RC24683 (W0810-164) October 29, 2008
Mathematics

IBM Research Report

On the Location and p-Median Polytopes

Mourad Baïou
CNRS
LIMOS Complexe Scientifique des Cezeaux
63177 Aubière
France

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

ON THE LOCATION AND P-MEDIAN POLYTOPES

MOURAD BAÏOU AND FRANCISCO BARAHONA

Abstract. We revisit classical systems of linear inequalities associated with location problems and with the p-median problem. We present an overview of the cases for which these linear systems define integral polytopes. We also give polynomial time algorithms to recognize these cases.

1. Introduction

Facility location and p-median are among the most well-studied problems in combinatorial optimization. They are both NP-hard, so there is not much hope of having a complete polyhedral characterization of them. The linear programming relaxations that we use have been known since the 60's and have been the basis for many heuristics, branch and bound algorithms, and approximation algorithms. Despite all this work, very little is known about special cases where these formulations give integral polytopes, and also there are not many special cases where the associated polytope has been completely characterized. We have found a characterization of the graphs for which these linear relaxations define polytopes with all extreme points being integral. Here we present an overview of all these cases. We also give polynomial time algorithms to recognize these classes of graphs. Our characterization shows the basic structures that a graph contains when the polytope has fractional extreme points.

We first deal with location problems, we show that the linear relaxation gives an integral polytope if and only the graph does not contain a certain type of "odd" cycles. Then we deal with the p-median problem. We show that there are five configurations that should be forbidden in order to have an integral polytope. Here the proof consists of three parts as follows. First we show the result for the so-called Y-free graphs. We denote by Y some basic configuration in the graph. The result on Y-free graphs is used to start an induction proof for oriented graphs. These are directed graphs where between any two nodes u and v, at most one of the arcs (u, v) and (v, u) exists. Here the induction is done on the number of Y configurations. The third part consists of extending our result to general directed graphs. Here the induction is done on the number of pairs of nodes u and v such that both (u, v) and (v, u) exist. The initial step of the induction is given by the result on oriented graphs.

This paper is organized as follows. Section 2 contains some definitions. Section 3 deals with location problems. Section 4 covers the p-median problem. In Section 5 we give an algorithm to recognize the graphs defined in Section 4. Section 6 is devoted to some extensions.

Date: June 25, 2008.
Key words and phrases. location problems, uncapacitated facility location, p-median.

2. Preliminary definitions

A directed graph $G = (V, A)$ is called *oriented* if $(u, v) \in A$ implies $(v, u) \notin A$. For a directed graph $G = (V, A)$ and a set $W \subset V$, we denote by $\delta^+(W)$ the set of arcs $(u, v) \in A$, with $u \in W$ and $v \in V \setminus W$. Also we denote 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. If there is a risk of confusion we use δ^+_G and δ^-_G. A node u with $\delta^+(u) = \emptyset$ is called a *pendent* node.

A simple cycle C is an ordered sequence

$$v_0, a_0, v_1, a_1, \ldots, a_{p-1}, v_p,$$

where

- v_i, $0 \le i \le p-1$, are distinct nodes,
- a_i, $0 \le i \le p-1$, are distinct arcs,
- either v_i is the tail of a_i and v_{i+1} is the head of a_i, or v_i is the head of a_i and v_{i+1} is the tail of a_i, for $0 \le i \le p-1$, and
- $v_0 = v_p$.

By setting $a_p = a_0$, we associate with C three more sets as below.

- We denote by $\hat{C}$ the set of nodes v_i, such that v_i is the head of a_{i-1} and also the head of a_i, $1 \le i \le p$.
- We denote by $\dot{C}$ the set of nodes v_i, such that v_i is the tail of a_{i-1} and also the tail of a_i, $1 \le i \le p$.
- We denote by $\tilde{C}$ the set of nodes v_i, such that either v_i is the head of a_{i-1} and also the tail of a_i, or v_i is the tail of a_{i-1} and also the head of a_i, $1 \le i \le p$.

Notice that $|\hat{C}| = |\dot{C}|$. A cycle will be called *odd* if $p + |\dot{C}|$ (or $|\tilde{C}| + |\dot{C}|$) is odd, otherwise it will be called *even*. A cycle C with $\dot{C} = \emptyset$ is a *directed* cycle. The set of arcs in C is denoted by $A(C)$.

If we do not require $v_0 = v_p$ we have a *path* P. In a similar way we define $\dot{P}$, $\hat{P}$ and $\tilde{P}$, excluding v_0 and v_p. We say that P is *odd* if $p + |\dot{P}|$ is odd, otherwise it is *even*. For the path P, the nodes $v_1, \ldots, v_{p-1}$ are called *internal*.

If G is a connected graph and there is a node u such that its removal disconnects G, we say that u is an *articulation point*. A graph is said to be *two-connected* if at least two nodes should be removed to disconnect it. For simplicity, sometimes we use z to denote the vector (x, y), i.e., $z(u) = y(u)$ and $z(u, v) = x(u, v)$. Also for $S \subseteq V \cup A$ we use $z(S)$ to denote $z(S) = \sum_{a \in S} z(a)$.

A *polyhedron* P is a set defined by a system of linear inequalities, i.e., $P = \{x \mid Ax \le b\}$. A *face* of P is obtained by setting into equation some of these inequalities. An *extreme point* of P is given by a face that contains a unique element. In other words, some inequalities are set to equation so that this system has a unique solution. A *polytope* is a bounded polyhedron. A polyhedron is called *integral* if all its extreme points are integral.

3. Location problems

Let $G = (V, A)$ be a directed graph, not necessarily connected, where each arc and each node has weight associated with it. We study a “prize collecting” version of a *location problem* (LP) as follows. A set of nodes is selected, usually called *centers*, and then each non-selected node can be assigned to a center. The weight of a node is the revenue obtained by opening a facility at that location, minus the cost of building the facility. The weight of an arc (i, j) is the revenue obtained by assigning the location i to the location j, minus the cost originated by this assignment. The goal is to maximize the sum of the weights of the selected nodes plus the sum of the weights yielded by the assignment. The linear system below defines a linear programming relaxation.

$$\max \sum w(u,v)x(u,v) + \sum w(v)y(v)$$

$$\sum_{(u,v)\in A} x(u,v) + y(u) \le 1 \quad \forall u \in V, \tag{1}$$

$$x(u,v) \le y(v) \quad \forall (u,v) \in A, \tag{2}$$

$$0 \le y(v) \le 1 \quad \forall v \in V, \tag{3}$$

$$x(u,v) \ge 0 \quad \forall (u,v) \in A. \tag{4}$$

For each node u, the variable $y(u)$ takes the value 1 if the node u is selected and 0 otherwise. For each arc (u, v) the variable $x(u, v)$ takes the value 1 if u is assigned to v and 0 otherwise. Inequalities (1) express the fact that either node u can be selected or it can be assigned to another node. Inequalities (2) indicate that if a node u is assigned to a node v then this last node should be selected. The set of integer vectors that satisfy (1)-(4) corresponds to a *transitive packing* as defined in [21].

Let $P(G)$ be the polytope defined by (1)-(4), and let $LP(G)$ be the convex hull of $P(G) \cap \{0,1\}^{|V|+|A|}$. Clearly

$$LP(G) \subseteq P(G).$$

Here we characterize the graphs G for which $LP(G) = P(G)$. More precisely, we show that $LP(G) = P(G)$ if and only if G does not contain an odd cycle. We also give a polynomial algorithm to recognize the graphs in this class.

The *Uncapacitated Facility Location Problem* (UFLP) is a variation where V is partitioned into V_1 and V_2. The set V_1 corresponds to the customers, and the set V_2 corresponds to the potential facilities. Each customer in V_1 should be assigned to an opened facility in V_2. This is obtained by considering $A \subseteq V_1 \times V_2$, fixing to zero the variables y for the nodes in V_1 and setting into equation the inequalities (1) for the nodes in V_1. More precisely, the linear programming relaxation for this case is

$$\min \sum c(u,v)x(u,v) + \sum d(v)y(v)$$

$$\sum_{(u,v)\in A} x(u,v) = 1 \quad \forall u \in V_1, \tag{5}$$

$$x(u,v) \le y(v) \quad \forall (u,v) \in A, \tag{6}$$

$$0 \le y(v) \le 1 \quad \forall v \in V_2, \tag{7}$$

$$x(u,v) \ge 0 \quad \forall (u,v) \in A. \tag{8}$$

Here we also characterize the cases for which (5)-(8) defines an integral polytope.

We omit the proofs of several technical lemmas, the full details appear in [3]. The facets of the uncapacitated facility location polytope have been studied in [18], [15], [9], [10], [7]. In [2] we gave a description of $LP(G)$ for Y-free graphs. The UFLP has also been studied from the point of view of approximation algorithms in [23], [11], [25], [6] and others. Other references on this problem are [14] and [20]. The relationship between location polytopes and the stable set polytope has been studied in [15], [9], [10], [17], and others. It would be interesting to know if our results also have an equivalent in terms of stable set polytopes, but so far we have not found the right transformation.

3.1. **Decomposition.** In this subsection we consider a graph $G = (V, A)$ that decomposes into two graphs $G_1 = (V_1, A_1)$ and $G_2 = (V_2, A_2)$, with $V = V_1 \cup V_2$, $V_1 \cap V_2 = \{u\}$, $A = A_1 \cup A_2$, $A_1 \cap A_2 = \emptyset$. We define G'_1 that is obtained from G_1 after replacing u by u'. We also define G'_2, obtained from G_2 after replacing u by u''. The theorem below shows that we have to concentrate on two-connected graphs.

Theorem 1. *Suppose that the system*

$$Az' \leq b \tag{9}$$

$$z'\left(\delta^+_{G'_1}(u')\right) + z'(u') \leq 1 \tag{10}$$

describes $LP(G'_1)$. *Suppose that* (9) *contains the inequalities* (1)-(4) *except for* (10). *Similarly suppose that*

$$Cz'' \leq d \tag{11}$$

$$z''\left(\delta^+_{G'_2}(u'')\right) + z''(u'') \leq 1 \tag{12}$$

describes $LP(G'_2)$. *Also* (11) *contains the inequalities* (1)-(4) *except for* (12). *Then the system below describes an integer polytope.*

$$Az' \leq b \tag{13}$$

$$Cz'' \leq d \tag{14}$$

$$z'\left(\delta^+_{G'_1}(u')\right) + z''\left(\delta^+_{G'_2}(u'')\right) + z'(u') \leq 1 \tag{15}$$

$$z'(u') = z''(u'') \tag{16}$$

We have the following corollary.

Corollary 2. *The polytope $LP(G)$ is defined by the system (13)-(16) after identifying the variables $z'(u')$ and $z''(u'')$.*

This last corollary shows that if $LP(G'_1)$ and $LP(G'_2)$ are defined by (1)-(4), then $LP(G)$ is also defined by (1)-(4). Thus we have to concentrate on graphs that are two-connected. A result analogous to Theorem 1, for the stable set polytope, has been given in [12].

3.2. **Graph Transformations.** First we plan to prove that if G has no odd cycle then $LP(G) = P(G)$. The proof consists of assuming that $\bar{z}$ is a fractional extreme point of $P(G)$ and arriving at a contradiction. Below we give several assumptions that can be made about $\bar{z}$ and G, they will be used in the next subsection. The proofs of the lemmas below consist of modifying the graph and the vector $\bar{z}$ so that we obtain a new extreme point associated with a new graph satisfying the assumptions below.

Lemma 3. *We can assume that G consists of only one connected component.*

Lemma 4. *If $0 < \bar{z}(u, v) < \bar{z}(v)$, we can assume that v is a pendent node with $|\delta^-(v)| = 1$ and $\bar{z}(v) = 1$.*

Lemma 5. *We can assume that $0 < \bar{z}(u, v) < 1$ for all $(u, v) \in A$.*

Lemma 6. *We can assume that G is either two-connected or it consists of a single arc.*

If the graph G consists of a single arc it is fairly easy to see that $LP(G) = P(G)$, so now we have to deal with the two-connected components. This is treated in the next subsection.

3.3. **Treating two-connected graphs.** In this subsection we assume that the graph G is two-connected and it has no odd cycle. Let $\bar{z}$ be a fractional extreme point of $P(G)$, we are going to assign labels l to the nodes and arcs and define $z'(u,v) = \bar{z}(u,v) + l(u,v)\epsilon$, $z'(u) = \bar{z}(u) + l(u)\epsilon$, $\epsilon > 0$, for each arc (u,v) and each node u. We shall see that every constraint that is satisfied with equality by $\bar{z}$ is also satisfied with equality by z'. This is the required contradiction.

Given a path $P = v_0, a_0, \ldots, a_{p-1}, v_p$. Assume that the label of a_0, $l(a_0)$ has the value 1 or -1. We define the *labeling procedure* as follows.

For $i = 1$ to $p-1$ do

- If v_i is the head of a_{i-1} and it is the tail of a_i then $l(v_i) = l(a_{i-1})$, $l(a_i) = -l(a_{i-1})$.
- If v_i is the head of a_{i-1} and it is the head of a_i then $l(v_i) = l(a_{i-1})$, $l(a_i) = l(a_{i-1})$.
- If v_i is the tail of a_{i-1} and it is the head of a_i then $l(v_i) = -l(a_{i-1})$, $l(a_i) = -l(a_{i-1})$.
- If v_i is the tail of a_{i-1} and it is the tail of a_i then $l(v_i) = 0$, $l(a_i) = -l(a_{i-1})$.

Notice that the labels of v_0 and v_p were not defined.

We have to study several cases as follows.

Case 1. G contains a directed cycle $C = v_0, a_0, \ldots, a_{p-1}, v_p$. Assume that the head of a_0 is v_1, set $l(v_0) = -1$, $l(a_0) = 1$ and extend the labels as above.

Case 2. G contains a cycle $C = v_0, a_0, \ldots, a_{p-1}, v_p$ and $\dot{C} \neq \emptyset$. Assume $v_0 \in \dot{C}$. Set $l(v_0) = 0$, $l(a_0) = 1$ and extend the labels.

The lemma below is needed to show that for v_0, the constraints that were satisfied with equality by $\bar{z}$ remain satisfied with equality.

Lemma 7. *After labeling as in Cases 1 and 2 we have* $l(a_{p-1}) = -l(a_0)$.

Notice that after the first cycle has been labeled as in Cases 1 or 2, the properties below hold, we shall see that these properties hold throughout the entire labeling procedure.

Property 1. If a node has a nonzero label, then it is the tail of at most one labeled arc.

Property 2. If a node has a zero label, then it is the tail of exactly two labeled arcs.

Once a cycle C has been labeled as in Cases 1 or 2, we have to extend the labeling as follows.

Case 3. Suppose that $l(v_0) \neq 0$ for $v_0 \in C$, (v_0 is the head of a labeled arc), and there is a path $P = v_0, a_0, v_1, a_1, \ldots,\ a_{p-1}, v_p$ in G such that:

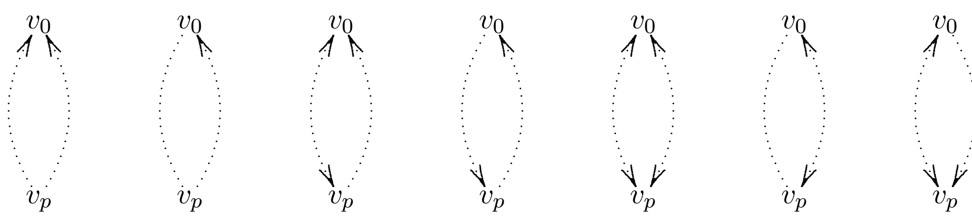
- v_0 is the head of a_0,
- $v_p \in C$,
- $\{v_1, \ldots, v_{p-1}\}$ is disjoint from C.

We set $l(a_0) = l(v_0)$ and extend the labels. Case 3 is needed so that any inequality (2) associated with v_0 that is satisfied with equality, remains satisfied with equality.

We have to see that the label $l(a_{p-1})$ is such that constraints associated with v_p that were satisfied with equality remain satisfied with equality. This is discussed in the next lemma.

Lemma 8. *If v_p is the head of a_{p-1} then $l(a_{p-1}) = l(v_p)$. If v_p is the tail of a_{p-1} then $l(a_{p-1}) = -l(v_p)$.*

Proof. cf. [3]. Notice that $v_0 \notin \dot{C}$, in Figure 1 we represent the possible configurations for the paths in C between v_0 and v_p. In this figure we show whether v_0 and v_p are the head or the tail of the arcs in C incident to them. These two paths are denoted by P_1 and P_2.



(1) (2) (3) (4) (5) (6) (7)

FIGURE 1. Possible paths in C between v_0 and v_p. It is shown whether v_0 and v_p are the head or the tail of the arcs in C incident to them.

Consider configuration (1), these two paths should have different parity. When adding the path P, an odd cycle is created with either P_1 or P_2. So configuration (1) will not occur. The same happens with configuration (2).

Now we discuss configuration (3). These two paths should have the same parity. If v_p is the tail of a_{p-1} then P would create an odd cycle with either P_1 or P_2. If v_p is the head of a_{p-1} then P should have the same parity as P_1 and P_2. Then $l(a_{p-1}) = l(v_p)$.

The study of configuration (4) is similar. The two paths should have the same parity. If v_p is the tail of a_{p-1} then P would create an odd cycle with either P_1 or P_2. If v_p is the head of a_{p-1} then P should have the same parity as P_1 and P_2, and $l(a_{p-1}) = l(v_p)$.

For configuration (5) again the two paths should have the same parity. If v_p is the head of a_{p-1} then P should have the same parity as P_1 and P_2, and $l(a_{p-1}) = l(v_p)$. If v_p is the tail of a_{p-1} then P should have the same parity as P_1 and P_2, and $l(a_{p-1}) = -l(v_p)$.

Also in configuration (6) the paths P_1 and P_2 should have the same parity. If v_p is the tail of a_{p-1} then P would form an odd cycle with either P_1 or P_2. If v_p is the head of a_{p-1} then P should have the same parity as P_1 and P_2, and $l(a_{p-1}) = l(v_p)$.

In configuration (7) also the two paths should have the same parity. If v_p is the head of a_{p-1} then P should have the same parity as P_1 and P_2, and $l(a_{p-1}) = l(v_p)$. If v_p is the tail of a_{p-1} then P should have the same parity as P_1 and P_2, and $l(a_{p-1}) = -l(v_p)$. $\square$

Based on this the labels are extended successively. Denote by G_l the subgraph defined by the labeled arcs. This is a two-connected graph, so for any two nodes v_0 and v_p it contains a cycle going through these two nodes. Thus we can check if Case 3 applies and extend the labels adding each time a path to the graph G_l. The two lemmas below show that Properties 1 and 2 remain satisfied.

Lemma 9. *Let v_p be a node with $l(v_p) \neq 0$. If v_p is the tail of an arc in G_l, then in Case 3 it cannot be the tail of a_{p-1}. Thus Property 1 remains satisfied.*

Lemma 10. *Let v_p be a node with $l(v_p) = 0$, thus v_p is the tail of exactly two arcs in G_l. Then in Case 3 it cannot be the tail of a_{p-1}. Therefore Property 2 remains satisfied.*

Once Case 3 has been exhausted we might have some nodes in G_l that are not pendent in G and that are only the head of labeled arcs. For such nodes we have to ensure that inequalities (1) that were satisfied as equality remain satisfied as equality. This is treated in the following.

Case 4. Suppose that v_0 is only the head of labeled arcs, ($l(v_0) \neq 0$), v_0 is not pendent. We have that $\delta^+(v_0) \neq \emptyset$ thus there is a cycle C in G_l and there is a path $P = v_0, a_0, v_1, a_1, \dots, a_{p-1}, v_p$ in G such that:

- $v_0 \in C$ is the tail of a_0,
- $v_p \in C$,
- $\{v_1, \dots, v_{p-1}\}$ is disjoint from G_l.

We set $l(a_0) = -l(v_0)$ and extend the labels. We have to see that the label $l(a_{p-1})$ is such that constraints associated with v_p, that were satisfied with equality, remain satisfied with equality. This is discussed below.

Lemma 11. *In Case 4 we have that v_p is the tail of a_{p-1} and $l(a_{p-1}) = -l(v_p)$. Also Properties 1 and 2 continue to hold.*

To summarize, the labeling algorithm consists of the following steps.

- Step 1. Identify a cycle C in G and treat it as in Cases 1 or 2. Set $G_l = C$.
- Step 2. For as long as needed label as in Case 3. Each time add to G_l the new set of labeled nodes and arcs.
- Step 3. If needed, label as in Case 4. Each time add to G_l the new set of labeled nodes and arcs. If some new labels have been assigned in this step go to Step 2, otherwise stop.

At this point we can discuss the properties of the labeling procedure. The labels are such that any inequality (2) that was satisfied with equality by $\bar{z}$ is also satisfied with equality by z'. To see that inequalities (1) that were tight remain tight, we need two observations about G_l:

- Any node that has a nonzero label is the tail of exactly one labeled arc having the opposite label.
- If u is a node with $l(u) = 0$, then there are exactly two labeled arcs having opposite labels and whose tail is u.

Finally we give the label "0" to all nodes and arcs that are unlabeled, this completes the definition of z'. Lemma 5 shows that inequalities (4) will not be violated. The fact that nodes v with $\bar{z}(v) = 0$ or $\bar{z}(v) = 1$ receive a zero label, shows that inequalities (3) will not be violated. Any constraint that is satisfied with equality by $\bar{z}$ is also satisfied with equality by z', this contradicts the assumption that $\bar{z}$ is an extreme point. We can state the main result of this subsection.

Theorem 12. *If the graph G is two-connected and has no odd cycle then $LP(G) = P(G)$.*

This implies the following.

Theorem 13. *If G is a graph with no odd cycle, then $LP(G) = P(G)$.*

Theorem 14. *For graphs with no odd cycle, the uncapacitated facility location problem is polynomially solvable.*

3.4. **Odd cycles.** In this subsection we study the effect of odd cycles in $P(G)$. Let C be an odd cycle. We can define a fractional vector $(\bar{x}, \bar{y}) \in P(G)$ as follows:

$$\bar{y}(u) = 0 \quad \text{for all nodes } u \in \dot{C}, \tag{17}$$

$$\bar{y}(u) = 1/2 \quad \text{for all nodes } u \in C \setminus \dot{C}, \tag{18}$$

$$\bar{x}(a) = 1/2 \quad \text{for } a \in A(C), \tag{19}$$

$$\bar{y}(v) = 0 \quad \text{for all other nodes } v \notin C, \tag{20}$$

$$\bar{x}(a) = 0 \quad \text{for all other arcs.} \tag{21}$$

Below we show a family of inequalities that separate the vectors defined above from $LP(G)$. We call them *odd cycle* inequalities.

Lemma 15. *The following inequalities are valid for $LP(G)$.*

$$\sum_{a \in A(C)} x(a) - \sum_{v \in \hat{C}} y(v) \leq \frac{|\tilde{C}| + |\hat{C}| - 1}{2} \tag{22}$$

for every odd cycle C.

These inequalities are $\{0, 1/2\}$-Chvatal-Gomory cuts, using the terminology of [8]. A separation algorithm can be obtained from the results of [8]. In [3] we gave an alternative separation algorithm.

Now we can present the following result.

Theorem 16. *Let G be a directed graph, then $LP(G) = P(G)$ if and only if G does not contain an odd cycle.*

Proof. cf. [3]. If G contains and odd cycle C, then we can define a vector $(\bar{x}, \bar{y}) \in P(G)$ as in (17)-(21). We have

$$\sum_{a \in A(C)} \bar{x}(a) - \sum_{v \in \hat{C}} \bar{y}(v) = \frac{|\tilde{C}| + |\hat{C}|}{2}.$$

Lemma 15 shows that $\bar{z} \notin LP(G)$.

Then the theorem follows from Theorem 13. □

3.5. **Detecting odd cycles.** Now we study how to recognize the graphs G for which $LP(G) = P(G)$. We start with a graph G and several transformations are needed.

The first transformation consists of building an undirected graph $H = (N, E)$. For every node $u \in G$ we have the nodes u' and u'' in N, and the edge $u'u'' \in E$. For every arc $(u, v) \in G$ we have an edge $u'v'' \in E$. See Figure 2.

Consider a cycle C in G, we build a cycle C_H in H as follows.

- If (u, v) and (u, w) are in C, then the edges $u'v''$ and $u'w''$ are taken.
- If (u, v) and (w, v) are in C, then the edges $u'v''$ and $v''w'$ are taken.
- If (u, v) and (v, w) are in C, then the edges $u'v''$, $v''v'$, and $v'w''$ are taken.

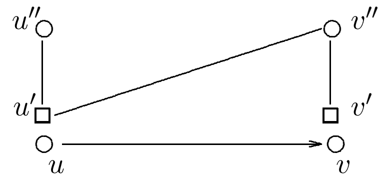


FIGURE 2

On the other hand, a cycle in H corresponds to a cycle in G. Thus there is a one to one correspondence among cycles of G and cycles of H. Moreover, if the cycle in H has cardinality $2q$, then $q = |\dot{C}| + |\tilde{C}|$, where C is the corresponding cycle in G. Therefore an odd cycle in G corresponds to a cycle in H of cardinality $2(2p+1)$ for some positive integer p. See Figure 3.

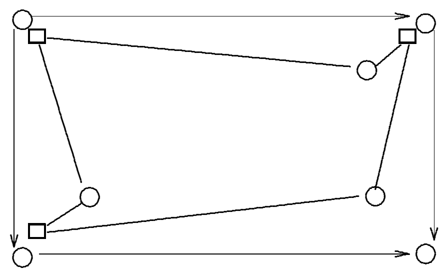


FIGURE 3. An odd cycle in G and the corresponding cycle in H. The nodes of H close to a node $u \in G$ correspond to u' or u''.

In other words, finding an odd cycle in G reduces to finding a cycle of cardinality $2(2p+1)$, for some positive integer p, in the bipartite graph H.

For this question, a linear time algorithm was given in [28], a simple $O(|V||A|^2)$ has been given in [13].

3.6. **Uncapacitated Facility Location.** Now we assume that V is partitioned into V_1 and V_2, $A \subseteq V_1 \times V_2$, and we deal with the system

$$\sum_{(u,v)\in A} x(u,v) = 1 \quad \forall u \in V_1, \tag{23}$$

$$x(u,v) \le y(v) \quad \forall (u,v) \in A, \tag{24}$$

$$0 \le y(v) \le 1 \quad \forall v \in V_2, \tag{25}$$

$$x(u,v) \ge 0 \quad \forall (u,v) \in A. \tag{26}$$

If the variables x and y are constrained to be integer, then we have the uncapacitated facility location problem (UFLP). We denote by $\Pi(G)$ the polytope defined by (23)-(26). Notice that $\Pi(G)$ is a face of $P(G)$. Let $\bar{V}_1$ be the set of nodes $u \in V_1$ with $|\delta^+(u)| = 1$. Let $\bar{V}_2$ be the set of nodes in V_2 that are adjacent to a node in $\bar{V}_1$. It is clear that the variables associated with nodes in $\bar{V}_2$ should be fixed, i.e., $y(v) = 1$ for all $v \in \bar{V}_2$. Let us denote by $\bar{G}$ the subgraph induced by $V \setminus \bar{V}_2$. In this section we prove that $\Pi(G)$ is an integer polytope if and only if $\bar{G}$ has no odd cycle.

Let us first assume that $\bar{G}$ has no odd cycle. As before, we suppose that $\bar{z}$ is a fractional extreme point of $\Pi(G)$. The analogues of lemmas 3, 4 and 5 apply here. Thus we can assume that we deal with a connected component G'. Lemma 4 implies that any node in $\bar{V}_2$ is not in a cycle of G'. Therefore G' has no odd cycle and $P(G')$ is an integer polytope. Since $\Pi(G')$ is a face of $P(G')$, we have a contradiction.

Now let C be an odd cycle of $\bar{G}$. We can define a fractional vector as follows:

$$\begin{aligned}
&\bar{y}(v) = 1/2 \quad \text{for all nodes } v \in V_2 \cap V(C),\\
&\bar{x}(a) = 1/2 \quad \text{for } a \in A(C),\\
&\bar{y}(v) = 1 \quad \text{for all nodes } v \in V_2 \setminus V(C).
\end{aligned}$$

For every node $u \in V_1 \setminus V(C)$, we look for an arc $(u, v) \in \delta^+(u)$. If $\bar{y}(v) = 1$ we set $\bar{x}(u, v) = 1$. If $\bar{y}(v) = 1/2$, then there is another arc $(u, w) \in \delta^+(u)$ such that $\bar{y}(w) = 1/2$ or $\bar{y}(w) = 1$. We set $\bar{x}(u, v) = \bar{x}(u, w) = 1/2$. Finally we set $\bar{x}(a) = 0$ for each remaining arc a. This vector satisfies (23)-(26), but it violates the inequality (22) associated with C. This shows that in this case (23)-(26) does not define an integer polytope. Thus we can state our main results.

Theorem 17. *The system* (23)-(26) *defines an integral polytope if and only if $\bar{G}$ has no odd cycle.*

Theorem 18. *The UFLP is polynomially solvable for graphs G such that $\bar{G}$ has no odd cycle.*

This class of graphs can be recognized in polynomial time as described in Subsection 3.5.

4. The p-median problem

The p-median problem is closely related to the uncapacitated facility location problem. Here we need to select a specific number of centers. Formally, let $G = (V, A)$ be a directed graph, not necessarily connected. We assume that G is simple, i.e., between any two nodes u and v there is at most one arc directed from u to v. Also for each arc $(u, v) \in A$ and node $v \in V$ there is an associated cost $c(u, v)$ and $w(v)$, respectively. The *p-median problem* (pMP) consists 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 of the selected nodes plus 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 [14], placement of web proxies in a computer network [27], semistructured data bases [26, 22].

The following define an integer linear programming formulation for the pMP:

$$\min \sum_{(u,v)\in A} c(u, v)x(u, v) + \sum_{v\in V} d(v)y(v) \tag{27}$$

$$\sum_{v\in V} y(v) = p, \tag{28}$$

$$\sum_{v:(u,v)\in A} x(u, v) + y(u) = 1 \quad \forall u \in V, \tag{29}$$

$$x(u, v) \le y(v) \quad \forall (u, v) \in A, \tag{30}$$

$$0 \le y(v) \le 1 \quad \forall v \in V, \tag{31}$$

$$x(u, v) \ge 0 \quad \forall (u, v) \in A. \tag{32}$$

Denote by $P_p(G)$ the polytope defined by (28)-(32), this give a linear programming relaxation of the pMP. Let $pMP(G)$ be the convex hull of $P_p(G) \cap \{0,1\}^{|A|+|V|}$.

The facets of $pMP(G)$ have been studied in [1] and [16]. In [1], new facets have been presented using a reduction to the stable set problem in the *intersection graph of* G. The intersection graph of G is defined as follows: its nodes are the arcs of G and there is an edge between two nodes (u,v) and (w,t) if $u = w$ or $v = w$. If we associate the cost $c(u,v)$ with each node (u,v) of the intersection graph, then the p-median problem in G, when the cost associated with the nodes of G is zero, is equivalent to find a stable set with minimum weight of cardinality $|V| - p$ in the intersection graph of G. In [16], other class of facets have been presented in the class of bipartite graphs.

In this section we characterize all directed graphs such that $P_p(G) = pMP(G)$. To state our main result we need some definitions.

In Figure 4, we show four directed graphs and for each of them a fractional extreme point of $P_p(G)$. The numbers near the nodes correspond to the variables y, all the arcs variables are equal to $\frac{1}{2}$.

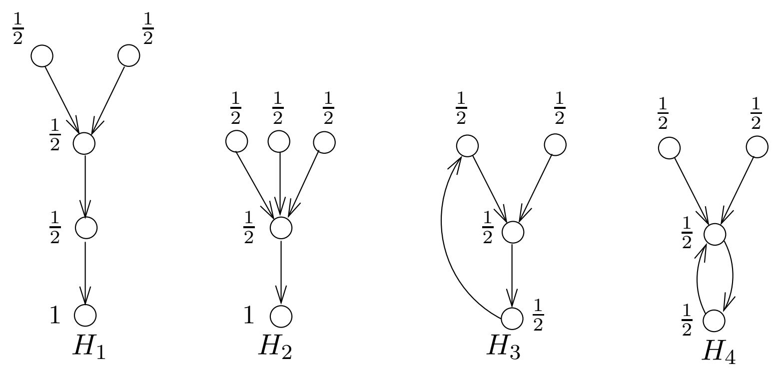


FIGURE 4. Fractional extreme points of $P_p(G)$.

Definition 19. *A simple cycle C is called a Y-cycle if for every $v \in \hat{C}$ there is an arc $(v,\bar{v})$, where $\bar{v}$ is in $V \setminus \dot{C}$.*

In Figure 5 we show a fractional extreme point of $P_p(G)$ different from those given in Figure 4. It consists of an odd Y-cycle with an arc having both of its endnodes outside the cycle. The values reported near each node represent the node variables, the arc variables are all equal to $\frac{1}{2}$. These values form a fractional extreme point of $P_p(G)$, with $p = 4$.

The theorem below is the main result of this section. It shows that the configurations in Figures 4 and 5 are the only configurations that should be forbidden in order to have an integral polytope.

Theorem 20. *Let $G = (V,A)$ be a directed graph, then $P_p(G)$ is integral if and only if*

- *(i) it does not contain as a subgraph any of the graphs H_1, H_2, H_3 or H_4 of Figure 4, and*
- *(ii) it does not contain an odd Y-cycle C and an arc (u,v) with neither u nor v in $V(C)$.*

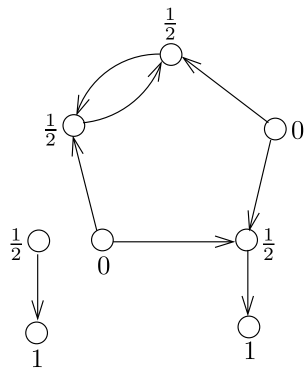


FIGURE 5. An odd Y-cycle with an arc outside the cycle .

The proof of this theorem consists of three parts presented in Subsections 4.2, 4.3 and 4.4. The last two parts are the subject of two papers, see [4, 5], each requires more than twenty pages. For these reasons, here we only present an overview of the proof.

In the first part of the proof, Subsection 4.2, we show that $P_p(G)$ is integral in Y-free graphs with no odd directed cycles. A *Y-free graph* is an oriented graph that does not contain as a subgraph the graph Y of Figure 6. This class of graphs has been introduced in [2].

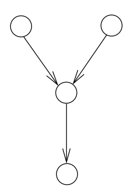


FIGURE 6. The graph Y.

In the second part, Subsection 4.3, we prove Theorem 20 when restricted to oriented graphs. This proof uses an induction on the number of subgraphs Y. The last part is devoted to the proof of Theorem 20 in general directed graphs and uses the result in oriented graphs as starting point. We will only present the sufficiency proof. The necessity proof is illustrated in Figures 4 and 5. The fractional extreme points given in these figures can be easily extended to any graph that does not satisfy Conditions (i) and (ii) of Theorem 20. Thus the graphs we consider do not contain as a subgraph any of the graphs H_1, H_2, H_3 or H_4 of Figure 4.

4.1. **Preliminaries.** Let $G = (V, A)$ be a directed graph. Let $l : V \cup A \rightarrow \{0, -1, 1\}$ be a labeling function that associates to each node and arc of G a label 0, -1 or 1.

A vector $(x, y) \in P_p(G)$ will be denoted by z, i.e., $z(u) = y(u)$ for all $u \in V$ and $z(u, v) = x(u, v)$ for all $(u, v) \in A$. Given a vector z and a labeling function l, we define a new vector z_l from z as follows: $z_l(u) = z(u) + l(u)\epsilon$, for all $u \in V$, and $z_l(u, v) = z(u, v) + l(u, v)\epsilon$, for all $(u, v) \in A$, where ϵ is a sufficiently small positive scalar.

The labeling procedure for even cycles. Let $C = v_0, a_0, v_1, a_1, \ldots, a_{p-1}, v_p$ be an even cycle, not necessarily a Y-cycle.

- If C is a directed cycle, assume that v_0 is the tail of a_0, then set $l(v_0) \leftarrow 1$; $l(a_0) \leftarrow -1$. Otherwise, assume $v_0 \in \dot{C}$ and set $l(v_0) \leftarrow 0$; $l(a_0) \leftarrow 1$.
- Extend the labels as in Subsection 3.3.

Remark 21. *If C is a directed even cycle, then $l(a_{p-1}) = l(v_0)$ and $\sum l(v_i) = 0$.*

This remark is easy to see. The second property is given in the following lemma and it concerns non-directed cycles.

Lemma 22. *If C is a non-directed even cycle, then $l(a_{p-1}) = -l(a_0)$ and $\sum l(v_i) = 0$.*

We are going to deal with a vector z that is a fractional extreme point of $P_p(G)$.

Recall that the graph G we consider in Subsection 4.2 is Y-free and with no odd directed cycles and the graph G in Subsections 4.3 and 4.4 do not contain as a subgraph any of the graphs H_1, H_2, H_3 or H_4 of Figure 4. In these graphs the following two lemmas hold:

Lemma 23. *We may assume that $z(u, v) > 0$ for all $(u, v) \in A$.*

Proof. cf. [4]. Let G' be the graph obtained after removing all arcs (u, v) with $\bar{z}(u, v) = 0$. The graph G' has the same properties as G. Let z' be the restriction of $\bar{z}$ on G'. Then z' is a fractional extreme point of $P_p(G')$. □

Lemma 24. *We may assume that $|\delta^-(v)| \leq 1$ for every pendent node v in G.*

Proof. cf. [4]. If v is a pendent node in G and $\delta^-(v) = \{(u_1, v), \ldots, (u_k, v)\}$, we can split v into k pendent nodes $\{v_1, \ldots, v_k\}$ and replace every arc (u_i, v) with (u_i, v_i). Then we define z' such that $z'(u_i, v_i) = z(u_i, v)$, $z'(v_i) = 1$, for all i, and $z'(u) = z(u)$, $z'(u, w) = z(u, w)$ for every other node and arc. Let G' be this new graph. The graph G' has the same properties as G. Moreover, it is easy to check that z' is a fractional extreme point of $P_{p+k-1}(G')$. □

4.2. **Y-free graphs.** In [2], we characterized the fractional extreme points of $P_p(G)$ for Y-free graphs. Then we showed that by adding the family of odd cycle inequalities associated with each directed odd cycle in G we obtain an integral polytope. An alternate proof of this result based on matching theory is given in [24].

To prove our main result we do not need the description of $pMP(G)$ in Y-free graphs. We need its description in a smaller class described by those Y-free graphs with no odd directed cycle. In this restricted class of graphs $P_p(G)$ is integral, this is a directed consequence of Theorem 14 in [2]. Below we give a proof based on the matching polytope in bipartite graphs, which is along the same lines of the proof given in [24].

Theorem 25. *If $G = (V, A)$ is a Y-free graph with no odd directed cycle, then for any p the polytope $P_p(G)$ is integral.*

Proof. Let $G = (V, A)$ be a Y-free graph with no odd directed cycle. Assume the contrary, and let $z = (x, y)$ be an extreme fractional point of $P_p(G)$.

Using Fourier-Motzkin elimination, we obtain the following system of linear inequalities, that defines the projection of $P_p(G)$ onto the arc variables space; call it $Q_p(G)$.

$$\sum_{(u,v)\in A} x(u,v) = |V| - p, \tag{33}$$

$$x(w,u) + \sum_{v:(u,v)\in A} x(u,v) \leq 1 \quad \forall (w,u) \in A, \tag{34}$$

$$x(u,v) \geq 0 \quad \forall (u,v) \in A. \tag{35}$$

Remark that by Lemma 24 and the fact that G is a Y-free graph, we have that $|\delta^-(v)| \leq 1$ for all $v \in V$. Hence if we omit the orientation of the arcs in G we obtain a undirected graph $I(G) = (V, E)$, and inequalities (34) and (35) are equivalent to

$$x(\delta_{I(G)}(v)) \leq 1 \quad \forall v \in V, \tag{36}$$

$$x(e) \geq 0 \quad \forall e \in E. \tag{37}$$

Combining Lemma 24 and the fact that G does not contain an odd directed cycle, we obtain that $I(G)$ is a bipartite graph and hence the polytope defined by inequalities (36) and (37) is the matching polytope of a bipartite graph, so it is integral. Now by adding the equality $\sum_{e\in E} x(e) = |V| - p$ to the linear system defined by (36) and (37) the resulting polytope still integral, this is a well known property of the matching polytope, see for instance [19]. This proves that $Q_p(G)$ is integral.

To finish the proof of our theorem it suffices to see that if $z = (x, y)$ is an extreme point of $P_p(G)$, then x is an extreme point of $Q_p(G)$, which is easy to verify. □

4.3. **Oriented graphs.** Let $G = (V, A)$ be an oriented graph that satisfies Conditions (i) and (ii) of Theorem 20. First we study the case when G has no odd Y-cycle, and in the second case we assume that G has an odd Y-cycle.

4.3.1. *G does not contain an odd Y-cycle.* Let $t \in V$. The node t is called a *Y-node in* $G = (V, A)$ if there are three different nodes u_1, u_2, w in V such that (u_1, t), (u_2, t) and (t, w) belong to A. Denote by Y_G the set of Y-nodes in G.

The proof is done by induction on the number of Y-nodes. If $|Y_G| = 0$ then, the graph is Y-free with no odd directed cycle, it follows from Theorem 25 in Subsection 4.2 that $P_p(G)$ is integral. Assume that $P_p(G')$ is integral for any positive integer p and for any oriented graph G', with $|Y_{G'}| < |Y_G|$, that satisfies Condition (i) and does not contain an odd Y-cycle. Now we suppose that $z = (x, y)$ is a fractional extreme point of $P_p(G)$ and we plan to obtain a contradiction. The next lemma we need is as follows.

Lemma 26. *G does not contain a cycle.*

Proof. (Sketch), cf. [4]. The proof of this lemma is a direct application of the labeling procedure of Section 4.1. We assume that there is a cycle, using Lemma 24, we can derive an even Y-cycle C, and we assign labels to the nodes and arcs of C following the labeling procedure of Section 4.1. Extend the labels as follows: for each node $v \in \hat{C}$, choose an arc $(v, \bar{v})$, $\bar{v} \notin V(C)$, and assign the label $-l(v)$ to it. Assign a zero label to all remaining nodes and arcs. In the last step, using Lemma 23 we show that any constraint that is satisfied with equality by z is also satisfied with equality by z_l. This contradicts the fact that z is an extreme point of $P_p(G)$. □

The graph G must contain at least one Y-node t with its incident arcs (u_1, t), (u_2, t), (t, w). Using Lemma 26 we can prove that V can be partitioned into W_1 and W_2 so that $\{u_1, t, w\} \subseteq W_1$ and $u_2 \in W_2$, and that the only arc in G between W_1 and W_2 is (u_2, t).

Next we show that $z(t) = \frac{1}{2}$. We have that $Q(G)$, the the polytope defined by (29)-(32), is a face of the polytope $P(G)$ defined by (1)-(4)) studied in Section 3. And by Theorem 13, we know that $P(G)$ is integral when G does not contain an odd cycle, which is the case here. Thus $Q(G)$ is also integral. The polytope $P_p(G)$ is obtained from $Q(G)$ by adding exactly one equation. A simple polyhedral fact is that if $Q(G)$ is integral, then the values of z are in $\{0, 1, \alpha, 1-\alpha\}$, for some number $\alpha \in [0, 1]$. But since $z(t) = \frac{1}{2}$ we have that all fractional values of z are equal to $\frac{1}{2}$.

Define $p_1 = \sum_{v \in W_1} z(v)$ and $p_2 = \sum_{v \in W_2} z(v)$, so $p = p_1 + p_2$. We distinguish two cases: p_1 and p_2 are integer; and they are not.

If the numbers p_1 and p_2 are integer, we define the graphs G^1 and G^2 as follows. Let $A(W_1)$ and $A(W_2)$ be the set of arcs in G having both endnodes in W_1 and W_2, respectively. Let $G^1 = (W_1, A(W_1))$ and $G^2 = (W_2 \cup \{t', v', w'\}, A(W_2) \cup \{(u_2, t'), (t', v'), (v', w')\})$.

Let z_1 be the restriction of z to G^1. Clearly $z_1 \in P_{p_1}(G^1)$. Define z_2 as follows, $z_2(u_2, t') = z(u_2, t) = \frac{1}{2}$, $z_2(t') = \frac{1}{2}$, $z_2(t', v') = \frac{1}{2}$, $z_2(v') = \frac{1}{2}$, $z_2(v', w') = \frac{1}{2}$, $z_2(w') = 1$ and $z_2(u) = z(u)$, $z_2(u, v) = z(u, v)$ for all other nodes and arcs of G^2. We have that $z_2 \in P_{p_2+2}(G^2)$.

Both graphs G^1 and G^2 satisfy Condition (i) of Theorem 20 and do not contain an odd Y-cycle. Moreover, $|Y_{G^1}| < |Y_G|$ and $|Y_{G^2}| < |Y_G|$. Since z_1 and z_2 are both fractional, the induction hypothesis implies that they are not extreme points of $P_{p_1}(G^1)$ and $P_{p_2+2}(G^2)$, respectively. Thus there must exist a 0-1 vector $z'_1 \in P_{p_1}(G^1)$ with $z'_1(t) = 0$ so that the same constraints that are tight for z_1 are also tight for z'_1. Also there must exist a 0-1 vector $z'_2 \in P_{p_2+2}(G^2)$ with $z'_2(t') = 0$ such that the same constraints that are tight for z_2 are also tight for z'_2. Now by combining z'_1 and z'_2 one can define a solution $z' \in P_p(G)$ that satisfies as equality each constraint that is satisfied as equality by z.

In the case where the numbers p_1 and p_2 are not integer, we use the same idea as above but applied for new graphs G^1 and G^2, where $G^1 = \big(W_1 \cup \{u'_1\}, (A(W_1) \setminus \{(u_1, t)\}) \cup \{(u_1, u'_1), (u'_1, t)\}\big)$ and $G^2 = \big(W_2 \cup \{t', w'\}, A(W_2) \cup \{(u_2, t'), (t', w')\}\big)$.

Notice that $\sum_{v \in W_1} z(v) = p_1 = \alpha + \frac{1}{2}$ and $\sum_{v \in W_2} z(v) = p_2 = \beta - \frac{1}{2}$, where α and β are integers and $\alpha + \beta = p$. Thus the number of nodes to be selected in G^1 (resp. G^2) is $\alpha + 1$ (resp. $\beta + 1$). This concludes the proof.

4.3.2. *G contains an odd Y-cycle.* We assume that G satisfies conditions (i) and (ii) of Theorem 20 and contains an odd Y-cycle. We also assume that z is a fractional extreme point of $P_p(G)$. The first lemma we need is the following.

Lemma 27. *The graph G contains exactly one odd Y-cycle.*

Let C be the unique odd Y-cycle in G. We showed in [4] that in this case G has the following special structure. The node set V is partitioned into three subsets, $V(C)$, V' and V''. Each node in V' has exactly one arc incident to it, this arc is directed into a node in $\dot{C}$. Also, each node in V'' has exactly one arc incident to it, this arc is directed away from a node in $V(C)$. Each node in $V(C)$ is adjacent to at most one node in $V' \cup V''$. We denote by A' (resp. A'') the set of arcs incident to the nodes in V' (resp. V''). These arcs together with $A(C)$ define the arc-set of G.

Lemma 28. *We may assume that $z(u,v) = z(v)$ for each arc in $A(C)$.*

Proof. cf. [4]. If we have an arc $(u,v) \in A(C)$ with $z(u,v) < z(v)$, then we remove this arc and add a new arc (u,v') and we assign the value $z(u,v)$ to the arc (u,v') and 1 to the node v'. From Lemma 27, this new graph does not contain an odd Y-cycle and the associated solution is fractional extreme point. But from subsection 4.3.1, this is impossible. □

Next we concentrate on $Q(G)$, the polytope defined by (29)-(32). Notice that since $P_p(G)$ is obtained from $Q(G)$ by adding one equation, then an extreme point of $P_p(G)$ is either an extreme point of $Q(G)$ or a convex combination of two extreme points of $Q(G)$. We omit the proof of the following lemma.

Lemma 29. *If z is a fractional extreme point of $Q(G)$ with $z(u,v) = z(v)$ for each $(u,v) \in A(C)$, then $z(u,v) = \frac{1}{2}$ for each arc $(u,v) \in A(C)$, $z(v) = \frac{1}{2}$ for each node $v \in \hat{C} \cup \tilde{C}$ and $z(v) = 0$ for each node $v \in \dot{C}$.*

From this lemma and the definition of G we obtain the following corollary.

Corollary 30. *z cannot be an extreme point of $Q(G)$.*

Proof. cf. [4]. Assume that z is an extreme point of $Q(G)$. By definition we have $z(v) = 1$ for each node $v \in V''$. From Lemma 29, $z(v) = 0$ if $v \in \dot{C}$, so by the definition of V' we have $z(v) = 1$ for each $v \in V'$. Again from Lemma 29 we have $z(v) = \frac{1}{2}$ if $v \in \tilde{C} \cup \hat{C}$. Hence $\sum_{v \in V} z(v) = |V'| + |V''| + \frac{|\tilde{C}|+|\hat{C}|}{2}$ but $|\tilde{C}| + |\hat{C}|$ is odd, so $\sum_{v \in V} z(v)$ is not an integer, a contradiction. □

The corollary above implies that the extreme point z of $P_p(G)$ is a convex combination of two extreme points of $Q(G)$, they are $\tilde{z}$ and $\hat{z}$. Thus $z = \alpha\tilde{z} + (1-\alpha)\hat{z}$, with $0 < \alpha < 1$.

Denote by A''_1 the arcs in A'' that are incident to a node in $\dot{C} \cup \tilde{C}$. Also let $\dot{C}^+$ be the set of nodes $v \in \dot{C}$ with $z(v) > 0$. Using the structure of the graph and the fact that z is a fractional extreme point of $P_p(G)$, the proof reduces to the following four cases: (1) $A' = \{(u,v)\}$, $A''_1 = \emptyset$, $\dot{C}^+ = \{v\}$; (2) $A''_1 = \{(u,v)\}$, $A' = \emptyset$, $\dot{C}^+ = \emptyset$; (3) $\dot{C}^+ = \{v\}$, $A' \cup A''_1 = \emptyset$; (4) $\dot{C}^+ = \emptyset$, $A' \cup A''_1 = \emptyset$.

In each of theses cases, we show that z cannot be an extreme point of $P_p(G)$ if both $\tilde{z}$ and $\hat{z}$ are fractional or both are integral. It remains the case when one is integral and the other is fractional. Notice that if $\tilde{z}$ or $\hat{z}$ is fractional, then it satisfies the conditions of Lemma 29, this follows from Lemma 28 where $z(u,v) = z(v)$ if $(u,v) \in A(C)$. Using this together with Lemma 29 the contradiction we obtain is that $\sum_{v \in V} z(v) = q + \frac{\alpha}{2}$, where q is an integer.

4.4. **General directed graphs.** Let us redefine a Y-cycle in this context, that is in the graphs that do not contain any of the graphs of Figure 4 as a subgraph. With this we can distinguish the nodes in $\hat{C}$ that do not satisfy Definition 19, which is useful in the proof of Lemma 33.

Definition 31. *A simple cycle C is called a Y-cycle if for every $v \in \hat{C}$ at least one of the following hold:*

- *(i) there exists an arc $(v, \bar{v}) \notin A(C)$, $\bar{v} \notin V(C)$, or*

- *(ii) there exists an arc* $(v, \bar{v}) \notin A(C)$, $\bar{v} \in \tilde{C}$ *and* $\bar{v}$ *is one of the two neighbors of* v *in* C.

For a simple cycle C, denote by $\hat{C}_{(i)}$ the set of nodes in $\hat{C}$ that satisfy Condition (i) of the above definition. Notice that we may have nodes in $\hat{C}$ that satisfy both (i) and (ii).

We study two cases as follows.

4.4.1. *G does not contain an odd Y-cycle.* We assume that $G = (V, A)$ is a directed graph that does not contain any of the graphs H_1, H_2, H_3 or H_4 of Figure 4 as a subgraph. Also we assume that G does not contain an odd Y-cycle.

Assume that z is a fractional extreme point of $P_p(G)$. The first step in the proof is to show the following lemma, that is an analogue of Lemma 26. Its proof is harder than the proof of Lemma 26 and requires new definitions and notions. This proof illustrates one of the main differences between the oriented and the directed case.

Definition 32. *Let C be a Y-cycle in a directed graph $G = (V, A)$. A node $v \in V(C)$ is called a blocking node if one of the following hold:*

(i) $v \in \tilde{C}$, $(v, u) \in A(C)$, $(u, v) \in A \setminus A(C)$ *and* $u \in \tilde{C}$, *or*
(ii) $v \in \hat{C}$, $(u, v) \in A(C)$, $(w, v) \in A(C)$, $(v, u) \in A \setminus A(C)$, $(v, w) \in A \setminus A(C)$ *and both* u *and* w *are in* $\tilde{C}$.

Lemma 33. *If $z(u, v) = z(v)$, for all (u, v) with v not a pendent node, then G does not contain a cycle of size at least three.*

Proof. cf. [5]. Assume the contrary. Suppose that G admits such a cycle. The first step is to derive an even Y-cycle. Let $C' = v_0, a_0, v_1, a_1, \ldots, a_{p-1}, v_p$, be a simple cycle with $p \geq 3$. Suppose that C' is not a Y-cycle. Then we can show that there is a node $v_i \in \hat{C}'$ with $\delta^+(v_i) = \{(v_i, v_{i-1}), (v_i, v_{i+1})\}$, where v_{i-1} and v_{i+1} are the two neighbors of v_i in C' and they belong to $\dot{C}$. Now it suffices to define C from C', recursively, following the procedure below:

Step 1. $A(C) \leftarrow A(C')$, $V(C) \leftarrow V(C')$, $C \leftarrow C'$.
Step 2. If there exist $v_i \in \hat{C}$, a node not satisfying Definition 31 (i) and (ii), go to Step 3. Otherwise stop, C is a Y-cycle.
Step 3. $A(C) \leftarrow (A(C) \setminus \{(v_{i-1}, v_i), (v_{i+1}, v_i)\}) \cup \{(v_i, v_{i-1}), (v_i, v_{i+1})\}$. C is the new cycle defined by $A(C)$. Go to Step 2.

Each Step 3 decreases by one the number of nodes in $\hat{C}$. Thus the procedure must end with a Y-cycle C.

The next step is to apply the labeling procedure to an even Y-cycle. Call this labeling l. We extend l to all other nodes and arcs in order to get a solution z_l that satisfies as equality each constraint satisfied as equality by z. The extension of l is possible only when the Y-cycle does not contain a blocking node. We show that if we choose a Y-cycle C with $|C_{(i)}|$ maximum, then this cycle does not contain a blocking node. This completes the proof of this lemma. □

The lemma above is used to prove the following:

Lemma 34. *We cannot have $z(u, v) = z(v)$, for all (u, v) with v not a pendent node.*

Proof. cf. [5]. Denote by $Pair(G)$ the set of pair of nodes $\{u, v\}$ such that both arcs (u, v) and (v, u) belong to A.

The proof is by induction on $|Pair(G)|$. If $|Pair(G)| = 0$ then G is an oriented graph that satisfies Conditions (i) and (ii) of Theorem 20. Thus from subsection 4.3.1, $P_p(G)$ has no fractional extreme point so the lemma is true. Suppose now that $|Pair(G)| = m + 1$, for $m \geq 0$.

Suppose that $z(u, v) = z(v)$ for each arc (u, v) with v not a pendent node. Notice that Lemma 33 applies, so G does not contain a cycle. Let (u, v) and (v, u) be two arcs in G. Denote by $G(u, v)$ the graph obtained from G by removing the arc (u, v) and adding a new arc (u, t), where t is a new pendent node. Define $\tilde{z} \in P_{\tilde{p}}(G(u, v))$, $\tilde{p} = p + 1$, to be $\tilde{z}(u, t) = z(u, v)$, $\tilde{z}(t) = 1$ and $\tilde{z}(r) = z(r)$, $\tilde{z}(r, s) = \bar{z}(r, s)$ for every other node and arc.

The graph $G(u, v)$ is directed with no multiple arcs and satisfies Condition (i) of Theorem 20. Since G does not contain a cycle, we have that $G(u, v)$ has no odd Y-cycle. Moreover $|Pair(G(u, v))| \leq m$, hence the induction hypothesis applies for $G(u, v)$. We have that $\tilde{z}$ is a fractional vector in $P_{\tilde{p}}(G(u, v))$ with $\tilde{z}(u, v) = \tilde{z}(v)$ for each arc (u, v), with v not pendent. By the induction hypothesis $\tilde{z}$ is not an extreme point. Thus, there exists a set of extreme points of $P_{\tilde{p}}(G(u, v))$, $z^1, \ldots, z^k$, where each constraint that is tight for $\tilde{z}$ is also tight for each of $z^1, \ldots, z^k$, and $\tilde{z}$ is a convex combination of $z^1, \ldots, z^k$. We can show that all these extreme points are in 0-1.

Let z^1, with $z^1(v, u) = 1$. Define $z'' \in P_p(G)$ as follows: $z''(u, v) = z^1(u, t)$ and $z''(r, s) = z^1(r, s)$, $z''(r) = z^1(r)$, for all other nodes and arcs. All constraints that are tight for z are also tight for z''. To see this, it suffices to remark that $z''(v) = z^1(v) = 0$ and $z''(u, v) = z^1(u, t) = 0$. This contradicts the fact that z is an extreme point of $P_p(G)$. □

Let v a node in G. We call v a *knot* if $\delta^-(v) = \{(u, v), (w, v)\}$, $u \neq w$ and both (v, u) and (v, w) belong to $\delta^+(v)$.

Suppose that G does not contain a knot. From Lemma 34 we may assume that there is an arc (u, v) with $z(u, v) < z(v)$ and v is not a pendent node. We may assume that G' the graph obtained from G by removing (u, v) and adding a new pendent node v' and the arc (u, v') contains an odd Y-cycle C. Otherwise, instead of considering G with z, we consider G' with z', $z'(u, v') = z(u, v)$, $z'(v') = 1$, and $z'(s, t) = z(s, t)$, $z'(r) = z(r)$ for all other arcs and nodes.

Since G contains no knot, this implies that $\delta^+_G(u) = \{(u, v)\}$ and $\delta^-_G(u) = \{(s, u), (v, u)\}$, where s and v are the nodes that are adjacent to u in C. Remark that v must be in $\dot{C}$, otherwise C is also an odd Y-cycle in G, which is not possible.

We have that $\delta^-(v) = \{(u, v)\}$. In fact, since $v \in \dot{C}$ we must have an arc (v, w) in C. Because G has no knot this implies that the arc (w, v) cannot exist. So suppose (w', v) is an arc of G with $w' \neq w$, $w' \neq u$. Since w is in C, it is not a pendent node and hence G does not satisfies Condition (i) of Theorem 20.

We must have $z(v, u) = z(u)$. Otherwise, we can construct the pair G' and z' as above. But in this case, one can check that equation (29) with respect to v is violated.

This permit us to apply an induction on the number of knots in G to finish the proof.

4.4.2. *G contains an odd Y-cycle.* Let C be an odd Y-cycle in G. Assume that z is a fractional extreme point of $P_p(G)$.

Lemma 35. *The node set of any cycle of size at least three in G coincides with $V(C)$.*

Lemma 36. *Let $G = (V, A)$ be a directed graph and (u, v) and (v, u) two arcs in A. If $P_p(G)$ admits a fractional extreme point $\bar{z}$ with $\bar{z}(v, u) > 0$, then $P_{\tilde{p}}(G(u, v)) \neq \tilde{p}MP(G(u, v))$, where $\tilde{p} = p + 1$. The graph $G(u, v)$ was defined in the last subsection.*

Proof. cf. [5]. Let $\bar{z}$ be a fractional extreme point of $P_p(G)$ with $\bar{z}(v, u) > 0$. Suppose that $P_{\tilde{p}}(G(u, v)) = \tilde{p}MP(G(u, v))$. Define $\tilde{z} \in P_{\tilde{p}}(G(u, v))$ to be $\tilde{z}(u, t) = \bar{z}(u, v)$, $\tilde{z}(t) = 1$ and $\tilde{z}(r) = \bar{z}(r)$, $\tilde{z}(r, s) = \bar{z}(r, s)$ for all other nodes and arcs. The solution $\tilde{z}$ is fractional, so $\tilde{z}$ is not an extreme point of $P_{\tilde{p}}(G(u, v))$. Since $P_{\tilde{p}}(G(u, v))$ is integral, there is a 0-1 vector $z^* \in P_{\tilde{p}}(G(u, v))$ with $z^*(v, u) = 1$, so that the same constraints that are tight for $\tilde{z}$ are also tight for z^*. From z^* define $z'' \in P_p(G)$ as follows: $z''(u, v) = z^*(u, t)$ and $z''(r) = z^*(r)$, $z''(r, s) = z^*(r, s)$, for all other nodes and arcs. All constraints that are tight for $\bar{z}$ are also tight for z''. To see this, it suffices to remark that $z''(v) = z^*(v) = 0$ and $z''(u, v) = z^*(u, t) = 0$. This contradicts the fact that $\bar{z}$ is an extreme point of $P_p(G)$. □

The proof of Theorem 20 in this case is by induction on $|Pair(G)|$, the number of pairs of nodes $\{u, v\}$ with both (u, v) and (v, u) in A. If $|Pair(G)| = 0$ then G is an oriented graph that satisfies Conditions (i) and (ii) of Theorem 20. Hence the result follows from Subsection 4.3.

Let (u, v) and (v, u) be two arcs in A. Lemma 23 implies $z(v, u) > 0$, so Lemma 36 applies and implies that

$$P_{\tilde{p}}(G(u, v)) \neq \tilde{p}MP(G(u, v)). \tag{38}$$

Using Lemma 35 and the definition of G, we can see that $G(u, v)$ satisfies Conditions (i) and (ii) of Theorem 20. Since $|Pair(G(u, v))| < |Pair(G)|$, we can apply the induction hypothesis so $P_{\tilde{p}}(G(u, v)) = \tilde{p}MP(G(u, v))$. This contradicts (38). Thus the proof of our main result is complete.

5. Recognizing the graphs defined in Theorem 20

In this section we show how to decide if a graph satisfies conditions (i) and (ii) of Theorem 20. Clearly Condition (i) can be tested in polynomial time. Thus we assume that we have a graph satisfying Condition (i), then we split all pendent nodes as in Lemma 24, then we pick an arc (u, v), we remove u and v, and look for an odd Y-cycle in the new graph. We repeat this for every arc. It remains to show how to find an odd Y-cycle.

In Subsection 3.5 we gave a procedure that finds an odd cycle if there is any. We remind the reader that a cycle C is odd if $|V(C)| + |\hat{C}|$ is odd. Since an odd cycle is not necessarily a Y-cycle, we are going to modify the graph so that an odd cycle in the new graph gives an odd Y-cycle in the original graph. The main difficulty resides in how to deal with nodes that satisfy condition (ii) of Definition 31. Such a node should appear in a pair $\{(u, v), (v, u)\}$. Instead of working with such a pair we are going to work with a maximal bidirected path $P = v_1, \ldots, v_q$, this is a path where the arcs (v_{i+1}, v_i) and (v_i, v_{i+1}) belong to G, for $i = 1, \ldots, q - 1$. Notice that if the graph contains a bidirected cycle (when $v_1 = v_q$), then it is easy to derive an odd Y-cycle. So in what follows we assume that there is no bidirected cycle. The transformation is based on the following two remarks.

Remark 37. *There is at most one arc (u, v_1), $u \notin P$, and at most one arc (v, v_q), $v \notin P$. Otherwise the graph H_4 is present.*

Remark 38. *If the arc (u, v_1) is in A, $u \notin P$, and there is an arc (v_1, w) also in A, $w \notin P$, then w is a pendent node. Otherwise we obtain one of the graphs in Figure 4.*

Let C be a Y-cycle that goes through P. We have three cases to study.

Case 1. $\delta^-(P) = \{(u, v_1), (v, v_q)\}$. In this case C contains all nodes in P and also the arcs (u, v_1) and (v, v_q). Since C contains all nodes from P, the only variable that can change the parity of C is the parity of $|\hat{C} \cap P|$.

Notice that if $q \geq 5$ and if there is a Y-cycle going through P then we can always change the parity of it if needed. In fact, we can always join the nodes v_1 and v_q using arcs of P in such a way that $|\hat{C} \cap P| = 1$ as shown in Figure 7 (a), or $|\hat{C} \cap P| = 2$ as shown in Figure 7 (b). It follows that if there is a cycle C' going through P then there is a cycle C of the same parity, whose nodes in $|\hat{C} \cap P|$ satisfy Definition 31 (ii).

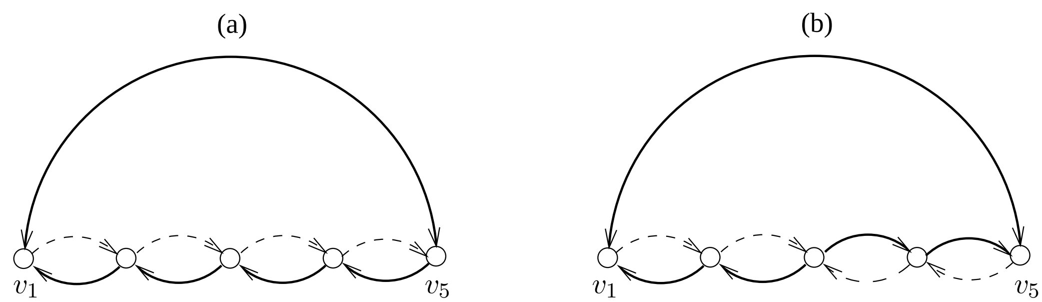


FIGURE 7. Case 1, $q \geq 5$. In bold the Y-cycle C. In dashed line the other arcs of P.

It remains to analyze the cases when $q \leq 4$. The only cases when a transformation is required, are the following two:

- $q = 4$ and neither v_1 nor v_4 is adjacent to a pendent node. In this case we should have $|\hat{C} \cap P| = 1$. To impose that when looking for an odd cycle, we replace P by a bidirected path with two nodes. See Figure 8.

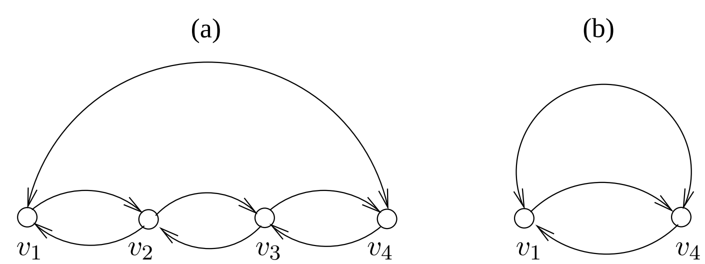


FIGURE 8. Case 1, $q = 4$. (a): before transformation. (b): after transformation.

Let P' the new bidirected path. Any cycle C' with $|\hat{C}' \cap P'| = 1$ can be extended to a cycle C with $|\hat{C} \cap P| = 1$ and where the node in $\hat{C} \cap P$ satisfies Definition 31 (ii).

- $q = 3$ and at most one of v_1 or v_3 is adjacent to a pendent node. Also here we have $|\hat{C} \cap P| = 1$. To impose that when looking for an odd cycle, we remove the arc (v_2, v_3).

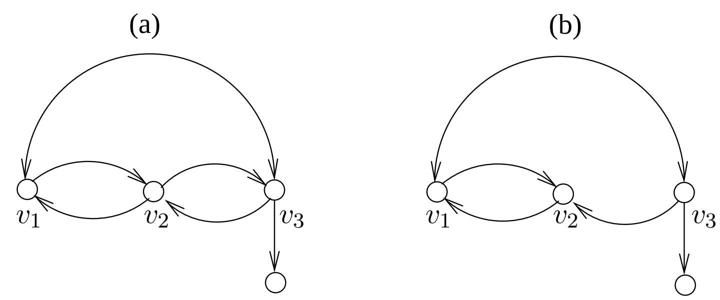


FIGURE 9. Case 1, $q = 3$. (a): before transformation. (b): after transformation.

In Figure 9, we supposed that v_3 is adjacent to a pendent node and v_1 is not.

The two remaining cases below follow the same philosophy as above.

Case 2. $\delta^-(P) = \{(u, v_1)\}$. In this case C contains (u, v_1), all the nodes in P and one arc (v_q, v), $v \notin P$. Here we have two cases to analyze.

- $q \geq 3$ or $q = 2$ and v_1 is adjacent to a pendent node. If $|\hat{C} \cap P|$ is even, we can assume that $|\hat{C} \cap P| = 0$. If $|\hat{C} \cap P|$ is odd, we can assume that $|\hat{C} \cap P| = 1$. Here no transformation is needed.
- $q = 2$ and v_1 is not adjacent to a pendent node. Here we should have $|\hat{C} \cap P| = 0$. To impose that when looking for an odd cycle, we remove (v_2, v_1).

Case 3. $\delta^-(P) = \emptyset$. In this case C contains an arc (v_1, u), $u \notin P$, all nodes in P, and an arc (v_q, v), $v \notin P$. Again we have two cases to analyze.

- $q \neq 3$ or $q = 3$ and v_2 is adjacent to a pendent node. If $|\hat{C} \cap P|$ is even, we can assume that $|\hat{C} \cap P| = 0$. If $|\hat{C} \cap P|$ is odd, we can assume that $|\hat{C} \cap P| = 1$. Here no transformation is needed.
- $q = 3$ and v_2 is not adjacent to a pendent node. Here we should have $|\hat{C} \cap P| = 0$. To impose that when looking for an odd cycle, we remove (v_1, v_2) and (v_3, v_2).

After preprocessing the graph as in Cases 1, 2, and 3, we look for an odd cycle; if there is one, it gives an odd Y-cycle in the original graph.

6. Related polyhedra

In many applications, the underlying graph associated with the p-median problem may be a bipartite (oriented) or an undirected graph. Also, when the cost function is positive, then the problem always reduces to an instance in a bipartite graph. In this section we will consider the polyhedra associated with these two special cases. We will show how Theorem 20 is used to study the integrality of the linear relaxation of the p-median problem in bipartite and undirected graphs.

6.1. **Bipartite graphs.** This is the standard case when V is partitioned into V_1 and V_2 and $A \subseteq V_1 \times V_2$. The customers are the nodes in V_1 and the potential locations are the nodes in V_2. Here we deal with the system

$$\sum_{v \in V_2} y(v) = p, \tag{39}$$

$$\sum_{(u,v) \in A} x(u,v) = 1 \quad \forall u \in V_1, \tag{40}$$

$$x(u,v) \le y(v) \quad \forall (u,v) \in A, \tag{41}$$

$$y(v) \ge 0 \quad \forall v \in V_2, \tag{42}$$

$$y(v) \le 1 \quad \forall v \in V_2, \tag{43}$$

$$x(u,v) \ge 0 \quad \forall (u,v) \in A. \tag{44}$$

Let $\Pi_p(G)$ be the polytope defined by (39)-(44), in this section we characterize the bipartite graphs for which $\Pi_p(G)$ is an integral polytope.

Let us recall some notations introduced in Subsection 3.6 when dealing with bipartite graphs. Let $\bar{V}_1$ be the set of nodes $u \in V_1$ with $|\delta^+(u)| = 1$. Let $\bar{V}_2$ be the set of nodes in V_2 that are adjacent to a node in $\bar{V}_1$. It is clear that the variables associated with nodes in $\bar{V}_2$ should be fixed, i.e., $y(v) = 1$ for all $v \in \bar{V}_2$. Let $\bar{G}$ be the graph induced by $V \setminus \bar{V}_2$.

Let H be a graph with node set $\{u_1, u_2, u_3, v_1, v_2, v_3, v_4\}$ and arc set

$$\{(u_1, v_1), (u_2, v_2), (u_3, v_3), (u_1, v_4), (u_2, v_4), (u_3, v_4)\}.$$

If the graph $\bar{G}$ contains H as a subgraph then we can construct a fractional extreme point as follows: Assign the value $\frac{1}{2}$ to each arc in H and to each node v_i, $i = 1, \ldots, 4$, set to zero all other node and arc variable of G. If $\bar{G}$ contains an odd cycle and one extra node in $V_2 \setminus \bar{V}_2$, we can also construct a fractional extreme point. Now we prove that these are only configurations that should be forbidden in order to have an integral polytope.

Theorem 39. *The polytope $\Pi_p(G)$ is integral if and only if*

- *(i) $\bar{G}$ does not contain the graph H as a subgraph, and*
- *(ii) $\bar{G}$ does not contain an odd cycle C and one extra node in $V_2 \setminus \bar{V}_2$.*

So let G be a graph such that $\bar{G}$ does not contain these two configurations. We assume that z is a fractional extreme point of $\Pi_p(G)$. As in Subsection 4.1, we can assume that $z(u,v) > 0$ for every arc $(u,v) \in A$.

Lemma 40. *We can assume that $z(u,v) = z(v)$ for each arc (u,v) such that $v \in V_2 \setminus \bar{V}_2$.*

Proof. cf. [5]. Suppose that $z(u,v) < z(v)$ for an arc (u,v) and $v \in V_2 \setminus \bar{V}_2$. We can add the nodes u', v', the arcs (u', v'), (u, v') and remove the arc (u,v). Then define $z'(u', v') = z(v') = 1$, $z'(u, v') = z(u,v)$, and $z'(s,t) = z(s,t)$, $z'(w) = z(w)$, for all other nodes and arcs. Let G' be the new graph. Then z' is an extreme point of $\Pi_{p+1}(G')$. The graph G' satisfies the hypothesis of Theorem 39. $\square$

The proof of Theorem 39 is divided into the following three cases:

(1) $\bar{G}$ does not contain an odd cycle nor the graph H.
(2) $\bar{G}$ does not contain H and contains an odd cycle C that includes all nodes in $V_2 \setminus \bar{V}_2$, and $|V_2 \setminus \bar{V}_2| \ge 5$.

(3) $\bar{G}$ does not contain H and contains an odd cycle C that includes all nodes in $V_2 \setminus \bar{V}_2$, and $|V_2 \setminus \bar{V}_2| = 3$.

We treat these three cases below.

6.1.1. *$\bar{G}$ does not contain an odd cycle nor the graph H.*

Lemma 41. *For all $u \in V_1$ we have $|\delta^+(u)| \leq 2$.*

Proof. cf. [5]. Since $\bar{G}$ has no odd cycle, the polytope defined by (40)-(44) is integral, this is Theorem 17 in Subsection 3.6. Thus z is a convex combination of two integral vectors satisfying (40)-(44). Therefore $|\delta^+(u)| \leq 2$. □

Now we build an auxiliary undirected graph G' whose node-set is $V_2 \setminus \bar{V}_2$. For each node $u \in V_1$ such that $\delta^+(u) = \{(u,s),(u,t)\}$, $\{s,t\} \subseteq V_2 \setminus \bar{V}_2$, we have an edge in G' between s and t. This could create parallel edges. Notice that any node v in G' is adjacent to at most two other nodes. If v was adjacent to three other nodes, we would have the sub-graph H in $\bar{G}$.

Lemma 40 implies that if $z(v) = 1$ for $v \in V_2 \setminus \bar{V}_2$, then v is not adjacent to any other node in G'. A node $v \in V_2 \setminus \bar{V}_2$ is called fractional if $0 < z(v) < 1$. So G' consists of a set of isolated nodes, and a set of cycles and paths. We have to study the four cases below.

- If G' contains a cycle, it should be even, because $\bar{G}$ has no odd cycle. For a cycle in G' we can label the nodes with $+1$ and -1 so that adjacent nodes in the cycle have opposite labels. This labeling translates into a labeling in G as follows: If s and t have the labels $+1$ and -1 respectively, and the arcs (u,s) and (u,t) are in G, then (u,s) receives the label $+1$ and (u,t) receives the label -1. If s has the label $l(s)$ and the arcs (u,s) and (u,t) are in G with $t \in \bar{V}_2$, then (u,s) receives the label $l(s)$ and (u,t) receives the label $-l(s)$. All other nodes and arcs receive the label 0. This defines a new vector that satisfies with equality the same constraints that z satisfies with equality.
- If there is a path with an even number of fractional nodes we label them as before. This translates into a labeling in G as follows. If s and t have the labels $+1$ and -1 respectively, and the arcs (u,s) and (u,t) are in G, then (u,s) receives the label $+1$ and (u,t) receives the label -1. If s has the label $l(s)$ and the arcs (u,s) and (u,t) are in G with $t \in \bar{V}_2$, then (u,s) receives the label $l(s)$ and (u,t) receives the label $-l(s)$. All other nodes and arcs receive the label 0. This defines a new vector that satisfies with equality the same constraints that z satisfies with equality.
- If G' has two paths with an odd number of fractional nodes then again we can label the fractional nodes in these two paths and proceed as before.
- It remains the case where G' contains just one path with an odd number of fractional nodes. Let $v_1, \ldots, v_{2q+1}$ be the ordered sequence of nodes in this path. We should have $z(v_i) = \alpha$ if i is odd, and $z(v_i) = 1 - \alpha$ if i is even, with $0 < \alpha < 1$. This implies $\sum_{v \in V_2} z(v) = r + \alpha$ where r is an integer. We have then a contradiction.

6.1.2. *$\bar{G}$ does not contain H and contains an odd cycle C that includes all nodes in $V_2 \setminus \bar{V}_2$, and $|V_2 \setminus \bar{V}_2| \geq 5$.* Here we use several transformations to obtain a new graph $\tilde{G}$ that satisfies conditions (i) and (ii) of Theorem 20, and we use the fact that $P_p(\tilde{G})$ is an integral polytope.

Lemma 42. *Let $u, v \in V(C)$, then there is no arc $(u, v) \notin A(C)$.*

Proof. cf. [5]. If such an arc exists, then the graph H would be present. □

Lemma 43. *A node $u \in (V_1 \setminus \bar{V}_1)$ cannot be adjacent to more than one node in $\bar{V}_2$.*

Proof. cf. [5]. Suppose that the arcs (u, v_1) and (u, v_2) exist with v_1 and v_2 in $\bar{V}_2$. We can add and subtract ϵ to $z(u, v_1)$ and $z(u, v_2)$ to obtain a new vector that satisfies with equality the same constraints that z does. □

Lemma 44. *We can assume that $(V_1 \setminus \bar{V}_1) \setminus V(C) = \emptyset$*

Proof. cf. [5]. Consider a node $u \in (V_1 \setminus \bar{V}_1) \setminus V(C)$ and suppose that the arcs (u, v_1) and (u, v_2) exist, with $v_1, v_2 \in V(C)$. If both paths in C between v_1 and v_2 contain another node in V_2, then there is an odd cycle in $\bar{G}$ and an extra node in $V_2 \setminus \bar{V}_2$. Then we can assume that there is a node $w \in V(C)$ and $(w, v_1), (w, v_2) \in A(C)$. If there is another node $v_3 \in V(C)$ such that the arc (u, v_3) exists, then the graph H is present, this is because $|V_2 \setminus \bar{V}_2| \geq 5$. Thus u cannot be adjacent to any other node in $V(C)$. Lemma 40 implies

$$z(u, v_1) = z(w, v_1), \tag{45}$$

$$z(u, v_2) = z(w, v_2). \tag{46}$$

Then we remove the node u and study the vector z' that is the restriction of z to $G \setminus u$. If there is another vector z'' that satisfies with equality the same constraints that z' does, we can extend z'' using equations (45) and (46), to obtain a vector that satisfies with equality the same constraints that z does.

If there is a node $u \in (V_1 \setminus \bar{V}_1) \setminus V(C)$ that is adjacent to exactly one node $v \in V(C)$, then u is adjacent also to a node $w \in \bar{V}_2$. It follows from Lemma 43 that the node in $\bar{V}_2$ is unique. Lemma 40 implies

$$z(u, v) = z(v), \tag{47}$$

and we also have

$$z(u, v) + z(u, w) = 1. \tag{48}$$

Then we remove the node u and study the vector z' that is the restriction of z to $G \setminus u$. If there is another vector z'' that satisfies with equality the same constraints that z' does, we can extend z'' using equations (47) and (48), to obtain a vector that satisfies with equality the same constraints that z does.

The resulting graph does not contain H and contains the odd cycle C. □

Now consider a node $u \in \bar{V}_1$ that is adjacent to $v \in \bar{V}_2$. We should have $z(u, v) = 1$ and $z(v) = 1$. We remove u from the graph and keep v with $z(v) = 1$.

Finally we add slack variables to the inequalities (43) for each node in $V_2 \setminus \bar{V}_2$. For that we add a node v' and the arc (v, v'), for each node $v \in V_2 \setminus \bar{V}_2$. Then we add the constraints

$$\begin{aligned}
&z(v) + z(v, v') = 1, \\
&z(v, v') \leq z(v'), \\
&z(v') = 1, \\
&z(v, v') \geq 0.
\end{aligned}$$

Let $\tilde{G}$ be this new graph, and $\tilde{p} = p + |V_2 \setminus \bar{V}_2|$. It follows from Lemmas 42, 43 and 44 that $\tilde{G}$ is a graph satisfying conditions (i) and (ii) of Theorem 20. Here we have a face of $P_{\tilde{p}}(\tilde{G})$; because $z(v) = 0$ for all $v \in V_1$. Since $P_{\tilde{p}}(\tilde{G})$ is an integral polytope, there is an integral vector $\tilde{z}$ that satisfies with equality the same constraints that z does. From $\tilde{z} \in P_{\tilde{p}}(\tilde{G})$ one can easily derive $\tilde{z}' \in P_p(G)$ that satisfies with equality the same constrains that $z \in P_p(G)$ satisfies with equality.

6.1.3. *$\bar{G}$ does not contain H and contains an odd cycle C that includes all nodes in $V_2 \setminus \bar{V}_2$, and $|V_2 \setminus \bar{V}_2| = 3$.* Let $p' = p - |\bar{V}_2|$. If $p' = 3$, we should have $z(v) = 1$ for all $v \in V_2$. Then it is easy to see that we have an integral polytope. So we assume that $p' \leq 2$. Let $V_2 \setminus \bar{V}_2 = \{v_1, v_2, v_3\}$.

Consider first $p' = 2$. If z is fractional, then at most one variable $z(v_i)$ can take the value one, so assume that

$$\begin{aligned}
&z(v_1) = 1,\\
&1 > z(v_2) > 0,\\
&1 > z(v_3) = 1 - z(v_2) > 0.
\end{aligned}$$

We give the label $l(v_2) = +1$ to v_2, the label $l(v_3) = -1$ to v_3, and $l(v) = 0$ for every other node in V_2. Then for each arc (u, v) with $z(u, v) = z(v)$, we give it the label $l(u, v) = l(v)$. If there is a node $u \in V_1$ that has only one arc (u, v) incident to it that is labeled, pick another arc (u, w) with $z(u, w) > 0$ and give it the label $l(u, w) = -l(u, v)$. For all the other arcs give the label 0. These labels define a new vector that satisfies with equation the same constraints that z does.

Now suppose that

$$\begin{aligned}
&1 > z(v_1) > 0,\\
&1 > z(v_2) > 0,\\
&1 > z(v_3) > 0.
\end{aligned}$$

Then for every node $u \in V_1$ there is at most one arc (u, v) such that $z(u, v) = z(v)$. Otherwise there is a node $w \in V_2 \setminus \bar{V}_2$ with $z(w) = 1$. Let us define a new vector z' as follows. Start with $z' = 0$. Set $z'(v_1) = z'(v_2) = 1$, $z'(v_3) = 0$, and $z'(v) = 1$ for all $v \in \bar{V}_2$. Then for each arc (u, v_1) with $z(u, v_1) = z(v_1)$ set $z'(u, v_1) = 1$. Also for each arc (u, v_2) with $z(u, v_2) = z(v_2)$ set $z'(u, v_2) = 1$. For each node u with $\sum_{(u,v)\in\delta^+(u)} z'(u, v) = 0$, pick an arc (u, v) with $v \neq v_3$ and set $z'(u, v) = 1$. This new vector satisfies with equality all the constraints that z does.

Finally suppose $p' = 1$ and

$$\begin{aligned}
&z(v_1) > 0,\\
&z(v_2) > 0,\\
&z(v_3) > 0.
\end{aligned}$$

We define a new vector z' as below. We set $z'(v_1) = 1$, $z'(v_2) = z'(v_3) = 0$, and $z'(v) = 1$ for $v \in \bar{V}_2$. For each node $u \in V_1$, if the arc (u, v_1) exists, we set $z(u, v_1) = 1$; otherwise there is a node $v \in \bar{V}_2$ such that the arc (u, v) exists, we set $z'(u, v) = 1$. We set $z'(s, t) = 0$ for every other arc. Every constraint that is satisfied with equality by z is also satisfied with equality by z'.

6.2. **Undirected graphs.** For a undirected graph $G = (V, E)$ we denote by $\overleftrightarrow{G} = (V, A)$ the directed graph obtained from G by replacing each edge $uv \in E$ by two arcs (u, v) and (v, u).

Theorem 45. *Let G be a connected undirected graph. Then $P_p(\overleftrightarrow{G})$ is integral for all p if and only if G is a path or a cycle.*

Proof. cf. [5]. If G is a path or a cycle, then $\overleftrightarrow{G}$ satisfies conditions (i) and (ii) of Theorem 20 and so $P_p(\overleftrightarrow{G})$ is integral.

Suppose G is not a path nor a cycle. Then G contains a node of degree at least 3. Thus $\overleftrightarrow{G}$ contains H_4 as a subgraph. Again Theorem 20 implies that $P_p(\overleftrightarrow{G})$ is not integral for all p. □

Acknowledgments

We are grateful to the referees for several comments that helped us to improve the presentation.

References

[1] P. Avella and A. Sassano, *On the p-median polytope*, Math. Program., 89 (2001), pp. 395–411.
[2] M. Baïou and F. Barahona, *On the p-median polytope of Y-free graphs*, Technical Report RC23636, IBM T. J. Watson Research Center, 2005. Available at http://www.optimization-online.org.
[3] M. Baïou and F. Barahona, *On the integrality of the uncapacitated facility location polytope*, Technical Report RC24101, IBM Watson Research Center, 2006.
[4] ———, *On the linear relaxation p-median problem I: oriented graphs*, technical report, IBM Watson Research Center, 2007.
[5] ———, *On the linear relaxation p-median problem II: directed graphs*, technical report, IBM Watson Research Center, 2007.
[6] J. Byrka and K. Aardal, *The approximation gap for the metric facility location problem is not yet closed*, Oper. Res. Lett., 35 (2007), pp. 379–384.
[7] L. Cánovas, M. Landete, and A. Marín, *On the facets of the simple plant location packing polytope*, Discrete Appl. Math., 124 (2002), pp. 27–53. Workshop on Discrete Optimization (Piscataway, NJ, 1999).
[8] A. Caprara and M. Fischetti, $\{0, \frac{1}{2}\}$*-Chvátal-Gomory cuts*, Math. Programming, 74 (1996), pp. 221–235.
[9] D. C. Cho, E. L. Johnson, M. Padberg, and M. R. Rao, *On the uncapacitated plant location problem. I. Valid inequalities and facets*, Math. Oper. Res., 8 (1983), pp. 579–589.
[10] D. C. Cho, M. W. Padberg, and M. R. Rao, *On the uncapacitated plant location problem. II. Facets and lifting theorems*, Math. Oper. Res., 8 (1983), pp. 590–612.
[11] F. A. Chudak and D. B. Shmoys, *Improved approximation algorithms for the uncapacitated facility location problem*, SIAM J. Comput., 33 (2003), pp. 1–25.
[12] V. Chvátal, *On certain polytopes associated with graphs*, J. Combinatorial Theory Ser. B, 18 (1975), pp. 138–154.
[13] M. Conforti and M. R. Rao, *Structural properties and recognition of restricted and strongly unimodular matrices*, Math. Programming, 38 (1987), pp. 17–27.
[14] 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.
[15] G. Cornuejols and J.-M. Thizy, *Some facets of the simple plant location polytope*, Math. Programming, 23 (1982), pp. 50–74.
[16] I. R. de Farias, Jr., *A family of facets for the uncapacitated p-median polytope*, Oper. Res. Lett., 28 (2001), pp. 161–167.
[17] C. De Simone and C. Mannino, *Easy instances of the plant location problem*, Technical Report R. 427, IASI, CNR, 1996.

[18] M. Guignard, *Fractional vertices, cuts and facets of the simple plant location problem*, Math. Programming Stud., (1980), pp. 150–162. Combinatorial optimization.
[19] E. L. Lawler, *Combinatorial optimization: networks and matroids*, Holt, Rinehart and Winston, New York, 1976.
[20] P. B. Mirchandani and R. L. Francis, eds., *Discrete location theory*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons Inc., New York, 1990.
[21] R. Müller and A. S. Schulz, *Transitive packing: a unifying concept in combinatorial optimization*, SIAM J. Optim., 13 (2002), pp. 335–367 (electronic).
[22] 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.
[23] D. B. Shmoys, É. Tardos, and K. Aardal, *Approximation algorithms for facility location problems (extended abstract)*, in 29th ACM Symposium on Theory of Computing, 1997, pp. 265–274.
[24] G. Stauffer, *The p-median polytope of Y-free graphs: an application of the matching theory*, technical report, 2007. To appear in Op. Res. Letters.
[25] M. Sviridenko, *An improved approximation algorithm for the metric uncapacitated facility location problem*, in Integer programming and combinatorial optimization, vol. 2337 of Lecture Notes in Comput. Sci., Springer, Berlin, 2002, pp. 240–257.
[26] F. Toumani, *Personal communication*. 2002.
[27] 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.
[28] M. Yannakakis, *On a class of totally unimodular matrices*, Math. Oper. Res., 10 (1985), pp. 280–304.

(M. Baïou) CNRS, LIMOS Complexe scientifique des Cezeaux, 63177 Aubière, 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`