then:*

$$|S| < \frac{|W|}{B(1-\varepsilon)} \cdot \frac{\text{Cost}(\mathbf{y})}{\theta - 1 - 2\varepsilon}.$$

*Proof.* Given the local optimum  $\mathbf{x}$  with  $\sum_{j \in C_\ell} s_j x_j = B$ . Let  $M_1 = \{j \in C \mid x_j + \frac{\delta}{s_j} \leq y_j\}$  and  $M_2 = \{j \in C \mid x_j < y_j < x_j + \frac{\delta}{s_j}\}$ . Note that  $M_1 \cup M_2 \subseteq C_\ell$ , since for all  $j \in C_s$ , we have  $x_j = 1 \geq y_j$ .

Let  $\mathbf{y}'$  be the allocation where  $y'_j = y_j$  for  $j \in M_1$ , and  $y'_j = x_j$  otherwise. Note that when moving from  $\mathbf{x}$  to  $\mathbf{y}'$ , utility increase for  $i \in S$  can only come from allocation increase of candidates in  $M_1$ . Further, note that the direction  $\mathbf{y}' - \mathbf{x}$  is non-negative, since  $y_j > x_j$  for  $j \in M_1$  and  $y'_j = x_j$  for  $j \notin M_1$ .

By applying [Lemma 3.1](#) to the multilinear function  $U_i$  and observing that  $\frac{\partial U_i}{\partial x_j} \geq 0$ , we have

$$\sum_{j \in M_1} \frac{\partial U_i(\mathbf{x})}{\partial x_j} \cdot y_j + U_i(\mathbf{x}) \geq \sum_{j \in C} \frac{\partial U_i(\mathbf{x})}{\partial x_j} \cdot (y'_j - x_j) + U_i(\mathbf{x}) \geq U_i(\mathbf{y}'),$$

where the last inequality follows from the concavity of  $U_i$  in non-negative directions.

Consider now the allocation  $\mathbf{z}$ , where  $z_j = y_j$  if  $j \in M_1 \cup M_2$ , and  $z_j = x_j$  otherwise. Note that  $\mathbf{z} \geq \mathbf{y}$ , so that  $U_i(\mathbf{z}) \geq U_i(\mathbf{y})$ . Further, the direction  $\mathbf{z} - \mathbf{y}'$  is non-negative, since  $z_j = y_j \geq x_j = y'_j$  for  $j \in M_2$ , and  $z_j = y'_j$  otherwise. Applying [Lemma 3.1](#) again gives

$$U_i(\mathbf{z}) \leq U_i(\mathbf{y}') + \sum_j \frac{\partial U_i(\mathbf{y}')}{\partial y'_j} \cdot (z_j - y'_j) = U_i(\mathbf{y}') + \sum_{j \in M_2} \frac{\partial U_i(\mathbf{y}')}{\partial x_j} \cdot (y_j - y'_j).$$

Combining the relations above and observing that  $U_i(\mathbf{z}) \geq U_i(\mathbf{y})$ , we have

$$\sum_{j \in M_1} \frac{\partial U_i(\mathbf{x})}{\partial x_j} \cdot y_j + U_i(\mathbf{x}) \geq U_i(\mathbf{y}') \geq U_i(\mathbf{y}) - \sum_{j \in M_2} \frac{\partial U_i(\mathbf{y}')}{\partial x_j} \cdot (y_j - y'_j).$$

By [Lemma 3.2](#), we have  $\frac{\partial U_i(\mathbf{y}')}{\partial x_j} \leq 1$  for all  $j \in M_2$ . Further,  $y_j - y'_j = y_j - x_j \leq \frac{\delta}{s_j} \leq \frac{\varepsilon^6}{312m^5}$ , since  $s_j \geq \frac{\varepsilon}{m}b$  and  $\delta = \frac{\varepsilon^7 b}{312m^6}$ . Therefore, for any  $i \in S$ , we have:

$$\sum_{j \in M_1} \frac{\partial U_i(\mathbf{x})}{\partial x_j} \cdot y_j \geq U_i(\mathbf{y}) - U_i(\mathbf{x}) - \frac{\varepsilon^6}{312m^4} > (\theta - 1)U_i(\mathbf{x}) - \frac{\varepsilon^6}{312m^4},$$

where we have used that  $U_i(\mathbf{y}) > \theta U_i(\mathbf{x})$  for all  $i \in S$ .

Note now that

$$(3.2) \quad U_i(\mathbf{x}) \geq u_i^{\max} \cdot \underline{x}_j = \frac{B \cdot \varepsilon}{\sum_{j \in C_\ell} s_j} \geq \frac{\varepsilon^2 \cdot b/m}{b \cdot m} = \frac{\varepsilon^2}{m^2},$$

where we have used that  $s_j \leq b$  for all  $j \in C$ , and  $B \geq \frac{\varepsilon}{5m}b$ . Combining the two inequalities above gives

$$\sum_{j \in M_1} \frac{\partial U_i(\mathbf{x})}{\partial x_j} \cdot y_j \geq (\theta - 1 - \varepsilon) \cdot U_i(\mathbf{x}).$$

Summing this inequality over all candidates in  $S$ , we have

$$\sum_{i \in S} \sum_{j \in M_1} \frac{\partial U_i(\mathbf{x})}{\partial x_j} \cdot \frac{y_j}{U_i(\mathbf{x})} \geq |S| \cdot (\theta - 1 - \varepsilon).$$

Since  $\frac{\partial \phi_i(\mathbf{x})}{\partial x_j} = \frac{1}{U_i(\mathbf{x})} \cdot \frac{\partial U_i(\mathbf{x})}{\partial x_j}$ , we have

$$\begin{aligned} \frac{\sum_{j \in M_1} \left( y_j \cdot s_j \cdot \frac{\partial \phi(\mathbf{x})}{\partial x_j \cdot s_j} \right)}{\sum_{j \in M_1} y_j \cdot s_j} &\geq \frac{\sum_{j \in M_1} \left( \sum_{i \in S} \frac{\partial \phi_i(\mathbf{x})}{\partial x_j \cdot s_j} \cdot y_j \cdot s_j \right)}{\sum_{j \in M_1} y_j \cdot s_j} \geq \frac{\sum_{i \in S} \sum_{j \in M_1} \frac{\partial U_i(\mathbf{x})}{\partial x_j} \cdot \frac{y_j}{U_i(\mathbf{x})}}{\sum_{j \in M_1} y_j \cdot s_j} \\ &\geq \frac{|S| \cdot (\theta - 1 - \varepsilon)}{\text{Cost}(\mathbf{y})}. \end{aligned}$$

Therefore,

$$\max_{j \in M_1} \frac{\partial \phi(\mathbf{x})}{\partial x_j \cdot s_j} > \frac{|S| \cdot (\theta - 1 - \varepsilon)}{\text{Cost}(\mathbf{y})} \geq \frac{|S| \cdot (\theta - 1 - 2\varepsilon)}{\text{Cost}(\mathbf{y})} + \varepsilon/b,$$

where we have used that  $\text{Cost}(\mathbf{y}) \leq b$  and  $|S| \geq 1$ . Let  $j_1$  denote the  $j \in M_1$  that achieves the maximum in the previous inequality.

Now, let  $R_1 = \{j \in C_\ell \mid x_j \geq \underline{x}_j + \frac{\delta}{s_j}\}$ . Applying [Lemma 3.1](#) along the positive direction  $\mathbf{x}$  gives  $\sum_{j \in C} x_j \cdot \frac{\partial U_i(\mathbf{x})}{\partial x_j} \leq U_i(\mathbf{x})$ . Since  $\frac{\partial U_i(\mathbf{x})}{\partial x_j} \geq 0$  for all  $j \in C$ , we have

$$\sum_{j \in R_1} x_j \cdot \frac{\partial U_i(\mathbf{x})}{\partial x_j} \leq U_i(\mathbf{x}).$$

Summing this over all  $i \in W$ , we have:

$$\sum_{i \in W} \sum_{j \in R_1} x_j \cdot s_j \cdot \frac{\partial \phi_i(\mathbf{x})}{\partial x_j \cdot s_j} = \sum_{i \in W} \frac{\sum_{j \in R_1} x_j \cdot \frac{\partial U_i(\mathbf{x})}{\partial x_j}}{U_i(\mathbf{x})} \leq \sum_{i \in W} \frac{U_i(\mathbf{x})}{U_i(\mathbf{x})} = |W|.$$

Therefore,

$$\frac{\sum_{j \in R_1} x_j \cdot s_j \cdot \frac{\partial \phi(\mathbf{x})}{\partial x_j \cdot s_j}}{\sum_{j \in R_1} x_j \cdot s_j} = \frac{\sum_{i \in W} \sum_{j \in R_1} x_j \cdot s_j \cdot \frac{\partial \phi_i(\mathbf{x})}{\partial x_j \cdot s_j}}{\sum_{j \in R_1} x_j \cdot s_j} \leq \frac{|W|}{\sum_{j \in R_1} x_j \cdot s_j} \leq \frac{|W|}{B(1 - \varepsilon)}.$$

The last inequality holds since  $\sum_{j \in C_\ell \setminus R_1} x_j \cdot s_j \leq \sum_{j \in C_\ell \setminus R_1} \frac{\delta}{s_j} \cdot s_j \leq m\delta \leq \varepsilon B$ , where we have used that  $\delta = \frac{\varepsilon^7 b}{312m^6}$  and  $B \geq \frac{\varepsilon}{5m} b$ . This implies  $\sum_{j \in R_1} s_j x_j \geq B(1 - \varepsilon)$ .

Therefore,

$$\min_{j \in R_1} \frac{\partial \phi(\mathbf{x})}{\partial x_j \cdot s_j} \leq \frac{|W|}{B(1 - \varepsilon)}.$$

Let  $j_2$  denote the  $j \in R_1$  that achieves this minimum.

Note that  $x_{j_1} < y_{j_1}$  and  $x_{j_2} > y_{j_2}$  by assumption. Further, since  $\mathbf{x}$  is locally optimal, [Eq. \(3.1\)](#) cannot hold for  $j = j_2$  and  $\ell = j_1$ . This means  $\frac{\partial \phi(\mathbf{x})}{\partial x_{j_1} \cdot s_{j_1}} \leq \frac{\partial \phi(\mathbf{x})}{\partial x_{j_2} \cdot s_{j_2}} + \varepsilon/b$ . Therefore, we have

$$\frac{|S| \cdot (\theta - 1 - 2\varepsilon)}{\text{Cost}(\mathbf{y})} < \frac{\partial \phi(\mathbf{x})}{\partial x_{j_1} \cdot s_{j_1}} - \varepsilon/b \leq \frac{\partial \phi(\mathbf{x})}{\partial x_{j_2} \cdot s_{j_2}} \leq \frac{|W|}{B(1 - \varepsilon)}.$$

Rearranging completes the proof.  $\square$

As an aside, the following corollary is immediate if all one wants is an approximate fractional core on the utilities  $\{U_i\}$  and budget  $b$ . Note that the proof trivially extends to *any* continuous, concave utilities  $\{U_i\}$ , since these are concave in positive directions.

**COROLLARY 3.1.** *The solution  $\mathbf{x}$  of  $\text{NW}(C, V, \{U_i\}, b)$  is in a 2-approximate fractional core for any continuous concave utilities  $\{U_i\}$ .*

*Proof.* Setting  $\theta = 2$  and  $\varepsilon = 0$  in [Theorem 3.1](#), if there exists an  $S$  and a solution  $\mathbf{y}$  such that  $U_i(\mathbf{y}) > 2U_i(\mathbf{x})$  for all  $i \in S$ , then  $|S| < \frac{|W|}{b} \cdot \text{Cost}(\mathbf{y})$ . This contradicts [Definition 2.1](#) and thus completes the proof.  $\square$

**3.3 Randomized Rounding and Satisfied Voters** We now present a randomized rounding scheme on the fractional solution constructed above. Since the utilities of voters are different but the allocation is the same, we will need the rounding of the fractional allocation to be *oblivious* to the utilities. The natural approach is therefore randomized rounding. Clearly, if we randomly round, the expected utility of a voter in the integral committee is at least that in the fractional core. However, the same cannot be said for an arbitrary *realization* of the randomization: there can be many voters having lower-than-expected utilities, and they can form a coalition and deviate. Further, the size of the resulting committee can exceed the bound  $B$ . Therefore, we cannot directly argue that the resulting integral committee will lie in the core.

Our contribution in this section is to show a rounding procedure that ensures that with constant probability, a constant fraction of voters obtain at least a constant fraction of the utilities they obtained in the fractional solution. We will use this as a subroutine in our main procedure in [Section 3.4](#) in order to argue that these voters will not form part of any deviating coalition.

**3.3.1 Rounding Procedure** To begin with, in the sequel, we will distinguish between items that are fractionally and integrally allocated as follows:

**DEFINITION 3.2.** (FULLY AND FRACTIONALLY ALLOCATED ITEMS) *An item  $j \in C$  in the fractional solution is fully allocated if  $x_j \geq 1$ ; we call it fractionally allocated if  $x_j \in (0, 1)$ . Note that all items in  $C_s$  are fully allocated.*

The rounding procedure has a parameter  $\kappa \leq 1$ , and executes as shown in [Algorithm 1](#).

---

**Algorithm 1** Randomized Rounding of Fractional Allocation  $\mathbf{x}$ 


---

```

1: function ROUND( $C, W, \{U_i\}, B$ )
2:    $\mathbf{x} \leftarrow \text{NW}(C, W, \{U_i\}, \kappa B)$ 
3:    $O \leftarrow C_s$ 
4:    $T_2 \leftarrow \{j \in C_\ell \mid s_j \leq \kappa B\}$ 
5:   Include  $j \in T_2$  in  $O$  independently with probability  $\min(1, x_j)$ 
6:   return  $O$ 
7: end function

```

---

**3.3.2 Satisfied Voters** We now define what it means for a voter to be *satisfied*, and present a bound on the number of satisfied voters after the procedure ROUND is applied.

**DEFINITION 3.3.** ( $\gamma$ -SATISFIED) *Let  $\mathbf{x}$  be the fractional solution of  $\text{NW}(C, W, \{U_i\}, \kappa B)$ , and  $O$  be the resulting integer committee found by  $\text{ROUND}(C, W, \{U_i\}, B)$ . For  $\gamma \geq 2$ , we say that voter  $i$  is  $\gamma$ -satisfied by  $O$  w.r.t. solution  $\mathbf{x}$  if there is a  $q \in C$  such that  $u_i(O \cup \{q\}) \geq \frac{U_i(\mathbf{x})}{\gamma}$ .*

Our main result in this section is the following theorem.

**THEOREM 3.2.** (CONSTANT FRACTION OF CONSTANT-SATISFIED VOTERS FOR SUBMODULAR UTILITIES) *Given the fractional solution  $\mathbf{x}$  produced by  $\text{NW}(C, W, \{U_i\}, \kappa B)$  where  $|W| = n'$ , there is a integral committee  $O$  produced by  $\text{ROUND}(C, W, \{U_i\}, B)$  with at least  $(1 - \beta)n'$   $\gamma$ -satisfied voters, where*

$$(3.3) \quad \beta = (\kappa e^{1-\kappa})^{\frac{1}{\kappa}} + (\gamma - 1)e^{2-\gamma}.$$

Further, this committee is feasible, so that  $\sum_{j \in O \cap C_\ell} s_j \leq B$ .

For the proof, we will need the following lemma, which is the analog of [Lemma 4.1](#) for submodular functions.

**LEMMA 3.3.** (LOWER TAIL BOUND, [\[16\]](#)) *Assume  $f(\{j\}) \in [0, 1]$  for all  $j \in C$ , and let  $\mu_0 = F(\mathbf{x})$ , where  $F$  is the multilinear extension of  $f$ . Let  $O$  denote the outcome of independent randomized rounding applied to  $\mathbf{x} \in [0, 1]^m$ . Then for any  $\rho > 1$ :*

$$\Pr \left[ f(O) \leq \frac{\mu_0}{\rho} \right] \leq (\rho e^{1-\rho})^{\frac{\mu_0}{\rho}}.$$

*Proof.* (Proof of Theorem [3.2](#)) We show that for each voter  $i$ , the probability that  $i$  is  $\gamma$ -satisfied is at least  $1 - \beta$ . This implies the number of  $\gamma$ -satisfied voters is at least  $(1 - \beta)n'$  in expectation, and thus it is at least  $(1 - \beta)n'$  for some realization produced by ROUND.

Let  $T_1 = \{j \in C_\ell \mid s_j \geq \kappa B\}$  and  $T_2 = C \setminus T_1$ . Fix voter  $i$  with utility function  $u_i$ , and multilinear extension  $U_i$ .

Let  $G_1$  denote the multilinear extension restricted only to candidates in  $T_1$ , that is,

$$G_1(\mathbf{x}) = \sum_{T \subseteq T_1} u_i(T) \left( \prod_{j \in T} x_j \right) \left( \prod_{j \in T_1 \setminus T} (1 - x_j) \right),$$

and similarly,  $G_2$  denote the multilinear extension restricted to  $T_2$ , that is,

$$G_2(\mathbf{x}) = \sum_{T \subseteq T_2} u_i(T) \left( \prod_{j \in T} x_j \right) \left( \prod_{j \in T_2 \setminus T} (1 - x_j) \right).$$

By sub-additivity of  $u_i$ , we have  $U_i(\mathbf{x}) \leq G_1(\mathbf{x}) + G_2(\mathbf{x})$ , and by monotonicity of  $u_i$ , we have  $U_i(\mathbf{x}) \geq \max(G_1(\mathbf{x}), G_2(\mathbf{x}))$ .

We now split the analysis into several cases:

**Case (1):** Suppose  $G_1(\mathbf{x}) \geq \frac{1}{\gamma}U_i(\mathbf{x})$ . In this case, we make two observations. First,

$$G_1(\mathbf{x}) \leq \sum_{T \subseteq T_1} \left( \sum_{j \in T} u_i(\{j\}) \right) \left( \prod_{j \in T} x_j \right) \left( \prod_{j \in T_1 \setminus T} (1 - x_j) \right) \leq \sum_{j \in T_1} x_j u_i(\{j\}).$$

Further, since  $\sum_{j \in T_1} s_j x_j \leq \kappa B$ , and since  $s_j \geq \kappa B$  for all  $j \in T_1$ , we have  $\sum_{j \in T_1} x_j \leq 1$ . Putting these together, we have

$$G_1(\mathbf{x}) \leq \max_{j \in T_1} u_i(\{j\}).$$

Therefore, using  $j^* = \operatorname{argmax}_{j \in T_1} u_i(\{j\})$  as additament makes  $i$  achieve utility at least  $1/\gamma$  fraction of  $U_i(\mathbf{x})$ . Therefore,  $i$  is  $\gamma$ -satisfied just by the additament.

**Case (2):** We now assume  $G_1(\mathbf{x}) \leq \frac{1}{\gamma}U_i(\mathbf{x})$  and thus  $G_2(\mathbf{x}) \geq (1 - 1/\gamma)U_i(\mathbf{x})$ . Suppose there exists  $\ell \in T_2$  such that  $u_i(\{\ell\}) \geq \frac{1}{\gamma-1}G_2(\mathbf{x})$ . Then, just using  $\ell$  as additament causes the utility of  $i$  to be at least  $1/\gamma$  fraction of  $F_i(\mathbf{x})$  and voter  $i$  is  $\gamma$ -satisfied by the additament  $\ell$ .

**Case (3):** In the final case, we have  $G_2(\mathbf{x}) \geq (1 - 1/\gamma)U_i(\mathbf{x})$ , and we assume that for all  $j \in T_2$ ,  $u_i(\{j\}) < \frac{1}{\gamma-1}G_2(\mathbf{x})$ . Let  $O' = O \cap C_\ell$ .

First note that  $\mathbb{E} \left[ \sum_{j \in O'} s_j \right] \leq \kappa B$ , since  $O'$  is the result of randomized rounding of  $\mathbf{x}$  and since  $\sum_{j \in C_\ell} s_j x_j \leq \kappa B$ . Further,  $O' \subseteq T_2 \cap C_\ell$ , and for all  $j \in T_2 \cap C_\ell$ , we have  $s_j \leq \kappa B$ . Therefore, a standard application of Chernoff bounds yields:

$$\Pr \left[ \sum_{j \in O'} s_j > B \right] \leq (\kappa e^{1-\kappa})^{\frac{1}{\kappa}}.$$

Further, note that for all  $j \in T_2$ , we have that the marginal  $u_i(\{j\}) < \frac{1}{\gamma-1}G_2(\mathbf{x})$ . Applying [Lemma 3.3](#) with  $\rho = \gamma - 1$ , we have:

$$\Pr \left[ u_i(O) \leq \frac{G_2(\mathbf{x})}{\gamma-1} \right] \leq (\gamma-1)e^{2-\gamma}.$$

Therefore, by union bounds, with probability at least  $1 - \beta$ , we have both events: (1) The committee  $O'$  is feasible for size  $B$ ; and (2)  $u_i(O) \geq \frac{G_2(\mathbf{x})}{\gamma-1} \geq \frac{U_i(\mathbf{x})}{\gamma}$ , implying  $i$  is  $\gamma$ -satisfied by  $O$ . This completes the proof.  $\square$

The following corollary shows that this step can be implemented in polynomial time.

**COROLLARY 3.2.** For  $\beta$  defined in [Eq. \(3.3\)](#),  $n' = |W|$  and any  $\varepsilon \in (0, 1)$ , a committee  $O \subseteq C$  with  $(1 - \beta - \varepsilon)n'$   $\gamma$ -satisfied voters can be computed with probability  $1 - \frac{1}{\operatorname{poly}(m, n)}$  in time  $\operatorname{poly}(n, m, 1/\varepsilon)$ .

*Proof.* Let  $X$  be the random variable indicating the number of  $\gamma$ -satisfied voters returned by  $\operatorname{ROUND}(C, W, \{U_i\}, B)$ . By [Theorem 3.2](#), we have  $\mathbb{E}[X] = (1 - \beta)n'$ . Further,  $X \leq n'$ . Therefore,  $\Pr[X \geq (1 - \beta - \varepsilon)n] \geq \frac{\varepsilon}{\beta + \varepsilon} \geq \frac{\varepsilon}{2}$ . A standard application of sampling bounds now completes the proof.  $\square$

**3.4 The Constant Approximation to the Core** In this section, we design an algorithm for the overall problem, and prove that it returns a committee in an  $O(1)$ -core. The algorithm will repeatedly construct fractional solutions and round them, using the algorithms developed above as subroutines. The analysis critically requires the local optimality property of the Nash Welfare objective, captured in [Theorem 3.1](#).

**3.4.1 Algorithm** [Algorithm 2](#) finds an approximate core solution in the following way: It iteratively computes the fractional local optimum to Nash Welfare on the remaining voters with a scaled-down budget, rounds it, eliminates voters that are  $\gamma$ -satisfied with respect to the solution at this iteration, scales down the budget again, and iterates on the remaining voters with the smaller budget. The scaling parameter  $\omega$  for the budget will be determined later. The overall structure of the algorithm is similar to that in [\[32\]](#), though the details of constructing the fractional solution, and the resulting proof of correctness are entirely different.

---

**Algorithm 2** Iterative Rounding of the Nash Welfare Solution

---

```

1: function ITERROUND( $C, V, \{u_{ij}\}, b$ )
2:    $t \leftarrow 0$ 
3:    $V_0 \leftarrow V$ 
4:    $T^* \leftarrow C_s$ 
5:    $b_0 \leftarrow (1 - \varepsilon)(1 - \omega)b$ 
6:   while  $b_t \geq \frac{\varepsilon}{m}b$  do
7:      $\mathbf{x}_t \leftarrow \text{NW}(C, V_t, \{U_i\}, \kappa b_t)$ 
8:      $O_t \leftarrow \text{Solution of ROUND}(C, W, \{U_i\}, b_t)$  that satisfies Corollary 3.2
9:      $W_t \leftarrow \text{Voters in } V_t \text{ that are } \gamma\text{-satisfied by } O_t \text{ with respect to the solution } \mathbf{x}_t$ 
10:     $V_{t+1} \leftarrow V_t \setminus W_t$ 
11:     $T^* \leftarrow T^* \cup O_t$ 
12:     $b_{t+1} \leftarrow \omega b_t$ 
13:     $t \leftarrow t + 1$ 
14:  end while
15:  return  $T^*$  as the final integer solution
16: end function

```

---

**3.4.2 Analysis** Let  $n = |V|$ , and recall that  $b$  is the initial budget. We will show that the solution  $T^*$  lies in the  $\alpha$ -core for the set  $V$  of voters with size constraint  $b$ . First note that since  $\sum_{j \in C_s} s_j \leq \varepsilon b$ , and since each  $O_t$  is feasible for budget  $b_t$ , we have:

$$\sum_{j \in T^*} s_j \leq \varepsilon b + \sum_t \sum_{j \in O_t} s_j \leq \varepsilon b + \sum_t b_t = \varepsilon b + (1 - \varepsilon)(1 - \omega) \sum_{t \geq 0} \omega^t b \leq b.$$

Therefore, the solution  $T^*$  is feasible for the size  $b$ .

First note that when  $b_t < \frac{\varepsilon}{m}b$ , since all items in  $C_\ell$  has  $s_j > \frac{\varepsilon}{m}b$ , we have  $\sum_{j \in C_\ell} x_{jt} < 1$ . Since  $C_s \subseteq T^*$ , for any fractional solution  $\mathbf{x}_t$ , we have  $\max_{j \in C} u(\{j\}) \geq U_i(\mathbf{x}_t)$ , so that all voters are 1-satisfied. This implies  $V_t = \emptyset$  at termination, so that any voter  $i \in V$  belongs to  $W_{t'}$  for some  $t'$ .

For the purpose of contradiction, we assume the resulting solution  $T^*$  is not in the  $\alpha$ -core (Definition 1.2), where  $\alpha \geq 1$  is a quantity we will determine later. Let  $S$  denote the set of voters that deviate, and let  $A$  denote the set of items they deviate to. We have  $\sum_{j \in A} s_j \leq \frac{|S|}{n} \cdot b$ , and  $A$  provides an  $\alpha$ -factor larger utility to voters in  $S$  even after including any additament.

Consider the voters in  $W_t$ , and let  $S_t = S \cap W_t$ . These voters are  $\gamma$ -satisfied by  $O_t$  with respect to the fractional solution  $\mathbf{x}_t$  (Definition 3.3). Therefore, if  $i \in S_t$  deviates to  $A$  to obtain an  $\alpha$ -factor larger utility, it must be that  $U_i(A) \geq \frac{\alpha}{\gamma} \cdot U_i(\mathbf{x}_t)$ . Let  $\theta = \frac{\alpha}{\gamma}$ . We will assume  $\theta > 1$  below.

Using Theorem 3.1 since  $\mathbf{x}_t$  is a local optimum, using  $\mathbf{y}$  as the set  $S_t$ , and observing that all agents  $i \in S_t$  have  $U_i(\mathbf{y}) > \theta U_i(\mathbf{x}_t)$ , where  $\theta = \frac{\alpha}{\gamma}$ , we have:

$$|S_t| \leq \frac{n_t}{\kappa b_t \cdot (1 - \varepsilon)} \cdot \frac{\sum_{j \in A} s_j}{\theta - 1 - 2\varepsilon}.$$

Summing this over all  $t$ , and using  $\sum_{j \in A} s_j \leq \frac{|S|}{n}b$ , we have:

$$|S| = \sum_t |S_t| \leq \frac{\sum_{j \in A} s_j}{(\theta - 1 - 2\varepsilon)(1 - \varepsilon)} \cdot \sum_t \frac{n_t}{\kappa b_t} \leq \frac{|S|}{n} \cdot \frac{b}{(\theta - 1 - 2\varepsilon)(1 - \varepsilon)} \cdot \sum_t \frac{n_t}{\kappa b_t}.$$

Therefore, for a blocking coalition to exist, we need:

$$(3.4) \quad \frac{b}{n} \cdot \sum_t \frac{n_t}{\kappa b_t} \geq (\theta - 1 - 2\varepsilon)(1 - \varepsilon) = \left( \frac{\alpha}{\gamma} - 1 - 2\varepsilon \right) (1 - \varepsilon).$$

We will now set the parameters  $\omega, \gamma, \alpha$  so that the above inequality is false. First note by [Corollary 3.2](#) that  $n_{t+1} \leq (\beta + \varepsilon)n_t$  where  $\beta$  satisfied [Eq. \(3.3\)](#). Further,  $b_{t+1} = \omega b_t$ . Therefore,

$$\frac{n_{t+1}}{b_{t+1}} \leq \frac{\beta + \varepsilon}{\omega} \cdot \frac{n_t}{b_t}$$

with  $\frac{n_0}{b_0} = \frac{n}{(1-\omega)(1-\varepsilon)b}$ . Therefore,

$$(3.5) \quad \frac{(1-\varepsilon)b}{n} \cdot \sum_t \frac{n_t}{\kappa b_t} \leq \frac{1}{(1-\omega)\kappa} \cdot \sum_{t \geq 0} \left( \frac{\beta + \varepsilon}{\omega} \right)^t = \frac{\omega}{(1-\omega)(\omega - \beta - \varepsilon)\kappa}.$$

Combining [Eq. \(3.4\)](#) and [Eq. \(3.5\)](#), for a blocking coalition to exist, we need

$$\alpha \leq \frac{\omega\gamma}{\kappa(1-\omega) \left( \omega - (\gamma - 1)e^{2-\gamma} - (\kappa e^{1-\kappa})^{\frac{1}{\kappa}} - \varepsilon \right) (1-\varepsilon)^2} + (1 + 2\varepsilon) \cdot \gamma.$$

For an  $\alpha$  slightly larger than the right-hand side, a blocking coalition will therefore not exist. Then The right-hand side of the above inequality is approximately minimized when  $\omega = 0.23$ ,  $\gamma = 7.435$ ,  $\kappa = 0.21$  and  $\varepsilon \rightarrow 0$ , yielding  $\alpha < 67.37$ <sup>1</sup>. This finally yields the following theorem; the only missing detail is the running time of local search in [Section 3.2](#), which we address in [Section 3.5](#).

**THEOREM 3.3. (MAIN THEOREM)** *For monotone submodular utilities, a 67.37-core is always non-empty. One such solution can be computed in polynomial time.*

**3.5 The Running Time of Local Search** We now show that [Algorithm 2](#) runs in polynomial time. This requires showing that the local search procedure in [Section 3.2](#) runs in polynomial time; the rest of the steps in [Algorithm 2](#) can easily be implemented efficiently. We will show this in two parts: The partial derivatives  $\frac{\partial \phi_i(\mathbf{x})}{\partial x_j}$  can be approximately computed efficiently; and the number of iterations (finding the candidate pair  $(j, \ell)$  and updating the allocations) performed by the local search procedure is polynomially bounded.

**3.5.1 Estimating the Gradient of the Nash Welfare Objective** We first show how to estimate the derivative of  $\phi(\mathbf{x})$  in the procedure in [Section 3.2](#). First, the estimation procedure for  $U_i$  and its derivative is the same as that in the work of [\[14\]](#): Each time we compute  $U_i(\mathbf{x})$ , we pretend  $\mathbf{x}$  is a product distribution over candidates, and sample it  $H$  times. Denote the samples as the committees  $O_h \sim \mathbf{x}$  where  $h \in \{1, 2, \dots, H\}$ . We calculate the quantity  $\hat{U}_i(\mathbf{x}) = \frac{1}{H} \cdot \sum_{h=1}^H u_i(O_h)$  as the estimation of  $U_i(\mathbf{x})$ . The following lemma gives a bound on the additive error of estimation on  $U_i(\mathbf{x})$  if  $H$  is sufficiently large.

**LEMMA 3.4. (CHERNOFF BOUND, [\[14\]](#))**  $\Pr \left[ \left| \hat{U}_i(\mathbf{x}) - U_i(\mathbf{x}) \right| > \Delta \right] < 2e^{-\frac{\Delta^2 \cdot H}{m^2}}.$

Since the derivative of  $U_i$  is given in [Lemma 3.2](#), its error is bounded by estimating the multilinear function twice. This yields the following lemma:

**LEMMA 3.5.** *In the local search procedure in [Section 3.2](#), suppose the total number of times the multilinear function  $U_i$  and its derivative is evaluated is  $N$ . Then, using  $\text{poly}(N, m, \frac{1}{\Delta})$  samples, the additive error in each estimate is bounded by  $\Delta$  with probability  $1 - \frac{1}{\text{poly}(m, N)}$ .*

In the sequel, we will set  $\Delta = \frac{\varepsilon^6}{64n \cdot m^5}$ . We denote the estimated derivative of the multilinear function as  $\frac{\partial \hat{U}_i(\mathbf{x})}{\partial x_i}$ , and the estimated derivative of  $\phi_i(\mathbf{x})$  and  $\phi(\mathbf{x})$  as  $\frac{\partial \phi_i(\mathbf{x})}{\partial x_j} = \frac{1}{U_i} \cdot \frac{\partial U_i(\mathbf{x})}{\partial x_j}$  and  $\frac{\partial \phi(\mathbf{x})}{\partial x_j} = \sum_{i=1}^n \frac{\partial \phi_i(\mathbf{x})}{\partial x_j}$  respectively. We have the following lemma for bounding the error of estimating  $\frac{\partial \phi(\mathbf{x})}{\partial x_j} \cdot \frac{1}{s_j}$ . For this proof and subsequent ones, we crucially need that the Nash Welfare program in [Section 3.2](#) lower bounds the allocations as  $x_j \geq \underline{x}_j$ .

<sup>1</sup>Note that for this choice of  $\kappa$ , the subroutine  $\text{NW}(C, V_t, \{U_i\}, \kappa b_t)$  is run with budget at least  $\kappa b_t \geq \kappa \frac{\varepsilon}{m} b \geq \frac{\varepsilon}{5m} b$ , so that the precondition of Theorem [3.1](#) holds.

LEMMA 3.6. Let  $\Delta = \frac{\varepsilon^6}{64n \cdot m^5}$ , and suppose  $x_j \geq \underline{x}_j$  for all candidates  $j \in C_\ell$ , where  $\underline{x}_j$  is as defined in Section 3.2. If for any  $j \in C_\ell$  and all  $i \in V$ , we have:  $|\hat{U}_i(\mathbf{x}) - U_i(\mathbf{x})| < \Delta$  and  $|\frac{\partial \hat{U}_i(\mathbf{x})}{\partial x_j} - \frac{\partial U_i(\mathbf{x})}{\partial x_j}| < \Delta$ , then it holds that  $\frac{1}{s_j} \cdot \left| \frac{\partial \phi(\mathbf{x})}{\partial x_j} - \frac{\partial \phi(\mathbf{x})}{\partial x_j} \right| \leq \frac{\varepsilon}{8b}$ .

*Proof.* We have

$$\begin{aligned} \left| \frac{\partial \phi_i(\mathbf{x})}{\partial x_j} - \frac{\partial \phi_i(\mathbf{x})}{\partial x_j} \right| &= \left| \frac{1}{\hat{U}_i(\mathbf{x})} \cdot \frac{\partial \hat{U}_i(\mathbf{x})}{\partial x_j} - \frac{1}{U_i(\mathbf{x})} \cdot \frac{\partial U_i(\mathbf{x})}{\partial x_j} \right| \\ &\leq \frac{1}{U_i(\mathbf{x}) - \Delta} \cdot \left( \frac{\partial U_i(\mathbf{x})}{\partial x_j} + \Delta \right) - \frac{1}{U_i(\mathbf{x})} \cdot \frac{\partial U_i(\mathbf{x})}{\partial x_j} \\ &\leq \left| \frac{\Delta \cdot \left( U_i(\mathbf{x}) + \frac{\partial U_i(\mathbf{x})}{\partial x_j} \right)}{(U_i(\mathbf{x}) - \Delta)^2} \right|. \end{aligned}$$

By Lemma 3.2, we have  $\frac{\partial U_i(\mathbf{x})}{\partial x_j} \leq 1$ . If  $\Delta \leq \frac{\varepsilon^2}{2m^2}$ , then by Eq. (3.2) we have  $2\Delta \leq U_i(\mathbf{x})$ , thus:

$$\left| \frac{\partial \phi_i(\mathbf{x})}{\partial x_j} - \frac{\partial \phi_i(\mathbf{x})}{\partial x_j} \right| = \left| \frac{\Delta \cdot \left( U_i(\mathbf{x}) + \frac{\partial U_i(\mathbf{x})}{\partial x_j} \right)}{(U_i(\mathbf{x}) - \Delta)^2} \right| \leq \left| \frac{\Delta \cdot (U_i(\mathbf{x}) + 1)}{U_i(\mathbf{x})^2/4} \right| \leq \left| \frac{\Delta \cdot 2}{\frac{\varepsilon^4}{m^4}/4} \right| \leq \left| \frac{8\Delta \cdot m^4}{\varepsilon^4} \right|.$$

Since  $s_j \geq \frac{\varepsilon}{m}b$  for  $j \in C_\ell$ , we have

$$\frac{1}{s_j} \cdot \left| \frac{\partial \phi(\mathbf{x})}{\partial x_j} - \frac{\partial \phi(\mathbf{x})}{\partial x_j} \right| \leq \frac{1}{s_j} \sum_{i=1}^n \left| \frac{\partial \phi_i(\mathbf{x})}{\partial x_j} - \frac{\partial \phi_i(\mathbf{x})}{\partial x_j} \right| \leq \left| \frac{8\Delta \cdot m^4}{\varepsilon^4} \right| \cdot n \cdot \frac{m}{b \cdot \varepsilon} \leq \frac{8\Delta \cdot m^5 \cdot n}{\varepsilon^5 \cdot b} = \frac{\varepsilon}{8b},$$

finishing the proof.  $\square$

**3.5.2 Number of Iterations in Local Search** The local search procedure in Section 3.2 iteratively finds a pair of candidates  $(j, \ell)$  such that

$$(3.6) \quad \frac{\partial \phi(\mathbf{x})}{\partial x_j} \cdot \frac{1}{s_j} > \frac{\partial \phi(\mathbf{x})}{\partial x_\ell} \cdot \frac{1}{s_\ell} + \frac{3\varepsilon}{4b}.$$

Each time it finds such a pair, it increases  $x_j$  by  $\delta \cdot \frac{1}{s_j}$  and decreases  $x_\ell$  by  $\delta \cdot \frac{1}{s_\ell}$ . Note that at stopping, Eq. (3.6) does not hold, which by Lemma 3.6 implies Eq. (3.1) also does not hold with high probability. Therefore, the termination condition in Section 3.2 is satisfied.

In order to bound the number of iterations of this procedure, we first show that the first and second partial derivatives of  $\phi$  are bounded from above.

LEMMA 3.7. If  $\phi_i(\mathbf{x}) = \log U_i(\mathbf{x})$  and  $x_j \geq \underline{x}_j$  for all candidates  $j \in C_\ell$ , where  $\underline{x}_j$  is as defined in Section 3.2, then we have the following bounds for all  $j, k \in C_\ell$  and  $i \in W$ :

$$0 \leq \frac{\partial \phi_i(\mathbf{x})}{\partial x_j} = O\left(\frac{m^2}{\varepsilon^2}\right) \quad \text{and} \quad \left| \frac{\partial^2 \phi_i(\mathbf{x})}{\partial x_j \partial x_k} \right| = O\left(\frac{m^4}{\varepsilon^4}\right).$$

*Proof.* Since  $\phi_i(\mathbf{x}) = \log U_i(\mathbf{x})$ , using Lemma 3.2, we have  $\frac{\partial \phi_i(\mathbf{x})}{\partial x_j} \leq \frac{1}{U_i}$ . Combining this with Eq. (3.2) gives  $\frac{\partial \phi_i(\mathbf{x})}{\partial x_j} \leq \frac{5m^2}{\varepsilon^2}$ .

We next bound the second order derivatives as follows. Here,  $\mathbf{x}_{-j,k}$  is  $\mathbf{x}$  with the  $j^{\text{th}}$  and  $k^{\text{th}}$  dimension removed, and  $T \sim \mathbf{x}_{-j,k}$  means that  $T$  is chosen by including  $\ell \in C \setminus \{j, k\}$  independently with probability  $x_\ell$ .

The first inequality below uses [Eq. \(3.2\)](#).

$$\begin{aligned}
\left| \frac{\partial^2 \phi_i(\mathbf{x})}{\partial x_j \partial x_k} \right| &= \left| -\frac{1}{U_i(\mathbf{x})^2} \cdot \frac{\partial U_i(\mathbf{x})}{\partial x_j} + \frac{1}{U_i(\mathbf{x})} \cdot \frac{\partial^2 U_i(\mathbf{x})}{\partial x_j \partial x_k} \right| \\
&\leq \frac{25m^4}{\varepsilon^4} + \left| \frac{5m^2}{\varepsilon^2} \cdot \sum_{T \subseteq C \setminus \{j,k\}} \Pr_{T \sim \mathbf{x}_{-j,k}} [T] \cdot (u_i(T \cup \{j,k\}) - u_i(T \cup \{j\}) - u_i(T \cup \{k\}) + u_i(T)) \right| \\
&\leq \frac{25m^4}{\varepsilon^4} + \frac{5m^2}{\varepsilon^2} \cdot \sum_{T \subseteq C \setminus \{j,k\}} \Pr_{T \sim \mathbf{x}_{-j,k}} [T] \cdot (u_i(T \cup \{k\}) - u_i(T) - u_i(T \cup \{j,k\}) + u_i(T \cup \{j\})) \\
&\leq \frac{25m^4}{\varepsilon^4} + \frac{5m^2}{\varepsilon^2} \cdot \sum_{T \subseteq C \setminus \{j,k\}} \Pr_{T \sim \mathbf{x}_{-j,k}} [T] \cdot 1 \leq \frac{26m^4}{\varepsilon^4}.
\end{aligned}$$

This finishes the proof.  $\square$

The next lemma now follows from a standard application of first order Taylor approximation. We assume the local search procedure in [Section 3.2](#) iteratively finds a pair of candidates  $(j, \ell)$  such that [Eq. \(3.6\)](#) holds.

LEMMA 3.8. *Assuming all the estimates on  $\frac{\partial \phi(\mathbf{x})}{\partial x_j}$  during execution of local search are within  $\pm \frac{\varepsilon}{8b}$  of the true values, the total number of iterations is  $\text{poly}(m, n, \frac{1}{\varepsilon})$ .*

*Proof.* Since all estimates of  $\frac{\partial \phi(\mathbf{x})}{\partial x_j}$  during execution of local search are within  $\pm \frac{\varepsilon}{8b}$  of the true values, if [Eq. \(3.6\)](#) holds for  $j, k \in C_\ell$ , then:

$$\frac{\partial \phi(\mathbf{x})}{\partial x_j} \cdot \frac{1}{s_j} > \frac{\hat{\partial} \phi(\mathbf{x})}{\hat{\partial} x_j} \cdot \frac{1}{s_j} - \frac{\varepsilon}{8b} \geq \frac{\hat{\partial} \phi(\mathbf{x})}{\hat{\partial} x_\ell} \cdot \frac{1}{s_\ell} + \frac{3\varepsilon}{4b} - \frac{\varepsilon}{8b} \geq \frac{\partial \phi(\mathbf{x})}{\partial x_\ell} \cdot \frac{1}{s_\ell} + \frac{\varepsilon}{2b}.$$

We lower bound the increase in  $\phi(\mathbf{x})$  when  $\mathbf{x} = (\mathbf{x}_{-j,k}, x_j, x_k)$  moves to  $\mathbf{x}' = (\mathbf{x}_{-j,k}, x_j + \frac{\delta}{s_j}, x_\ell - \frac{\delta}{s_\ell})$ . By Taylor approximation, we have:

$$\phi(\mathbf{x}') - \phi(\mathbf{x}) \geq (\mathbf{x}' - \mathbf{x}) \cdot \nabla \phi(\mathbf{x}) - \sum_{r=1}^m (x'_r - x_r)^2 \frac{M}{2}$$

where  $M$  is an upper bound on the absolute value of the second derivatives of  $\phi$ . Since  $M = O\left(\frac{m^4}{\varepsilon^4}\right)$  by [Lemma 3.7](#), this implies:

$$\phi(\mathbf{x}') - \phi(\mathbf{x}) \geq \delta \cdot \left( \frac{\partial \phi(\mathbf{x})}{\partial x_j} \cdot \frac{1}{s_j} - \frac{\partial \phi(\mathbf{x})}{\partial x_\ell} \cdot \frac{1}{s_\ell} \right) - \frac{26m^4}{\varepsilon^4} \cdot \frac{3\delta^2}{s_{\min}^2},$$

where  $s_{\min} = \min_{\ell \in C_\ell} s_\ell \geq \frac{\varepsilon}{m} b$  and  $\delta = \frac{\varepsilon^7 b}{312m^6}$ . Plugging these values in and simplifying, we have:

$$\phi(\mathbf{x}') - \phi(\mathbf{x}) = \Omega\left(\frac{\varepsilon \delta}{b}\right) = \text{poly}\left(\varepsilon, \frac{1}{n}, \frac{1}{m}\right).$$

Since  $\phi(\mathbf{x})$  lies in  $\left(\log\left(\frac{\varepsilon^2}{5m^2}\right), n \log m\right)$ , the number of iterations is  $\text{poly}(m, n, \frac{1}{\varepsilon})$ .  $\square$

Combining the previous lemmas, the following theorem is now immediate by union bounds, where we assume  $\varepsilon > 0$  is a small enough constant. This completes the proof of Theorem [1.1](#).

THEOREM 3.4. *With probability at least  $1 - \frac{1}{\text{poly}(n, m, \frac{1}{\varepsilon})}$ , the local search algorithm in [Section 3.2](#) has running time  $\text{poly}(n, m, \frac{1}{\varepsilon})$ .*

*Proof.* By Lemma 3.8, if we estimate all the  $U_i(\mathbf{x})$  and  $\frac{\partial U_i(\mathbf{x})}{\partial x_j}$  within  $\pm\Delta$  additive error, the algorithm ends in  $\text{poly}(n, m, \frac{1}{\epsilon})$  iterations. Since within each iteration we estimate  $U_i$  and  $\frac{\partial U_i(\mathbf{x})}{\partial x_j}$  for  $\text{poly}(m)$  times, the total number of evaluations is bounded by  $N = \text{poly}(n, m, \frac{1}{\epsilon})$ . Then by Lemma 3.5, we need at most  $\text{poly}(N, m, \frac{1}{\Delta})$  samples for all additive errors to be bounded by  $\pm\Delta$  with probability  $1 - \frac{1}{\text{poly}(N)}$ . Since  $\Delta = \frac{\epsilon^6}{64n \cdot m^5} = \frac{1}{\text{poly}(n, m, \frac{1}{\epsilon})}$  and the number of iterations is bounded by  $\text{poly}(n, m, \frac{1}{\epsilon})$ , the total running time including sampling steps is  $\text{poly}(n, m, \frac{1}{\epsilon})$ .  $\square$

#### 4 Improved Approximation for Additive Utilities

In this section, we present a 9.27-core for the special case of additive utilities. Recall that for such utilities,  $u_i(S) = \sum_{j \in S} u_{ij}$ , where  $S \subseteq C$  is a subset of candidates and  $i \in W$  is a voter. Though we could use the approach in the previous section, we lose constant factors first because the local optimum to the Nash Welfare objective only finds an approximate fractional core, and secondly because the randomized rounding needs to scale down budgets to satisfy the size constraints.

We address the first issue by using an exact core solution to the fractional problem. We do this via the classic Lindahl equilibrium that we describe in Section 4.1. To address the second issue, we use dependent rounding that preserves the budget constraint with probability one for additive utilities, and we describe this in Section 4.2. This yields a 9.27-core, though we do not know how to implement the resulting algorithm in polynomial time, since the fractional solution is now via a fixed point argument.

**4.1 Fractional Solution: Lindahl Equilibrium** Our algorithm for constructing a committee in the approximate core will make use of the Lindahl equilibrium [34, 28]. This equilibrium yields a fractional committee that lies in the fractional core (Definition 2.1 for  $\alpha = 1$ ).

We follow the approach in [25] for specifying the Lindahl equilibrium. Let  $x_j \geq 0$  denote the fraction to which candidate  $j$  is chosen. Here, we will assume for technical reasons that this can be a quantity greater than 1. We assign endowment  $\frac{b}{n}$  to each voter  $i \in V$ , and a price  $p_{ij}$  of  $j \in C$  for  $i \in V$ . The Lindahl equilibrium is now defined as follows.

**DEFINITION 4.1. (LINDAHL EQUILIBRIUM)** Let  $p_{ij}$  be the price of  $j \in C$  for  $i \in V$ , and let  $x_j \geq 0$  be the fraction with which item  $j$  is allocated. The prices and allocations constitute a Lindahl Equilibrium if:

1. For all  $i \in V$ , suppose the voter computes allocation  $\mathbf{y} \geq 0$  that maximizes her utility  $U_i(\mathbf{y})$  subject to her endowment constraint  $\sum_{j \in C} p_{ij} y_j \leq \frac{b}{n}$ , then we have that  $\mathbf{y} = \mathbf{x}$ .
2. The profit of the allocation given by  $\sum_{i \in V} \sum_{j \in C} y_j p_{ij} - \sum_{j \in C} s_j y_j$ , is maximized when  $\mathbf{y} = \mathbf{x}$ .

The main result in [28] is the following theorem proved by a fixed point argument. The setting in [28] is much more general. Therefore, for the purpose of completeness, we present a simple and direct proof of this theorem and its consequences via complementarity theory [22] in Appendix A.

**THEOREM 4.1. ([28])** If the utility functions  $U_i(\mathbf{x})$  have continuous derivatives, are strictly increasing, and are strictly concave, then a Lindahl equilibrium always exists.

**4.1.1 Properties of the Lindahl Equilibrium** We will need the following corollary to Theorem 4.1, which for the purpose of completeness, we also prove in Appendix A.

**COROLLARY 4.1.** It holds for a Lindahl equilibrium that:

1. For all voters  $i \in V$ , we have  $\sum_j p_{ij} x_j = \frac{b}{n}$ .
2. For each candidate with  $x_j > 0$ , we have  $\sum_i p_{ij} = s_j$ , and for each candidate with  $x_j = 0$ , we have  $\sum_i p_{ij} \leq s_j$ .
3.  $\sum_{j \in C} s_j x_j = b$ .

Given Corollary 4.1, it is easy and instructive to see that a Lindahl equilibrium lies in the core. Since this proof idea will be crucial to our analysis, we present it for completeness.

COROLLARY 4.2. ([28]) *The Lindahl equilibrium lies in the fractional core (Definition 2.1) for  $\alpha = 1$ .*

*Proof.* Given the equilibrium  $\mathbf{x}$ , suppose for the purpose of contradiction that there exists a subset  $S \subseteq V$  of  $t$  voters that can deviate and choose a committee  $\mathbf{z}$  of size  $\sum_j s_j z_j \leq \frac{t}{n}b$  such that  $U_i(\mathbf{z}) > U_i(\mathbf{x})$  for all  $i \in V$ . Since  $\mathbf{x}$  is utility maximizing for the voter  $i$  at endowment  $\frac{b}{n}$ , and since by Item 1 of Corollary 4.1, this endowment is spent exactly, the committee  $\mathbf{z}$  must cost more than  $b/n$ . Therefore, for all  $i \in V$ , we have:

$$\sum_j p_{ij} z_j > \frac{b}{n}.$$

However, by Item 2 of Corollary 4.1,  $\sum_i p_{ij} \leq s_j$  for all  $i \in S, j \in C$ . Summing the previous inequality over all  $i \in S$ , and applying  $\sum_i p_{ij} \leq s_j$ , we obtain:

$$\sum_j s_j z_j \geq \sum_{j,i \in V} p_{ij} z_j \geq \sum_{i \in S} \left( \sum_j p_{ij} z_j \right) > \sum_{i \in S} \frac{b}{n} = t \frac{b}{n}.$$

This contradicts the fact that  $\mathbf{z}$  could have been purchased with the endowments of  $i \in S$ , that is,  $\sum_j s_j z_j \leq t \frac{b}{n}$ , and thus completes the proof.  $\square$

**4.1.2 Lindahl Equilibrium for Additive Utilities** So far, we have presented the Lindahl equilibrium in its generality for continuous, concave, non-decreasing utilities. We now specialize it to additive utilities. Given the additive utility function  $u_i$ , we make it continuous and unconstrained using the following natural definition:

$$U_i(\mathbf{x}) := \sum_{j \in C} \min(1, x_j) \cdot u_i(\{j\}).$$

It is clearly a concave function.<sup>2</sup> Given this utility function, the Lindahl equilibrium is defined exactly as in Definition 4.1, and this implies Corollary 4.1 holds as is. The Lindahl equilibrium computation will be encapsulated by a subroutine below.

DEFINITION 4.2. (SUBROUTINE LINDAHL( $C, W, \{U_i\}, B$ )) *Given the set of candidates  $C$ , a subset  $W \subseteq V$  of voters with continuous and concave utilities  $\{U_i\}$ , and a size constraint  $B$ , this procedure finds a fractional solution  $\mathbf{x} \geq 0$  over  $C$  with  $\sum_{j \in C} s_j x_j \leq B$ , such that this solution is a Lindahl equilibrium.*

Note that the size  $B$  could be different from  $b$ . Using Definition 2.1 the solution  $\mathbf{x}$  of this subroutine satisfies the following condition: There is no subset  $S \subseteq W$  of voters who can find a committee  $\mathbf{z} \geq 0$  over  $E$ , with  $\sum_{j \in E} s_j z_j \leq \frac{|S|}{|W|} \cdot B$  such that  $U_i(\mathbf{z}) > U_i(\mathbf{x})$  for all  $i \in S$ .

**4.2 Randomized Rounding and Satisfied Voters** Note that if in the Lindahl equilibrium, we have  $x_j > 1$  for some  $j \in C$ , then all items must be integrally allocated. This is because we assumed  $U_i(\mathbf{x}) = u_i^+(\mathbf{y})$  where  $y_j = \min(1, x_j)$ . This means that if  $x_j > 1$ , then it can be reduced to 1 without affecting any utilities. Therefore, if some item is fractionally allocated, we can decrease the allocation of  $j$  with  $x_j > 1$  and increase the allocation of a fractional item, improving the utility of some agent. This violates the core condition. We will therefore assume throughout that  $x_j \in [0, 1]$  for all  $j \in C$ .

**4.2.1 Rounding Procedure** ROUND( $C, W, \{u_i\}, B$ ) Assume we are given a fractional solution  $\mathbf{x} \geq 0$  to LINDAHL( $C, W, \{u_{ij}\}_{i,j}, B$ ). The procedure ROUND( $C, W, \{u_{ij}\}, B$ ) uses the algorithm of [12] to the fractionally allocated items in this solution. This produces a committee  $(O, \ell)$ , where the candidates in  $O$  are chosen integrally, at most candidate  $\ell$  is chosen fractionally, and the total size is always at most  $B$ .

<sup>2</sup>As a technicality, to make the function satisfy the preconditions in Theorem 4.1, we perturb  $U_i$  slightly to make it be strictly increasing, strictly concave, and have continuous derivative. This perturbation preserves  $U_i$  to within a  $(1 + \epsilon)$  approximation for  $\epsilon > 0$  being arbitrary small, which will suffice for our approximation guarantees.

Let  $X_j$  be the random variable denoting whether candidate  $j$  is selected:  $X_j = 1$  if  $j \in O$ ,  $X_j = 0$  if  $j \notin O \cup \{\ell\}$ , and  $X_j$  is the fractional weight if  $j = \ell$ . Then the procedure ROUND preserves the marginals:  $\mathbb{E}[X_j] = \min(1, x_j)$  for all  $j \in C$ , and the  $\{X_j\}$  are negatively correlated [30, 38]. Note that the candidates with  $X_j = 1$  define  $O$ , while the one with  $X_\ell \in (0, 1)$  is the fractionally chosen candidate (it may not exist).

Note that

$$(4.7) \quad \mathbb{E} \left[ u_{i\ell} + \sum_{j \in O} u_{ij} \right] \geq \mathbb{E} \left[ \sum_{j \in C} u_{ij} X_j \right] = \sum_{j \in C} u_{ij} \min(1, x_j) = U_i(\mathbf{x})$$

where the expectation is over the random choice of  $O$ . Further, the size constraint is never violated, so that:

$$\sum_{j \in O} s_j \leq \sum_{j \in C} s_j X_j \leq B$$

regardless of the outcome of the rounding procedure. Note finally that any candidate that is fully allocated by  $\mathbf{x}$ , that is, with  $x_j \geq 1$ , must be present in  $O$ .

**4.2.2 Satisfied Voters** The procedure  $\text{ROUND}(C, W, \{u_{ij}\}, B)$  returns the committee  $O$ . Given the definition of  $\gamma$ -satisfied from Definition 3.3, we can strengthen Theorem 3.2 as follows:

**THEOREM 4.2.** (CONSTANT FRACTION OF CONSTANT-SATISFIED VOTERS) *Given the fractional solution  $\mathbf{x}$  produced by  $\text{LINDAHL}(C, W, \{u_{ij}\}, B)$  where  $|W| = n'$ , there is a integral committee  $O$  produced by  $\text{ROUND}(C, W, \{u_{ij}\}, B)$  with at least  $(1 - \beta)n'$   $\gamma$ -satisfied voters, where  $\beta = \gamma e^{1-\gamma}$ .*

Proving Theorem 4.2 requires a concentration bound for the sum of negatively correlated weighted Bernoulli random variables.

**LEMMA 4.1.** ([38]) *Let  $B_1, B_2, \dots, B_k$  be  $k$  negatively correlated Bernoulli random variables. Let  $X = \sum_{i=1}^k \beta_i B_i$ , where  $\beta_i \in [0, 1]$ . Let  $\mathbb{E}[X] \geq \mu$ . Then for any constant  $\delta \in (0, 1)$  we have:*

$$\Pr[X < (1 - \delta)\mu] < \left( \frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}} \right)^\mu.$$

Using the above bound, we complete the proof of Theorem 4.2

*Proof.* (Proof of Theorem 4.2) Let  $I \subseteq C$  be the set of fully allocated candidates in  $\text{LINDAHL}(C, W, \{u_{ij}\}_{i,j}, B)$  (i.e.  $I = \{j \in C \mid x_j \geq 1\}$ ), and let  $F \subseteq C$  be the set of fractionally allocated ones (i.e.  $F = \{j \in C \mid x_j \in (0, 1)\}$ ). Let  $u_i(F) = \sum_{j \in F} x_j u_{ij}$ . There are two cases:

**Case (1):** Suppose there is some candidate  $q \in F$  so that  $u_{iq} \geq \frac{u_i(F)}{\gamma}$ . In this case, voter  $i$  is  $\gamma$ -satisfied with probability 1, since

$$u_{iq} + \sum_{j \in O} u_{ij} \geq \frac{\sum_{j \in F} x_j u_{ij}}{\gamma} + \sum_{j \in I} u_{ij} \geq \frac{\sum_{j \in E} u_{ij} \min(1, x_j)}{\gamma},$$

where the first step uses that  $I \subseteq O$ .

**Case (2):** Suppose for every candidate  $j \in F$ ,  $u_{ij} < \frac{u_i(F)}{\gamma}$ . In this case, we consider the candidate  $\ell$  from the procedure ROUND as the additament, and invoke our concentration bound of Lemma 4.1. Notice  $X_j$ 's are negatively correlated Bernoulli random variables. Let  $\mu = \mathbb{E} \left[ \sum_{j \in F} u_{ij} \cdot \frac{\gamma}{u_i(F)} \cdot X_j \right] = \gamma$  and  $\delta = 1 - \frac{1}{\gamma}$ . By Lemma 4.1, we have

$$\Pr \left[ \sum_{j \in F} u_{ij} \cdot \frac{\gamma}{u_i(F)} \cdot X_j < 1 \right] < \left( \frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}} \right)^\mu = \gamma e^{1-\gamma}.$$

If the event  $\sum_{j \in F} u_{ij} \cdot \frac{\gamma}{u_i(F)} \cdot X_j < 1$  does not happen, the utility of  $i$  with additament  $\ell$  is at least

$$\begin{aligned} u_{i\ell} + \sum_{j \in O} u_{ij} &\geq u_{i\ell} X_\ell + \sum_{j \in F \setminus \{\ell\}} u_{ij} X_j + \sum_{j \in I} u_{ij} = \sum_{j \in F} u_{ij} X_j + \sum_{j \in I} u_{ij} \\ &\geq \frac{u_i(F)}{\gamma} + \sum_{j \in I} u_{ij} \geq \frac{\sum_{j \in E} u_{ij} \min(1, x_j)}{\gamma}, \end{aligned}$$

where the first step again uses that  $I \subseteq O$ .  $\square$

**4.3 The Constant Approximation to the Core** In this section, we modify the algorithm in [Section 3.4](#) and prove that it returns a committee in an  $O(1)$ -core. The analysis of our algorithm now critically requires the market-clearing properties of the Lindahl equilibrium presented in [Definition 4.1](#), [Corollary 4.1](#) and [Corollary 4.2](#).

**Algorithm.** The algorithm is the same as [Algorithm 2](#) except the following lines:

- **Line 7:**  $\mathbf{x}_t \leftarrow \text{LINDAHL}(C, V_t, \{U_i\}, b_t)$ .
- **Line 8:**  $O_t \leftarrow \text{Solution of ROUND}(C, V_t, \{u_i\}, b_t)$  that satisfies Theorem [4.2](#).

**Analysis.** The analysis follows the same outline as that in [Section 3.4.2](#). First, using the same argument as in that section, the solution  $T^*$  is feasible for the size  $b$ .

As before, we proceed to show a contradiction. Let  $S$  denote the set of voters that deviate, and let  $A$  denote the set of items they deviate to. We have  $\sum_{j \in A} s_j \leq \frac{|S|}{n} \cdot b$ , and  $A$  provides an  $\alpha$ -factor larger utility to voters in  $S$  even after including any additament. Consider the voters in  $W_t$ , and let  $S_t = S \cap W_t$ . These voters are  $\gamma$ -satisfied by  $O_t$  with respect to the fractional solution  $\mathbf{x}_t$  ([Definition 3.3](#)). Therefore, if  $i \in S_t$  deviates to  $A$  to obtain an  $\alpha$ -factor larger utility, it must be that  $U_i(A) \geq \frac{\alpha}{\gamma} \cdot U_i(\mathbf{x}_t)$ . Let  $\theta = \frac{\alpha}{\gamma}$ . We will assume  $\theta > 1$  below.

We will now show the analog of Theorem [3.1](#) using the prices computed by the Lindahl equilibrium. Let  $p^t$  denote the prices computed by  $\text{LINDAHL}(C, V_t, \{U_i\}, b_t)$ , and let  $n_t = |V_t|$ . The following lemma generalizes [Corollary 4.2](#) and bounds the price of the set  $A$  via the optimality conditions of the Lindahl equilibrium  $\mathbf{x}_t$ .

**LEMMA 4.2.** *For all  $i \in S_t$ , we have  $\sum_{j \in A} p_{ij}^t \geq \theta \cdot \frac{b_t}{n_t}$ .*

*Proof.* First note that  $x_{tj} \leq 1$  for all items  $j$ . If any  $x_{tj} > 1$ , then all items must be integrally allocated (see [Section 4.1](#)), which means  $A$  cannot achieve a  $\theta$ -factor larger utility. Note that  $\mathbf{x}$  is the utility maximizing solution to a packing problem for voter  $i$  where the “size” of item  $j$  is  $p_{ij}^t$  and the “size” constraint is  $\frac{b_t}{n_t}$ . This constraint is exactly satisfied by [Item 1](#) of [Corollary 4.1](#). Since  $U_i$  is concave, any solution that produces  $\theta$  factor more utility must have “size” at least  $\theta$  times larger. Therefore, the price of  $A$  is  $\theta$  times larger than that of  $\mathbf{x}_t$ , completing the proof.  $\square$

We now bound the size of set  $S_t$  as follows:

**LEMMA 4.3.** (ANALOG OF THEOREM [3.1](#))  $|S_t| \leq \frac{n_t}{b_t} \cdot \frac{\sum_{j \in A} s_j}{\theta}$ .

*Proof.* Summing the bound in [Lemma 4.2](#) over  $i \in S_t$ , we have

$$\sum_{i \in S_t} \sum_{j \in A} p_{ij}^t \geq |S_t| \cdot \theta \cdot \frac{b_t}{n_t}.$$

By [Item 2](#) of [Corollary 4.1](#), we have

$$\sum_{i \in V_t} p_{ij}^t \leq s_j \quad \forall j \quad \implies \quad \sum_{i \in S_t} \sum_{j \in A} p_{ij}^t \leq \sum_{j \in A} s_j.$$

Combining these inequalities completes the proof.  $\square$

This proves the analog of Theorem 3.1. Continuing as before, we sum Lemma 4.3 over all  $t$ , and using  $\sum_{j \in A} s_j \leq \frac{|S|}{n}b$ , we have:

$$|S| = \sum_t |S_t| \leq \frac{\sum_{j \in A} s_j}{\theta} \cdot \sum_t \frac{n_t}{b_t} \leq \frac{|S|}{n} \cdot \frac{b}{\theta} \cdot \sum_t \frac{n_t}{b_t}.$$

Therefore, for a blocking coalition to exist, we need:

$$(4.8) \quad \frac{b}{n} \cdot \sum_t \frac{n_t}{b_t} \geq \theta = \frac{\alpha}{\gamma}.$$

We will now set the parameters  $\omega, \gamma, \alpha$  so that the above inequality is false. First note by Theorem 4.2 that  $n_{t+1} \leq \beta n_t$  where  $\beta = \gamma e^{1-\gamma}$ . Further,  $b_{t+1} = \omega b_t$ . Therefore,

$$\frac{n_{t+1}}{b_{t+1}} \leq \frac{\beta}{\omega} \cdot \frac{n_t}{b_t}$$

with  $\frac{n_0}{b_0} = \frac{n}{(1-\omega)b}$ . Therefore,

$$(4.9) \quad \frac{b}{n} \cdot \sum_t \frac{n_t}{b_t} \leq \frac{1}{(1-\omega)} \cdot \sum_{t \geq 0} \left(\frac{\beta}{\omega}\right)^t = \frac{\omega}{(1-\omega)(\omega-\beta)}.$$

Combining Eq. (4.8) and Eq. (4.9), for a blocking coalition to exist, we need

$$\alpha \leq \frac{\omega\gamma}{(1-\omega)(\omega-\beta)} = \frac{\omega\gamma}{(1-\omega)(\omega-\gamma e^{1-\gamma})},$$

where the last step uses  $\beta = \gamma e^{1-\gamma}$  according to Theorem 4.2.

For an  $\alpha$  slightly larger than the right-hand side, a blocking coalition will therefore not exist. Plugging  $\omega = 0.15$  and  $\gamma = 6.7$  shows  $\alpha < 9.27$ , which yields the following theorem:

**THEOREM 4.3. (ADDITIVE UTILITIES)** *For additive utilities, a 9.27-core is always non-empty.*

## 5 Lower Bounds

In this section, we provide lower bound examples for general monotone utilities and monotone submodular utilities. The former result rules out extending our constant factor bound to general monotone utilities, while the latter shows that for submodular utilities, there is a lower bound on approximation of an absolute constant  $c > 1$ .

### 5.1 General Monotone Utilities

**THEOREM 5.1. (GENERAL UTILITY LOWER BOUND)** *For general monotone utilities, for any function  $\varphi : \mathbb{Z}^+ \times \mathbb{Z}^+ \rightarrow \mathbb{R}^+$ , a  $\varphi(n, m)$ -core can be empty.*

Theorem 1.3 is proved by the following example. (The same structure of 2 groups of 3 cyclically symmetric voters appears in [26, 40].)

*Example.* We have  $n = 6$  voters and  $m = 30$  candidates. The candidates are grouped into 6 disjoint sets, each of which contains 5 candidates and is called a “gadget”. We name the gadgets  $g_1, \dots, g_6$ . Each voter  $i$  has a favorite gadget  $g_{f_i}$  and a second favorite gadget  $g_{s_i}$ , given by:

$$\begin{aligned} f_1 &= 1, f_2 = 2, f_3 = 3, f_4 = 4, f_5 = 5, f_6 = 6; \\ s_1 &= 2, s_2 = 3, s_3 = 1, s_4 = 5, s_5 = 6, s_6 = 4. \end{aligned}$$

For any committee  $E$ , let  $x_i(E) = \frac{1}{5}|E \cap g_{f_i}|$  and  $y_i(E) = \frac{1}{5}|E \cap g_{s_i}|$ , denoting the fraction of candidates in the favorite / second favorite gadget of voter  $i$  being selected into  $E$ , respectively. The utility of voter  $i$  on  $E$  is given by

$$u_i(E) = (\alpha + 1) \cdot \mathbb{1}[x_i(E) = 1] + \mathbb{1}[y_i(E) = 1].$$

Here  $\alpha = \varphi(n, m) = \varphi(6, 30)$ . Her utility is monotone and supermodular.

Each candidate is of unit size 1 and the budget  $b = 15$ . For any feasible committee  $E$ , there must be at least 3 gadgets  $g$ 's with  $|E \cap g| \leq 3$  – otherwise the committee has at least  $4 \cdot 4 = 16 > b$  candidates. Therefore, either  $\{g_1, g_2, g_3\}$  or  $\{g_4, g_5, g_6\}$  includes at least 2 gadgets with  $|E \cap g| \leq 3$ . Without loss of generality, assume  $\{g_1, g_2, g_3\}$  does and  $|E \cap g_1| \leq 3$ ,  $|E \cap g_2| \leq 3$ . In this case, voters 1 and 2 can deviate and buy  $g_2$ , as they have a budget of  $b \cdot \frac{2}{n} = 5$ . For any additaments  $q$  and  $q'$ :

$$\begin{aligned} u_1(g_2) &= 1, & u_1(E \cup \{q\}) &= 0; \\ u_2(g_2) &= \alpha + 1, & u_2(E \cup \{q'\}) &\leq 1. \end{aligned}$$

We have  $u_1(g_2) > \alpha u_1(E \cup \{q\})$  and  $u_2(g_2) > \alpha u_2(E \cup \{q'\})$ .  $\square$

**5.2 Submodular Utilities** Next, we modify [Section 5.1](#) to show a lower bound for monotone submodular utilities.

**THEOREM 5.2. (SUBMODULAR LOWER BOUND)** *For monotone submodular utilities, a 1.015-core can be empty.*

*Example.* We use the same setting as [Section 5.1](#) except the utility of each voter  $i$  is given by

$$u_i(E) = x_i(E) + z \cdot (1 - x_i(E)) \cdot y_i(E),$$

where  $z \in (0, 1)$  is a constant to be determined later.

**LEMMA 5.1.** *The function  $u_i$  is monotone and submodular.*

*Proof.* Fix any  $E \subseteq C$  and  $t \notin E$  and consider  $u_i(E \cup \{t\}) - u_i(E)$ . Since  $g_{f_i} \cap g_{s_i} = \emptyset$ , this  $t$  lies in one of these two sets but not both. Suppose  $t \in g_{s_i}$ . Then,

$$u_i(E \cup \{t\}) - u_i(E) = z \cdot (1 - x_i(E)) \cdot (y_i(E \cup \{t\}) - y_i(E)) \geq 0.$$

Similarly, if  $t \in g_{f_i}$ , we have

$$u_i(E \cup \{t\}) - u_i(E) = (1 - z y_i(E)) \cdot (x_i(E \cup \{t\}) - x_i(E)) \geq 0,$$

where we have used that  $y_i(E) \leq 1$  and  $z \in [0, 1]$ . This shows that  $u_i$  is a monotone function.

Similarly, if  $E \subseteq E'$ , then  $1 - x_i(E) \geq 1 - x_i(E')$  since the coverage function is monotone. Further, by the submodularity of the coverage function, we have

$$y_i(E \cup \{t\}) - y_i(E) \geq y_i(E' \cup \{t\}) - y_i(E').$$

Therefore, if  $t \in g_{s_i}$ , then

$$\begin{aligned} u_i(E \cup \{t\}) - u_i(E) &= z \cdot (1 - x_i(E)) \cdot (y_i(E \cup \{t\}) - y_i(E)) \\ &\geq z \cdot (1 - x_i(E')) \cdot (y_i(E' \cup \{t\}) - y_i(E')) \\ &= u_i(E' \cup \{t\}) - u_i(E'). \end{aligned}$$

A similar argument for the case where  $t \in g_{f_i}$  completes the proof of submodularity.  $\square$

For any feasible committee  $E$ , again without loss of generality, assume  $|E \cap g_1| \leq 3$ ,  $|E \cap g_2| \leq 3$ . In this case, voters 1 and 2 can deviate and buy  $g_2$ . For any additaments  $q$  and  $q'$ :

$$\begin{aligned} u_1(g_2) &= z, & u_1(E \cup \{q\}) &\leq 0.8 + z \cdot 0.2 \cdot 0.6; \\ u_2(g_2) &= 1, & u_2(E \cup \{q'\}) &\leq 0.8 + z \cdot 0.2 \cdot 1. \end{aligned}$$

When  $z = \frac{\sqrt{689}-17}{10} \approx 0.925$ , the gap is  $\min\left(\frac{u_1(g_2)}{u_1(E \cup \{q\})}, \frac{u_2(g_2)}{u_2(E \cup \{q'\})}\right) \geq \frac{5\sqrt{689}-115}{16} > 1.015$ .  $\square$

## 6 Conclusions and Open Questions

Our work brings up several open questions. First is the existence of polynomial time computable pricing rules that approximate the core. The work of [40] shows a price increase process that runs in polynomial time and provides a logarithmic approximation to the core for additive utilities. However, the prices are *common* to the voters, as opposed to the per-voter market clearing prices of the Lindahl equilibrium, and we do not know how to compute the latter in polynomial time. Is there an intermediate price tatonnement scheme that not only runs in polynomial time, but also achieves a constant approximation?

Next, our approximation bounds are far from tight. In particular, we have not ruled out the existence of a 1-core for additive utility, or an  $O(1)$ -core for subadditive or XOS utilities. To address the latter, we would need to understand the approximability of the core when utilities of voters are continuous and the maximum of linear functions. Such utilities are convex, but have a specific form that may be amenable to better approximation results than the worst case illustrated in Theorem [1.3]. We leave this as an interesting open question.

### A Existence of Lindahl Equilibrium

For the purpose of completeness, we present a direct proof of the existence of Lindahl equilibrium, showing Theorem [4.1] and Corollary [4.1]. We restate Theorem [4.1] below.

**THEOREM A.1.** ([28]) *If the utility functions  $U_i(\mathbf{x})$  have continuous derivatives, are strictly increasing, and are strictly concave, then a Lindahl equilibrium always exists.*

Following [25], define

$$f_j(\mathbf{x}) = - \sum_{i \in V} \frac{\frac{\partial U_i(\mathbf{x})}{\partial x_j}}{\sum_{\ell \in C} x_\ell \frac{\partial U_i(\mathbf{x})}{\partial x_\ell}}.$$

Applying the constrained complementarity theorem of [22] to this continuous function  $\{f_j\}$  with the constraint  $\sum_{j \in C} s_j x_j \leq b$ , there exists  $\mathbf{x}, \mathbf{y} \geq 0$ , and scalar  $z \geq 0$  such that:

- For all  $j \in C$ ,  $f_j + z s_j = y_j$ ; and  $\mathbf{x} \cdot \mathbf{y} = 0$ .
- $\sum_{j \in C} x_j s_j \leq b$ , and  $z(b - \sum_{j \in C} x_j s_j) = 0$ .

In the sequel, we will focus on this solution. We will first show that  $z > 0$  so that  $\sum_{j \in C} s_j x_j = b$ . Otherwise, for all  $j \in C$ , we have  $f_j = y_j$ . However,  $f_j < 0$  since the utilities are strictly increasing, while  $y_j \geq 0$  by assumption. This is a contradiction.

Let  $C' = \{j \in C \mid x_j > 0\}$ . For  $j \in C'$ , we have  $y_j = 0$  so that  $f_j = -z s_j$ . Multiplying by  $x_j$  and summing, we have

$$\sum_{j \in C'} x_j f_j = - \sum_{i \in V} \frac{\sum_{j \in C'} x_j \frac{\partial U_i(\mathbf{x})}{\partial x_j}}{\sum_{\ell \in C} x_\ell \frac{\partial U_i(\mathbf{x})}{\partial x_\ell}} = - \sum_{i \in V} 1 = -n = -z \sum_{j \in C'} s_j x_j = -z b.$$

Therefore,  $z = n/b$ .

Set  $p_{ij} = \frac{b}{n} \frac{\frac{\partial U_i(\mathbf{x})}{\partial x_j}}{\sum_{\ell \in C} x_\ell \frac{\partial U_i(\mathbf{x})}{\partial x_\ell}}$  for all  $i \in V, j \in C$ . Note that  $\sum_i p_{ij} = -\frac{b}{n} f_j$ . This implies the following:

- For all  $i \in V$ , we have  $\sum_j p_{ij} x_j = \frac{b}{n} \frac{\sum_{j \in C} x_j \frac{\partial U_i(\mathbf{x})}{\partial x_j}}{\sum_{\ell \in C} x_\ell \frac{\partial U_i(\mathbf{x})}{\partial x_\ell}} = \frac{b}{n}$ .
- For all  $j \in C'$ , since  $f_j + z s_j = 0$  and  $z = \frac{b}{n}$ , we have  $\sum_i p_{ij} = s_j$ .
- For all  $j \in C$ , since  $f_j + z s_j = y_j \geq 0$ , we have  $\sum_i p_{ij} \leq s_j$ .
- For all  $i \in V$  and for all  $j, \ell \in C$ , we have  $p_{ij} \frac{\partial U_i}{\partial x_\ell} = p_{i\ell} \frac{\partial U_i}{\partial x_j}$ .

The first three consequences, along with  $\sum_{j \in C} s_j x_j = b$  prove Corollary [4.1]. By simple gradient optimality, these conditions also show Item [2] of Definition [4.1]. To see this, note that Definition [4.1] does not constrain the allocation  $\mathbf{y}$ . Therefore, for the profit to be finite, we have  $\sum_i p_{ij} - s_j \leq 0$  for all  $j$ . Further, if any of these

inequalities is strict, the profit is only larger if  $x_j = 0$ . Therefore, if  $x_j > 0$ , it must force this inequality to be tight. These are exactly the conditions we derived above, which means this solution satisfies [Item 2](#). Similarly, since  $U_i$  is strictly increasing and concave, the last condition derived above is the gradient optimality condition for [Item 1](#) in [Definition 4.1](#). To see this, note that the gradient optimality condition of [Item 1](#) can be written as  $\frac{\partial U_i}{\partial x_j} = \lambda_i p_{ij}$  for all  $i \in V, j \in C$ . Therefore,  $p_{ij} \frac{\partial U_i}{\partial x_j} = p_{il} \frac{\partial U_i}{\partial x_l}$ , so that any solution satisfying the latter and with  $\sum_j p_{ij} x_j = \frac{b}{n}$  must satisfy [Item 1](#).

## References

- [1] The Stanford Participatory Budgeting Platform. <https://pbstanford.org>.
- [2] Shipra Agrawal, Yichuan Ding, Amin Saberi, and Yinyu Ye. Correlation robust stochastic optimization. In *SODA*, page 1087–1096, 2010.
- [3] Kenneth J. Arrow and Gerard Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, 22(3): 265–290, 1954.
- [4] Haris Aziz and Nisarg Shah. Participatory budgeting: models and approaches. In *Pathways Between Social Science and Computational Social Science*, pages 215–236. 2021.
- [5] Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2):461–485, 2017.
- [6] Haris Aziz, Edith Elkind, Shenwei Huang, Martin Lackner, Luis Sánchez Fernández, and Piotr Skowron. On the complexity of extended and proportional justified representation. In *AAAI*, pages 902–909, 2018.
- [7] Haris Aziz, Felix Brandt, Edith Elkind, and Piotr Skowron. *Computational Social Choice: The First Ten Years and Beyond*, pages 48–65. Springer International Publishing, 2019.
- [8] Siddharth Barman, Sanath Kumar Krishnamurthy, and Rohit Vaish. Finding fair and efficient allocations. In *EC*, page 557–574, 2018.
- [9] William C. Brainard and Herbert E. Scarf. How to compute equilibrium prices in 1891. *American Journal of Economics and Sociology*, 64(1):57–83, 2005.
- [10] Steven J. Brams, D. Marc Kilgour, and M. Remzi Sanver. A minimax procedure for electing committees. *Public Choice*, 132(3):401–420, 2007.
- [11] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. *Handbook of Computational Social Choice*. Cambridge University Press, USA, 1st edition, 2016. ISBN 1107060435.
- [12] Jaroslaw Byrka, Thomas W. Pensyl, Bartosz Rybicki, Aravind Srinivasan, and Khoa Trinh. An improved approximation for  $k$ -median and positive correlation in budgeted optimization. *ACM Trans. Algorithms*, 13(2): 23:1–23:31, 2017.
- [13] Yves Cabannes. Participatory budgeting: a significant contribution to participatory democracy. *Environment and urbanization*, 16(1):27–46, 2004.
- [14] Gruia Călinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM J. Comput.*, 40(6):1740–1766, 2011.
- [15] John R. Chamberlin and Paul N. Courant. Representative deliberations and representative decisions: Proportional representation and the borda rule. *The American Political Science Review*, 77(3):718–733, 1983.
- [16] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Dependent randomized rounding via exchange properties of combinatorial structures. In *FOCS*, pages 575–584, 2010.
- [17] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Submodular function maximization via the multilinear relaxation and contention resolution schemes. *SIAM J. Comput.*, 43(6):1831–1879, 2014.
- [18] Xingyu Chen, Brandon Fain, Liang Lyu, and Kamesh Munagala. Proportionally fair clustering. In *ICML*, pages 1032–1041, 2019.
- [19] Yu Cheng, Zhihao Jiang, Kamesh Munagala, and Kangning Wang. Group fairness in committee selection. *ACM Trans. Economics and Comput.*, 8(4):23:1–23:18, 2020.
- [20] Richard Cole and Vasilis Gkatzelis. Approximating the nash social welfare with indivisible items. *SIGecom Exch.*, 14(1):84–88, November 2015.
- [21] H. R. Droop. On methods of electing representatives. *Journal of the Statistical Society of London*, 44(2):141–202, 1881.
- [22] B. C. Eaves. On the basic theorem of complementarity. *Mathematical Programming*, 1(1):68–75, 1971.
- [23] Edmund Eisenberg and David Gale. Consensus of subjective probabilities: The pari-mutuel method. *The Annals of Mathematical Statistics*, 30(1):165–168, 1959.
- [24] Ulle Endriss. *Trends in Computational Social Choice*. Lulu.com, 2017. ISBN 1326912097.
- [25] Brandon Fain, Ashish Goel, and Kamesh Munagala. The core of the participatory budgeting problem. In *WINE*, pages 384–399, 2016.

- [26] Brandon Fain, Kamesh Munagala, and Nisarg Shah. Fair allocation of indivisible public goods. In *EC*, pages 575–592, 2018.
- [27] Luis Sánchez Fernández, Edith Elkind, Martin Lackner, Norberto Fernández García, Jesús Arias-Fisteus, Pablo Basanta-Val, and Piotr Skowron. Proportional justified representation. In *AAAI*, pages 670–676, 2017.
- [28] Duncan K Foley. Lindahl’s solution and the core of an economy with public goods. *Econometrica*, pages 66–72, 1970.
- [29] Eric J. Friedman, Vasilis Gkatzelis, Christos-Alexandros Psomas, and Scott Shenker. Fair and efficient memory sharing: Confronting free riders. In *AAAI*, pages 1965–1972, 2019.
- [30] Rajiv Gandhi, Samir Khuller, Srinivasan Parthasarathy, and Aravind Srinivasan. Dependent rounding and its applications to approximation algorithms. *J. ACM*, 53(3):324–360, 2006.
- [31] Ashish Goel, Anilesh K. Krishnaswamy, Sukolsak Sakshuwong, and Tanja Aitamurto. Knapsack voting for participatory budgeting. *ACM Trans. Econ. Comput.*, 7(2), July 2019.
- [32] Zhihao Jiang, Kamesh Munagala, and Kangning Wang. Approximately stable committee selection. In *STOC*, pages 463–472, 2020.
- [33] Mayuresh Kunjir, Brandon Fain, Kamesh Munagala, and Shivnath Babu. ROBUST: fair cache allocation for data-parallel workloads. In *SIGMOD*, pages 219–234, 2017.
- [34] Erik Lindahl. Just taxation—a positive solution. In *Classics in the theory of public finance*, pages 168–176. 1958.
- [35] Burt L. Monroe. Fully proportional representation. *The American Political Science Review*, 89(4):925–940, 1995.
- [36] Kamesh Munagala, Zeyu Shen, and Kangning Wang. Optimal algorithms for multiwinner elections and the chamberlin-courant rule. In *EC*, pages 697–717, 2021.
- [37] John F. Nash. The bargaining problem. *Econometrica*, 18(2):155–162, 1950.
- [38] Alessandro Panconesi and Aravind Srinivasan. Randomized distributed edge coloring via an extension of the chernoff-hoeffding bounds. *SIAM J. Comput.*, 26(2):350–368, 1997.
- [39] Dominik Peters and Piotr Skowron. Proportionality and the limits of welfarism. In *EC*, pages 793–794, 2020.
- [40] Dominik Peters, Grzegorz Pierczyński, and Piotr Skowron. Proportional participatory budgeting with additive utilities. In *NeurIPS*, 2021.
- [41] Thorvald N Thiele. Om flerfoldsvalg. *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*, 1895:415–441, 1895.
- [42] Nicolaus Tideman and Daniel Richardson. Better voting methods through technology: The refinement-manageability trade-off in the single transferable vote. *Public Choice*, 103(1):13–34, 2000.
- [43] Jan Vondrak. Optimal approximation for the submodular welfare problem in the value oracle model. In *STOC*, page 67–74, 2008.
- [44] Qiqi Yan. Mechanism design via correlation gap. In *SODA*, page 710–719, 2011.