Packing paths in digraphs

Richard C. Brewster*† Pavol Hell*‡ Sarah H. Pantel §

Romeo Rizzi¶‖ Anders Yeo **

Abstract

Let $\mathcal{G}$ be a fixed set of digraphs. Given a digraph H, a $\mathcal{G}$-packing in H is a collection $\mathcal{P}$ of vertex disjoint subgraphs of H, each isomorphic to a member of $\mathcal{G}$. A $\mathcal{G}$-packing $\mathcal{P}$ is *maximum* if the number of vertices belonging to members of $\mathcal{P}$ is maximum, over all $\mathcal{G}$-packings. The analogous problem for undirected graphs has been extensively studied in the literature. The purpose of this paper is to initiate the study of digraph packing problems. We focus on the case when $\mathcal{G}$ is a family of directed paths. We show that unless $\mathcal{G}$ is (essentially) either $\{\vec{P}_1\}$, or $\{\vec{P}_1, \vec{P}_2\}$, the $\mathcal{G}$-packing problem is NP-complete.

When $\mathcal{G} = \{\vec{P}_1\}$, the $\mathcal{G}$-packing problem is simply the matching problem. We treat in detail the one remaining case, $\mathcal{G} = \{\vec{P}_1, \vec{P}_2\}$. We give in this case a polynomial algorithm for the packing problem. We also give the following positive results: a Berge type augmenting configuration theorem, a min-max characterization, and a reduction to bipartite matching. These results apply also to packings by the family $\mathcal{G}$ consisting of all directed paths and cycles. We also explore weighted variants of the problem and include a polyhedral analysis.

*The authors gratefully acknowledge the support of the Natural Science and Engineering Research Council of Canada.

†Department of Computer Science, Bishop's University, Lennoxville, Québec, Canada, J1M 1Z7, rbrewste@ubishops.ca

‡School of Computing Science, Simon Fraser University, Burnaby, B.C., Canada, V5A 1S6, pavol@cs.sfu.ca

§Department of Mathematics and Statistics, Simon Fraser University, Burnaby, B.C., Canada, V5A 1S6

¶financial support from project TMR-DONET nr. ERB FMRX-CT98-0202 of the European Community.

‖CWI, Amsterdam, The Netherlands, romeo@science.unitn.it

**Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom, anders@sartre.cs.rhbnc.ac.uk

1 Introduction

A matching can be viewed as a collection of vertex disjoint copies of K_2. A natural generalization of this concept is a collection of vertex disjoint copies of an arbitrary fixed graph G, or of members of an arbitrary fixed family $\mathcal{G}$ of graphs. Specifically, a *$\mathcal{G}$-packing* of a graph H is a collection of vertex disjoint subgraphs of H, each isomorphic to some member of $\mathcal{G}$. Given such a packing $\mathcal{P}$, a vertex of H is called *covered* by $\mathcal{P}$ if it belongs to one of the subgraphs in $\mathcal{P}$ (and *exposed* otherwise).

The $\mathcal{G}$-packing problem has received much attention [5, 8, 9, 14] in the case of undirected graphs. In particular, [5] and [9] identified the role of hypomatchable graphs G in the family $\mathcal{G}$. This work was continued by [14], who additionally identified another family of graphs playing an important role for the problem: A *propeller* is a graph which can be obtained from a hypomatchable graph B by the addition of two new vertices c, r, any nonempty set of new edges connecting c to B, and the edge cr. Note that, in particular, the path of length two is a propeller. In [14] the authors obtained a complete classification of the complexity of $\mathcal{G}$-packing with any family of the form $\mathcal{G} = \{K_2, G\}$: The $\mathcal{G}$-packing problem is polynomial time solvable when G has a perfect matching, or is hypomatchable, or is a propeller, and is NP-complete otherwise.

The purpose of this paper is to initiate the study of digraph packings. Note that if $\mathcal{G}$-packing is NP-complete for a family of undirected graphs $\mathcal{G}$, then $\mathcal{G}'$-packing is NP-complete for any family $\mathcal{G}'$ obtained from $\mathcal{G}$ by orienting the arcs of each graph in $\mathcal{G}$. In other words, NP-complete undirected problems yield NP-complete directed problems.

Our focus is on families $\mathcal{G}$ of directed paths. We show that unless $\mathcal{G}$ can be reduced (as described below) to either $\{\vec{P}_1\}$, or $\{\vec{P}_1, \vec{P}_2\}$, the $\mathcal{G}$-packing problem is NP-complete, while in these two cases it is polynomial time solvable. Other results on digraphs are found in [3], where $\mathcal{G}$-packings with families $\mathcal{G}$ of oriented stars are explored, and in [2], which gives a complete classification of the complexity of $\mathcal{G}$-packing with any family of the form $\mathcal{G} = \{\vec{P}_1, G\}$. In turns out that $\{\vec{P}_1, G\}$-packing is polynomial when the underlying graph of the digraph G has a perfect matching, or is hypomatchable. Of all the cases when the underlying graph is a propeller, we only have a polynomial time algorithm for the $\{\vec{P}_1, \vec{P}_2\}$-packing problem. All other digraphs G (including orientations of propellers other than $\vec{P}_2$) yield NP-complete $\{\vec{P}_1, G\}$-packing problems. This further identifies the $\{\vec{P}_1, \vec{P}_2\}$-packing problem as an interesting case. The third author presents a survey of graph packings in her M.Sc. Thesis [15] including an introduction to the directed case. (See also [10].)

We now describe $\mathcal{G}$-packing as a formal decision problem and introduce the concept of reducing one family to another. Let $\mathcal{G}$ be a fixed family of digraphs. (To avoid trivialities we assume $\mathcal{G}$ is nonempty.)

$\mathcal{G}$-packing.

Instance: A digraph H and an integer k.

Question: Does H admit a $\mathcal{G}$-packing that covers at least k vertices?

Clearly, we can also view this problem as an optimization problem. Given a host graph H, a $\mathcal{G}$-packing is *maximum* if it covers the maximum number of vertices in H taken over all $\mathcal{G}$-packings of H. A packing is *perfect* if it covers all the vertices of H.

We will show that when $\mathcal{G}$ is (essentially) a family of directed paths other than $\{\vec{P}_1\}$ or $\{\vec{P}_1, \vec{P}_2\}$, the $\mathcal{G}$-packing problem is NP-complete. The qualifier *essentially* refers to the concept of *reducibility.* Specifically, suppose $\mathcal{G}$ is a family of digraphs and $G \in \mathcal{G}$. Further suppose that there exists a $\mathcal{G}\backslash\{G\}$-packing of G which covers every vertex of G. Then for any digraph H and any $\mathcal{G}$-packing $\mathcal{P}$ of H, there exists a $\mathcal{G}\backslash\{G\}$-packing of H covering the same vertices as $\mathcal{P}$, since the use of G can be avoided. Such a family $\mathcal{G}$ containing a redundant element G is called *reducible.* A family which is not reducible is called *irreducible.* Each family $\mathcal{G}$ contains a unique irreducible subset called the *kernel* of $\mathcal{G}$. Clearly, all families of directed paths containing both $\vec{P}_1$ and $\vec{P}_2$ have $\{\vec{P}_1, \vec{P}_2\}$ as their kernel. Similarly, $\{\vec{P}_1\}$ is the kernel of any family of directed paths containing $\vec{P}_1$ but no paths of even length.

Our main result follows.

Theorem 1.1 *Let $\mathcal{G} \subseteq \{\vec{P}_1, \vec{P}_2, \vec{P}_3, \ldots\}$ be an irreducible family. If $\mathcal{G}$ is neither $\{\vec{P}_1\}$, nor $\{\vec{P}_1, \vec{P}_2\}$, then $\mathcal{G}$-packing is NP-complete. Both $\{\vec{P}_1\}$-packing and $\{\vec{P}_1, \vec{P}_2\}$-packing are polynomial.*

The proof of the NP-completeness result follows in the next section. We note that when $\mathcal{G} = \{\vec{P}_1\}$, a $\mathcal{G}$-packing of a digraph H is just ordinary matching of the underlying graph of H. Thus, the $\{\vec{P}_1\}$-packing problem is polynomial. In the remainder of the paper, the case when $\mathcal{G} = \{\vec{P}_1, \vec{P}_2\}$ is studied from various perspectives. We present a Berge type theorem which states that a packing is not maximum if and only if there exists an *augmenting configuration.* On the other hand, we also present a Hall type min-max characterization stating a packing is maximum if and only if there exists a subset of the vertices which satisfies a certain neighbourhood condition. Together, these two results give us an algorithm which simply consists of constructing a search tree for augmenting configurations. It is very similar to the *bipartite* matching algorithm. (In particular, the more complicated blossom algorithm [7, 16] for nonbipartite matching is not required.)

Rather than discussing our algorithm in detail, we also present a reduction to the bipartite matching problem. The reduction allows us to exploit the rich theory of matchings including algorithmic, polyhedral, and matroid results. Consequently, we have many complexity and polyhedral results which follow directly from matching theory, or whose proofs are similar to standard proofs from matching theory. We state these results here, and direct the reader to a technical report [4] (avaliable online) for the details. The report also contains a direct $O(mn)$ algorithm which is a labeling counterpart to the above augmenting configurations approach.

We use the following terminology: Given a digraph H and two vertices u, v of H, such that uv or vu is an arc in H, we use the term *edge* $\{u, v\}$ for the appropriate arc (either one of them if both are present). We use the term edge when the orientation is unknown or unimportant. A *trail of length* k in H is a sequence of vertices and edges $u_0, e_1, u_1, e_2, u_2, \ldots, e_k, u_k$ of H, in which u_{i-1}, u_i are the two vertices incident with edge e_i, and $e_i \neq e_j$ for $i \neq j$. We will only list the vertices in a trail when the sequence of vertices suffices to identify the trail. A trail in which additionally all vertices are distinct is called a *path*, denoted P_k. We also use the notation $\vec{P}_k$ for the *directed path of length* k, i.e., the path $u_0, u_1, \ldots, u_k$ in which all arcs are oriented from u_{i-1} to u_i for $i = 1, 2, \ldots k$. Given a vertex v, the *in-neighbourhood* of v, denoted $N^+(v)$, respectively *out-neighbourhood* of v, denoted $N^-(v)$, is the set $\{u|uv$ is an arc of $H\}$, respectively $\{u|vu$ is an arc of $H\}$. Given a set of vertices S, $N^+(S)$ is the union of $N^+(v)$ taken over all vertices v in S. The set $N^-(S)$ is analogously defined using $N^-(v)$.

An arc uv is *in the packing* $\mathcal{P}$, denoted $uv \in \mathcal{P}$, if uv belongs to a subgraph in $\mathcal{P}$.

Given a digraph H and a packing $\mathcal{P}$, we extend our neighbourhood definitions by letting $N^+_{\mathcal{P}}(v) = \{u|uv \in \mathcal{P}\}$, and similarly for $N^-_{\mathcal{P}}(v)$, $N^+_{\mathcal{P}}(S)$, and $N^-_{\mathcal{P}}(S)$.

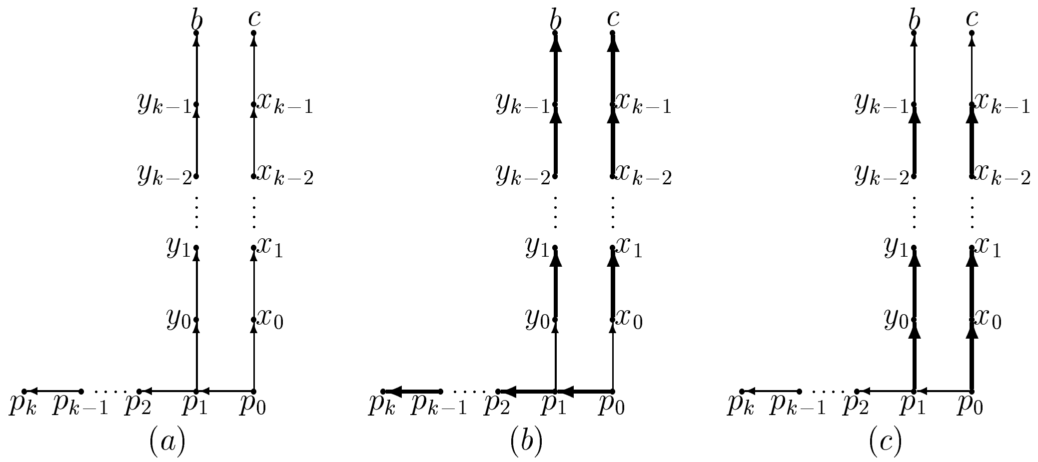

Figure 1: (a): The gadget G_1. (b): A $\mathcal{G}$-packing of G_1 covering the connectors. (c): A $\mathcal{G}$-packing of G_1 leaving the connectors exposed.

2 The NP-completeness proof

In this section, we establish the NP-completeness of the $\mathcal{G}$-packing problem when $\mathcal{G}$ is a family of directed paths whose kernel is neither $\{\vec{P}_1\}$ nor $\{\vec{P}_1, \vec{P}_2\}$.

We reduce 3-dimensional matching to $\mathcal{G}$-packing. Let $\mathcal{T} \subseteq A \times B \times C$ be an instance of 3-dimensional matching. We assume that $|A| = |B| = |C|$ as otherwise $\mathcal{T}$ does not admit a 3-dimensional matching.

Assume that $\mathcal{G}$ is an irreducible family of directed paths other than $\{\vec{P}_1\}$ or $\{\vec{P}_1, \vec{P}_2\}$. Let k be the length of the shortest path in $\mathcal{G}$. We have two cases.

Case 1: $\mathbf{k \geq 2}$. We create an instance of $\mathcal{G}$-packing, say H, as follows. For each point in A, we create a path $Q_i = q_2^i q_3^i \dots q_k^i$ (for $i = 1, 2, \dots, |A|$). For each point in B, we create a vertex b_i (for $i = 1, 2, \dots, |B|$). For each point in C, we create a vertex c_i (for $i = 1, 2, \dots, |C|$). For each triple in (a_i, b_j, c_k) in $\mathcal{T}$, we take a copy of G_1 (see Figure 1 (a)) and identify the path $p_2 p_3 \dots p_k$ with the path $q_2^i q_3^i \dots q_k^i$ in H. Similarly we identify the vertex b in G_1 with b_j in H, and we identify the vertex c in G_1 with c_k in H. We call the path $p_2 p_3 \dots p_k$, and the vertices b, and c the *connectors* of G_1. We show that H admits a perfect $\mathcal{G}$-packing if and only if $\mathcal{T}$ admits a 3-dimensional matching.

Assume $\mathcal{T}$ admits a 3-dimensional matching. Given a triple in the matching, pack the corresponding copy of G_1 as shown in Figure 1 (b). (Thick arcs are in the packing.) For each triple not in the matching, pack the corresponding copy of G_1 as shown in Figure 1 (c). It is easy to see that this yields a perfect $\mathcal{G}$-packing of H.

On the other hand, assume that there exists a perfect $\mathcal{G}$-packing, say $\mathcal{P}$, of H. If the path $p_0p_1\dots p_k$ in a gadget G_1 is in $\mathcal{P}$, then $\mathcal{P}$ restricted to the gadget also covers b and c (see Figure 1 (b)). (The vertex y_0 must be covered by some directed path, and $\vec{P}_k$ is the shortest path in $\mathcal{G}$.) As $|A| = |B| = |C|$ this implies that $\mathcal{P}$ restricted to a G_1 must either cover all of the connectors or none of the connectors. The collection of triples which correspond to gadgets whose connectors are covered forms a 3-dimensional matching. This completes the first case.

Case 2: $\mathbf{k = 1}$. We claim that there must exist an integer $l \geq 2$, such that $\vec{P}_{2l} \in \mathcal{G}$. This is so, since $\mathcal{G} \neq \{\vec{P}_1\}, \mathcal{G} \neq \{\vec{P}_1, \vec{P}_2\}$, and is irreducible. Indeed, an irreducible family containing $\vec{P}_1$ does not contain any other odd length paths. We furthermore assume that l is chosen as small as possible. We now consider the gadget G_2 depicted in Figure 2 (a). We note that if a $\mathcal{G}$-packing of G_2 covers a and p_{2l}, then it also covers b and c (see Figure 2 (b)). Also, there exists a $\mathcal{G}$-packing of G_2 that covers all of G_2, except a, b and c (see Figure 2 (c)). The proof is now analogous to that of case 1, where we use the vertex a instead of the path $p_2p_3\dots p_k$ (and the vertex p_{2l} instead of the vertex y_0). ∎

3 The min-max theorem via augmenting configurations

In the remainder of the paper, we study the $\{\vec{P}_1, \vec{P}_2\}$-packing problem. Thus from now on, the term *packing* without further specification always refers to a $\{\vec{P}_1, \vec{P}_2\}$-packing. Let H be a digraph and let $\mathcal{P}$ be a packing in H. Then $exp(\mathcal{P})$ denotes the number of vertices of H left exposed by $\mathcal{P}$. For any set S of vertices from H, the *deficiency* of S is defined as

$$def(S) := |S| - |N^+(S)| - |N^-(S)|.$$

Clearly, $exp(\mathcal{P}) \geq def(S)$ for every set S and every packing $\mathcal{P}$, since for every vertex of $s \in S$ covered by $\mathcal{P}$ there is an arc sv with $v \in N^-(S)$ or an arc vs

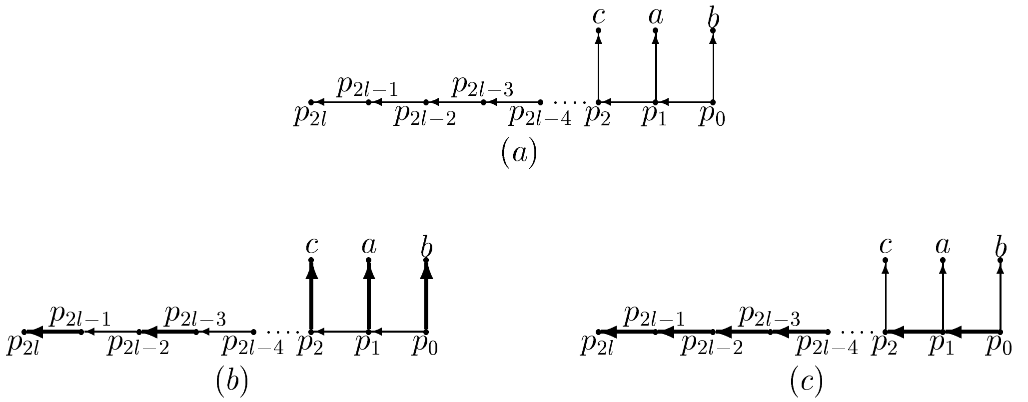

Figure 2: (a): The gadget G_2. (b): A $\mathcal{G}$-packing of G_2 covering the connectors. (c): A $\mathcal{G}$-packing of G_2 leaving the connectors exposed.

with $v \in N^+(S)$ in the packing. The following min-max characterization is a main result of our analysis.

Theorem 3.1 *Let H be a digraph. Then*

$$\min exp(\mathcal{P}) = \max def(S)$$

where the minimum is taken over all packings $\mathcal{P}$ of H and the maximum is taken over all set S of vertices of H.

3.1 Augmenting Configurations

Let H be a digraph and $\mathcal{P}$ a packing in H. An *alternating trail* in H (with respect to $\mathcal{P}$) is an even length trail $u_0, u_1, u_2, \ldots, u_{2k-1}, u_{2k}$ ($k \geq 0$) such that u_0 is exposed, with edges alternately not in $\mathcal{P}$ and in $\mathcal{P}$. For each $i = 0, 1, \ldots, k-1$ we have

- the edges $\{u_{2i}, u_{2i+1}\}$ and $\{u_{2i+1}, u_{2i+2}\}$ are both oriented towards u_{2i+1} or both oriented away from u_{2i+1}; and
- u_{2i+2} is incident with exactly one arc of $\mathcal{P}$ (namely, $\{u_{2i+1}, u_{2i+2}\}$).

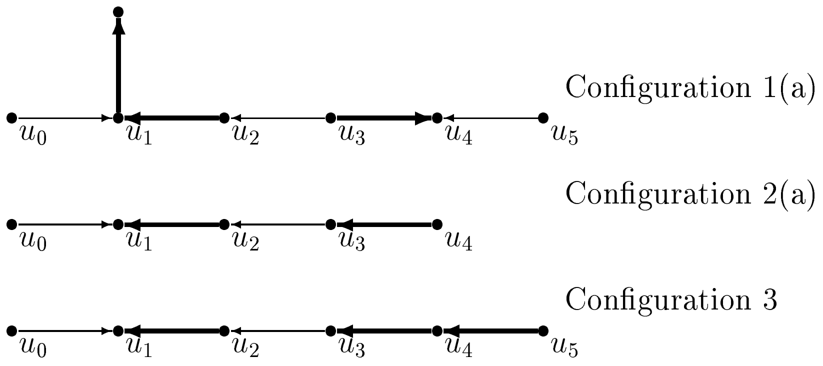

Figure 3: Examples of non-self-intersecting augmenting configurations

We emphasize here that the trails can revisit vertices but not arcs. However, the starting vertex u_0 can only appear once, and each other vertex at most twice (each occurrence requires a different arc from $\mathcal{P}$). In fact, (for any i) a vertex u_{2i} cannot appear twice, and if a vertex u_{2i+1} equals u_{2j+1} (for some $i \neq j$), then both edges incident with u_{2i+1} are oriented towards it and both edges incident with u_{2j+1} away from it, or vice versa.

An *augmenting configuration* is one of the following types of trails:

TYPE 1. (An alternating trail plus a single arc.)

A trail $u_0, u_1, u_2, \ldots, u_{2k}, u_{2k+1}$ where $u_0, u_1, u_2, \ldots, u_{2k}$ is an alternating trail T and

(a) u_{2k+1} does not belong to T and is exposed in $\mathcal{P}$; or

(b) $u_{2k+1} = u_{2i+1}$ for some i with $0 \leq i \leq k-1$, the path u_{2k}, u_{2i+1}, u_{2i} is a directed $\vec{P}_2$, and u_{2i+1}, u_{2i+2} is a $\vec{P}_1$ in $\mathcal{P}$; or

(c) $u_{2k+1} = u_{2i}$ for some i with $0 \leq i \leq k-1$, the path u_{2k}, u_{2i}, u_{2i+1} is a directed $\vec{P}_2$, and u_{2i+1}, u_{2i+2} is a $\vec{P}_1$ in $\mathcal{P}$.

TYPE 2. (An alternating trail plus two arcs.)

A trail $u_0, u_1, u_2, \ldots, u_{2k}, u_{2k+1}, u_{2k+2}$ where $u_0, u_1, u_2, \ldots, u_{2k}$ is an alternating trail T, with the edge $\{u_{2k}, u_{2k+1}\}$ not in $\mathcal{P}$, the edge $\{u_{2k+1}, u_{2k+2}\}$ in $\mathcal{P}$; and

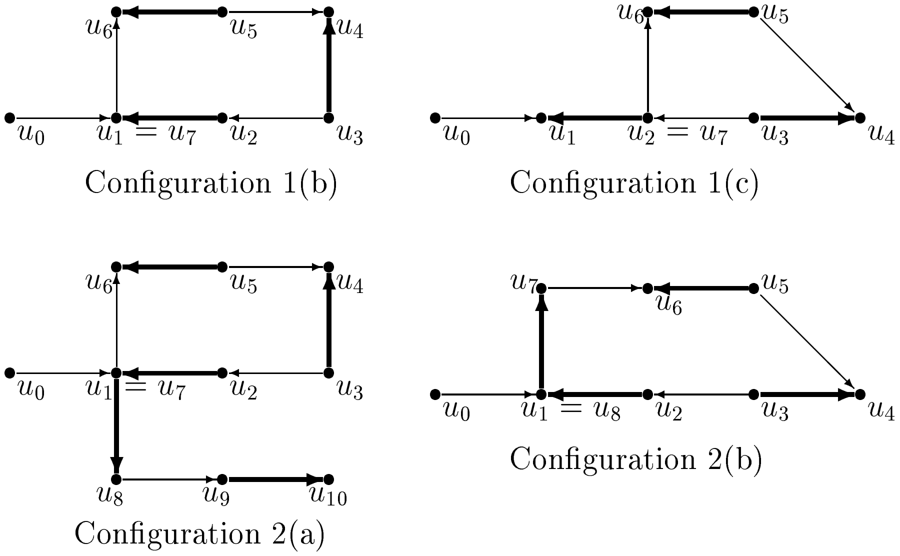

Figure 4: Examples of self-intersecting augmenting configurations

(a) neither u_{2k+1} nor u_{2k+2} belong to T, $u_{2k}, u_{2k+1}, u_{2k+2}$ is a directed $\vec{P}_2$, and u_{2k+1}, u_{2k+2} is a $\vec{P}_1$ in $\mathcal{P}$; or

(b) u_{2k+1} does not belong to T, and $u_{2k+2} = u_{2i+1}$ for some i with $0 \leq i \leq k-1$.

TYPE 3. (An alternating trail plus three arcs.)

A trail $u_0, u_1, u_2, \ldots, u_{2k}, u_{2k+1}, u_{2k+2}, u_{2k+3}$ where $u_0, u_1, u_2, \ldots, u_{2k}$ is an alternating trail T; none of $u_{2k+1}, u_{2k+2}, u_{2k+3}$ belong to T; and the edge $\{u_{2k}, u_{2k+1}\}$ does not belong to $\mathcal{P}$ but both $\{u_{2k+1}, u_{2k+2}\}$ and $\{u_{2k+2}, u_{2k+3}\}$ do.

We now observe that the presence of any of the above augmenting configurations signals that the packing $\mathcal{P}$ can be modified to increase the number of covered vertices. For instance, in a configuration of type 1(a) we may simply take the edges $\{u_{2i+1}, u_{2i+2}\}$ (for all $i = 0, 1, \ldots, k-1$) out of the packing and put the edges $\{u_{2i}, u_{2i+1}\}$ into the packing (for all $i = 0, 1, \ldots, k$). The definition of alternating trail assures that we again obtain a packing. The other configurations are similar – only the types 2(a) and 3 are exceptional in that the last arc is not removed from $\mathcal{P}$.

Figure 3 contains examples of non-self-intersecting augmenting configurations. Figure 4 contains examples of augmenting configurations that do self-intersect. Thick arcs in the drawing belong to $\mathcal{P}$. Note that the role odd cycles play in the examples of type 1(c) and 2(b) suggests why blossoming is not required in our algorithm.

We next state our min-max result which incorporates the concise certificate for maximality of packings and the augmenting configuration theorem for increasing packings.

Theorem 3.2 *Suppose H is a digraph and $\mathcal{P}$ a packing in H. The following statements are equivalent:*

(a) $\mathcal{P}$ is a maximum packing;

(b) $\mathcal{P}$ does not admit an augmenting configuration;

(c) there is a set $S \subseteq V(H)$ with $def(S) = exp(\mathcal{P})$.

Proof: We have already observed that (a) implies (b) - each augmenting configuration allows us modify the packing to increase the number of covered vertices. The inequality $def(S) \leq exp(\mathcal{P})$, derived earlier, implies that for any S and $\mathcal{P}$ with $def(S) = exp(\mathcal{P})$, we have that $def(S)$ is maximum and $exp(\mathcal{P})$ is minimum. Hence (c) implies (a).

It remains to prove that (b) implies (c). Assume that $\mathcal{P}$ does not admit an augmenting configuration, and define the set S as follows:

$$S = \{u : u_0 u_1 \ldots u_{2k-1} u_{2k} \text{ is an alternating trail where } u = u_{2k}\}$$

That is, S is the set of all vertices reachable from an exposed vertex by an even length alternating trail. We now claim:

1. no vertex of S is the center vertex of a $\vec{P}_2$ in $\mathcal{P}$;
2. $N^+(S) = N^+_{\mathcal{P}}(S)$ and $N^-(S) = N^-_{\mathcal{P}}(S)$;
3. $def(S) = exp(\mathcal{P})$.

Since no augmenting configuration exists, no vertex in S is incident with two packing edges; otherwise, we have a configuration of type 3.

We now prove that $N^+(S) = N^+_{\mathcal{P}}(S)$. Let $u \in S$ and $v \in N^+(u)$. If $vu \in \mathcal{P}$, then we are done. On the other hand, if $vu \notin \mathcal{P}$, then by definition of S, there is an alternating trail, $T : u_0, u_1, \ldots u_{2k-1} u_{2k} (= u)$ (where u_0 is by definition exposed). If v is exposed and v does not belong to T, then T together with vu forms an augmenting configuration, contrary to our assumption. (The case that v belongs to T is below.)

Therefore, assume that v is incident with a packing edge say $\{v, w\}$. Begin by assuming that v and w do not belong to T. Since $\mathcal{P}$ does not admit an augmenting configuration, without loss of generality v dominates w, and the trail T, extended by v, w, is an alternating trail. Thus $w \in S$ and $v \in N^+_{\mathcal{P}}(S)$.

Next assume v does not belong to T, but w does belong to T. In this case, $w = u_{2i+1}$ for some $i = 0 \ldots k-1$ and we have an augmenting configuration of type 2(b), a contradiction.

Thus, we assume v belongs to T. If $v = u_{2i+1}$ for some $i = 0, \ldots, k-1$, then either an augmenting configuration of type 1(b) exists, or without loss of generality $w \in S$ and v dominates w. Hence $v \in N^+_{\mathcal{P}}(S)$.

Finally we assume that $v = u_{2i}$ for some $i = 0, \ldots k-1$. (This includes the case above where v is the exposed vertex u_0.) Consider the trail $T' = u_0, \ldots, u_{2i}(= v), u_{2k}(= u), u_{2k-1}$. Since we do not have any augmenting configurations, it must be the case that the arc $u_{2k-1}u_{2k}$ is a $\vec{P}_1$ in $\mathcal{P}$. (Specifically u_{2k-1} dominates u_{2k}). Since T is an alternating trail, it must be that u_{2k-1} also dominates u_{2k-2}. Consider $T'' = u_0, \ldots, u_{2i}, u_{2k}, u_{2k-1}, u_{2k-2}, u_{2k-3}$. Arguing as above, we see that u_{2k-3}, u_{2k-2} is a $\vec{P}_1$ in $\mathcal{P}$ and u_{2k-3} dominates both u_{2k-2} and u_{2k-4}. Continuing in this manner, we can conclude u_{2i+1}, u_{2i+2} is a $\vec{P}_1$ in $\mathcal{P}$ where u_{2i+1} dominates $u_{2i} = v$ and u_{2i+2}. This yields an augmenting configuration of type 1(c), a contradiction.

A similar argument shows $N^-(S) = N^-_{\mathcal{P}}(S)$. Hence, we can count $|N^+(S)|$ and $|N^-(S)|$ using only packing edges. Since each vertex in S is either exposed or incident with exactly one packing edge, we have $|N^+(S)| + |N^-(S)| = |S| - exp(\mathcal{P})$. That is,

$$def(S) = |S| - |N^+(S)| - |N^-(S)| = exp(\mathcal{P}).$$

∎

A different proof of the equivalence of (a) and (b) appears in [15]. This proof is by induction and does not use the min-max condition (c).

We close this section by deducing the following Tutte-type condition for the existence of perfect $\{\vec{P}_1, \vec{P}_2\}$-packings.

Corollary 3.3 *Let H be a digraph with vertex set V. Then the following are equivalent:*

(a) H admits a perfect $\{\vec{P}_1, \vec{P}_2\}$-packing;

(b) $|S| \leq |N^+(S)| + |N^-(S)|$ for all $S \subseteq V$.

$\blacksquare$

Theorem 3.1 applies to two additional families: $\mathcal{DPC}$, the family of all directed paths and cycles; and $\mathcal{DP}$, be the family of all directed paths. This follows from the fact that $\{\vec{P}_1, \vec{P}_2\}$ is the kernel of both $\mathcal{DPC}$ and $\mathcal{DP}$. Specifically, given a digraph H, we have $\min exp(\mathcal{P}) = \max def(S)$, where the minimum is taken over all $\mathcal{DPC}$-packings, respectively $\mathcal{DP}$-packings, and the maximum is taken over all subsets S of the vertex set. The analogue to Corollary 3.3 follows:

Corollary 3.4 *Let H be a digraph with vertex set V. The following are equivalent:*

(a) H admits a perfect $\mathcal{DPC}$-packing;

(b) H admits a perfect $\mathcal{DP}$-packing;

(c) $|S| \leq |N^+(S)| + |N^-(S)|$ for all $S \subseteq V$.

$\blacksquare$

4 Algorithmic considerations

Using Theorem 3.2, it is possible to construct a polynomial time algorithm for the $\{\vec{P}_1, \vec{P}_2\}$-packing problem. We outline this algorithm in subsection 4.1; however, in subsection 4.2, we show how to reduce the packing problem to bipartite matching. This reduction implies that for the packing problem we can achieve the same results as for bipartite matching [11] or max-flow [6] in terms of running time.

4.1 An algorithm based on augmenting configurations

Given some packing $\mathcal{P}$ (in a digraph H), we seek to increase the number of covered vertices using Theorem 3.2, i.e. we seek an augmenting configuration. We begin with a set S consisting of the exposed vertices (with respect to $\mathcal{P}$) and proceed to increase S through the examination of vertices in its neighbourhood. The set S corresponds to the set in Theorem 3.2. As each vertex is examined one of three things happens: First, we can discover an augmenting configuration. In this case we increase the packing and restart the algorithm with this larger packing. Second, we can increase the set S.

In this case we continue with the examination of vertices. Finally, if neither of the first two cases occur, then our set S is a certificate that the packing is maximum, i.e. $def(S) = exp(\mathcal{P})$, and the algorithm terminates. The details of this are fairly straightforward, and the resulting algorithm resembles the bipartite matching algorithm (in particular, there is no blossoming).

4.2 A reduction to bipartite matching

Instead of giving the details of the above algorithm we show how to reduce the problem to the bipartite matching problem. Let H be a digraph with vertex set V and arc set A. We define the *bipartite graph G associated with H* (denoted by $G = G(H)$), as follows: Let V^+, V^* and V^- be three distinct copies of V, with $u^+ \in V^+, u^* \in V^*$, and $u^- \in V^-$ denoting the vertices corresponding to $u \in V$ respectively. Let $E^+ = \{u^+v^* \ : uv \in A\}$ and $E^- = \{\, u^*v^- \ : uv \in A\}$. The graph G has the vertex set $W = V^+ \cup V^* \cup V^-$ and the edge set $E = E^+ \cup E^-$. Note that G is indeed a bipartite graph with bipartition $V^*, V^+ \cup V^-$. We shall describe a correspondence between sets of vertices in H that can be covered by packings of H, and sets of vertices in V^* that can be covered by matchings of G.

Lemma 4.1 *Let G be the bipartite graph associated with a digraph H. For every packing $\mathcal{P}$ of H, there exists a matching M of G such that $u \in V$ is covered by $\mathcal{P}$ if and only if u^* is covered by M.*

Proof: For each arc uv which forms a $\vec{P}_1$ in $\mathcal{P}$, we put in M the edges u^+v^* and u^*v^-. For each pair of arcs uv, vw which form a $\vec{P}_2$ in $\mathcal{P}$, we put in M the edges u^+v^*, u^*v^-, and v^+w^*. ∎

Lemma 4.2 *Let G be the bipartite graph associated with a digraph H. For every matching M of G, there exists a packing $\mathcal{P}$ of H such that $u \in V$ is covered by $\mathcal{P}$ whenever u^* is covered by M.*

Proof: Every nontrivial directed path or cycle admits a perfect $\{\vec{P}_1, \vec{P}_2\}$-packing; thus, it suffices to find a packing of H by directed paths and cycles covering the appropriate vertices. However, the subgraph of H which naturally corresponds to M may contain vertices incident with three arcs (one arc for each of u^-, u^*, u^+). Thus, we shall process M to form a new matching M' covering the same vertices of V^* as M, but which corresponds to a collection of directed paths and cycles in H.

Consider the bipartite graph F on the vertex set $V^+ \cup V^-$, with the edges u^+v^- for all arcs uv of H. Consider the matching $M_1 := M \cap E^+$: by replacing each asterisk with a plus (in the superscripts), we can view M_1 to be a matching of F. Similarly, we can view $M_2 := M \cap E^-$ as a matching of F by replacing all asterisks with a minus sign. Note that v^* is covered by M if and only if v^- is covered by M_1 or v^+ is covered by M_2. By the Dulmage-Mendelsohn theorem [13], we can find (in linear time) a matching M' in F which covers all vertices of V^+ covered by M_1 and also all vertices of V^- covered by M_2. To this matching M' in F there corresponds in H a set of disjoint directed paths and cycles, covering a vertex v whenever v^* is covered by M. ∎

Recall that the size of a packing of H is the number of covered vertices. While the size of a matching of G is formally the number of *edges* in the matching, we note that this equals the number of vertices in V^* covered by the matching.

Theorem 4.3 *Let G be the bipartite graph associated with a digraph H. Then the size of a maximum packing of H equals the size of a maximum matching of G.* ∎

These results (together with the bipartite matching algorithm and a linear-time algorithm inherent in the Dulmage-Mendelsohn theorem) yield a polynomial time algorithm to find a maximum packing.

We remark the above theorem also gives an alternative derivation of our min-max formula. Trivially, for $S \subseteq V^*$, M will leave at least $|S| - |N^+(S)| - |N^-(S)|$ vertices exposed in V^*. Note that this fully corresponds to deficiency in the digraph H. Moreover, by Hall's theorem for bipartite graphs, we know that when M is a maximum matching, there exists a set of vertices S with exactly $def(S)$ exposed vertices.

4.3 Polyhedral considerations

If we put weights on the vertices, then we can find the maximum possible total weight of a set coverable by a $\{\vec{P}_1, \vec{P}_2\}$-packing, by reducing the problem to weighted bipartite matching [4].

However, if we put weights on the arcs, the situation is quite different. In [4] we show that even the "cardinality case" (all arc weights are 1) is already difficult:

Problem 4.4 *Given a digraph H and an integer k, does there exist a maximum $\{\vec{P}_1, \vec{P}_2\}$-packing $\mathcal{P}$ with at least k arcs.*

(Note that maximizing the number of arcs in a packing is equivalent to maximizing the number of $\vec{P}_2$'s.) This makes it unlikely that we can find a polyhedral description of maximum packings. This also explains why post-processing (Dulmage-Mendelsohn) was needed in our reduction. In [4] we also show that for planar digraphs (and even for planar digraphs with in- and out-degrees bounded by 3) the arc weighted version remains NP-complete even if the weights are all 0 or 1.

Now we briefly mention some facts that can be established in standard ways (often by following existing proofs of results on bipartite matchings). Full proofs are available in [4], which is accessible electronically.

Consider the system below

$$\begin{aligned} &\max cx \\ &\begin{cases} x(\delta^+(v)) \leq & 1 \quad \forall v \in V \\ x(\delta^-(v)) \leq & 1 \quad \forall v \in V \\ \qquad\quad x \geq & 0 \end{cases} \end{aligned} \tag{1}$$

The system has only integral vertices and describes the convex hull of characteristic vectors of packings of directed paths and cycles. Hence we can find packings of directed paths and cycles of maximum weight.

Also the vertices of the dual, below, are all integral.

$$\begin{aligned} &\min 1y^+ + 1y^- \\ &\begin{cases} y^+(u) + y^-(v) \geq & c_{(u,v)} \quad \forall (u,v) \in A \\ \qquad\qquad\quad y \geq & 0 \end{cases} \end{aligned} \tag{2}$$

References

[1] J. Akiyama and M. Kano, Path factors of a graph, in: *Graphs and Applications,* (eds. F. Harary and J. S. Maybee), John Wiley & Sons, New York (1985), 1-21.

[2] R. C. Brewster, R. Rizzi, Digraph packings with hypomatchable graphs and propellers, *Information Processing Letters*, to appear.

[3] R. C. Brewster, P. Hell and R. Rizzi, Digraph packings with stars, *in preparation.*

[4] R. C. Brewster, P. Hell and R. Rizzi, Sidepath results on packing P_1's and P_2's. Technical report, *in preparation.*

[5] G. Cornuéjols, D. Hartvigsen, and W. Pulleyblank, Packing subgraphs in a graph, *Operations Research Letters,* **1**(1982), 139-143.

[6] A. Goldberg, and S. Rao, Flows in undirected unit capacity networks, *SIAM J. Discrete Math.,* **12** no. 1 (1999), 1-5.

[7] J. Edmonds. Paths, trees, and flowers. *Canadian Journal of Mathematics,* **17**(1965), 449-467.

[8] P. Hell and D. G. Kirkpatrick. On generalized matching problems. *Information Processing Letters,* **12**(1981), 33-35.

[9] P. Hell and D. G. Kirkpatrick. Packings by cliques and by finite families of graphs. *Discrete Mathematics,* **49**(1984), 45-59.

[10] P. Hell. Graph Packings. *Proceedings of the 6$^{\text{ème}}$ Colloque International de Théorie des Graphes, Electornic Notes in Discrete Mathematics,* **5**.

[11] J. E. Hopcroft and R. M. Karp. An $n^{5/2}$ algorithm for maximum matching in bipartite graphs. *SIAM Journal on Computing,* **2**(1973), 255-231.

[12] D. Lichtenstein, Planar formulae and their uses. *SIAM Journal on Computing,* **11**(1982), no. 2, 329–343.

[13] L. Lovász and M.D. Plummer, Matching Theory, (North-Holland, Amsterdam, 1986).

[14] M. Loebl and S. Poljak. Efficient subgraph packing. *Journal of Combinatorial Theory, Series B,* **59**(1993), 106-121.

[15] S. H. Pantel. Graph Packing Problems, M.Sc. Thesis, Simon Fraser University, 1998.

[16] R. E. Tarjan. *Data Structures and Network Algorithms,* Society for Industrial and Applied Mathematics, 1983.