An LP-Rounding $2\sqrt{2}$-Approximation for Restricted Maximum Acyclic Subgraph

Fabrizio Grandoni[a], Tomasz Kociumaka[b], Michał Włodarczyk[b]

[a] *IDSIA, University of Lugano, Switzerland*
[b] *Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Poland*

Abstract

In the classical Maximum Acyclic Subgraph problem (MAS), given a directed-edge weighted graph, we are required to find an ordering of the nodes that maximizes the total weight of forward-directed edges. MAS admits a 2-approximation, and this approximation is optimal under the Unique Game Conjecture.

In this paper we consider a generalization of MAS, the Restricted Maximum Acyclic Subgraph problem (RMAS), where each node is associated with a list of integer labels, and we have to find a labeling of the nodes so as to maximize the weight of edges whose head label is larger than the tail label. The interest in RMAS is mostly due to its connections with the Vertex Pricing problem (VP). VP is known to be $(2-\epsilon)$-hard to approximate via a reduction from RMAS, and the best known approximation factor for both problems is 4 (which is achieved via fairly simple algorithms).

In this paper we present a non-trivial LP-rounding algorithm for RMAS with approximation ratio $2\sqrt{2} \approx 2.828$. Our result shows that, in order to prove a 4-hardness of approximation result for VP (if possible), one should consider reductions from harder problems. Alternatively, our approach might suggest a different way to design approximation algorithms for VP.

Keywords: approximation algorithms, directed graphs, vertex pricing, linear programming, derandomization

---

1. Introduction

In the classical Maximum Acyclic Subgraph problem (MAS) we are given a directed graph $G = (V, E)$, with edge weights $\{w_e\}_{e \in E}$, and we look for an ordering of the nodes so as to maximize the total weight of forward-oriented edges. MAS admits a 2-approximation, which is optimal under the Unique Games Conjecture (UGC) [13].

In this paper we consider the following generalization of MAS. In the Restricted Maximum Acyclic Subgraph problem (RMAS) we are given the same input as for MAS, plus a set L_v of integer labels for each node v. We assume that the lists L_v are given explicitly. Our goal is to find a labeling $\{\ell(v)\}_{v \in V}$ of the nodes, $\ell(v) \in L_v$, that maximizes the weight of edges going from a lower label to a higher one. In other words, the objective function is

$$\sum_{\substack{e=(u,v)\in E\\ \ell(u)<\ell(v)}} w_e.$$

Note that we are allowed to assign the same label to multiple nodes. MAS is a special case of RMAS where $L_v = \{1, \ldots, |V|\}$ for all nodes v. Clearly, the 2-hardness of approximation for MAS extends to RMAS. Moreover, the Maximum Directed Cut problem is a special case of RMAS where $L_v = \{0, 1\}$. Consequently, due the result of Landis et al. [15], RMAS is APX-hard even for acyclic input graphs, which suggests that in full generality it might be harder to approximate than MAS. The best known approximation ratio for RMAS is 4, which is achieved with an almost trivial algorithm.

In this paper we investigate the approximability of RMAS and we present an improved $2\sqrt{2}$-approximation for the problem. Our result combines the trivial 4-approximation algorithm with a novel LP-rounding algorithm.

1.1. Related Work

Our interest in RMAS is motivated by the following Vertex Pricing problem (VP): we are given an undirected (multi-) graph $G = (V, E)$ with positive edge budgets $\{B_e\}_{e\in E}$. Our goal is to assign a non-negative price $p(v)$ to each node v so as to maximize the sum $p(u) + p(v)$ over the edges $e = \{u, v\}$ such that $p(u) + p(v) \le B_e$. Khandekar et al. [14] proved, via a reduction from a special case of RMAS, that VP is $(2-\varepsilon)$-hard to approximate for any constant $\varepsilon > 0$ (under UGC). Their reduction exploits instances of RMAS where labels are non-negative, each L_v contains 0, and $L_u \cap L_v = \{0\}$ for any distinct $u, v \in V$. Note that such instances still generalize MAS.

VP is APX-hard even on bipartite graphs [8]. The best known approximation for VP is 4 [2], and improving on that (if possible) is a well-known challenging open problem. Interestingly enough, the mentioned approximation is obtained with an algorithm analogous to the best-known 4-approximation for RMAS. So it is natural to wonder whether RMAS and VP are equally hard to approximate. Our result suggests that RMAS might actually be an easier problem than VP. Alternatively, it might suggest a way to design improved approximation algorithms for VP (though generalizing our approach to VP does not seem easy).

VP belongs to a broader family of *pricing* problems, which recently attracted a lot of attention. In particular, in the (single-minded unlimited-supply) Item Pricing problem (IT), we are given a (multi-) hyper-graph $G = (V, E)$ with hyper-edge positive budgets $\{B_e\}_{e\in E}$. We have to assign a non-negative price $p(v)$ to each node, so as to maximize the objective function $\sum_{e\in E: p(e)\le B_e} p(e)$, where $p(e) = \sum_{v\in e} p(v)$. This problem admits a $O(\log m + \log n)$-approximation, where n is the number of nodes and m the number of hyper-edges [12] (see also [4] for a refinement of this result). On the negative side, Chalermsook et al. [5] showed that this problem is hard to approximate within $\log^{1-\varepsilon} n$ for any constant $\varepsilon > 0$, unless $NP \subseteq ZPTIME(n^{\log^{\delta} n})$, where $\delta > 0$ is a constant depending on ε. Better approximation algorithms are known for the special case where the maximum size k of any hyper-edge is small. In particular, an $O(k)$-approximation is given in [2]. As recently shown [6], for constant values of k the latter result is (asymptotically) essentially the best possible unless $P = NP$.

VP is a special case of IT where all hyper-edges have size precisely 2. Another well-studied special case of IT is the Highway problem. Here one is given a path P on the node-set V, and hyper-edges are forced to induce subpaths of P. This problem was shown to be weakly NP-hard by Briest and Krysta [4] and strongly NP-hard by Elbassioni, Raman, Ray, and Sitters [8]. Balcan and Blum [2] gave an $O(\log n)$-approximation for the problem. Gamzu and Segev [10] improved the approximation factor to $O(\log n/\log\log n)$. Elbassioni, Sitters, and Zhang [9] developed a QPTAS, exploiting the profiling technique introduced by Bansal et al. [3]. Finally, a PTAS was given by Grandoni and Rothvoß [11].

The Tree Tollbooth problem is another special case of IT that generalized the Highway problem. Here hyper-edges are forced to induce paths in a given tree T. An $O(\log n)$-approximation was developed in [8], which was slightly improved to $O(\log n/\log\log n)$ by Gamzu and Segev [10]. For the case of uniform budgets an $O(\log\log n)$-approximation was given by Cygan et al. [7]. Tree Tollbooth is APX-hard [12].

2. An Improved Approximation Algorithm for RMAS

In this section we present our improved approximation algorithm for RMAS. In Section 2.1 we revisit the folklore 4-approximation for the problem, that is one of our building blocks. In Section 2.2 we present and analyze a novel LP-based algorithm. Finally, in Section 2.3 we discuss the derandomization of both algorithms and conclude with our main result.

In the following $W = \sum_{e\in E} w_e$ is the sum of all the weights. Observe that trivially W is an upper bound on the profit of the optimum solution. For a given node $u \in V$, let $\ell_{u,\min} = \min\{\ell : \ell \in L_u\}$ and $\ell_{u,\max} = \max\{\ell : \ell \in L_u\}$. Without loss of generality, we can assume that, for any edge $e = (u, u') \in E$, one has $\ell_{u,\min} < \ell_{u',\max}$. Otherwise, e can be filtered out without changing the value of the optimum solution.

2.1. A Simple Randomized Algorithm

Consider the following simple algorithm for RMAS: independently for each $u \in V$, set $\ell(u) = \ell_{u,\min}$ with probability $\frac{1}{2}$ and $\ell(u) = \ell_{u,\max}$ otherwise.

Lemma 1. *The above algorithm computes a solution of expected profit at least* $\frac{W}{4}$.

Proof. Fix an edge $e = (u, u') \in E$. By the initial filtering, $\ell_{u,\min} < \ell_{u',\max}$. Therefore, with probability at least $\frac{1}{4}$ one has $\ell(u) < \ell(u')$ and we benefit from e. By linearity of expectation the total expected profit is at least $\sum_e \frac{w_e}{4} = \frac{W}{4}$. □

The above result provides a 4-approximation for RMAS. This analysis of the approximation ratio turns out to be tight, by exploiting a result in [1] (see also [14]).

Lemma 2. *The above algorithm has approximation factor at least* 4.

Proof. Observe that any solution returned by the trivial approximation algorithm can be seen as a bipartition of the set of vertices $V = V_{\min} \cup V_{\max}$ depending on whether a vertex u was assigned $\ell_{u,\min}$ or $\ell_{u,\max}$. Moreover, if label lists $L_u = L$ are uniform for all $u \in V$, such solutions may only benefit from edges from $V_{\min}$ to $V_{\max}$, and consequently values of these solutions cannot exceed the maximum weight of a directed cut in the underlying weighted graph.

Alon et al. [1] constructed a family of directed acyclic graphs with maximal directed cut of size $\frac{m}{4} + o(m)$ where m is the number of edges. We use such a graph as an RMAS instance with unit weights and $L_u = \{1, \ldots, |V|\}$ for every vertex u. Acyclic graphs admit a topological order, which can be used to construct the optimal solution setting positions in this order as labels. The value of this solution is m which is, by the argument above, $4 - o(1)$ larger than any solution the considered algorithm can produce. $\square$

2.2. An LP-Rounding Algorithm

Observe that the algorithm from Section 2.1 provides an approximation strictly better than 4 whenever the optimum solution has profit strictly smaller than W. In this subsection we present a better algorithm for the opposite case.

For an instance D of RMAS, consider the following LP-relaxation $LP(D)$ of D. Let $L = \bigcup_{v \in V} L_v$.

$$
\begin{array}{llll}
\max & \displaystyle\sum_{e=(u,u') \in E} \sum_{\ell < \ell'} w_e y_{uu'}(\ell, \ell') & & \\
s.t. & \displaystyle\sum_{\ell \in L} x_u(\ell) = 1 & \forall\, u \in V & \\
& x_u(\ell) = 0 & \forall\, u \in V, \forall\, \ell \in L \setminus L_u & (1) \\
& \displaystyle\sum_{\ell' \in L} y_{uu'}(\ell, \ell') = x_u(\ell) & \forall\, u, u' \in V, \forall\, \ell \in L & (2) \\
& y_{uu'}(\ell, \ell') = y_{u'u}(\ell', \ell) & \forall\, u, u' \in V, \forall\, \ell, \ell' \in L & (3) \\
& x_u(\ell),\ y_{uu'}(\ell, \ell') \geq 0 & \forall\, u, u' \in V, \forall\, \ell, \ell' \in L &
\end{array}
$$

In the above LP, variable $x_u(\ell)$ denotes whether a vertex u has label ℓ, and variable $y_{uu'}(\ell, \ell')$ denotes whether simultaneously u has label ℓ and u' has label ℓ'. For the sake of presentation, we defined the variables $x_u(\ell)$ and $y_{uu'}(\ell, \ell')$ also for unfeasible label assignments. Constraint (1) guarantees that such variables are set to zero.

Consider the following natural randomized LP-rounding algorithm. Let (x, y) be an optimal solution to $LP(D)$. Observe that for a fixed vertex v variables $x_u(\ell)$, $\ell \in L_u$, define a probability distribution. We draw $\ell(u)$ from this distribution, independently for each $u \in V$. Then $\ell(u) = \ell$ with probability $x_u(\ell)$.

Lemma 3. *The above algorithm computes a solution of expected cost at least $\frac{\mathsf{LP}^2}{2W}$, where LP is the value of the optimal fractional solution to $LP(D)$.*

In order to prove the above lemma, we need the following technical result.

Lemma 4. *Let $A = [a_{ij}]$ be an $n \times n$ matrix with $a_{ij} \in \mathbb{R}_{\geq 0}$. Let $r_i = \sum_j a_{ij}$ and $c_j = \sum_i a_{ij}$ be the sum of entries in the i-th row and j-th column, respectively. Then*

$$\sum_{i<j} r_i c_j \geq \frac{1}{2}\left(\sum_{i<j} a_{ij}\right)^2.$$

Proof. We use Iverson notation: $[\phi]$ is 1 if ϕ is satisfied and 0 otherwise. By symmetry between (i,j) and (i',j') we have

$$\left(\sum_{i<j} a_{ij}\right)^2 = \sum_{i,i',j,j'} [i<j][i'<j']a_{ij}a_{i'j'} \leq 2\sum_{i,i',j,j'} [i<j][i'<j'][i \leq i']a_{ij}a_{i'j'}. \tag{4}$$

Clearly $(i \leq i' \wedge i' < j') \Rightarrow i < j'$, and consequently

$$\begin{aligned}\sum_{i,i',j,j'} [i<j][i'<j'][i \leq i']a_{ij}a_{i'j'} &\leq \sum_{i,i',j,j'} [i'<j'][i \leq i']a_{ij}a_{i'j'} \leq \sum_{i,i',j,j'} [i<j']a_{ij}a_{i'j'} \\ &= \sum_{i<j'}\left(\sum_j a_{ij}\right)\left(\sum_{i'} a_{i'j'}\right) = \sum_{i<j'} r_i c_{j'}.\end{aligned} \tag{5}$$

The claim follows by combining (4) and (5). □

Proof. (of Lemma 3) For an edge $e = (u,u') \in E$, define $p_e = \sum_{\ell<\ell'} x_u(\ell)x_{u'}(\ell')$ and $q_e = \sum_{\ell<\ell'} y_{u,u'}(\ell,\ell')$. Note that the expected profit from e equals $p_e w_e$, while the profit of the LP solution for the same edge is $q_e w_e$. In particular, $\mathsf{LP} = \sum_{e\in E} q_e w_e$.

For each $e = (u,u')$ we apply Lemma 4 to the $|L| \times |L|$ matrix A with $a_{ij} = y_{uu'}(i,j)$. By (2) the sum r_i of the entries in the i-th row is equal to $x_u(i)$. Moreover, combining (2) and (3), one has that the sum c_j of the entries in the j-th column is $x_{u'}(j)$. We conclude that

$$p_e = \sum_{i<j} x_u(i)x_{u'}(j) = \sum_{i<j} r_i c_j \overset{\text{Lem. 4}}{\geq} \tfrac{1}{2}\left(\sum_{i<j} a_{i,j}\right)^2 = \tfrac{1}{2}\left(\sum_{i<j} y_{uu'}(i,j)\right)^2 = \tfrac{1}{2}q_e^2.$$

As function $f(x) = x^2$ is convex, by Jensen's inequality with coefficients $\frac{w_e}{W}$ we obtain that the expected profit $\sum_e w_e p_e$ of the approximate solution satisfies:

$$\sum_e w_e p_e \geq \tfrac{W}{2}\sum_e \tfrac{w_e}{W} q_e^2 \geq \tfrac{W}{2}\left(\sum_e \tfrac{w_e}{W} q_e\right)^2 = \frac{\mathsf{LP}^2}{2W}.$$ □

2.3. Derandomization and Conclusions

We start by observing that both the mentioned algorithms can be easily derandomized. Although this is a rather standard application of the method of conditional expectations, for completeness we provide a short description below.

Lemma 5. *The algorithm from Lemma 1 can be derandomized.*

Proof. Let Z be a random variable equal to the value of the integer solution computed by the randomized algorithm. We consider nodes in an arbitrary order $v_1, v_2, \ldots, v_{|V|}$. At each iteration $i = 1, \ldots, |V|$, we have already fixed labels ℓ_j for each node v_j, $j < i$, such that the invariant $\mathbb{E}[Z|\ell(v_j) = \ell_j, j = 1, \ldots, i-1] \geq \mathbb{E}[Z]$ holds. In the considered iteration we fix the label ℓ_i for node v_i as follows. We compute the two quantities $\mathbb{E}[Z|\ell(v_j) = \ell_j, j = 1, \ldots, i-1 \text{ and } \ell(v_i) = \ell_{v_i,\min}]$ and $\mathbb{E}[Z|\ell(v_j) = \ell_j, j = 1, \ldots, i-1 \text{ and } \ell(v_i) = \ell_{v_i,\max}]$. Note that these quantities can be easily computed in polynomial time. Observe also that at least one of the two quantities is lower bounded by $\mathbb{E}[Z|\ell(v_j) = \ell_j, j = 1, \ldots, i-1]$, hence by $\mathbb{E}[Z]$ because of the invariant. We set ℓ_i to the label in $\{\ell_{v_i,\min}, \ell_{v_i,\max}\}$ that achieves the larger conditional expectation. It follows that the resulting deterministic algorithm computes a solution of profit at least $\mathbb{E}[Z] \geq \frac{W}{4}$. □

Let us remark that the lower bound from Lemma 2 also applies to the derandomized variant in the above lemma.

Lemma 6. *The algorithm from Lemma 3 can be derandomized.*

Proof. We use the method of conditional expectation similarly to the proof of Lemma 5. Let Z be a random variable equal to the value of the solution computed by the randomized algorithm. Recall from the proof of Lemma 3 that

$$\mathbb{E}[Z] = \sum_{e \in E} p_e w_e, \quad \text{where} \quad p_e = \sum_{\ell < \ell'} x_u(\ell) x_{u'}(\ell'). \tag{6}$$

Let us choose some vertex v_1. Observe that for some $\ell_1 \in L_{v_1}$ it must be $\mathbb{E}[Z \,|\, l(v_1) = \ell_1] \geq \mathbb{E}[Z]$. To compute $\mathbb{E}[Z \,|\, \ell(v_1) = \ell]$ we can set $x_{v_1}(\ell) = 1$ and $x_{v_1}(\ell') = 0$ for $\ell' \neq \ell$ and use (6). Therefore such ℓ_1 may be computed in polynomial time.

We fix $\ell(v_1) = \ell_1$ and repeat this procedure on the remaining nodes considered in any order $v_2, \ldots, v_{|V|}$ until a label ℓ_i is chosen for each node v_i. The conditional expected value never decreases so the value of the resulting solution is at least $\mathbb{E}[Z] \geq \frac{\mathsf{LP}^2}{2W}$. □

We now have all the ingredients to prove the main result in this paper.

Theorem 1. *RMAS admits a deterministic $2\sqrt{2}$-approximation algorithm.*

Proof. Let OPT be the value of the optimal solution, $\mathsf{OPT} \leq \mathsf{LP} \leq W$, and assume w.l.o.g. $\mathsf{OPT} > 0$. Consider the algorithm that returns the better solution among the ones computed by the algorithms from Lemmas 5 and 6. The profit of the constructed solution is bounded from below by

$$\max\left\{\frac{W}{4}, \frac{\mathsf{LP}^2}{2W}\right\} = \mathsf{LP} \cdot \max\left\{\frac{W}{4\mathsf{LP}}, \frac{\mathsf{LP}}{2W}\right\} \geq \frac{\mathsf{LP}}{2\sqrt{2}} \geq \frac{\mathsf{OPT}}{2\sqrt{2}}.$$

□

We conclude with a few remarks about the integrality gap of the considered LP relaxation for RMAS. On one hand, the proof of Theorem 1 shows that this gap is at most $2\sqrt{2}$. On the other hand, this gap is at least 2. To see that, consider the same construction as in Lemma 2, but with label sets $L_u = \{0,1\}$. Here the optimal integral solution has value at most $\frac{m}{4} + o(m)$ (i.e., the size of the maximum directed cut). However, a feasible fractional solution with value $\frac{W}{2} = \frac{m}{2}$ is obtained by setting $x_u(\ell_{u,min}) = x_u(\ell_{u,max}) = \frac{1}{2}$ and $y_{uu'}(\ell_{u,min}, \ell_{u',max}) = y_{uu'}(\ell_{u,max}, \ell_{u',min}) = \frac{1}{2}$. Determining the exact value of the integrality gap is left as an open problem.

3. Acknowledgements

This work has been partially supported by the ERC Starting Grant NEWNET 279352 and by Foundation for Polish Science grant HOMING PLUS/2012-6/2.

References

[1] Noga Alon, Béla Bollobás, András Gyárfás, Jenő Lehel, and Alex Scott. Maximum directed cuts in acyclic digraphs. *Journal of Graph Theory*, 55(1):1–13, 2007.

[2] Maria-Florina Balcan and Avrim Blum. Approximation algorithms and online mechanisms for item pricing. In *ACM Conference on Electronic Commerce (EC)*, pages 29–35, 2006.

[3] Nikhil Bansal, Amit Chakrabarti, Amir Epstein, and Baruch Schieber. A quasi-PTAS for unsplittable flow on line graphs. In *ACM Symposium on Theory of Computing (STOC)*, pages 721–729, 2006.

[4] Patrick Briest and Piotr Krysta. Single-minded unlimited supply pricing on sparse instances. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1093–1102, 2006.

[5] Parinya Chalermsook, Julia Chuzhoy, Sampath Kannan, and Sanjeev Khanna. Improved hardness results for profit maximization pricing problems with unlimited supply. In *International Workshop on Approximation, Randomization, and Combinatorial Optimization (APPROX-RANDOM)*, pages 73–84, 2012.

[6] Parinya Chalermsook, Bundit Laekhanukit, and Danupon Nanongkai. Independent set, induced matching, and pricing: Connections and tight (subexponential time) approximation hardnesses. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 370–379, 2013.

[7] Marek Cygan, Fabrizio Grandoni, Stefano Leonardi, Marcin Pilipczuk, and Piotr Sankowski. A path-decomposition theorem with applications to pricing and covering on trees. In *European Symposium on Algorithms (ESA)*, pages 349–360, 2012.

[8] Khaled M. Elbassioni, Rajiv Raman, Saurabh Ray, and René Sitters. On profit-maximizing pricing for the highway and tollbooth problems. In *Symposium on Algorithmic Game Theory (SAGT)*, pages 275–286, 2009.

[9] Khaled M. Elbassioni, René Sitters, and Yan Zhang. A quasi-ptas for profit-maximizing pricing on line graphs. In *European Symposium on Algorithms (ESA)*, pages 451–462, 2007.

[10] Iftah Gamzu and Danny Segev. A sublogarithmic approximation for highway and tollbooth pricing. In *International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 582–593, 2010.

[11] Fabrizio Grandoni and Thomas Rothvoß. Pricing on paths: A PTAS for the highway problem. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 675–684, 2011.

[12] Venkatesan Guruswami, Jason D. Hartline, Anna R. Karlin, David Kempe, Claire Kenyon, and Frank McSherry. On profit-maximizing envy-free pricing. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1164–1173, 2005.

[13] Venkatesan Guruswami, Rajsekar Manokaran, and Prasad Raghavendra. Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 573–582, 2008.

[14] Rohit Khandekar, Tracy Kimbrel, Konstantin Makarychev, and Maxim Sviridenko. On hardness of pricing items for single-minded bidders. In *International Workshop on Approximation, Randomization, and Combinatorial Optimization (APPROX-RANDOM)*, pages 202–216, 2009.

[15] Michael Lampis, Georgia Kaouri, and Valia Mitsou. On the algorithmic effectiveness of digraph decompositions and complexity measures. *Discrete Optimization*, 8(1):129–138, 2011.