Uniform Budgets and the Envy-Free Pricing Problem

Patrick Briest*

Dept. of Computer Science, University of Liverpool, UK.
patrick.briest@liverpool.ac.uk

Abstract

We consider the unit-demand min-buying pricing problem, in which we want to compute revenue maximizing prices for a set of products $\mathcal{P}$ assuming that each consumer from a set of consumer samples $\mathcal{C}$ will purchase her cheapest affordable product once prices are fixed. We focus on the special uniform-budget case, in which every consumer has only a single non-zero budget for some set of products. This constitutes a special case also of the unit-demand envy-free pricing problem.

We show that, assuming specific hardness of the balanced bipartite independent set problem in constant degree graphs or hardness of refuting random 3CNF formulas, the unit-demand min-buying pricing problem with uniform budgets cannot be approximated in polynomial time within $\mathcal{O}(\log^{\varepsilon} |\mathcal{C}|)$ for some $\varepsilon > 0$. This is the first result giving evidence that unit-demand envy-free pricing, as well, might be hard to approximate essentially better than within the known logarithmic ratio.

We then introduce a slightly more general problem definition in which consumers are given as an explicit probability distribution and show that in this case the envy-free pricing problem can be shown to be inapproximable within $\mathcal{O}(|\mathcal{P}|^{\varepsilon})$ assuming NP $\not\subseteq \bigcap_{\delta>0}$ BPTIME($2^{\mathcal{O}(n^{\delta})}$). Finally, we briefly argue that all the results apply to the important setting of pricing with single-minded consumers as well.

1 Introduction

Inspired by the possibility of gathering large amounts of data about the preferences and budgets of a company's potential customers by web sites designed for this purpose, Rusmevichientong [18] and Glynn et al. [13] introduced a class of so called *multi-product pricing* problems that aim at computing optimal pricing schemes for a company's product range. In the original version of the problem each consumer is represented by a budget and a set of products she is interested in. Given fixed prices for the products, she decides to buy one of the products she is interested in with a price not

*Supported by DFG grant Kr 2332/1-2 within Emmy Noether program.

exceeding her budget. The decision is made corresponding to either the *min-buying*, *max-buying*, or *rank-buying* model, where the consumer buys the product with lowest price not exceeding the budget, highest price not exceeding the budget, or highest rank according to some consumer specific ranking, respectively. All these problems are usually referred to as *unit-demand pricing*, as consumers will decide to buy exactly one product if they can afford to do so.

Aggarwal et al. [1] extend the problem definition and allow consumers with different budgets for the different products they are interested in. Assuming that a *price ladder constraint*, i.e., a predefined order on the prices of all products, is given, they derive a polynomial time approximation scheme (PTAS) for the max-buying and rank-buying (under another reasonable assumption) models. They also show how to obtain logarithmic approximation ratios for all three models if no price ladder is given. Briest and Krysta [7] show that both of these algorithms are essentially best possible.

Guruswami et al. [14] consider a different selection rule, which has already been proposed in [1]. In the *max-gain* model, a consumer buys the product maximizing her personal utility, i.e., the difference between the product's price and her respective budget. In the case of limited product supply, the definition in [14] additionally requires that each consumer must obtain the product she desires most whenever she can afford any product at all. Thus, the resulting pricing scheme must be *envy-free* and we obtain the *unit-demand envy-free pricing* problem, which has received a lot of attention. Guruswami et al. present an algorithm with logarithmic approximation guarantee for this problem and prove APX-hardness. Chawla et al. [8] consider the situation in which instead of having consumer samples, consumers are drawn from an explicit probability distribution. They show that for the case of product distributions (i.e., consumers' budgets are drawn independently for different products), results from optimal auction theory [17] yield constant approximation guarantees.

Another problem introduced in [14] is so called *single-minded pricing*, which is inspired by single-minded combinatorial auction design. In this scenario each consumer has a single budget value and buys the whole set of products she is interested in if the sum of prices does not exceed her budget. Among other results, Guruswami et al. show that techniques similar to those of [1] yield a logarithmic approximation for this problem, which is proven to be close to best possible by Demaine et al. [10]. Balcan and Blum [3] and Briest and Krysta [6] present improved approximation results for several different restricted versions of the problem.

Finally, Balcan et al. [4] show how algorithmic pricing feeds back into incentive-compatible auction design and present competitive auctions based on pricing algorithms combined with an appropriate random sampling procedure.

1.1 Preliminaries

Most of this paper will be focused on the unit-demand min-buying (or envy-free) pricing problem (UDP-MIN) with uniform budgets, which is defined formally below. Throughout the paper we will assume that all products are available in unlimited supply and have zero marginal cost.

Definition 1 *In* uniform-budget UDP-MIN *we are given products* $\mathcal{P}$ *and consumer samples* $\mathcal{C}$ *consisting of budgets* $b_c \in \mathbb{R}_0^+$ *and product sets* $S_c \subseteq \mathcal{P}$ *for all* $c \in \mathcal{C}$*. We want to find prices* $p : \mathcal{P} \rightarrow \mathbb{R}_0^+$ *that maximize*

$$rev_{\mathit{min}}(p) = \sum_{c \in \mathcal{A}(p)} \min\{p(e) \,|\, e \in S_c \wedge p(e) \leq b_c\},$$

where $\mathcal{A}(p) = \{c \in \mathcal{C} \,|\, \exists e \in S_c : p(e) \leq b_c\}$ *denotes the set of consumers that can afford to buy any product given prices* p*.*

Unit-demand pricing models the situation that products are *strict substitutes* and each consumer is interested in purchasing exactly one product out of a set of alternatives. The other extreme is reached if products constitute *strict complements* and every consumer seeks to purchase some specific set of products rather than a single alternative. In the single-minded pricing problem (SMP) we assume that each consumer is interested in exactly one such product set, which she purchases if the sum of prices does not exceed her budget.

Definition 2 *Given products* $\mathcal{P}$ *and consumer samples* $\mathcal{C}$ *consisting of budgets* $b_c \in \mathbb{R}_0^+$ *and product sets* $S_c \subseteq \mathcal{P}$*,* SMP *asks for prices* $p : \mathcal{P} \rightarrow \mathbb{R}_0^+$ *maximizing*

$$rev_{\mathit{smp}}(p) = \sum_{c \in \mathcal{A}(p)} \sum_{e \in S_c} p(e),$$

where $\mathcal{A}(p) = \{c \in \mathcal{C} \mid \sum_{e \in S_c} p(e) \leq b_c\}$*.*

A natural extension of both problems is obtained if we assume that our knowledge of consumer preferences does not stem from some sampling procedure, but that we know the explicit probability distribution over the space $\mathcal{C}^*$ of all possible consumers, which is a widely spread assumption in economics (see, e.g., [8] or [17]). Thus, in the *economist's version* of these problems we are given a probability distribution $\mathcal{D}$ over consumer space $\mathcal{C}^*$. In this situation our aim is to find prices p maximizing the expected revenue from a sale to a single consumer drawn according to distribution $\mathcal{D}$. In order to avoid additional complications (which are of no interest to this paper) we restrict ourselves to finite support distributions, i.e., consumer sets $\mathcal{C}$ with a discrete distribution $\mathcal{D}$ defined on $\mathcal{C}$.

1.2 New Results

We first focus on the sampling-based version of uniform-budget UDP-MIN and prove that assuming specific hardness of refuting random 3SAT-instances or approximating the balanced bipartite independent set problem (BBIS) in constant degree graphs, this problem does not allow approximation guarantees essentially beyond the known logarithmic ratios. The connection between BBIS and UDP-MIN is made via so-called *maximum expanding sequences* (MES), which are a combinatorial formulation of the interaction between different price levels in UDP-MIN and potentially also of independent interest. In order to show hardness of sampling-based UDP-MIN we need hardness of very sparse MES instances, which we obtain from constant degree BBIS by the careful application of derandomized graph products [2] to scale hardness to the desired level. Unfortunately, no explicit hardness of approximation results are known for constant degree BBIS, although the problem has been receiving considerable attention. We show that hardness of constant degree BBIS can be derived from a hypothesis about the average case complexity of refuting random 3SAT-instances, which is almost identical to the one originally put forward by Feige [11] in a similar context. Since UDP-MIN with uniform budgets is a special case of the envy-free pricing problem from [14], for which previously only APX-hardness was known, our results yield the first (strong) evidence that this problem might be hard to approximate within $\mathcal{O}(\log^{\varepsilon} |\mathcal{C}|)$ for some $\varepsilon > 0$. Turning to the economist's version of envy-free pricing we obtain strong hardness of approximation under standard assumptions, in which case the reduction from BBIS to MES yields inapproximability within $\mathcal{O}(|\mathcal{P}|^{\varepsilon})$ for some $\varepsilon > 0$. Similar bounds in terms of the number of products can be shown for the sampling-based version of the problem, if we strengthen the underlying hypothesis a little further.

Finally, we point out that with a few minor modifications the same reductions yield similar hardness results for SMP as well. Even though SMP is known to be hard to approximate within semi-logarithmic ratios [10] in the number of consumer samples, this has some new and interesting implications. First, we obtain the first near-tight hardness results for approximation guarantees expressed in the number of products rather than the number of consumer samples. Second, we obtain lower bounds for the economist's problem version, which can provably not be derived from previous reductions. This also yields evidence that maximum expanding sequences are a combinatorial problem that is implicitly present in quite different combinatorial pricing problems.

Independently of this paper, Chuzhoy et al. [9] have quite recently considered uniform-budget UDP-MIN in a network setting with unit-sized flows, obtaining semi-logarithmic lower bounds on the approximability in terms of the network size. While their result holds under standard complexity theoretic assumptions even for unit-sized flows, our results imply stronger bounds for the non-unit flow case, which corresponds to the economist's version of UDP-MIN.

The rest of this paper is organized as follows. We proceed by giving an exposition of our results on UDP-MIN in Section 2. Section 3 briefly describes the application of our results to SMP. Section 4 concludes.

2 Unit-Demand Pricing

As the main result of this section we describe a reduction from the *Balanced Bipartite Independent Set Problem* (BBIS) in constant degree bipartite graphs to uniform-budget UDP-MIN. This will prove that, assuming there are no randomized polynomial time algorithms approximating constant degree BBIS within arbitrarily small constant factors, there are no polynomial time algorithms approximating UDP-MIN within $\mathcal{O}(\log^{\varepsilon} |\mathcal{C}|)$ for some $\varepsilon > 0$. At the very end of the section we state similar (yet stronger) results for the economist's version of the problem, which hold under standard complexity theoretic assumptions.

Up to now, no explicit hardness results have been proven for BBIS in constant degree graphs, although the problem has been receiving a lot of attention. The first result for general BBIS using a quite moderate complexity theoretic assumption was obtained by Khot [15]. Previous results by Feige [11] and Feige and Kogan [12] are deriving hardness of BBIS under more specific assumptions. In [11] Feige shows an interesting connection between the average case complexity of refuting 3CNF-formulas and the worst case approximation complexity of several notorious optimization problems including BBIS. We are going to formulate a slightly stronger version of the hypothesis in [11] and show that this is enough for our purposes.

Remember that a 3CNF-formula is a conjunction of clauses, each of which is the disjunction of 3 literals over variables $x_1, \ldots, x_n$, where a literal is a variable or its negation. Before stating the hypothesis we need to describe the random sampling procedure used to obtain 3CNF formulas in [11]. Given n variables we create formulas consisting of $m = \Delta n$ clauses for some large constant $\Delta \in \mathbb{N}$ by independently picking each literal of every clause uniformly at random. When Δ is large enough, every truth assignment satisfies roughly $(7/8)m$ clauses of such a random 3CNF formula. Thus, a *typical* random 3CNF formula does not have significantly more than $(7/8)m$ simultaneously satisfiable clauses. On the other hand, for a sufficiently small $\varepsilon > 0$, formulas with $(1-\varepsilon)m$ simultaneously satisfiable clauses can be considered *exceptional*. Hypothesis 1 states that it is hard to detect exceptional formulas on average.

Hypothesis 1 *For every fixed $\varepsilon > 0$ and sufficiently large constant $\Delta \in \mathbb{N}$, there is no (randomized) algorithm that runs in time $\mathcal{O}(t(n))$ and, given a random 3CNF formula with n variables and $m = \Delta n$ clauses, outputs* typical *with probability at least $1/2$ (randomization over input), but outputs* exceptional *on every formula with $(1-\varepsilon)m$ simultaneously satisfiable clauses with probability at least $1 - 1/2^{poly(n)}$ (randomization over algorithm's coin flips).*

Choosing $t(n) = \text{poly}(n)$ the only difference between Hypothesis 1 and the hypothesis in [11] is that we also exclude randomized algorithms that have exponentially small error probability when it comes to detecting exceptional formulas. We need this stronger version as a result of our reduction from BBIS to uniform-budget UDP-MIN, which is partially based on a random construction that introduces an exponentially small one-sided error probability for detecting large independent sets. We are mostly interested here in the case of $t(n) = \text{poly}(n)$. However, going to other subexponential time bounds will allow us to obtain lower bounds for differently parametrized approximation goals. In analogy to [11] we define a notion of hardness based on Hypothesis 1. We use slightly different notation compared to [11] to reflect the difference in the underlying hypotheses.

Definition 3 *A problem is said to be* R3SAT$^\star(t(n))$*-hard, if the existence of a (randomized) polynomial time algorithm (with exponentially small failure probability) for it refutes Hypothesis 1.*

Most importantly, R3SAT$^\star(t(n))$-hard problems do not allow polynomial time algorithms if we believe that Hypothesis 1 is true for the given choice of $t(n)$. As a byproduct of the fact that Hypothesis 1 also excludes certain randomized algorithms, R3SAT$^\star(t(n))$-hardness rules out the existence of this type of algorithm, too. We continue by giving a formal definition of BBIS, the base problem of our reduction.

Definition 4 *In the* Balanced Bipartite Independent Set Problem *(*BBIS*) we are given a bipartite graph $G = (V, W, E)$. We want to find maximum cardinality subsets of vertices $V' \subset V$, $W' \subset W$ with $|V'| = |W'|$, such that $\{v, w\} \notin E$ for all $v \in V'$, $w \in W'$.*

A slightly refined version of the analysis presented in [11] can be used to obtain R3SAT*(poly(n))-hardness of BBIS in constant degree graphs. We point out that this part of the reduction can be replaced by the following weaker hypothesis, which states that the gap variant of BBIS in constant degree graphs does not have randomized polynomial time algorithms with one-sided error (i.e., the decision variant does not belong to class RP). More formally, let $\mathcal{G}^-(a,d)$, $\mathcal{G}^+(b,d)$ be two families of bipartite graphs on $2n$ vertices with constant degree $d \in \mathbb{N}$ and maximum balanced independent set of size at most an or at least bn, respectively. Given $0 < a < b < 1$ and $d \in \mathbb{N}$ the problem BBIS(a,b,d) requires deciding whether $G \in \mathcal{G}^-(a,d)$ or $G \in \mathcal{G}^+(b,d)$ for a given graph $G \in \mathcal{G}^-(a,d) \cup \mathcal{G}^+(b,d)$. For our purposes Hypothesis 2 is fully sufficient.

Hypothesis 2 *There exist constants $0 < a < b < 1$ and $d \in \mathbb{N}$, such that* BBIS*(a,b,d) $\notin$ RP.*

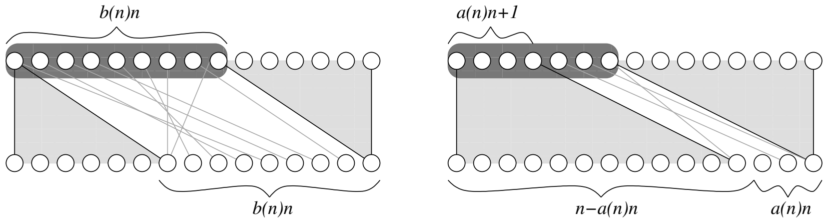
Without expressing too much of an opinion about the validity of Hypothesis 2, it should be noted that it is in accordance with our current knowledge and backed by the fact that strong super-constant approximability thresholds have been proven for general BBIS [15]. Having hardness of constant degree BBIS we apply the method of derandomized graph products [2] to obtain hardness of approximation within $\mathcal{O}(f(n)^\varepsilon)$ for BBIS in graphs with maximum degree $\mathcal{O}(f(n))$, where the appropriate choice for $f(n)$ will become apparent later on. The main part of our result consists of the reduction to UDP-MIN. As an intermediate step in the reduction we modify the BBIS instance by adding a number of random edges and interpret vertices on one side of the bipartition as sets. The connection to UDP-MIN is made by considering sequences of these sets that have a certain expansion property. This is formalized in the following definition.

Definition 5 *In the* Maximum Expanding Sequence Problem *(*MES*) we are given an ordered collection $S_1, \ldots, S_m$ of sets. An expanding sequence $\phi = (\phi(1) < \cdots < \phi(\ell))$ of length $|\phi| = \ell$ is a selection of sets $S_{\phi(1)}, \ldots, S_{\phi(\ell)}$, such that*

$$S_{\phi(j)} \nsubseteq \bigcup_{i=1}^{j-1} S_{\phi(i)}$$

for $2 \leq j \leq \ell$. MES *asks for finding such a sequence of maximum length.*

We are not aware that MES has been considered explicitly before. We briefly point out that a reduction similar to the one given in the proof of Lemma 1 below yields hardness of approximation under a standard assumption, which is formally stated in Theorem 4. In order to reduce MES to UDP-MIN with consumer samples we have to focus our attention on severely restricted problem instances. BBIS instances with bounded maximum degree yield MES instances that exhibit a nicely sparse structure. Definition 6 formalizes our notion of *sparse*.



(a) (b)

Figure 1: Reducing BBIS to MES. (a) Adding random edges with probability $1/(b(n)n)$ each implants an expanding sequence of expected size $\Omega(b(n)n)$ into an independent set of size $b(n)n$. (b) An independent set of size $a(n)n$ allows for expanding sequences of at most twice that size.

Definition 6 *We say that an* MES *instance* $S_1, \ldots, S_m$ *is* κ-separable *if it can be partitioned into* κ *subsequences* $\mathcal{C}_1, \ldots, \mathcal{C}_\kappa$*,* $\mathcal{C}_j = \{S_{k(j)}, S_{k(j)+1}, \ldots, S_{\ell(j)}\}$*, where* $k(1) = 1$*,* $\ell(\kappa) = m$*,* $k(j+1) = \ell(j) + 1$ *for* $1 \leq j \leq \kappa - 1$ *and each* $\mathcal{C}_j$ *contains only non-intersecting sets.*

Our starting point to prove hardness of approximation for sufficiently sparse MES instances is Theorem 1, which can be derived by applying the method of derandomized graph products [2] to constant degree BBIS, and states super-constant approximability thresholds for BBIS parametrized in the graph's super-constant maximum degree $f(n)$.

Theorem 1 *Let* $f : \mathbb{N} \longrightarrow \mathbb{R}^+$ *be non-decreasing with* $f(n) \leq n$ *and* $f(n^c) \leq f(n)^c$ *for all* $c \geq 1$*,* $n \in \mathbb{N}$*. Let* $\mathcal{G}^-(a(n), f(n))$ *and* $\mathcal{G}^+(b(n), f(n))$ *be the families of balanced bipartite graphs on* $2n$ *vertices, with maximum degree bounded by* $f(n)$ *and maximum* BBIS *of size at most* $a(n)n$ *or at least* $b(n)n$*, respectively. There exist* $0 < a(n) < b(n) < 1$ *with* $b(n)/a(n) = \Omega(f(n)^\varepsilon)$ *for some* $\varepsilon > 0$*, such that given* $G \in \mathcal{G}^-(a(n), f(n)) \cup \mathcal{G}^+(b(n), f(n))$ *it is* R3SAT*(poly(n))*-hard to decide whether* $G \in \mathcal{G}^-(a(n), f(n))$ *or* $G \in \mathcal{G}^+(b(n), f(n))$*.*

Lemma 1 *There exists* $\varepsilon > 0$*, such that* MES *with* $f(m)$*-separable instances (*f *as in Theorem 1) is* R3SAT*(poly(n))*-hard to approximate within* $\mathcal{O}(f(m)^\varepsilon)$*.*

Proof: Let some $G \in \mathcal{G}^-(a(n), f(n)) \cup \mathcal{G}^+(b(n), f(n))$, $G = (V, W, E)$, $|V| = |W| = n$, with $a(n)$, $b(n)$ and $f(n)$ as in Theorem 1 be given. We will reduce the problem of deciding whether $G \in \mathcal{G}^-(a(n), f(n))$ or $G \in \mathcal{G}^+(b(n), f(n))$ to solving a separable instance of MES, essentially by implanting large expanding sequences into large balanced independent sets. As neither the independent set nor its size are known at reduction time, we do this by adding a number of random edges to the graph, which create a long expanding sequence in expectation if a large balanced independent set exists (see Fig. 1).

More precisely, every possible edge is independently added to G with probability $(b(n)n)^{-1}$ if it is not already present in the original graph. We then remove vertices whose degree has become too high. In expectation the random procedure above tries to add $b(n)^{-1}$ edges to every vertex $v \in V \cup W$. We remove a vertex v and all its incident edges if more than $c \cdot b(n)^{-1}$ edges are added to it, where c is some sufficiently large constant to be determined later. Let A_v be the random variable counting the number of edges added to v. Applying the Chernoff bound [16] we obtain $\Pr(v \text{ is removed}) = \Pr(A_v \geq c \cdot b(n)^{-1}) \leq \mathrm{e}^{-c/b(n)}$ for any constant $c \geq 3\mathrm{e} - 1$. We denote the modified graph by $G' = (V', W', E')$. For every vertex $v_i \in V'$ we define a corresponding set S_i by $S_i = \{w_j \in W' \mid \{v_i, w_j\} \in E'\}$, i.e., vertices V' will correspond to sets over the universe W' in our MES instance.

In order to obtain a feasible MES instance we need to define an order on sets S_i, which we do next. Observe that vertices in G' have degree at most $f'(n) \leq f(n) + c \cdot b(n)^{-1} = \mathcal{O}(f(n))$, where we use the fact that bipartite graphs with bounded degree $f(n)$ have a balanced independent set of size at least $n/(f(n)+1)$ and, thus, it must be the case that $b(n)^{-1} = \mathcal{O}(f(n))$ in order for the problem to be non-trivial. Furthermore, if the maximum degree of G' is $f'(n)$, then the sets S_i can be partitioned into $f'(n)^2$ classes, such that sets in each class do not intersect, since every set contains at most $f'(n)$ elements, each of which is contained in at most $f'(n) - 1$ further sets. Reordering the sets appropriately we obtain an $\mathcal{O}(f(n)^2)$-separable MES instance.

Soundness: Let $G \in \mathcal{G}^+(b(n), f(n))$. Assume for the moment that no vertices are removed from G and let $\mathcal{S}^* = \{S_{\phi(1)}, \ldots, S_{\phi(\ell)}\}$, $\ell = \lceil b(n)n \rceil$, be the sets in the MES instance corresponding to vertices from V that belong to a maximum balanced bipartite independent set. Analogously, let $W^* \subset W$ denote the vertices from W belonging to the balanced bipartite independent set. For $1 \leq j \leq \ell/2$ consider set $S_{\phi(j)}$. We say that $S_{\phi(j)}$ is *successful* if we can use it to construct a large expanding sequence or, more formally, if the following conditions are satisfied:

- A_j: $|S_{\phi(j)} \cap W^*| = 1$, i.e., $S_{\phi(j)}$ contains exactly one element from W^*.
- B_j: $S_{\phi(j)} \cap S_{\phi(i)} \cap W^* = \emptyset$ for all $1 \leq i \leq \ell/2$, $i \neq j$, i.e., the intersection of $S_{\phi(j)}$ with any other set from $\mathcal{S}^*$ lies outside W^*.
- C_j: The vertex corresponding to set $S_{\phi(j)}$ is not removed due to the degree constraint.
- D_j: None of the vertices in $S_{\phi(j)} \cap W^*$ are removed due to the degree constraint.

It is not difficult to check that successful sets belong to the MES-instance and form an expanding sequence, since their corresponding vertices are not removed from the graph and each set covers a unique element in W^*, which yields the necessary expansion property. Let us now determine the probability that set $S_{\phi(j)}$ is successful. We can write that

$$\begin{aligned}
\Pr(S_{\phi(j)} \text{ is successful}) &= 1 - \Pr(\overline{A}_j \vee \overline{B}_j \vee \overline{C}_j \vee \overline{D}_j) \\
&= 1 - \Pr(\overline{A}_j) - \Pr(\overline{B}_j | A_j) \cdot \Pr(A_j) \\
&\quad - \Pr(\overline{C}_j | A_j \wedge B_j) \cdot \Pr(A_j \wedge B_j) \\
&\quad - \Pr(\overline{D}_j | A_j \wedge B_j \wedge C_j) \cdot \Pr(A_j \wedge B_j \wedge C_j).
\end{aligned}$$

We first consider event A_j and obtain

$$\begin{aligned}\Pr(A_j) &= \sum_{w \in W^*} \Pr\bigl(S_{\phi(j)} \cap W^* = \{w\}\bigr) = \sum_{w \in W^*} \frac{1}{b(n)n}\left(1 - \frac{1}{b(n)n}\right)^{\lceil b(n)n \rceil - 1} \\ &\approx b(n)n \frac{1}{\mathrm{e}b(n)n} \approx \frac{1}{\mathrm{e}},\end{aligned}$$

where the above holds with arbitrary precision for large values of n. Let us then consider $\Pr(\overline{B}_j|A_j)$. Sets $S_{\phi(i)}$ and $S_{\phi(j)}$ contain every element from W^* with equal probability $1/(b(n)n)$. Furthermore, $S_{\phi(i)} \cap W^*$ and $S_{\phi(j)} \cap W^*$ are independent by construction. Applying the union bound yields

$$\Pr(\overline{B}_j|A_j) \leq \sum_{i=1}^{\ell/2} \sum_{w \in W^*} \Pr(w \in S_{\phi(i)}) \Pr(w \in S_{\phi(j)}) \leq \frac{\lceil b(n)n \rceil^2}{2} \frac{1}{(b(n)n)^2} \approx \frac{1}{2},$$

again with arbitrary precision for large n. We have already seen that the probability of any specific vertex being removed due to the degree constraint is bounded above by $\mathrm{e}^{-c/b(n)}$. We conclude that

$$\Pr(\overline{C}_j|A_j \wedge B_j) \cdot \Pr(A_j \wedge B_j) \leq \Pr(\overline{C}_j) \leq \mathrm{e}^{-c/b(n)},$$

and the same estimate holds for $\Pr(\overline{D}_j|A_j \wedge B_j \wedge C_j) \cdot \Pr(A_j \wedge B_j \wedge C_j)$. Thus,

$$\Pr(S_{\phi(j)} \text{ is successful}) \geq 1 - \left(1 - \frac{1}{\mathrm{e}}\right) - \frac{1}{2}\frac{1}{\mathrm{e}} - 2\mathrm{e}^{-c/b(n)} \approx \frac{1}{2\mathrm{e}}$$

for a sufficiently large constant c. Let Y denote the number of successful sets. By linearity of expectation and the above bounds it holds that $\mathrm{E}[Y] \geq (1/4\mathrm{e})b(n)n$. Using that the value of Y is bounded above by $b(n)n$ and applying a Markov type inequality then yields that $\Pr(Y \leq 1/(8\mathrm{e})b(n)n) \leq 1 - 1/(8\mathrm{e})$. This implies that with probability $\Omega(1)$ there exists an expanding sequence of length $\Omega(b(n)n)$.

Completeness: Let $G \in \mathcal{G}^-(a(n), f(n))$ and consider any expanding sequence ϕ in $S_1, \ldots, S_m$. Since the maximum balanced bipartite independent set in G is of size $a(n)n$, every selection of $a(n)n + 1$ vertices from V must be adjacent to all but $a(n)n$ vertices from W. Thus, the first $a(n)n + 1$ sets from ϕ leave at most $a(n)n$ elements uncovered. Since the expansion property requires that every further set in the sequence must contain a previously uncovered element, it follows that $|\phi| \leq 2a(n)n + 1$.

We have shown a randomized reduction with constant one-sided error probability. By repeating the algorithm a polynomial number of times, we obtain error probabilities that are exponentially close to 0. This proves Lemma 1. □

To encode MES in terms of UDP-MIN we translate sets into collections of consumers with exponentially decreasing budgets. Reducing $\log m$-separable MES ensures the resulting instances are of polynomial size.

Theorem 2 *There exists a constant $\varepsilon > 0$, such that it is* R3SAT*(poly(n))*-hard to approximate uniform-budget* UDP-MIN *within $\mathcal{O}(\log^{\varepsilon} |\mathcal{C}|)$. Hardness of approximation holds even under the weaker assumption of Hypothesis 2.*

Proof: Let MES instance $S_1, \dots, S_m$ be separable into $\mathcal{C}_1, \dots, \mathcal{C}_\kappa$, $\kappa = \mathcal{O}(f(m))$. For each element e in the universe of the MES instance we have a corresponding product e. For every set S_i in class $\mathcal{C}_k$ we define a collection of 2^{k-1} identical consumers $C_i = \{c_i^1, c_i^2, \dots, c_i^{2^{k-1}}\}$. Each of these consumers has budget $b_i = 2^{1-k}$ and is interested in products $e \in S_i$. Note that the total number of consumer samples in this construction is bounded above by $m2^{\mathcal{O}(f(m))}$.

Soundness: Let $\phi = (\phi(1) < \cdots < \phi(\ell))$ be an expanding sequence of length ℓ. For every $1 \leq i \leq \ell$ let $N_{\phi(i)}$ denote the elements that are newly covered by $S_{\phi(i)}$. Now, for $i = 1, \dots, \ell$, determine $N_{\phi(i)}$ and set the prices of all products $e \in N_{\phi(i)}$ to b_i. For consumers $C_{\phi(i)}$ it then holds that $p(e) = b_i$ for all $e \in N_{\phi(i)}$, $p(e) > b_i$ for all $e \in S_{\phi(i)} \backslash N_{\phi(i)}$. As a result, all 2^{k-1} consumers belonging to a set $S_{\phi(i)}$ in the expanding sequence will buy at their budget value $b_i = 2^{1-k}$ and jointly contribute revenue 1. Thus, the overall revenue from consumers corresponding to the expanding sequence is at least ℓ.

Completeness: Assume that we are given a price assignment resulting in overall revenue r. First observe that w.l.o.g. all prices are from the set of distinct budget values, i.e., all prices are powers of 2. Then note that w.l.o.g. revenue at least $r/2$ is due to consumers buying at their budget values, since otherwise we could increase overall revenue by doubling all prices. Finally, it is not difficult to see that consumers buying at their budget values form an expanding sequence, as each such consumer must be purchasing a product that none of the consumers with higher budgets is interested in. It follows that we obtain an expanding sequence ϕ of length at least $r/2$. Choosing $f(m) = \log m$ yields the theorem. □

Our reduction is flexible enough to yield inapproximability results also in the maximum number ℓ of non-zero budgets per consumer and, allowing UDP-MIN instances of arbitrary subexponential size, we can stretch the construction to the limit and obtain lower bounds on the approximability in terms of the number of products $|\mathcal{P}|$ as well.

Theorem 3 *There exist constants $\ell_0 \in \mathbb{N}$ and $\varepsilon > 0$, such that for every $\ell \geq \ell_0$ it is* R3SAT*(poly(n))*-hard to approximate uniform-budget* UDP-MIN *with at most ℓ non-zero budgets per consumer within ℓ^ε. Furthermore, for every $\delta > 0$ there exists $\varepsilon > 0$, such that it is* R3SAT*($2^{\mathcal{O}(n^\delta)}$)*-hard to approximate uniform-budget* UDP-MIN *within $\mathcal{O}(|\mathcal{P}|^\varepsilon)$.*

Next, let us consider the economist's versions of uniform-budget UDP-MIN and the unit-demand envy-free pricing problem. As mentioned before, a reduction similar to the one given in the proof of Lemma 1 in combination with the known hardness results for general BBIS from [15] yields a strong hardness result for general MES. Applying the reduction from the proof of Theorem 2 we obtain inapproximability results for uniform-budget UDP-MIN (economist's version) under standard complexity theoretic assumptions. These immediately extend to the more general unit-demand envy-free pricing problem.

Theorem 4 MES *is inapproximable within $\mathcal{O}(m^\varepsilon)$ for some $\varepsilon > 0$, assuming that* NP $\not\subseteq \bigcap_{\delta>0}$ BPTIME($2^{\mathcal{O}(n^\delta)}$).

Theorem 5 UDP-MIN *(economist's version) with uniform budgets is hard to approximate within $\mathcal{O}(|\mathcal{P}|^{\varepsilon})$ for some $\varepsilon > 0$, if* NP $\not\subseteq \bigcap_{\delta>0}$ BPTIME($2^{\mathcal{O}(n^{\delta})}$).

Finally, let us mention that it is not difficult to achieve approximation guarantee $\mathcal{O}(|\mathcal{P}|)$ for the economist's envy-free pricing problem, e.g., by using the known single-price algorithm [1].

3 Single-Minded Pricing

We can adapt the reduction in Theorem 2 to work for single-minded consumers as well. In fact, all we need to do is to choose price levels as powers of $2|\mathcal{P}|$ rather than 2 and define consumers in the opposite direction, i.e., start on the lowest price level and work our way up. Again only consumers corresponding to an expanding sequence allow extraction of full revenue. Demaine et al. [10] prove an approximation threshold of $\Omega(\log^{\varepsilon}|\mathcal{C}|)$ for SMP under standard complexity theoretic assumptions. Our reduction yields these and asymptotically stronger bounds in the number of products based on the notion of R3SAT*-hardness.

Theorem 6 *For every $\delta > 0$ there exists $\varepsilon > 0$, such that it is R3SAT*($2^{\mathcal{O}(n^{\delta})}$)-hard to approximate* SMP *within $\mathcal{O}(|\mathcal{P}|^{\varepsilon})$.*

Once more, turning to the economist's version of the problem, we obtain strong inapproximability results under standard assumptions. We point out that it is not possible to achieve these bounds by previous reductions.

Theorem 7 SMP *(economist's version) is hard to approximate within $\mathcal{O}(|\mathcal{P}|^{\varepsilon})$ for some $\varepsilon > 0$, if* NP $\not\subseteq \bigcap_{\delta>0}$ BPTIME($2^{\mathcal{O}(n^{\delta})}$).

4 Conclusions

In this paper we have made progress towards understanding the difficulty of different combinatorial pricing problems. First, we have shown that assuming specific hardness of constant degree BBIS or hardness on average of refuting random 3CNF-formulas, the unit-demand min-buying pricing problem with uniform budgets, which constitutes a special case also of unit-demand envy-free pricing, does not allow sub-logarithmic approximation guarantees. Secondly, we have shown that our techniques apply to the case of single-minded pricing as well, which indicates that expanding sequences are a common source of hardness for quite different combinatorial pricing problems.

5 Acknowledgments

The author thanks Piotr Krysta for insightful discussions and several anonymous referees for their valuable comments.

References

[1] G. Aggarwal, T. Feder, R. Motwani, and A. Zhu. Algorithms for Multi-Product Pricing. In *Proc. of 31st International Colloquium on Automata, Languages, and Programming (ICALP)*, 2004.

[2] N. Alon, U. Feige, A. Wigderson, and D. Zuckerman. Derandomized Graph Products. *Computational Complexity*, **5**(1): 60–75, 1995.

[3] N. Balcan and A. Blum. Approximation Algorithms and Online Mechanisms for Item Pricing. In *Proc. of 7th ACM Conference on Electronic Commerce (EC)*, 2006.

[4] N. Balcan, A. Blum, J. Hartline, and Y. Mansour. Mechanism Design via Machine Learning. In *Proc. of 46th IEEE Symposium on Foundations of Computer Science (FOCS)*, 2005.

[5] P. Berman and G. Schnitger. On the Complexity of Approximating the Independent Set Problem. *Information and Computation*, **96**(1): 77–94, 1992.

[6] P. Briest and P. Krysta. Single-Minded Unlimited-Supply Pricing on Sparse Instances. In *Proc. of 17th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2006.

[7] P. Briest and P. Krysta. Buying Cheap is Expensive: Hardness of Non-Parametric Multi-Product Pricing. In *Proc. of 18th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2007.

[8] S. Chawla, J. Hartline, and R. Kleinberg. Algorithmic Pricing via Virtual Valuations. In *Proc. of 8th ACM Conference on Electronic Commerce (EC)*, 2007.

[9] J. Chuzhoy, S. Kannan, and S. Khanna. Network Pricing for Multicommodity Flows. Unpublished Manuscript, 2007.

[10] E.D. Demaine, U. Feige, M.T. Hajiaghayi, and M.R. Salavatipour. Combination Can Be Hard: Approximability of the Unique Coverage Problem. In *Proc. of 17th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2006.

[11] U. Feige. Relations between Average Case Complexity and Approximation Complexity. In *Proc. of 34th ACM Symposium on Theory of Computing (STOC)*, 2002.

[12] U. Feige and S. Kogan. Hardness of Approximation of the Balanced Complete Bipartite Subgraph Problem. *Technical Report MCS04-04*, Dept. of Computer Science and Applied Math., The Weizmann Institute of Science, 2004.

[13] P. Glynn, P. Rusmevichientong, and B. Van Roy. A Non-Parametric Approach to Multi-Product Pricing. *Operations Research*, **54**(1):82–98, 2006.

[14] V. Guruswami, J.D. Hartline, A.R. Karlin, D. Kempe, C. Kenyon, and F. McSherry. On Profit-Maximizing Envy-Free Pricing. In *Proc. of 16th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2005.

[15] S. Khot. Ruling out PTAS for Graph Min-Bisection, Densest Subgraph and Bipartite Clique. In *Proc. of 45th IEEE Symposium on Foundations of Computer Science (FOCS)*, 2004.

[16] R. Motwani and P. Raghavan. *Randomized Algorithms.* Cambridge University Press, 1995.

[17] R. Myerson. Optimal Auction Design. *Mathematics of Operations Research*, 6, 58–73, 1981.

[18] P. Rusmevichientong. *A Non-Parametric Approach to Multi-Product Pricing: Theory and Application.* Ph.D. dissertation, Stanford University, 2003.

A Constant Degree BBIS: Missing Proofs

We will show R3SAT*(poly(n))-hardness of constant degree BBIS by a reduction from MAX-3AND. Given a collection of clauses, each of which contains 3 (not necessarily distinct) literals and is satisfied if all 3 literals are assigned the boolean value true, we want to determine the maximum number of simultaneously satisfiable clauses. The remainder of this part of the proof is roughly identical to the one in [11], except for the fact that a small change in the reduction yields graphs of constant degree. Lemma 2 is explicitly stated in [11] for the case of their underlying hypothesis and extends easily to our notion of R3SAT*(poly(n))-hardness. We note that if we talk about random MAX-3AND instances, we assume the sampling procedure as described in Section 2.

Lemma 2 ([11]) *For every fixed $\varepsilon > 0$ and sufficiently large $\Delta \in \mathbb{N}$, the following problem is* R3SAT*(poly(n))*-hard. Given a random* 3AND *formula with n variables and $m = \Delta n$ clauses, output* typical *with probability at least $1/2$, but output* exceptional *on every formula with $(1/4 - \varepsilon)m$ simultaneously satisfiable clauses.*

We want to show that if we have some good approximation algorithm for BBIS in constant degree graphs, then we can use it to design a refutation algorithm for MAX-3AND, which contradicts Hypothesis 1. Before doing this, we introduce the following technical lemma, which states an upper bound on the probability of a random variable with bounded range falling far below its expectation, similar to the Markov inequality [16].

Lemma 3 *Let $X \in [0, s]$ be random variable with $\mathrm{E}[X] \geq \eta s$ for some $0 < \eta < 1$. Then*

$$\Pr\left(X \leq \frac{\eta s}{t}\right) \leq \frac{1-\eta}{1-\frac{\eta}{t}}$$

for any $t > 1$.

Proof: Towards a contradiction, assume that the claim does not hold. We may then write that

$$\begin{aligned}
\mathrm{E}[X] &\leq \Pr\left(X \leq \frac{\eta s}{t}\right) \cdot \frac{\eta s}{t} + \Pr\left(X > \frac{\eta s}{t}\right) \cdot s \\
&< \frac{1-\eta}{1-\frac{\eta}{t}} \cdot \frac{\eta s}{t} + \left(1 - \frac{1-\eta}{1-\frac{\eta}{t}}\right) s \\
&= \frac{\eta}{1-\frac{\eta}{t}} \left(\frac{1}{t} - \frac{\eta}{t} + 1 - \frac{1}{t}\right) s \\
&= \eta s,
\end{aligned}$$

a contradiction. □

Let us now have a closer look at the random 3AND formulas we are given as an input. Clearly, in expectation each literal will appear $(3/2)\Delta$ times in the formula.

Now let V_i be a random variable counting the number of occurrences of literal ℓ_i. Applying the Chernoff bound [16] we have that

$$\Pr\left[(1-\delta)\frac{3}{2}\Delta \le V_i \le (1+\delta)\frac{3}{2}\Delta\right] \ge 1 - 2\mathrm{e}^{-(3/4)\delta^2\Delta}$$

for any $0 < \delta < 1$. For every literal we define an additional random variable $X_i \in \{0,1\}$ that indicates whether the above condition is satisfied and let $X = X_1 + \cdots + X_{2n}$. By linearity of expectation it obviously holds that

$$\mathrm{E}[X] \ge \left(1 - 2\mathrm{e}^{-(3/4)\delta^2\Delta}\right)2n.$$

This implies that

$$\Pr\left[X < (1-\sqrt{2}\mathrm{e}^{-(3/8)\delta^2\Delta})2n\right] \le \sqrt{2}\mathrm{e}^{-(3/8)\delta^2\Delta},$$

by Lemma 3 with $\eta = 1 - 2\mathrm{e}^{-(3/4)\delta^2\Delta}$,

$$t = (1 - 2\mathrm{e}^{-(3/4)\delta^2\Delta})/(1-\sqrt{2}\mathrm{e}^{-(3/8)\delta^2\Delta})$$

and $s = 2n$. Now fix any $\gamma > 0$ and observe that by choosing Δ sufficiently large we can ensure that $\sqrt{2}\mathrm{e}^{-(3/8)\delta^2\Delta} \le \gamma$.

Fact 1 *With probability* $1-\gamma$ *a* $(1-\gamma)$*-fraction of the literals appear between* $(1-\delta)\frac{3}{2}\Delta$ *and* $(1+\delta)\frac{3}{2}\Delta$ *times in a random* MAX-3AND *formula.*

The first step of our refutation algorithm for MAX-3AND consists of checking the above condition. If too many literals deviate from their expected number of occurrences, the algorithm outputs *exceptional*. If this is not the case, we continue by removing the few problematic literals from the formula. More precisely, we remove every clause that contains a literal appearing more than $(1+\delta)(3/2)\Delta$ times.

Let $\mu = 3(\delta+\gamma)$. We know that $(1-\gamma)2n$ *good* literals appear at least $(1-\delta)(3/2)\Delta$ times within the formula. Thus, a total number of at least

$$(1-\gamma)2n(1-\delta)(3/2)\Delta \ge (1-\delta-\gamma)3\Delta n$$

literal occurrences belong to *good* literals. This leaves at most $(\delta+\gamma)3\Delta n = \mu m$ literal occurrences belonging to *bad* literals and, consequently, gives an upper bound on the number of clauses that are removed from the formula. For the rest of the reduction to BBIS we need two more facts. Fact 3 is explicitly proven in [11]. Fact 2 is immediate from the above.

Fact 2 *If the original* MAX-3AND *formula had* $(1/4-\varepsilon)m$ *satisfiable clauses, the number of satisfiable clauses in our modified formula is bounded below by* $(1/4-\varepsilon-\mu)m$.

Fact 3 *For every* $\varepsilon > 0$*, sufficiently large* $\Delta \in \mathbb{N}$ *and* n *large enough, the following holds. With high probability every set of* $(1/8+\varepsilon)m$ *clauses in a random* MAX-3AND *formula with* $m = \Delta n$ *clauses contains at least* $n+1$ *different literals.*

We transform the modified formula into an instance of BBIS as follows. On both sides of the bipartition we have a vertex for every clause of the formula. Vertices on opposite sides are connected by an edge, if the corresponding clauses contain conflicting literals, i.e., if some variable appears in positive form in one clause and in negative form in the other. Thus, two vertices are connected if and only if the corresponding clauses cannot be satisfied simultaneously.

It is straightforward to argue that $(1/4 - \varepsilon - \mu)m$ satisfiable clauses result in a balanced bipartite independent set of at least the same size, since for any given truth assignment we can select the vertices corresponding to satisfied clauses on both sides of the bipartition as a balanced bipartite independent set. On the other hand, for random formulas the size of the maximum balanced bipartite independent set is bounded above by $(1/8 + \varepsilon)m$ with high probability, since by Fact 3 every selection of $(1/8 + \varepsilon)m$ clauses contains at least $n + 1$ distinct literals with high probability and, thus, is not satisfiable because one literal must appear in both positive and negative form. Additionally we know that, since every clause contains 3 literals and every literal appears at most $(1 + \delta)(3/2)\Delta$ times, the resulting bipartite graph has a maximum degree of at most $(1 + \delta)(9/2)\Delta$.

Assume now we had some polynomial time algorithm that can distinguish the two cases with an error probability exponentially close to 0. By applying this algorithm to the above BBIS instance we immediately obtain a polynomial time refutation algorithm for MAX-3AND with exponentially small failure probability for detecting exceptional formulas. If the BBIS algorithm returns a balanced bipartite independent set larger than $(1/8 + \varepsilon)m$, we output *exceptional*. Otherwise, we output *typical*. The failure probability for detecting typical formulas is dominated by the probability that the formula has too many literals deviating from their expected number of occurrences and, thus, can be made an arbitrarily small constant. Hence, we have shown the following lemma.

Lemma 4 *Let $\mathcal{G}^-(a, d)$, $\mathcal{G}^+(b, d)$ be the families of bipartite graphs on $2n$ vertices with maximum degree bounded by $d \in \mathbb{N}$ and a maximum balanced bipartite independent set of size at most an or at least bn, respectively. There exist $0 < a < b < 1$ and $d \in \mathbb{N}$, such that deciding whether a given graph $G \in \mathcal{G}^-(a, d) \cup \mathcal{G}^+(b, d)$ belongs to $\mathcal{G}^-(a, d)$ or $\mathcal{G}^+(b, d)$ is* R3SAT*(poly(n))*-hard.*

B Graph Products: Missing Proofs

For a bipartite graph $G = (V, W, E)$, $|V| = |W| = n$, let $\alpha(G)$ refer to the size of a maximum balanced bipartite independent set in G. Let $\mathcal{G}^-(a, d)$ and $\mathcal{G}^+(b, d)$ be two families of bipartite graphs with maximum degree bounded by d and $\alpha(G) \leq an$ for $G \in \mathcal{G}^-(a, d)$, $\alpha(G) \geq bn$ for $G \in \mathcal{G}^+(b, d)$. From Section A we know that we can choose constants a, b and d, such that deciding whether a given graph is from $\mathcal{G}^-(a, d)$ or $\mathcal{G}^+(b, d)$ is hard assuming Hypothesis 1 holds. The following definition is in analogy to the definition in [5].

Definition 7 *Let $G = (V, W, E)$, $|V| = |W| = n$, be a bipartite graph and $k \in \mathbb{N}$. The k-*fold graph product *$G^k = (V^k, W^k, E_k)$ is defined by Cartesian products V^k,*

W^k and $\{(v_1,\ldots,v_k),(w_1,\ldots,w_k)\} \in E_k$ if and only if $\{v_1,\ldots,v_k,w_1,\ldots,w_k\}$ is not a bipartite independent set in graph G.

We briefly describe the application of the technique of derandomized graph products due to Alon et al. [2] to bipartite graphs. Given $G = (V,W,E)$, $|V| = |W| = n$, we construct a non-bipartite δ-regular Ramanujan graph H on n vertices with constant degree δ (depending only on constants a and b). Vertices V^k and W^k of the derandomized graph product DG^k are obtained by choosing a vertex of H uniformly at random and taking a random walk of length $k-1$ starting at this vertex. For $k = O(\log n)$ the number $n\delta^{k-1}$ of such random walks is polynomial and, thus, DG^k can be constructed deterministically in polynomial time. The edges of DG^k are defined as in Definition 7.

An analysis similar to the standard case of amplifying general independent sets in non-bipartite graphs yields Theorem 1 (see Section 2). We want to remark that by construction the constant degree graphs obtained by the reduction in Section A are symmetric in the sense that we can rename vertices $V = \{v_1,\ldots,v_n\}$ and $W = \{w_1,\ldots,w_n\}$, such that $\{v_i,w_j\} \in E$ if and only if $\{v_j,w_i\} \in E$. This property is not lost during gap amplification, since we can use the same expander graph to obtain the vertices on both sides of the graph product.

C Unit-Demand Pricing: Missing Proofs

Theorem 3 *There exist constants $\ell_0 \in \mathbb{N}$ and $\varepsilon > 0$, such that for every $\ell \geq \ell_0$ it is* R3SAT*(poly(n))*-hard to approximate uniform-budget* UDP-MIN *(computer scientist's version) with at most ℓ non-zero budgets per consumer within ℓ^ε. Furthermore, for every $\delta > 0$ there exists $\varepsilon > 0$, such that it is* R3SAT*($2^{\mathcal{O}(n^\delta)}$)*-hard to approximate uniform-budget* UDP-MIN *(computer scientist's version) within $\mathcal{O}(|\mathcal{P}|^\varepsilon)$.*

Proof: To obtain inapproximability for instances with a bounded number of non-zero budgets per consumer we have to start from BBIS in constant degree graphs again. As shown in [2] for the independent set problem, BBIS in graphs of degree at most Δ is R3SAT*(poly(n))-hard to approximate within a factor of Δ^ε for some $\varepsilon > 0$ and all $\Delta \geq \Delta_0$, where Δ_0 is constant d from Lemma 4 (see Appendix A). We then apply our reduction as described above and obtain uniform-budget UDP-MIN (computer scientist's version) instances with $\ell = \Delta$ non-zero budgets per consumer and inapproximability within a factor of $(1/(16\mathrm{e}))\Delta^\varepsilon$, where the term $1/16\mathrm{e}$ stems from the fact that the randomized reduction from BBIS to MES might blow up small independent sets by a factor of 2 or shrink large independent sets by a factor of $1/8\mathrm{e}$. Choosing $\ell_0 \in \mathbb{N}$ sufficiently large ensures that $\ell^{\varepsilon-\delta} \leq (1/(16\mathrm{e}))\ell^\varepsilon$ for all $\ell \geq \ell_0$.

R3SAT*($2^{\mathcal{O}(n^\delta)}$)-hardness of approximation for uniform-budget UDP-MIN within $\mathcal{O}(n^\varepsilon)$ follows by choosing $f(m) = m^\varepsilon$ for arbitrary $\varepsilon > 0$. $\square$

D Single-Minded Pricing: Missing Proofs

Theorem 6 *There exist constants $\ell_0 \in \mathbb{N}$ and $\varepsilon > 0$, such that for every $\ell \geq \ell_0$ it is R3SAT*(poly(n))-hard to approximate* SMP *with sets of maximum cardinality ℓ within ℓ^ε. Furthermore, for every $\delta > 0$ there exists $\varepsilon > 0$, such that it is R3SAT*($2^{\mathcal{O}(n^\delta)}$)-hard to approximate* SMP *within $\mathcal{O}(n^\varepsilon)$.*

Proof: We only prove the second part of the theorem. Let a given MES instance $S_1, \ldots, S_m$ be separable into $\mathcal{C}_1, \ldots, \mathcal{C}_\kappa$ with $\kappa = \mathcal{O}(f(m))$. For each element e in the universe of the MES instance we have a corresponding product e. We assume that the MES instance is defined on a universe of size n and $n = m$. This is w.l.o.g. due to the reduction in Section 2. For every set S_i in class $\mathcal{C}_j$ we define a collection of $(2n)^{\kappa-j}$ identical consumers $C_i = \{c_i^1, c_i^2, \ldots, c_i^{2^{\kappa-j}}\}$. Each of these consumers has budget $b_i = (2n)^{j-\kappa}$ and wants to purchase all the products from set S_i. Note that the total number of consumer samples in this construction is bounded above by $n(2n)^{\mathcal{O}(f(n))}$.

Soundness: Let $\phi = (\phi(1) < \cdots < \phi(\ell))$ be an expanding sequence of length ℓ. For every $1 \leq i \leq \ell$ let $N_{\phi(i)}$ denote the elements that are newly covered by $S_{\phi(i)}$. Now we repeat the following for $i = 1, \ldots, \ell$. Determine $N_{\phi(i)}$ and let

$$\xi_{\phi(i)} = \sum_{e \in S_{\phi(i)} \setminus N_{\phi(i)}} p(e)$$

denote the sum of prices of previously covered elements contained in $S_{\phi(i)}$. Let $S_{\phi(i)} \in \mathcal{C}_j$. By the fact that sets from class $\mathcal{C}_j$ do not intersect and budgets increase exponentially between different classes, it follows that every $e \in S_{\phi(i)} \setminus N_{\phi(i)}$ has price $p(e) \leq b_{\phi(i)}/(2n)$ and we get that $\xi_{\phi(i)} < b_{\phi(i)}/2$. Thus, $b_{\phi(i)} - \xi_{\phi(i)} > 0$ and by setting the prices of all $e \in N_{\phi(i)}$ to

$$p(e) = \frac{b_{\phi(i)} - \xi_{\phi(i)}}{|N_{\phi(i)}|}$$

we extract revenue $|C_{\phi(i)}| b_{\phi(i)} = (2n)^{\kappa-j}(2n)^{j-\kappa} = 1$ from consumers $C_{\phi(i)}$. Consequently, overall revenue is at least ℓ.

Completeness: Assume that we are given a price assignment p resulting in overall revenue r. W.l.o.g. revenue at least $r/2$ is due to consumers buying at a price that is at least half their budget value, since otherwise we could increase overall revenue by multiplying all prices by 2.

Consider consumers C_i with product set S_i. For consumers C_k, $k < i$, with product sets S_k and $S_i \cap S_k \neq \emptyset$ it must be true that $b_k \leq b_i/(2n)$. Thus, if S_i is contained in the union of product sets of consumers with smaller indices contributing to the revenue, it follows that $p(e) \leq b_i/(2n)$ for all $e \in S_i$. Consequently, revenue from any consumer $c_i^k \in C_i$ is at most $|S_i| b_i/(2n) < b_i/2$, where we use that w.l.o.g. $|S_i| < n$ for all but the last set in any MES instance. Conversely, product sets of consumers that yield revenue equal to at least half their budget values must adhere to the expansion property and we obtain an expanding sequence ϕ of length at least $r/2$.

Finally, for any given value of $\delta > 0$ fix $\delta' < \delta$ and let $f(n) = n^{\delta'}$. We obtain MES instances of size $2^{\mathcal{O}(n^{\delta})}$, which are hard to approximate within $\mathcal{O}(n^{\varepsilon})$ for some $\varepsilon > 0$. The first part of the theorem follows similar to our argumentation in Section C. $\square$