Journal of Combinatorial Optimization, 9, 35–47, 2005

Requiring Connectivity in the Set Covering Problem

J. ORESTES CERDEIRA*,† orestes@isa.utl.pt
Instituto Superior de Agronomia, Univ. Técnica de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal

LEONOR S. PINTO‡ lspinto@iseg.utl.pt
Instituto Superior de Economia e Gestão, Univ. Técnica de Lisboa, Rua do Quelhas 6, 1200-781 Lisboa, Portugal

Received May 8, 2003; Revised July 26, 2004; Accepted September 21, 2004

Abstract. Given a bipartite graph with bipartition V and W, a cover is a subset $C \subseteq V$ such that each node of W is adjacent to at least one node in C. The set covering problem seeks a minimum cardinality cover. Set covering has many practical applications. In the context of reserve selection for conservation of species, V is a set of candidate sites from a reserve network, W is the set of species to be protected, and the edges describe which species are represented in each site. Some covers however may assume spatial configurations which are not adequate for conservational purposes. Indeed, for sustainability reasons the fragmentation of existing natural habitats should be avoided, since this is recognized as being disruptive to the species adapted to the habitats. Thus, connectivity appears to be an important issue for protection of biological diversity. We therefore consider along with the bipartite graph, a graph G with node set V, describing the adjacencies of the elements of V, and we look for those covers $C \subseteq V$ for which the subgraph of G induced by C is connected. We call such covers connected covers. In this paper we introduce and study some valid inequalities for the convex hull of the set of incidence vectors of connected covers.

Keywords: set covering, graphs, connected components, integer polytopes

MSC2000: 90C10, 90C57

1. Introduction

Given a bipartite graph with bipartition V and W, a cover is a subset $C \subseteq V$ such that each node of W is adjacent to at least one node in C. The set covering problem seeks a minimum cardinality cover. The problem has been extensively studied, and several papers have focused on the polyhedral structure of the convex hull of the set of incidence vectors of covers. Balas and Ng (1989a, b) characterize facets with coefficients and right hand sides in {0, 1, 2}. Sánchez-García et al. (1998) extend this work to {0, 1, 2, 3}. Other classes of facets are study in Cornuéjols and Sassano (1989), Nobili and Sassano (1989), and Sassano (1989).

Set covering has many practical applications (Vemuganti, 1998). It appears as a basic model in reserve selection for conservation of species. In this context, V is a set of candidate

*Corresponding author.
†This author's research was financially supported by the Portuguese Foundation for Science and Technology (FCT).
‡This paper is part of this author's Ph.D. research.

sites from a reserve network, W is the set of species to be protected, and the edges of the graph describe which species are represented in each site.

Some covers however may have spatial configurations which, for sustainability and management cost reasons, are not appropriate (Nalle et al., 2002). In particular, fragmentation is considered an undesirable design attribute. Biologists defend that habitat fragmentation may precipitate population decline and extinction by dividing an existing widespread population into subpopulations in a restricted area. These smaller populations are then more vulnerable to inbreeding depression, genetic drift and other problems associated with small population size (Primack, 2002). The establishment of wildlife corridors is used to counteract fragmentation (Primack, 2002; Williams, 1998). In fact, according to the overview experiments from Debinski and Holt (2000), the presence of corridors enhance movement for at least some species and increase species richness in fragments.

Connectivity thus appears to be an important issue for protection of biological diversity, and several methods have been recently proposed for reserve selection which incorporate connectivity. These include iterative procedures in which a target parameter called connectivity value is sequentially updated depending on which sites have been selected in the previous iterations (Siitonen et al., 2002, 2003), algorithms to minimize the distances between pairs of sites to be included in a network (Nicholls and Margules, 1993; Briers, 2002; Önal and Briers, 2002), the summed distance between all pairs of sites (Briers, 2002; Nalle et al., 2002), the boundary length of a network (Önal and Briers, 2003), or a combination of boundary length and total area (Possingham et al., 2000; McDonnell et al., 2002).

All of these approaches treat connectivity as a quantified target incorporated somehow in the objective function. We develop an alternative approach in which connectivity is explicitly addressed as part of the model constraints, so that only connected networks are feasible solutions. We therefore consider the following extension of the set covering. Along with the bipartite graph, we are given a graph G with node set V, describing the adjacencies of the elements of V, and we look for those covers $C \subseteq V$ for which the subgraph of G induced by C is connected. We call such covers connected covers.

In this paper we identify some classes of facets for the connected cover polytope, i.e., the convex hull of the set of incidence vectors of connected covers. We start by giving some notation in Section 2. In each of Sections 3, 4, 5, 6 we introduce a certain class of inequalities, and give conditions for each inequality to be valid and to describe facets for the connected cover polytope.

2. Notation

Consider a bipartite graph with bipartition V and W, and a connected graph G with vertex set V. A *cover* is a subset $C \subseteq V$ such that each node in W is adjacent to at least one node of C. The *cover polytope* is the convex hull of the set of incidence vectors of covers. A cover which induces a connected subgraph of G is called a *connected cover*. The *connected cover polytope*, denoted by $\mathcal{P}$, is the convex hull of the set of incidence vectors of connected covers. Note that when G is complete every cover is a connected cover.

In what follows we will assume that, for every $v \in V$, $V \backslash \{v\}$ is a connected cover. This is equivalent to assuming that the bipartite graph is such that each vertex of W is adjacent to more than one vertex in V, and that graph G is 2-connected. The assumption of 2-connectedness will permit to simplify some statements and proofs, and without any associated loss of generality. Indeed from a connected graph we can easily construct a 2-connected graph by duplicating each articulation vertex v, i.e., adding vertex v' with the same neighbours as v. Obviously, every minimal connected cover does not include both v and v'. With these assumptions it results immediately that $\mathcal{P}$ is full dimensional, and that, for every $v \in V$, $x_v \leq 1$ induces a facet of $\mathcal{P}$.

Consider any subset $S \subseteq V$. Denote by $\mathcal{C}(S)$ the set of connected components of the subgraph of G induced by $V \backslash S$. If $v \in V \backslash S$, let C_v denote the component of $\mathcal{C}(S)$ which includes v, and S_v be the set of vertices of S which are adjacent to at least one vertex in C_v. If $s \in S$, C_s denotes the set of vertices of the components of $\mathcal{C}(S)$ which include at least one vertex adjacent to s. Finally, for $x \in \mathbb{R}^V$ we use $x(S)$ for $\sum_{s \in S} x_s$ and assume $x(\emptyset) = 0$.

3. Inequalities of type $x(S) \geq 1$

For the cover polytope, Balas and Ng (1989a) identify among the valid inequalities of type

$$x(S) \geq 1, \quad S \subseteq V \tag{1}$$

those that are facet defining.

The inequality (1) is valid for the cover polytope iff $V \backslash S$ is not a cover. This holds when, in the bipartite graph, S contains the set of all vertices which are adjacent to some vertex of W.

Balas and Ng (1989a) state the following.

Theorem 1 (Balas and Ng, 1989a)**.** *If the inequality* (1) *is valid for the cover polytope*, *it induces a facet iff*

(a) *for every* $s \in S$, $V \backslash S \cup \{s\}$ *is a cover*, *and*

(b) *for every* $v \in V \backslash S$ *there exists* $s \in S$ *such that* $(V \backslash S) \backslash \{v\} \cup \{s\}$ *is a cover.*

Note that if the set S of a valid inequality satisfies (a), S is a minimal set of all vertices adjacent to some vertex of W. Hence, no more than $|W|$ inequalities of type (1) define facets of the cover polytope.

As Balas and Ng noticed, Theorem 1 follows directly from a result of Hammer et al. (1975) and Wolsey (1975) on independence systems. We state this result for any collection $\mathcal{F}$ of subsets of V with $V \in \mathcal{F}$ and satisfying $B \in \mathcal{F}$, whenever $A \in \mathcal{F}$ and $B \supseteq A$. Clearly, (1) being valid means now $V \backslash S \notin \mathcal{F}$.

Theorem 2 (Hammer et al., 1975; Wolsey, 1975)**.** *If* (1) *is valid for the convex hull of the set of incidence vectors of members of* $\mathcal{F}$, *it defines a facet of this polytope iff*

(c) *for every* $s \in S$, $V \backslash S \cup \{s\} \in \mathcal{F}$, *and*

(d) *for every* $v \in V \backslash S$ *there exists* $s \in S$ *such that* $(V \backslash S) \backslash \{v\} \cup \{s\} \in \mathcal{F}$.

If we identify the minimal elements of $\mathcal{F}$ with minimal connected covers, Theorem 2 characterizes the valid inequalities of type (1) which induce facets of the convex hull $\mathcal{Q} \supseteq \mathcal{P}$ of the set of incidence vectors of the subsets of V that include connected covers.

Inequality (1) being valid for $\mathcal{Q}$ (and $\mathcal{P}$) means that no component of $\mathcal{C}(S)$ is a cover. For valid inequalities, condition (c) states in terms of $\mathcal{Q}$ that

(e) for every $s \in S$, the connected component $C_s \cup \{s\}$ is a cover.

This is obviously necessary for (1) to define a facet of $\mathcal{Q}$ and $\mathcal{P}$, since if for vertex $s \in S$, $C_s \cup \{s\}$ is not a cover, $x(S \backslash \{s\}) \geq 1$ would be valid for $\mathcal{Q}$ and $\mathcal{P}$.

Condition (d) is not sufficient to construct a list of $|V|$ linearly independent members of $\mathcal{P}$ satisfying (1) with equality. Indeed, we will show that the following strongest condition must hold.

(f) for every $v \in V \backslash S$ there exists $s \in S_v$ such that $C_s \backslash \{v\} \cup \{s\}$ includes a connected cover.

Conditions (d) and (f) differ as where to find s in S. Condition (f) is more restrictive, requiring, for each $v \in V \backslash S$, s to be adjacent to some vertex in C_v.

If, for $v \in V \backslash S$, condition (d) holds while (f) fails (i.e., $S \neq S_v$ and, for every $s \in S_v$, no connected cover is included in $C_s \backslash \{v\} \cup \{s\}$), then the two inequalities

$$x(S_v) \geq x_v \tag{2}$$

$$x(S_v) + 2x(S \backslash S_v) + x_v \geq 2 \tag{3}$$

would be valid for $\mathcal{P}$. Since (1) is obtained halfing the sum of (2) and (3), this shows that inequality (1) defines no facet of $\mathcal{P}$.

Note that if $S = S_v$ and condition (f) is not satisfied, then inequality $x(S) + x_v \geq 2$, which implies (1), is valid for $\mathcal{P}$ (as well as for $\mathcal{Q}$).

The arguments above allow to conclude the following.

Lemma 3. *Conditions* (e) *and* (f) *are necessary for the inequality* (1), *if valid for* $\mathcal{P}$, *to be facet inducing.*

We proceed proving sufficiency. We will do this exhibiting, for each S satisfying (e) and (f), a linearly independent set of $|V|$ incidence vectors of connected covers for which (1) holds with equality. For this purpose we describe an algorithm to (totally) order the vertices of each component of $\mathcal{C}(S)$, and to define, for $v \in V \backslash S$, a connected cover M_v contained in $C_s \backslash \{v\} \cup \{s\}$, with $s \in S_v$, that includes every vertex $u \in C_v$ for which $u < v$.

Let C be a component of $\mathcal{C}(S)$, T be any spanning tree of C, and denote by $A(v)$ the set of vertices of graph (C, T) adjacent to v. The algorithm starts with all vertices of C unmarked. In each step an unmarked vertex v is selected, its order position with respect to every other vertex of C is defined, and vertex v is marked. This ordering is carried out in two phases.

Phase one

1. If every vertex is marked, the procedure finishes. Otherwise, select an unmarked vertex v such that no more than one vertex in $A(v)$ is unmarked. In consequence, all the unmarked vertices different from v belong to a same component of $T \backslash \{v\}$. The order position of v will depend on vertex $s \in S_v$, for which $C_s \backslash \{v\} \cup \{s\}$ includes a connected cover.
2. Let M be the maximal connected cover contained in $C_s \backslash \{v\} \cup \{s\}$. Note that since all the unmarked vertices different from v belong to a same component of $T \backslash \{v\}$, M either contains all the unmarked vertices of $C \backslash \{v\}$, or none of them.
3. If M contains all the unmarked vertices of $C \backslash \{v\}$, let $M_v := M$; define $u < v$, for every unmarked vertex u and $v < w$, for every marked vertex w; consider v as marked and proceed to 1.
4. If M does not include unmarked vertices, then vertex v is defined as the minimum element of C and proceed to Phase two.

Phase two

1. Mark vertex v.
2. If every vertex is marked, the procedure finishes. Otherwise, select an unmarked vertex v such that $A(v)$ includes a vertex marked during Phase two.
3. Settles $u < v$ for every u marked on Phase two, $v < w$, for every $w \neq v$ unmarked or marked on Phase one and mark v.
4. Let $M_v := M_u \cup \{u\}$, where u is the vertex which was marked immediately before v, and proceed to 2 (of Phase two).

If condition (f) holds, the algorithm defines, for each component $C \in \mathcal{C}(S)$ and for each $v \in C$, a connected cover M_v, which is such that $v \notin M_v$ and $u \in M_v$, for every $u \in C$, with $u < v$. If condition (e) holds, for every $s \in S$, $M_s = C_s \cup \{s\}$ is a connected cover. Each of these covers includes exactly one vertex from S.

To show that the set of the incidence vectors of these connected covers is linearly independent we construct a matrix M whose rows are the incidence vectors, and prove that M is nonsingular.

The first $|S|$ columns of M are indexed by S. The remaining $|V| - |S|$ columns are partitioned in $|\mathcal{C}(S)|$ blocks of consecutive columns, each block corresponding to a connected component of $\mathcal{C}(S)$. The columns of each block are arranged from left to right indexed by increasing vertices of the corresponding component, with respect to the order established by the above algorithm.

The first $|S|$ rows of M consist of the incidence vectors of M_s, $s \in S$, arranged in such a way that the left most $|S| \times |S|$ principal submatrix of M is the identity matrix. The remaining $|V| - |S|$ rows are partitioned in blocks in the same way that the corresponding columns were partitioned. The rows in each block are filled from up to down with the incidence vectors of M_v, by ascending order of the vertices v of the corresponding component.

We use matrix M to prove the following.

Lemma 4. *If conditions* (e) *and* (f) *hold, there is a linearly independent set of* $|V|$ *incidence vectors of connected covers, each satisfying* (1) *with equality.*

Proof: It amounts to show that matrix M is nonsingular.

Consider any row $i > |S|$ and let $v \in V \backslash S$ be the vertex corresponding to i. Row i is therefore the incidence vector of the connected cover $M_v \subseteq C_s \backslash \{v\} \cup \{s\}$, with $s \in S_v$. Let $j \leq |S|$ be the row of the incidence vector of $M_s = C_s \cup \{s\}$. The vector which results subtracting row j from row i has -1 as its first nonzero entry, and this occurs in the position i. Hence, it is possible to perform elementary row operations on M to obtain an upper triangular matrix with every entry of the main diagonal equal to either 1 (on the first $|S|$ entries) or -1 (on every other entry). □

We can now state the result that generalizes Theorem 1 for the connected cover polytope.

Theorem 5. *If inequality* (1) *is valid for* $\mathcal{P}$, *it induces a facet iff conditions* (e) *and* (f) *hold.*

Proof: The result follows from Lemmas 3 and 4, and the assumption that $\dim \mathcal{P} = |V|$. □

4. Inequalities of type $x(S^1) + 2x(S^2) \geq 2$

Balas and Ng (1989a) also characterize the inequalities of type

$$x(S^1) + 2x(S^2) \geq 2, \quad S = S^1 \cup S^2 \subseteq V, \quad \text{with } S^1 \cap S^2 = \emptyset \tag{4}$$

that induce facets of the cover polytope. Their result can be straightforwardly generalized for the polytope associated with the collection $\mathcal{F}$ of subsets of V defined in Section 3. We state the result for $\mathcal{Q}$, the convex hull of the set of incidence vectors of subsets of V which include connected covers. Inequality (4) is valid for $\mathcal{Q}$ and $\mathcal{P}$ iff, for every $s \in S^1$, $V \backslash S \cup \{s\}$ does not contain a connected cover. Let us call 2-connected cover graph the graph $G^1 = (S^1, E^1)$, having vertex set S^1 and edge set $E^1 = \{[s', s''] : C_{s'} \cup C_{s''} \cup \{s', s''\}$ is a connected cover$\}$.

Theorem 6 (Balas and Ng, 1989a). *A valid inequality* (4) *defines a facet of* $\mathcal{Q}$ *iff*

(g) *for every* $s \in S^2$, $C_s \cup \{s\}$ *is a connected cover*,

(h) *there is an odd cycle in every component of the* 2*-connected cover graph* G^1, *and*

(i) *for every* $v \in V \backslash S$ *there exists either*

  (i1) $s \in S^2$ *such that* $(V \backslash S) \backslash \{v\} \cup \{s\}$ *includes a connected cover*, *or*

  (i2) $s', s'' \in S^1$ *such that* $(V \backslash S) \backslash \{v\} \cup \{s', s''\}$ *includes a connected cover.*

It is easy to derive valid inequalities implying (4) whenever any of these conditions fails.

Lemma 7. *Conditions* (g), (h) *and* (i) *are necessary for the valid inequality* (4) *to define a facet of* ($\mathcal{Q}$ *and*) $\mathcal{P}$.

Proof: If condition (g) does not hold, i.e., if for some $s \in S^2$, $C_s \cup \{s\}$ is not a connected cover, inequality $x(S^1 \cup \{s\}) + 2x(S^2 \backslash \{s\}) \geq 2$, which implies (4), would be valid.

To see (h) is necessary, suppose some component of G^1 is a bipartite graph with bipartition A and B. Inequalities

$$1.5\,x(S^1 \backslash (A \cup B)) + x(A) + 2x(B) + 3x(S^2) \geq 3 \quad \text{and}$$
$$1.5\,x(S^1 \backslash (A \cup B)) + 2x(A) + x(B) + 3x(S^2) \geq 3$$

would hold for $\mathcal{Q}$ and $\mathcal{P}$, and their sum equals 3 times (4).

If, for $v \in V \backslash S$, both (i1) and (i2) are not satisfied, inequality $x(S^1) + 2x(S^2) + x_v \geq 3$ is valid and implies (4). □

We were not able to characterize when a valid inequality (4) induces a facet of $\mathcal{P}$. Yet we can prove the following results.

Lemma 8. *If a valid inequality* (4) *satisfies* (g), (h), *and*

(j) *for every* $v \in V \backslash S$ *there exists either*
- (j1) $s \in S^2 \cap S_v$ *such that* $C_s \backslash \{v\} \cup \{s\}$ *includes a connected cover*, *or*
- (j2) $s' \in S^1 \cap S_v$ *and* $s'' \in S^1$ *such that* $C_{s'} \cup C_{s''} \backslash \{v\} \cup \{s', s''\}$ *includes a connected cover*

it induces a facet for $\mathcal{P}$.

Proof: As we did when proving Lemma 4, we define a matrix M whose rows are the incidence vectors of connected covers satisfying (4) with equality, and prove that the rank of M is $|V|$.

The first $|S|$ columns of M are indexed by S^1 and S^2, in this order. The remaining $|V| - |S|$ columns are partitioned in $|\mathcal{C}(S)|$ blocks of consecutive columns, each block corresponding to a connected component of $\mathcal{C}(S)$.

The first rows are the incidence vectors of $M_{[s',s'']} = C_{s'} \cup C_{s''} \cup \{s', s''\}$, for each edge $[s', s'']$ of the 2-connected cover graph G^1, whose definition ensures to be connected covers. Note that the submatrix consisting of these $|E^1|$ rows and of the columns indexed by S^1 has rank $|S^1|$, since it is the edge-node incidence matrix of a graph which according to condition h) has no bipartite components. (The rank of the incidence matrix of a connected graph with p vertices is equal to $p - 1$, if it is bipartite, and equal to p otherwise (Sachs, 1967; Van Nuffelen, 1973).)

The next $|S^2|$ rows consist of the incidence vectors of $M_s = C_s \cup \{s\}$, $s \in S^2$, arranged in such a way that the $|S^2| \times |S^2|$ submatrix lying on these rows and on the columns indexed by S^2 is the identity matrix. Condition (g) ensures that each M_s is a connected cover.

From the above construction it is clear that the set of the first $|S|$ columns of M is linearly independent.

The remaining $|V| - |S|$ rows are partitioned in $|\mathcal{C}(S)|$ blocks of consecutive rows, in such a way that the rows of the i-th block refer to the same component of $\mathcal{C}(S)$ as the columns of the i-th block.

To order the vertices of each component C of $\mathcal{C}(S)$, and to define, for each $v \in C$, a connected cover M_v including either a single vertex of S^2 or exactly two vertices of S^1, and no other vertex from S, the algorithm of Section 3 will be used. However, using the algorithm requires some minor changes regarding the vertices of S^1. Every edge $[s', s'']$ of G^1 will be treated as a vertex $s_{[s',s'']}$, $C_{s_{[s',s'']}}$ being the set of vertices of the components of $\mathcal{C}(S)$ which include at least one vertex adjacent to either s' or s''. If $v \in V \backslash S$, $s_{[s',s'']} \in S_v$ iff either s' or s'' is adjacent to one vertex in C_v. Thus, $S_v := (S^2 \cap S_v) \cup \{s_{[s',s'']} : [s', s''] \in E^1$ with either s' or s'' adjacent to one vertex of $C_v\}$. Condition (j) now reads: for every $v \in V \backslash S$ there exists $s \in S_v$ such that $C_s \backslash \{v\} \cup \{s\}$ includes a connected cover. Observe that, when $s = s_{[s',s'']}$, $C_s \backslash \{v\} \cup \{s\}$ is equal to $C_{s'} \cup C_{s''} \backslash \{v\} \cup \{s', s''\}$.

The columns (rows) of M corresponding to each component C of $\mathcal{C}(S)$, are arranged from left to right (up to down) indexed by increasing vertices of C, with respect to the order established by the algorithm. For $v \in V \backslash S$, the incidence vector of M_v is in the row indexed by v. Recall that M_v is contained in either $M_s = C_s \backslash \{v\} \cup \{s\}$, for some s in S^2, or $M_{[s',s'']} = C_{s'} \cup C_{s''} \backslash \{v\} \cup \{s', s''\}$, for some edge $[s', s'']$ of G^1, whose incidence vector is one of the $|E^1| + |S^2|$ first rows of M. It is now possible to perform elementary row operations as in the proof of Lemma 4 to obtain an upper triangular submatrix with -1 in all principal entries lying in the last $|V| - |S|$ rows and columns. This allows to conclude that M is a full column rank matrix. □

Lemma 9. *Suppose that condition* (g) *is verified, and each component of the* 2*-connected cover graph* $G^1 = (S^1, E^1)$ *has a unique cycle which is odd. If* (j) *fails a valid inequality* (4) *does not induce a facet of* $\mathcal{P}$.

Proof: We prove that every incidence vector of a connected cover that satisfies

$$x(S^1) + 2x(S^2) = 2 \tag{5}$$

also satisfies a linear equation which is not a scalar multiple of (5).

Suppose G^1 has only one component, and let $[s', s'']$ be an edge of its unique cycle.

Every incidence vector of a connected cover that satisfies (5) and that does not use both s' and s'' also satisfies

$$x(A) = x(B), \tag{6}$$

for bipartition A and B of graph $(S^1, E^1 \setminus [s', s''])$.

Let $v \in V \backslash S$ be such that neither (j1) nor (j2) hold and suppose there is $\alpha = (\alpha_s)$, $s \in S^1$, such that for every edge $[i, j] \in E^1 \setminus [s', s'']$, $\alpha_i + \alpha_j = 1$ if either i or j are in $S^1 \cap S_v$, and $\alpha_i + \alpha_j = 0$, otherwise. Then

$$\sum_{s \in S^1} \alpha_s x_s + x(S^2 \cap S_v) = x_v \tag{7}$$

is satisfied by every incidence vector of a connected cover verifying (5), and not including both s' and s''. The following algorithm shows how to define α. Choose any vertex $s \in S^1$ and let $\alpha_s := 0$. While there are vertices $s \in S^1$ for which α_s is not settled, take any edge $[i, j]$ of G^1, with α_j not yet settled and α_i settled already and define $\alpha_j := 1 - \alpha_i$ if either i or j are in $S^1 \cap S_v$, and $\alpha_j := -\alpha_i$, otherwise.

Every incidence vector of a connected cover that satisfies (5) and that does not use both s' and s'', also verifies the Eqs. (6) and (7). These three equations are linearly independent, and the coefficients associated to vertices not in $S \cup \{v\}$ are all equal to zero. The other connected covers whose incidence vectors also satisfy (5), all use the same vertices of $S \cup \{v\}$ (besides s' and s'', either all include v, or none includes v). This allows to conclude the existence of a linear equation (with the coefficients associated to vertices not in $S \cup \{v\}$ all equal to zero), which is not a escalar multiple of (5), that is satisfied by all the incidence vectors of connected covers that satisfy (5).

To extend the proof for the case where G^1 has $k > 1$ components, consider an edge $[s'_i, s''_i]$ of the unique cycle of the connected component C^1_i, and let A_i and B_i be the sets of the bipartition resulting from deleting $[s'_i, s''_i]$ from C^1_i. Every incidence vector of a connected cover that satisfies (5) and that does not use both s'_i and s''_i in any C_i also satisfies

$$x(A_i) = x(B_i), \quad i = 1, \ldots, k. \tag{8}$$

The algorithm above can be easily adapted to run for each connected component C^1_i of G^1 to derive values of α such that (7) is satisfied by the incidence vector of every the connected cover that does not use any pair of vertices s'_i, s''_i simultaneously. The $k+2$ Eqs. (5), (7) and (8) are linearly independent, and the coefficients associated to vertices not in $S \cup \{v\}$ are all equal to zero. Since the connected covers whose incidence vectors satisfy (5) and that use both s'_i, s''_i, all have the same vertices of $S \cup \{v\}$ (besides s'_i and s''_i, either all include v, or none includes v), there has to be a linear equation (with the coefficients associated to vertices not in $S \cup \{v\}$ all equal to zero), which is not a escalar multiple of (5), that is satisfied by all the incidence vectors of connected covers that satisfy (5). □

Lemma 10. *If each component of the* 2*-connected cover graph* $G^1 = (S^1, E^1)$ *has a unique cycle, a valid inequality* (4) *defines a facet of* $\mathcal{P}$ *iff conditions* (g), (h) *and* (j) *hold.*

Proof: The result follows directly from Lemmas 7, 8 and 9. □

5. Inequalities of type $x(S) \geq x_a$

Here we consider inequalities of type

$$x(S) \geq x_a, \quad S \subset V, \quad \text{and} \quad a \in V \backslash S \tag{9}$$

Unlike (1) and (4), inequalities of type (9) do not define facets of $\mathcal{Q}$. Indeed, whenever (9) is valid for $\mathcal{Q}$ (i.e., no component in $\mathcal{C}(S)$ is a cover), the stronger inequality $x(S) \geq 1$ also holds.

For $\mathcal{P}$, (9) is a valid inequality iff the component C_a is not a cover. We will show that the conditions for a valid inequality (9) to define a facet of $\mathcal{P}$ are (e) and

(k) $S = S_a$,
(l) for every $v \in V \backslash S$, $v \neq a$, either

  (l1) there exists $s \in S_v$ such that $C_s \backslash \{v\} \cup \{s\}$ includes a connected cover to which a belongs, or
  (l2) C_v is a cover and v is the unique vertex in C_v that does not satisfy (l1), and

(m) at least one component $C \in \mathcal{C}(S)$ is a cover and every vertex $v \in C$ satisfies (l1).

Lemma 11. *Conditions* (e), (k), (l) *and* (m) *are necessary for a valid inequality* (9) *to induce a facet of* $\mathcal{P}$.

Proof: Conditions (e) and (k) must hold since if either there is s in S such that $C_s \cup \{s\}$ is not a cover, or we can find s in $S \backslash S_a$, then $x(S \backslash \{s\}) \geq x_a$ would be valid for $\mathcal{P}$.

If $v \in V \backslash S$, $v \neq a$, does not satisfy (l1), then $x(S_v) + 2x(S \backslash S_v) + x_v \geq 2x_a$ is valid. If C_v is not a cover, $x(S_v) \geq x_v$ is also valid, and these two inequalities imply (9). If C_v is a cover and for some vertex $u \in C_v$, $u \neq v$, there exists no $s \in S_v$ such that $C_s \backslash \{u\} \cup \{s\}$ includes a connected cover, 2-connectedness of G insures that $x(S) + x_v \geq x_a + x_u$ and $x(S) + x_u \geq x_a + x_v$ hold. The sum of the two inequalities is equal to two times (9).

At least one component of $\mathcal{C}(S)$ must be a cover, otherwise every point of $\mathcal{P}$ would satisfy $x(S) \geq 1$ which implies (9). Let U be the set of vertices $v \in V \backslash S$ not satisfying (l1). Note that if $v \neq v' \in U$ (which only occurs if v and v' are in different components of $\mathcal{C}(S)$), then $S_v \cap S_{v'} = \emptyset$, since otherwise any $s \in S_v \cap S_{v'}$ is such that $C_s \backslash \{v\} \cup \{s\}$ includes the connected cover $C_{v'} \cup C_a \cup \{s\}$. Therefore, if S_U is the set of vertices in S adjacent to at least one component of $\mathcal{C}(S)$ that includes one vertex of U, then $x(S_U) \geq x(U) + x_a - 1$ is valid for $\mathcal{P}$. Indeed, every connected cover including $k > 1$ vertices of U, has to include k vertices from S_U, and every connected cover including a and a single vertex of U, has to include one vertex from S_U. If (m) fails, i.e., every component of $\mathcal{C}(S)$ which is a cover has one vertex in U, then $x(S_U) + 2x(S \backslash S_U) + x(U) \geq x_a + 1$ would be valid for $\mathcal{P}$. Halfing the sum of the last two inequalities we get (9), showing that it would not define a facet of $\mathcal{P}$.

□

Lemma 12. *Conditions* (e), (k), (l) *and* (m) *are sufficient for a valid inequality* (9) *to be facet inducing for* $\mathcal{P}$.

Proof: We display an affinely independent set of $|V|$ incidence vectors of connected covers satisfying (9) with equality.

Conditions (e) and (k) ensure that, for each $s \in S$, $M_s = C_s \cup \{s\}$ is a connected cover with $a \in M_s$.

The remaining $|V| - |S|$ connected covers are obtained using the algorithm of Section 3.

For each component $C \in \mathcal{C}(S)$, $C \neq C_a$, such that every vertex $v \in C$ satisfies (l1), let the algorithm order the vertices v of C and define M_v, the maximal connected cover contained in $C_s \backslash \{v\} \cup \{s\}$, with $s \in S_v$, to which a belongs.

When using the algorithm for C_a, take in consideration that when selecting an unmarked vertex in step 1 of Phase one, if there are two or more unmarked vertices, there is always more than one candidate for selection. Select always a candidate u different from a. Condition (l1) ensures that the connected cover M_u will include every unmarked vertex, and consequently Phase one will continue until a is the only unmarked vertex in C_a. When this happens define $a < u$, for every $u \in C_a \backslash \{a\}$ and let $M_a = C$, where $C \in \mathcal{C}(S)$ is any component to which (m) refers. Note that for $u \in C_a$, $u \neq a$, M_u includes a and a single vertex from S, and M_a includes neither a nor any vertex from S. Hence, the incidence vectors of all these connected covers satisfy (9) with equality.

When using the algorithm for each component C_v, where v is the unique vertex in C_v to which the condition (l2) refers, proceed as for C_a, always selecting an unmarked vertex $u \neq v$. Condition (l1) ensures that the connected cover M_u includes every unmarked vertex, and consequently Phase one will continue until v is the only unmarked vertex in C_v. When this happens define $v < u$, for every $u \in C_v \backslash \{v\}$ and $M_v = C_v$. Note that for $u \in C_v$, $u \neq v$, M_u includes a and a single vertex from S, and M_v includes neither a nor any vertex from S. Thus, the incidence vectors of these connected covers also satisfy (9) with equality.

The $|V|$ incidence vectors of M_s, $s \in S$ and M_v, $v \in V \backslash S$ are arranged in a matrix M', exactly in the same way that was used in Section 3 to obtain the matrix M. In order to show that the rows of M' are affinely independent, we now add to M' an extra column e in position $|V| + 1$ with all elements equal to 1, and represent the augmented matrix by $[M' \mid e]$. Matrices M and M' only differ in the row corresponding to vertex a and in the rows corresponding to vertices $v \in V \backslash S$ for which (l2) holds. For every row $i > |S|$ of M' different from these perform on $[M' \mid e]$ the row operation described in the proof of Lemma 4, and let $[M'' \mid e']$ be the resulting matrix. The left most $|S| \times |S|$ principal submatrix of M'' is the identity matrix. The submatrix of M'' with rows and columns indexed by the vertices of $(V \backslash S) \backslash \{a\}$ is a block diagonal matrix, where the blocks (corresponding to the components of $\mathcal{C}(S)$) are upper triangular matrices with diagonal entries equal to 1 on the rows not used in the elementary operations and equal to -1 elsewhere. Note that every row of $[M'' \mid e']$ whose first nonzero element is -1 has a zero entry in column e'. The rows indexed by the vertices of M_a, which is a component mentioned in condition (m), have leading elements equal to -1. The row corresponding to a (i.e., the incidence vector of M_a) has all entries equal to zero, except on the columns indexed by the vertices of M_a and on column e', where they are equal to 1. Hence, we can use the rows indexed by M_a to perform elementary operations to change the row indexed by a into a row with all elements equal to zero except for the element of column e' which remains equal to 1. This shows that $[M' \mid e]$ is a full row rank matrix. □

Theorem 13. *If inequality* (9) *is valid for* $\mathcal{P}$, *it induces a facet iff conditions* (e), (k), (l) *and* (m) *hold.*

Proof: Follows from Lemmas 11 and 12. □

6. Inequalities of type $x(S) \geq x_a + x_b - 1$

The inequalities of type

$$x(S) \geq x_a + x_b - 1, \quad S \subset V, \quad \text{and} \quad a, b \in V \backslash S \tag{10}$$

are valid for $\mathcal{Q}$ iff no component of $\mathcal{C}(S)$ is a cover. In such case $x(S) \geq 1$ is also valid, showing that (10) does not define facets of $\mathcal{Q}$.

For $\mathcal{P}$, inequality (10) is valid iff a and b are not both in a component of $\mathcal{C}(S)$ which is a cover.

Theorem 14. *If inequality* (10) *is valid for* $\mathcal{P}$, *it induces a facet iff*
(n) $S = S_a = S_b$, *and*
(o) C_a *and* C_b *are both covers.*

Proof: To conclude that (n) is necessary observe that if a vertex s could be found in $S \backslash S_a$ or in $S \backslash S_b$, then $x(S \backslash \{s\}) \geq x_a + x_b - 1$ would also be valid. To show that (o) is necessary, note that if C_a or C_b is not a cover, then every point in $\mathcal{P}$ would satisfy $x(S) \geq x_a$ or $x(S) \geq x_b$ respectively, and each of these inequalities implies (10).

To prove sufficiency we will exhibit a linearly independent set of $|V|$ incidence vectors of connected covers verifying (10) with equality.

The conditions (n) and (o) imply that for $s \in S$, $M_s = C_s \cup \{s\}$ is a connected cover including both a and b.

Since G is 2-connected, (n) and (o) also imply that for $v \in V \backslash S$, there exists $s \in S_v$ such that $C_s \backslash \{v\} \cup \{s\}$ includes a connected cover. If $v \notin C_a \cup C_b$, every maximal connected cover contained in $C_s \backslash \{v\} \cup \{s\}$ includes both a and b. We can therefore use the algorithm of Section 3 to order the vertices of every component $C \neq C_a, C_b$ and to define, for each $v \in C$, the connected cover M_v which includes a, b and a unique $s \in S$.

For $v \in C_a$ or $v \in C_b$, $v \neq a, b$, 2-connectedness ensures that s can be chosen so that the maximal connected cover contained in $C_s \backslash \{v\} \cup \{s\}$ also includes a and b. We can therefore implement this choice when using the algorithm for C_a and C_b, as well as keeping a and b unmarked during every step of phase one, as described in the proof of Lemma 12. In this way, for $v \in C_a \cup C_b$, $v \neq a, b$, the connected cover M_v defined by the algorithm includes a, b and a unique $s \in S$. When a and b are the only unmarked vertices in their components, define $a < v$ for every $v \in C_a \backslash \{a\}$, $b < v$ for every $v \in C_b \backslash \{b\}$, $M_a = C_a$ and $M_b = C_b$.

It is straightforward, following the proof of Lemma 12, to see that the set of incidence vectors of these $|V|$ connected covers is linearly independent. □

Acknowledgments

We are grateful to I. Faria for discussion and assistance.

References

E. Balas and S.M. Ng, "On the set covering polytope: I. all facets with coefficients in {0, 1, 2}," *Mathematical Programming*, vol. 43, pp. 57–69, 1989a.

E. Balas and S.M. Ng, "On the set covering polytope: II. lifting the facets with coefficients in {0, 1, 2}," *Mathematical Programming*, vol. 45, pp. 1–20, 1989b.

R.A. Briers, "Incorporating connectivity into reserve selection procedures," *Biological Conservation*, vol. 103, pp. 77–83, 2002.

D.M. Debinski and R.D. Holt, "A survey and overview of habitat fragmentation," *Conservation Biology*, vol. 14, pp. 342–355, 2000.

G. Cornuéjols and A. Sassano, "On the 0, 1 facets of the set covering polytope," *Mathematical Programming*, vol. 43, pp. 45–55, 1989.

P.L. Hammer, E.L. Johnson, and U.N. Peled, "Facets of regular 0-1 polytopes," *Mathematical Programming*, vol. 8, pp. 179–206, 1975.

M.D. McDonnell, H.P. Possingham, I.R. Ball, and E.A. Cousins, "Mathematical methods for spatially cohesive reserve design," *Environmental Modeling and Assessment*, vol. 7, pp. 107–114, 2002.

D.J. Nalle, J.L. Arthur, and J. Sessions, "Designing compact and contiguous reserve networks with a hybrid heuristic approach," *Forest Science*, vol. 48, pp. 59–68, 2002.

A.O. Nicholls and C.R. Margules, "An upgraded reserve selection algorithm," *Biological Conservation*, vol. 64, pp. 165–169, 1993.

P. Nobili and A. Sassano, "Facets and lifting procedures for the set covering polytope," *Mathematical Programming*, vol. 45, pp. 111–137, 1989.

C. Van Nuffelen, "On the rank of the incidence matrix of a graph," *Cahiers du Centre d' Etudes de Recherche Opérationnelle*, vol. 15, pp. 363–365, 1973.

H. Önal and R.A. Briers, "Incorporating spatial criteria in optimum reserve network selection," in *Proceedings of the Royal Society*, *London B*, 2002, vol. 269, pp. 2437–2441.

H. Önal and R.A. Briers, "Selection of a minimum-boundary reserve network using integer programming," in *Proceedings of the Royal Society*, *London B*, 2003, vol. 270, pp. 1487–1491.

H. Possingham, I. Ball, and S. Andelman, "Mathematical methods for identifying representative reserve networks," in *Quantitative Methods for Conservation Biology*, S. Ferson and M. Burgman (Eds.), Springer-Verlag: New York, 2000, pp. 291–306.

R. Primack, *Essentials of Conservation Biology*, Sinauer Associates, Sunderland, 2002.

M. Sánchez-García, M.I. Sobrón, and B. Victoriano, "On the set covering polytope: Facets with coefficients in {0, 1, 2, 3}," *Annals of Operations Research*, vol. 81, pp. 343–356, 1998.

H. Sachs, "Über teiler, faktoren und charakteristische polynome von graphen. II," *Wiss Z. Tech. Hochsch Ilmenau*, vol. 13, pp. 405–412, 1967.

A. Sassano, "On the facial structure of the set covering polytope," *Mathematical Programming*, vol. 44, pp. 181–202, 1989.

P. Siitonen, A. Tanskanen, and A. Lehtinen, "Method for selection of old-forest reserves," *Conservation Biology*, vol. 16, pp. 1398–1408, 2002.

P. Siitonen, A. Tanskanen, and A. Lehtinen, "Selecting forest reserves with a multiobjective spatial algorithm," *Environmental Science & Policy*, vol. 6, pp. 301–309, 2003.

R.R. Vemuganti, "Applications of the set covering, set packing and partitioning models: A survey," in *Handbook of Combinatorial Optimization*, vol. 1, D.-Z. Du and P. Pardalos (Eds.), Kluwer Academic Publishers: Boston, 1998, pp. 573–746.

J.C. Williams, "Delineating protected wildlife corridors with multi-objective programming," *Environmental Modeling and Assessment*, vol. 3, pp. 77–86, 1998.

L. Wolsey, "Faces for a linear inequality in 0-1 variables," *Mathematical Programming*, vol. 8, pp. 165–178, 1975.