Greedy Approximation via Duality for Packing, Combinatorial Auctions and Routing

Piotr Krysta*

Department of Computer Science, Dortmund University
Baroper Str. 301, 44221 Dortmund, Germany
piotr.krysta@cs.uni-dortmund.de

Abstract. We study simple greedy approximation algorithms for general class of integer packing problems. We provide a novel analysis based on the duality theory of linear programming. This enables to significantly improve on the approximation ratios of these greedy methods, and gives a unified analysis of greedy for many packing problems. We show matching lower bounds on the ratios of such greedy methods. Applications to some specific problems, including mechanism design for combinatorial auctions, are also shown.

1 Introduction

Combinatorial auctions (CAs) is the canonical problem motivated by applications in electronic commerce and game theoretical treatment of the Internet. A seminal paper of Lehmann et al. [24] identified a class of greedy approximation algorithms for the set packing problem as having certain monotonicity properties. These properties proved crucial in obtaining approximate non-VCG mechanisms for truthful CAs. This is one of our main motivations to study greedy algorithms.

Greedy algorithms for combinatorial optimization problems are very simple and efficient. However, their performance analysis can be difficult, and there are no general, unified tools known. We study simple greedy approximation algorithms for general integer packing problems, and provide a technique for analyzing their performance via the duality theory of linear programming. This significantly improves upon known approximation ratios of greedy methods for a class of integer packing problems.

We are not aware about any existing work of analyzing greedy approximation algorithms for integer packing problems via duality (an exception is [6] – see end of Sec. 1.1). The situation is completely different for the integer covering problems, where starting from the seminal work of Lovász [25] and Chvátal [8], such analyzes were performed and generalized [9, 29]. One of our purposes is to initiate filling this gap. In fact our technique is fairly general in that it can even be extended to some routing problems.

A class of packing integer programs [31], (PIP), reads: $\max\{cx : Ax \leq b, x \in \{0,1\}^n\}$, where A is an $m \times n$ matrix with non-negative entries, $b \in \mathsf{R}^m_{\geq 1}$, $c \in \mathsf{R}^n_{\geq 0}$. Our results can be extended to allow $x \in \{0, 1, \ldots, u\}^n$ for some $u \in \mathsf{N}$. When all entries in A are $0/1$, (PIP) is called (0,1)-PIP, and generalizes many weighted problems, e.g., maximum weighted independent sets, max-cliques, hypergraph b-matching,

* The author is supported by DFG grant Kr 2332/1-1 within "Aktionsplan Informatik" of the Emmy Noether program.

k-dimensional matching, (multi-) set (multi-) packing, edge-disjoint paths, etc. If A has non-negative entries, we can also model multicommodity unsplittable flow (UFP), and multi-dimensional knapsack problems.

We reformulate (PIP) as a generalized set packing. Let U be a set of m elements, and $\mathcal{S} \subseteq 2^U$ a family of n subsets of U. Each set $S \in \mathcal{S}$ has cost c_S, and can in fact be a multi-set: let $q(e,S) \in \mathsf{N}_{\geq 0}$ be the number of copies of e in S. Each element $e \in U$ has an upper bound $b_e \in \mathsf{N}$ on the number of times it can appear in the packing. A feasible packing is a subfamily of $\mathcal{S}$, where the total number of copies of each element $e \in U$ in all sets of the subfamily is at most b_e. The problem is to find a feasible packing with maximum total cost. These assumptions can be relaxed to $q(e,S) \in \mathsf{R}_{\geq 0}$ and $b_e \in \mathsf{R}_{\geq 1}$. In terms of (PIP), if $A = (a_{ij})$ and $S \in \mathcal{S}$, then $a_{eS} = q(e,S)$. Thus, (PIP) is now:

$$\max \quad \sum_{S \in \mathcal{S}} c_S x_S \tag{1}$$
$$\text{s.t.} \sum_{S: S \in \mathcal{S}, e \in S} q(e,S) \cdot x_S \leq b_e \;\; \forall e \in U \tag{2}$$
$$x_S \in \{0,1\} \qquad \forall S \in \mathcal{S}. \tag{3}$$

Let $b_{min} = \min\{b_e : e \in U\}$, and $d = \max\{|S| : S \in \mathcal{S}\}$, i.e., $d =$ max. number of non-zero entries in any column of A. Let $\phi = \max\{b_e/b_f : \exists S \in \mathcal{S} \text{ s.t. } e, f \in S\}$.

1.1 Our contributions in general

We present a novel analysis of greedy algorithms for general PIPs by using the duality theory of linear programming (LP). We employ dual-fitting, and many new ingredients. Two fractional, possibly infeasible, dual solutions are defined. One during the execution of the algorithm, and the second one after it stops. By the dual LP, the solutions must have high values on any set in the problem instance. To achieve this we treat one of these solutions as a *back-up*, and prove that if the first solution is not high enough, the second one is. We combine these two solutions by taking their convex combination, and prove that a suitable scaling gives a feasible dual solution. By weak duality, this combined dual solution is an upper bound on the value of an optimal primal integral solution, thus implying the approximation ratio. An interesting aspect of this analysis is that we do not lose any constant factor when the two dual solutions are combined. Thus, we emphasize here that our constant in front of the approximation ratio is precisely 1. Our analysis results in provably best possible approximation ratios for a large class of integer packing problems. We also show that it gives best possible approximation ratios in the natural class of *oblivious* greedy algorithms (cf. Sec. 2).

Our analysis gives significant improvements on the approximation ratios known for simple (monotone, cf. Sec. 3) greedy methods for general PIPs, and improves the approximation ratios for many specific packing problems by constant factors. It is quite flexible and general. The largest improvements are obtained for general PIPs, and for b-matching in hypergraphs. We also slightly improve (by constant factors) the approximation ratios for truthful CAs and routing UFP problems. An additional advantage of our analysis is that it also implies bounds on the integrality gaps of the LP relaxations.

An LP duality-based (dual-fitting) analysis was previously known for the weighted set cover problem due to Chvátal [8], and for generalized set cover by Rajagopalan & Vazirani [29]. We are not aware of such analyzes for PIPs. (Except a recent primal-dual algorithm and analysis by Briest, Krysta and Vöcking [6], but this analysis does not apply to the simple greedy algorithms that we study.)

1.2 Previous work and our improvements

To compare the previous and our results, we discuss them for (0,1)-PIP and assume $b_{min} \geq 1$. (In fact our bounds are more general – see further sections for details.)

The best approximation ratios, $O\left(\min\left\{m^{1/(b_{min}+1)}, d^{1/b_{min}}\right\}\right)$ [28, 31, 30], for (0,1)-PIP were obtained by solving an LP relaxation and performing randomized rounding of the fractional solution. They are better than the ratios of combinatorial methods, but the main drawback is the need of solving LPs. This is inefficient in many cases, and also does not guarantee monotonicity properties (cf. Sec. 3). (A monotone randomized rounding algorithm is known [2], but it applies to restricted bidders and supplies of goods, and is truthful only in a probabilistic sense – see Sec. 3 for the definitions.)

Our main results. Lehmann et al. [24] have analyzed the following simple greedy algorithm for (0,1)-PIP assuming $b_e = 1$ for all $e \in U$.

$\mathcal{P}$:= $\emptyset$; `let` $\mathcal{S} = \{S_1, S_2, \ldots, S_n\}$ s.t. $\frac{c_{S_1}}{\sqrt{|S_1|}} \geq \frac{c_{S_2}}{\sqrt{|S_2|}} \geq \cdots \geq \frac{c_{S_n}}{\sqrt{|S_n|}}$
`for` $S' = S_1, S_2, \ldots, S_n$ `do if` $\mathcal{P} \cup \{S'\}$ `fulfills` (2) & (3) `then` $\mathcal{P} := \mathcal{P} \cup \{S'\}$
`output packing` $\mathcal{P}$

(A more general version of this algorithm, called Greedy-2, is given in (14).) Lehmann et al. proved that it gives a $\sqrt{m}$-approximation for (0,1)-PIP with all $b_e = 1$. Gonen & Lehmann [10] have shown that Greedy-2 is a $\sqrt{\sum_{e \in U} b_e}$-approximation algorithm for (0,1)-PIP. We use our duality-based analysis to improve this ratio significantly, showing that Greedy-2 is a $(\sqrt{\sum_{e \in U} b_e / b_{min}} + 1)$-approximation algorithm for (0,1)-PIP. This, for instance, implies a ratio of $(\sqrt{m} + 1)$ for (0,1)-PIP when all $b_e = b$ for some value b, which is not necessarily one. We also show a corresponding lower bound, by proving that this ratio is essentially best possible for this problem in the class of oblivious greedy algorithms – which basically captures all natural greedy algorithms for (0,1)-PIP.

We give another greedy algorithm for (0,1)-PIP, called Greedy-1, and show it is a $(\sqrt{\phi m} + 1)$-approximation using our duality-based technique. A third presented algorithm, Greedy-3, is a $(d+1)$-approximation for (0,1)-PIP. Obtaining a ratio better than $\sqrt{m}$ (even if $\phi = 1$) is not possible, unless $\mathsf{NP} = \mathsf{ZPP}$ [17], and it is NP-hard to obtain a ratio of $O(d/\log d)$ [18]. Thus, our analysis implies essentially best possible ratios.

It is possible to modify a combinatorial greedy algorithm for the unsplittable flow problem, presented by Kolman and Scheideler [23], to obtain an $O(\sqrt{m})$-approximation algorithm for (0,1)-PIP. However, our greedy algorithms and analysis have the following advantages over that modified algorithm: our algorithm is monotone, which is needed for mechanism design (cf. Sec. 3); our duality-based analysis implies bounds on the integrality gaps of PIPs; and, finally, we do not lose constant factors in the ratio.

The integrality gaps of LP relaxations for (0,1)-PIPs were proved before by Aharoni et al. [1] (for the unweighted (0,1)-PIP), Raghavan [27], and Srinivasan [31]. Our duality-based analysis improves these bounds by constant factors.

Further known results vs ours. Some combinatorial algorithms are known for (0,1)-PIP. E.g., an algorithm of Bartal, Gonen and Nisan [4], following ideas from [3]. If

$b_e = b$, $\forall e \in U$, the algorithm of [4] achieves the best ratio of $O(b \cdot (m)^{1/(b-1)})$ for (0,1)-PIP. A very recent result is a primal-dual $O(m^{1/(b+1)})$-approximation algorithm for (0,1)-PIP, by Briest, Krysta and Vöcking [6]. Our contribution here is the greedy $(\sqrt{m}+1)$-approximate algorithm, which is faster and much simpler than the other two combinatorial algorithms in [4, 6], and tightens the big-O constants.

An additional motivation for simple greedy method here comes from the branch-and-bound heuristics for CAs. Gonen and Lehmann [10] proved that algorithm Greedy-2 (see (14)) gives the best method of ordering the bids in the branch and bound heuristics for CAs. Their experiments [11] support this claim in practice. Our improved ratio for Greedy-2 might be a step towards theoretical explanation of this good performance.

Combinatorial approximation algorithms are known for special problems modeled by PIPs. Halldórsson et al. [13] gave a greedy $b\sqrt{m}$-approximation for unweighted b-matching in hypergraphs (all $c_S = 1$ and all $b_e = b$). For the same weighted problem [10] gives a greedy $\sqrt{bm}$-approximation. Thus, we improve the ratio of a greedy approximation for the problem to $\sqrt{m} + 1$. A simple greedy $\sqrt{m}$-approximation for unweighted set packing ($b = 1$) [13], and a $2\sqrt{m}$-approximation for weighted set packing [15] are known. Our ratio $\sqrt{m} + 1$ applies here as well. For weighted set packing ($c_S \geq 0$; all $b_e = 1$), Hochbaum [19] gave a greedy d-approximation, and Berman & Krysta [5] show a local search $\frac{2}{3}d$-approximation. Our ratio $d + 1$ applies to a more general problem. See survey [14] for other results on related approximations.

Further consequences of our analysis. Our results can be applied to obtain truthful mechanisms for combinatorial auctions with slightly improved approximation factors. This is discussed in detail in Section 3.

A related problem is the routing unsplittable flow problem (UFP). We can cast UFP by slightly generalizing PIP, but in fact UFP is less general than PIPs. There are greedy and other combinatorial approximation algorithms for UFP and related problems, e.g., [20, 12, 7, 23, 22, 3]. Their ratios, usually, look like $\sqrt{m}$, sometimes with additional logarithmic factors or factors depending on capacities and demands, with m denoting the number of edges in the graph. Our duality-based analysis can be extended to these problems. See Section 4 for the definitions and details on our improvements.

2 Oblivious greedy algorithms

We study a class of algorithms for (PIP) that we call *oblivious greedy algorithms*. This class was studied in context of truthful CAs in [24]. Besides their simplicity, truthful CAs are the main motivation to study these algorithms. The crucial feature of this class is that of *monotonicity*, which implies a truthful mechanism for CAs (cf. Section 3).

An *oblivious greedy algorithm* for (PIP) is a polynomial time algorithm $\mathcal{A}$ with a rank function, say $\rho : \mathcal{S} \times \mathsf{R}_{\geq 0} \longrightarrow \mathsf{R}_{\geq 0}$, assigning a real number $\rho(S, c_S)$ to each pair (S, c_S), $S \in \mathcal{S}$. We assume, after [24], that having any $S \in \mathcal{S}$ fixed, $\rho(S, x)$ is strictly increasing as a function of x. Having $\rho(\cdot, \cdot)$, $\mathcal{A}$ sorts all the sets in $\mathcal{S}$ w.r.t. non-increasing numbers $\rho(S, c_S)$. Then $\mathcal{A}$ scans the sets in this order once, and picks them to the solution one by one, maintaining feasibility.

The monotonicity of $\rho(S, \cdot)$ is a natural assumption for CAs: Let a buyer offer some amount of money for a product, so that the seller wants to sell the product for that money. If, now, the buyer offers even more money, the seller, obviously, is also willing to sell the same product. We will write $\rho(S)$ for short, instead of $\rho(S, c_S)$.

2.1 The upper bound

We give a greedy algorithm, Greedy-1, for (PIP), having the best possible approximation ratio for (0,1)-PIP when $b = \min_e b_e$ is small. Greedy-1 will also be shown to have the best possible ratio in the class of all oblivious greedy algorithms for (0,1)-PIP.
Generic greedy algorithm and dual-fitting analysis. We can assume w.l.o.g., that given $S \in \mathcal{S}$, we have $q(e, S) \leq b_e$ for each $e \in S$. Let, for any $S \in \mathcal{S}$, a rank value $\rho(S) \in \mathsf{R}_{\geq 0}$ be given. Let $\mathcal{P} \subseteq \mathcal{S}$ be a given packing. We say that element $e \in U$ is *saturated* or *sat* w.r.t. $\mathcal{P}$ if there is a set $S' \in \mathcal{S} \setminus \mathcal{P}$ with $e \in S'$, such that $q(e, S') + \sum_{S:S\in\mathcal{P},e\in S} q(e, S) > b_e$. Our generic greedy algorithm, Greedy, is then:

01. $\mathcal{P} := \emptyset$
02. `let` $\mathcal{S} = \{S_1, S_2, \ldots, S_n\}$ s.t. $\rho(S_1) \geq \rho(S_2) \geq \cdots \geq \rho(S_n)$
03. `for` $S' = S_1, S_2, \ldots, S_n$ `do`
04. `if` $\mathcal{P} \cup \{S'\}$ `fulfills` (2) and (3) `then`
05. $\mathcal{P} := \mathcal{P} \cup \{S'\}$
06. `output packing` $\mathcal{P}$

We now present our analysis via dual fitting. Our original primal problem is given by the integer linear program (1)–(3). The LP relaxation of this integer program is:

$$\max \quad \sum_{S\in\mathcal{S}} c_S x_S \tag{4}$$
$$\text{s.t.} \sum_{S:S\in\mathcal{S},e\in S} q(e, S)x_S \leq b_e \ \forall e \in U \tag{5}$$
$$0 \leq x_S \leq 1 \qquad \forall S \in \mathcal{S}. \tag{6}$$

Its corresponding dual linear program can be written as:

$$\min \quad \sum_{e\in U} b_e y_e + \sum_{S\in\mathcal{S}} z_S \tag{7}$$
$$\text{s.t. } z_S + \sum_{e\in S} q(e, S)y_e \geq c_S \ \forall S \in \mathcal{S} \tag{8}$$
$$z_S, y_e \geq 0 \qquad \forall S \in \mathcal{S}, e \in U. \tag{9}$$

Dual variable z_S corresponds to $x_S \leq 1$. We present below a performance proof of a greedy with a specific rank function ρ_1, which will be used as *generic proof.*

$$\text{Given a set } S \in \mathcal{S}, \text{ let: } \rho_1(S) = \frac{c_S}{\sqrt{\sum_{e\in S} \frac{q(e,S)}{b_e}}} \ .$$

We call algorithm Greedy with the rank function $\rho = \rho_1$, Greedy-1.

Theorem 1. *Algorithm* Greedy-*1 has an approximation ratio of* $\frac{q_{max}}{q_{min}}\sqrt{\phi m} + 1$ *for the generalized set packing problem, and for* (*PIP*)*, assuming* $\phi = \max\{b_e/b_f : e, f \in S, S \in \mathcal{S}\}$*, and for each* $S \in \mathcal{S}$*,* $e \in S$*, we have* $q_{min} \leq q(e, S) \leq q_{max}$ *or* $q(e, S) = 0$*.*

Due to the generality of (PIP), there is no better approximation ratio than $\sqrt{m}$, unless NP $=$ ZPP [17]. This hardness holds even when $\frac{q_{max}}{q_{min}} = 1$, and $\phi = 1$.

Proof. (Theorem 1) This is a sketch of the **generic proof**. Suppose Greedy-1 terminated and output solution $\mathcal{P}$. Let $SAT_{\mathcal{P}} = \{e \in U : e \text{ is sat w.r.t. } \mathcal{P}\}$. Notice, for each set $S \in \mathcal{S} \setminus \mathcal{P}$, there is an $e \in SAT_{\mathcal{P}}$, called a *witness*: That is, when S was considered in line 04, and $\mathcal{P}$ was the current (partial) solution, $e \in S$ was an element such that $q(e,S) + \sum_{S':S' \in \mathcal{P}, e \in S'} q(e,S') > b_e$. For each set $S \in \mathcal{S} \setminus \mathcal{P}$ we keep in $SAT_{\mathcal{P}}$ one (arbitrary) witness for S, and discard the remaining elements from $SAT_{\mathcal{P}}$.
Defining two dual solutions. We define two fractional dual solutions, y^1 and y^2. y^1 is defined after Greedy-1 has terminated. Let us define the following:

$$\sigma = \sum_{S \in \mathcal{P}} \frac{c_S \cdot q_S}{\sqrt{\sum_{e' \in \mathcal{P}(S)} \frac{q(e',S) \cdot \max_{e'' \in \mathcal{P}(S)} \{b_{e''}\}}{b_{e'}}}} \quad , \quad \text{where}$$

$\mathcal{P}(S) = S \cap SAT_{\mathcal{P}}$ if $S \cap SAT_{\mathcal{P}} \neq \emptyset$ and $\mathcal{P}(S) = S$ if $S \cap SAT_{\mathcal{P}} = \emptyset$; and $q_S = \sqrt{|\mathcal{P}(S)|}$. For each $e \in U$, define y^1 as: $y^1_e = \frac{\sigma}{b_e \cdot m}$. Solution y^2 is defined during the execution of Greedy-1. We need to know $\mathcal{P}$ to define y^2, which is needed only for analysis. In line 01 of Greedy-1 we initialize: $y^2_e := 0, z^2_S := 0$, for all $e \in U$, and $S \in \mathcal{S}$. The following is added in line 05 of Greedy-1:

$$y^2_e := y^2_e + \Delta^{S'}_e \quad , \text{ for all } e \in \mathcal{P}(S') \quad , \quad \text{where}$$

$$\Delta^{S'}_e = \frac{c_{S'}}{b_e \cdot q_{S'} \cdot \sqrt{\sum_{e' \in \mathcal{P}(S')} \frac{q(e',S') \max_{e'' \in \mathcal{P}(S')} \{b_{e''}\}}{b_{e'}}}} \quad , \text{ for } e \in \mathcal{P}(S') \quad .$$

Note, that for $e \in S' \setminus SAT_{\mathcal{P}}$ the value of y^2_e is not updated and remains zero. We also add the following in line 05 of Greedy-1: $z^2_{S'} := z^2_{S'} + c_{S'}$. Obviously, if for an $e \in U$, values of y^1, y^2 or z^1, z^2 have not been defined, they are zero.
Dual lower bound on the solution. We now argue that both dual solutions (appropriately scaled) provide lower bounds on the cost of the output solution:

$$\sum_{e \in U} b_e \cdot y^i_e \leq \frac{1}{\sqrt{q_{min}}} \sum_{S \in \mathcal{P}} c_S, \quad \text{for} \quad i = 1, 2 \quad . \tag{10}$$

Final dual solution. We will now show that there exists a dual solution y such that the scaled solution $(\frac{q_{max}}{\sqrt{q_{min}}} \sqrt{\phi m} \cdot y, z)$ is feasible for the dual LP, i.e., constraints (8) are fulfilled for all sets $S \in \mathcal{S}$. Thus, we have to show that, for *each* set $S \in \mathcal{S}$,

$$z_S + \frac{q_{max}}{\sqrt{q_{min}}} \sqrt{\phi m} \sum_{e \in S} q(e,S) y_e \geq c_S \quad . \tag{11}$$

We will prove (11) by using both y^1 and y^2. The main idea is to use the second solution as a *back-up*: whenever on some set $S \in \mathcal{S}$, y^1 is not high enough we will prove that y^2 is sufficiently high. We define a new solution y as a convex combination of y^1, y^2: $y_e = \frac{1}{2}\left(y^1_e + y^2_e\right)$, for each $e \in U$; also, define z as $z_S = z^2_S$ for each $S \in \mathcal{S}$.
Proving (11). Suppose first that $S \in \mathcal{S} \setminus \mathcal{P}$. The reason that S' has not been included in solution $\mathcal{P}$ must be an $e \in SAT_{\mathcal{P}}$ such that $e \in S'$. This means that adding set S' to solution $\mathcal{P}$ would violate constraint (2). Let $\mathcal{E} \subseteq \mathcal{P}$ be the family of all sets in the solution that contain element e. Observe, that $|\mathcal{E}| \geq 1$.

Lower-bounding y^1. Using our greedy selection rule, and lower-bounding appropriately σ, we are able to show the following bound:

$$\sum_{e' \in S'} q(e', S') y^1_{e'} \geq \frac{c_{S'}}{m\gamma} \cdot \left(\sum_{S \in \mathcal{E}} q_S \right) \cdot \sqrt{\sum_{e' \in S'} \frac{\gamma q(e', S')}{b_{e'}}} \quad . \tag{12}$$

Lower-bounding y^2. Again using our greedy selection rule, a well known inequality between the arithmetic and harmonic means [16], we can lower-bound appropriately parameters Δ to obtain ($|\mathcal{E}| = p$):

$$\sum_{e' \in S'} q(e', S') y^2_{e'} \geq \frac{c_{S'}}{b_e} \frac{q_{min} p^2}{\left(\sum_{S \in \mathcal{E}} q_S\right) \sqrt{\sum_{e' \in S'} \frac{\gamma q(e', S')}{b_{e'}}}} \quad . \tag{13}$$

Lower-bounding y. Using (12) and (13) we can write

$$\sum_{e' \in S'} q(e', S') y_{e'} = \frac{1}{2} \sum_{e' \in S'} \left(q(e', S') y^1_{e'} + q(e', S') y^2_{e'} \right) \geq$$

$$\frac{1}{2} \left(\frac{c_{S'}}{m\gamma} \cdot \left(\sum_{S \in \mathcal{E}} q_S \right) \cdot \sqrt{\sum_{e' \in S'} \frac{\gamma q(e', S')}{b_{e'}}} + \frac{c_{S'} q_{min} p^2}{b_e \left(\sum_{S \in \mathcal{E}} q_S\right) \sqrt{\sum_{e' \in S'} \frac{\gamma q(e', S')}{b_{e'}}}} \right) =$$

$$\frac{c_{S'}}{2} \left(\frac{x}{m\gamma} + \frac{q_{min} p^2}{b_e x} \right), \quad \text{where} \quad x = \left(\sum_{S \in \mathcal{E}} q_S \right) \cdot \sqrt{\sum_{e' \in S'} \frac{\gamma q(e', S')}{b_{e'}}} \quad .$$

Consider now function $f(x) = \frac{x}{m\gamma} + \frac{q_{min} p^2}{b_e x}$ for $x > 0$. We can show that $f(x) \geq 2\sqrt{\frac{q_{min} p^2}{\gamma m b_e}}$ for all $x \in (0, \infty]$. Observing that $|\mathcal{E}| = p \geq \frac{b_e}{q_{max}}$, and $\frac{\gamma}{b_e} \leq \phi$, we obtain $f(x) \geq 2\sqrt{\frac{q_{min}}{q^2_{max} \phi m}}$ for all $x \in (0, \infty]$. This proves claim (11) when set $S' \in \mathcal{S} \setminus \mathcal{P}$. If $S' \in \mathcal{P}$, then claim (11) follows from the definition of $z_{S'}$.

Finishing the proof. We have shown that the dual solution $(\frac{q_{max}}{\sqrt{q_{min}}} \sqrt{\phi m} \cdot y, z)$ is feasible for the dual linear program and so by weak duality $\sum_{S \in \mathcal{S}} z_S + \frac{q_{max}}{\sqrt{q_{min}}} \sqrt{\phi m} \cdot \sum_{e \in U} b_e y_e$ is an upper bound on the value of the optimal integral solution to our problem. By (10), we have that $\sum_{e \in U} b_e y_e \leq \frac{1}{\sqrt{q_{min}}} \sum_{S \in \mathcal{P}} c_S$. Thus, we obtain that

$$opt \leq \sum_{S \in \mathcal{S}} z_S + \frac{q_{max}}{\sqrt{q_{min}}} \sqrt{\phi m} \cdot \sum_{e \in U} b_e y_e = \sum_{S \in \mathcal{P}} z_S + \frac{q_{max}}{\sqrt{q_{min}}} \sqrt{\phi m} \cdot \sum_{e \in U} b_e y_e$$

$$\leq \sum_{S \in \mathcal{P}} c_S + \frac{q_{max}}{q_{min}} \sqrt{\phi m} \cdot \left(\sum_{S \in \mathcal{P}} c_S \right) = \left(\frac{q_{max}}{q_{min}} \sqrt{\phi m} + 1 \right) \cdot \left(\sum_{S \in \mathcal{P}} c_S \right) \quad . \qquad \square$$

Other greedy selection rules. The next two theorems – Theorem 2 and 3 can easily be shown by using our generic proof (of Theorem 1).

$$\text{Given a set } \; S \in \mathcal{S}, \quad \text{let:} \quad \rho_2(S) = \frac{c_S}{\sqrt{\sum_{e \in S} q(e, S)}}. \tag{14}$$

Greedy-2 is Greedy with the greedy selection rule ρ_2.

Theorem 2. *Algorithm* Greedy-*2 has an approximation ratio of* $\frac{q_{max}}{q_{min}}\sqrt{\frac{\sum_{e\in U} b_e}{b_{min}}}+1$ *for the generalized set packing problem, and for* (*PIP*)*, assuming that* $b_{min} = \min\{b_e : e \in U\}$*, and for each* $S \in \mathcal{S}$*,* $e \in S$*, we have* $q_{min} \leq q(e,S) \leq q_{max}$ *or* $q(e,S) = 0$*.*

Suppose that $q_{max}/q_{min} = 1$. We improve the best known approximation ratio for a combinatorial (greedy) algorithm for (PIP) from $\sqrt{\sum_{e\in U} b_e/q_{min}}$ (Gonen & Lehmann [10]) to $\min\left\{\sqrt{\sum_{e\in U} b_e/b_{min}}+1, \sqrt{\phi m}+1\right\}$ (Theorems 1, 2). This ratio is always better than [10], since w.l.o.g. $q_{min} \leq 1$ and $b_{min} \geq 1$ (see Srinivasan [31]). Also if $\phi = 1$, then our ratio is $O\left(\sqrt{m}\right)$, and that of [10] is still $\sqrt{\sum_{e\in U} b_e}$ (for $q_{min} = 1$).

$$\text{Given a set} \quad S \in \mathcal{S}, \quad \text{let:} \quad \rho_3(S) = c_S/|S| \ .$$

We call algorithm Greedy with the greedy selection rule ρ_3, Greedy-3.

Theorem 3. *Algorithm* Greedy-*3 has an approximation ratio of* $\frac{q_{max}}{q_{min}}d+1$ *for the generalized set packing problem, and for* (*PIP*)*, assuming that* $|S| \leq d$ *for each* $S \in \mathcal{S}$*. Moreover, for each* $S \in \mathcal{S}$*,* $e \in S$*, we have* $q_{min} \leq q(e,S) \leq q_{max}$ *or* $q(e,S) = 0$*.*

Observe that the approximation ratio is close to best possible. Let $\frac{q_{max}}{q_{min}} = 1$ and all $q(e,S) \in \{0,1\}$; then our ratio is $d+1$. The considered PIP can express the unweighted set packing problem for which obtaining a ratio of $O(\frac{d}{\log d})$ is NP-hard [18].

2.2 The lower bound

In this section we consider (0,1)-PIP. By Theorem 1 we obtain the following.

Corollary 1. *Algorithm* Greedy-*1 is an oblivious greedy* $(\sqrt{\phi m}+1)$*-approximation algorithm for the* (*0,1*)*-PIP problem, where* $\phi = \max\{b_e/b_f : e,f \in S, S \in \mathcal{S}\}$*.*

We show below, by modifying an argument in [10], that this upper bound can be matched by a lower bound in the class of oblivious greedy algorithms.

Proposition 1. *Let us consider* (*0,1*)*-PIP problem, assuming that* $b_e = b \in \mathsf{N}_{\geq 1}$ *for each* $e \in U$*. Then any oblivious greedy polynomial time algorithm for this problem has an approximation ratio of at least* $\sqrt{m} - \varepsilon$*, for any* $\varepsilon > 0$*.*

Note, that this lower bound above is meaningful, since there is a polynomial time $O(m^{1/(b+1)})$-approximation to the described problem via LP randomized rounding.

3 An application to truthful combinatorial auctions

We will use now our results to give truthful approximation mechanisms for combinatorial auctions (CAs) with single-minded bidders. A seller (auctioneer) wants to sell m kinds of goods U, to n potential customers (bidders). A good $e \in U$ is available in $b_e \in \mathsf{N}_{\geq 1}$ units (supply). Suppose each bidder j can valuate subsets of goods: a valuation $v_j(S) \in \mathsf{R}_{\geq 0}$ for a subset $S \subseteq U$ means the maximum amount of money j is willing to pay for getting S. For simplicity, assume that bidders can bid for 0 or single unit of a good, i.e., $q(e, S) \in \{0, 1\}$ for all $e \in U$, and $S \subseteq U$. An allocation of goods to bidders is a packing $S^1, \ldots, S^n \subseteq U$ w.r.t. the defined (PIP) with $\mathcal{S} = 2^U$, i.e., bidder j gets set S^j, and e appears at most b_e times in $S^1, \ldots, S^n$. The objective is to find an allocation with maximum *social welfare*, $\sum_j v_j(S^j)$. Each v_j is only known to bidder j. Our bidders are *single-minded* [24], i.e., for each bidder j there exists a set $S_j \subseteq U$ she prefers and a $v_j^* \geq 0$, such that $v_j(S) = v_j^*$ if $S_j \subseteq S$ and $v_j(S) = 0$ otherwise.

An *auction mechanism* (seller) is an algorithm which first collects the bids from the bidders, i.e., their declarations (S'_j, v'_j), where S'_j is supposed to mean the preferred set S_j, and v'_j the valuation v_j for j. The mechanism then determines the allocation and a payment p_j for each bidder j. The utility of bidder j is $u_j = v_j(S) - p_j$ for winning set S. Note that v_j is the true valuation function. We assume the mechanism is *normalized*, i.e., $p_j = 0$, when bidder j is not allocated any set. Our allocation problem is to maximize the social welfare; this corresponds to approximating our (PIP) problem.

Each bidder aims at maximizing her own utility. It may be profitable for bidder j to lie and report $v'_j \neq v_j$ and $S'_j \neq S_j$ to increase her utility. A mechanism is *truthful* (*incentive compatible*) if declaring truth, i.e., $v'_j = v_j$ and $S'_j = S_j$, is a *dominant* strategy for each bidder j. That is for any fixed set of bids of all bidders except j, if j does not declare the truth, then this may only decrease j's utility. Our goal is a truthful approximate mechanism.

An allocation algorithm is *monotone* if whenever bidder j declares (S'_j, v'_j) (given other bidders' declarations) and wins, i.e., gets set S'_j allocated, then declaring (S''_j, v''_j) s.t. $S''_j \subseteq S'_j, v'_j \leq v''_j$ results also in winning set S''_j. It is well known that if an allocation (approximate) algorithm is monotone and exact (i.e., a bidder gets exactly her declared set or nothing), then there is a payment scheme which together with the allocation algorithm is a truthful (approximate) mechanism (see, e.g., [26, 24]). It is easy to see that all our greedy algorithms are monotone and exact. We can also modify the payment scheme in [24] to serve our purposes. Thus, using Theorems 1, 2 and 3, this gives the following result.

Theorem 4. *Suppose we have m kinds of goods U, each good $e \in U$ available in $b_e \in \mathsf{N}_{\geq 1}$ units, and $b_{min} = \min\{b_e : e \in U\}$. Suppose each bidder bids only on at most $d \in \mathsf{N}_{\geq 1}$ goods, and for each bid (S'_j, v'_j) we have $\max\{b_e/b_f : e, f \in S'_j\} \leq \phi$. There is a truthful mechanism for CAs with single-minded bidders with an approximation ratio*

$$\min\left\{1 + \sqrt{\phi m},\ \ 1 + \sqrt{\sum_{e \in U} b_e/b_{min}},\ \ 1 + d\right\}.$$

Note, that we assume d is known to the mechanism. Theorem 4 improves on the mechanism of Lehmann et al. [24], where they assume $b_e = 1$ for each $e \in U$, and their ratio is $\sqrt{m}$. We achieve the same ratio $(+1)$ for a more general setting, where the supplies of the goods are given arbitrary numbers. The best known truthful mechanism for the problem with $b_e = b\ \forall e \in U$, is $5.44 \cdot (m)^{1/b}$-approximate, see Briest, Krysta and Vöcking [6]. Thus, our ratio for the same problem is slightly better for $b = 2$. Our ratio is also very good if d is small – a very natural assumption for bidders.

4 An application to the unsplittable flow problem

We show that our dual fitting analysis can be extended to deal with a difficult routing problem – multicommodity unsplittable flow problem (UFP).

Let $G = (V, E)$ be a given graph ($|E| = m$), and $C = \{(s_i, t_i) : s_i, t_i \in V, i = 1, \dots, k\}$ be k source-sink pairs, or commodities. For each commodity $i \in \{1, \dots, k\}$, we are given a demand $d_i \in \mathsf{N}_{\geq 1}$, and a profit $p_i \in \mathsf{R}_{\geq 0}$. For each edge $e \in E$, $b_e \in \mathsf{N}_{\geq 0}$ denotes its capacity. Given $i \in \{1, \dots, k\}$, let C_i be the set of all simple s_i-t_i-paths in G, such that all edge capacities on these paths are at least d_i. The multicommodity unsplittable flow problem (UFP) is to route for each commodity i demand d_i along a single path in C_i, respecting edge capacities. The objective is to maximize the sum of the profits of all commodities that can be simultaneously routed.

We give now an LP relaxation of this problem (see Guruswami et al. [12]). The ground set is $U = E$, and the set family is $\mathcal{S} = \cup_{i=1}^k C_i$. Each set $S \in \mathcal{S} \cap C_i$, that is an s_i-t_i-path, has cost $c_S = c_i = p_i$. The LP relaxation of UFP is:

$$\max \ \sum_{i=1}^k c_i \cdot \left(\sum_{S \in C_i} x_S\right) \tag{15}$$

$$\text{s.t.} \sum_{S: S \in \mathcal{S}, e \in S} d_S x_S \leq b_e \ \forall e \in U \tag{16}$$

$$\sum_{S \in C_i} x_S \leq 1 \qquad \forall i \in \{1, \dots, k\} \tag{17}$$

$$x_S \geq 0 \qquad \forall S \in \mathcal{S}, \tag{18}$$

where $d_S = d_i$ iff $S \in C_i$. The corresponding dual linear program reads then:

$$\min \sum_{e \in U} b_e y_e + \sum_{i=1}^k z_i \tag{19}$$

$$\text{s.t.} \ \ z_i + \sum_{e \in S} d_i y_e \geq c_i \ \ \forall i \in \{1, \dots, k\} \ \ \forall S \in C_i \tag{20}$$

$$z_i, y_e \geq 0 \qquad \forall i \in \{1, \dots, k\} \ \ \forall e \in U. \tag{21}$$

In this dual linear program, dual variable z_i corresponds to the constraint (17).

We will use algorithm Greedy-1 to approximate UFP. We will say how to realize it in the case of UFP. Given a commodity i, the greedy selection rule is:

$$\max_{S \in C_i} \rho_1(S) = \max_{S \in C_i} \frac{c_i}{\sqrt{\sum_{e \in S} \frac{d_i}{b_e}}} \ . \tag{22}$$

We replaced $q(e, S)$ with d_i, since $S \in C_i$. (22) is same as $\min_{S \in C_i} \left(\sum_{e \in S} 1/b_e\right) d_i / c_i^2$, and such an $S \in C_i$ can be found by a shortest path computation. The implementation of Greedy-1 is as follows. We maintain the current edge capacities b'_e. We declare all commodities *unsatisfied*, and put $b'_e := b_e$ for each $e \in E$. Perform the following until all commodities are *satisfied*. Find a commodity i and $S \in C_i$ that minimize $\left(\sum_{e \in S} 1/b_e\right) d_i / c_i^2$ among all unsatisfied commodities: when computing the shortest path for i use only edges e with $d_i \leq b'_e$. If s_i and t_i are disconnected by such edges, then declare i satisfied. Otherwise, let i_0 and $S_0 \in C_{i_0}$ be the shortest path and commodity. Route demand d_{i_0} along S_0, put $b'_e := b_e - d_{i_0}$ for all $e \in S_0$, and declare i_0 satisfied. (We define dual variable $z_{i_0} := c_{i_0}$ as in our generic proof.)

The definitions of other dual variables remain the same as in the generic proof (proof of Theorem 1). This proof basically goes through. Observe, that now $q_{min} = \min\{d_i :$

$d_i > 0,\ i = 1, \ldots, k\} = d_{min}$ and $q_{max} = \max\{d_i : i = 1, \ldots, k\} = d_{max}$. Thus, by employing the generic proof we obtain the following theorem for the UFP problem.

Theorem 5. *There is a greedy* $(\frac{d_{max}}{d_{min}}\sqrt{\phi m} + 1)$*-approximation algorithm for the unsplittable flow problem (UFP), where* $\phi = \max\{b_e/b_f : \exists i\ \exists S \in C_i \text{ s.t. } e, f \in S\}$.

Guruswami et al. [12] prove that a repeated use of a similar greedy algorithm gives a $(\frac{2d_{max}}{d_{min}}\sqrt{m})$-approximation for unweighted UFP (i.e., $c_i = 1, \forall i$). We now present some applications of Theorem 5 but the proofs are omitted, due to the lack of space.

Theorem 5 gives a $(\sqrt{m}+1)$-approximation when specialized to the weighted edge-disjoint paths (EDP) problem: this is just UFP problem with $b_e = 1$ for each $e \in E$, and $d_i = 1$ for each $i \in \{1, \ldots, k\}$. Our greedy when specialized to this case, reduces to the greedy algorithm of Kolliopoulos & Stein [21]. They prove a similar approximation ratio for this algorithm for the unweighted EDP problem. There is also a greedy $O(\min\{\sqrt{m}, |V|^{2/3}\})$-approximation for the same problem by Chekuri and Khanna [7]. Such a result is not possible for (0,1)-PIP, since obtaining ratio $O(d/\log d)$ is NP-hard [18]; note that d in (0,1)-PIP corresponds to $|V|$. Notice, that our $(\sqrt{m}+1)$-approximation holds also for the weighted and more general (than EDP) problem where we allow each edge to be used up to b times, where $b_e = b$.

The previous best known algorithm for general UFP, assuming $d_{max} \leq b_{min}$, is a combinatorial $32\sqrt{m}$-approximation algorithm of Azar and Regev [3]. If $\phi = 1$, then Theorem 5 implies a combinatorial $(2 + \epsilon)\sqrt{m}$-approximation to the UFP, assuming $d_{max} \leq b_{min}$, for any $\epsilon > 0$. On uniform capacity networks (a stronger assumption than $\phi = 1$), there is a combinatorial $O(\min\{\sqrt{m}, |V|^{2/3}\})$-approximation [7].

Another application is a $(\sqrt{|V|} + 1)$-approximation algorithm for the weighted vertex-disjoint paths problem, improving on the previous best known ratio of roughly $18\sqrt{|V|}$ due to Kolliopoulos & Stein [21]. We get this result by using ideas from [21] and our algorithm from Theorem 5 as a subroutine.

Note, that results in [22, 23] hold, unlike ours, for the unweighted UFP, i.e., unit profits problem. We would like to point out that we are able to extend our analysis to UFP thanks to the use of LP duality theory. For instance, it seems difficult to extend the combinatorial analysis in [10] to such problems.

References

1. R. Aharoni, P. Erdös and N. Linial. Optima of dual linear programs. *Combinatorica*, **8**, pp. 13–20, 1988.
2. A. Archer, C.H. Papadimitriou, K. Talwar and É. Tardos. An approximate truthful mechanism for combinatorial auctions with single parameter agents. In the *Proc. 14th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2003.
3. Y. Azar and O. Regev. Strongly polynomial algorithms for the unsplittable flow problem. In the *Proc. 8th Conference on Integer Programming and Combinatorial Optimization (IPCO)*, Springer LNCS, 2001.
4. Y. Bartal, R. Gonen, and N. Nisan. Incentive Compatible Multi-Unit Combinatorial Auctions. In the *Proc. 9th conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, Bloomington, IN, USA, June, 2003.
5. P. Berman and P. Krysta. Optimizing misdirection. In the *Proc. 14th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 192–201, 2003.

6. P. Briest, P. Krysta and B. Vöcking. Approximation Techniques for Utilitarian Mechanism Design. In the *Proc. 37th ACM Symposium on Theory of Computing (STOC)*, 2005.
7. C. Chekuri and S. Khanna. Edge-disjoint paths revisited. In the *Proc. 14th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2003.
8. V. Chvátal. A greedy heuristic for the set-covering problem. *Mathematics of Operations Research*, **4**: 233–235, 1979.
9. G. Dobson. Worst-case analysis of greedy heuristics for integer programming with non-negative data. *Mathematics of Operations Research*, **7**, pp. 515–531, 1982.
10. R. Gonen, D.J. Lehmann. Optimal solutions for multi-unit combinatorial auctions: branch and bound heuristics. In *Proc. 2nd ACM Conference on Electronic Commerce (EC)*, 2000.
11. R. Gonen, D.J. Lehmann. Linear Programming Helps Solve Large Multi-Unit Combinatorial Auctions. In the *Proceedings of INFORMS 2001*, November 2001.
12. V. Guruswami, S. Khanna, R. Rajagopalan, B. Shepherd, and M. Yannakakis. Near-optimal hardness results and approximation algorithms for edge-disjoint paths and related problems. In the *Proc. 31st ACM Symposium on Theory of Computing (STOC)*, 1999.
13. M.M. Halldórsson, J. Kratochvíl, and J.A. Telle. Independent sets with domination constraints. In the *Proc. ICALP*. Springer LNCS, 1998.
14. M.M. Halldórsson. A survey on independent set approximations. In the *Proc. APPROX*, Springer LNCS, **1444**, pp. 1–14, 1998.
15. M.M. Halldórsson. Approximations of Weighted Independent Set and Hereditary Subset Problems. *Journal of Graph Algorithms and Applications*, **4**(1), pp. 1–16, 2000.
16. G. Hardy, J.E. Littlewood, G. Polya. *Inequalities.* 2nd Edition, Cambridge Univ. Press, 1997.
17. J. Hastad. Clique is hard to approximate within $n^{1-\varepsilon}$. In the *Proc. IEEE FOCS*, 1996.
18. E. Hazan, S. Safra and O. Schwartz. On the hardness of approximating k-dimensional matching. In the *Proc. APPROX*, Springer LNCS, 2003.
19. D.S. Hochbaum. Efficient bounds for the stable set, vertex cover, and set packing problems. *Discrete Applied Mathematics*, **6**, pp. 243–254, 1983.
20. J. Kleinberg. *Approximation algorithms for disjoint paths problems.* PhD thesis, MIT, 1996.
21. S.G. Kolliopoulos and C. Stein. Approximating disjoint-path problems using greedy algorithms and packing integer programs. In *Proc. 6th IPCO*, LNCS, **1412**, pp. 153–162, 1998.
22. P. Kolman. A note on the greedy algorithm for the unsplittable flow problem. *Information Processing Letters*, **88**(3), pp. 101–105, 2003.
23. P. Kolman and Ch. Scheideler. Improved bounds for the unsplittable flow problem. In the *Proc. 13th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2002.
24. D. Lehmann, L. Ita O'Callaghan and Y. Shoham. Truth revelation in rapid, approximately efficient combinatorial auctions. In the *Proc. 1st ACM Conference on Electronic Commerce (EC)*, 1999. Journal version in: *Journal of the ACM*, **49**(5): 577–602, 2002.
25. L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Mathematics*, **13**, pp. 383–390, 1975.
26. A. Mu'alem and N. Nisan. Truthful Approximation Mechanisms for Restricted Combinatorial Auctions. In the *Proc. 18th National AAAI Conference on Artificial Intelligence*, 2002.
27. P. Raghavan. Probabilistic construction of deterministic algorithms: Approximating packing integer programs. *Journal of Computer and System Sciences*, **37**, pp. 130–143, 1988.
28. P. Raghavan and C.D. Thompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica*, **7**: 365–374, 1987.
29. S. Rajagopalan and V.V. Vazirani. Primal-dual RNC approximation algorithms for set cover and covering integer programs. *SIAM Journal on Computing*, **28**(2), 1998.
30. A. Srinivasan. A extension of the Lovász Local Lemma and its applications to integer programming. In the *Proc. 7th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 1996.
31. A. Srinivasan. Improved Approximation Guarantees for Packing and Covering Integer Programs, *SIAM Journal on Computing*, **29**, pp. 648–670, 1999.