



HAL
open science

The labeled perfect matching in bipartite graphs

Jérôme Monnot

► **To cite this version:**

| Jérôme Monnot. The labeled perfect matching in bipartite graphs. 2006. <hal-00116834>

HAL Id: hal-00116834

<https://hal.science/hal-00116834v1>

Preprint submitted on 28 Nov 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

The Labeled perfect matching in bipartite graphs

Jérôme Monnot*

Abstract

In this paper, we deal with both the complexity and the approximability of the labeled perfect matching problem in bipartite graphs. Given a simple graph $G = (V, E)$ with $|V| = 2n$ vertices such that E contains a perfect matching (of size n), together with a color (or label) function $L : E \rightarrow \{c_1, \dots, c_q\}$, the labeled perfect matching problem consists in finding a perfect matching on G that uses a minimum or a maximum number of colors.

Key words : Labeled matching, bipartite graphs, **NP**-complete, approximate algorithms.

1 Introduction

Let Π be a **NPO** problem accepting simple graphs $G = (V, E)$ as instances, edge-subsets $E' \subseteq E$ verifying a given polynomial-time decidable property $Pred$ as solutions, and the solutions cardinality as objective function; the labeled problem associated to Π , denoted by **LABELED** Π , seeks, given an instance $I = (G, L)$ where $G = (V, E)$ is a simple graph and L is a mapping from E to $\{c_1, \dots, c_q\}$, in finding a subset E' verifying $Pred$ that optimizes the size of the set $L(E') = \{L(e) : e \in E'\}$. Note that two versions of **LABELED** Π may be considered according to the optimization goal: **LABELED Min** Π that consists in minimizing $|L(E')|$ and **LABELED Max** Π that consists in maximizing $|L(E')|$. Roughly speaking, the mapping L corresponds to assigning a color (or a label) to each edge and the goal of **LABELED Min** Π (resp., **Max** Π) is to find an edge subset using the fewest (resp., the most) number of colors. If a given **NPO** problem Π is **NP**-hard, then the associated labeled problem **LABELED** Π is clearly **NP**-hard (consider a distinct color

*LAMSADE, Université Paris-Dauphine, place du Maréchal de Lattre de Tassigny, 75775 Paris cedex 16, France. monnot@lamsade.dauphine.fr

per edge). For instance, the LABELED Longest path problem or the LABELED maximum induced matching problem are both **NP**-hard. Moreover, if the decision problem associated to Π is **NP**-complete, (the decision problem aims at deciding if a graph G contains an edge subset verifying $Pred$), then LABELED *Min* Π can not be approximated within performance ratio better than $2 - \varepsilon$ for all $\varepsilon > 0$ unless $\mathbf{P}=\mathbf{NP}$, even if the graph is complete. Indeed, if we color the edges from $G = (V, E)$ with a lonely color and then we complete the graph, adding a new color per edge, then it is **NP**-complete to decide between $opt(I) = 1$ and $opt(I) \geq 2$, where $opt(I)$ is the value of an optimal solution. Notably, it is the case of the LABELED traveling salesman problem or the LABELED minimum partition problem into paths of length k for any $k \geq 2$.

Thus, labeled problems have been mainly studied, from a complexity and an approximability point of view, when Π is polynomial, [5, 6, 7, 9, 14, 18, 19]. For example, the first labeled problem introduced in the literature is the LABELED minimum spanning tree problem, which has several applications in communication network design. This problem is **NP**-hard and many complexity and approximability results have been proposed in [5, 7, 9, 14, 18, 19]. On the other hand, the LABELED maximum spanning tree problem has been shown polynomial in [5]. Very recently, the LABELED path and the LABELED cycle problems have been studied in [6]; in particular, the authors prove that the LABELED minimum path problem is **NP**-hard.

In this paper, we go thoroughly into the investigation of the complexity and the approximability of labeled problems, with the analysis of the matching problem in bipartite graphs. The maximum matching problem is one of the most known combinatorial optimization problem and arises in several applications such as images analysis, artificial intelligence or scheduling. It turns out that a problem very closed to it has been studied in the literature, which is called in [12] the restricted perfect matching problem. This latter aims at determining, given a graph $G = (V, E)$, a partition E_1, \dots, E_k of E and k positive integers r_1, \dots, r_k , weather there exists a perfect matching M on G satisfying for all $j = 1, \dots, k$ the restrictions $|M \cap E_j| \leq r_j$. This problem has some relationship with the timetable problem, since a solution may be seen as a matching between classes and teachers that satisfies additional restrictions (for instance, no more that r labs at the same time). The restricted perfect matching problem is proved to be **NP**-complete in [12], even if (i) $|E_j| \leq 2$, (ii) $r_j = 1$, and (iii) G is a bipartite graph. On the other hand, it is shown in [20] that the restricted perfect matching problem is polynomial when G is a complete bipartite graph and $k = 2$. A perfect matching M only verifying condition (ii) (that is to say $|M \cap E_i| \leq 1$) is called good in [8]. Thus, we deduce that the LABELED maximum perfect matching problem is **NP**-hard in bipartite graph since $opt(I) = n$ iff G contains a good matching.

In section 2, we analyze both the complexity and the approximability of the LABELED minimum perfect matching problem and the LABELED maximum perfect matching problem in 2-regular bipartite graphs. Finally, section 3 focuses on the case of complete

bipartite graphs.

Now, we introduce some terminology and notations that will be used in the paper. A matching M on a graph $G = (V, E)$ is a subset of edges that are pairwise non adjacent; M is said perfect if it covers the vertex set of G . In the labeled perfect matching problem (LBELED PM in short), we are given a simple graph $G = (V, E)$ on $|V| = 2n$ vertices which contains a perfect matching together with a color (or label) function $L : E \rightarrow \{c_1, \dots, c_q\}$ on the edge set of G . For $i = 1, \dots, q$, we denote by $L^{-1}(\{c_i\}) \subseteq E$ the set of edges of color c_i . The goal of LBELED $Min PM$ (resp., $Max PM$) is to find a perfect matching on G using a minimum (resp., a maximum) number of colors. An equivalent formulation of LBELED $Min PM$ could be the following: if $G[\mathcal{C}]$ denotes the subgraph induced by the edges of colors $\mathcal{C} \subseteq \{c_1, \dots, c_q\}$, then LBELED $Min PM$ aims at finding a subset \mathcal{C} of minimum size such that $G[\mathcal{C}]$ contains a perfect matching. The restriction of LBELED PM to the case where each color occurs at most r times in $I = (G, L)$ (i.e., $|L^{-1}(\{c_i\})| \leq r$ for $i = 1, \dots, q$) will be denoted by LBELED PM_r .

We denote by $opt(I)$ and $apx(I)$ the value of an optimal and an approximate solution, respectively. We say that an algorithm \mathcal{A} is an ε -approximation of LBELED $Min PM$ with $\varepsilon \geq 1$ (resp., $Max PM$ with $\varepsilon \leq 1$) if $apx(I) \leq \varepsilon \times opt(I)$ (resp., $apx(I) \geq \varepsilon \times opt(I)$) for any instance $I = (G, L)$.

2 The 2-regular bipartite case

In this section, we deal with a particular class of graphs that consist in a collection of pairwise disjoint cycles of even length; note that such graphs are 2-regular bipartite graphs.

Theorem 2.1 LBELED $Min PM_r$ is **APX**-complete in 2-regular bipartite graphs for any $r \geq 2$.

Proof. Observe that any solution of LBELED $Min PM_r$ is an r -approximation. The rest of the proof will be done via an approximation preserving reduction from the minimum balanced satisfiability problem with clauses of size at most r , MIN BALANCED r -SAT for short. An instance $I = (\mathcal{C}, X)$ of MIN BALANCED r -SAT consists in a collection $\mathcal{C} = (C_1, \dots, C_m)$ of clauses over the set $X = \{x_1, \dots, x_n\}$ of boolean variables, such that each clause C_j has at most r literals and each variable appears positively as many time as negatively; let B_i denotes this number for any $i = 1, \dots, n$. The goal is to find a truth assignment f satisfying a minimum number of clauses. MIN BALANCED 2-SAT where $2 \leq B_i \leq 3$ has been shown **APX**-complete by the way of an L -reduction from MAX BALANCED 2-SAT where $B_i = 3$, [4, 13].

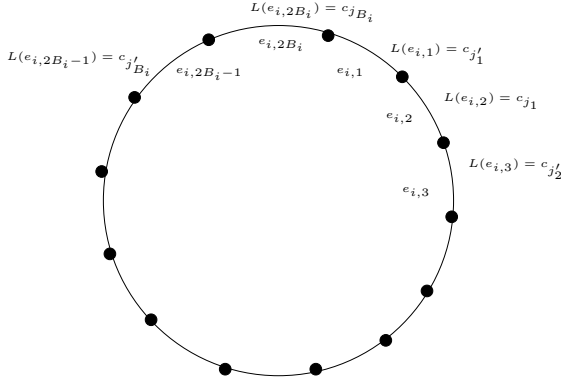


Figure 1: The gadget $H(x_i)$ and the color of its edges.

We only prove the case $r = 2$. Let $I = (\mathcal{C}, X)$ be an instance of MIN BALANCED 2-SAT on m clauses $\mathcal{C} = \{C_1, \dots, C_m\}$ and n variables $X = \{x_1, \dots, x_n\}$ such that each variable x_i has either 2 occurrences positive and 2 occurrences negative, or 3 occurrences positive and 3 occurrences negative. We build the instance $I' = (H, L)$ of LABELED *Min PM*₂ where H is a collection of pairwise disjoint cycles $\{H(x_1), \dots, H(x_n)\}$ and L colors edges of H with colors $c_1, \dots, c_j, \dots, c_m$, by applying the following process:

- For each variable x_i , create $2B_i$ -long cycle $H(x_i) = \{e_{i,1}, \dots, e_{i,k}, \dots, e_{i,2B_i}\}$.
- Color the edges of $H(x_i)$ as follows: if x_i appears positively in clauses $C_{j_1}, \dots, C_{j_{B_i}}$ and negatively in clauses $C_{j'_1}, \dots, C_{j'_{B_i}}$, then set $L(e_{i,2k}) = c_{j_k}$ and $L(e_{i,2k-1}) = c_{j'_k}$ for $k = 1, \dots, B_i$.

Figure 1 provides an illustration of the gadget $H(x_i)$. Clearly, H is made of n disjoint cycles and is painted with m colors. Moreover, each color appears exactly twice.

Let f^* be an optimal truth assignment on I satisfying m^* clauses and consider the perfect matching $M = \cup_{i=1}^n M_i$ where $M_i = \{e_{i,2k} | k = 1, \dots, B_i\}$ if $f(x_i) = \text{true}$, $M_i = \{e_{i,2k-1} | k = 1, \dots, B_i\}$ otherwise; M uses exactly m^* colors and thus:

$$\text{opt}(I) \leq m^* \tag{1}$$

Conversely, let M' be a perfect matching on H using $\text{apx}(I) = m'$ colors; if one sets $f'(x_i) = \text{true}$ if $e_{i,2} \in M'$, $f'(x_i) = \text{false}$ otherwise, we can easily observe that the truth assignment f' satisfies m' clauses.

$$apx(I) = val(f') \tag{2}$$

Hence, using inequalities (1) and (2) the result follows.

Trivially, the problem becomes obvious when each color is used exactly once. We now show that we have a 2-approximation in 2-regular bipartite graphs, showing that the restriction of LABELED *Min PM* to 2-regular bipartite graphs is as hard as approximate as MINSAT.

Theorem 2.2 *There exists an approximation preserving reduction from LABELED *Min PM* in 2-regular bipartite graphs to MINSAT of expansion $c(\varepsilon) = \varepsilon$.*

Proof. The result comes from the reciprocal of the previous transformation. Let $I = (G, L)$ be an instance of LABELED *Min PM* where $G = (V, E)$ is a collection $\{H_1, \dots, H_n\}$ of disjoint cycles of even length and $L(E) = \{c_1, \dots, c_m\}$ defines the label set, we describe every cycle H_i as the union of two matchings M_i and \overline{M}_i . We construct an instance $I' = (\mathcal{C}, X)$ of MINSAT where $\mathcal{C} = \{C_1, \dots, C_m\}$ is a set of m clauses and $X = \{x_1, \dots, x_n\}$ is a set of n variables, as follows. The clause set \mathcal{C} is in one to one correspondence with the color set $L(E)$ and the variable set X is in one to one correspondence with the connected components of G ; a literal x_i (resp., \overline{x}_i) appears in C_j iff $c_j \in L(M_i)$ (resp., $c_j \in L(\overline{M}_i)$). We easily deduce that any truth assignment f on I' that satisfies k clauses can be converted into a perfect matching M_f on I that uses k colors.

Using the 2-approximation of MINSAT [15] and the Theorem 2.2, we deduce:

Corollary 2.3 *LABELED *Min PM* in 2-regular bipartite graphs is 2-approximable.*

Dealing with LABELED *Max PM_r*, the result of [12] shows that computing a *good matching* is NP-hard even if the graph is bipartite and each color appears at most twice; a good matching M is a perfect matching using $|M|$ colors. Thus, we deduce from this result that LABELED *Max PM_r* is NP-hard for any $r \geq 2$. We strengthen this result using a reduction from MAX BALANCED 2-SAT.

Theorem 2.4 *LABELED *Max PM_r* is APX-complete in collection of pairwise disjoint cycles of even length for any $r \geq 2$.*

In the same way, using the approximate result for MAXSAT [2], we obtain

Corollary 2.5 *LABELED *Max PM* in 2-regular bipartite graphs is 0.7846-approximable.*

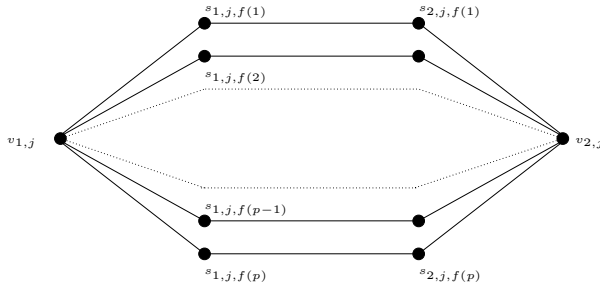


Figure 2: The gadget $H(x_j)$.

3 The complete bipartite case

When considering complete bipartite graphs, we obtain several results:

Theorem 3.1 *LABELLED $Min PM_r$ is APX-complete in bipartite complete graphs $K_{n,n}$ for any $r \geq 6$.*

Proof. We give an approximation preserving L -reduction (cf. Papadimitriou & Yannakakis [16]) from the set cover problem, MINSC for short. Given a family $\mathcal{S} = \{S_1, \dots, S_{n_0}\}$ of subsets and a ground set $X = \{x_1, \dots, x_{m_0}\}$ (we assume that $\cup_{i=1}^{n_0} S_i = X$), a set cover of X is a sub-family $\mathcal{S}' = \{S_{f(1)}, \dots, S_{f(p)}\} \subseteq \mathcal{S}$ such that $\cup_{i=1}^p S_{f(i)} = X$; MINSC is the problem of determining a minimum-size set cover $\mathcal{S}^* = \{S_{f^*(1)}, \dots, S_{f^*(q)}\}$ of X . Its restriction MINSC₃ to instances where each set is of size at most 3 and each element x_j appears in at most 3 different sets has been proved APX-complete in [16].

Given an instance $I_0 = (\mathcal{S}, X)$ of MINSC, its characteristic graph $G_{I_0} = (L_0, R_0; E_{I_0})$ is a bipartite graph with a left set $L_0 = \{l_1, \dots, l_{n_0}\}$ that represents the members of the family \mathcal{S} and a right set $R_0 = \{r_1, \dots, r_{m_0}\}$ that represents the elements of the ground set X ; the edge-set E_{I_0} of the characteristic graph is defined by $E_{I_0} = \{[l_i, r_j] : x_j \in S_i\}$. Note that G_{I_0} is of maximum degree 3 iff I_0 is an instance of MINSC₃. From I_0 , we construct the instance $I = (K_{n,n}, L)$ of LABELLED $Min PM_6$. First, we start from a bipartite graph having m_0 connected components $H(x_j)$ and $n_0 + m_0$ colors $\{c_1, \dots, c_{n_0+m_0}\}$, described as follows:

- For each element $x_j \in X$, we build a gadget $H(x_j)$ that consists in a bipartite graph of $2(d_{G_{I_0}}(r_j) + 1)$ vertices and $3d_{G_{I_0}}(r_j)$ edges, where $d_{G_{I_0}}(r_j)$ denotes the degree of vertex $r_j \in R$ in G_{I_0} . The graph $H(x_j)$ is illustrated in Figure 2.

- Assume that vertices $\{l_{f(1)}, \dots, l_{f(p)}\}$ are the neighbors of r_j in G_{I_0} , then color $H(x_j)$ as follows: for any $k = 1, \dots, p$, $L(v_{1,j}, s_{1,j,f(k)}) = L(v_{2,j}, s_{2,j,f(k)}) = c_{f(k)}$ and $L(s_{1,j,f(k)}, s_{2,j,f(k)}) = c_{n_0+j}$.
- We complete $H = \cup_{x_j \in X} H(x_j)$ into $K_{n,n}$, by adding a new color per edge.

Clearly, $K_{n,n}$ is complete bipartite and has $n = 2 \sum_{r_j \in R} (d_{G_{I_0}}(r_j) + 1) = 2|E_{I_0}| + 2m_0$ vertices. Moreover, each color is used at most 6 times.

Let \mathcal{S}^* be an optimal set cover on I_0 . From \mathcal{S}^* , we can easily construct a perfect matching M^* on I using exactly $|\mathcal{S}^*| + m_0$ colors and thus:

$$opt_{\text{Labeled Min PM}_6}(I) \leq opt_{\text{MinSC}_3}(I_0) + m_0 \tag{3}$$

Conversely, we show that any perfect matching M may be transformed into a perfect matching M'' using the edges of H and verifying: $|L(M'')| \leq |L(M)|$. Let M be a perfect matching on I and consider M_1 the subset of edges from M that link two different gadgets $H(x_j)$; we denote by G the multi-graph of vertex set $\cup_j V(H(x_j))$ and of edge set M_1 . Remark that each connected component of G is eulerian. Each cycle C on G may be completed into a $2|C|$ -long cycle C' on $K_{n,n}$ in such a way that the two endpoints of each edge from $C' \setminus C$ do belong to the same gadget $H(x_j)$. If one swaps the edges from each cycle C by the edges from $C' \setminus C$, we obtain a new perfect matching M' of which every edge has its two endpoints in a same gadget $H(x_j)$ and that verifies $|L(M')| = |L(M)|$. Now consider for any j the set M'_j of edges from $M' \cap H(x_j)$, we set $M''_j = \{[v_{1,j}, s_{1,j,f(k)}], [v_{2,j}, s_{2,j,f(k)}]\} \cup \{[s_{1,j,f(i)}, s_{2,j,f(i)}] | i = 1, \dots, p\}$ for some k such that $[v_{1,j}, s_{1,j,f(k)}] \in M'_j$ or $[v_{2,j}, s_{2,j,f(k)}] \in M'_j$ (if there does not exist such a k , set $k = 1$). In any case, $M'' = (M' \setminus M'_j) \cup M''_j$ is a perfect matching that uses no more colors than M' does. Applying this procedure for any $j = 1, \dots, m_0$, we obtain the expected matching M'' with value $apx(I)$. From such a matching, we may obtain a set cover $\mathcal{S}'' = \{S_k | c_k \in L(M'')\}$ on I_0 verifying:

$$|\mathcal{S}''| = apx(I) - m_0 \tag{4}$$

Using (3) and (4), we deduce $opt_{\text{Labeled Min PM}_6}(I) = opt_{\text{MinSC}_3}(I_0) + m_0$ and $|\mathcal{S}''| - opt_{\text{MinSC}_3}(I_0) \leq |L(M'')| - opt_{\text{Labeled Min PM}_6}(I)$. Finally, since $opt_{\text{MinSC}_3}(I_0) \geq \frac{m_0}{3}$ the result follows.

Applying the same kind of proof to the vertex cover problem in cubic graphs [1], we obtain that **Labeled Min PM_r** in $K_{n,n}$ is **APX**-complete for any $r \geq 3$. In order to establish this fact and starting from a cubic graph $G = (V, E)$, we associate to each edge $e = [x, y] \in E$ a 4-long cycle $\{a_{1,e}, a_{2,e}, a_{3,e}, a_{4,e}\}$ together with a coloration L given by: $L(a_{1,e}) = c_x$, $L(a_{2,e}) = c_y$ and $L(a_{3,e}) = L(a_{4,e}) = c_e$. We complete this graph

into a complete bipartite graph, adding a new color per edge. Each color c_x ($\forall x \in V$) appears 3 times, c_e ($\forall e \in E$) twice and any other color, once. Hence, the application of the proof that was made in Theorem 3.1 leads to the announced result. Unfortunately, we can not apply the proof of Theorem 2.2 since in this latter, on the one hand, we have some cycles of size 6 and, on the other hand, a color may occurs in different gadgets. One open question concerns the complexity of Labeled $Min PM_2$ in bipartite complete graphs. Moreover, from Theorem 3.1, we can also obtain a stronger inapproximability result: one can not compute in polynomial-time an approximate solution that uses less than $(1/2 - \varepsilon)\ln(opt_{\text{Labeled } Max PM}(I))$ colors in complete bipartite graphs where $opt_{\text{Labeled } Max PM}(I)$ is the value of an optimal solution of Labeled $Max PM$, i.e., the maximum number of colors used by a perfect matching.

Corollary 3.2 *For any $\varepsilon > 0$, Labeled $Min PM$ is not $(\frac{1}{2} - \varepsilon) \times \ln(n)$ approximable in complete bipartite graphs $K_{n,n}$, unless $\mathbf{NP} \subset \mathbf{DTIME}(n^{\log \log n})$.*

Proof. First, we apply the construction made in Theorem 3.1, except that $I_0 = (S, X)$ is an instance of MINSC such that the number of elements m_0 is strictly larger than the number of sets n_0 . From I_0 , we construct n_0 instances I'_1, \dots, I'_{n_0} of Labeled $Min PM$ where $I'_i = (H, L_i)$. The colors $L_i(E)$ are the same than $L(E)$, except that we replace colors $c_{n_0+1}, \dots, c_{n_0+m_0}$ by c_i .

Let \mathcal{S}^* be an optimal set cover on I_0 and assume that $S_i \in \mathcal{S}^*$, we consider the instance I_i of Labeled $Min PM$. From \mathcal{S}^* , we can easily construct a perfect matching M_i^* of I_i that uses exactly $|\mathcal{S}^*|$ colors. Conversely, let M_i be a perfect matching on I_i ; by construction, the subset $\mathcal{S}' = \{S_k : c_k \in L(M_i)\}$ of \mathcal{S} is a set cover of X using $|L(M_i)|$ sets. Finally, let A be an approximate algorithm for Labeled $Min PM$, we compute n_0 perfect matchings M_i , applying A on instances I_i . Thus, if we pick the matching that uses the minimum number of colors, then we can polynomially construct a set cover on I_0 of cardinality this number of colors.

Since $n_0 \leq m_0 - 1$, the size n of a perfect matching of $K_{n,n}$ verifies: $n = |E_{I_0}| + m_0 \leq n_0 \times m_0 + m_0 \leq m_0(m_0 - 1) + m_0 = m_0^2$. Hence, from any algorithm A solving Labeled $Min PM$ within a performance ratio $\rho_A(I) \leq \frac{1}{2} \times \ln(n)$, we can deduce an algorithm for MINSC that guarantees the performance ratio $\frac{1}{2} \ln(n) \leq \frac{1}{2} \ln(m_0^2) = \ln(m_0)$. Since the negative result of [10] holds when $n_0 \leq m_0 - 1$, i.e., MINSC is not $(1 - \varepsilon) \times \ln(m_0)$ approximable for any $\varepsilon > 0$, unless $\mathbf{NP} \subset \mathbf{DTIME}(n^{\log \log n})$, we obtain a contradiction.

On the other hand, dealing with Labeled $Max PM_r$ in $K_{n,n}$, the result of [8] shows that the case $r = 2$ is polynomial, whereas it becomes \mathbf{NP} -hard when $r = \Omega(n^2)$. Indeed, it is proved in [8] that, on the one hand, we can compute a good matching in $K_{n,n}$ within polynomial-time when each color appears at most twice and, on the other hand, there always exists a good matching in such a graph if $n \geq 3$. An interesting question is

to decide the complexity and the approximability of LABELED $Max PM_r$ when r is a constant greater than 2.

3.1 Approximation algorithm for LABELED $Min PM_r$

Let us consider the greedy algorithm for LABELED $Min PM_r$ in complete bipartite graphs that iteratively picks the color that induces the maximum-size matching in the current graph and delete the corresponding vertices. Formally, if $L(G')$ denotes the colors that are still available in the graph G' at a given iteration and if $G'[c]$ (resp., $G'[V']$) denotes the subgraph of G' that is induced by the edges of color c (resp., by the vertices V'), then the greedy algorithm consists in the following process:

Greedy

- 1 Set $C' = \emptyset$, $V' = V$ and $G' = G$;
 - 2 While $V' \neq \emptyset$ do
 - 2.1 For any $c \in L(G')$, compute a maximum matching M_c in $G'[c]$;
 - 2.2 Select a color c^* maximizing $|M_{c^*}|$;
 - 2.3 $C' \leftarrow C' \cup \{c^*\}$, $V' \leftarrow V' \setminus V(M_{c^*})$ and $G' = G[V']$;
 - 3 output C' ;
-

Theorem 3.3 *Greedy is an $\frac{H_r+r}{2}$ -approximation of LABELED $Min PM_r$ in complete bipartite graphs where H_r is the r -th harmonic number $H_r = \sum_{i=1}^r \frac{1}{i}$, and this ratio is tight.*

Proof. Let $I = (G, L)$ be an instance of LABELED $Min PM_r$, we denote by C'_i for $i = 1, \dots, r$ be the set of colors of the approximate solution which appears exactly i times in C' and by p_i its cardinality; finally, let M_i denote the matching with colors C'_i . If $apx(I) = |C'|$, then we have:

$$apx(I) = \sum_{i=1}^r p_i \tag{5}$$

Let \mathcal{C}^* be an optimal solution corresponding to the perfect matching M^* of size $opt(I) = |\mathcal{C}^*|$; we denote by E_i the set of edges from M^* that belong to $G[\cup_{k=1}^i V(M_k)]$, the subgraph induced by $\cup_{k=1}^i V(M_k)$ and we set $q_i = |E_i \setminus E_{i-1}|$ (where we assume that $E_0 = \emptyset$). For any $i = 1, \dots, r - 1$, we get:

$$opt(I) \geq \frac{1}{i} \sum_{k=1}^i q_k \tag{6}$$

Indeed, $\sum_{k=1}^i q_k = |E_i|$ and by construction, each color appears at most i times in $G[\cup_{k=1}^i V(M_k)]$.

We also have the following inequality for any $i = 1, \dots, r - 1$:

$$opt(I) \geq \frac{1}{r} \left(2 \sum_{k=1}^i k \times p_k - \sum_{k=1}^i q_k \right) \tag{7}$$

Since M^* is a perfect matching, the quantity $2 \sum_{k=1}^i k \times p_k - \sum_{k=1}^i q_k$ counts the edges of M^* of which at least one endpoint belongs to $G[\cup_{k=1}^i V(M_k)]$. Because each color appears on at most r edges, the result follows.

Finally, since $\sum_{k=1}^r k \times p_k$ is the size of a perfect matching of G , the following inequality holds:

$$opt(I) \geq \frac{1}{r} \sum_{k=1}^r k \times p_k \tag{8}$$

Using equality (5) and adding inequalities (6) with coefficient $\alpha_i = \frac{1}{2(i+1)}$ for $i = 1, \dots, r - 1$, inequalities (7) with coefficient $\beta_i = \frac{r}{2i(i+1)}$ for $i = 1, \dots, r - 1$ and inequality (8), we obtain:

$$apx(I) \leq \left(\frac{H_r + r}{2} \right) opt(I) \tag{9}$$

Indeed, $\sum_{i=1}^{r-1} \alpha_i = \frac{1}{2}H_r - \frac{1}{2}$ and $\sum_{i=1}^{r-1} \beta_i = \frac{r}{2} - \frac{1}{2}$. Thus, $\sum_{i=1}^{r-1} (\alpha_i + \beta_i) + 1 = \frac{H_r+r}{2}$.

The quantity p_j appears in inequality (8) and inequality (7) for $i = j, \dots, r - 1$. Its total contribution is: $\frac{1}{r}j \times p_j + \frac{2}{r} \left(\sum_{i=j}^{r-1} \beta_i \right) j \times p_j = p_j$. The quantity q_j appears in inequality (6) for $i = j, \dots, r - 1$ and inequality (7) for $i = j, \dots, r - 1$. We have: $\left(\sum_{i=j}^{r-1} \frac{\alpha_i}{i} \right) - \frac{1}{r} \left(\sum_{i=j}^{r-1} \beta_i \right) q_j = 0$. Thus, using equality (5), the inequality (9) holds.

In order to show the tightness of this bound, consider the instance $I = (K_{n,n}, L)$ where the left set A and the right set B of vertices of the complete bipartite graph are given by $A = \{a_{i,j} : i = 1, \dots, r, j = 1, \dots, n_i\}$ and $B = \{b_{i,j} : i = 1, \dots, r, j = 1, \dots, n_i\}$, with $n_1 = (r + 1)!$ and $n_i = r!$ for $i = 2, \dots, r$. Moreover, the edge coloration verifies:

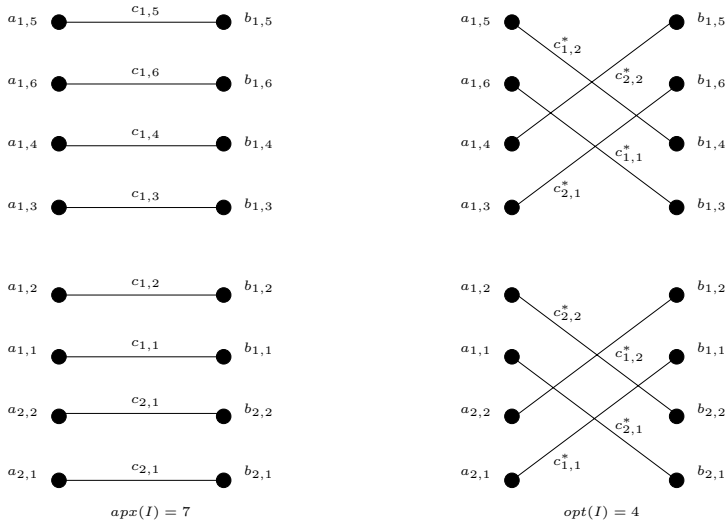


Figure 3: The instance I when $r = 2$.

- For any $i = 1, \dots, r$ and for any $j = 1, \dots, n_i$, $L(a_{i,j}, b_{i,j}) = c_{i, \lceil \frac{j}{r} \rceil}$.
- For any $i = 2, \dots, r$ and for any $j = 1, \dots, r!$, $L(a_{i,j}, b_{1, i-1+(r-1)(j-1)}) = c_{1,j}^*$ and $L(b_{i,j}, a_{1, i-1+(r-1)(j-1)}) = c_{2,j}^*$.
- For any $j = 1, \dots, r!$, $L(b_{1, j+(r-1) \times r!}, a_{1, (r+1)!-j+1}) = c_{1,j}^*$ and $L(a_{1, j+(r-1) \times r!}, b_{1, (r+1)!-j+1}) = c_{2,j}^*$.
- We associate a new color to each missing edge.

I is clearly an instance of *LABELED Min PM_r*. The set of colors $\mathcal{C}' = \{c_{i, \lceil \frac{j}{r} \rceil} : i = 1, \dots, r, j = 1, \dots, n_i\}$ is the approximate solution outputted by *Greedy* and it uses $apx(I) = (H_r + r) \times r!$ colors, whereas $\mathcal{C}^* = \{c_{i,j}^* : i = 1, 2, j = 1, \dots, r!\}$ is the set of colors that are used by an optimal solution; this latter verifies $opt(I) = 2 \times r!$. The Figure 3 describes the instance I for $r = 2$.

We conjecture that *LABELED Min PM* is not $O(n^\varepsilon)$ -approximable in complete bipartite graphs. Thus, a challenge will be to give better approximate algorithms or to improve the lower bound.

References

- [1] P. Alimonti and V. Kann. Some APX-completeness results for cubic graphs. *Theoretical Computer Science* 237:123-134, 2000.
- [2] T. Asano and D. P. Williamson Improved Approximation Algorithms for MAX SAT. *Journal of Algorithms* 42(1):173-202, 2002.
- [3] P. Berman, M. Karpinski, and A. D. Scott. Approximation Hardness of Short Symmetric Instances of MAX-3SAT. *ECCC TR-03-049*, 2003.
- [4] P. Berman and M. Karpinski. On Some Tighter Inapproximability Results. *ECCC TR-05-029*, 1998.
- [5] H. Broersma and X. Li. Spanning trees with many or few colors in edgecolored graphs. *Discussiones Mathematicae Graph Theory*, 17(2):259-269, 1997.
- [6] H. Broersma X. Li, G. J. Woeginger and S. Zhang. Paths and cycles in colored graphs. *Australasian J. Combin.*, 31:, 2005.
- [7] T. Brüggemann, J. Monnot, and G. J. Woeginger. Local search for the minimum label spanning tree problem with bounded color classes. *Operations Research Letters* 31(3):195-201, 2003.
- [8] K. Cameron. Coloured matchings in bipartite graphs. *Discrete Mathematics*, 169:205-209, 1997.
- [9] R-S. Chang and S-J. Leu. The minimum labeling spanning trees. *Information Processing Letters*, 63:277-282, 1997.
- [10] U. Feige. A threshold of for approximating set cover. *J. ACM*, 45:634-652, 1998.
- [11] M. R. Garey and D. S. Johnson. Computers and intractability. a guide to the theory of NP-completeness. *CA, Freeman*, 1979.
- [12] A. Itai, M. Rodeh, and S. Tanimoto. Some matching problems in bipartite graphs. *J. ACM*, 25(4):517-525, 1978.
- [13] M. Karpinski. Personnal communication. 2005.
- [14] S. O. Krumke and H-C. Wirth. On the minimum label spanning tree problem. *Information Processing Letters*, 66:81-85, 1998.
- [15] M. V. Marathe, S. S. Ravi. On Approximation Algorithms for the Minimum Satisfiability Problem. *Information Processing Letters*, 58(1):23-29 1996.

- [16] C.H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *J. of Comp. and Sys. Sci.*, 43:425-440, 1991.
- [17] R. Raz and S. Safra. A sub-constant error-probability low-degree test, and sub-constant error-probability PCP characterization of NP. *Proc. 29th Ann. ACM Symp. on Theory of Comp., ACM*, 475-484, 1997.
- [18] Y. Wan, G. Chen, and Y. Xu. A note on the minimum label spanning tree. *Information Processing Letters*, 84:99-101, 2002.
- [19] Y. Xiong, B. Golden and E. Wasil. Worst-case behavior of the MVCA heuristic for the minimum labeling spanning tree problem. *Operations Research Letters* 33(1):77-80, 2005.
- [20] T. Yi, K. G. Murty, C. Spera. Matchings in colored bipartite networks. *Discrete Applied Mathematics* 121(1-3):261-277, 2002.