Mechanisms for Fair Attribution

Eric Balkanski* Yaron Singer†

Abstract

We propose a new framework for optimization under fairness constraints. The problems we consider model procurement where the goal is to optimize a buyer's utility while paying sellers in a manner that reflects their contribution to the buyer's utility. The payment rules we consider are natural interpretations of fairness based on concepts such as Shapley values and the nucleolus from cooperative game theory. The question in this setting is whether the outcome (measured in terms of the buyer's utility) produced by mechanisms that enforce fair payments is competitive with the outcome of a mechanism that simply pays agents their costs and is not committed to fair payments. Our main result shows that there exists a mechanism which guarantees a solution whose value is at least one third of the optimal unfair solution for any submodular utility function, and that this ratio is optimal. We discuss several special cases for which this approximation ratio can be improved and natural extensions.

1 Introduction

In a typical procurement setting, a buyer aims to purchase goods or services from agents. The agents associate some cost for their goods, and the buyer aims to optimize some objective function in a cost-effective manner. In standard mechanism design settings, the costs are private information and agents are strategic. In such scenarios a reasonable approach is to design truthful mechanisms that have desirable guarantees [28, 19, 6, 7, 10, 4, 5, 1]. In this paper we consider settings in which the agents' costs are known to the mechanism designer, and truthfulness is therefore not the design objective. Instead, we investigate cases in which the mechanism designer seeks to enforce payments that are *fair*, or more precisely, payments that reflect the sellers' contribution to the buyer's overall utility.

For concreteness, consider a set of agents, each selling a single good associated with some known cost, and a buyer who has a utility function over goods she aims to optimize under some given budget. As an extreme example consider two agents, one with some small cost $\epsilon > 0$ and some arbitrarily high utility to the buyer and another agent with cost $B - \epsilon$ and utility ϵ. A buyer which has a budget B and an additive utility function can apply the following mechanism: select both agents and pay each agent her cost. While it would be individually rational (every agent is paid at least her cost), there is something fundamentally unfair about this mechanism since it allocates the budget disproportionally to the agents' contribution.

*School of Engineering and Applied Sciences, Harvard University. ericbalkanski@g.harvard.edu.

†School of Engineering and Applied Sciences, Harvard University. yaron@seas.harvard.edu. This research was supported by NSF grant CCF-1301976 and a Google Faculty Research Award.

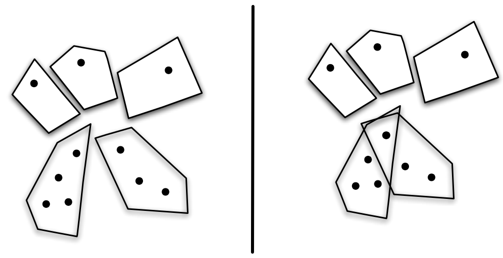

Figure 1: Example to illustrate fairness. Each agent is represented as a set and each agent's cost is one unit.

In a wide range of applications one seeks *fair attribution mechanisms*: mechanisms that attribute the budget in a manner that reflects the agents' contribution to the buyer's utility. Scenarios where this may arise include cases in which government agencies aim to procure goods or services and enforce fairness criteria. Other interesting examples occur in online advertising when several agencies advertise to a single user who makes a conversion after seeing multiple ads, and the reward from the conversion is split amongst the different advertisers proportionally to their contribution to the user's conversion [33].

Fair attribution mechanism design. In this paper we employ natural models of fairness and quantify the performance of mechanisms that use payments based on these concepts. For concreteness, we present the results through a natural notion of fairness that relies on the *Shapley value* from cooperative game theory. The results we describe in this paper however, extend to a far more general class of solution concepts, as we will later discuss. At a high level, we are motivated by the following question:

What is the price of fairness in procurement?

More concretely, we are interested in understanding whether the performance (in terms of buyer's utility) of a mechanism which pays agents "fairly" is comparable to that of a mechanism that optimizes the buyer's utility and simply pays each agent her cost.

1.1 Shapley-fair mechanisms

To illustrate the notion of fairness we intend to employ, consider the two caricature procurement problems in Figure 1. In this example each agent is represented as a set of elements, illustrated as black balls. The utility of the buyer from a set of agents is the total number of elements in the union of their sets. In the example on the left, if we select the two sets at the bottom the total value of the solution is 7 units, since one set contributes 4 units and the other contributes 3 and all sets are disjoint. Thus, if we select the two agents associated with the bottom sets and pay them in a 4 : 3 ratio that would encapsulate an intuitive notion of fairness. The example on the right is slightly more involved as the sets are no longer disjoint. A natural extension when selecting the two bottom sets would be to pay in a ratio of 3.5 : 2.5. This is derived from the following rationale: the bottom left set covers three elements uniquely (this contributes 3) and another element that is covered by another set in the solution (this contributes 0.5). Its contribution can therefore be interpreted as 3.5. Similarly, the contribution of the bottom right set can be interpreted as 2.5. This natural notion extends to a fundamental notion of fairness in cooperative game theory known as *Shapley values*.

Fairness through Shapley values. Given a utility function and some solution (a set of agents), the *Shapley value* of an agent in the solution is defined as her expected marginal contribution to the set of agents before her in a random permutation of the solution [27]. The Shapley values are often used in the *cost-sharing* literature to prescribe the fair amount an agent should pay when it shares a resource with a set of agents. In order to employ this concept for procurement, we define a concept of Shapley-fair payments that uses the Shapley values. For a given budget, solution, and utility function, the Shapley payment to an agent is her Shapley value normalized such that the sum of the payments to the agents in the solution is equal to the budget. By the efficiency property of Shapley values, which states that the sum of Shapley values is equal to the value of the solution, this normalizing term for all agents is the budget divided by the value of the solution.

We say that a mechanism is Shapley-fair (or for short, *fair*) if it pays the agents in the solution that it selects their Shapley payments. A solution is then *feasible* if these payments are individually rational. That is, the payment to any agent in the solution is at least her cost. Naturally, we seek Shapley-fair mechanisms that pick a feasible solution that maximizes the utility function of the buyer.

Example: coverage functions. For a simple example consider again the instance illustrated in Figure 1. We use the instance on the right where there is an intersection between the sets on the bottom, and assume that every agent has a cost of one unit and that we have a budget of three units. A Shapley-fair mechanism can select the bottom two agents and pay them their Shapley payments. Since the budget is three and the value of the set consisting of these two agents is six, the payment to the bottom left agent would be $3/6 \cdot (3 + 1/2) = 1.75$ and the payment to the bottom right agent would be $3/6 \cdot (2 + 1/2) = 1.25$. For both agents these payments are individually rational (above one unit), and thus this solution is feasible. It is easy to check that this solution is optimal: no other agent can be added since the remaining agents' Shapley values are all one, but since the budget divided by the value of this new solution is strictly less than one, the payment to another agent would not be individually rational.

The optimal *unfair* solution is the solution that maximizes utility by selecting agents whose total cost is under the budget and then pays them their cost. In this example, the two bottom sets and one of the top sets is an optimal unfair solution. This solution respects the budget and has total value of 7. This example illustrates a simple but important idea: regardless of computational constraints, the optimal fair solution can lead to qualitatively worse outcomes. Formally, we say that a Shapley-fair mechanism is an α-approximation if it outputs a fair solution that is an α-approximation to the optimal unfair solution. As we will now see, the performance largely depends on the utility function.

In general, fairness is unobtainable. It is not hard to see that there are simple examples of utility functions for which no fair solution with a finite approximation ratio exists. Suppose the utility function f is such that $f(\{a_1\}) = f(\{a_2\}) = 0$, and $f(\{a_1, a_2\}) = 1$, the budget is $B = 1$, and the agents' costs are $c_1 = 1$ and $c_2 = 0$. The optimal unfair solution is $S^* = \{a_1, a_2\}$. However, a fair solution cannot select both agents since it would need to pay both agents equally as they have the same contribution. The approximation ratio in this case is therefore unbounded. This example leads to the following question.

Are there utility functions for which fair attribution mechanisms perform well?

1.2 Main Results

In this paper we show that there are rich classes of utility functions for which fair attribution mechanisms perform well. In particular, we prove the following theorem.

Theorem 1. *For any submodular utility function there exists a fair mechanism which is a 3-approximation. Furthermore, this bound is tight as no fair mechanism can obtain an approximation better than* $3-\epsilon$*, for any fixed* $\epsilon > 0$*.*

We also show that when one allows for randomization, large market assumptions, or when the agents have uniform costs, the approximation ratio can be improved to two, and in all cases we show this is tight as well.

Beyond Shapley-fairness. Although Shapley payments capture the notion of fairness quite well, there may be other notions of fairness one may be interested in considering. One example is that of the *nucleolus* which is also standard in cooperative game theory. In Section 6 we extend our results to very broad classes of fairness that need to respect a very simple and intuitive property. We show that the nucleolus respects this criterion.

1.3 Paper organization.

We begin with a short preliminary section to formally introduce the model and our objectives. In Section 3 we describe an axiomatic characterization that justifies the use of Shapley values. We give our main result in Section 4 where we present and analyze a fair attribution mechanism that obtains an approximation ratio of three to the problem. We then continue to discuss special cases in which the approximation ratio can be improved. In Section 6 we extend the notion of fairness to broader concepts than Shapley values. In Section 7 we discuss interesting extensions and applications of our results to related scenarios. We conclude with a brief discussion and state some open questions that follow from this line of work.

1.4 Related work

Procurement has been studied in algorithmic mechanism design for over a decade. There have been two major directions of investigation: frugality and budget feasible mechanism design. Frugality considers the task of designing truthful mechanisms that aim to minimize payments. This was first studied by Archer and Tardos [3] and has been well developed in the past decade [34, 11, 17, 19, 6]. An alternative notion is that of budget feasible mechanism design initiated by Singer [28] where the goal is to optimize a utility function under a hard budget constraint. This concept enables surprisingly positive results and has been studied in algorithmic mechanism design [7, 10, 4, 5, 1] and used in various application domains that include social network analysis [29], crowdsourcing [36, 30, 32] and privacy auctions [9, 31]. The setting that we study most resembles that of budget feasible mechanism design. The crucial difference is that in our setting we do not seek truthfulness but fairness.

The main notion of fairness we use is that of Shapley values from his seminal paper [27]. Shapley gives an axiomatic characterization and shows that for convex games this value is in the core. The notion of core is a fundamental notion in cooperative games, first introduced by von Neumann and Morenstern [35]. Shapley values are studied in algorithmic mechanism design primarily in the context of cost-sharing [22, 13, 2, 15, 21]. For a survey of results in algorithmic mechanism design see [16].

The concept of Fairness has been heavily studied in social choice theory. In particular, there has been a great deal of interest in fair division problems [23, 18, 8, 25]. For a recent survey see [24].

The concept of attribution has been studied by Sun and Sundararajan [33]. Their main result shows that the Shapley-Shubik method (the continuous analogue of the Shapley values) is the only method which under reasonable conditions appropriately attributes change in the value of a continuous multilinear function. Their motivation for studying the problem is that of "beyond-the-last-click" in online marketing. Our work is in many ways a continuation to that of Sun and Sundararajan. While their work deals with characterization, in our work we provide a mechanism and guarantees that can be directly applied to the problems considered in their work.

2 Preliminaries

Setup. We consider a model in which there is a set of agents $N = \{a_1, \dots, a_n\}$ and a buyer. Each agent a_i has a cost $c_i \geq 0$ for the item she is selling, and this cost is known. The buyer has some budget $B \geq 0$ and a utility function $f : 2^N \to \mathbb{R}^+$ defined over sets of agents. Without loss of generality we assume that $\forall i \in [n] : \ c_i \leq B$. We will mainly consider submodular functions in this paper.

Submodularity. A function is *submodular* if for all $S \subseteq T$ and $a \in N \setminus T$ it has diminishing marginal returns:

$$f(S \cup a) - f(S) \geq f(T \cup a) - f(T)$$

Note that we abuse notation and write $S \cup a$ instead of $S \cup \{a\}$ when $a \in N$. Similarly, we write $S \setminus a$ instead of $S \setminus \{a\}$. An alternative definition is that for every S, T it holds that: $f(S \cup T) \leq f(S) + f(T) - f(S \cap T)$. We use $f_T(S)$ to denote the marginal contribution of $S \subseteq N$ given some $T \subseteq N$:

$$f_T(S) := f(S \cup T) - f(T).$$

Shapley Fairness. The Shapley value of agent a_i with respect to a set of agents $S \subseteq N$ such that $a_i \in S$ is defined as:

$$\phi_i(S) = \sum_{T \subseteq S \setminus a_i} \frac{|T|!(|S| - |T| - 1)!}{|S|!} f_T(a_i).$$

With value functions such as the Shapley value $\phi(\cdot)$, we emphasize that the *contribution* of agent a_i to $f(S)$, $\phi_i(S)$, is different than the *marginal contribution* of agent a_i to $f(S)$, $f_S(a_i)$. We note that from a computational perspective computing Shapley values *exactly* requires summing over exponentially many sets though we discuss methods to approximate this value with high precision in a computationally efficient manner in Section 4. For a given solution $S \subseteq N$, an agent $a_i \in S$, a budget B, and a utility function $f(\cdot)$, the Shapley payment to a_i denoted as p_i is:

$$p_i = \frac{B \cdot \phi_i(S)}{f(S)}.$$

We say that a mechanism is Shapley-fair (or for short, *fair*) if it pays the agents in the solution that it selects their Shapley payments. A solution is then *feasible* if these payments are individually rational. That is, the payment to any agent in the solution is at least her cost. Note that the budget constraint is always satisfied since the payments are defined to be such that their sum is exactly equal to the budget. Naturally we seek Shapley-fair mechanisms that pick a feasible solution that maximizes the utility function of the buyer:

$$\max \Big\{ f(S) : \frac{B \cdot \phi_i(S)}{f(S)} \geq c_i \text{ for all } a_i \in S, S \subseteq N \Big\}.$$

Approximation. We wish to obtain fair mechanisms which produce a solution that is comparable with that of the standard solution in terms of the utility to the buyer. In particular, a mechanism is an α-approximation if it outputs a solution S such that $\alpha f(S) \geq f(S^*)$ where S^* is an unfair solution given to the mechanism. That is, S^* is a solution to the following optimization problem:

$$\max \Big\{ f(S) : \sum_{a_i \in S} c_i \leq B, S \subseteq N \Big\}.$$

3 Axiomatic Attribution of the Budget

In this section, we describe natural axioms of fairness and show that mechanisms with Shapley payments are the only ones that respect these axioms. For a given utility function $f : 2^N \rightarrow \mathbb{R}^+$ and solution $S \subseteq N$ selected by the mechanism, then the payments to agents by a fair mechanism must satisfy:

- **Efficiency:** The entire budget must be split among agents, $\sum_{i \in S} p_i = B$.
- **Dummy:** An agent which does not contribute to $f(S)$ should not be paid: if $f(T) = f(T \cup a_i)$ for all $T \subseteq S$ such that $a_i \notin T$, then $p_i = 0$.
- **Anonymity:** If two agents a_i and a_j are equivalent in their contribution to $f(S)$, then they should receive the same payment: if $f(T \cup a_i) = f(T \cup a_j)$ for all $T \subseteq S$ which do not contain a_i and a_j, then $p_i = p_j$.
- **Proportionality:** The function $(f + g)(T) = f(T) + g(T)$ should attribute a payment to an agent which is proportional to the payment she would receive with the functions f and g separately with coefficients $f(S)$ and $g(S)$ respectively, $p_i(f + g) = \alpha(f(S)p_i(f) + g(S)p_i(g))$ for some α which only depends on f and g and not on a_i.

In the cost sharing literature, the Shapley value is the classical attribution method of a function f where the total gain $f(S)$ is divided among the agents. The Shapley value is considered to be the fair attribution method in such a situation since it is the only attribution method to satisfy the Shapley axioms: Efficiency, Dummy, Anonymity and Additivity, where efficiency divides $f(S)$ rather than B. In our model, the payments still need to satisfy Efficiency, Dummy and Anonymity but the Additivity axiom has been replaced by Proportionality: to divide the budget fairly in terms of contribution, the budget is split proportionally between $f(S)$ and $g(S)$.

The following theorem shows that our attribution method is the unique attribution method which satisfy the four axioms listed above. The proof is deferred to the full version of the paper.

Theorem 2. *The Shapley payments are the only payments that satisfy the Efficiency, Dummy, Anonymity, and Proportionality axioms.*

4 Fair optimization

In this section we prove our main result: for any submodular function, there is a set of agents who can be paid by Shapley payments and the value of the solution induced by these agents is no less than a third of the value of the optimal unfair solution. Furthermore, we show that this bound is tight in the sense that no mechanism that respects Shapley-fairness can obtain a better approximation ratio.

It is important to note that these results do not take computational restrictions into consideration. As we later discuss, when computation is involved, a computationally efficient mechanism can be implemented at the expense of losing an approximation ratio of $e/(e-1)$ and the Shapley payments inevitably need to be approximated to arbitrarily fine precision. Since we are interested in making the distinction between limitations due to fairness and those due to computation, we consider efficient computation as another restriction on the mechanism.

We will begin by introducing a natural greedy mechanism and show its approximation ratio cannot be better than $\Omega(\sqrt{n})$ (interestingly, this mechanism is a $O(\sqrt{m})$-approximation for coverage utility functions with universe size m, though we defer the analysis to the full version of the paper). The idiosyncrasies of the greedy mechanism lead to the SIEVE mechanism which is the working-horse in this paper. We then describe and analyze this mechanism and prove our main result.

4.1 The greedy mechanism is an $\Omega(\sqrt{n})$-approximation

In standard optimization of a monotone submodular function under a budget constraint, one often applies variants of a simple greedy algorithm to obtain constant factor approximation guarantees. The greedy algorithm simply adds the element which maximizes the bang-per-buck (i.e. the element whose marginal contribution-per-cost is greatest), until it runs out of budget. Variations like taking the maximum between the element with the largest utility and the solution returned by the greedy algorithm are used to protect from cases in which there is an element whose marginal-contribution-per-cost is always smaller than those of the rest of the elements though has substantially larger utility then the rest of the elements (e.g. an element with arbitrarily large utility $M >> B$ and cost B and elements with arbitrarily small cost and utility $\epsilon > 0$). The analogue of the greedy algorithm to the fair-design problem would be to sequentially add the agent whose marginal-contribution-per-cost is greatest until the cost of some participating agent exceeds its fair payment.

Lower bound on the greedy mechanism. Consider the instance in which every agent a_i has cost of $1 - \epsilon_i$ where $\epsilon_i > \epsilon_{i+1}$, and all $\{\epsilon_i\}_{i=1}^n$ are arbitrarily small. The budget is $B = n$. The function f is the coverage function where a_1 covers two elements e and e'. Agents a_i for $i > 1$ where i is even cover two elements as well, e and a unique element to each agent e_i, while agents a_i for $i > 1$ where i is odd cover e' and a unique element e_i. The greedy mechanism begins by selecting agent a_1, and then $a_2, a_3, \ldots$ sequentially. It is easy to check that after $2\sqrt{n} - 1$ steps, for the set of selected agents S we have that $f(S) = 2\sqrt{n}$, and $\phi_1(S) = 2/\sqrt{n}$. Thus:

$$p_1 = \phi_1(S) \cdot B/f(S) = 2/\sqrt{n} \cdot B/2\sqrt{n} = 1.$$

If the mechanism tries to add another agent into S after $2\sqrt{n} - 1$ steps this will break individual rationality since that would decrease $\phi_1(S)$ strictly below 1, and since ϵ_1 is arbitrarily small, this would make the payment to fall below a_1's cost. Since the optimal unfair solution that picks all the agents has value $2(n-1)$, the greedy mechanism is an $\Omega(\sqrt{n})$-approximation, even in the case of coverage functions. Let m be the size of the universe (the number of elements), then this instance also gives us that the greedy mechanism is an $\Omega(\sqrt{m})$-approximation.

Upper bound on the greedy mechanism. Interestingly, the greedy mechanism is an $O(\sqrt{m})$-approximation for coverage functions with universe size m. We defer the analysis to the full version of the paper. In conclusion, we have:

Theorem 3. *The greedy mechanism is a fair $\Omega(\sqrt{n})$-approximation for submodular functions and a fair $\Theta(\sqrt{m})$-approximation for coverage functions where m is the size of the universe.*

4.2 The main result: fairness is a 3-approximation

Intuitively, the greedy mechanism fails because it may select an agent whose Shapley payment quickly decrease. This is due to the fact that the Shapley payments monotonically decrease as the size of the set grows. We can therefore consider an opposite approach to that of the greedy mechanism: start from a feasible solution and greedily remove agents until all agents can be paid fairly. Surprisingly, this simple mechanism which we call SIEVE essentially leads to the main result in this paper.

Sieve. The SIEVE mechanism receives as input a set $S^* \in \text{argmax}_{T:c(T)\leq B} f(T)$ where $c(T)$ is the cost of set T: $c(T) := \sum_{a_i \in T} c_i$. At every step, the mechanism removes the agent whose marginal contribution-per-cost is lowest, and does so until the budget can be attributed fairly amongst the agents in the remaining set. Finally, the mechanism chooses between the remaining set and the element with the largest utility from the original set of agents, S^*. We describe the mechanism formally below.

ALGORITHM 1: SIEVE

Input: $S^* \in \text{argmax}_{T:c(T)\leq B} f(T)$

$S \leftarrow S^*$
while S not feasible **do**
  $S \leftarrow S \setminus \text{argmin}_{a_j \in S} f_{S\setminus a_j}(a_j)/c_j$
end while
if $f(a_i) \geq f(S)$ for some $a_i \in S^*$ **then**
  return $\{a_i\}$
end if
return S

Note that this mechanism is similar to Moulin's group-strategyproof mechanism [22] in the sense that it iteratively removes agents from the solution. However, two main differences are that the condition on which agents are removed are very different (Moulin's mechanism may even remove multiple agents at a single iteration) and that our algorithm does not start with the entire set of agents.

An alternative formulation. The main idea behind the analysis of the mechanism's approximation ratio is to consider its performance against an alternative optimization problem. In the original formulation, the payments are uniquely defined such that the budget constraint holds exactly. A solution is then feasible if all the payments are greater than the agents' costs:

$$\max\Big\{f(S)\ :\ p_i > c_i, \forall a_i \in S,\quad S \subseteq N\Big\} \tag{1}$$

Alternatively, the payments can be defined to be the minimum payments such that the agents are still paid proportionally to their contribution measured by their Shapley value and such that the agents' cost constraints are satisfied. The cost constraint of at least one agent will then be met exactly. With this definition, a solution is feasible when the budget constraint is satisfied. In this formulation, we therefore define the payment to an agent a_i to be $\gamma_S \cdot \phi_i(S)$ where $\gamma_S := \max_{a_i \in S} c_i/\phi_i(S)$ and consider the following optimization problem:

$$\max\Big\{f(S)\ :\ \sum_{a_i \in S} \gamma_S \cdot \phi_i(S) \leq B,\quad S \subseteq N\Big\} \tag{2}$$

The benefit of this new formulation is that there is one condition, instead of n, that needs to hold for a solution to be feasible. The following lemma suggests that for the purposes of our analysis, the problems are equivalent.

Lemma 1. *Let S be as defined as in* Sieve*, the following always holds: S is feasible according to problem 1 if and only if it is feasible according to problem 2.*

Proof. The proof proceeds in two steps. The first is to show that if $\mathcal{S}_1$ and $\mathcal{S}_2$ are the collections of feasible solutions to problems 1 and 2 respectively, then $\mathcal{S}_1 = \mathcal{S}_2$ if $\phi_i(S) \geq 0$ for all $S \in \mathcal{S}_2$ and for all $a_i \in S$. Then, what remains to show is that $\phi_i(S) \geq 0$ holds for any S in Sieve and for all $a_i \in S$.

We start by showing that $\mathcal{S}_1 \subseteq \mathcal{S}_2$. Let $S \in \mathcal{S}_1$ be a feasible solution to problem 1. According to problem 1, the payments to the selected agents are above the costs of these agents, so $p_i = B\phi_i(S)/f(S) \geq c_i$ for all $a_i \in S$. According to problem 2, the sum of the payments is therefore:

$$\sum_{a_i \in S} \phi_i(S) \max_{a_j \in S} \frac{c_j}{\phi_j(S)} \leq \frac{B}{f(S)} \sum_{a_i \in S} \phi_i(S) = B$$

where the last equality follows from the efficiency property of Shapley values. Therefore S is also a feasible solution to problem 2.

We next show that $\mathcal{S}_2 \subseteq \mathcal{S}_1$ when $\phi_i(S) \geq 0$ for all $S \in \mathcal{S}_2$ and for all $a_i \in S$. Let S be a feasible solution to problem 2 and assume that $\phi_i(S) \geq 0$ for all $a_i \in S$. According to problem 2, the sum of the payments to the agents which are served is at most the budget, so $\sum_{a_i \in S} \phi_i(S) \max_{a_j \in S} \frac{c_j}{\phi_j(S)} \leq B$. By contradiction, assume there exists some agent k that has payment, according to problem 1, under its cost, so $\frac{B\phi_k(S)}{f(S)} < c_k$. Then,

$$\sum_{a_i \in S} \phi_i(S) \max_{a_j \in S} \frac{c_j}{\phi_j(S)} \geq \sum_{a_i \in S} \phi_i(S) \frac{c_k}{\phi_k(S)} > \frac{B}{f(S)} \sum_{a_i \in S} \phi_i(S) = B$$

where the second inequality follows from $\frac{B\phi_k(S)}{f(S)} < c_k$ and from $\phi_k(S) \geq 0$. This is a contradiction with S being a feasible solution to problem 2.

Finally, consider some input S^* to Sieve. The marginal contribution of any agent $a_i \in S^*$ satisfies $f_{S^* \setminus a_i}(a_i) \geq 0$ otherwise S^* would not be optimal. By submodularity, we then get that $f_{S \setminus a_i}(a_i) \geq 0$ for any $S \subseteq S^*$. From the definition of Shapley values, we then obtain that $\phi_i(S) \geq 0$ holds for any S in Sieve and for all $a_i \in S$. ☐

Useful property. Using the formulation of problem 2, the analysis of this mechanism is now based on quantifying the relationship between the marginal contribution and the Shapley value of an agent in the solution. We show that for a given set of agents, the marginal contribution of any agent to that set is at most its Shapley value.

Lemma 2. *Let $S \subset N$ and $a_i \notin S$, then $\phi_i(S \cup a_i) \geq f_S(a_i)$ for any submodular $f : 2^N \to \mathbb{R}$.*

Proof. First, observe that for any $S \subseteq N$

$$\sum_{T \subseteq S} \frac{|T|!(|S|-|T|)!}{(|S|+1)!} = \frac{1}{|S|+1} \sum_{k=0}^{|S|} \binom{|S|}{k} \frac{k!(|S|-k)!}{|S|!} = \frac{1}{|S|+1} \sum_{k=0}^{|S|} 1 = 1.$$

Therefore,

$$\phi_i(S \cup a_i) = \sum_{T \subseteq S} \frac{|T|!(|S|-|T|)!}{(|S|+1)!} f_T(a_i) \geq \sum_{T \subseteq S} \frac{|T|!(|S|-|T|)!}{(|S|+1)!} f_S(a_i) = f_S(a_i)$$

where the inequality is due to submodularity. □

The approximation ratio. We now have the necessary tools to prove the approximation ratio. We begin with the following lemma.

Lemma 3. *For any submodular function $f : 2^N \to \mathbb{R}$, let S^* be the input to* Sieve*, S be the feasible set obtained by* Sieve *after removing agents from S^*, and a_i be the last agent removed from S by* Sieve*, then:*

$$f(S^*) - f(S) \leq \frac{f_S(a_i)}{c_i} \sum_{a_k \in S^* \setminus S} c_k.$$

Proof. Let S_k be the set S at the end of the iteration of Sieve where a_k was removed.

$$f(S^*) - f(S) = \sum_{a_k \in S^* \setminus S} \frac{f_{S_k}(a_k)}{c_k} c_k \leq \sum_{a_k \in S^* \setminus S} \frac{f_{S_k \cup a_k \setminus a_i}(a_i)}{c_i} c_k \leq \frac{f_S(a_i)}{c_i} \sum_{a_k \in S^* \setminus S} c_k$$

The first inequality is due to the fact that a_i was removed from S after a_k. The second inequality follows from submodularity. □

The inequality from the above lemma nicely quantifies the difference in the objective function in terms of costs. We can now use the fact that according to our alternative formulation of problem 2, when a_i is the last agent to be sieved, this implies that the cost for $S \cup a_i$ exceeds the budget. We then apply the relation between marginal contributions and Shapley values from Lemma 2 to get the desired result.

Theorem 4. *For any submodular $f : 2^N \to \mathbb{R}$,* Sieve *is a 3-approximation.*

Proof. Let a_i be the last agent removed from S by Sieve and let $a_\ell \in \arg\max_{a_j \in S \cup a_i} c_j / \phi_j(S \cup a_i)$. Since $S \cup a_i$ is not feasible, the cost of $S \cup a_i$ exceeds the budget according to problem 2. This implies:

$$
\begin{aligned}
B &< \sum_{a_k \in S \cup a_i} \phi_k\Big(S \cup a_i\Big) \max_{a_j \in S \cup a_i} \frac{c_j}{\phi_j(S \cup a_i)} \\
&= f(S \cup a_i) \frac{c_\ell}{\phi_\ell(S \cup a_i)} \\
&\leq f(S \cup a_i) \frac{c_\ell}{f_{S \setminus \{a_\ell\} \cup a_i}(a_\ell)} \\
&\leq \Big(f(S) + f_S(a_i)\Big) \frac{c_i}{f_S(a_i)}
\end{aligned}
$$

where the second inequality follows from lemma 2 and the third from the fact that $a_i \in \arg\max_{a_j \in S \cup \{i\}} c_j / f_{S \cup a_i \setminus a_j}(a_j)$ since a_i was removed by the algorithm. Combining with lemma 3 we get:

$$
f(S^*) - f(S) \leq \frac{f_S(a_i)}{c_i} B < f(S) + f_S(a_i).
$$

Recall that S^* is the optimal solution to the unfair problem. If $f_S(a_i) \geq \frac{1}{3} f(S^*)$, then $3f(a_i) \geq f(S^*)$ by submodularity. Otherwise, $f(S^*) - f(S) < f(S) + \frac{1}{3} f(S^*)$, so $3f(S) \geq f(S^*)$. Therefore, if $3f(S) < f(S^*)$, then Sieve returns $\{a_j\}$ for some j such that $3f(a_j) \geq f(S^*)$ □

The lower bound. The natural question now is whether there exists a mechanism that can somehow do better than Sieve, or alternatively if Sieve can be better analyzed. We show that the approximation ratio obtained by analyzing Sieve is optimal: no Shapley-fair mechanism can do better than a factor of $3 - \epsilon$, *independently of any computational assumptions*. We show this holds even for the restricted case of *additive* functions. Recall that a function is additive if $f(S) = \sum_{a \in S} f(a), \ \ \forall S \subseteq N$.

Lemma 4. *No fair mechanism can obtain an approximation ratio better than* $3 - \epsilon$ *for any* $\epsilon \in (0, 1)$*, even for additive functions.*

Proof. Let $\epsilon \in (0, 1)$. Suppose $f(S)$ is additive, $f(a_1) = f(a_2) = 1$, $f(a_3) = 1 - \epsilon$, $B = 1$ while agents' costs are $c_1 = 0$, $c_2 = 1/2 + \epsilon/5$ and $c_3 = 1/2 - \epsilon/5$.

In the unfair case, the optimal solution can be obtained by selecting the three agents. Any Shapley-fair mechanism must pay agents a_1 and a_2 the same since their contribution is identical. But since a_2 needs to be paid at least $\frac{1}{2} + \frac{\epsilon}{5}$ for the solution to be individually rational and budget feasible, agents a_1 and a_2 cannot be both selected. Now consider selecting a_3 and either a_1 and a_2. The payments are then $p_3 = (1/(2-\epsilon)) \cdot (1-\epsilon)$ and $p_i = (1/(2-\epsilon)) \cdot 1$ where $i = 1$ or $i = 2$. Observe that $(1/2 - \epsilon/5)(2-\epsilon) > 1 - \epsilon/2 - 2\epsilon/5 > 1 - \epsilon$, which implies that $c_3 > p_3$ and that this solution is not feasible. It is therefore impossible to serve more than one agent. □

Theorem 5. *For any submodular function* $f : 2^N \to \mathbb{R}^+$ *there exists a fair mechanism that obtains an approximation ratio of* 3*. Furthermore, for any* $\epsilon > 0$ *no mechanism can obtain an approximation better than* $3 - \epsilon$*.*

4.3 Computation

The above result teaches us about the price of fairness . If we consider computational limitations, there are two aspects to consider, as described below.

Seeding Sieve. Note that Sieve takes as input the optimal solution to the unfair problem. In the case of submodular function, this optimal solution cannot be found in polynomial time unless P=NP [12]. However, the input to Sieve can be approximated by the $\frac{e}{e-1}$-approximation algorithm for the optimization of monotone functions under a budget constraint [20]. In the case of non-monotone submodular functions, an e-approximation algorithm is known [14]. Note that for the proof of lemma 1, we need the input S^* to Sieve to be such that $\phi_i(S) \geq 0$ for all $S \subseteq S^*$ and for all $a_i \in S$. It is easy to see that this condition holds for monotone submodular functions. In the full version of the paper, we show how this condition can be satisfied for non-monotone submodular functions by a preprocessing step that modifies S^* into some set S_p^* that is given as an input to Sieve. The advantage of S_p^* is that f is monotone over any of its subsets.

Computing Shapley values. For some utility functions like cover functions (as in the examples used throughout this paper), computing Shapley values is easy. In general, however, Shapley values cannot be computed in polynomial time. At each iteration of Sieve and for each agent that has not yet been removed by the algorithm, its Shapley value has to be computed to decide if the remaining agents are a feasible solution. Thankfully, however, Shapley values of a submodular and monotone function can be approximated arbitrarily well. Note that the monotone condition is not an issue since even for non-monotone f, the input set to Sieve is such that f behaves monotonically over that input thanks to the preprocessing step explained in the full version of the paper. We also describe a Fully Polynomial-Time Randomized Approximation Scheme (FPRAS) in the full version of the paper, which provides an oracle that computes an $(1+\epsilon)$-approximation to the Shapley value by generating $4n(n-1)/\epsilon^2$ random permutations of the agents. This is shown by extending a recent result by Liben-Nowell et al. [21].

Lemma 5. *There exists a polynomial time oracle that computes an $(1+\epsilon)$-approximation of the Shapley value with high probability for monotone submodular functions.*

5 Special cases

In this section, we explore interesting cases for which the approximation ratio can be improved. In particular, we consider large markets, randomization, and uniform costs. For these three distinct settings, we show that the approximation ratio can be improved from three to be arbitrarily close to two, and that in all settings this is tight.

5.1 Large markets

The notion of large markets in budget feasible mechanism design was recently introduced by Anari et al. [1]. We formally defined large markets as follow.

Definition 1. *A market is θ-large if the instance respects $B \geq \left(1 + \frac{1}{\theta}\right) \max_{i \in [n]} c_i$.*

So for large markets where $B >> c_i$, we get that $\theta \to 0$.

Lemma 6. *For any submodular function* SIEVE *is a* $(2+\theta)$*-approximation when the market is* θ*-large.*

Proof. Let a_i be the last agent removed from S by SIEVE. Using the same argument as in the first part of the proof of Theorem 4,

$$B < \frac{c_i f(S)}{f_S(a_i)} + c_i.$$

By the assumption on the budget,

$$\begin{aligned} B &\geq \frac{1+\theta}{\theta} \max_{i\in[n]} c_i \\ \frac{(-1+(1+\theta))B}{1+\theta} &\geq c_i \\ B &\geq \frac{B}{1+\theta} + c_i. \end{aligned}$$

Combining our upper and lower bounds on B, we get

$$\frac{B}{1+\theta} < \frac{c_i f(S)}{f_S(i)} \leq \frac{f(S)\sum_{a_k\in S^*\setminus S} c_k}{f(S^*)-f(S)} < \frac{f(S)B}{f(S^*)-f(S)}$$

where the second inequality follows from lemma 3. Since $\frac{B}{1+\theta} < \frac{f(S)B}{f(S^*)-f(S)}$, we conclude that $f(S^*) < (2+\theta)f(S)$ □

Lemma 7. *No fair mechanism can obtain an approximation ratio better than two when the market is* θ*-large, for any* $\theta > 0$*, even for additive functions.*

Proof. Pick n sufficiently large such that $n \geq 2 + \frac{1}{\theta}$. Suppose $f(S)$ is additive, $f(a_i) = \frac{1}{n-1}$ for $i < n$, $f(a_n) = 1$, $B = 1$ while agents' costs are $c_i = \frac{1}{n-1}$ for $i < n$ and $c_n = 0$. The market is θ-large since $1 \geq (1+\frac{1}{\theta})\frac{1}{n-1}$ by our assumption on n. In the unfair case, the optimal solution can be obtained by selecting all the agents, and the value of this solution is 2. Note however, that for any fair mechanism, agent a_n must be paid $n-1$ times more than any other agents since its contribution is $n-1$ times greater. Picking agent a_n and at least one other agent would infer a cost that is therefore at least $\frac{1}{n-1} + 1 > 1$. So a fair mechanism cannot pick another agent if it picks a_n. □

Theorem 6. *For any submodular function and* θ*-large market,* SIEVE *is a* $(2+\theta)$*-approximation and no fair mechanism can obtain a strictly better approximation ratio.*

5.2 Randomization

The mechanisms and concepts we considered until now were deterministic, and it seems natural to explore the power of randomization in the context of fairness. To do so, we begin by extending our definitions so that we can select agents with probabilities. For each agent, define an allocation variable $x_i \in [0,1]$ indicating the probability with which an agent is selected. A natural extension for a submodular function is the multilinear extension:

$$F(x) = \sum_{S\subseteq N} \prod_{a_j\in S} x_j \prod_{a_j\notin S} (1-x_j) f(S).$$

This extension is the expected performance of a set where each agent is picked independently with probability x_i. Similarly, fairness is extended to be fairness in expectation: the Shapley payment to a_i being defined as

$$p_i = \Phi_i(x) \cdot \frac{B}{F(x)}$$

where:

$$\Phi_i(x) = \sum_{S \subseteq N} \prod_{a_j \in S} x_j \prod_{a_j \notin S} (1 - x_j) \phi_i(S).$$

Individual rationality in this case is defined as having payments $p_i \geq x_i c_i$. The main question is whether randomization allows an improvement on previous results. We first show that randomization cannot do better than in the case of large markets.

Lemma 8. *No fair mechanism can obtain an approximation ratio better than two using randomization, even for additive functions.*

Proof. Suppose $f(S)$ is additive, $f(a_1) = f(a_2) = 1$, $B = 1$ while agents' costs are $c_1 = 1$ and $c_2 = 0$. In the unfair case, the optimal solution can be obtained by selecting both agents. Consider picking the agents fractionally according to x_1 and x_2. Since the function is additive, each agent's contribution in expectation is x_i and $F(x) = x_1 + x_2$. The Shapley payment to p_1 is then $p_1 = x_1 \cdot \frac{1}{x_1 + x_2}$. To satisfy individual rationality, this payment must be at least x_1, implying $x_1 + x_2 \leq 1$. $\square$

Randomization helps. Even though randomization does not allow improvement on the result for large markets, we show that in the general case, the approximation ratio can be improved to two, making the results for this section tight. For this positive result, we compare ourselves to the optimal unfair integral solution. Note that since we are interested in the question of how much can randomization help, the benchmark remains the optimal integral unfair solution (and in particular, not the optimal unfair *fractional* solution). The intuition is similar as to why fractional knapsack performs better than integer knapsack: the relaxation allows us to pick an additional agent with some probability in addition to the integral agents.

R-Sieve. Consider the following modification to Sieve. Once S is feasible, before termination, construct the vector x as follow and return x. Let a_i be the last agent removed from S by Sieve, then $x_j = 1$ for all $a_j \in S$, $x_j = 0$ for all $a_j \in N \setminus \{S \cup a_i\}$, and x_i be such that $B = \sum_{a_k \in S \cup a_i} \Phi_k(x) \max_{a_j \in S \cup \{i\}} \frac{x_j c_j}{\Phi_k(x)}$. It is easy to see that problem 2 also extends according to Φ and F. With x, the budget is tight according to both problem 1 and problem 2. To avoid ambiguity, we refer to this new algorithm as R-Sieve.

Lemma 9. *For any submodular function and input S^*,* R-Sieve *returns a solution x such that $2 \cdot F(x) \geq f(S^*)$.*

Proof. Let x be the output of R-Sieve. Let S be the set of integral agents and a_i be the fractional agent according to x. First, observe that $F(x) = f(S) + x_i f_S(a_i)$ and that $\Phi_j(x) = x_i \phi(S \cup a_i) + (1 - x_i)\phi(S)$ and $\Phi_i(x) = x_i \phi(S \cup a_i)$.

Since we picked x_i such that the budget constraint is tight according to problem 2 and similarly as in the proof of Theorem 4,

$$B = \sum_{a_k \in S \cup \{i\}} \Phi_k(x) \max_{a_j \in S \cup \{i\}} \frac{x_j c_j}{\Phi_k(x)} \leq \Big(f(S) + x_i f_S(a_i) \Big) \frac{c_i}{f_S(a_i)}.$$

Combining with lemma 3, we get that

$$f(S^*) - f(S) \leq \frac{f_S(a_i)}{c_i}\Big(\sum_{a_k \in S^* \setminus S} c_k\Big) \leq \frac{f_S(a_i)}{c_i} B = F(x). \quad \square$$

Observing that $F(x) \geq f(S)$ then concludes our proof. □

Theorem 7. *For any submodular function there is a randomized fair mechanism that achieves an approximation ratio of two, and no randomized fair mechanism can obtain an approximation ratio that is strictly better.*

5.3 Uniform costs

Finally, we consider the case in which all agents have uniform costs, i.e. without loss of generality $\forall i \in [n]$ we have that $c_i = 1$. The upper bound can be improved here too.

Lemma 10. *Let f be as submodular function and assume all agents' costs are uniform, then* Sieve *achieves a 2-approximation.*

Proof. Let c be the cost for any agent. Let t be the number of agents in S^*, so $B \geq tc$. By lemma 3:

$$f(S^*) - f(S) \leq \frac{f_S(a_i)}{c} \sum_{a_k \in S^* \setminus S} c \leq f_S(a_i) \cdot (t-1) \leq f_S(a_i)(B/c - 1)$$

The second inequality follows from the observation that at least one agent can always be selected. Then, as in the proof for Theorem 4, $B < \frac{cf(S)}{f_S(i)} + c$. Combining this bound on the budget with the previous inequality, $f(S^*) - f(S) \leq f(S)$. □

Lemma 11. *No fair mechanism can obtain an approximation ratio better than $2-\epsilon$ in the case of uniform costs for any $\epsilon > 0$, even when the function is additive.*

Proof. Let $\epsilon \in (0,2)$ and pick $\delta > 0$ such that $2-\epsilon < \frac{2}{\delta+1}$. Suppose $f(S)$ is additive, $f(a_1) = 1+\delta$, $f(a_2) = 1-\delta$, $B = 2$, while agents' costs are $c_1 = c_2 = 1$. In the unfair case, the optimal solution is to pick both agents. In the fair case, picking both agents would imply paying a_1 more than a_2, which is not possible while staying feasible. Therefore the optimal fair solution is to pick a_1. Observe that $(2-\epsilon)f(a_1) < \frac{2}{\delta+1} f(a_1) = f(\{a_1, a_2\})$. □

Theorem 8. *For any submodular function, when agents have uniform costs,* Sieve *is a 2-approximation, and no randomized fair mechanism can obtain an approximation ratio that is strictly better.*

6 Beyond Shapley Payments

Throughout the paper we restricted ourselves to fairness defined through Shapley values. A closer inspection into the mechanics of the proofs shows that any value function that satisfies Lemma 2 can be used instead of Shapley values and the same approximation guarantees hold. That is, in order for one to implement Sieve with a value function $\varphi : 2^N \to \mathbb{R}^{|N|}$ and the analysis of the approximation ratio to go through, it is enough to show that for a desired function $f : 2^N \to \mathbb{R}^+$ we wish optimize, we have that for any $a_i \in N$ and $S \subseteq N$:

$$\varphi_i(S) \geq f_{S \setminus a_i}(a_i).$$

This relation is intuitively a natural notion of fairness for submodular functions since the marginal contribution of an agent a_i to S' for any $S' \subseteq S \setminus a_i$ is at least the marginal contribution of a_i to $S \setminus a_i$. Therefore $f_{S \setminus a_i}(a_i)$ should be a lower bound on the contribution of a_i to $f(S)$.

Fairness via nucleolus values. The nucleolus is an interesting value function also heavily used in cooperative game theory. We now show that this value function can also be used to define a notion of fairness, and that the approximation ratio holds for this concept as well. Let $\sigma : 2^N \to \mathbb{R}^{|N|}$ and $\theta(\sigma(S))$ be the vector of excesses of $\sigma(S)$ where

$$\theta_{S_j}(\sigma(S)) = f(S_j) - \sum_{a_i \in S_j} \sigma_i(S)$$

in non-increasing order over all non trivial subsets of S: $\theta_{S_1} \geq \theta_{S_2} \geq \cdots \geq \theta_{S_{2^{|S|}-2}}$.

The excess of a subset S_j is the difference between the utility of S_j according to f and the sum of the values of the agents in S_j according to $\sigma(S)$. An intuitive notion of fairness is that this difference should not be too large, otherwise there is some subset S_j that has a contribution according to $\sigma(S)$ that is not representative of its true utility $f(S_j)$. The definition of the nucleolus $\sigma(S)$ of f is based on this intuition: $\sigma(S)$ is the efficient, meaning that the nucleolus values sum to $f(S)$ over agents, function that lexicographically minimizes the ordering $\theta_{S_1} \geq \theta_{S_2} \geq \cdots \geq \theta_{S_{2^{|S|}-2}}$. This value function was first introduced by Schmeidler [26].

Lemma 12. *Let $S \subseteq N$ and $a_i \in S$, then $\sigma_i(S) \geq f_{S \setminus a_i}(a_i)$ for all submodular $f : 2^N \to \mathbb{R}^+$.*

Proof. Assume by contradiction that there exists S and $a_i \in S$ such that $\sigma_i(S) < f_{S \setminus a_i}(a_i)$. Define $k \geq 1$ to be the number of subsets that have excess equal to θ_{S_1}, so that $\theta_{S_j} = \theta_{S_1}$ for all $j \leq k$. Let S_j be such that $a_i \notin S_j$, then the excess of $\sigma(S)$ for coalition $S_j \cup a_i$ is greater than the excess for S_j:

$$f(S_j \cup a_i) - \sum_{a_l \in S_j \cup a_i} \sigma_l(S) \geq f(S_j) + f_{S \setminus a_i}(a_i) - \sigma_i(S) - \sum_{a_l \in S_j} \sigma_l(S) > f(S_j) - \sum_{a_l \in S_j} \sigma_l(S)$$

where the first inequality is by submodularity and the second is by our initial assumption. It follows that $a_i \in S_j$ for all $j \leq k$.

Pick some agent $a_{i'}$ such that $a_{i'} \notin S_j$ for some $j \leq k$. Note that such an agent must exist since $S_1 \subset S$. Consider the alternate payoff vector x to $\sigma_i(S)$ such that

$$x_l = \begin{cases} \sigma_l(S) + \epsilon & \text{if } l = i \\ \sigma_l(S) - \epsilon & \text{if } l = i' \\ \sigma_l(S) & \text{if } l \neq i, l \neq i' \end{cases}$$

where $\epsilon < \theta_{S_k}(\sigma(S)) - \theta_{S_{k+1}}(\sigma(S))$. Observe that $\theta_{S_j}(x) \leq \theta_{S_j}(\sigma(S))$ for all $j \leq k$ since $a_i \in S_j$ and that $\theta_{S_j}(x) < \theta_{S_j}(\sigma(S))$ for some $j \leq k$ since $i' \notin S_j$ for some S_j. Since $\epsilon < \theta_{S_k}(\sigma(S)) - \theta_{S_{k+1}}(\sigma(S))$, $\theta_{S_j}(x) < \theta_{S_k}(\sigma(S))$ for all $j > k$. Therefore the vector of excesses of x in non increasing order has a lower lexicographical order than $\theta(\sigma(S))$, contradiction with $\sigma(S)$ being the nucleolus. □

Other solution concepts in cooperative game theory have been studied such as the stable set, the core, and the kernel. However these solution concepts do not admit a unique value function over the agents. In order to have well-defined payments, we have therefore focused our attention on the nucleolus and the Shapley value, which admit a unique value function over agents for any function f over subsets of agents.

7 Extensions

We conclude with several natural extensions where one can use fair mechanisms to obtain related objectives.

7.1 Frugality

The problem of frugality is the dual of the budget feasible problem: rather than placing a hard budget constraint and aim to maximize the utility function, frugality places some threshold constraint τ on the value of the utility function and the goal is to minimize payments. The goal can be formulated as:

$$\min\Big\{\sum_{a_i \in S} p_i \ : \ f(S) \geq \tau \ \ S \subseteq N\Big\} \tag{3}$$

Here as well p_i are the fair Shapley payments. We first show that in general, even for additive functions when considering the concept of frugality no bounded approximation ratio is obtainable through a fair mechanism.

Lemma 13. *Let B^* and B' be the optimal budgets in the unfair and fair cases. There is no α such that $\alpha B' \geq B^*$, even for additive functions.*

Proof. Consider the case of two agents a_1 and a_2 where for some positive $\delta < 1$ we have that $f(a_1) = 1 - \delta$, $f(a_2) = \delta$, and $f(\{a_1, a_2\}) = 1$. The costs are uniform, $c_1 = c_2 = 1$ and the threshold is $\tau = 1$. Then $B^* = 2$ and $B' = 1/\delta$. □

Conditions for frugality. Although there are no approximation ratio that can be obtained in the case of frugality in general, approximation ratios can be obtained by adding some assumption. Let $f : 2^N \to \mathbb{R}^+$ be a submodular function, τ be the desired threshold, and B^* be the optimal budget to achieve τ in the unfair case. Assume that in the unfair case, threshold $\alpha\tau$ can be reached with budget B where $\alpha \geq \min(2 + \theta, 3)$ in the case of a θ-large market. In such cases, there exists a $\frac{B}{B^*}$-approximation to the frugality problem with threshold τ. To see this, we can consider running the fair budget feasible mechanism with budget B. We know that the fair solution that is obtained has value that is at least τ. We therefore get a $\frac{B}{B^*}$-approximation to the frugality problem since B^* is the optimal budget to achieve τ.

7.2 Multidimensional Constraints

Another possible extension is to have not one budget constraint on all the agents, but multiple budget constraints over different groups of agents. With multiple groups of agents, each with a budget constraint, one can consider two reasonable notions of fairness. One notion is that of *local fairness* which requires that within each group, all the agents are paid fairly, and the payments to two agents that belong to different groups do not need to be fair. The other notion we can consider is that of *global fairness* which would require fairness over all agents – the mechanism must treat any two agents fairly, regardless of which groups they are in.

Local Fairness. Consider the following scenario: There are m groups of agents, and each group of agent has a budget constraint for the payments to the agents in that group. Assume that the fairness constraint is local in that two agents must be paid proportionally according to their contribution only if they are in a same group.

The objective function f is said to be *composable* if f can be decomposed into $f_1, \ldots, f_m$ such that $f = \sum_i f_i$ and that f_i is the objective function over group $i \in [m]$. If the groups of agents are distinct, it is easy to see that the problem can be decomposed into m fair problems and we obtain the same results as for the problem above. If the group of agents are not distinct or if the objective function is not composable, then the problem gets harder. We conjecture that there exist solutions that achieve the same guarantees as for the regular problem, however finding these solutions is hard.

Global Fairness. When all agents are required to be treated fairly and there are multiple budget constraints, a negative result can be easily obtained.

Lemma 14. *Let S^* and S be the optimal unfair and fair solutions respectively. In the case of multiple budget constraints with global fairness, there is no α such that $\alpha f(S) \geq S^*$, even for additive functions.*

Proof. Consider n agents such that $c_i = i$, $f(a_i) = 1$ with f being additive. There are n budget constraints, $B_i = i$ is over the set consisting of the single agent $\{a_i\}$. Then S^* contain all the agents and S can contain at most one agent because of the fairness constraints and because of the budgets being tight with the cost of each agent. □

8 Discussion and Open Questions

In this paper we introduced the notion of fair attribution mechanisms. We used the notion of Shapley values to define fairness, and showed that other natural concepts can be used as well such as the nucleolus. In general, saw that for submodular functions, there is a simple mechanism that obtains a 3-approximation, and that no mechanism can do better.

The results we showed are tight for submodular functions. One interesting open question that remains is whether positive results are obtainable for more general classes of functions. In the introduction we saw that even for some very simple super-additive functions, fairness is not obtainable. For general subadditive functions we currently do not know how to obtain anything better than the trivial n- approximation which simply selects the single agent with the highest utility and surrenders all the budget to her.

Acknowledgements. The authors would like to thank Ariel Procaccia for helpful discussions.

References

[1] Anari, N., Goel, G., and Nikzad, A. (2014). Mechanism design for crowdsourcing: An optimal 1-1/e competitive budget-feasible mechanism for large markets. In *55th IEEE Annual Symposium*

on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014, pages 266–275.

[2] Anshelevich, E., Dasgupta, A., Kleinberg, J. M., Tardos, É., Wexler, T., and Roughgarden, T. (2004). The price of stability for network design with fair cost allocation. In *45th Symposium on Foundations of Computer Science (FOCS 2004), 17-19 October 2004, Rome, Italy, Proceedings*, pages 295–304.

[3] Archer, A. and Tardos, É. (2007). Frugal path mechanisms. *ACM Transactions on Algorithms*, 3(1).

[4] Badanidiyuru, A., Kleinberg, R., and Singer, Y. (2012). Learning on a budget: posted price mechanisms for online procurement. In *ACM Conference on Electronic Commerce*, pages 128–145.

[5] Bei, X., Chen, N., Gravin, N., and Lu, P. (2012). Budget feasible mechanism design: from prior-free to bayesian. In *STOC*, pages 449–458.

[6] Chen, N., Elkind, E., Gravin, N., and Petrov, F. (2010). Frugal mechanism design via spectral techniques. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 755–764.

[7] Chen, N., Gravin, N., and Lu, P. (2011). On the approximability of budget feasible mechanisms. In *SODA*, pages 685–699.

[8] Chen, Y., Lai, J. K., Parkes, D. C., and Procaccia, A. D. (2013). Truth, justice, and cake cutting. *Games and Economic Behavior*, 77(1):284–297.

[9] Dandekar, P., Fawaz, N., and Ioannidis, S. (2013). Privacy auctions for recommender systems. In *ACM Transactions on Economics and Computation (TEAC)*.

[10] Dobzinski, S., Papadimitriou, C., and Singer, Y. (2011). Mechanisms for complement-free procurement. In *EC*.

[11] Elkind, E., Sahai, A., and Steiglitz, K. (2004). Frugality in path auctions. In *SODA*, pages 701–709.

[12] Feige, U. (1998). A threshold of ln n for approximating set cover. *Journal of the ACM*, 45(4):634–652.

[13] Feigenbaum, J., Papadimitriou, C. H., and Shenker, S. (2001). Sharing the cost of multicast transmissions. *J. Comput. Syst. Sci.*, 63(1):21–41.

[14] Feldman, M., Naor, J., and Schwartz, R. (2011). A unified continuous greedy algorithm for submodular maximization. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 570–579. IEEE.

[15] Immorlica, N., Mahdian, M., and Mirrokni, V. S. (2008). Limitations of cross-monotonic cost-sharing schemes. *ACM Transactions on Algorithms*, 4(2).

[16] Jain, K. and Mahdian, M. (2007). Cost sharing. In Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V., editors, *Algorithmic Game Theory*. Cambridge University Press.

[17] Karlin, A. R., Kempe, D., and Tamir, T. (2005). Beyond VCG: Frugality of truthful mechanisms. In *FOCS*, pages 615–626.

[18] Kash, I. A., Procaccia, A. D., and Shah, N. (2013). No agent left behind: dynamic fair division of multiple resources. In *International conference on Autonomous Agents and Multi-Agent Systems, AAMAS '13, Saint Paul, MN, USA, May 6-10, 2013*, pages 351–358.

[19] Kempe, D., Salek, M., and Moore, C. (2010). Frugal and truthful auctions for vertex covers, flows and cuts. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 745–754.

[20] Khuller, S., Moss, A., and Naor, J. S. (1999). The budgeted maximum coverage problem. *Inf. Process. Lett.*, 70(1):39–45.

[21] Liben-Nowell, D., Sharp, A., Wexler, T., and Woods, K. (2012). Computing shapley value in supermodular coalitional games. In *Computing and Combinatorics - 18th Annual International Conference, COCOON 2012, Sydney, Australia, August 20-22, 2012. Proceedings*, pages 568–579.

[22] Moulin, H. (1999). Incremental cost sharing: Characterization by coalition strategy-proofness. *Social Choice and Welfare*, 16(2):279–320.

[23] Parkes, D. C., Procaccia, A. D., and Shah, N. (2012). Beyond dominant resource fairness: extensions, limitations, and indivisibilities. In *ACM Conference on Electronic Commerce, EC '12, Valencia, Spain, June 4-8, 2012*, pages 808–825.

[24] Procaccia, A. D. (2013). Cake cutting: not just child's play. *Commun. ACM*, 56(7):78–87.

[25] Procaccia, A. D. and Wang, J. (2014). Fair enough: guaranteeing approximate maximin shares. In *ACM Conference on Economics and Computation, EC '14, Stanford , CA, USA, June 8-12, 2014*, pages 675–692.

[26] Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on applied mathematics*, 17(6):1163–1170.

[27] Shapley, L. S. (1953). A value for n-person games. *Contributions to the theory of games*, 2:307–317.

[28] Singer, Y. (2010). Budget feasible mechanisms. In *FOCS*, pages 765–774.

[29] Singer, Y. (2012). How to win friends and influence people, truthfully: influence maximization mechanisms for social networks. In *WSDM*, pages 733–742.

[30] Singer, Y. and Mittal, M. (2013). Pricing mechanisms for crowdsourcing markets. In *WWW*, pages 1157–1166.

[31] Singla, A. and Krause, A. (2013a). Incentives for privacy tradeoff in community sensing. In *HCOMP*.

[32] Singla, A. and Krause, A. (2013b). Truthful incentives in crowdsourcing tasks using regret minimization mechanisms. In *WWW*, pages 1167–1178.

[33] Sun, Y. and Sundararajan, M. (2011). Axiomatic attribution for multilinear functions. In *Proceedings 12th ACM Conference on Electronic Commerce (EC-2011), San Jose, CA, USA, June 5-9, 2011*, pages 177–178.

[34] Talwar, K. (2003). The price of truth: Frugality in truthful mechanisms. In *STACS*, pages 608–619.

[35] von Neumann, J. and Morenstern, O. (1944). Theory of Games and Economics Behavior.

[36] Yang, D., Xue, G., Fang, X., and Tang, J. (2012). Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, Mobicom '12, pages 173–184, New York, NY, USA. ACM.

APPENDIX

Missing discussion and proof from Section 3

Discussion of axioms. To motivate why the Additive axiom is not listed, consider the following situation where agents j, k are fixed: $f(T) = 1$ for all $T \subseteq S$ such that $j \in T$ and $f(T) = 0$ otherwise, $g(T) = 100$ for all $T \subseteq S$ such that $k \in T$ and $g(T) = 0$ otherwise. By our three other axioms, $p_j(f) = p_k(g) = B$ and $p_j(g) = p_k(f) = 0$. First, observe that we cannot have $p_i(f+g) = p_i(f) + p_i(g)$ otherwise we would not satisfy Efficiency. We need an α normalizing term for $p_i(f+g)$ to stay in the budget, however the issue with $p_i(f+g) = \alpha(p_i(f) + p_i(g))$ is that in our situation these payments would be $p_j(f+g) = p_k(f+g) = \alpha B$. Since $(f+g)(T) = 101$ if j and k in T, $(f+g)(T) = 1$ if j in T and not k, and $(f+g)(T) = 100$ if k in T and not j, it is clear that in this situation k contributes more than j to $f+g$ but both j and k get paid the same amount.

Proof of Theorem 2. For convenience we restate the theorem.

Theorem 1. *The Shapley payments are the only payments that satisfy the Efficiency, Dummy, Anonymity, and Proportionality axioms.*

Proof. This proof is similar to the uniqueness proof of uniqueness of Shapley values. Define:

$$w_{S,c}(T) = \begin{cases} c & \text{if } S \subseteq T \\ 0 & \text{otherwise} \end{cases}.$$

By Dummy, $p_i(w_{S,c}) = 0$ if $i \notin S$. By Anonymity, $p_i(w_{S,c}) = p_j(w_{S,c})$ if $a_i, a_j \in S$. By Efficiency, $\sum_i p_i(w_{S,c}) = B$. It follows that $p_i(w_{S,c}) = B/|S|$ for all $a_i \in S$.

We therefore weigh $p_i(f)$ and $p_i(g)$ by $f(S)$ and $g(S)$ so that an agent which receives a high payment for a function which contributes mostly to $(f+g)(S)$ is rewarded with a payment which depends mostly on his payment for this function. We then need to multiply by α to get $p_i(f+g)$ so payments sum to the budget.

From the proof of the uniqueness of Shapley values, we know that f can be decomposed uniquely as $f = \sum_{S \subseteq M} c_S w_S$ where M are all the agents selected. By proportionality, $p_i(f) = \alpha \sum_{S \subseteq M} p_i(c_S, w_S) c_S w_S(M)$. Since $p_i(c_S w_S) = B/|S|$ if $a_i \in S$ and 0 otherwise, and since $c_S w_S(M) = c_S$,

$$p_i(f) = \alpha B \sum_{S \subseteq N, a_i \in S} \frac{c_S}{|S|}$$

where α is the unique constant such that the sum of payments equal the budget.

The previous paragraph shows that there is at most one vector of payments which satisfies these four axioms. To conclude the proof, we show that $p_i = B\phi_i(S)/f(S)$ satisfies these four axioms. Dummy and Anonymity follow directly from the Dummy and Anonymity properties of Shapley values. By the Efficiency property of Shapley values, $\sum_{i \in S} \phi_i(S) = f(S)$, therefore $\sum_{i \in S} p_i = B$. Finally we show that Proportionality holds:

$$p_i(f+g) = \frac{B\phi_{f+g,S}(i)}{(f+g)(S)} = \frac{B\phi_{f,S}(i) + B\phi_{g,S}(i)}{f(S) + g(S)} = \frac{p_i(f)f(S) + p_i(g)g(S)}{f(S) + g(S)}.$$

The first and third equality follow from our definition of p_i and the second follows from the Additivity property of Shapley values. □

Missing discussions and proofs from Section 4

The greedy Shapley mechanism. The greedy algorithm continues to include agents as long as it can pay the fair payments. We showed that no such mechanism can obtain an approximation better than $\Omega(\sqrt{m})$ for coverage functions, where m is the size of the universe. We will now show that this lower bound is tight for a modest variation of this mechanism. In the statement below we use S to denote the solution returned by the greedy Shapley mechanism and S^* denote the optimal unfair solution.

Theorem 8.1. *$O(\sqrt{m}) \cdot \max\{f(\arg\max_{i\in[n]} a_i), f(S)\} \geq f(S^*)$.*

Proof. Let $a_1, \ldots, a_\ell$ be the agents returned by the standard greedy algorithm, i.e. the algorithm which at every step takes the agent whose ratio between its marginal contribution to the existing solution and cost is the largest, until exhausting the budget. Since the greedy Shapley mechanism considers agents by their marginal contribution per cost ordering, the solution retuned this procedure is a subset of the agents in the standard greedy algorithm.

Let k be the index for which the stopping condition of the mechanism was invoked, i.e. k is the smallest index according to the marginal-contribution-per cost sorting for which adding a_k would have violated the individual rationality condition if all agents $a_1, \ldots, a_k$ were paid by the fair Shapley payment. Thus, $\{a_1, \ldots, a_{k-1}\}$ is the set of agents returned by the Shapley greedy mechanism. For convenience, for every $i \in [k]$ let v_i denote the marginal contribution of agent a_i at the stage in which it was added to the solution, and let $S_k = \{a_1, \ldots, a_k\}$.

Using an analogous reasoning as in the proof of Lemma 3, we can show that for any $j \in [k]$ we have that:

$$f(S^*) - f(S_{k-1}) \leq \frac{v_j B}{c_j}. \tag{4}$$

If $f(S_{k-1}) \geq f(S^*)/2$, then we are done. Otherwise, $f(S^*) - f(S_{k-1}) \geq f(S^*)/2$ and from conditions (4)

$$c_j \leq \frac{2 v_j B}{f(S^*)}. \tag{5}$$

Let a_i be the agent in $[k]$ for which individual rationality is violated at stage k. In the worst case, every element in the universe it covers is also covered by k other agents, and thus her Shapley value is at least v_i/k. From the fact that the fair payment violates individual rationality we get:

$$\frac{B}{f(S_k)} \cdot \frac{v_i}{k} \leq c_i \tag{6}$$

By conditions (5) and (6) we get that:

$$k \cdot f(S_k) \geq \frac{f(S^*)}{2}$$

In the case where $k > f(S_k)$, an agent with no marginal contribution was added to the solution by the greedy algorithm since $f(\cdot)$ is a coverage function. This only happens in the case where all of the universe is already covered by the solution. So in that case, we obtain the optimal solution and we are done. Otherwise, $k \leq f(S_k)$ and the above bound implies that:

$$f(S_k)^2 \geq \frac{f(S^*)}{2} \iff f(S_k) \geq \sqrt{f(S^*)/2} \iff \frac{f(S_k)}{f(S^*)} \geq \frac{1}{\sqrt{2f(S^*)}} \geq \frac{1}{\sqrt{2m}}$$

where the last inequality uses the fact that in a universe with m elements $f(S^*)$ is upper bounded by m. We therefore have that $f(S_k)$ is a $O(\sqrt{m})$-approximation. Recall that $S = S_k \setminus \{a_k\}$. Let $a^* \in \arg\max_{i \in [n]} f(a_i)$. Since $f(a_k) \leq f(a^*)$, taking $\arg\max\{f(a^*), f(S)\}$ is a 2-approximation to $f(S_k)$, and this concludes our proof. □

Preprocessing step for non-monotone submodular functions. Note that for the proof of lemma 1, we need the input S^* to SIEVE to be such that $\phi_i(S) \geq 0$ for all $S \subseteq S^*$ and for all $a_i \in S$. It is easy to see that this condition holds for monotone submodular functions. To obtain this property in the non-monotone case, consider the following greedy preprocessing step: remove agents with negative marginal contribution until all marginal contributions are non-negative. Let S_p^* be the set obtained after the preprocessing, then by submodularity, $f_S(a_i) \geq 0$ for all $S \subseteq S_p^*$ and all $a_i \in S_p^*$. By the definition of Shapley values, it then follows that $\phi_i(S) \geq 0$ for all $S \subseteq S_p^*$ and all $a_i \in S_p^*$. This preprocessing step can only improve the performance of the input, so the performance of SIEVE compared to S^* is at least as good as compared to S_p^*.

Proof of lemma 5. For convenience we restate the lemma.

Lemma 15. *There exists a polynomial time oracle that computes an $(1+\epsilon)$-approximation of the Shapley value with high probability for submodular monotone functions.*

Proof. Consider the following oracle based fully polynomial-time randomized approximation scheme (FPRAS) that generates $4n(n-1)/\epsilon^2$ random permutations of the players. For each permutation, the marginal contribution of an agent a_i is the marginal contribution of a_i to the set of agents before a_i in the permutation. The $(1+\epsilon)$-approximation of the Shapley value of agent a_i is then its average marginal contribution over all permutations. Liben-Nowell et al. [21] showed this result for supermodular and monotone games. The only property of supermodular functions used for this result is that the maximum marginal contribution of a player in a random permutation occurs at least in a $1/n$ fraction of permutations, which is at least when the agent is the last player in the permutation. In the case of submodular functions, this maximum marginal contribution occurs when the agent is the first player in a permutation, so it also happens in at least a $1/n$ fraction of the permutations. Therefore, the result also holds for submodular and monotone functions. □