RC23636 (W0506-085) June 22, 2005
Mathematics

IBM Research Report

On the *p*-Median Polytope of a Special Class of Graphs

Mourad Baïou, Francisco Barahona
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

ON THE p-MEDIAN POLYTOPE OF A SPECIAL CLASS OF GRAPHS

MOURAD BAÏOU AND FRANCISCO BARAHONA

ABSTRACT. In this paper we consider a well known class of valid inequalities for the p-median and the uncapacitated facility location polytopes, the odd cycle inequalities. It is known that their separation problem is polynomially solvable. We give a new polynomial separation algorithm based on a reduction from the original graph. Then, we define a nontrivial class of graphs, where the odd cycle inequalities together with the linear relaxations of both the p-median and uncapacitated facility location problems, suffice to describe the associated polytope. To do this, we first give a complete description of the fractional extreme points of the linear relaxation for the p-median polytope in that class of graphs.

1. Introduction

Let $G = (V, E)$ be a directed graph, not necessarily connected, where each arc $(u, v) \in E$ has an associated cost $c(u, v)$. The *p-median problem* (pMP) consist of selecting p nodes, usually called *centers*, and then assign each non-selected node to a selected node. The goal is to select p nodes that minimize the sum of the costs yield by the assignment of the non-selected nodes. This problem has several applications such as location of bank accounts [9], placement of web proxies in a computer network [22], semistructured data bases [21, 18]. When the number of centers is not specified and each opened center induces a given cost, this is called the *uncapacitated facility location problem* (UFLP).

In this paper we study the so called *odd cycle inequalities*. We give a new separation algorithm, and we show that when they are added to the linear programming relaxation of the pMP, a complete description of the polytope is obtained for the so called Y-free graphs. To accomplish that, first we have to characterize the extreme points of this linear programming relaxation for the class of Y-free graphs. Finally and for the same class of graphs, we show that when we add the odd cycle inequalities to the linear programming relaxation of the UFLP, we also obtain an integral polytope. We do not know of any other class of graphs for which the p-median polytope has been characterized.

The pMP is NP-hard in general [16]. For some particular cases it has been shown that it is polynomially solvable. This is the case when the underlying graph is an undirected tree and the cost function c is defined by $c(u, v) = w(u)d(u, v)$, where $w(u)$ is a positive weight associated with each node $u \in V$, and $d(u, v)$ is the length of the unique path in the tree from u to v, and it satisfies the triangle inequalities. An $O(p|V|^2)$ algorithm has been presented in [20] improving a previous algorithms in the same class of graphs, $O(p^2|V|^2)$ in [16] and $O(p|V|^3)$ in [14]. When the tree is directed, an $O(p^2L)$ algorithm based on dynamic programming is given in [22], where L is the path length of the tree.

Many heuristics without guaranties on the value of the solution have been given (see [9] for references). In [9] some known heuristics were evaluated using Lagrangian relaxation. This study has been used in [13] for evaluating heuristics based on a reduction from the set-covering problem. Later, α-approximation algorithms were developed, where α is not a constant factor, see [4, 5, 6]. The first constant-factor approximation algorithm was given in [8], with $\alpha = 6\frac{2}{3}$. Algorithms that achieve a performance within a ratio of 6 and 4 were presented in [15] and [7], respectively. Most of approximation algorithms are based on rounding the optimal fractional solution of the following natural linear programming relaxation for the pMP:

$$\text{minimize} \sum_{(u,v)\in E} c(u,v)x(u,v), \tag{1}$$

$$\sum_{v:(u,v)\in E} x(u,v) = 1 - y(u) \qquad \forall u \in V, \tag{2}$$

$$\sum_{v\in V} y(v) = p, \tag{3}$$

$$x(u,v) \leq y(v) \quad \forall (u,v) \in E, \tag{4}$$

$$0 \leq y(v) \leq 1 \qquad \forall v \in V, \tag{5}$$

$$x(u,v) \geq 0 \quad \forall (u,v) \in E. \tag{6}$$

If in addition the variables are 0-1, then we obtain an integer linear programming formulation. The 0-1 variable $y(v)$, $v \in V$ is 1 if the node v is selected as a center and 0 otherwise. The 0-1 variable $x(u,v)$ takes a the value 1 if a non-selected node u is assigned to a selected node v. Constraints (2) ensure that each node must be assigned to some center, constraints (3) ensure that exactly p centers must be selected and constraints (4) indicate that if a node v is not selected as center then no node u may be assigned to v.

Denote by $P_p(G)$ the polytope defined by constraints (2)-(6), and let $pMP(G)$ be the convex hull of $P_p(G) \cap \{0,1\}^{|E|+|V|}$, this is called the *$p$-median polytope of G*.

Let $Q(G)$ the polytope defined by constraints (2), (4), (5) and (6). Let $UFLP(G)$ be the convex hull of $Q(G) \cap \{0,1\}^{|E|+|V|}$, this is the *uncapacitated facility location polytope of G*.

Some polyhedral properties of $P_p(G)$ and the integrality gap are discussed in [23], when G is restricted to be a tree. This relaxation has also been studied for trees in [11]. In [12], an extended formulation by adding $O(|V|^{p-1})$ variables characterize the p-median polytope for any fixed p; also the 2-median polytope is characterized in the original set of variables.

A formulation based on the arc variables is studied in [2]. Then the relationship with the stable set problem is exploited. This approach was also used in [10] for the UFLP. Also in [2], they remark that the p-median polytope when $p = 2$ is completely described using a result in [17]. In [3] a Branch-and-Cut-and-Prize algorithm is developed to solve large instances for the p-median polytope.

We conclude this introduction with a few definitions. For a vector $x \in \mathbb{R}^S$ and a subset $A \subseteq S$, we denote $\sum_{a\in A} x(a)$ by $x(A)$. For a set $W \subset V$, we denote by $\delta^+(W)$, the set of arcs $(u,v) \in E$, with $u \in W$ and $v \in V \setminus W$, and by $\delta^-(W)$ the set of arcs (u,v), with $v \in W$ and $u \in V \setminus W$. We write $\delta^+(v)$ and $\delta^-(v)$ instead of $\delta^+(\{v\})$ and $\delta^-(\{v\})$, respectively. For a solution (x,y) of $P_p(G)$, define, for simplicity, $z = (x,y)$ as $z(u,v) = x(u,v)$ for $(u,v) \in E$ and $z(v) = y(v)$ for $v \in V$. If an inequality (4) is satisfied as equation by z with respect to an arc (u,v), we say that (u,v) is *saturated* by z, otherwise it is *non saturated* by z.

A directed graph $G = (V, E)$, not necessarily connected, is called a *Y-free graph* if $(u, v) \in E$ implies $(v, u) \notin E$ and it does not contain as induced subgraph the graph of Figure 1. This class of graphs may contain cycles and it contains the class of 1-rooted directed trees, Figure 2 shows a Y-free graph.

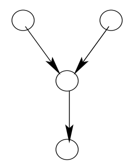

FIGURE 1. The graph Y.

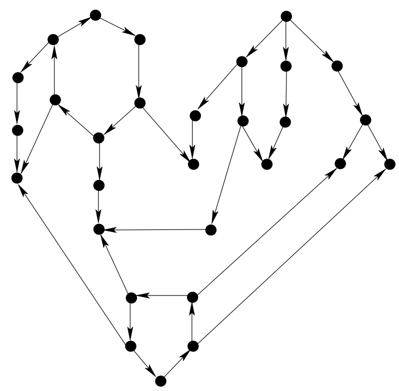

FIGURE 2. A Y-free graph.

This paper is organized as follows. In Section 2 we give a separation algorithm for the odd cycle inequalities. In Section 3 a complete description of the fractional extreme points of $P_p(G)$ is given for Y-free graphs. In Section 4 we show that when we add the odd cycle inequalities to (2)-(6), we obtain a complete description of $pMP(G)$ when G is Y-free. In Section 5 it is proved that when equation (3) is removed from the description of $pMP(G)$ and G is Y-free, then the polytope is still integral.

2. Odd cycle inequalities

In this section we describe this class of inequalities and their separation algorithm. Let C be the cycle

$$C = \{(v_i, v_{i+1}) \,|\, i = 1, \dots, 2l\} \cup \{(v_{2l+1}, v_1)\}.$$

The inequality

$$x(C) \leq \frac{|C| - 1}{2} \tag{7}$$

is called an *odd cycle* inequality.

Lemma 1. *Inequality* (7) *is valid for $pMP(G)$ and for $UFLP(G)$.*

Proof. The combination of inequalities (2) and (4) gives

$$x(u_i, u_{i+1}) + x(\delta^+(u_{i+1})) \leq 1, \text{ for } i = 1, \dots, 2l,$$
$$x(u_{2l+1}, u_1) + x(\delta^+(u_1)) \leq 1.$$

Adding these inequalities and non-negativity constraints we obtain

$$2x(C) \leq 2l + 1 = |C|.$$

Dividing by 2 and rounding down the right hand side yields inequality (7). $\square$

Odd cycle inequalities are a special case of Wq inequalities introduced in [3] with $q = 1$. Since we might have an exponential number of odd cycle inequalities, it is important to have an efficient algorithm to solve the *separation problem*: Given a vector $\bar{x}$ satisfying (2)-(6), find a violated odd cycle inequality if there is any, or prove that none exists. We describe such a procedure below.

The odd cycle inequalities we consider here, are equivalent to the odd cycle inequalities that are valid for the stable set polytope in the *intersection graph*. The intersection graph has a node for every arc in E, and for any two arcs of the form (u, v), (v, w) there is an edge in the intersection graph. Their separation can be reduced to $|E|$ shortest path problems in an auxiliary graph with positive arc-weights obtained from the intersection graph, see Theorem 68.1 in [19] for instance. Below, we give a reduction to $|V|$ shortest path problem in an auxiliary graph with no negative cycle, obtained from the original graph.

Inequality (7) can be written as

$$|C| - 2x(C) \geq 1. \tag{8}$$

To find a violated inequality, if there is any, we create an auxiliary graph as follows. For each node u, we make two copies u', u''. For each arc (u, v) we create arcs (u', v'') and (u'', v') with weights $1 - 2\bar{x}(u, v)$. Then for each node u we find a shortest path P from u' to u''. We identify every node in P with its copy, this gives a union of cycles, and and least 1 of them is odd. If the weight of P is less than 1, then we have found an odd cycle of weight less than 1. On the other hand, if for every node u the weight of a shortest path from u' to u'' is at least 1, then there is no violated odd cycle inequality.

Since the arc-weights could be negative, we should apply Bellman-Ford algorithm for finding a shortest path, see [1]. We have to see that this graph has no negative cycle.

Lemma 2. *The auxiliary graph has no cycle of negative weight.*

Proof. Let (u, v) and (v, w) be two consecutive arcs in a cycle C. It follows from (2) and (4) that $\bar{x}(u, v) + \bar{x}(v, w) \leq 1$. This implies

$$\bar{x}(C) \leq \frac{|C|}{2}. \tag{9}$$

If C is a cycle of negative weight, we have

$$|C| - 2\bar{x}(C) < 0,$$

or

$$\bar{x}(C) > \frac{|C|}{2},$$

this contradicts (9) $\square$

Since each shortest path computation takes $O(|V|^3)$ time, the entire separation algorithm takes $O(|V|^4)$ time.

Remark that the number of odd cycles in a Y-free graph is polynomially bounded. Indeed in a Y-free graph no two cycles can intersect, so an arc can belong to at most one cycle. Hence for this class of graphs the system defined by (2)-(6) and (7) has a polynomial number of inequalities.

3. The extreme points of $P_p(G)$, when G is a Y-free graph

In this section we give a description of the extreme points of $P_p(G)$, when G is a Y-free graph. Let $G = (V, E)$ be a Y-free graph. Let $z = (x, y) \in P_p(G)$. Let $E_z = \{(u, v) \in E : 0 < z(u, v) < 1\}$ and denote by $G_z = (V_z, E_z)$ the subgraph of G induced by E_z. Denote by $V_z^<$ the set of nodes v with $z(v) > 0$ such that either $z(u, v) < z(v)$ for $(u, v) \in E_z$ or $|\delta^-_{G_z}(v)| = 0$ or $z(v) = 1$. Call a node v with $z(v) = 1$ a *pendent node*. A directed path from v to w is denoted by P^v_w. Define the size of P^v_w to be the number of its inner nodes (nodes different from v and w). If the size of P^v_w is even (resp. odd) , we say that P^v_w is an *even path* (resp. *odd path*). Two paths are said to be *node-disjoint*, if the sets of their inner nodes are disjoint.

For a solution $z = (x, y)$ and a path $P^{v_1}_{v_k} = v_1, \ldots, v_k$, define $z[P^{v_1}_{v_k}]^\delta \in I\!R^{|E|+|V|}$ to be:

$$
\begin{array}{ll}
z[P^{v_1}_{v_k}]^\delta(v_1) = z(v_1) & \text{if } z(v_1) = 0,\\
z[P^{v_1}_{v_k}]^\delta(v_1) = z(v_1) - \delta & \text{if } z(v_1) > 0,\\
z[P^{v_1}_{v_k}]^\delta(v_i) = z(v_i) + (-1)^i\delta & \text{for } i = 2, \ldots, k-1\\
z[P^{v_1}_{v_k}]^\delta(v_k) = z(v_k), & \\
z[P^{v_1}_{v_k}]^\delta(v_i, v_{i+1}) = z(v_i, v_{i+1}) + (-1)^{i+1}\delta & \text{for } i = 1, \ldots, k-1,
\end{array}
$$

and $z[P^v_w]^\delta(u, v) = z(u, v)$, $z[P^v_w]^\delta(u) = z(u)$, for all other arcs and nodes (not in the path).

Let $G^1_z, \ldots, G^q_z$ be the connected components of G_z. In the following, we shall study the structure of these connected components when z is a fractional extreme point of $P_p(G)$. For this purpose, let us consider $G^1_z = (V^1_z, E^1_z)$.

Remark that for every node $v \in V$ that is not pendent, we have $|\delta^-_{G_z}(v)| \leq 1$. This remark is used implicitly in all the proofs of this section. Let $z = (x, y)$ be an extreme fractional point of $P_p(G)$.

Lemma 3. *G^1_z does not contain an odd path P^v_w with $v, w \in V^<_z$.*

Proof. Suppose that $P^v_w = v, v_1, v_2, \ldots, v_k, w$ is such a path, with k odd. Then the same constraints that are tight for z are also tight for $z[P^v_w]^{-\epsilon}$. This contradicts the fact that z is an extreme point of $P_p(G)$. □

Lemma 4. *G^1_z cannot contains two node-disjoint paths P^v_u and P^v_w having the same parity, where u, w are in $V^<_z$ and are not necessarily different.*

Proof. Let $P^v_u = v, u_1, u_2, \ldots, u_{k_1}, u$ and $P^v_w = v, w_1, w_2, \ldots, w_{k_2}, w$ be two node-disjoint paths such that k_1 and k_2 are of the same parity. Let $z_1 = z[P^v_u]^\epsilon$ and $z_2 = z_1[P^v_w]^{-\epsilon}$. Then the same constraints that are tight for z are also tight for z_2, a contradiction. □

Lemma 5. *Let $w, t \in V^1_z$ be two non necessarily different pendent nodes. If G^1_z contains two node-disjoint paths P^u_w and P^v_t, such that there is no saturated arc directed into u and v, then $z(u) = z(v) = 0$ and $u = v$.*

Proof. Let $P^u_w = u, u_1, u_2, \ldots, u_{k_1}, w$ and $P^v_t = v, v_1, v_2, \ldots, v_{k_2}, t$. Three cases are distinguished, as described by the figure below, (a) u and v in $V^<_z$, (b) $v \in V^<_z$ and $z(u) = 0$ and (c) $z(u) = z(v) = 0$ and $u \neq v$.

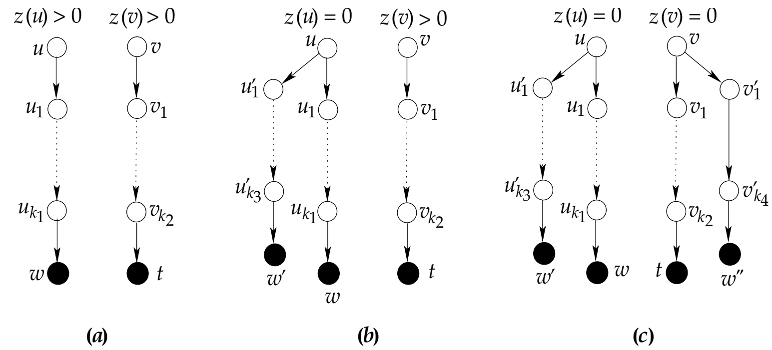

FIGURE 3.

(a) From Lemma 3, these two paths are even. And from Lemma 4 $u \neq v$. Let $z_1 = z[P^u_w]^\epsilon$. Then the same constraints that are tight for z are also tight for $z_1[P^v_t]^{-\epsilon}$, a contradiction.

(b) By definition, there must exist a path $P^u_{w'} = u, u'_1, \ldots, u'_{k_3}, w'$ from u to a pendent node w', w' may coincide with w and t, (see Figure 3 (b)). By Lemma 4, k_3 and k_1 are of different parity, and by Lemma 3, k_2 is even. Let $z_1 = z[P^u_{w'}]^\epsilon$, $z_2 = z_1[P^u_w]^{-\epsilon}$, $z_3 = z_2[P^v_t]^{-\epsilon}$, and $z_4 = z_2[P^v_t]^\epsilon$. Then the same constraints that are tight for z are also tight for either z_3 or z_4, again a contradiction.

(c) There must exist paths $P^u_{w'} = u, u'_1, \ldots, u'_{k_3}, w'$ and $P^v_{w''} = v, v'_1, \ldots, v'_{k_4}, w''$, where w' and w'' are pendent nodes, w, t, w' and w'' are not necessarily different, (see Figure 3 (c)). From Lemma 4, k_1 with k_3 and k_2 with k_4 are of different parity. Again we are going to construct a new vector z^* such that all constraints that are tight for z are also tight for z^*. Suppose for simplicity that k_3 and k_2 are odd and that k_1 and k_4 are even. The other cases may be treated similarly. Let $z_1 = z[P^u_{w'}]^\epsilon$, $z_2 = z_1[P^u_w]^{-\epsilon}$, $z_3 = z_2[P^v_t]^{-\epsilon}$, finally $z_3[P^v_{w''}]^\epsilon$ has the desired property. □

Lemma 6. *Let $C = \{(v_i, v_{i+1}) \,|\, i = 1, \ldots, 2l\} \cup \{(v_{2l+1}, v_1)\}$ be an odd cycle of G^1_z. Then at most one arc of C is not saturated.*

Proof. Suppose we have two arcs (v_i, v_{i+1}) and (v_j, v_{j+1}) not saturated, that is,

$$z(v_i, v_{i+1}) < z(v_{i+1}), \text{ and}$$

$$z(v_j, v_{j+1}) < z(v_{j+1}),$$

so v_{i+1} and v_{j+1} are in $V^<_z$. And since C is odd then, either $P^{v_{i+1}}_{v_{j+1}}$ or $P^{v_{j+1}}_{v_{i+1}}$ is odd. This contradicts Lemma 3. □

We say that $r \in V^1_z$ is a *root* if $|\delta^-_{G^1_z}(r)| = 0$.

Lemma 7. *If G^1_z contains a root r then it contains a directed path from r to a pendent node.*

Proof. Starting from r perform a depth-first search for a pendent node using the arcs of G^1_z. Since there is no Y the search should end on a pendent node. □

Lemma 8. *If G^1_z contains a root then it contains exactly one.*

Proof. Let r_1 and r_2 be two roots in G_z^1. Since G_z^1 has no Y, there is a directed path $P_{w_1}^{r_1}$ from r_1 to a pendent node w_1. Also there is a directed path $P_{w_2}^{r_2}$ from r_2 to a pendent node w_2 . These two paths are disjoint because there is no Y. Lemma 5 implies $r_1 = r_2$. □

Lemma 9. *In G_z^1 there is a path from r to any pendent node.*

Proof. Consider a pendent node v in G_z^1. There is an undirected path P from r to v. Let P_1 be the maximal directed path directed away from r using only arcs in P. Let w be the last node in P_1, since there is no Y, w is a pendent node. If $w = v$ we are done, otherwise let P_2 be the maximal path directed into v, included in P. Let s be the first node in P_2. Let P_3 the other maximal path directed away from s included in P. Since there is no Y, the last node in P_3 is pendent. From Lemma 4 we have that P_2 and P_3 have different parity. Let $z_1 = z[P_2]^{\epsilon}$, $z_2 = z_1[P_3]^{-\epsilon}$. If $z(r) > 0$, then from Lemma 3, P_1 must be even. Hence $z_2[P_1]^{\epsilon}$ or $z_2[P_1]^{-\epsilon}$ is a vector that satisfies as equation the same constraints as z.

Otherwise, if $z(r) = 0$, then there must be a path P_1' from r to pendent node, where P_1 and P_1' are node-disjoint. From Lemma 4, P_1 and P_1' have different parity. Define $z_3 = z_2[P_1]^{\epsilon}$ and $z_4 = z_2[P_1]^{-\epsilon}$, then either $z_3[P_1']^{-\epsilon}$ or $z_4[P_1']^{\epsilon}$ is a vector that satisfies as equation the same constraints as z. □

Now we have to study two cases as follows.

Case 1. G_z^1 does not contain a directed cycle.

In this case, G_z^1 must contain a root r. From Lemma 8, the root r is unique. From Lemma 4, $|\delta^+_{G_z^1}(r)| \le 2$. Otherwise, since G_z^1 is Y-free, there must exist two node-disjoint paths having the same parity from r to two pendent nodes w_1 and w_2.

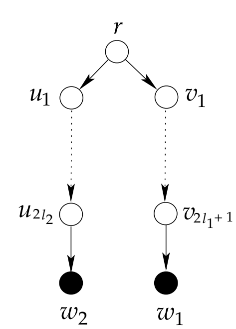

FIGURE 4. Two node-disjoint paths, $P_{w_1}^r$ and $P_{w_2}^r$, where $k_1 = 2l_1 + 1$, $k_2 = 2l_2$ and $P_{w_1}^r$ may be empty.

Proposition 10. *G_z^1 consists either of*

- *two paths $P_{w_1}^r = r, v_1, \dots, v_{k_1}, w_1$ and $P_{w_2}^r = r, u_1, \dots, u_{k_2}, w_2$ of different parity, or*
- *just the even path $P_{w_2}^r = r, u_1, \dots, u_{k_2}, w_2$, see Figure 4.*

Moreover, all the arcs, different from (v_{k_1}, w_1) *and* (u_{k_2}, w_2)*, are saturated by* z*, and*

$$\sum_{i=1}^{k_1} z(v_i) + \sum_{i=1}^{k_2} z(u_i) + z(r) = l_1 + l_2 + z(u_{k_2}), \text{ if } G_z^1 \text{ consists of two paths,} \tag{10}$$

$$\sum_{i=1}^{k_2} z(u_i) + z(r) = l_2 + z(u_{k_2}), \text{ if } G_z^1 \text{ consists of an even path.} \tag{11}$$

Proof. If $|\delta^+(r)| = 2$, then there are two node-disjoint paths to some pendent nodes. If $|\delta^+(v)| \geq 2$ for a node v in any of these paths, then there are two node-disjoint paths P_1 and P_2 from v to some pendent nodes. Lemma 4 implies that P_1 and P_2 have different parity. This implies that there are two node-disjoint paths of the same parity from r to some pendent nodes, a contradiction.

If $|\delta^+(r)| = 1$, then there is a directed path from r to a pendent node. If $|\delta^+(v)| \geq 2$ for a node v in this path, then there are two node-disjoint paths P_1 and P_2 from v to some pendent nodes. Lemma 4 implies that P_1 and P_2 have different parity. In this case $z(r) > 0$, thus Lemma 3 yields a contradiction.

Let S be the set of nodes in V_z^1 that does not belong to any path from r to a pendent node. Since G_z^1 is connected, there must exist a node $s \in S$ incident to one of the nodes in a path from r to a pendent node. Since G is Y-free, we must have an arc (s, w) and w is a pendent node. Since s is not a root, there must exist an arc directed into s, repeating this process we must end up with the root r, which is a contradiction since s does not belong to any path from r to a pendent node, or we must have a directed cycle, which is impossible.

We have exactly one root and $|\delta^+_{G_z^1}(v)| = 1$ for all non-pendent nodes $v \in V_z^1 \setminus \{r\}$. It follows that, if $|\delta^+_{G_z^1}(r)| = 2$, then G_z^1 consists of two node-disjoint paths from r to two pendent nodes or to the same pendent node. From Lemma 4, these two paths must be of different parity. If $|\delta^+_{G_z^1}(r)| = 1$, G_z^1 is a path from r to a pendent node. Since $z(r) > 0$, it follows from Lemma 3 that this path is even.

Let $k_1 = 2l_1 + 1$ and $k_2 = 2l_2$. Remark that if $|\delta^+_{G_z^1}(r)| = 2$ then $z(r) = 0$, otherwise this will contradict Lemma 3. Suppose $z(v_l, v_{l+1}) < z(v_{l+1})$, for $l \in \{1, \ldots, k_1 - 1\}$. So $v_{l+1} \in V_z^<$. It follows from Lemma 3 that the path $P_{w_1}^{l+1}$ has to be even. Since k_1 is odd then, both l and k_2 are even, which contradicts Lemma 4. In the same manner it may be shown that all the arcs, different from (v_{k_1}, w_1) and (u_{k_2}, w_2), are saturated by z.

Now, let us verify (10). From above we have,

$$z(r, v_1) = z(v_1), \tag{12}$$
$$z(r, u_1) = z(u_1), \tag{13}$$
$$z(v_i, v_{i+1}) = z(v_{i+1}) \quad \text{for } i = 1, \ldots, k_1 - 1, \tag{14}$$
$$z(u_i, u_{i+1}) = z(u_{i+1}) \quad \text{for } i = 1, \ldots, k_2 - 1, \tag{15}$$

and equalities (2) with respect to r, v_i, $i = 1, \ldots, k_1 - 1$, and u_i, $i = 1, \ldots, k_2 - 1$, give

$$z(r, v_1) + z(r, u_1) = 1 - z(r), \tag{16}$$
$$z(v_i, v_{i+1}) = 1 - z(v_i) \quad \text{for } i = 1, \ldots, k_1 - 1, \tag{17}$$
$$z(u_i, u_{i+1}) = 1 - z(u_i) \quad \text{for } i = 1, \ldots, k_2 - 1. \tag{18}$$

The combination of equations (14) with (17) and (15) with (18) gives,

$$z(v_{i+1}) = 1 - z(v_i) \quad \text{for } i = 1, \dots, k_1 - 1, \tag{19}$$

$$z(u_{i+1}) = 1 - z(u_i) \quad \text{for } i = 1, \dots, k_2 - 1. \tag{20}$$

The sum of equations (12)-(15) is equal to the sum of equations (16)-(18), hence

$$\sum_{i=1}^{k_1} z(v_i) + \sum_{i=1}^{k_2} z(u_i) = k_1 + k_2 - 1 - z(r) - \sum_{i=1}^{k_1-1} z(v_i) - \sum_{i=1}^{k_2-1} z(u_i). \tag{21}$$

Recall that $k_1 = 2l_1 + 1$ and $k_2 = 2l_2$. Now by considering (19) and (20), equation (21) may be rewritten as follows

$$\sum_{i=1}^{k_1} z(v_i) + \sum_{i=1}^{k_2} z(u_i) + z(r) = 2l_1 + 2l_2 - l_1 - l_2 + z(u_{k_2}),$$

hence

$$\sum_{i=1}^{k_1} z(v_i) + \sum_{i=1}^{k_2} z(u_i) + z(r) = l_1 + l_2 + z(u_{k_2}).$$

Equation (11) can be obtained in a similar way. □

Case 2. G_z^1 contains a directed cycle $C = \{(v_i, v_{i+1}) \,|\, i = 1, \dots, k-1\} \cup \{(v_k, v_1)\}$.

If $k = 2l$ consider the vector $\bar{z}$ defined below. The same constraints that are tight for z are also tight for $\bar{z}$.

$$\begin{array}{ll} \bar{z}_{v_i} = z(v_i) + (-1)^i \epsilon & \text{for } i = 1, \dots, k, \\ \bar{z}(v_i, v_{i+1}) = z(v_i, v_{i+1}) + (-1)^{i+1}\epsilon & \text{for } i = 1, \dots, k-1, \\ \bar{z}(v_k, v_1) = z(v_k, v_1) - \epsilon, & \end{array}$$

and $\bar{z}(u,v) = z(u,v)$, $\bar{z}(u) = z(u)$ for all other nodes and arcs not in C.

So we can assume that C is odd. Let $k = 2l + 1$ and $V(C) = \{v_1, \dots, v_{2l+1}\}$. From Lemma 6, we have to consider two sub cases:

Case 2.1. C contains exactly one non saturated arc.

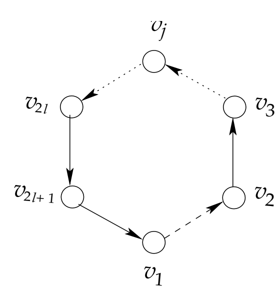

FIGURE 5. Odd cycle C, $\delta^+(V(C)) = \emptyset$. A dashed arc means a non saturated arc.

Proposition 11. *We should have that $\delta^+_{G^1_z}(V(C)) = \emptyset$, and if (v_1, v_2) is the non saturated arc then*

$$\sum_{i=1}^{2l+1} z(v_i) = l + z(v_2). \tag{22}$$

Proof. Note that $v_2 \in V_z^<$. Let v_m a node of $V(C)$ such that $|\delta^+_{G^1_z}(v_m)| \geq 2$, v_m may coincide with v_1 or v_2. Since G is Y-free, there is a path $P_w^{v_m}$ having all its nodes not in $V(C)$ with w a pendent node. We have four cases to consider, $P_w^{v_m}$ is odd or even, and m is odd or even. Suppose that $P_w^{v_m}$ is odd, either m is odd, and in this case the paths $P_{v_2}^{v_m}$ and $P_w^{v_m}$ are both odd, which contradicts Lemma 4, or m is even and in this case the path defined by the junction of $P_{v_m}^{v_2}$ and $P_w^{v_m}$ is odd, which contradicts Lemma 3. The same arguments hold when $P_w^{v_m}$ is even. So we have proved that $\delta^+(V(C)) = \emptyset$.

From equations (1),

$$\begin{aligned} &z(v_i, v_{i+1}) = 1 - z(v_i) \qquad &&\text{for } i = 2, \dots, 2l \\ &z(v_{2l+1}, v_1) = 1 - z(v_{2l+1}) && \end{aligned}$$

Also, $z(v_i, v_{i+1}) = z(v_{i+1})$ for each arc of C different from (v_1, v_2). The combination of these equations gives

$$\sum_{i=2}^{2l+1} (1 - z(v_i)) = \sum_{i=1, i\neq 2}^{2l+1} z(v_i) \tag{23}$$

we also have,

$$\begin{aligned} &z(v_{i+1}) = 1 - z(v_i) \qquad &&\text{for } i = 2, \dots, 2l \\ &z(v_1) = 1 - z(v_{2l+1}), && \end{aligned}$$

these imply that $(1 - z(v_i)) + (1 - z(v_{i+1})) = 1$, for $i = 1, \dots, 2l$. Hence the equation (23) is equivalent to

$$\sum_{i=1, i\neq 2}^{2l+1} z(v_i) = l.$$

□

Case 2.2 All the arcs of C are saturated.

Note that in this case, if one of the node-variables or arc-variables of C has a value $1/2$, then all the other node-variables and arc-variables have the same value and that G^1_z consists of only the cycle C. In general, the structure of G^1_z is described by the following proposition, see the figure below.

Proposition 12. *We have that $|\delta^+_{G^1_z}(V(C))| \leq 1$. And if $|\delta^+_{G^1_z}(V(C))| = 1$, then G^1_z consists of the odd cycle C and a path $P_w^{v_m} = v_m, u_1, \dots, u_k, w$, going from a node $v_m \in V(C)$ to a pendent node w, $m \in \{1, \dots, 2l+1\}$. Moreover, all the arcs different from (u_k, w) are saturated and*

$$\sum_{i=1}^{2l+1} z(v_i) + \sum_{i=1}^{k} z(u_i) = l + l_1 + z(v_m) \qquad \text{if } k = 2l_1, \tag{24}$$

$$\sum_{i=1}^{2l+1} z(v_i) + \sum_{i=1}^{k} z(u_i) = l + l_1 + 1 - z(v_m, v_{m+1}) \quad \text{if } k = 2l + 1. \tag{25}$$

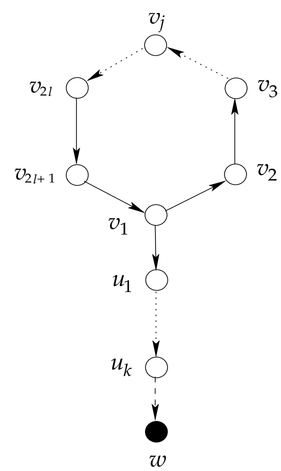

FIGURE 6. Odd cycle C, $|\delta^+(V(C))| = 1$, all arcs are saturated except the dashed arc .

Proof. Suppose that $|\delta^+_{G^1_{\bar z}}(V(C))| \geq 2$. Let (v_{i_1}, u) and (v_{i_2}, u') be two arcs in $\delta^+_{G^1_{\bar z}}(V(C))$. We must have two node-disjoint paths: P_1 from v_{i_1} to a pendent node w_1, containing (v_{i_1}, u), and P_2 from v_{i_2} to a pendent node w_2, containing (v_{i_2}, u'). If $i_1 \neq i_2$, then one may define two node-disjoint paths having the same parity, going from v_{i_1} (or v_{i_2}), to w_1 and w_2, which contradicts Lemma 4. If $i_1 = i_2$, then by Lemma 4, P_1 and P_2 are of different parity. For simplicity let $i_1 = i_2 = 1$. Let P'_1 and P'_2 be the portions of P_1 and P_2, from u to w_1 and from u' to w_2, respectively. Define $z_1 = z[P'_1]^{+\epsilon}$, $z_2 = z_1[P'_2]^{+\epsilon}$, and $\bar z$ to be

$$
\begin{array}{ll}
\bar z(v_1, u) = z_2(v_1, u) - \epsilon, & \\
\bar z(v_1, u') = z_2(v_1, u') - \epsilon, & \\
\bar z(v_1) = z_2(v_1) + \epsilon, & \\
\bar z(v_1, v_2) = z_2(v_1, v_2) + \epsilon, & \\
\bar z(v_i) = z_2(v_i) + (-1)^i \epsilon & \text{for } i = 2, \dots, 2l+1, \\
\bar z(v_i, v_{i+1}) = z_2(v_i, v_{i+1}) + (-1)^{i+1} \epsilon & \text{for } i = 2, \dots, 2l, \\
\bar z(v_{2l+1}, v_1) = z_2(v_{2l+1}, v_1) + \epsilon, &
\end{array}
$$

and $\bar z(u) = z_2(u)$ (resp. $\bar z(u, v) = z_2(u, v)$) for all other nodes (resp. arcs).

Since P_1 and P_2 are of different parity, $\bar z$ also satisfies (3). We have that the constraints that are tight for z are also tight for $\bar z$, which contradicts the fact that z is an extreme point of $P_p(G)$.

Thus it may be assumed that $|\delta^+_{G^1_{\bar z}}(V(C))| = 1$. Let (v_m, u_1) be the unique arc of $\delta^+_{G^1_{\bar z}}(V(C))$. For simplicity, take $m = 1$. There must exist a path from v_1 to a pendent node, call it $P^{v_1}_w = v_1, u_1, \dots, u_k, w$. Suppose there exists a node u_r, $1 \leq r \leq k$, with $|\delta^+_{G^1_{\bar z}}(u_r)| \geq 2$, then there are two node-disjoint paths P_1 and P_2 from u_r to some pendent nodes. Lemma 4 implies that P_1 and P_2 have different parity. Suppose r odd. Define

$z_1 = z[P^{u_1}_{u_r}]^{-2\epsilon}$, $z_2 = z_1[P_1]^{-\epsilon}$, $z_3 = z_2[P_2]^{-\epsilon}$ and let . Define $\bar{z}$ to be

$$\begin{array}{ll}
\bar{z}(v_1, u_1) = z_3(v_1, u_1) + 2\epsilon, & \\
\bar{z}(v_1) = z_3(v_1) - \epsilon, & \\
\bar{z}(v_1, v_2) = z_3(v_1, v_2) - \epsilon, & \\
\bar{z}(v_i) = z_3(v_i) + (-1)^{i+1}\epsilon & \text{for } i = 2, \dots, 2l+1, \\
\bar{z}(v_i, v_{i+1}) = z_3(v_i, v_{i+1}) + (-1)^{i}\epsilon & \text{for } i = 2, \dots, 2l, \\
\bar{z}(v_{2l+1}, v_1) = z_3(v_{2l+1}, v_1) - \epsilon, &
\end{array}$$

and $\bar{z}(u) = z_3(u)$ (resp. $\bar{z}(u, v) = z_3(u, v)$), for all other nodes (resp. arcs). Thus $\bar{z}$ and z satisfy the same constraints as equations, this contradicts again the fact that z is an extreme point. Using the same ideas, one obtain the same contradiction when r is even.

Thus, it may be assumed that G^1_z consists of only the cycle C and the path $P^{v_1}_w$. Suppose (u_r, u_{r+1}) is non saturated by z, for $0 \le r \le k-1$, where $u_0 = v_1$. It follows from Lemma 3 that $P^{u_{r+1}}_w$ is even.

Define $z_1 = z[P^{u_1}_{u_{r+1}}]^{-2\epsilon}$, $z_2 = z_1[P^{u_{r+1}}_w]^{+\epsilon}$ if $P^{u_1}_{u_{r+1}}$ is odd, otherwise $z_2 = z_1[P^{u_{r+1}}_w]^{-\epsilon}$ and let $\bar{z}$ to be

$$\begin{array}{ll}
\bar{z}(v_1, u_1) = z_2(v_1, u_1) + 2\epsilon, & \\
\bar{z}(v_1) = z_2(v_1) - \epsilon, & \\
\bar{z}(v_1, v_2) = z_2(v_1, v_2) - \epsilon, & \\
\bar{z}(v_i) = z_2(v_i) + (-1)^{i+1}\epsilon & \text{for } i = 2, \dots, 2l+1, \\
\bar{z}(v_i, v_{i+1}) = z_2(v_i, v_{i+1}) + (-1)^{i}\epsilon & \text{for } i = 2, \dots, 2l, \\
\bar{z}(v_{2l+1}, v_1) = z_2(v_{2l+1}, v_1) - \epsilon, &
\end{array}$$

and $\bar{z}(u) = z_2(u)$ (resp. $\bar{z}(u, v) = z_2(u, v)$), for all other nodes (resp. arcs). We have that constraints (2)-(6) that are tight for z are also tight for $\bar{z}$, which contradicts the fact that z is an extreme point.

Thus every arc, different from (u_k, w), is saturated, so

$$z(v_1, u_1) = z(u_1), \tag{26}$$
$$z(u_{i-1}, u_i) = z(u_i) \quad \text{for } i = 2, \dots, k, \tag{27}$$
$$z(v_i, v_{i+1}) = z(v_{i+1}) \quad \text{for } i = 1, \dots, 2l, \tag{28}$$
$$z(v_{2l+1}, v_1) = z(v_1). \tag{29}$$

Now we have to see that (24) and (25) hold. Equations (2) with respect to u_i, $i = 1, \dots, k-1$, and v_i, $i = 1, \dots, 2l+1$, give

$$z(v_1, u_1) + z(v_1, v_2) = 1 - z(v_1), \tag{30}$$
$$z(v_i, v_{i+1}) = 1 - z(v_i) \quad \text{for } i = 2, \dots, 2l, \tag{31}$$
$$z(v_{2l+1}, v_1) = 1 - z(v_{2l+1}), \tag{32}$$
$$z(u_i, u_{i+1}) = 1 - z(u_i) \quad \text{for } i = 1, \dots, k-1. \tag{33}$$

The combination of equations (26)-(29) with (30)-(33) gives,

$$z(u_1) + z(v_2) = 1 - z(v_1), \tag{34}$$
$$z(v_{i+1}) = 1 - z(v_i) \text{ for } i = 1, \dots, 2l, \tag{35}$$
$$z(u_{i+1}) = 1 - z(u_i) \text{ for } i = 1, \dots, k-1. \tag{36}$$

The sum of equations (26)-(29) is equal to the sum of equations (30)-(33), hence

$$\sum_{i=1}^{2l+1} z(v_i) + \sum_{i=1}^{k} z(u_i) = (2l+1) - \sum_{i=1}^{2l} z(v_i) - z(v_{2l+1}) + (k-1) - \sum_{i=1}^{k-1} z(u_i). \tag{37}$$

By considering (34)-(36), equation (37) may be rewritten as follows.
If $k = 2l_1$:

$$\begin{aligned} \sum_{i=1}^{2l+1} z(v_i) + \sum_{i=1}^{k} z(u_i) &= (2l+1) - (1 - z(u_1)) - (l-1) - z(v_{2l+1}) \\ &+ (k-1) - z(u_1) - (l_1 - 1), \\ &= l + l_1 + 1 - z(v_{2l+1}), \\ &= l + l_1 + z(v_1). \end{aligned}$$

If $k = 2l_1 + 1$:

$$\begin{aligned} \sum_{i=1}^{2l+1} z(v_i) + \sum_{i=1}^{k} z(u_i) &= (2l+1) - (1 - z(u_1)) - (l-1) - z(v_{2l+1}) \\ &+ 2l_1 - l_1, \\ &= l + l_1 + 1 + z(u_1) - z(v_{2l+1}), \\ &= l + l_1 + z(u_1) + z(v_1), \\ &= l + l_1 + 1 - z(v_1, v_2). \end{aligned}$$

□

Theorem 13. *Let G be a Y-free graph. Let $z = (x, y)$ be a fractional extreme point of $P_p(G)$. Then the following hold:*

- *(i) G_z contains q connected components, $G_z^1, \ldots, G_z^q$, with $q \geq 2$,*
- *(ii) G_z contains at most one component that corresponds to one of the graphs of Figures 4, 5. The others components are all odd cycles where each arc is saturated by z.*
- *(iii) The values of z are 0, 1 or $1/2$.*

Proof. (i) If G_z is connected. Then from Propositions 10, 11 and 12, G_z is one of the graphs of Figures 4, 5 and 6 or an odd cycle where each node and arc is associated with the value $1/2$. These propositions also show that $\sum_{v \in V_z} z(v)$ is fractional, since $z(v) = 0$ or 1 for every $v \notin V_z$, we have that $\sum_{v \in V} z(v)$ is fractional, which is impossible.

(ii) Let G_z^1 and G_z^2 be two connected components of G_z. Consider the case where G_z^1 and G_z^2 are the graphs of Figures 4 and 5, respectively, the other cases may be treated similarly. Recall that G_z^1 consists of two node-disjoint paths $P_{w_1}^r$ and $P_{w_2}^r$ having different parity and G_z^2 consists of an odd cycle $C = \{(v_i, v_{i+1}) \mid i = 1, \ldots, 2l\} \cup \{(v_{2l+1}, v_1)\}$ where (v_1, v_2) is non saturated by z. Define recursively a new vector $\bar{z}$ as follows. $z_1 = z[P_{w_1}^r]^{-\epsilon}$, $z_2 = z_1[P_{w_2}^r]^{+\epsilon}$, $\bar{z} = z_2[P_{v_2}^{v_2}]^{-\epsilon}$. Any arc saturated by z remains saturated by $\bar{z}$. Moreover, equation 3 holds. In fact,

$$\begin{aligned}
\sum_{v\in V}\bar z(v) &= \sum_{v\in V_z^1}\bar z(v) + \sum_{v\in V_z^2}\bar z(v) + \sum_{v\in V\setminus(V_z^1\cup V_z^2)}\bar z(v)\\
&= \sum_{v\in V_z^1} z(v) - \epsilon + \sum_{v\in V_z^2} z(v) + \epsilon + \sum_{v\in V\setminus(V_z^1\cup V_z^2)} z(v)\\
&= \sum_{v\in V} z(v) = p.
\end{aligned}$$

We conclude that every constraint that is tight for z is also tight for $\bar z$, which contradicts the fact that z is an extreme point.

Suppose that G_z^1 is the graph of Figure 6, that is G_z^1 is an odd cycle $C = \{(v_i, v_{i+1}) \mid i = 1,\dots,2l\} \cup \{(v_{2l+1}, v_1)\}$ and a path $P_w^{v_m} = v_m, u_1, \dots, u_k, w$ for $v_m \in V(C)$. If G_z^2 is one of the graphs of Figures 4, 5 or 6, then as above we can define a solution $\bar z$ such that the same constraints that are tight for z are also tight for $\bar z$. Thus all other connected components, $G_z^2, \dots, G_z^q$ of G_z consist of an odd cycle where each arc is saturated and that $z(v) = 1/2$ for all $v \in (V_z \setminus V_z^1)$. By Proposition 12 $\sum_{v\in V_z^1} z(v)$ must be fractional. Hence $q-1$ must be odd, and since $\sum_{v\in V_z} z(v)$ is integer, by Proposition 12 we have $z(v_m) = 1/2$ if k is even, otherwise $z(v_m, v_{m+1}) = 1/2$. In both cases, since all the arcs of C are saturated, we must have $z(v_m, u_1) = 0$. Thus G_z^1 is also an odd cycle where each arc is saturated.

(iii) By the definition of G_z, z takes the values 0 or 1 for any arc or node not in G_z. Now (iii) is a straightforward consequence of (ii) and Propositions 10, 11, 12. □

Corollary 14. *If G is a Y-free graph without odd cycles, then $pMP(G)$ is completely described by constraints (2)-(6). That is, $P_p(G)$ is integral.*

4. Description of $pMP(G)$, when G is a Y-free graph

We show that the addition of the odd cycle inequalities (7) to (2)-(6), completely describes $pMP(G)$ when G is a Y-free graph.

Call $PC_p(G)$ the polytope described by constraints (2)-(6) and inequalities (7).

Theorem 15. *If G is a Y-free graph then, $pMP(G) = PC_p(G)$.*

Proof. The proof is by induction on the number of arcs. Obviously, the theorem is true for small graphs, with no more than 2 arcs. Suppose it is true for any Y-free graph with no more than m arcs and let G contain exactly $m+1$ arcs. Suppose $pMP(G) \neq PC_p(G)$, and let $z = (x, y)$ be a fractional extreme point of $PC_p(G)$.

Claim 1. $0 < z(u,v) < 1$ *for all* $(u,v) \in E$.

Proof.

i) Let $(u,v) \in E$ with $z(u,v) = 0$. Let $G' = \big(V, E \setminus (u,v)\big)$, and z' the restriction of z to G'. It is clear that $z' \in PC_p(G')$. Suppose that $z' = 1/2 z_1 + 1/2 z_2$ where $z_1, z_2 \in PC_p(G')$, $z_1 \neq z_2$. Let $\bar z_1$ (resp. $\bar z_2$) be the vector obtained by adding a zero component to z_1 (resp. z_2). We have that $z = 1/2\bar z_1 + 1/2\bar z_2$.

Now let us see that $\bar z_1$ and $\bar z_2$ are in $PC_p(G)$. Clearly they satisfy (2)-(6), so we just have to see that constraints (7) are satisfied. Consider the odd cycle

$$C = \{(w_i, w_{i+1}) \mid i = 1,\dots,2l\} \cup \{(u,v)\},$$

where $u = w_{2l+1}$ and $v = w_1$. We have that

$$z_1(w_{2i-1}, w_{2i}) + z_1(w_{2i}, w_{2i+1}) \leq 1, \text{ for } i = 1, \ldots, l.$$

This implies $\bar{z}_1(C) \leq l$. The same is true for $\bar{z}_2$. Therefore $\bar{z}_1$ and $\bar{z}_2$ are in $PC_p(G)$.

We have then a contradiction because z is an extreme point. So z' must be an extreme point of $PC_p(G')$ and because of the induction hypothesis, it must be integral.

ii) Let $(u, v) \in E$ with $z(u, v) = 1$. This implies $z(u) = 0$. From i) we have that u is incident only to (u, v). Let $G' = \big(V \setminus \{u\}, E \setminus (u, v)\big)$, and z' the restriction of z to G'. It is clear that $z' \in PC_p(G')$. Suppose that $z' = 1/2z_1 + 1/2z_2$ where $z_1, z_2 \in PC_p(G')$, $z_1 \neq z_2$. Let $\bar{z}_1$ (resp. $\bar{z}_2$) be the vector obtained by adding a zero and a one component to z_1 (resp. z_2). We have that $z = 1/2\bar{z}_1 + 1/2\bar{z}_2$.

First we should see that $\bar{z}_1$ and $\bar{z}_2$ are in $PC_p(G)$. Clearly they satisfy (2)-(6), and since (u, v) does not belong to any cycle, constraints (7) are satisfied.

We have then a contradiction because z is an extreme point. So z' must be an extreme point of $PC_p(G')$, and because of the induction hypothesis it must be integral. □

In what follows, we will show that G contains no odd cycle. Let us assume the contrary, and let $C = \{(v_i, v_{i+1}) \,|\, i = 1, \ldots, 2l\} \cup \{(v_{2l+1}, v_1)\}$ be an odd cycle.

Remark 16. $0 < z(v_i) < 1$*, for* $i = 1, \ldots, 2l + 1$.

This remark follows from Claim 1.

Remark 17. *Since G is Y-free, if C_1 and C_2 are two directed cycles in G, then $C_1 \cap C_2 = \emptyset$.*

Claim 2. *At most one arc of C is non saturated by z.*

Proof. Suppose that we have two arcs (v_i, v_{i+1}) and (v_j, v_{j+1}) that are not saturated by z, that is,

$$z(v_i, v_{i+1}) < z(v_{i+1}), \text{ and}$$
$$z(v_j, v_{j+1}) < z(v_{j+1}).$$

Since C is odd, then either $P^{v_{i+1}}_{v_{j+1}}$ or $P^{v_{j+1}}_{v_{i+1}}$ is odd. Suppose $P^{v_{i+1}}_{v_{j+1}}$. By Claim 1, $0 < z(u, v) < 1$ for every arc (u, v) in $P^{v_{i+1}}_{v_{j+1}}$ and by Remark 16, $0 < z(v) < 1$ for every node v in $V(C)$. Also by Remark 17, the nodes in $P^{v_{i+1}}_{v_{j+1}}$ do not belong to another cycle. It follows that $z[P^{v_{i+1}}_{v_{j+1}}]^{+\epsilon}$ tight the same tight constraints by z, which contradicts the fact that z is an extreme point of $PC_p(G)$. □

Claim 3. *If P_1 and P_2 are two paths going from $V(C)$ to some pendent nodes, where all the inner nodes of both P_1 and P_2 are not in $V(C)$, then P_1 and P_2 cannot have the same parity.*

Proof. Let $P^{v_k}_{u_{l_1}} = v_k, u_1, \ldots, u_{l_1}$, $P^{v_m}_{u'_{l_2}} = v_m, u'_1, \ldots, u'_{l_2}$, be two paths having the same parity, with $z(u_{l_1}) = z(u'_{l_2}) = 1$, $k \leq m$. Notice that (v_k, u_1) and (v_m, u'_1) are in $\delta^+(V(C))$. Where l_1 and l_2 may be equal to 1, and u_{l_1}, u'_{l_2} may coincide. Remark that from Claim 1, we have $0 < z(v) < 1$ for any node v in P_1 and P_2 (v is not a pendent node).

Let $P^{v_k}_{u'_{l_2}}$ be the path obtained by joining the path in C from v_k to v_m and $P^{v_m}_{u'_{l_2}}$ and let $P^{v_m}_{u_{l_1}}$ be obtained by joining the path in C from v_m to v_k and $P^{v_k}_{u_{l_1}}$. Since C is odd, then either $P^{v_k}_{u'_{l_2}}$ and $P^{v_k}_{u_{l_1}}$ or $P^{v_m}_{u_{l_1}}$ and $P^{v_m}_{u'_{l_2}}$, are of the same parity.

Suppose that $P^{v_k}_{u'_{l_2}}$ and $P^{v_k}_{u_{l_1}}$ have the same parity. Consider $z_1 = z[P^{v_k}_{u'_{l_2}}]^{\epsilon}$ and $z_2 = z_1[P^{v_k}_{u_{l_1}}]^{-\epsilon}$, then z_2 and z satisfy the same constraints with equality. Notice that since G is Y-free, the nodes and arcs in C, $P^{v_k}_{u'_{l_2}}$ and $P^{v_k}_{u_{l_1}}$ cannot appear in any other cycle. A similar proof can be used if $P^{v_m}_{u_{l_1}}$ and $P^{v_m}_{u'_{l_2}}$ have the same parity. □

From Claim 2, we distinguish two cases:

(a). All arcs of C are saturated by z. By Claim 3 we have that $|\delta^+(V(C))| \leq 2$. Suppose that $|\delta^+(V(C))| = 2$. We can assume that $\delta^+(V(C)) = \{(v_1, u), (v_{2k+1}, w)\}$. We have

$$x(v_{2i-1}, v_{2i}) + x(v_{2i}, v_{2i+1}) = 1,$$

for $i = 1, \dots, l$. which implies $x(C) = l + x(v_{2l+1}, v_1)$. Thus one of inequalities (7) is violated. If $|\delta^+(V(C))| \leq 1$ a similar proof can be used.

(b). There is exactly one arc in C that is not saturated by z. First we are going to prove that any path from $V(C)$ to a pendent node is even.

Let (v_{2l+1}, v_1) be the non saturated arc in C. Let $P^{v_{2k}}_w$ be an odd path from v_{2k} to a pendent node w. Let $z_1 = z[P^{v_{2k}}_w]^{\epsilon}$, and $z_2 = z_1[P^{v_{2k}}_{v_1}]^{-\epsilon}$. Here $P^{v_{2k}}_{v_1}$ denotes the path in C from v_{2k} to v_1. Since G is Y-free, the nodes and arcs in C and $P^{v_{2k}}_w$ cannot appear in any other cycle. We have that z and z_2 satisfy the same constraints as equation, a contradiction.

Now let $P^{v_{2k+1}}_w$ be an odd path from v_{2k+1} to a pendent node w. Let P be the path obtained by joining the path in C from v_1 to v_{2k+1} and $P^{v_{2k+1}}_w$. Let $z_1 = z[P]^{\epsilon}$, the vectors z and z_1 satisfy the same constraints as equation.

This shows that any path from $V(C)$ to a pendent node is even, and by Claim 3 we have that $|\delta^+(V(C))| \leq 1$.

If $\delta^+(V(C)) = \{(v_{2k+1}, w)\}$, we have that

$$x(v_{2i-1}, v_{2i}) + x(v_{2i}, v_{2i+1}) = 1,$$

for $i = 1, \dots, l$, which implies $x(C) = l + x(v_{2l+1}, v_1)$. Thus one of inequalities (7) is violated.

If $\delta^+(V(C)) = \{(v_{2k}, w)\}$, we have that

$$\begin{aligned} &x(v_{2i-1}, v_{2i}) + x(v_{2i}, v_{2i+1}) = 1, \text{ for } i = 1, \dots, k-1, \\ &x(v_{2i}, v_{2i+1}) + x(v_{2i+1}, v_{2i+2}) = 1, \text{ for } i = k, \dots, l-1, \\ &x(v_{2l}, v_{2l+1}) + x(v_{2l+1}, v_1) = 1. \end{aligned}$$

Therefore $x(C) = l + x(v_{2k-1}, v_{2k})$. Thus one of inequalities (7) is violated.

Finally if $\delta^+(V(C)) = \emptyset$, in a similar way we obtain $x(C) = l + x(v_{2l+1}, v_1)$.

Thus G contains no odd cycle. From Corollary 14, z is integer. A contradiction. □

5. The Uncapacitated Facility Location polytope

Now we study the case when equation (3) is removed from the definition of $PC_p(G)$. Let $G = (V, E)$ be a graph, and let G' be the graph obtained by adding to G a new component consisting of the nodes u and v and the arc (u, v). For a vector z associated with G' let z^G be the restriction of z to G. Let $\Pi(G)$ be the polytope defined by (2), (4), (5), (6) and (7). Let $\Phi_p(G')$ be the polytope defined by

$$z^G \in \Pi(G) \tag{38}$$

$$z(V \cup \{u, v\}) = p \tag{39}$$

$$z(u, v) = 1 - z(u) \tag{40}$$

$$0 \le z(u, v) \le z(v) = 1. \tag{41}$$

Lemma 18. *If z^* is an extreme point of $\Pi(G)$, then $\bar{z}$ is is an extreme point of $\Phi_q(G')$, where*

$$\begin{aligned}
&q = \lceil z^*(V) \rceil + 1\\
&\bar{z}(w) = z^*(w) \textit{ for } w \in V,\\
&\bar{z}(w,t) = z^*(w,t) \textit{ for } (w,t) \in E,\\
&\bar{z}(u) = q - z^*(V) - 1\\
&\bar{z}(u,v) = 1 - \bar{z}(u)\\
&\bar{z}(v) = 1.
\end{aligned}$$

Proof. Clearly $\bar{z} \in \Phi_q(G')$. Suppose that $\bar{z} = 1/2z_1 + 1/2z_2$, with $z_1, z_2 \in \Phi_q(G')$. Then $z^* = 1/2z_1^G + 1/2z_2^G$, and $z_1^G = z_2^G$ since z^* is an extreme point. We have $z_1(v) = z_2(v) = 1$. Equation (39) implies $z_1(u) = z_2(u)$ and $z_1(u, v) = z_2(u, v)$. Thus $\bar{z}$ is an extreme point. □

The lemma above shows that any fractional extreme point of $\Pi(G)$ can be completed to a fractional extreme point of $\Phi_q(G')$. If G is Y-free then G' is also Y-free and $\Phi_q(G')$ is integral. This shows the following.

Theorem 19. *If G is Y-free then the polytope $\Pi(G)$ is integral.*

Theorem 20. *The pMP and the UFLP are polynomially solvable for the class of Y-free graphs.*

Acknowledgments

Part of this work was done while the first author was visiting the T. J. Watson Research Center of IBM and while the second author was visiting the Laboratoire d'Econométrie de L'Ecole Polytechnique in Paris. The financial support of both institutions is greatly appreciated.

References

[1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network flows*, Prentice Hall Inc., Englewood Cliffs, NJ, 1993.

[2] P. Avella and A. Sassano, *On the p-median polytope*, Math. Program., 89 (2001), pp. 395–411.

[3] P. Avella, A. Sassano, and I. Vasil'ev, *Computational study of large-scale p-median problems*, Technical Report 08-03, DIS, Universita di Roma "La Sapienza", 2003. Available at http://www.optimization-online.org.

[4] Y. Bartal, *Probabilistic approximation of metric spaces and its algorithmic applications*, in 37th Annual Symposium on Foundations of Computer Science (Burlington, VT, 1996), IEEE Comput. Soc. Press, Los Alamitos, CA, 1996, pp. 184–193.
[5] ———, *On approximating arbitrary metrices by tree metrics*, in STOC '98 (Dallas, TX), ACM, New York, 1999, pp. 161–168.
[6] M. Charikar, C. Chekuri, A. Goel, and S. Guha, *Rounding via trees: deterministic approximation algorithms for group Steiner trees and k-median*, in STOC '98 (Dallas, TX), ACM, New York, 1999, pp. 114–123.
[7] M. Charikar and S. Guha, *Improved combinatorial algorithms for the facility location and k-median problems*, in 40th Annual Symposium on Foundations of Computer Science (New York, 1999), IEEE Computer Soc., Los Alamitos, CA, 1999, pp. 378–388.
[8] M. Charikar, S. Guha, É. Tardos, and D. B. Shmoys, *A constant-factor approximation algorithm for the k-median problem*, J. Comput. System Sci., 65 (2002), pp. 129–149. Special issue on STOC, 1999 (Atlanta, GA).
[9] G. Cornuejols, M. L. Fisher, and G. L. Nemhauser, *Location of bank accounts to optimize float: an analytic study of exact and approximate algorithms*, Management Sci., 23 (1976/77), pp. 789–810.
[10] G. Cornuejols and J.-M. Thizy, *Some facets of the simple plant location polytope*, Math. Programming, 23 (1982), pp. 50–74.
[11] S. de Vries, M. Posner, and R. Vohra, *Polyhedral properties of the k -median problem on a tree*, Tech. Rep. 1367, Northwestern University, Center for Mathematical Studies in Economics and Management Science, Apr. 2003. available at http://ideas.repec.org/p/nwu/cmsems/1367.html.
[12] M. Goemans, *Notes on the k-median polytope on trees*. 1992.
[13] D. S. Hochbaum, *Heuristics for the fixed cost median problem*, Math. Programming, 22 (1982), pp. 148–162.
[14] W. L. Hsu, *The distance-domination numbers of trees*, Oper. Res. Lett., 1 (1981/82), pp. 96–100.
[15] K. Jain and V. V. Vazirani, *Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation*, J. ACM, 48 (2001), pp. 274–296.
[16] O. Kariv and S. L. Hakimi, *An algorithmic approach to network location problems. II. The p-medians*, SIAM J. Appl. Math., 37 (1979), pp. 539–560.
[17] J.-F. Maurras, *Convex hull of the edges of a graph and near bipartite graphs*, Discrete Math., 46 (1983), pp. 257–265.
[18] S. Nestorov, S. Abiteboul, and R. Motwani, *Extracting schema from semistructured data*, in SIGMOD '98: Proceedings of the 1998 ACM SIGMOD international conference on Management of data, New York, NY, USA, 1998, ACM Press, pp. 295–306.
[19] A. Schrijver, *Combinatorial optimization. Polyhedra and efficiency. Vol. B*, vol. 24 of Algorithms and Combinatorics, Springer-Verlag, Berlin, 2003. Matroids, trees, stable sets, Chapters 39–69.
[20] A. Tamir, *An $O(pn^2)$ algorithm for the p-median and related problems on tree graphs*, Oper. Res. Lett., 19 (1996), pp. 59–64.
[21] F. Toumani, *Personal communication*. 2002.
[22] A. Vigneron, L. Gao, M. J. Golin, G. F. Italiano, and B. Li, *An algorithm for finding a k-median in a directed tree*, Inform. Process. Lett., 74 (2000), pp. 81–88.
[23] R. Ward, P. Lemke, and O. A, *Properties of the k-median linear programming relaxation*, technical report, Rensselaer Polytechnic Institute, 1988.

(M. Baïou) Cust Université Clermont II and Laboratoire d'Econométrie de l'Ecole Polytechnique, 1 rue Descartes , 75005 Paris, France

E-mail address, M. Baïou: baiou@custsv.univ-bpclermont.fr

(F. Barahona) IBM T. J. Watson research Center, Yorktown Heights, NY 10589, USA

E-mail address, F. Barahona: barahon@us.ibm.com