Mechanisms for Hiring a Matroid Base without Money

Emmanouil Pountourakis[1] and Guido Schäfer[2]

[1] Northwestern University, USA
manolis@u.northwestern.edu
[2] CWI and VU University Amsterdam, The Netherlands
g.schaefer@cwi.nl

Abstract. We consider the problem of designing mechanisms for hiring a matroid base without money. In our model, the elements of a given matroid correspond to agents who might misreport their actual costs that are incurred if they are hired. The goal is to hire a matroid base of minimum total cost. There are no monetary transfers involved. We assume that the reports are binding in the sense that an agent's cost is equal to the maximum of his declared and actual costs. Our model encompasses a variety of problems as special cases, such as computing a minimum cost spanning tree or finding minimum cost allocation of jobs to machines.

We derive a polynomial-time randomized mechanism that is truthful in expectation and achieves an approximation ratio of $(m-r)/2+1$, where m and r refer to the number of elements and the rank of the matroid, respectively. We also prove that this is best possible by showing that no mechanism that is truthful in expectation can achieve a better approximation ratio in general. If the declared costs of the agents are bounded by the cost of a socially optimal solution, we are able derive an improved approximation ratio of $3\sqrt{m}$. For example, this condition is satisfied if the costs constitute a metric in the graphical matroid.

Our mechanism iteratively extends a partial solution by adding feasible elements at random. As it turns out, this algorithm achieves the best possible approximation ratio if it is equipped with a distribution that is optimal for the allocation of a single task to multiple machines. This seems surprising given that matroids allow for much richer combinatorial structures than the assignment of a single job.

1 Introduction

The task of designing algorithms that are resilient to manipulations of strategic agents in large, distributed systems (such as the Internet) has become a major challenge in recent years. For example, in online marketplaces (such as eBay or eBid) auction formats are desired that incentivize truth revelation of the bidders' valuations for the items on auction. In online workplaces (like Elance, oDesk or Guru) that match freelance experts to clients mechanisms are sought to prevent unjustified declarations of costs.

The classical approach to incite truth telling in strategic environments is to use *mechanism design* (see, e.g., [12, 11]). Here the basic idea is to issue payments to the agents in order to convince them to behave truthfully. Typically, these payments are used to compensate for the advantage that an agent could obtain by lying. Mechanism design is a powerful approach that gave rise to several enlightening results in the past and still is a very active research area with many intriguing open questions.

However, there are many applications in which monetary transfers (as used in the traditional setting) are infeasible. As a result, researchers have more recently started to look into what is known as *mechanism design without money*. Here the basic question one asks is: Can one incite agents to behave truthfully without the use of monetary transfers? Unfortunately, classical results in voting theory show that the answer to this question is "No!" in general. In particular, the well-known Gibbard-Satterthwaite theorem [6, 15] states that for unrestricted domains and at least three outcomes the only mechanism enforcing truthfulness without monetary transfers is *dictatorial*, i.e., the outcome is determined by a single agent. In particular, this also rules out the possibility of approximating any interesting objective in such a setting.

In light of this strong intractability result, there has recently been a large interest in studying more restrictive settings of mechanism design without money. A partial list of proposals that have been addressed in the literature includes the limitation of the agents' preferences [14], changing the social choice model using imposition [13] or binding reports [8].

Our model In this paper, we study the problem of selecting a minimum cost matroid base in a strategic environment. Here the elements of the matroid correspond to agents who might misreport their actual costs. The intuition behind our model is that a certain task can be accomplished only through the collaboration of certain groups of agents. These groups correspond to the bases of the given matroid. Each agent i declares a cost c_i for performing the task, which is not necessarily equal to his actual cost. Based on the declared costs, the mechanism designer wants to "hire" a matroid base at the cheapest possible cost. There are no monetary transfers between the mechanism designer and the agents.

As an example, suppose that the mechanism designer wants to hire a spanning tree in a given network in order to establish connectivity between all nodes at the lowest possible cost. Here the agents are the edges and each edge declares a cost that it incurs for establishing connectivity between its endpoints. This problem falls into our matroid model simply by using the graphic matroid whose bases correspond to the spanning trees of the given graph.

Another example is the problem of scheduling n jobs on m unrelated machines (possibly with restrictions). Every machine i declares for each job $j \in [n]$ it can execute a processing time p_{ij}. The goal is to determine an assignment of jobs to machines such that the total processing time is minimized. It is not hard to see that this problem is a special case of finding a minimum cost basis in a partitioning matroid and is therefore captured by our matroid model.

Binding reports The latter problem was studied by Koutsoupias [8] under the assumption that the reports are *binding*. Basically, this means that one can detect whether an agent overstates his actual cost. The motivation for this assumption is that in many situations costs are “observable” and thus declaring a cost that is larger than the actual one can be punished. On the other hand, if an agent understates his cost then his actual cost remains unaffected through this false declaration. For example, in the scheduling problem mentioned above binding reports means that the mechanism can enforce that the machine is busy for at least the declared processing time.

Koutsoupias [8] settles the problem of assigning one job to m machines completely. He designs a randomized algorithm that is truthful in expectation and achieves an approximation ratio of $(m + 1)/2$ (which he shows is best possible). He also extends these results to the case of scheduling n jobs on m machines. The crucial insight in [8] that enables him to derive these results is a characterization of the distributions for the assignment of a single job that guarantee truthfulness in expectation. Given this characterization, he then determines a distribution that achieves the best possible approximation ratio.

Our contributions Here we continue this line of research. We consider the problem of designing mechanisms without money for the more general model of hiring a matroid base under binding reports. Our main contributions are as follows:

1. We give a randomized algorithm that is truthful in expectation and achieves an approximation ratio of $(m-r)/2+1$, where m and r refer to the number of elements and the rank of the underlying matroid, respectively.
2. We prove that this approximation ratio is best possible. More specifically, we show that no (randomized) mechanism that is truthful in expectation can achieve a better approximation ratio.
3. We then show that an improved approximation ratio of $3\sqrt{m}$ can be achieved if the declared costs of the agents are bounded by the cost of a socially optimal solution. For example, this condition is satisfied if the costs constitute a metric in the graphic matroid.

Our techniques Our results are based on a natural extension of the greedy algorithm for the computation of a minimum cost basis of a matroid. The algorithm iteratively extends a partial solution by adding elements that maintain feasibility. However, because of truthfulness we cannot enforce that a minimum cost element is chosen in each iteration (as in the standard greedy algorithm). Instead, we have to ensure that in each iteration each feasible addition of an element is chosen with some positive probability such that the resulting probability of picking an element meets certain monotonicity properties.

Although we have some freedom to choose these distributions, their choice impacts the resulting approximation ratio of the mechanism. Intuitively, we would like to tailor these distributions in such a way that the minimum cost element is chosen with some good probability, while the required monotonicity properties are still satisfied. Here the insights obtained by Koutsoupias [8] for assigning a single job to m machines turn out to be very useful.

Our findings show that an appropriate composition of the distribution that is proven to be optimal for the single task assignment in [8] also delivers the best possible results in the more general setting of hiring a matroid base. We find this somewhat surprising because matroids allow for combinatorially much richer structures than the assignment of a single job. In fact, the problem of optimally assigning a single job to m machines is equivalent to computing a minimum cost basis of a 1-uniform matroid (which is one of the most trivial matroids). For this special case our mechanism coincides with the one of Koutsoupias.

In order to bound the approximation ratio of our mechanism we crucially exploit properties of the matroid. However, there are many approximation algorithms that follow a similar design paradigm of iteratively extending a partial solution in a greedy manner (e.g., the greedy algorithm for the set cover problem). We conjecture that our findings might be extended to a broader context of greedy-like approximation algorithms which gives rise to some intriguing questions for follow-up research.

Additional related work The design of mechanism that do not use monetary transfers has recently received considerable attention in the literature on economics and computation. Procaccia and Tennenholtz [14] initiated the study of approximate mechanism design without payments for combinatorial problems by studying facility location problems. Their studies triggered several follow-up articles on this topic (see, e.g., [1, 10, 9, 3, 4]). Dughmi and Gosh [2] derived approximate mechanisms without money for several variants of the assignment problems. Guo and Conitzer [7] studied the problem of selling items without payments for the case of two agents.

The idea of binding reports is also related to *mechanisms with verification*, whose study was first proposed by Nisan and Ronen [12]. However, the notion of verification is much stronger than the notion of binding reports that we consider here. In particular, mechanisms with verification may defer the issuing of payments to the agents until they learned the actual outcome. As a result, these mechanisms can punish misreports a posteriori by imposing very high penalties for lying.

Mechanism with binding reports are related to the notion of *imposition* proposed by Nissim et al. [13]. In the context of the facility location problem, agents might be forced to connect to the facility that is closest to their declared position instead of the one that is closest to their actual position. This approach was further pursued by Fotakis and Tzamos [5].

2 Preliminaries

In this section, we give a formal definition of the model that we consider in this paper and introduce some basic concepts.

2.1 Matroids

We first formally introduce the notion of a *matroid*:

Definition 1. *A* matroid $\mathcal{M} = (E, F)$ *is defined by a finite set* E *of elements and a set* $F \subseteq 2^E$ *of subsets of* E *satisfying*

1. $\emptyset \in F$ *(non-emptiness),*
2. *if* $S \in F$ *and* $S' \subseteq S$ *then* $S' \in F$ *(downward closure),*
3. *if* $S, T \in F$ *and* $|S| > |T|$ *then there exists some* $i \in S \setminus T$ *such that* $T + i \in F$ *(exchange property).*[3]

The sets in F *are called* independent sets. *An inclusion-wise maximal independent set* $B \in F$ *is a* basis *of* $\mathcal{M}$.[4] *The common size of all bases of* $\mathcal{M}$ *is called the* rank *of* $\mathcal{M}$ *and will be denoted by* $r(\mathcal{M})$.[5]

Throughout this paper, we assume that the matroid $\mathcal{M} = (E, F)$ is implicitly represented by an *independent set oracle*: given a set $S \subseteq E$, the oracle specifies whether S is an independent set or not. Unless specified otherwise, we identify the elements in E with the first m natural numbers, i.e., $E = [m]$. We assume that every element $i \in E$ constitutes and independent set, i.e., $i \in F$.[6] Note that this assumption is without loss of generality because we can remove all elements from E that do not occur in any independent set.

Example 1. A typical example of a matroid is the *graphic matroid*. Given a graph $G = (V, E)$, we let the edges E of G be the elements of the matroid and each subset $S \subseteq E$ of edges that does not contain a cycle in G constitutes an independent set in F. It is easy to verify that Properties 1–3 of Definition 1 are satisfied. The bases of $\mathcal{M} = (E, F)$ correspond to the spanning trees of G. The rank of $\mathcal{M}$ is $r(\mathcal{M}) = n - 1$, where n is the number of vertices in G.

Example 2. One of the simplest matroids is the *r-uniform matroid*. Here we are given a set $E = [m]$ and each subset $S \subseteq E$ of at most r elements constitutes and independent set in F. It is easy to see that $\mathcal{M} = (E, F)$ is a matroid of rank $r(\mathcal{M}) = r$.

Example 3. Another example is the so-called *partition matroid*. Here the set E of elements is partitioned into n sets $E_1, \dots, E_n$ and we are given some parameters $k_1, \dots, k_n$. The set of independent sets contains all subsets $S \subseteq E$ such that S contains at most k_i elements from each set E_i of the partition; more formally, $F = \{S \subseteq E \mid |S \cap E_i| \leq k_i \text{ for every } i \in [n]\}$. The rank of $\mathcal{M}$ is $r(\mathcal{M}) = \sum_{i \in [n]} k_i$. Note that the partition matroid generalizes the one-sided matching matroid.

[3] For ease of notation, for a set $T \subseteq E$ and an element $i \in E$ we also use $T + i$ and $T - i$ as a short for $T \cup \{i\}$ and $T \setminus \{i\}$, respectively.

[4] Subsequently, by "maximal" we mean "inclusion-wise maximal", i.e., B is maximal if for every $i \in E \setminus B$, $B + i$ is not an independent set.

[5] Using the properties of Definition 1, it is not hard to show that all bases of a matroid $\mathcal{M}$ have the same size.

[6] We slightly abuse notation here and write $i \in F$ instead of $\{i\} \in F$ for notational convenience.

2.2 Hiring a matroid base

Let $\mathcal{M} = (E, F)$ be a matroid. In our model, we associate an agent with each element $i \in E$ of the matroid. Each agent $i \in E$ has a non-negative cost $\bar{c}_i \in \mathbb{R}_+$. Intuitively, by choosing agent $i \in E$ a cost of $\bar{c}_i$ is incurred. Our goal is to select (or hire) a base of the matroid of minimum total cost. The intuition behind our model is that the bases of the underlying matroid represent groups of agents that together can perform a certain task.

Example 4. In order to establish connectivity among all nodes in a given graph $G = (V, E)$ one may want to determine a minimum cost spanning tree of G. Here each edge $i \in E$ corresponds to an agent and selecting an edge incurs a cost of $\bar{c}_i$. Our goal then is to select a minimum cost basis of the graphic matroid.

Example 5. Another example is to schedule n jobs on m unrelated machines (possibly with restrictions). Every machine i has for each job $j \in [n]$ it can execute a processing time p_{ij}. The goal is to determine an assignment of jobs to machines such that the total processing time is minimized. It is not hard to see that this is equivalent to selecting a minimum cost basis of the one-sided matching matroid. Here the bipartite graph $G = (L \cup R, E)$ is given by $L = [n]$, $R = [m]$ and edge (j, i) is part of E iff machine i is available for job j. The cost $c_{(j,i)}$ of edge (j, i) is equal to p_{ij}. Note that for the special case of scheduling a single job, the corresponding one-sided matching matroid reduces to a 1-uniform matroid.

Subsequently, we use $\bar{\mathbf{c}} = (\bar{c}_1, \ldots, \bar{c}_m) \in \mathbb{R}^m_+$ to refer to the vector of *actual* costs. These costs are considered to be "private" in the sense that $\bar{c}_i$ is only known to agent i.

2.3 Binding reports

We assume that agents might misreport their costs, i.e., each agent $i \in E$ declares a cost c_i, which is possibly different from his actual cost $\bar{c}_i$. Based on the matroid $\mathcal{M}$ and the declared costs $\mathbf{c} = (c_1, \ldots, c_m)$, the mechanism selects a basis of the underlying matroid. We consider mechanisms *without money*, i.e., the mechanism does not receive/issue any payments from/to the agents.

In order to achieve truthfulness it will turn out to be crucial to allow for random selections of agents, i.e., we consider *randomized* mechanisms. Our mechanism outputs a basis of the underlying matroid with probability 1, i.e., randomization will not affect the feasibility of the solution. Subsequently, we use $p_i(\mathbf{c})$ to refer to the probability that our (random) mechanism picks element $i \in E$, given the reported costs $\mathbf{c}$.

We assume that the reports are *binding* as proposed by Koutsoupias [8]. More precisely, if agent i's reported cost is c_i then his actual cost is $\max\{\bar{c}_i, c_i\}$. That is, if agent i overstates his actual cost by reporting $c_i > \bar{c}_i$ and agent i is selected then his actual cost becomes c_i. Basically, this means that we can detect if agent i overstate his actual cost. The intuition behind this is that often these costs are "observable" (e.g., think of processing times to execute a job) and thus declaring a cost that is larger than the actual one can be punished. On the other hand, if agent i understates his actual cost $\bar{c}_i$ and is selected then his actual cost remains $\bar{c}_i$ because this is the cost incurred by i. Formally, we assume that each agent $i \in E$ strives to minimize his expected cost

$$C_i(\mathbf{c}) = \max\{\bar{c}_i, c_i\} p_i(\mathbf{c}).$$

Subsequently, we use $\mathbf{c} = (c_1, \ldots, c_m) \in \mathbb{R}_+^m$ to denote the vector of *declared* costs. Recall that the mechanism receives as input the matroid $\mathcal{M} = (E, F)$ and the vector $\mathbf{c}$ of declared costs.

2.4 Truthful mechanisms

We are interested in designing mechanisms that are *truthful in expectation*, which we define next. To this aim, we first need to introduce some standard notation. Let $\mathbf{c} = (c_1, \ldots, c_m) \in \mathbb{R}_+^m$ be a cost vector. Then we denote by $\mathbf{c}_{-i}$, $i \in [m]$, the $(m-1)$-dimensional vector with the ith coordinate removed, i.e.,

$$\mathbf{c}_{-i} = (c_1, c_2, \ldots, c_{i-1}, c_{i+1}, \ldots, c_m).$$

For a subset $T \subseteq [m]$, we will also use $\mathbf{c}_T$ to refer to the restriction of $\mathbf{c}$ to index set T, i.e., $\mathbf{c}_T = (c_{i_1}, c_{i_2}, \ldots, c_{i_{|T|}})$ with $T = \{i_1, \ldots, i_{|T|}\}$.

Definition 2 (Truthful mechanism). *A mechanism M is* truthful in expectation *if for every agent i and every vector $\boldsymbol{c}_{-i}$, the expected cost of i is minimized by declaring the actual cost truthfully, i.e., for every $i \in E$,*

$$C_i(\bar{c}_i, \boldsymbol{c}_{-i}) = \bar{c}_i p_i(\bar{c}_i, \boldsymbol{c}_{-i}) \leq \max\{\bar{c}_i, c_i\} p_i(\boldsymbol{c}) = C_i(\boldsymbol{c}).$$

There are stronger notions of truthfulness (e.g., truthfulness or universal truthfulness). However, it is easy to see that with these stronger notions of truthfulness no positive results are possible; see also [8].

2.5 Approximate social cost

The *social cost* function that we consider throughout this paper is the sum of the individual costs, i.e., $\mathsf{SC}(\mathbf{c}) = \sum_{i \in E} C_i(\mathbf{c})$. We use $\mathsf{OPT}(\mathbf{c})$ to refer to the cost of a socially optimal solution, i.e., the minimum cost of a base of $\mathcal{M} = (E, F)$ with respect to $\mathbf{c}$. Ideally, we would like to derive a truthful mechanism that computes a socially optimal outcome. However, this is impossible and we therefore relax the optimality condition and resort to approximate solutions.

Definition 3. *A mechanism M is* α-approximate *with $\alpha \geq 1$ if for every vector $\boldsymbol{c}$ of declared costs, the expected social cost satisfies*

$$\mathit{SC}(\boldsymbol{c}) = \sum_{i \in E} C_i(\boldsymbol{c}) = \sum_{i \in E} \max\{\bar{c}_i, c_i\} p_i(\boldsymbol{c}) \leq \alpha \mathit{OPT}(\boldsymbol{c}).$$

2.6 Koutsoupias' characterization

Koutsoupias [8] considers the problem of scheduling one job on m available machines. The actual cost incurred by machine i to schedule job j is $\bar{p}_i$ and each machine wants to minimize his cost. The overall objective is to determine an assignment of minimum total cost.[7] Recall that this is equivalent to computing a minimum cost basis in a 1-uniform matroid (see Example 2).

Koutsoupias characterizes the set of truthful mechanisms for this problem.

Proposition 1 ([8]). *Let $p_i(\boldsymbol{c})$ be the probability that element i is chosen by mechanism M given the vector of declared costs $\mathbf{c}$. Then M is truthful in expectation if and only if for every $i \in E$:*

1. *$p_i(c_i, \boldsymbol{c}_{-i})$ is non-increasing in c_i,*
2. *$c_i p_i(c_i, \boldsymbol{c}_{-i})$ is non-decreasing in c_i.*

Based on the above characterization result, Koutsoupias then derives a distribution that satisfies the above properties and whose expected social cost is at most $(m+1)/2$ times the optimal one. He also proves that this is best possible in the sense that no other truthful in expectation mechanism (without payments) can achieve a better approximation ratio.

Koutsoupias [8] also lifts all his results for a single job to n jobs that need to be processed on m machines. As described in Example 5, this scheduling problem can be modeled as a one-sided matching matroid.[8]

3 Greedy Mechanism and Truthfulness Conditions

In this section, we provide a general framework for constructing truthful mechanisms. Our framework is based on the greedy approach which iteratively extends a partial solution (i.e., independent set) by adding a least cost element. We parameterize our mechanism with a collection of distributions: for every $T \subseteq E$ we are given a distribution $d^T = \{d_i^T(\mathbf{c}_T) \mid i \in T\}$ over the elements i in T.[9]

Definition 4 (Greedy Mechanism). *Given a matroid $\mathcal{M} = (E, F)$ with a cost vector $\boldsymbol{c}$ and a collection of distributions $(d^T)_{T \subseteq E}$, the greedy algorithm is as follows:*

1. *Let $S \leftarrow \emptyset$*
2. *While S is not a base*
   (a) *Let $T = \{i \in E \setminus S \mid S + i \in F\}$*

[7] We note that in [8] also the objective of minimizing the makespan is considered.

[8] Koutsoupias does actually not mention the extension to the case of machine restrictions (which is captured by the partitioning matroid in Example 3), but it is trivial to see that all his results go through also in this more general case.

[9] The assumption that all these distributions are given is a conceptual one. Subsequently, it will become clear that we can generate the relevant distributions considered by the algorithm efficiently.

(b) Draw $i \in T$ with probablity $d_i^T(\boldsymbol{c}_T)$
(c) Set $S \leftarrow S + i$
3. Output S

Note that the set T in Step 2(a) contains all elements that can be added to the independent set S without rendering it infeasible. The main difference of our algorithm to the standard greedy algorithm for matroids is that we do not require that the element $i \in T$ added to S in Step 2(c) is of minimum cost. Indeed, such a mechanism would not be truthful because it failed to satisfy Condition (b) of Proposition 1. Instead, here we choose an element i from T with probability $d_i^T(\mathbf{c}_T)$. In particular, our algorithm coincides with the standard greedy algorithm if $d_i^T(\mathbf{c}_T) > 0$ only for the minimum cost elements in T.

We next establish some sufficient conditions for the distributions used by our greedy algorithm that ensure truthfulness.

Theorem 1. *The greedy mechanism is truthful in expectation if for every $T \subseteq E$ and every $i \in T$ it holds:*

1. *$d_j^T(c_i, \boldsymbol{c}_{T-i})$ is non-decreasing in c_i for every $j \in T - i$,*
2. *$d_i^T(c_i, \boldsymbol{c}_{T-i})c_i$ is non-decreasing in c_i.*

The proof of this theorem can be found in Appendix A.

4 Optimal Distributions and Approximation Ratio

We next show that distributions exist that satisfy Properties (1) and (2) of Theorem 1. Basically, the idea here is to lift Koutsoupias' distributions for the special case of 1-uniform matroids (see [8]) to the general matroid case. Then we show that the approximation ratio of the resulting greedy mechanism is $(m - r)/2 + 1$, where $r = r(\mathcal{M})$ is the rank of the underlying matroid $\mathcal{M}$. The proofs of this section are provided in Appendix B.

4.1 Optimal Distributions

A natural choice for a collection $(d^T)_{T\subseteq E}$ of distributions to be used by the greedy algorithm is to choose each element i from a given set T with probability that is inversely proportional to its cost c_i. More precisely, for every $T \subseteq E$ and every $i \in T$, we define

$$d_i^T(\mathbf{c}_T) = \frac{c_i^{-1}}{\sum_{k\in T} c_k^{-1}}. \tag{1}$$

The distribution d^T is also called the *proportional distribution*. It is not hard to show that these distributions satisfy Properties (1) and (2) of Theorem 1. However, the problem is that the greedy mechanism equipped with these distributions results in an approximation ratio which is arbitrarily close to m.

The following distribution was introduced by Koutsoupias [8] for scheduling a single job on m machines.

Definition 5 (Optimal Distribution). *Let $T \subseteq E$ be a subset of elements and assume without loss of generality that $T = \{1, \ldots, |T|\}$ such that $c_1 \leq c_2 \leq \cdots \leq c_{|T|}$. Define probabilities*

$$d_1^T(\boldsymbol{c}_T) = \frac{1}{c_1} \int_0^{c_1} \prod_{k \neq 1} \Big(1 - \frac{x}{c_k}\Big) dx$$

$$d_j^T(\boldsymbol{c}_T) = \frac{1}{c_1 c_j} \int_0^{c_1} \int_0^{y} \prod_{k \neq 1, j} \Big(1 - \frac{x}{c_k}\Big) dx dy \qquad \text{for } j \neq 2.$$

We show that their generalization to our setting leads to a greedy mechanism which is truthful in expectation and achieves the best possible approximation ratio.

Theorem 2. *The greedy mechanism equipped with the distributions of Definition 5 is truthful.*

4.2 Approximation Ratio

Koutsoupias [8] used the distribution d^T given in Definition 5 to handle the case of allocating a single job to m machines. Recall that this problem can be modeled as a 1-uniform matroid as described in Example 2. He showed that the resulting mechanism achieves an approximation ratio of $(m+1)/2$.

Here we prove that our greedy mechanism, equipped with the distributions in Definition 5, has an approximation ratio of $(m-r)/2+1$, where r is the rank of the matroid. Note that this matches the bound of Koutsoupias in the case of a 1-uniform matroid.

Theorem 3. *The greedy mechanism with distributions d^T as defined in Definition 5 has approximation ratio $(m-r)/2+1$.*

5 Lower bound

In this section, we provide a general lower bound on the approximation ratio of mechanisms for hiring a matroid base that are truthful in expectation. Our lower bound matches the upper bound of our greedy algorithm established in the previous section. We defer all proofs of this section to Appndeix C.

We show that for any given parameters m and r, we can always construct a matroid with m elements and rank r such that no mechanism that is truthful in expectation can achieve an approximation ratio better than $(m-r)/2+1$. Note this result does not necessarily imply that every truthful mechanism will perform poorly given any matroid set system with these parameters. For example, in the case of a partition matroid with $k_i = 1$ for all i, the optimal mechanism has an approximation ratio of $(\max_i |E_i| + 1)/2$.

Using the previous lemma we show that for every choice of m and r our upper bound is tight in the sense that there exists a matroid instance where any truthful mechanism has approximation ratio $(m-r)/2+1$.

Theorem 4. *Given m and r, there exists a matroid $\mathcal{M} = (E, F)$ with $|E| = m$ and $r(\mathcal{M}) = r$ for which no mechanism that is truthful in expectation can achieve an approximation ratio better than $(m - r)/2 + 1$.*

Finally, we address the question whether we can do better if we only consider graphical matroids. Since our previous worst case examples do not correspond to a graphical matroid we extend the result directly. We manage to show that there is a family of graphs where the worst case bound of $(m - r)/2 + 1$ occurs.

Theorem 5. *There is no mechanism that is truthful in expectation that achieves an approximation ratio better than $(m - r)/2 + 1$ for graphical matroids.*

6 Improved Approximation Ratio for Metrics

In this section, we show that we can derive an improved approximation ratio of $O(\sqrt{m})$ for our greedy algorithm if each agent's declared cost is at most the cost of a socially optimal solution, i.e., for every agent $i \in E$, $c_i \leq \mathsf{OPT}(\mathbf{c})$. Said differently, this condition requires that the cost of an arbitrary base of the matroid is at least as large as $\max_{i \in E} c_i$. We call a vector $\mathbf{c} = (c_1, \ldots, c_m)$ of declared costs *opt-bounded* if it satisfies this condition.

Note that in the case of a graphical matroid this property requires that no edge declares a cost that is larger than the cost of any spanning tree of the given graph (with respect to $\mathbf{c}$). In particular, if the declared cost vectors $\mathbf{c}$ are restricted to constitute a metric then this condition is trivially satisfied. It is interesting to note that we obtain this result for the greedy algorithm using the proportional distributions. Tthe proof of this theorem is given in Appendix D.

Theorem 6. *If the declared cost vector is opt-bounded then the greedy mechanism using the proportional distributions as defined in* (1) *is truthful in expectation and achieves an approximation ratio of $3\sqrt{m}$.*

7 Future work

There are a lot of open problems that arise from our work. We designed an algorithm that achieves an approximation ratio based on the size of the matroid and its rank. In Section 5 we proved a lower bound that was dependent on the substitutability of elements within the matroid's bases. It could be possible to provide a more refined upper bound using this parameter.

Also there are many questions still open in the case of graphical matroids when the costs constitute a metric. We analyzed only the propotional method which generaly performs worse than the distribution in Definition 5. We also have no matching lower bounds. Additionally, our iterative algorithm and generally our framework didn't depend on the matroid property of the set system to satisfy truthfulness. Thus, it will be interesting to analyze its performance in more general settings especially where the classic greedy has good approximation guarantees. Finally, we only considered social costs and not other social objecties like a minmax solution concept.

References

1. Noga Alon, Michal Feldman, Ariel D. Procaccia, and Moshe Tennenholtz. Strategyproof approximation of the minimax on networks. *Math. Oper. Res.*, 35(3):513–526, 2010.
2. Shaddin Dughmi and Arpita Ghosh. Truthful assignment without money. In David C. Parkes, Chrysanthos Dellarocas, and Moshe Tennenholtz, editors, *ACM Conference on Electronic Commerce*, pages 325–334. ACM, 2010.
3. Dimitris Fotakis and Christos Tzamos. On the power of deterministic mechanisms for facility location games. In Fedor V. Fomin, Rusins Freivalds, Marta Z. Kwiatkowska, and David Peleg, editors, *ICALP (1)*, volume 7965 of *Lecture Notes in Computer Science*, pages 449–460. Springer, 2013.
4. Dimitris Fotakis and Christos Tzamos. Strategyproof facility location for concave cost functions. In Michael Kearns, R. Preston McAfee, and Éva Tardos, editors, *ACM Conference on Electronic Commerce*, pages 435–452. ACM, 2013.
5. Dimitris Fotakis and Christos Tzamos. Winner-imposing strategyproof mechanisms for multiple facility location games. *Theor. Comput. Sci.*, 472:90–103, 2013.
6. Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, July 1973.
7. Mingyu Guo and Vincent Conitzer. Strategy-proof allocation of multiple items between two agents without payments or priors. In Wiebe van der Hoek, Gal A. Kaminka, Yves Lespérance, Michael Luck, and Sandip Sen, editors, *AAMAS*, pages 881–888. IFAAMAS, 2010.
8. Elias Koutsoupias. Scheduling without payments. In Giuseppe Persiano, editor, *SAGT*, volume 6982 of *Lecture Notes in Computer Science*, pages 143–153. Springer, 2011.
9. Pinyan Lu, Xiaorui Sun, Yajun Wang, and Zeyuan Allen Zhu. Asymptotically optimal strategy-proof mechanisms for two-facility games. In *Proceedings of the 11th ACM conference on Electronic commerce*, EC '10, pages 315–324, New York, NY, USA, 2010. ACM.
10. Pinyan Lu, Yajun Wang, and Yuan Zhou. Tighter bounds for facility games. In *Proceedings of the 5th International Workshop on Internet and Network Economics*, WINE '09, pages 137–148, 2009.
11. N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani. *Algorithmic Game Theory.* Cambridge University Press, 2007.
12. Noam Nisan and Amir Ronen. Algorithmic mechanism design. *Games and Economic Behavior*, 35(1-2):166–196, 2001.
13. Kobbi Nissim, Rann Smorodinsky, and Moshe Tennenholtz. Approximately optimal mechanism design via differential privacy. In Shafi Goldwasser, editor, *ITCS*, pages 203–213. ACM, 2012.
14. Ariel D. Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. In John Chuang, Lance Fortnow, and Pearl Pu, editors, *ACM Conference on Electronic Commerce*, pages 177–186. ACM, 2009.
15. Mark Allen Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, April 1975.

A Missing Proofs of Section 3

The proof of this proposition proceeds exactly along the same lines as the one in [8]. We repeat it here for the sake of completeness.

A.1 Proof of Proposition 1

Suppose properties (1) and (2) hold. Fix a player i. If $c_i \leq \bar{c}_i$ then agent i prefers cost $\bar{c}_i$ to c_i because of (1), i.e., $C_i(c_i, \mathbf{c}_{-i}) = \bar{c}_i p_i(c_i, \mathbf{c}_{-i})$ is minimized at $c_i = \bar{c}_i$. Similarly, if $c_i \geq \bar{c}_i$ then agent i prefers cost $\bar{c}_i$ to c_i because of (2), i.e., $C_i(c_i, \mathbf{c}_{-i}) = c_i p_i(c_i, \mathbf{c}_{-i})$ is minimized at $c_i = \bar{c}_i$.

Next suppose that (1) does not hold for some agent i, i.e., there exist x, y with $x < y$ and $p_i(x, \mathbf{c}_{-i}) < p_i(y, \mathbf{c}_{-i})$. Then agent i has an incentive to lie: when $\bar{c}_i = y$ he prefers to report x. Similarly, if (2) does not hold for some agent i then there exist x, y with $x < y$ such that $xp_i(x, \mathbf{c}_{-i}) > yp_i(y, \mathbf{c}_{-i})$. Then agent i has an incentive to misreport: when $\bar{c}_i = x$ he prefers to report y.

B Missing proofs of Section 4

Some of our proofs rely on the concept of *contraction*, which is defined as follows:

Definition 6. *Let $\mathcal{M} = (E, F)$ be a matroid. Then by $\mathcal{M}^{(i)}$ we refer to the matroid that we obtain from $\mathcal{M}$ by* contracting *element i. More formally, $\mathcal{M}^{(i)} = (E^{(i)}, F^{(i)})$, where $E^{(i)} = E - i$ and*

$$F^{(i)} = \{S - i \mid S \in F,\ i \in S\}.$$

For example, if $\mathcal{M} = (E, F)$ is a graphic matroid then $\mathcal{M}^{(i)}$ refers to the matroid that we obtain from $\mathcal{M}$ by contracting edge $i \in E$.

B.1 Proof Theorem 1

Fix a collection of distributions $(d^T)_{T \subseteq E}$ that satisfies Properties (1) and (2). Let $p_i(\mathbf{c})$ be the probability of picking element $i \in E$ after the execution of the mechanism. We need to show that Properties (1) and (2) of Proposition 1 are satisfied, i.e.,

1. $p_i(c_i, \mathbf{c}_{-i})$ is non-increasing in c_i,
2. $p_i(c_i, \mathbf{c}_{-i})c_i$ is non-decreasing in c_i.

We prove these by induction on the number m of elements in E.

If $m = 1$ then there is only one element to be picked and the properties clearly hold.

Suppose that the claim holds true for all element sets of size less than m. We show that it continues to hold for sets of size m. Recall that we use $\mathcal{M}^{(j)}$ to refer to the matroid that we obtain from $\mathcal{M}$ by contracting element j (see Definition 6).

Let $p_i^{(j)}(\mathbf{c}_{-j})$ be the probability of picking element i in the matroid $\mathcal{M}^{(j)}$. Note that $p_i^{(j)}(\mathbf{c}_{-j})$ is precisely the probability of picking element i conditional on the event that player j has been picked in the first round.

First property: Using Bayes rule, we obtain

$$p_i(\mathbf{c}) = d_i^E(\mathbf{c}) + \sum_{j \in E-i} d_j^E(\mathbf{c}) \cdot p_i^{(j)}(\mathbf{c}_{-j}) = 1 - \sum_{j \in E-i} d_j^E(\mathbf{c}) + \sum_{j \in E-i} d_j^E(\mathbf{c}) \cdot p_i^{(j)}(\mathbf{c}_{-j})$$
$$= 1 + \sum_{j \in E-i} d_j^E(\mathbf{c})[p_i^{(j)}(\mathbf{c}_{-j}) - 1].$$

By assumption, $d_j^E(\mathbf{c})$ is non-decreasing in c_i for every $j \neq i$. Also, by our induction hypothesis, $p_i^{(j)}(\mathbf{c}_{-j})$ is non-increasing in c_i. Thus, the product $d_j^E(\mathbf{c})[p_i^{(j)}(\mathbf{c})-1]$ is non-increasing in c_i. We conclude that $p_i(\mathbf{c})$ is non-increasing in c_i.

Second property: Using Bayes rule, we obtain

$$p_i(\mathbf{c})c_i = d_i^E(\mathbf{c})c_i + \sum_{j \in E-i} d_j^E(\mathbf{c})[p_i^{(j)}(\mathbf{c}_{-j})c_i]$$

By assumption and our induction hypothesis, $d_j^E(\mathbf{c})$ and $p_i^{(j)}(\mathbf{c}_{-j})c_i$ are non-decreasing in c_i for every $j \neq i$. Thus, their product is non-decreasing in c_i. By assumption, also $d_i^E(\mathbf{c})c_i$ is non-decreasing in c_i. We conclude that $p_i(\mathbf{c})c_i$ is non-decreasing in c_i.

B.2 Proof Theorem 2

By employing Theorem 1 we only need to show that for every $T \subseteq E$ and $i \in T$

1. $d_j^T(c_i, \mathbf{c}_{T-i})$ is non-decreasing in c_i for every $j \in T - i$,
2. $d_i^T(c_i, \mathbf{c}_{T-i})c_i$ is non-decreasing in c_i.

Koutsoupias [8, Lemma 1] showed that this distributions satisfies the following properties:

Proposition 2 ([8]). *The distribution d^T of Definition 5 satisfies that for every $i \in T$*

1. *$d_i^T(c_i, \mathbf{c}_{T-i})$ is non-increasing in c_i,*
2. *$d_i^T(c_i, \mathbf{c}_{T-i})c_i$ is non-decreasing in c_i.*

Proposition 2 implies that the second property holds, but it does not imply the first one.

Following the proof of Koutsoupias [8], because the distribution d^T is symmetric we only need to show that the function $d_j^T(c_i, \mathbf{c}_{T-i})$ is non-decreasing in c_i for every $j \in T - i$ when the relative order of the elements with respect to their costs remains the same.

As in Definition 5, assume that $T = \{1, 2, \ldots, |T|\}$ such that $c_1 \leq c_2 \leq \cdots \leq c_{|T|}$. We define for every $i, j \in T$

$$G_{i,j}(x, \mathbf{c}_T) = \prod_{k \neq 1,i,j} \Big(1 - \frac{x}{c_k}\Big).$$

Note that for $x \leq c_1$ we have $G_{i,j}(x, \mathbf{c}_T) \geq 0$.

Case 1: $j = 1$ and $i \neq 1$. We have that

$$d_j^T(c_i, \mathbf{c}_{T-i}) = \frac{1}{c_1} \int_0^{c_1} \Big(1 - \frac{x}{c_i}\Big) G_{i,j}(x, \mathbf{c}_T) dx.$$

Using the non-negativity of $G_{i,j}(x, \mathbf{c}_T)$ for $x \leq c_1$, we obtain that for $c_i' \geq c_i$

$$\Big(1 - \frac{x}{c_i}\Big) G_{i,j}(x, \mathbf{c}_T) \leq \Big(1 - \frac{x}{c_i'}\Big) G_{i,j}(x, \mathbf{c}_T),$$

which implies that

$$\frac{1}{c_1} \int_0^{c_1} \Big(1 - \frac{x}{c_i}\Big) G_{i,j}(x, \mathbf{c}_T) dx \leq \frac{1}{c_1} \int_0^{c_1} \Big(1 - \frac{x}{c_i'}\Big) G_{i,j}(x, \mathbf{c}_T) dx$$

and thus $d_j^T(c_i, \mathbf{c}_{T-i}) \leq d_j^T(c_i', \mathbf{c}_{T-i})$ as desired.

Case 2: $j \neq 1$ and $i \neq 1$. We have that

$$d_j^T(c_i, \mathbf{c}_{T-i}) = \frac{1}{c_1 c_j} \int_0^{c_1} \int_0^{y} \Big(1 - \frac{x}{c_i}\Big) G_{i,j}(x, \mathbf{c}_T) dx dy$$

As in the first case, exploiting the non-negativity of $G_{i,j}(x)$ for $x \leq c_1$ we obtain for $c_i' \geq c_i$:

$$\Big(1 - \frac{x}{c_i}\Big) G_{i,j}(x, \mathbf{c}_T) \leq \Big(1 - \frac{x}{c_i'}\Big) G_{i,j}(x, \mathbf{c}_T),$$

which implies that

$$\frac{1}{c_1} \int_0^{c_1} \Big(1 - \frac{x}{c_i}\Big) G_{i,j}(x, \mathbf{c}_T) dx \leq \frac{1}{c_1} \int_0^{c_1} \Big(1 - \frac{x}{c_i'}\Big) G_{i,j}(x, \mathbf{c}_T) dx$$

and thus $d_j^T(c_i, \mathbf{c}_{T-i}) \leq d_i^T(c_i', \mathbf{c}_{T-i})$.

Case 3: $j \neq 1$ and $i = 1$. Differentiating $d_j^T(\mathbf{c}_T)$ with respect to c_1 we get that

$$\frac{\partial d_j^T(\mathbf{c}_T)}{\partial c_1} = \frac{1}{c_1^2 c_j} \Big[c_1 \int_0^{c_1} G_{i,j}(x, \mathbf{c}_T) dx - \int_0^{c_1} \int_0^{y} G_{i,j}(x, \mathbf{c}_T) dx dy \Big]$$

In order that show that this partial derivative is positive we need to show that

$$c_1 \int_0^{c_1} G_{i,j}(x, \mathbf{c}_T) dx \geq \int_0^{c_1} \int_0^{y} G_{i,j}(x, \mathbf{c}_T) dx dy$$

Because $G_{i,j}(x, \mathbf{c}_T) \geq 0$ for every $x \in [0, c_1]$, the function $\int_0^y G_{i,j}(x, \mathbf{c}_T)dx$ is increasing in y for the interval $y \in [0, c_1]$. Thus,

$$\int_0^{c_1} G_{i,j}(x, \mathbf{c}_T)dx \geq \int_0^y G_{i,j}(x, \mathbf{c}_T)dx,$$

which implies that

$$\int_0^{c_1} \int_0^{c_1} G_{i,j}(x, \mathbf{c}_T)dxdy \geq \int_0^{c_1} \int_0^y G_{i,j}(x, \mathbf{c}_T)dxdy.$$

The latter is equivalent to

$$c_1 \int_0^{c_1} G_{i,j}(x, \mathbf{c}_T)dx \geq \int_0^{c_1} \int_0^y G_{i,j}(x, \mathbf{c}_T)dxdy,$$

which proves the claim.

B.3 Proof Theorem 3

We will make use of the following fact proved in [8]

Proposition 3 ([8]). *For every set $T = \{1, \ldots, |T|\}$ such that $c_1 \leq c_2 \leq \cdots \leq c_{|T|}$, it holds that for every $j \geq 2$ that $c_j d_j^T(\boldsymbol{c}_T) \leq \frac{1}{2}c_1$.*

Using this proposition we will show the following theorem:

We will prove this using induction on the number m of elements of the matroid. For $m = 1$ the algorithm is optimal.

Next suppose that the claim holds true for matroids having less than m elements. We show that it also holds for m elements. Let $\mathcal{M} = (E, F)$ be a matroid with $E = [m]$ and $\mathbf{c} = (c_1, \ldots, c_m)$ a cost vector. Let GREEDY denote the social cost of the solution computed by the greedy algorithm for $\mathcal{M}$ and $\mathbf{c}$. Further, let OPT be the optimal social cost with respect to $\mathcal{M}$ and $\mathbf{c}$ and let $B \in F$ be a minimum cost basis, i.e., $\sum_{i \in B} c_i = \mathsf{OPT}$. We assume without loss of generality that the minimum cost element in B is $i = 1$.

We use $\mathsf{GREEDY}^{(i)}$ and $\mathsf{OPT}^{(i)}$ to denote the respective social costs when the underlying matroid is $\mathcal{M}^{(i)} = (E^{(i)}, F^{(i)})$ and the cost vector is $\mathbf{c}_{-i}$ (see Definition 6).

Using the linearity of expectation and Bayes Rule, we can re-write GREEDY as

$$\begin{aligned} \mathsf{GREEDY} &= \sum_{i \in E} d_i^E(\mathbf{c})(c_i + \mathsf{GREEDY}^{(i)}) \\ &= \sum_{i \in B} d_i^E(\mathbf{c})(c_i + \mathsf{GREEDY}^{(i)}) + \sum_{i \in E \setminus B} d_i^E(\mathbf{c})(c_i + \mathsf{GREEDY}^{(i)}) \end{aligned}$$

We will bound the two sums in the expression above separately.

First sum: Consider some element $i \in B$. The optimality of B with respect to $\mathcal{M}$ and $\mathbf{c}$ implies optimality of $B - i$ for the matroid $\mathcal{M}^{(i)}$ and $\mathbf{c}_{-i}$. Thus, $\mathsf{OPT}^{(i)} = \mathsf{OPT} - c_i$. Note that the matroid $\mathcal{M}^{(i)}$ has $m-1$ elements and every basis has size $r-1$. Hence, the ratio $(m-r)/2+1$ remains constant. By the induction hypothesis, we thus have

$$\mathsf{GREEDY}^{(i)} \le \Big(\frac{m-r}{2}+1\Big)(\mathsf{OPT} - c_i)$$

Each term of the summation is therefore at most

$$d_i^E(\mathbf{c})\Big(c_i + \Big(\frac{m-r}{2}+1\Big)(\mathsf{OPT} - c_i)\Big)$$

Because $(m-r)/2 \ge 0$, the maximum term of the summation is attained when c_i is minimum. By our assumption, this is the case for $i = 1$. Therefore, for every $i \in B$ we have that

$$\begin{aligned} d_i^E(\mathbf{c})(c_i + \mathsf{GREEDY}^{(i)}) &\le d_i^E(\mathbf{c})\Big(c_i + \Big(\frac{m-r}{2}+1\Big)(\mathsf{OPT} - c_i)\Big) \\ &\le d_i^E(\mathbf{c})\Big(c_1 + \Big(\frac{m-r}{2}+1\Big)(\mathsf{OPT} - c_1)\Big) \end{aligned}$$

Summing over all elements in B and exploiting the fact that $\sum_{i\in B} d_i^E(\mathbf{c})c_1 \le c_1$, we conclude that

$$\sum_{i\in B} d_i^E(\mathbf{c})(c_i + \mathsf{GREEDY}^{(i)}) \le c_1 + \sum_{i\in B} d_i^E(\mathbf{c})\Big(\frac{m-r}{2}+1\Big)(\mathsf{OPT} - c_1) \qquad (2)$$

Second sum: We first argue that for every element $i \notin B$ the matroid $\mathcal{M}^{(i)}$ with cost vector $\mathbf{c}_{-i}$ attains a basis of cost at most $\mathsf{OPT} - c_1$. By the exchange property of the matroid definition, there exists some $j \in B$ such that $(B-j)+i$ is a basis of $\mathcal{M}$ with cost $\mathsf{OPT} - c_j + c_i$. By the definition of $\mathcal{M}^{(i)}$, $B - j$ is a basis of $\mathcal{M}^{(i)}$ and has cost $\mathsf{OPT} - c_j$. Exploiting that the minimum cost element in B is $i = 1$, we conclude that

$$\mathsf{OPT}^{(i)} \le \mathsf{OPT} - c_j \le \mathsf{OPT} - c_1.$$

Using induction hypothesis and Proposition 3 we have that

$$\begin{aligned} d_i^E(\mathbf{c})(c_i + \mathsf{GREEDY}^{(i)}] &\le d_i^E(\mathbf{c})c_i + d_i^E(\mathbf{c})\Big(\frac{m-r}{2}+1\Big)\mathsf{OPT}^{(i)} \\ &\le \frac{1}{2}c_1 + d_i^E(\mathbf{c})\Big(\frac{m-r}{2}+1\Big)(\mathsf{OPT} - c_1). \end{aligned}$$

Summing over al element in $E \setminus B$, we obtain

$$\sum_{i\in E\setminus B} d_i^E(\mathbf{c})(c_i + \mathsf{GREEDY}^{(i)}) \le \frac{m-r}{2}c_1 + \sum_{i\in E\setminus B} d_i^E(\mathbf{c})\Big(\frac{m-r}{2}+1\Big)(\mathsf{OPT} - c_1) \qquad (3)$$

Adding (2) and (3) and using the fact that $\sum_{i\in E} d_i^E(\mathbf{c}) = 1$ it follows that

$$\mathsf{GREEDY} \leq \Big(\frac{m-r}{2}+1\Big)c_1 + \sum_{i\in E} d_i^E(\mathbf{c})\Big(\frac{m-r}{2}+1\Big)(\mathsf{OPT}-c_1)\Big(\frac{m-r}{2}+1\Big)\mathsf{OPT}.$$

This concludes the proof.

C Missing proofs of Section 5

Definition 7. *Let $\mathcal{M} = (E, F)$ be a matroid, B be a base of the matroid and $i \in B$. We define $S(i, B)$ as the number of elements $j \notin B$ such that $(B + j) - i$ is also a base of the matroid. Let $S(\mathcal{M}) = \max_{i,B} S(i, B)$.*

Lemma 1. *Given any matroid $\mathcal{M} = (E, F)$ there is cost vector such that any mechanism that is truthful in expectation has approximation ratio at least $S(\mathcal{M})/2+1$.*

Proof. Let B and i be the corresponding elements of the matroid such that $S(i, B) = S(\mathcal{M})$. Set $c_j = 0$ for all $j \in B - i$, $c_i = 1$ and $c_j = L$ for all $j \notin B$, where L is a large number. Any optimal truthful mechanism will include all agents with cost zero simply because these agents are indifferent in being included in the outcome or not and electing them always minimizes the overall cost. The remaining problem corresponds exactly to the lower bound instance in [8]. Because the total number of elements we can choose from is $S(\mathcal{M}) + 1$ the bound becomes $S(\mathcal{M})/2 + 1$.

C.1 Proof of Theorem 4

Consider the case of the r-uniform matroid (see Example 2). The bases of this matroid are precisely all subsets of E that contain r elements. By definition, every element in a base can be substituted by any other element. Thus, we have that $S(\mathcal{M}) = m - r$. Using Lemma 1, the claim follows.

C.2 Proof of Theorem 5

Consider the complete bipartite graph with n vertices in each part, say L and R, of the partition. Then we connect all vertices in each part L and R, respectively, using a single path of length $n - 1$. The resulting graph has $2n$ vertices and $n^2 + 2n - 2$ edges. Let B be the base that contains the two paths that connect the vertices in L and R plus one additional edge of the n^2 many edges that connect vertices in L with vertices in R. Since we can substitute the last edge with any other of the n^2 edges we deduce that $S(\mathcal{M}) = n^2 - 1$, which yields a lower bound of $(n^2 - 1)/2 + 1$. Since in this case $m = n^2 + 2n - 2$ and $r = 2n - 1$ we get that $(n^2 - 1)/2 = (m - r)/2 + 1$ as desired.

D Missing Proofs of Section 6

D.1 Proof of Theorem 6

Following a similar line of arguments as in the proof of Theorem 2, it is easy to verify that the proportional distributions satisfy the conditions of Theorem 1. The greedy mechanism is therefore truthful in expectation.

As before, we use OPT and GREEDY to denote the cost of the optimal solution and the solution output by the greedy algorithm (with respect to $\mathbf{c}$). Let B denote a minimum cost base such that $\sum_{i \in B} c_i = \mathsf{OPT}$. Let T_k be the set of feasible extensions (corresponding to the set T in Step 2(a)) in the k-th iteration of the greedy algorithm.

Recall the we assume that the greedy algorithm uses the proportional distribution to choose an element from T_k. In particular, the probability that element $i \in T_k$ is chosen is equal to

$$d_i^{T_k}(\mathbf{c}_{T_k}) = \frac{c_i^{-1}}{\sum_{j \in T_k} c_j^{-1}} = \frac{c_i^{-1}}{\sum_{j \in T_k \cap B} c_j^{-1} + \sum_{j \in T_k \setminus B} c_j^{-1}}.$$

We next derive lower bounds for each summation in the denominator.

By assumption, we have $c_i \leq \mathsf{OPT}$ and thus

$$\sum_{j \in T_k \setminus B} c_j^{-1} \geq \frac{|T_k \setminus B|}{\mathsf{OPT}}. \tag{4}$$

Further, using the matroid property we know that in the k-th iteration of the algorithm there are at least $|B| - k + 1$ elements of B in T_k. Thus, $|T_k \cap B| \geq |B| - k + 1$.

Proposition 4. *Let $a_1, \ldots, a_n \in \mathbb{R}_+$ and define $A = \sum_{i \in [n]} a_i$. Then $\sum_{i \in [n]} a_i^{-1}$ is minimized when $a_i = A/n$ for every $i \in [n]$ and the summation becomes $\sum_{i \in [n]} a_i^{-1} = n^2/A$.*

Proof. We use induction on n. For $n = 1$ the result follows trivially. Assume it holds for n. We will show it holds for $n + 1$. For simplicity let $a_{n+1} = x \cdot A$ for some x with $0 < x < 1$. If we fix x then we have that $\sum_{i \in [n]} a_i = (1 - x)A$. By the induction hypothesis, we get that

$$\min \sum_{i \in [n]} a_i^{-1} = \frac{n^2}{(1 - x)A}.$$

Hence, we seek to find some x that minimizes

$$\frac{1}{A}\left(\frac{n^2}{(1 - x)} + \frac{1}{x}\right).$$

Taking the derivative of this function and setting it to zero, we obtain

$$n^2 = \left(\frac{x - 1}{x}\right)^2.$$

The only solution to this equality such that x is positive is $x = \frac{1}{n+1}$. Also, the second derivative is positive for $0 < x < 1$. Hence, $x = \frac{1}{n+1}$ minimizes the expression above. The sum becomes

$$\frac{n^2}{A\frac{n}{n+1}} + \frac{1}{A\frac{1}{n+1}} = \frac{(n+1)^2}{A},$$

which concludes the proof.

Using Proposition 4 and the fact that $\sum_{j \in T_k \cap B} c_j \leq \sum_{j \in B} c_j = \mathsf{OPT}$ we can bound the other summand of the denominator as follows

$$\sum_{j \in T_k \cap B} c_j^{-1} \geq \frac{|T_k \cap B|^2}{\sum_{j \in T_k \cap B} c_j} \geq \frac{|T_k \cap B|^2}{\mathsf{OPT}}. \tag{5}$$

Combininig (4) and (5) we conclude that the probability that agent $i \in T_k$ is chosen is

$$d_i^{T_k}(\mathbf{c}_{T_k}) = \frac{c_i^{-1}}{|T_k \setminus B| + |T_k \cap B|^2} \mathsf{OPT}.$$

We can therefore upper bound the expected contribution of elements in $T_k \setminus B$ to the overall cost by

$$\sum_{i \in T_k \setminus B} c_i d_i^{T_k}(\mathbf{c}_{T_k}) \leq \sum_{i \in T_k \setminus B} \frac{\mathsf{OPT}}{|T_k \setminus B| + |T_k \cap B|^2} = \frac{|T_k \setminus B| \cdot \mathsf{OPT}}{|T_k \setminus B| + |T_k \cap B|^2}.$$

Note that each term in the summation is at most 1. Further, we have $|T_k \setminus B| \leq m$. As argued above, in the k-th iteration we have $|T_k \cap B| \geq |B| - k + 1$. Combining these facts we get that

$$\begin{aligned} \frac{|T_k \setminus B| \cdot \mathsf{OPT}}{|T_k \setminus B| + |T_k \cap B|^2} &\leq \frac{m \cdot \mathsf{OPT}}{m + |T_k \cap B|^2} \\ &\leq \frac{m \cdot \mathsf{OPT}}{m + (|B| - k + 1)^2} \end{aligned}$$

In the final cost of greedy there is some expected contribution of elements outside B. We upper-bounded the contribution in round k by $\frac{m \cdot \mathsf{OPT}}{m + (|B| - k + 1)^2}$. The total contribution cannot exceed the summation of the maximum contribution in each step. Also the contribution of elements in B towards the cost of the solution is trivially upper-bounded by OPT. Hence, we conclude that

$$\mathsf{GREEDY} \leq \mathsf{OPT} + \mathsf{OPT} \sum_{k=1}^{|B|} \frac{m \cdot \mathsf{OPT}}{m + (|B| - k + 1)^2}$$

Define $R = \mathsf{GREEDY}/\mathsf{OPT}$. Then the above inequality implies that

$$R \le 1 + \sum_{k=1}^{|B|} \frac{m}{m + (|B| - k + 1)^2} = 1 + \sum_{l=1}^{|B|} \frac{m}{m + l^2},$$

where we substituted $l = |B| - k + 1$. We split the final sum as follows:

$$R \le 1 + \sum_{l=1}^{\sqrt{m}-1} \frac{m}{m + l^2} + \sum_{l=\sqrt{m}}^{|B|} \frac{m}{m + l^2} \le 1 + \sum_{l=1}^{\sqrt{m}-1} 1 + \sum_{l=\sqrt{m}}^{|B|} \frac{m}{l^2} = \sqrt{m} + m \sum_{l=\sqrt{m}}^{|B|} \frac{1}{l^2}.$$

Proposition 5. *For every $t \ge 1$, $\sum_{k=t}^{\infty} \frac{1}{k^2} \le 2/t$.*

Proof. Note that for every $k \in \mathbb{N}$ we have $k \ge \sqrt{k}$. Using this inequality, we obtain $k(k+1)/2 \le k^2$. Thus

$$\sum_{k=t}^{\infty} \frac{1}{k^2} \le \sum_{k=t}^{\infty} \frac{2}{k(k+1)} = \sum_{k=t}^{\infty} \frac{2}{k} - \frac{2}{k+1} = \frac{2}{t}.$$

Finally, using Proposition 5 we obtain

$$R \le \sqrt{m} + 2\sqrt{m} = 3\sqrt{m},$$

which completes the proof of the theorem.