The complexity of the Pk partition problem and related problems in bipartite graphs

Jérôme Monnot, Sophie Toulouse

The complexity of the P_k partition problem and related problems in bipartite graphs

Jérôme Monnot[1], Sophie Toulouse[2]

[1] Université Paris Dauphine, LAMSADE, CNRS UMR 7024, 75016 Paris, France
`monnot@lamsade.dauphine.fr`
[2] Université Paris 13, LIPN, CNRS UMR 7030, 93430 Villetaneuse, France,
`sophie.toulouse@lipn.univ-paris13.fr`

Abstract. In this paper, we continue the investigation made in [11] about the approximability of $\mathbf{P}_k$ partition problems, but focusing here on their complexity. Precisely, we aim at designing the frontier between polynomial and **NP**-complete versions of the $\mathbf{P}_k$ partition problem in bipartite graphs, according to both the constant k and the maximum degree of the input graph. We actually extend the obtained results to more general classes of problems, namely, the minimum k-path partition problem and the maximum $\mathbf{P}_k$ packing problem. Moreover, we propose some simple approximation algorithms for those problems.

Keywords: P_k-partition; maximum (weighted) P_k-packing; minimum k-path partition; bipartite graphs; **NP**-completeness; approximation algorithms.

1 Introduction

The $\mathbf{P}_k$ partitioning problem ($\mathbf{P}_k$PARTITION in short) consists, given a simple graph $G = (V, E)$ on $k \times n$ vertices, in deciding whether there exists or not a partition of V into vertex-disjoint $\mathbf{P}_k$, where a $\mathbf{P}_k$ is a path on k vertices (or, equivalently, a path of length $k-1$). This problem has been proven to be **NP**-complete [7] for any $k \geq 3$, polynomial otherwise. $\mathbf{P}_k$PARTITION has been widely studied in the literature, mainly because its **NP**-completeness also implies the **NP**-completeness of two famous problems, namely: the minimum k-path partition problem (MINk-PATHPARTITION) and the maximum $\mathbf{P}_k$ packing problem (MAX$\mathbf{P}_k$PACKING). This former consists in computing the smallest number q of vertex-disjoint paths of length *at most* $k-1$ that form a partition of V; this number is usually denoted by $\rho_{k-1}(G)$ for a given constant k, by $\rho(G)$ when no constraint occurs on the path length. This problem has been extensively studied in the literature, [13, 12, 16], because of its applications, as broadcasting in communication networks: if a communication network verifies $q = \rho_{k-1}(G)$, or, in other words, if it admits a partition of its vertices into q paths of length *at most* $k-1$, then a message may be broadcasted throughout the network within k units of time, using q message originators. The latter problem, MAX$\mathbf{P}_k$PACKING,

consists in finding a maximum number of vertex-disjoint $\mathbf{P}_k$. In its weighted version, MaxW$\mathbf{P}_k$Packing, the input graph $G = (V, E)$ is given together with a weight function $w : E \to \mathbb{N}$ on its edges; the aim is thus to compute a collection $\mathcal{P} = \{P_1, \ldots, P_q\}$ of vertex-disjoint $\mathbf{P}_k$ that is of maximum weight, where the weight $w(\mathcal{P})$ of a solution is given by the sum of the weights of the edges that takes part to the solution. Finally, the special case where the graph is complete on $k \times n$ vertices is denoted by MaxW$\mathbf{P}_k$Partition (note that the minimization version may also be considered); some approximation results concerning this latter may be found in [8, 9, 11]. Note that both problems MaxW$\mathbf{P}_k$Packing and MinW$\mathbf{P}_k$Partition have some relationship with the vehicle routing problem, [16, 3].

Here, we study the complexity of $\mathbf{P}_k$Partition in the case of bipartite graphs; mainly, we aim at deciding whether the restriction of $\mathbf{P}_k$Partition to bipartite graphs belongs to **P** or is **NP**-complete, according to the maximum degree of the graph. We first show that $\mathbf{P}_k$Partition is **NP**-complete for any $k \geq 3$ in graphs with maximum degree 3, and this even if the graph is planar when $k = 3$. We immediately derive the result to Min$k$-PathPartition and Max$\mathbf{P}_k$Packing. On the opposite, both $\mathbf{P}_k$Partition, Min$k$-PathPartition and Max$\mathbf{P}_k$Packing trivially become polynomial-time computable in graphs with maximum degree 2, and we prove that this fact also holds for MaxW$\mathbf{P}_k$Packing. Finally, we propose some approximation results for these problems; although these latter may seem quite simple and thus, are likely to be improved soon, they are, to our knowledge, the first published ones.

This paper is organized as follows: in the next section, we will briefly present previous related work about the hardness of solving bounded-size-path packing problems; then, the third part is dedicated to complexity results concerning $\mathbf{P}_k$Partition when restricting to bipartite graphs; finally, some approximation results concerning MaxW$\mathbf{P}_3$Packing and Min3-PathPartition are proposed in a fourth section.

The notations that will be used by the following are the usual ones according to graph theory. When dealing with approximation matters, $opt(I)$ (*resp.,* $apx(I)$) will represent the value of an optimal (*resp.,* an approximate) solution. An algorithm $\mathcal{A}$ will thus be said to be ε-approximate with $\varepsilon \geq 1$ for a minimization problem (resp., with $0 < \varepsilon \leq 1$ for a maximization problem) if $apx(I) \leq \varepsilon \times opt(I)$ (resp., $apx(I) \geq \varepsilon \times opt(I)$) for any instance I (for more details, see for instance [2]).

2   Previous related work

From a complexity point of view and because of its natural applications, the minimum k-path partition problem may be the most intensively studied path packing problem. It is obviously **NP**-hard in general graphs, and remains intractable on some special graph classes: recently, it has been proven that its restriction to comparability graphs is also **NP**-hard, [13], as well as its restriction to cographs, [12] and to bipartite chordal graphs, [13], when k is part of the input. Note that

most of the proofs of **NP**-hardness actually establish the **NP**-completeness of $\mathbf{P}_k$PARTITION. Nevertheless, the problem turns to be polynomial-time solvable for some particular graph topologies, such as trees, [16], cographs when k is fixed, [12] or bipartite permutation graphs, [13]. Note that one can also find in the literature several results about partitioning the graph into disjoints paths of length at least 2, [15, 10].

Concerning the approximability of related problems, Hassin and Rubinstein, [8] proposed a generic algorithm to approximate MAX$\mathbf{P}_4$PARTITION that guarantees an approximation ratio of $3/4$ for general distance function. More recently in [11], it has been proven that this algorithm also is a $9/10$-approximation for the $1, 2$-instances and, if we consider the minimization case, that it provides respectively a $3/2$- and a $7/6$-approximation for the metric and the $1, 2$-instances. In [9], the authors established, by the means of a randomized algorithm, a $35/67 - \varepsilon$ approximation for MAX$\mathbf{P}_3$PARTITION. Observe that these results do not hold if we consider MAXW$\mathbf{P}_k$PACKING, since the graph is no longer necessarily complete. To our knowledge, there is no specific approximation results for MAXW$\mathbf{P}_k$PACKING. However, using approximation results for the maximum weighted k-packing problem (mainly based on some local search techniques, [1]), we can obtain a $(\frac{1}{k-1} - \varepsilon)$-approximation for MAXW$\mathbf{P}_k$PACKING. Finally, there is, to our knowledge, no approximation result for MINk-PATHPARTITION. Nevertheless, if we consider as objective function the total number of edges used by the paths and if the goal is to maximize, then we can find some approximation results, in [14] for the general case, in [4] for dense graphs.

3 Complexity results

Theorem 1. $\mathbf{P}_k$PARTITION *is **NP**-complete in bipartite graphs with maximum degree 3, for any $k \geq 3$.*

Proof. The proof is based on a reduction from the k-dimensional matching problem, denoted by kDM, which is known to be **NP**-complete, [7]. An instance of kDM consists of a subset $\mathcal{C} = \{c_1, \ldots, c_m\} \subseteq X_1 \times \ldots \times X_k$ of k-tuples, where $X_1, \ldots, X_k$ are k pairwise disjoint sets of size n. A matching is then a subset $M \subseteq \mathcal{C}$ such that no elements in M agree in any coordinate, and the purpose of kDM is to answer the question: does there exist a perfect matching M on $\mathcal{C}$, that is, a matching of size n?

Given an instance $I = (\mathcal{C}, X_1 \times \ldots \times X_k)$ of kDM, we build an instance $G = (V, E)$ of $\mathbf{P}_k$PARTITION, where G is a bipartite graph of maximum degree 3, as follows:

case 1: k is odd.

• To each k-tuple $c_i \in \mathcal{C}$, we associate a gadget $H(c_i)$ that consists of a collection $\{P^{i,1}, \ldots, P^{i,k}\}$ of k vertex-disjoint $\mathbf{P}_k$ with $P^{i,q} = \{a_1^{i,q}, \ldots, a_k^{i,q}\}$ for $q = 1, \ldots, k$. We also put into $H(c_i)$ the edges $[a_1^{i,q}, a_1^{i,q+1}]$ for $q = 1$ to $k - 1$, in order to form a $(k + 1)$-th $\mathbf{P}_k$ $\{a_1^{i,1}, \ldots, a_1^{i,k}\}$ (see Figure 1 for an illustration when $k = 3$).
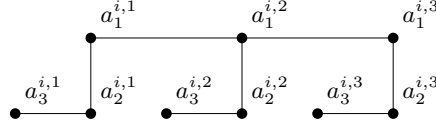
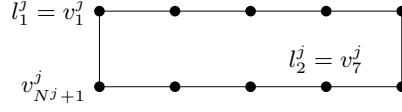**Fig. 1.** The gadget $H(c_i)$ when c_i is a 3-uplet.

Fig. 2. The gadget $H(e_j)$ for $k = 3$ and $d^j = 2$.

• For each element $e_j \in X_1 \cup \ldots \cup X_k$, let d^j denotes the number of k-tuples $c_i \in \mathcal{C}$ that contain e_j; the gadget $H(e_j)$ is then defined as a cycle $\{v_1^j, \ldots, v_{N^j+1}^j, v_1^j\}$ on $N^j + 1$ vertices, where $N^j = k(2d^j - 1)$. Furthermore, we denote by l_p^j for $p = 1$ to d^j the vertex of index $2k(p-1) + 1$ (see Figure 2 for an illustration of $H(e_j)$ when $k = 3$ and $d^j = 2$).

• Finally, for any couple (e_j, c_i) such that e_j is the value of c_i on q-th coordinate, the two gadgets $H(c_i)$ and $H(e_j)$ are connected by the means of an edge $[a_2^{i,q}, l_{p_i}^j]$. The indexes p_i of the vertices $l_{p_i}^j$ that will be linked to a given gadget $H(c_i)$ must be chosen in such a way that each vertex l_p^j from any $H(e_j)$ gadget will be connected to exactly one gadget $H(c_i)$ (what is possible since each $H(e_j)$ contains exactly d^j vertices l_p^j).

This construction obviously leads to a graph G of maximum degree 3, on $3k^2m + (1-k)kn$ vertices: consider, on the one hand, that each gadget $H(c_i)$ is a graph on k^2 vertices and, on the other hand, that $\sum_{j=1}^{kn} d^j = km$ (we may assume, wlog., that each element e_j appears at least once in $\mathcal{C}$). Thus, it requires a polynomial amount of time in the input size. Finally, one can weasily see that G is bipartite.

We claim that there exists a perfect matching $M \subseteq \mathcal{C}$ *iff* there exists a partition $\mathcal{P}^*$ of $V(G)$ into $\mathcal{P}_k$. The main argument lies in the following two properties:

Property 1.
(i) In any $\mathcal{P}_k$-partition $\mathcal{P}$ of $V(G)$, and for any $i = 1, \ldots, m$, $\mathcal{P}$ contains either $\mathcal{P}^i$ or $\mathcal{Q}^i$, where $\mathcal{P}^i$ and $\mathcal{Q}^i$ are defined as (see Figure 3 for an illustration from 3DM):

$\forall i = 1, \ldots, m, \forall q = 1, \ldots, k, P^{i,q} = \{a_k^{i,q}, \ldots, a_2^{i,q}, l^{i,q}\}, \; Q^{i,q} = \{a_k^{i,q}, \ldots, a_2^{i,q}, a_1^{i,q}\}$
$\forall i = 1, \ldots, m, \; \mathcal{P}^i = \cup_{q=1}^k P^{i,q} \cup \{a_1^{i,1}, a_1^{i,2}, \ldots, a_1^{i,k}\}, \; \mathcal{Q}^i = \cup_{q=1}^k Q^{i,q}$
where $l^{i,q}$ denotes the vertex from some $H(e_j)$ to which $a_2^{i,q}$ is connected.

(ii) In any $\mathbf{P}_k$-partition $\mathcal{P}$ of $V(G)$, and for any $j = 1, \ldots, kn$, $\mathcal{P}$ contains one of the collections $\{\mathcal{P}_p^j\}_{p=1,\ldots,d^j}$, where $\mathcal{P}_p^j$ is defined as (see Figure 4 for an
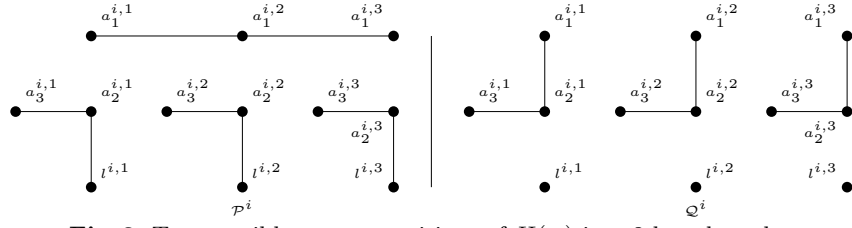
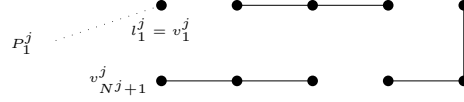**Fig. 3.** Two possible vertex partitions of $H(c_i)$ into 2-length paths.

Fig. 4. One of the d^j possible vertex partitions $\mathcal{P}_p^j$ of $H(e_j)$ into 2-length paths.

illustration):

$$\forall j = 1, \ldots, kn, \forall p = 1, \ldots, d^j, \ \mathcal{P}_p^j = P_p^j \cup \mathcal{Q}_p^j$$

where P_p^j denotes the path $P^{i,q}$ from some $H(c_i)$ to which l_p^j belongs and $\mathcal{Q}_p^j$ is the unique possible $\mathcal{P}_k$-partition of $V(H(e_j)) \setminus \{l_p^j\}$.

For (i): mainly, given a vertex $a_k^{i,q}$ from some $H(c_i)$ gadget, consider that there exist only two $\mathcal{P}_k$, namely $P^{i,q}$ and $Q^{i,q}$, passing by $a_k^{i,q}$. For (ii): since $H(e_i)$ contains a number $N^j = k(2d^j - 1) + 1$ of vertices that is not a multiple of k, at least one vertex l_j has to be covered by the means of some path that involves external vertices, that is, vertices from some $H(c_i)$ gadget; hence, from item (i), this path must be some $P^{i,q}$ path. Now, on the one hand, if exactly one l_p vertex from a gadget $H(e_j)$ is covered by the means of some $P^{i,q}$ path, then the $N^j = k(2d^j - 1)$ vertices that remain uncovered may easily be covered using a sequence of $(2d^j - 1)$ $\mathcal{P}_k$; on the other hand, if two consecutive vertices l_p and $l_{p'}$, $p' > p$ are both covered using some $P^{i,q}$ path, then the subchain of $H(e_j)$ between l_p and $l_{p'}$ contains $2k(p' - p) - 1$ vertices, which is not a multiple of k.

Let M be a perfect matching on $\mathcal{C}$; we build a packing $\mathcal{P}$ applying the following rule: if a given element c_i belongs to M, then use $\mathcal{P}^i$ to cover $H(c_i)$, use $\mathcal{Q}^i$ otherwise. At this stage, because M is a perfect matching, exactly one vertex l_p per gadget $H(e_j)$ is already covered by the means of some $P^{i,q}$ path and thus, one may use the corresponding partition $\mathcal{P}_p^j$ in order to partition $V(H(e_j))$.

Conversely, let $\mathcal{P}^* = \{P_1, \ldots, P_r\}$ be a partition of $V(G)$ into $\mathbf{P}_k$. From the first item of Property 1, we know that every $H(c_i)$ gadget is covered either by collection $\mathcal{P}^i$, or by collection $\mathcal{Q}^i$. Furthermore, item (ii) implies that, in order to cover the vertices of a given $H(e_j)$, $\mathcal{P}^*$ uses a single $P^{i,q}$ path and we deduce that the $H(c_i)$ gadget that is concerned with this $P^{i,q}$ path is covered by $\mathcal{P}^i$. Hence, by setting $M = \{c_i \mid \mathcal{P}^i \subseteq \mathcal{P}^*\}$, we define a perfect matching, and the proof is complete.
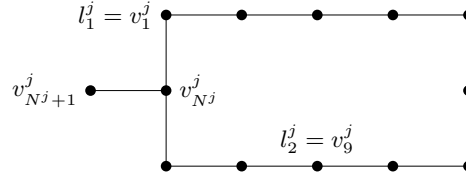
case 2: k is even.

Fig. 5. The gadget $H(e_j)$ for $k = 4$ and $d^j = 2$.

The proof is quite identical, except the construction of the $H(e_j)$ gadgets: $H(e_j)$ is no longer a cycle on $N^j + 1$ vertices, but a cycle $\{v_1^j, \ldots, v_{N^j}^j, v_1^j\}$ on N^j vertices, plus an additional edge $[v_{N^j}^j, v_{N^j+1}^j]$ (see Figure 5 for an illustration when k and d^j are respectively worth 4 and 2). The special vertices l_p^j are defined as well as the odd case as $l_p^j = v_{2k(p-1)+1}^j$ for $p = 1$ to d^j (note that $l_{d^j}^j$ never matches $v_{N^j}^j$). We can easily see that $H(e_j)$ still is bipartite (consider that now k is even, so is N^j). Furthermore, as the same as the odd case, the only valid $\mathbf{P}_k$ packings of $V(H(e_j))$ are the collections $\mathcal{P}_p'^j$ where $\mathcal{P}_p'^j$ contains the $P^{i,q}$ path P_p^j, plus the unique possible $\mathbf{P}_k$ partition $\mathcal{Q}_p'^j$ of the two chains $\{v_1^j, \ldots, v_{2k(p-1)}^j\}$ and $\{v_{2k(p-1)+2}^j, \ldots, v_{N^j}^j, v_{N^j+1}^j\}$.

Corollary 1. MAX$\mathbf{P}_k$PACKING *and* MINk-PATHPARTITION *both are* **NP**-*complete in bipartite graphs with maximum degree 3, for any* $k \geq 3$.

If we decrease the maximum degree of the graph down to 2, we can easily prove that MAX$\mathbf{P}_k$PARTITION, MAX$\mathbf{P}_k$PACKING and MINk-PATHPARTITION are polynomial-time computable. The same fact holds for MAX$W\mathbf{P}_k$PACKING, although its establishment is a little bit complicated.

Proposition 1. MAXW$\mathbf{P}_k$PACKING *is polynomial in graphs with maximum degree 2, for any* $k \geq 3$.

Proof. We reduce the problem of computing an optimum solution of MAXW$\mathbf{P}_k$PACKING in graphs with maximum degree 2 to the problem of computing a maximum weight independent set (MAXWIS) in a chordal graph, which is known to be polynomial [6]. Let $I = (G, w)$ be such an instance of MAXW$\mathbf{P}_k$PACKING, we may suppose wlog., that G is a path. Actually, a graph of maximum degree 2 is a collection of disjoint paths, cycles and isolated edges; thus, on the one hand, each connected component may be solved separately and, on the other hand, a cycle $C = \{e_1, \ldots, e_\ell\}$ may be optimally solved by picking the best solution among k optimum packing computed on the paths $P_j = C \setminus \{e_j\}$, $j = 1, \ldots, k$ (consider that an optimum solution may not incorporate the whole k-length path $\{e_1, \ldots, e_k\}$!). Thus, let $G = \{v_1, \ldots, v_n\}$ be such a path; we build the instance (H, w) of MAXWIS where the vertex set of H represents the $\mathbf{P}_k$ of G, precisely: a vertex v_i' for $i = 1, \ldots, n - (k-1)$ corresponds to the path $P_i = \{v_i, \ldots, v_{i+k}\}$, its weight it settled to $w(v_i') = w(P_i)$, and two vertices $v_i' \neq v_j'$ are linked in

H iff the corresponding paths P_i and P_j share at least one common vertex in
the initial graph. We deduce that the set of independent sets in H corresponds
to the set of $\mathbf{P}_k$-packings in G. Let us now prove that H is chordal, or, equiv-
alently, that H is recursively simplicial. Starting from v_1', we observe that it is
simplicial in H (v_1' is connected to the set $\{v_2', \ldots, v_k'\}$ that forms a clique), we
then observe that v_2 is simplicial in the subgraph induced by $\{v_2, \ldots, v_{n-(k-1)}\}$
and so on, which concludes the proof.

On the other hand, even when k is worth 3, the previous proof of **NP**-
completeness extends to some restrictions of the problem. Precisely, thanks to
the topology of the graph that enabled to establish Theorem 1, we may deduce
from that latter the following stronger result when $k = 3$.

Theorem 2. $\mathbf{P}_3$PARTITION *is **NP**-complete in planar bipartite graphs with max-
imum degree 3.*

Proof. We apply the previous proof, except that we start from a restriction of the
3-dimensional matching problem, which is denoted by PLANAR 3DM-3. With
respect to this restriction, on the one hand, each element $e_j \in X_1 \cup X_2 \cup X_3$
appears in at most three distinct 3-tuples $c_i \in \mathcal{C}$ and, on the other hand, the
characterization bipartite graph $G(\mathcal{C})$ of the instance is planar. The left-hand-
side and the right-hand-side vertex sets of $G(\mathcal{C})$ respectively represent the 3-
tuples from $\mathcal{C}$ and the elements from $X_1 \cup X_2 \cup X_3$; thus, a left vertex l_i will be
linked to a right one r_j *iff* the corresponding 3-tuple c_i contains the corresponding
element e_j. It is well known that this restriction of 3DM is still **NP**-complete, [5].
In order to apply the previous construction properly, we have to link the $H(c_i)$
gadgets to the $H(e_j)$ gadgets in such a way that the final graph G is planar,
namely: for any couple $(H(c_i), H(e_j))$ such that $e_j \in c_i$, the choice of the vertex
l_p^j from $H(e_j)$ that will be linked to $H(c_i)$ is no longer free, but depends on the
characteristic graph $G(\mathcal{C})$ of the input instance.

Corollary 2. MAX$\mathbf{P}_3$PACKING *and* MIN3-PATHPARTITION *are **NP**-complete
in planar bipartite graphs with maximum degree 3.*

An interesting question concerns the status of $\mathbf{P}_k$PARTITION for $k \geq 4$ in planar
bipartite graphs with maximum degree 3.

4   Approximation results

We present some approximation results for MAXW$\mathbf{P}_3$PACKING and MIN3-PATHPARTITION,
which are mainly based on matching heuristics.

4.1   MaxWP₃Packing in bipartite graphs

For this problem, the best approximate algorithm known so far provides a ratio
of $(\frac{1}{2} - \varepsilon)$; this algorithm is deduced from the one proposed in [1] to approximate

the weighted packing problem with sets of size 3. We slightly improve this ratio up to $\frac{1}{2}$ for bipartite graphs with maximum degree 3; we then show that, in the unweighted case, this result holds without any constraint on the graph maximum degree.

From a bipartite graph $G = (L \cup R, E)$ of maximum degree 3, we build two weighted graphs (G_L, d_L) and (G_R, d_R) where $G_L = (L, E_L)$ and $G_R = (R, E_R)$. Two vertices $x \neq y$ from L are linked in G_L *iff* there exists in G a 2-length path $P_{x,y}$ from x to y, rigorously: $[x, y] \in E_L$ *iff* $\exists z \in R$ s.t. $[x, z], [z, y] \in E$. The distance $d_L(x, y)$ of the edge $[x, y]$ is settled to the weight of a maximum weight 2-length path from x to y. (G_R, d_R) is defined as the same (just reverse L and R). If G is of maximum degree 3, then the following fact holds:

Lemma 1. *From any matching M on G_L (resp., on G_R), one can deduce a $\mathbf{P}_3$ packing $\mathcal{P}_M$ of weight $w(\mathcal{P}_M) = d_L(M)$ (resp., $w(\mathcal{P}_M) = d_R(M)$), when G is of degree at most 3.*

Proof. Consider two edges $e_1 = [x_1, y_1] \neq e_2 = [x_2, y_2]$ from a given matching M on G_L and let respectively $P_{e_1} = \{x_1, z_1, y_1\}$ and $P_{e_2} = \{x_2, z_2, y_2\}$ be the corresponding $\mathbf{P}_3$ in the initial graph G. If P_{e_1} and P_{e_2} share a common vertex, then this vertex necessarily is the central vertex $z_1 = z_2$ (or M is not a matching!), which would contradict the fact that G is of maximum degree 3.

Weighted $\mathbf{P}_3$-Packing

1 Build the weighted graphs (G_L, d_L) and (G_R, d_R);
2 Compute a maximum weight matching M_L^* (resp., M_R^*) on (G_L, d_L) (resp., on (G_R, d_R));
3 Deduce from M_L^* (resp., M_R^*) a $\mathbf{P}_3$ packing $\mathcal{P}_L$ (resp., $\mathcal{P}_R$) according to Lemma 1;
4 Output the best packing $\mathcal{P}$ among $\mathcal{P}_L$ and $\mathcal{P}_R$.

Theorem 3. *Weighted $\mathbf{P}_3$-Packing provides a $1/2$-approximation for* MᴀxW$\mathbf{P}_3$Pᴀᴄᴋɪɴɢ *in bipartite graphs with maximum degree 3 and this ratio is tight. In the unweighted case, this results holds without any constraint on the degree of the graph.*

Proof. Let $\mathcal{P}^*$ be an optimum P_3-packing on $I = (G, w)$, we partition $\mathcal{P}^*$ into $\mathcal{P}_L^*$ and $\mathcal{P}_R^*$ according to the membership (to L or to R) of the endpoints of the paths. To $\mathcal{P}_L^*$ (*resp.*, to $\mathcal{P}_R^*$), there corresponds in G_L (*resp.*, in G_R) a matching M_L (*resp.*, M_R) of value at least $w(\mathcal{P}_L^*)$ (*resp.*, $w(\mathcal{P}_R^*)$). From M_L^* and M_R^* optimality, and using Lemma 1, we deduce:

$$w(\mathcal{P}_L) \geq w(\mathcal{P}_L^*), \ w(\mathcal{P}_R) \geq w(\mathcal{P}_R^*) \tag{1}$$

Now, the solution outputted by the algorithm verifies $w(\mathcal{P}) \geq 1/2(w(\mathcal{P}_L) + w(\mathcal{P}_R))$, which concludes the proof. The instance $I = (G, w)$ that provides the

Fig. 6. The graph G that reaches the approximation ratio.

tightness is depicted in Figure 6; it consists of a graph on $12n$ vertices on which we can easily see that, on the one hand, $w(\mathcal{P}_L) = w(\mathcal{P}_R) = 2n(n+2)$ and, on the other hand, $w(\mathcal{P}^*) = 2n(2n+2)$.

Concerning the unweighted case, we may obtain the same performance ratio without any restriction on graph maximum degree. The only difference with the previous algorithm lies on the construction of the graphs G_L and G_R. This former is defined as follows: starting from G, we duplicate each vertex $r_i \in R$ by the means of a new vertex r_i'; we then affect to r_i' the same neighborhood as r_i and finally, add the edge $[r_i, r_i']$. If R_L denotes the vertex set $\{r_i, r_i' | r_i \in R\}$, then the following lemma holds:

Lemma 2. *From any matching M on G_L, one can build a matching M' on G_L that saturates R_L, and of which cardinality verifies $|M'| \geq |M|$.*

Let M be a matching on G_L. If none of the two vertices r_i and r_i' for some i are saturated by M, then set $M' = M \cup \{[r_i, r_i']\}$. If exactly one of them is saturated by a given edge e from M, then set $M' = (M \setminus \{e\}) \cup \{[r_i, r_i']\}$. In any case, M' still is a matching of size at least $|M|$. Thus, the expected result is obtained by applying this process to each vertex from R_L.

Thus, consider a matching M on G_L and assume that M saturates R_L. Hence, if d denotes the number of edges $[r_i, r_i']$ within M, we have $|M| = |R_L| - d = 2|R| - d$. We deduce from M a $\mathbf{P_3}$ packing $\mathcal{P}_M$ on G, of which weight verifies: $w(\mathcal{P}_M) = |M| - d = 2(|M| - |R|)$. Conversely, any $\mathbf{P_3}$ packing $\mathcal{P}_L$ of $V(G)$ such that every $\mathbf{P_3}$ it contains has its two endpoints in L may be converted into a matching M on G_L of size $|M| = w(\mathcal{P}_L) + (|R| - w(\mathcal{P}_L)/2) = w(\mathcal{P}_L)/2 + |R|$. Thus, from a maximum matching M^* on G_L that verifies Lemma 2, we build a $\mathbf{P_3}$ packing $\mathcal{P}_{M^*}$ on G that verifies $w(\mathcal{P}_{M^*}) = 2(|M^*| - |R|) \geq \mathcal{P}_L^*$ and the proof is complete.

4.2   Min3-PathPartition in general graphs

To our knowledge, neither the approximation status of MINk-PATHPARTITION, nor the one of MINPATHPARTITION, have been studied so far. Here, we propose a 3/2-approximation for MIN3-PATHPARTITION, by the means of a quite simple algorithm. Note that, concerning MINPATHPARTITION (that is, the approximation of $\rho(G)$), we can trivially see that it is not $(2 - \varepsilon)$-approximable, from the fact that deciding whether $\rho(G) = 1$ or $\rho(G) \geq 2$ is **NP**-complete. Actually,

we can more generally establish that $\rho(G)$ is not in **APX**: otherwise, we could obtain a **PTAS** for the traveling salesman problem with weights 1 and 2 when $opt(I) = n$, which is not possible, unless **P=NP**.

    Minimum 3-PathPartition

---

1 Compute a maximum matching M_1 on G;
2 Build a bipartite graph $G_2 = (L, R; E_2)$ where L is isomorphic to M_1, R is isomorphic to $V \setminus V(M_1)$, and $[l_e, r_v] \in E_2$ *iff* the corresponding isolated vertex $v \notin V(M_1)$ is adjacent in G to the edge $e \in M_1$;
3 Compute a maximum matching M_2 on G_2;
4 Output $\mathcal{P}$ the 3-paths partition deduced from M_1, M_2 and $V \setminus V(M_1 \cup M_2)$.

Theorem 4. *Min3-PathPartition is $3/2$-approximable in general graphs and this ratio is tight.*

Proof. Let $\mathbf{P}^* = (\mathbf{P}_2^*, \mathbf{P}_1^*, \mathbf{P}_0^*)$ and $\mathbf{P} = (\mathbf{P}_2, \mathbf{P}_1, \mathbf{P}_0)$ respectively be an optimal solution and the approximate 3-Paths partition on G, where $\mathbf{P}_i^*$ and $\mathbf{P}_i$ denote for $i = 0, 1, 2$ the subset of i-length paths from the considered partition. Let $V_1 = V \setminus V(M_1)$ and $V_2 = V_1 \setminus \mathbf{P}_0^*$; we consider G_2' the subgraph of G_2 induced by $\mathbf{P}^*$. According to these notations (consider $|\mathbf{P}_2| = |M_2|$ and $|\mathbf{P}_1| = |M_1| - |M_2|$), we have:

$$apx(I) = \sum_{i=0}^{2} |\mathbf{P}_i| = |M_1| + |\mathbf{P}_0| \tag{2}$$

Because M_1 is a maximum matching on G and by construction of V_2, we deduce, on the one hand, that $d_{G_2'}(r_v) \geq 1$ for any $v \in V_2$ and, on the other hand, that $d_{G_2'}(l_e) \leq 2$ for any $e \in M_1$; hence, we get: $|M_1| \geq |V_2|/2$. Furthermore, from the fact that $\mathbf{P}^*$ is a 3-Paths partition, we deduce that G_2' contains a matching of size at least one-half $|V_2|$ and thus, $|M_2| \geq |V_2|/2$. Hence, the approximate and the optimal solution respectively verify:

$$|\mathbf{P}_0| \leq \frac{|V| + |\mathbf{P}_0^*|}{2} - |M_1^*| \tag{3}$$

$$opt(I) \geq \frac{|V| + |\mathbf{P}_0^*|}{3} \tag{4}$$

Using inequalities (2), (3) and (4), we obtain the expected result. The proof of the tightness is omitted.

References

1. E. Arkin, R. Hassin. On local search for weighted packing problems. *Mathematics of Operations Research*, 23: 640-648, 1998.
2. G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela and M. Protasi. Complexity and Approximation (Combinatorial Optimization Problems and Their Approximability Properties). *Springer, Berlin*, 1999.

3. C. Bazgan, R. Hassin, and J. Monnot. Approximation algorithms for some routing problems. *Discrete Applied Mathematics*, 146: 3-26, 2005.
4. B. Csaba, M. Karpinski, P. Krysta. Approximability of dense and sparse instances of minimum 2-connectivity, TSP and path problems. *SODA*, 74-83, 2002.
5. M. Dyer, A. Frieze. Planar 3DM is NP-complete. *J. Algorithms*, 7:174-184, 1986.
6. A. Frank. Some Polynomial Time Algorithms for Certain Graphs and Hypergraphs. *Proceedings of the 5th British Combinatorial Conference, Congressus Numerantium XV, Utilitas Mathematicae, Winnipeg*, 211-226, 1976.
7. M. R. Garey, D. S. Johnson. Computers and intractability. A guide to the theory of NP-completeness. *CA, Freeman*, 1979.
8. R. Hassin, S. Rubinstein. An Approximation Algorithm for Maximum Packing of 3-Edge Paths. *Inf. Process. Lett.*, 63: 63-67, 1997.
9. R. Hassin, S. Rubinstein. An Approximation Algorithm for Maximum Triangle Packing. *ESA* , LNCS 3221: 403-413, 2004.
10. A. Kaneko. A necessary and sufficient condition for the existence of a path factor every component of which is a path of length at least two. *Journal of Combinatorial Theory, Series B*, 88: 195-218, 2003.
11. J. Monnot, S. Toulouse. Approximation results for the weighted P_4 partition problem. *The symposia on Fundamentals of Computation Theory*, F.C.T.'2005, LNCS 3623, 377-385, 2005.
12. G. Steiner. On the k-Path partition problem in cographs. *Cong. Numer.*, 147:89-96, 2000.
13. G. Steiner. k-Path partitions in trees. *Theor. Comput. Sci.*, 290:2147-2155, 2003.
14. S. Vishwanathan. An Approximation Algorithm for the Asymmetric Travelling Salesman Problem with Distances One and Two. *Information Processing Letter*, 44(6): 297-302, 1992.
15. H. Wang. Path factors of bipartite graphs. *Journal of Graph Theory*, 18: 161-167, 1994.
16. J-H Yan, G. J. Chang, S. M. Hedetniemi, S.T. Hedetniemi. On the k-path partition of graphs. *Discrete Applied Mathematics*, 78:227-233, 1997.