

## COMBINATORIAL APPROXIMATION OF MAXIMUM $k$ -VERTEX COVER IN BIPARTITE GRAPHS WITHIN RATIO 0.7

VANGELIS TH. PASCHOS\*

**Abstract.** We propose and analyze a simple *purely combinatorial algorithm* for MAX  $k$ -VERTEX COVER in bipartite graphs, achieving approximation ratio 0.7. The only combinatorial algorithm currently known until now for this problem is the natural greedy algorithm, that achieves ratio  $\frac{(e-1)}{e} = 0.632$ .

**Mathematics Subject Classification.** 03D15, 05C70, 05C85, 68Q25, 68W25, 68W40

Received April 25, 2017. Accepted October 30, 2017.

### 1. INTRODUCTION

In the MAX  $k$ -VERTEX COVER problem, a graph  $G(V, E)$  with  $|V| = n$  and  $|E| = m$  is given together with an integer  $k \leq n$ . The goal is to find a subset  $K \subseteq V$  with  $k$  elements such that the total number of edges covered by  $K$  is maximized. This problem is strongly **NP**-hard even in bipartite graphs [2, 4] (in what follows, for clarity reasons, MAX  $k$ -VERTEX COVER problem in bipartite graphs will be denoted by BIPARTITE MAX  $k$ -VERTEX COVER).

The approximation of MAX  $k$ -VERTEX COVER has been originally studied in [5] and revisited in [6], where an approximation  $1 - (\frac{1}{e})$  was proved, achieved by the natural greedy algorithm. This ratio is tight in bipartite graphs [3]. In [1], using a sophisticated linear programming method, the approximation ratio for MAX  $k$ -VERTEX COVER is improved up to  $\frac{3}{4}$ . Finally, by an easy reduction from MIN VERTEX COVER, it can be shown that MAX  $k$ -VERTEX COVER in general graphs does not admit a polynomial time approximation schema (PTAS), unless **P** = **NP** [7]. The proof-schema of this negative result of [7] does not apply to BIPARTITE MAX  $k$ -VERTEX COVER since MIN VERTEX COVER is polynomial in bipartite graphs. Obviously, the result of [1] immediately applies to BIPARTITE MAX  $k$ -VERTEX COVER. More recently, [4] has improved this ratio for BIPARTITE MAX  $k$ -VERTEX COVER up to  $\frac{8}{9}$ , always using involved linear programming techniques,<sup>1</sup> but the existence of a PTAS for such graphs always remains open.

Finally, let us note that BIPARTITE MAX  $k$ -VERTEX COVER is polynomial in regular bipartite graphs (where all the vertices have the same degree) or in semi-regular ones (where the vertices of each color class have the same degree but the two degrees are different). Indeed, in the case of regular bipartite graphs, it suffices to

---

*Keywords and phrases:* Approximation algorithm, bipartite graph, max  $k$ -VERTEX cover.

Université Paris-Dauphine, PSL\* Research University, CNRS UMR 7243, LAMSADE, 75016 Paris, France.

\* Corresponding author: [paschos@lamsade.dauphine.fr](mailto:paschos@lamsade.dauphine.fr)

<sup>1</sup>The complexity of the algorithm is not given in [4]; a rough evaluation of it, gives a complexity of  $O((|V_1||V_2|)^{\frac{11}{2}})$  which is bounded above by  $O(n^{11})$ , where  $V_1$  and  $V_2$  are the two independent that form the vertex-set of the bipartite graph ( $|V_1| + |V_2| = n$ ).

chose  $k$  vertices in one color class,<sup>2</sup> while in the case of semi-regular ones, it suffices to chose  $k$  vertices in the color class with the maximum degree.

Our principal motivation for this paper is to study *in what extent combinatorial methods for BIPARTITE MAX  $k$ -VERTEX COVER compete with linear programming ones*. In other words, *what is the ratios' level, a purely combinatorial algorithm can guarantee?* In this purpose, we devise an  $O(n^3)$  algorithm that builds five distinct solutions and returns the best among them; for this algorithm, we prove a worst case 0.7-approximation ratio. Let us note that a similar issue is presented in [8] for MAX CUT where a 0.531-ratio combinatorial algorithm is given. Comparison of classes of methods with respect to their abilities to solve problems seems to be a very interesting research issue. This may bring new insights to both the problems handled and the methods themselves. Furthermore, such studies may exhibit interesting and funny mathematical problems.

## 2. PRELIMINARIES

Consider a bipartite graph  $B(V_1, V_2, E)$ , fix an optimal solution  $O$  for BIPARTITE MAX  $k$ -VERTEX COVER (i.e., a vertex-set on  $k$  vertices covering a maximum number of edges in  $E$ ) and denote by  $k_1$  and  $k_2$  the cardinalities of the subsets  $O_1$  and  $O_2$  of  $O$  lying in the color-classes  $V_1$  and  $V_2$ , respectively. W.l.o.g., we assume  $k_1 \leq k_2$  and we set:

$$k_1 = \mu \cdot k_2, \quad \mu \leq 1 \tag{2.1}$$

$$k = k_1 + k_2 = (1 + \mu) \cdot k_2 \tag{2.2}$$

Denote by  $\delta(V')$ ,  $V' \subseteq V = V_1 \cup V_2$ , the number of edges covered by  $V'$  and by  $\text{opt}(B)$  the value of an optimal solution (i.e., the number of edges covered by  $O$ ).

Let  $S_i$ ,  $i = 1, 2$ , be the  $k_i$  vertices of  $V_i$  that cover a maximum number of edges. Obviously,  $S_i$  is the set of the  $k_i$  largest degree vertices in  $V_i$  (breaking ties arbitrarily) and the following hold:

$$\begin{aligned} \delta(S_1) &\geq \delta(O_1) \\ \delta(S_2) &\geq \delta(O_2) \end{aligned} \tag{2.3}$$

In what follows, we call “best” vertices, a set of vertices that cover a maximum number of *uncovered* edges<sup>3</sup> in  $B$ . Furthermore, we will also use the following additional notations and conventions (we assume that vertices in both  $V_1$  and  $V_2$  are ordered in decreasing degree order), where all the used Greek letters imply parameters that are all smaller than, or equal to, 1:

- $\delta(O_1)$ : the number of edges covered by  $O_1$ ; for conciseness we set  $\delta(O_1) = \alpha \cdot \text{opt}(B)$ ;
- $\beta_1 \cdot \delta(O_1) = \beta_1 \cdot \alpha \cdot \text{opt}(B)$ : the number of edges covered by  $S_1 \cap O_1$ ;
- $\delta'(O_2)$ : the number of *private* edges covered by  $O_2$ , i.e., the edges already covered by  $O_1$  are not counted up to  $\delta'(O_2)$ ; obviously,  $\delta'(O_2) = (1 - \alpha) \cdot \text{opt}(B)$ ;
- $\theta \cdot \delta(O_1)$ : the number of edges (if any) from  $O_1$  that go “below”  $O_2$  (recall  $V_1$  and  $V_2$  are ordered in decreasing degree order);
- $\gamma \cdot \delta'(O_2)$ : symmetrically, it denotes the number of edges of  $O_2$  that go below the vertices of  $O_1$ ;
- $\zeta \cdot \delta(O_1)$ : suppose that after taking the  $k$  best vertices of  $V_1$ , there still remain, say,  $k'_1$  vertices of  $O_1$  that have not been encountered yet; then,  $\zeta \cdot \delta(O_1)$  is the number of edges covered by those vertices;
- $\lambda \cdot \delta'(O_2)$ : this is the symmetric of the quantity  $\zeta \cdot \delta(O_1)$  for the pair  $(V_2, O_2)$  (supposing that the number of vertices in  $O_2$  that have not been encountered is  $k'_2$ ).

<sup>2</sup>For the rest of the paper, following standard graph-theoretic vocabulary, given a bipartite graph  $B(V_1, V_2, E)$ , vertex-sets  $V_1$  and  $V_2$  will be referred as color classes of  $B$ .

<sup>3</sup>For instance, saying “we take  $S_1$  plus the  $k_2$  best vertices in  $V_2$ , this means that we take  $S_1$  and then  $k_2$  vertices of highest degree in  $B[(V_1 \setminus S_1), V_2]$ .”

In Figure 1, the edge-sets defined by the parameters above are illustrated. Heavy lines within rectangles  $V_1$  and  $V_2$  represent the borders of  $S_1$  and  $S_2$  (the upper ones) and those of the  $k$  best vertices (the lower ones). Edges from  $O_1$  ( $\arg(\delta(O_1))$ ) are not shown in the the figure. They can go everywhere in  $V_2$ . Private edges of  $O_2$  ( $\arg(\delta'(O_2))$ ) are shown as heavy lack lines (the set of edges  $\delta'_2$ ). They can go everywhere in  $V_1 \setminus O_1$ .

The basic idea of the algorithm is quite simple. It computes the best among five solutions built. It is presented and analyzed in Section 3.

### 3. COMBINATORIAL ALGORITHMS CAN LEAD TO AT LEAST 0.7-APPROXIMATION RATIO

Consider Algorithm 3.1. The sorting in the first line is performed in time  $O(n \log n)$ . The guessing in the second line can be performed by running it for all pairs of integers  $(k_1, k_2)$  such that  $k_1 + k_2 = k \leq n$ ; thus, the time needed for this is  $O(n^2)$ . Finally, since the vertices are ordered, each of the solutions SOL<sub>1</sub> to SOL<sub>4</sub> can be built in linear time. So, the overall complexity of Algorithm 3.1 is  $O(n^3)$ . Let us also note that since the algorithm runs for any value of  $k_1$  and  $k_2$ , it will run for  $k_1 = k$  and  $k_2 = k$ . So, it is optimal for the family of instances of [3], where the greedy algorithm attains the ratio  $\frac{(e-1)}{e}$ . For these instances, the optimal solution consists of taking  $k$  vertices from one of the two color classes of the input graph, while the greedy algorithm fails to do this. Finally, let us note that Step 2 in SOL<sub>4</sub> of Algorithm 3.1 may seem to be redundant. But, as it will hopefully become clear later in Lemma 3.5, it is useful in order to simplify its proof.

**Algorithm 3.1.** A combinatorial algorithm for BIPARTITE MAX  $k$ -VERTEX COVER.

---

**Input:** A bipartite graph  $B(V_1, V_2, E)$  and a constant  $k < |V_1| + |V_2|$   
**Output:** A  $k$ -VERTEX cover of  $B$   
 sort the vertices of  $V_1$  and  $V_2$  in decreasing degree-order;  
 guess  $k_1, k_2$  of  $O_1$  and  $O_2$ , respectively;  
 build the following solutions;  
 SOL<sub>1</sub>: take  $S_1$  plus the  $k_2$  remaining best vertices from  $V_2$ ;  
 SOL<sub>2</sub>: take  $S_2$  plus the  $k_1$  remaining best vertices from  $V_1$ ;  
 SOL<sub>3</sub>: take the  $k$  best vertices of  $V_2$ ;  
 SOL<sub>4</sub>: take the best between the following two solutions:  
     1. the  $k$  best vertices of  $V_1$ ;  
     2. the best  $2 \cdot k_1$  vertices of  $V_1$  plus the remaining  $k - 2 \cdot k_1$  best vertices of  $V_2$ ;  
**return** the best among the solutions computed.

---

In what follows, in Lemmas 3.2–3.5, we analyze the solutions built by the algorithm and provide several expressions for the ratios achieved by each of them. All these ratios are expressed as functions of the parameters specified in Section 2. In order to simplify notations from now on we shall write  $\text{opt}$  instead of  $\text{opt}(B)$ .

**Lemma 3.2.** *The approximation ratio achieved by solution SOL<sub>1</sub> is the maximum of the following quantities:*

$$1 - \alpha + \beta_1 \cdot \alpha \tag{3.1}$$

$$\alpha + \gamma \cdot (1 - \alpha) \tag{3.2}$$

Furthermore, if  $S_1$  and  $O_1$  coincide (i.e.,  $S_1 \cap O_1 = S_1$ ), SOL<sub>1</sub> is optimal.

*Proof.* For (3.1),  $S_1$  covers, by (2.3), more than  $\delta(O_1) = \alpha \cdot \text{opt}$  edges. Decompose this edge-set into a set  $X$  of edges covered by  $S_1 \setminus (S_1 \cap O_1)$  and the set of edges of size  $\beta_1 \cdot \alpha \cdot \text{opt}$  of edges covered by  $S_1 \cap O_1$ . On the other hand, the  $k_2$  remaining best vertices in  $V_2$  will cover more edges than the  $k_2$  remaining best vertices in  $O_2$ , that cover more than  $(1 - \alpha) \cdot \text{opt} - |X|$  edges, qed.

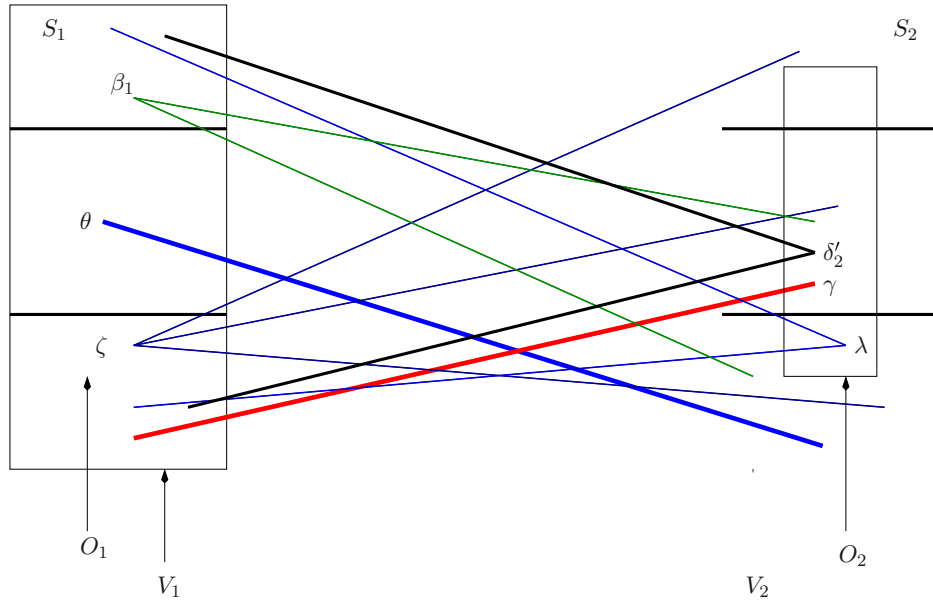


FIGURE 1. The edge-sets induced by the several parameters. (Color online.)

For (3.2), whenever \$S\_1\$ does not coincide with \$O\_1\$, there are vertices of \$O\_1\$ that are found below \$S\_1\$. Since \$\gamma \cdot \delta'(O\_2)\$ is the number of edges from \$O\_2\$ that go belong \$O\_1\$, these edges will be not counted up in the set of edges covered by \$S\_1\$.

Finally, if \$S\_1\$ and \$O\_1\$ coincide, then \$SOL\_1\$ will cover \$\alpha \cdot opt + (1 - \alpha) \cdot opt = opt\$ edges. □

**Lemma 3.3.** *The approximation ratio achieved by solution \$SOL\_2\$ is bounded below by:*

$$1 - \alpha + \alpha \cdot \theta \tag{3.3}$$

*Proof.* The proof is similar with the one of Lemma 3.2 for (3.2). □

**Lemma 3.4.** *The approximation ratio achieved by solution \$SOL\_3\$ is the maximum of the following quantities:*

$$1 - \lambda \cdot (1 - \alpha) - \alpha \cdot \theta \tag{3.4}$$

$$(1 - \alpha) \cdot (1 + \lambda \cdot \mu) \tag{3.5}$$

*Proof.* If after taking the \$k\$ best vertices of \$V\_2\$ the whole of \$O\_2\$ has been encountered, all but \$\theta \cdot \delta(O\_1)\$ edges of the optimum have been covered. In this case, an approximation ratio \$1 - \alpha \cdot \theta\$ is achieved.

Otherwise, by the definition of \$\lambda \cdot \delta'(O\_2)\$:

$$opt - \lambda \cdot \delta'(O_2) - \theta \cdot \delta(O_1) = opt \cdot (1 - \lambda \cdot (1 - \alpha) - \theta \cdot \alpha)$$

edges of the optimum are covered.

On the other hand, taking the \$k\$ best vertices of \$V\_2\$, consists in first taking \$S\_2\$ (covering \$(1 - \alpha) \cdot opt\$ edges) and then the \$k\_1\$ best vertices below it. Furthermore, below the \$k\$ best vertices, the group of the \$k'\_2\$ "worst" vertices of \$O\_2\$ has average degree at least \$\frac{\lambda \cdot \delta'(O\_2)}{k'\_2}\$. Since the algorithm takes \$k\_1\$ "better" vertices, they will cover

at least:

$$\frac{k_1}{k'_2} \cdot \lambda \cdot \delta'(O_2) \geq \frac{k_1}{k_2} \cdot \lambda \cdot (1 - \alpha) \cdot \text{opt} \stackrel{(2.1)}{=} \mu \cdot \lambda \cdot (1 - \alpha) \cdot \text{opt}$$

which proves (3.5).  $\square$

**Lemma 3.5.** *The approximation ratio achieved by solution  $\text{SOL}_4$  is the maximum of the following quantities:*

$$1 - \zeta \cdot \alpha - \gamma \cdot (1 - \alpha) \tag{3.6}$$

$$(2 - \beta_1) \cdot \alpha + \frac{(1 - \mu) \cdot \zeta \cdot \alpha}{\mu} \tag{3.7}$$

$$\frac{(1 - \mu) + \alpha \cdot (2 + \mu) - \alpha \cdot \beta_1}{2} + \frac{(1 - 2 \cdot \mu) \cdot \zeta \cdot \alpha}{2 \cdot \mu} \tag{3.8}$$

*Proof.* Let us first note that, if after taking the  $k$  best vertices in  $V_1$  all the vertices of  $O_1$  are captured, then the approximation ratio achieved is  $1 - \gamma \cdot (1 - \alpha)$ , since only  $\gamma \cdot \delta'(O_2) = \gamma \cdot (1 - \alpha) \cdot \text{opt}$  edges of the optimum are not covered. Suppose now that  $k'_1$  vertices of  $O_1$  are not captured. In this case, the  $k$  vertices taken from  $V_1$  cover:

$$\text{opt} - \zeta \cdot \delta(O_1) - \gamma \cdot \delta'(O_2) = (1 - \zeta \cdot \alpha - \gamma \cdot (1 - \alpha)) \cdot \text{opt}$$

For (3.7) and (3.8) now, observe first that the  $k$  vertices taken from  $V_1$  can be seen as the union of  $\frac{k}{k_1}$  consecutive  $k_1$ -groups (called clusters in what follows) and that, by (2.1) and (2.2),  $\frac{k}{k_1} = \frac{(1+\mu)}{\mu}$ . Assume also that the  $k - k'_1$  of  $O_1$  encountered among the  $k$  best vertices of  $V_1$  are included in the  $\pi$  first clusters. Denote by  $\kappa_i$  the number of vertices of  $O_1$  in the  $i$ th cluster,  $i = 1, \dots, \pi$ , and suppose that the ‘‘optimal’’  $\kappa_i$  vertices of cluster  $i$  cover  $\beta_i \cdot \delta(O_1) = \beta_i \cdot \alpha \cdot \text{opt}$  edges.

**Claim 3.6.** *Consider cluster  $i$  and denote by  $\bar{O}_{1,i}$  the part of  $O_1$  not captured by clusters  $1, 2, \dots, i - 1$  (thus,  $\bar{O}_{1,i} = \sum_{j=i}^{\pi} \kappa_j + k'_1$ ). Then, the vertices of cluster  $i$  will cover at least  $(1 - \sum_{j=1}^{i-1} \beta_j) \cdot \alpha \cdot \text{opt}$  edges.*

In order to prove Claim 3.6, observe that the part of  $\delta(O_1)$  covered by  $\bar{O}_{1,i}$  is:

$$\delta(\bar{O}_{1,i}) = \delta(O_1) - \sum_{j=1}^{i-1} \beta_j \cdot \delta(O_1) = \left(1 - \sum_{j=1}^{i-1} \beta_j\right) \cdot \alpha \cdot \text{opt}$$

and that the  $\delta(\bar{O}_{1,i})$  edges are covered by  $\sum_{j=i}^{\pi} \kappa_j + k'_1 = k_1 - (\sum_{j=1}^{i-1} \kappa_j) \leq k_1$  vertices, while cluster  $i$  contains exactly  $k_1$  vertices with degree at least as large as those of  $\bar{O}_{1,i}$ . An easy average argument derives then that the vertices of cluster  $i$  will cover at least:

$$k_1 \cdot \frac{\left(1 - \sum_{j=1}^{i-1} \beta_j\right) \cdot \alpha \cdot \text{opt}}{k_1 - \left(\sum_{j=1}^{i-1} \kappa_j\right)} \geq \left(1 - \sum_{j=1}^{i-1} \beta_j\right) \cdot \alpha \cdot \text{opt}$$

edges, qed.

Consider the two first groups clusters taken from  $V_1$ . The first of them ( $S_1$ ) covers more than  $\delta(O_1) = \alpha \cdot \text{opt}$  edges (by (2.3)) while, by Claim 3.6, the second one will cover more than  $\left(\frac{k_1}{(k_1 - \kappa_1)}\right) \cdot (1 - \beta_1) \cdot \delta(O_1) \geq (1 - \beta_1) \cdot \alpha \cdot \text{opt}$  edges. Observe also that, by (2.1) and (2.2),  $\frac{k}{k_1} = \frac{(1+\mu)}{\mu}$ . In any of the remaining  $\left(\frac{(1+\mu)}{\mu}\right) - 2 = \frac{(1-\mu)}{\mu}$

clusters, their vertices obviously cover more than  $\zeta \cdot \delta(O_1) = \zeta \cdot \alpha \cdot \text{opt}$  edges (indeed, by the average argument of Claim 3.6, more than  $k_1 \cdot \frac{\zeta \cdot \delta(O_1)}{k'_1} \geq \zeta \cdot \delta(O_1)$ ). Therefore:

$$|\text{SOL}_4| \geq \left[ (2 - \beta_1) + \frac{1 - \mu}{\mu} \cdot \zeta \right] \cdot \delta(O_1) = \left[ (2 - \beta_1) + \frac{1 - \mu}{\mu} \cdot \zeta \right] \cdot \alpha \cdot \text{opt}$$

that proves (3.7).

Let us now get some more insight in the value of  $\text{SOL}_4$ . By extending the discussion just above, the  $k_1$  vertices of cluster  $i$  will cover more than:

$$\frac{k_1}{k_1 - \sum_{j=1}^{i-1} \kappa_j} \cdot \left( \sum_{j=i}^{\pi} \beta_j + \zeta \right) \cdot \delta(O_1) \geq \left( 1 - \sum_{j=1}^{i-1} \beta_j \right) \cdot \delta(O_1) \quad (3.9)$$

Furthermore, as seen previously, all clusters below the  $\pi$  first ones containing the  $k_1 - k'_1$  captured vertices of  $O_1$ , will cover more than  $\zeta \cdot \delta(O_1)$  each.

Hence, summing (3.9) for  $i = 1$  to  $\pi$ , taking into account the remark just above, and setting  $\beta_0 = 0$ , the following relation holds:

$$\begin{aligned} |\text{SOL}_4| &\geq \left( \sum_{i=0}^{\pi-1} \left( 1 - \sum_{j=0}^i \beta_j \right) + \left( \frac{1 + \mu}{\mu} - \pi \right) \cdot \zeta \right) \cdot \delta(O_1) \\ &= \left( \pi - \sum_{i=1}^{\pi-1} (\pi - i) \cdot \beta_i + \left( \frac{1 + \mu}{\mu} - \pi \right) \cdot \zeta \right) \cdot \delta(O_1) \\ &= \left( \pi - \pi \cdot \sum_{i=1}^{\pi-1} \beta_i + \sum_{i=1}^{\pi-1} i \cdot \beta_i + \left( \frac{1 + \mu}{\mu} - \pi \right) \cdot \zeta \right) \cdot \delta(O_1) \end{aligned} \quad (3.10)$$

Observe now that:

$$\pi \cdot \sum_{i=1}^{\pi-1} \beta_i = \pi \cdot (1 - \beta_\pi - \zeta) \cdot \delta(O_1) \quad (3.11)$$

and combine (3.11) with (3.10). Then, the latter becomes:

$$\begin{aligned} |\text{SOL}_4| &\geq \left( \pi - \pi \cdot (1 - \beta_\pi - \zeta) + \sum_{i=1}^{\pi-1} i \cdot \beta_i + \left( \frac{1 + \mu}{\mu} - \pi \right) \cdot \zeta \right) \cdot \delta(O_1) \\ &= \left( \sum_{i=1}^{\pi} i \cdot \beta_i + \pi \zeta + \left( \frac{1 + \mu}{\mu} - \pi \right) \cdot \zeta \right) \cdot \delta(O_1) \\ &= \left( \sum_{i=1}^{\pi} i \cdot \beta_i + \left( \frac{1 + \mu}{\mu} \right) \cdot \zeta \right) \cdot \delta(O_1) \\ &= \left( \sum_{i=1}^{\pi} \beta_i + \sum_{i=2}^{\pi} \beta_i + \sum_{i=3}^{\pi} (i - 2) \cdot \beta_i + \frac{1 + \mu}{\mu} \cdot \zeta \right) \cdot \delta(O_1) \end{aligned}$$

$$\begin{aligned}
 &= \left( (1 - \zeta) + (1 - \beta_1 - \zeta) + \sum_{i=3}^{\pi} (i - 2) \cdot \beta_i + \frac{1 + \mu}{\mu} \cdot \zeta \right) \cdot \delta(O_1) \\
 &= \left( (2 - \beta_1) + \sum_{i=3}^{\pi} (i - 2) \cdot \beta_i + \frac{1 - \mu}{\mu} \cdot \zeta \right) \cdot \delta(O_1) \\
 &\geq \left( (2 - \beta_1) + \sum_{i=3}^{\pi} \beta_i + \frac{1 - \mu}{\mu} \cdot \zeta \right) \cdot \delta(O_1)
 \end{aligned} \tag{3.12}$$

Set  $\sum_{i=3}^{\pi} \beta_i \cdot \delta(O_1) = X$ . These edges are covered by both  $O_1$  and  $\text{SOL}_4$ . Then, (3.12) becomes:

$$|\text{SOL}_4| \geq \left( (2 - \beta_1) + \frac{1 - \mu}{\mu} \cdot \zeta \right) \cdot \delta(O_1) + X = \left( (2 - \beta_1) + \frac{1 - \mu}{\mu} \cdot \zeta \right) \cdot \alpha \cdot \text{opt} + X \tag{3.13}$$

On the other hand, consider Item 2 in  $\text{SOL}_4$ . The  $2 \cdot k_1$  best vertices from  $V_1$  cover  $(1 - \zeta) \cdot \delta(O_1) - X$  edges of the optimum. Let  $Y + [(1 - \zeta) \cdot \delta(O_1) - X]$  be the total number of edges covered by those vertices. Then, best  $k - 2 \cdot k_1$  vertices of  $V_2$  will cover at least as many edges as the  $k - 2 \cdot k_1$  best vertices of  $O_2$ , that will cover at least  $\left(\frac{k - 2 \cdot k_1}{k_2}\right) \cdot \delta'(O_2) - Y$ . Putting all this together, we get:

$$\begin{aligned}
 |\text{SOL}_4| &\geq Y + [(1 - \zeta)\delta(O_1) - X] + \frac{k - 2 \cdot k_1}{k_2} \cdot \delta'(O_2) - Y \\
 &= \left( (1 - \zeta) \cdot \alpha + \frac{k - 2 \cdot k_1}{k_2} \cdot (1 - \alpha) \right) \cdot \text{opt} - X \\
 &\stackrel{(2.1), (2.2)}{=} ((1 - \zeta) \cdot \alpha + (1 - \mu) \cdot (1 - \alpha)) \cdot \text{opt} - X
 \end{aligned} \tag{3.14}$$

Expression (3.13) is increasing with  $X$ , while (3.14) is decreasing. Equality of them, leads after some easy algebra to:

$$X = \left( \frac{(1 - \mu) - \alpha \cdot (2 - \mu) + \beta_1 \cdot \alpha}{2} - \frac{\zeta \cdot \alpha}{2 \cdot \mu} \right) \cdot \text{opt} \tag{3.15}$$

Embedding (3.15)–(3.13) and dividing the ratio obtained by  $\text{opt}$ , derives the ratio claimed by (3.8).  $\square$

The rest of this section is devoted to the proof of the following theorem.

**Theorem 3.7.** BIPARTITE MAX  $k$ -VERTEX COVER is combinatorially approximable within ratio 0.7.

*Proof.* For the proof we propose an exhaustive parameter-elimination method (very probably non-optimal) that has the advantage to be quite simple. It consists in subsequently eliminating parameters from the ratios proved in Lemmas 3.2–3.5 until two ratios that are only functions of  $\mu$  are got. These ratios have opposite monotonicities with respect to this parameter, hence, by equalizing them we determine a lower bound for the overall ratio of the algorithm.

### Elimination of $\theta$ : ratios (3.3) and (3.4)

Equalizing ratios given by (3.3) and (3.4) leads to  $2\alpha \cdot \theta = \alpha - \lambda \cdot (1 - \alpha) \Rightarrow \alpha \cdot \theta = \frac{(\alpha - \lambda \cdot (1 - \alpha))}{2}$  and embedding it in (3.3) derives:

$$\frac{2 - \alpha \cdot (1 - \lambda) - \lambda}{2} \tag{3.16}$$

**Elimination of  $\lambda$ : ratios (3.5) and (3.16)**

Equalizing ratios given by (3.5) and (3.16) gives  $\lambda = \frac{\alpha}{(1-\alpha) \cdot (1+2\mu)}$ . This, together with (3.5), derives:

$$1 - \alpha \cdot \frac{1 + \mu}{1 + 2 \cdot \mu} \quad (3.17)$$

**Elimination of  $\gamma$ : ratios (3.2) and (3.6)**

It gives  $\gamma = \frac{(1-\alpha-\zeta \cdot \alpha)}{(2 \cdot (1-\alpha))}$  and the ratio obtained is:

$$\frac{1 + \alpha - \zeta \cdot \alpha}{2} \quad (3.18)$$

**Elimination of  $\zeta$ : ratios (3.7) and (3.18)**

We have:

$$\begin{aligned} (2 - \beta_1) \cdot \alpha + \frac{(1 - \mu) \cdot \zeta \cdot \alpha}{\mu} &= \frac{1 + \alpha - \zeta \cdot \alpha}{2} \\ \Rightarrow \frac{2 - \mu}{\mu} \cdot \zeta \cdot \alpha &= 1 + \alpha - 2 \cdot (2 - \beta_1) \cdot \alpha = 1 - 3 \cdot \alpha + 2 \cdot \beta_1 \cdot \alpha \\ \Rightarrow \zeta &= \frac{\mu}{2 - \mu} \cdot \frac{1 - 3 \cdot \alpha + 2 \cdot \beta_1 \cdot \alpha}{\alpha} \end{aligned} \quad (3.19)$$

and embedding (3.19) in (3.7), we get:

$$\begin{aligned} (2 - \beta_1) \cdot \alpha + \frac{1 - \mu}{2 - \mu} \cdot (1 - \alpha \cdot (3 - 2 \cdot \beta_1)) \\ = \frac{(2 - \mu) \cdot (2 - \beta_1) \cdot \alpha + (1 - \mu) - \alpha \cdot (1 - \mu) \cdot (3 - 2 \cdot \beta_1)}{2 - \mu} \\ = \frac{(1 - \mu) + \alpha \cdot (1 + \mu - \mu \cdot \beta_1)}{2 - \mu} \end{aligned} \quad (3.20)$$

**First elimination of  $\beta_1$ : ratios (3.1) and (3.20)**

We have:

$$\begin{aligned} 1 - \alpha + \beta_1 \cdot \alpha &= \frac{(1 - \mu) + \alpha \cdot (1 + \mu - \mu \cdot \beta_1)}{2 - \mu} \\ \Rightarrow 2 \cdot \beta_1 \cdot \alpha &= (1 - \mu) + \alpha \cdot (1 + \mu) - (2 - \mu) + \alpha \cdot (2 - \mu) = -1 + 3 \cdot \alpha \\ \Rightarrow \beta_1 &= \frac{3 \cdot \alpha - 1}{2 \cdot \alpha} \end{aligned} \quad (3.21)$$

Now, combination of (3.1) and (3.21) derives:

$$1 - \alpha + \beta_1 \cdot \alpha = 1 - \alpha + \frac{3 \cdot \alpha - 1}{2} = \frac{1 + \alpha}{2} \quad (3.22)$$



**First ratio function of  $\mu$ : combination of ratios (3.17) and (3.22)**

Ratio (3.17) is decreasing with  $\alpha$ , while ratio (3.22) is increasing. Combination of them allows elimination of  $\alpha$  in order that a first ratio that is only a function of  $\mu$  is derived. Equalizing (3.17) and (3.22) gives:

$$\begin{aligned} 1 - \alpha \cdot \frac{1 + \mu}{1 + 2 \cdot \mu} &= \frac{1 + \alpha}{2} \Rightarrow \alpha \cdot (4 \cdot \mu + 3) = 2 \cdot \mu + 1 \\ \Rightarrow \alpha &= \frac{2 \cdot \mu + 1}{4 \cdot \mu + 3} \end{aligned} \quad (3.23)$$

and embedding (3.23) in (3.17) derives:

$$\begin{aligned} 1 - \alpha \cdot \frac{1 + \mu}{1 + 2 \cdot \mu} &= 1 - \frac{2 \cdot \mu + 1}{4 \cdot \mu + 3} \cdot \frac{1 + \mu}{1 + 2 \cdot \mu} \\ &= \frac{2 + 3 \cdot \mu}{3 + 4 \cdot \mu} \end{aligned} \quad (3.24)$$

**Second elimination of  $\beta_1$ : ratios (3.1) and (3.8)**

Revisit ratio (3.8) and observe that its last term  $\frac{((1-2\cdot\mu)\cdot\zeta\cdot\alpha)}{(2\cdot\mu)}$  is negative when  $\mu \geq \frac{1}{2}$ . On the other hand, ratio (3.24) is increasing with  $\mu$  and bounded below by 0.7 as long as  $\mu \geq \frac{1}{2}$ . We thus seek an “interesting” ratio when  $\mu \leq \frac{1}{2}$  and, in this case  $\frac{((1-2\cdot\mu)\cdot\zeta\cdot\alpha)}{(2\cdot\mu)} \geq 0$  and can be omitted.

Hence, combination of ratios (3.1) and (3.8), for  $\mu \leq \frac{1}{2}$ , leads to:

$$\begin{aligned} \frac{(1 - \mu) + \alpha \cdot (2 + \mu) - \alpha \cdot \beta_1}{2} &= 1 - \alpha + \beta_1 \cdot \alpha \\ \Rightarrow (1 - \mu) + \alpha \cdot (2 + \mu) - \alpha \cdot \beta_1 &= 2 - 2 \cdot \alpha + 2 \cdot \beta_1 \cdot \alpha \\ \Rightarrow 3 \cdot \beta_1 \cdot \alpha &= -(1 + \mu) + \alpha \cdot (4 + \mu) \Rightarrow \beta_1 = \frac{\alpha \cdot (4 + \mu) - (1 + \mu)}{3 \cdot \alpha} \end{aligned} \quad (3.25)$$

Then, combining (3.1) and (3.25), derives this time:

$$1 - \alpha + \beta_1 \cdot \alpha = \frac{3 - 3 \cdot \alpha + \alpha \cdot (4 + \mu) - (1 + \mu)}{3} = \frac{2 - \mu + \alpha \cdot (1 + \mu)}{3} \quad (3.26)$$

**Second ratio function of  $\mu$ : combination of ratios (3.26) and (3.17)**

Once again, ratio (3.17) is decreasing with  $\alpha$ , while ratio (3.26) is increasing. Combination of them allows elimination  $\alpha$  in order to get a second ratio exclusively function of  $\mu$ . Equalizing (3.17) and (3.26) gives:

$$\begin{aligned} 1 - \alpha \cdot \frac{1 + \mu}{1 + 2 \cdot \mu} &= \frac{2 - \mu + \alpha \cdot (1 + \mu)}{3} \Rightarrow 2 \cdot \alpha \cdot \frac{(1 + \mu) \cdot (2 + \mu)}{3 \cdot (1 + 2 \cdot \mu)} = \frac{1 + \mu}{3} \\ \Rightarrow \alpha &= \frac{1 + 2 \cdot \mu}{2 \cdot (2 + \mu)} \end{aligned} \quad (3.27)$$

and embedding (3.27) in (3.17) derives:

$$1 - \alpha \cdot \frac{1 + \mu}{1 + 2 \cdot \mu} = \frac{3 + \mu}{4 + 2 \cdot \mu} \quad (3.28)$$

## Final ratio

As noted above, ratio (3.24) increases with  $\mu$ , while (3.28) decreases. The value of  $\mu$  guaranteeing equality of these ratios also gives a lower bound for them. This value is  $\mu = \frac{1}{2}$  and, with this value, both ratios become 0.7.  $\square$

## 4. A FINAL REMARK

For reasons of simplicity, the proposed algorithm has been analyzed in the unweighted case of BIPARTITE MAX  $k$ -VERTEX COVER. It is easy to see that exactly the same analysis can be done when dealing with weighted BIPARTITE MAX  $k$ -VERTEX COVER where the edges of the input graph are weighted and the objective becomes to determine the  $k$  vertices that maximize the total weight of the edges covered. In this case, it suffices to consider that quantities  $\delta$  and  $\delta'$  defined in Section 2 do not express the number of edges covered, but rather the total weight of the edges covered. The rest of the analysis remain then exactly the same.

*Acknowledgements.* The very fruitful and stimulating discussions the author had with Edouard Bonnet and Georgios Stamoulis are gratefully acknowledged. Many thanks to Edouard Bonnet for his very pertinent comments on a first draft of the paper.

The very pertinent and useful suggestions and comments of an anonymous reviewer are also gratefully acknowledged.

## REFERENCES

- [1] A.A. Ageev and M. Sviridenko, Approximation algorithms for maximum coverage and max cut with given sizes of parts, in *Proc. Conference on Integer Programming and Combinatorial Optimization, IPCO'99*, edited by G. Cornuéjols, R.E. Burkard and G.J. Woeginger. Vol. 1610 of *Lecture Notes in Computer Science*. Springer-Verlag (1999) 17–30.
- [2] N. Apollonio and B. Simeone, The maximum vertex coverage problem on bipartite graphs. *Discrete Appl. Math.* **165** (2014) 37–48.
- [3] A. Badanidiyuru, R. Kleinberg and H. Lee, Approximating low-dimensional coverage problems, in *Proc. Symposium on Computational Geometry, SoCG'12*, edited by T.K. Dey and S. Whitesides. ACM, Chapel Hill, NC (2012) 161–170.
- [4] B. Caskurlu, V. Mkrtchyan, O. Parekh and K. Subramani, On partial vertex cover and budgeted maximum coverage problems in bipartite graphs, in *Proc. Theoretical Computer Science, IFIP TC 1/WG 2.2 International Conference, TCS'14*, edited by J. Diaz, I. Lanese and D. Sangiorgi. Vol. 8705 of *Lecture Notes in Computer Science*. Springer-Verlag (2014) 13–26.
- [5] G. Cornuejols, M.L. Fisher and G.L. Nemhauser, Location of bank accounts to optimize float: an analytic study of exact and approximate algorithms. *Manag. Sci.* **23** (1977) 789–810.
- [6] D.S. Hochbaum and A. Pathria, Analysis of the greedy approach in problems of maximum  $k$ -coverage. *Naval Res. Logist.* **45** (1998) 615–627.
- [7] E. Petrank, The hardness of approximation: gap location. *Comput. Complex.* **4** (1994) 133–157.
- [8] L. Trevisan, Max cut and the smallest eigenvalue, in *Proc. STOC'09* (2009) 263–272.