



**HAL**  
open science

# Polyhedral combinatorics of the $K$ -partitioning problem with representative variables

Zacharie Alès, Arnaud Knippel, Alexandre Pauchet

► **To cite this version:**

Zacharie Alès, Arnaud Knippel, Alexandre Pauchet. Polyhedral combinatorics of the  $K$ -partitioning problem with representative variables. *Discrete Applied Mathematics*, 2016, 211, pp.1 - 14. 10.1016/j.dam.2016.04.002 . hal-01759687

**HAL Id: hal-01759687**

**<https://hal.science/hal-01759687v1>**

Submitted on 29 Jun 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Polyhedral combinatorics of the $K$ -partitioning problem with representative variables

Zacharie ALES<sup>†,‡</sup>, Arnaud KNIPPEL<sup>†</sup>, Alexandre PAUCHET<sup>‡</sup>

<sup>†</sup>: LMI INSA Rouen (EA 3226) <sup>‡</sup>: LITIS INSA Rouen (EA 4051)

## Abstract

The  $K$ -partitioning problem consists in partitioning the vertices of a weighted graph in  $K$  sets in order to minimize a function related to the edge weights. We introduce a linear mixed integer formulation with edge variables and representative variables. We investigate the polyhedral combinatorics of the problem, study several families of facet-defining inequalities and evaluate their efficiency on the linear relaxation.

**Keyword:** combinatorial optimization, polyhedral approach, graph partitioning

## 1 Introduction

Graph partitioning consists in splitting the vertices of a graph in several sets called *clusters*, so that a given function of the edge weights is minimized or maximized. In many papers this function is linear hence minimizing the weight of the edges in the clusters is equivalent to maximizing the total weight of the multicut (*i.e.*, the total weight of the edges between two different clusters). For this reason different names are used in literature and the most frequent are *graph partitioning* problem [1, 2, 3] and *min-cut* problem [4]. The *max-cut* problem obtained when maximizing the cut of a graph with positive weights on its edges is known to be NP-complete [5]. The problem is sometimes called *clique partitioning* problem when the graph is complete [6, 7]. When dealing with general graphs, Chopra and Rao [1] note that additional edges can be considered to obtain a complete graph. For the sake of simplicity we adopt this point of view and only consider complete graphs. Specific valid inequalities such as the *star inequalities* or the *cycle with ear inequalities* can however be considered when dealing with sparse graphs [3].

Graph partitioning has many applications (*e.g.*, image segmentation or VLSI design [2]) and many variants, which in most cases are NP-hard [5]. The number of sets in a solution may be specified as a part of the problem definition [8, 9] or not [10, 11]. In this paper we consider the former case that we call the *K-partitioning* problem, where  $K$  is the number of sets.

Our motivation comes from a clustering problem for dialogues analysis in psychology [12]. Dialogues can be encoded using two-dimensional tables (or series of item-sets), among which dialogue patterns, representative of human behaviors, are repeated approximately. Partitioning a graph of dialogue patterns would enable to group similar instances and therefore to characterize significant behaviors. For this application, instances are in general complete graphs of 20 to 100 vertices that have to be partitioned in 6 to 10 sets. Even if big instances require approximate solutions, we are interested in improving the exact methods based on branch and bound (such as in [13, 14, 15]), by studying more precisely the polyhedral structure of a linear formulation in order to provide better bounds.

A general linear integer formulation using edge variables was proposed in [6], together with several facet-defining inequalities. We call this formulation *edge formulation* or *node-node formulation*, and it has experimentally proved to be stronger than the *node-cluster formulation* [15], although this may depend on the data sets. Contrary to the node-cluster formulation, the edge formulation however doesn't allow to fix the number of clusters easily. Some authors use a formulation with both edge variables and node-cluster variables. A formulation with an exponential number of constraints has been proposed in [16] – when a bound on the size of the clusters is considered – and applied to sparse graphs. More compact formulations have been proposed based on the linearization of the quadratic formulation [17, 13]. In [17] the weights of the edges are positive and the total edge weight is minimized, while in [13] it is maximized. When considering the triangle inequalities from both formulations [1], graph with arbitrary weights on the edges can be partitioned.

All these formulations have a common drawback: they contain a lot of symmetry (*i.e.*, equivalent solutions can be obtained by permuting variables indices) and this can considerably slow down methods based on branch-and-bound or branch-and-cut. One approach consists in dealing with the symmetry in the branching strategies. To this end, Kaibel et al [17] have proposed a general tool, called orbitopal fixing. Another way is to break the symmetry directly in the formulation. This approach has already been used in [18] for the vertex coloring problem. More recently a similar idea has been applied in the node-cluster formulation [15]. In this paper we

propose a formulation based on both edge variables of [6, 19] and vertex variables that we call *representative variables*. This allows not only to break the symmetry, but also to fix the number of clusters to  $K$ .

A side effect of breaking the symmetry is the complication of the polyhedral study. To simplify the presentation of the paper, the most technical proofs are skipped. In this paper we only include complete proofs related to the general clique inequalities and inequalities derived from our new variables (the strengthened triangle inequalities and the paw inequalities). For further details, the reader may refer to our technical report [20].

In Section 2, we introduce notations and a possible integer linear programming formulation of the problem. In Section 3, we study the dimension of the polyhedron  $P_{n,K}$  associated to our formulation. We characterize, in Section 4 all the facet-defining inequalities from our formulation. In Subsection 4.6, we strengthen the triangle inequalities that do not define facets. In Section 5, we study four families of inequalities (namely the 2-chorded cycle inequalities, the 2-partition inequalities, the general clique inequalities and the paw inequalities) and we determine cases for which they are facet defining of  $P_{n,K}$ . In the last section we illustrate the improvement on the linear relaxation value of our formulation for the facet-defining inequalities of the previous sections, on complete graphs with different kind of weights.

## 2 Notations and problem statement

Let  $V = \{1, \dots, n\}$  be a set of indexed vertices and  $G = (V, E)$  the complete graph induced by  $V$ . A  $K$ -partition  $\pi$  is a collection of  $K$  non-empty subsets  $C_1, C_2, \dots, C_K$ , called clusters, such that  $\forall i \neq j, C_i \cap C_j = \emptyset$  and  $\bigcup_{i=1}^K C_i = V$ . The aim of the  $K$ -partitioning problem is to find a  $K$ -partition which minimizes the total weight of the partition.

To each  $K$ -partition  $\pi$ , we associate a characteristic vector  $x^\pi \in \{0, 1\}^{|E|+|V|}$  such that:

- for each edge  $uv \in E$ ,  $x_{uv}^\pi$  (equivalent to  $x_{vu}^\pi$ ) is equal to 1 if  $u, v \in C_i$  for some vertex  $i$  in  $\{1, 2, \dots, K\}$  and 0 otherwise;
- for each vertex  $u \in V$ ,  $x_u^\pi = 1$  if  $u$  is the vertex with the smallest index of its cluster (in that case  $u$  is said to be the *representative* of its cluster) and 0 otherwise.

An edge  $uv \in E$  is said to be *activated* for a given partition  $\pi$  if  $x_{uv}^\pi = 1$ . In this context, vertex  $u$  is said to be *linked* to vertex  $v$  and vice versa. A vertex  $i$  is said to be *lower* than another vertex  $j$  (noted  $i < j$ ) if index  $i$  is lower than index  $j$ .

Let  $d_{ij}$  denote the cost of edge  $ij \in E$ . We consider the following formulation for the  $K$ -partitioning problem:

$$(P_{er}) \left\{ \begin{array}{ll} \min \sum_{ij \in E} d_{ij} x_{ij} & \\ x_{ik} + x_{jk} - x_{ij} \leq 1 & \forall i, j, k \in V, i \neq k, j \neq k, i < j \quad (1) \\ x_j + x_{ij} \leq 1 & \forall i, j \in V, i < j \quad (2) \\ x_j + \sum_{i=1}^{j-1} x_{ij} \geq 1 & \forall j \in V \quad (3) \\ \sum_{i=1}^n x_i = K & (4) \\ x_{ij} \in \{0, 1\} & ij \in E \quad (5) \\ x_i \in [0, 1] & i \in V \quad (6) \end{array} \right.$$

Constraints (1), called the *triangle inequalities*, ensure that if two incident edges  $ij$  and  $jk$  are activated then  $ik$  is also activated. Note that there are  $\frac{n(n-1)(n-2)}{2}$  many constraints (three for each triangle  $a, b, c$  of  $G = (V, E)$ ). Constraints (2), called the *upper representative inequalities*, ensure that every cluster contains no more than one representative. If vertex  $j$  is a representative then it is not linked to any lower vertex  $i$ . If vertex  $j$  is linked to such a lower vertex then it is not a representative. Constraints (3), called the *lower representative inequalities*, guarantee that a cluster contains at least one representative. Indeed, on one hand if vertex  $j$  is not a representative then it is linked to at least one lower vertex, on the other hand if vertex  $j$  is not linked to any of these vertices then it is a representative. Finally, constraint (4) ensures that the number of clusters is equal to  $K$ .

Note that in the above formulation the representative variables can be relaxed, as fixing all edge variables to 0 or 1 forces the representative variables to be in  $\{0, 1\}$  thanks to Equations (2) and (3). So we only have  $|E|$  binary variables.

As the polyhedron associated to the above formulation is not full-dimensional, we fix  $x_1$  to 1 (since vertex 1 is always a representative), substitute  $x_2$  by  $1 - x_{1,2}$  (since vertex 2 is a representative if and only if it is not in the same cluster than vertex 1) and  $x_3$  by  $K - 2 + x_{1,2} - \sum_{i=4}^n x_i$  using Equation (4). Therefore, the characteristic vector of a partition  $\pi$  now contains the  $n - 3$  remaining representative components followed by the  $|E|$  edges components:

$$(x^\pi)^T = (x_4, \dots, x_n, x_{1,2}, \dots, x_{1,n}, x_{2,3}, \dots, x_{n-1,n}).$$

In the following, we use the artificial variables  $x_1$ ,  $x_2$  and  $x_3$  instead of their substituted expression to simplify the notations. For a given hyperplane  $H = \{x \in \mathbb{R}^{|E|+|V|-3} \mid \alpha^T x = \alpha_0\}$  the coefficients  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  related to the three artificial variables may be mentioned in subsequent proofs. Since these coefficients are associated to artificial variables, they are equal to 0.

Let  $P_{n,K}$  be the convex hull of all integer points which are feasible for  $(P_{er})$ :

$$P_{n,K} = \text{conv}\{x \in \{0, 1\}^{|E|+|V|-3} \mid x \text{ satisfies (1), (2), (3), (4), (5), (6)}\}.$$

To simplify the notations, a singleton  $\{s\}$  may be denoted by  $s$ . Likewise for a given vector  $\alpha \in \mathbb{R}^{|E|+|V|-3}$  and a subset  $E_1$  of  $E$ , the term  $\alpha(E_1)$  is used to denote the sum of the  $\alpha$  components in  $E_1$  ( $\sum_{e \in E_1} \alpha_e$ ). Finally, if we consider two subsets of  $V$ ,  $V_1$  and  $V_2$ , the sum of the  $\alpha$  inter-set components  $\sum_{i \in V_1} \sum_{j \in V_2} \alpha_{ij}$  and the sum of the  $\alpha$  intra-set components  $\sum_{i,j \in V_1, i < j} \alpha_{ij}$  are respectively denoted by  $\alpha(V_1, V_2)$  and  $\alpha(V_1)$ .

## 2.1 Transformations

To study the dimension of a polytope  $P \subset \mathbb{R}^{|E|+|V|-3}$  (either  $P_{n,K}$  or one of its face), we identify the number of affinely independant hyperplanes  $H = \{x \in \mathbb{R}^{|E|+|V|-3} \mid \alpha^T x = \alpha_0\}$  which include  $P$ . In order to obtain relations between the coefficients of  $H$ , we successively consider pairs of valid  $K$ -partitions  $\pi^1$  and  $\pi^2$  in  $P$  and use the fact that

$$\alpha^T x^{\pi^1} = \alpha^T x^{\pi^2}. \quad (7)$$

To quickly identify the relation obtained after simplification of (7), we define an operator  $\mathcal{T}$  hereafter called *transformation*. Given two disjoint clusters  $C_1, C_2$  and a set of vertices  $R \subset C_1 \cup C_2$  we define  $\mathcal{T} : \{C_1, C_2, R\} \mapsto \{C'_1, C'_2\}$ , with  $C'_1 = (C_1 \setminus R) \cup (R \setminus C_1)$  and  $C'_2 = (C_2 \setminus R) \cup (R \setminus C_2)$ . The corresponding transformation is presented in Figure 1.

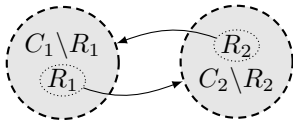


Figure 1: Representation of  $\mathcal{T}(C_1, C_2, R)$  with  $R_1 = R \cap C_1$  and  $R_2 = R \cap C_2$ .

A transformation  $\mathcal{T} : \{C_1, C_2, R\} \mapsto \{C'_1, C'_2\}$  is said to be *valid* for the polytope  $P$  if there exists two  $K$ -partitions  $\pi = \{C_1, C_2, C_3, \dots, C_K\}$  and  $\pi' = \{C'_1, C'_2, C_3, \dots, C_K\}$  such that  $x^\pi$  and  $x^{\pi'}$  are included in  $P$ .

## 3 Dimension of $P_{n,K}$

**Theorem 3.1.** *Depending on  $K$ , the dimension of  $P_{n,K}$  is:*

- (i)  $\dim(P_{n,2}) = |E| + n - 4$ ;
- (ii)  $\dim(P_{n,K}) = |E| + n - 3$ , for  $K \in \{3, 4, \dots, n - 2\}$  (i.e., it is full dimensional);
- (iii)  $\dim(P_{n,n-1}) = |E| - 1$ .

*Proof.* The complete proof of this theorem is available in [20, 21]. Here we present an alternative proof of case (ii).

We assume that  $P_{n,K}$  is included in a hyperplane  $H = \{x \in \mathbb{R}^{|E|+|V|-3} \mid \alpha^T x = \alpha_0\}$  and prove that all its coefficients are equal to 0.

Since  $K$  is in  $\{3, \dots, n-2\}$ , a transformation  $\mathcal{T}(C_1, C_2, R) \mapsto \{C'_1, C'_2\}$  is valid for  $P_{n,K}$  if  $C_1, C_2, C'_1$  and  $C'_2$  are non empty (otherwise less than  $K$  clusters are obtained) and if  $|C_1 \cup C_2| \leq 4$  (otherwise no partition with  $K$  clusters can be obtained if  $K$  is equal to  $n-2$ ).

Let  $i$  be a vertex of  $V \setminus \{1, 2, 3\}$  and let  $a$  and  $b$  be two vertices lower than  $i$ . Transformation  $\mathcal{T}(\{a, i\}, b, i)$  represented in Figure 2 gives:

$$\alpha_{ai} = \alpha_{bi}. \quad (8)$$

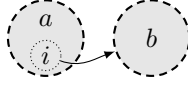


Figure 2:  $\mathcal{T}(\{a, i\}, b, i)$

Indeed,  $x_{ai}$  is equal to 1 only before the transformation and  $x_{bi}$  becomes equal to 1 after the transformation. No representative variable appears in (8), since vertices  $a$  and  $b$  remain the representative of their respective clusters during the transformation. Equation (8) shows that for each vertex  $i \in V \setminus \{1, 2, 3\}$  the value of  $\alpha_{hi}$  for all vertices  $h \in \{1, \dots, i-1\}$  is equal to a constant that we denote by  $\beta_i$ .

The transformation represented in Figure 3 leads to  $\beta_i = 0 \forall i \geq 4$ . Transformation  $\mathcal{T}(\{1, i\}, 2, 1)$  (see Figure 4) is used to prove that

$$0 = \alpha_i + \alpha_{1,2}. \quad (9)$$

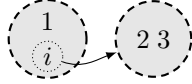


Figure 3:  $\mathcal{T}(\{1, i\}, \{2, 3\}, i)$

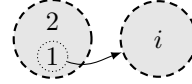


Figure 4:  $\mathcal{T}(\{1, 2\}, i, 1)$

Vertex  $i$  becomes a representative after the transformation, thus it appears on the right-hand side of (9). Vertex 2 is the lowest vertex of its cluster only before the transformation but it does not appear on the left-hand side of (9), since  $x_2$  is an artificial variable (*i.e.*, it is not associated to a coefficient in  $H$ ).

In the previous transformation,  $\{1, 2\}$  can be replaced by  $\{1, 3\}$  and  $\{2, 3\}$  to show that  $\alpha_{1,2}$ ,  $\alpha_{1,3}$  and  $\alpha_{2,3}$  are all equal to  $-\alpha_i$ .

Finally, the transformation represented in Figure 5 proves that  $\alpha_i$  is equal to 0.

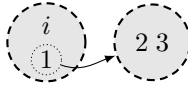


Figure 5:  $\mathcal{T}(\{1, i\}, \{2, 3\}, 1)$

□

Thereafter, we study facet-defining inequalities for  $P_{n,K}$  when it is full-dimensional (*i.e.*,  $K \in \{3, \dots, n-2\}$ ). For each studied face  $F = \{x \in P_{n,K} \mid \omega^T x = \omega_0\}$ , we consider a facet-defining inequality  $\alpha^T x \leq \alpha_0$  such that  $F \subseteq \{x \in P_{n,K} \mid \alpha^T x = \alpha_0\}$ . We then prove that  $F$  is facet-defining by highlighting, with reference to Theorem 3.6 in Section I.4.3 of [22], that  $(\alpha, \alpha_0)$  is proportional to  $(\omega, \omega_0)$ .

As mentioned in the introduction we refer to [20] for the most technical proofs.

## 4 Trivial inequalities

In this section, we show which of the inequalities from the integer formulation ( $P_{er}$ ) are facet-defining. We restrict our study to the general cases where  $P_{n,K}$  is full-dimensional (*i.e.*,  $K \in \{3, 4, \dots, n-2\}$ ).

## 4.1 The edge bound inequalities

**Remark** The inequalities  $x_{uv} \leq 1$  for all  $uv \in E$  are not facet-defining since they are induced by the two following inequalities:  $x_{uv} + x_{ui} - x_{vi} \leq 1$  and  $x_{uv} + x_{vi} - x_{ui} \leq 1$ , for all vertices  $i \in V \setminus \{u, v\}$ .

**Theorem 4.1.** *If  $P_{n,K}$  is full-dimensional,  $x_{uv} \geq 0$  is facet-defining if and only if  $uv \notin \{12, 13, 23\}$ .*

## 4.2 The representative bound inequalities

**Remark** The inequalities  $x_v \leq 1$ , for all vertices  $v \in \{4, 5, \dots, n\}$  are not facet-defining since the face induced by  $x_v = 1$  is contained in the hyperplanes  $\{x \in \mathbb{R}^{|E|+|V|-3} \mid x_{u,v} = 0\}$  for all vertices  $u \in \{1, 2, \dots, v-1\}$ . Indeed, if vertex  $v$  is the representative of a cluster  $C$ , it must be the lowest vertex of  $C$ . The dimension of the face induced by  $x_v \leq 1$  is consequently lower than or equal to  $\dim(P_{n,K}) - v + 1$ .

**Remark** For the same reason, the inequality  $\sum_{i=4}^n x_i - x_{1,2} \geq K - 3$  which corresponds to  $x_3 \leq 1$ , is not facet-defining. Neither is  $x_3 \geq 0$  since in that case  $x_{1,3} + x_{2,3} - x_{1,2} = 1$  (i.e., vertex 3 is not a representative so it is in the same cluster than 1 or 2).

**Theorem 4.2.** *If  $P_{n,K}$  is full-dimensional, the inequalities  $x_v \geq 0$ , for all vertices  $v \in \{4, 5, \dots, n\}$ , are facet-defining if and only if  $K \neq n - 2$ .*

## 4.3 The upper representative inequalities

**Theorem 4.3.** *If  $P_{n,K}$  is full-dimensional, then the inequalities  $x_{u,v} + x_v \leq 1$  for all vertices  $v \geq 4$  and  $u < v$  are facet-defining if and only if  $n \geq 6$  or  $\{u, v\} \neq \{4, 5\}$ .*

**Theorem 4.4.** *If  $P_{n,K}$  is full-dimensional, then the inequalities  $x_{1,2} + x_{a,3} - \sum_{i=4}^n x_i \leq 3 - K$  for vertex  $a \in \{1, 2\}$  - which correspond to  $x_{a,3} + x_3 \leq 1$  - are facet-defining.*

## 4.4 The lower representative inequalities

**Theorem 4.5.** *If  $P_{n,K}$  is full-dimensional, then the inequalities  $x_u + \sum_{i=1}^{u-1} x_{iu} \geq 1$  for all vertices  $u \geq 4$  are facet-defining.*

**Theorem 4.6.** *If  $P_{n,K}$  is full-dimensional, then the inequality  $x_{1,2} + x_{1,3} + x_{2,3} - \sum_{i=4}^n x_i \geq 3 - K$  for vertex  $a \in \{1, 2\}$  - which corresponds to  $x_3 + x_{1,3} + x_{2,3} \leq 1$  - is facet-defining.*

## 4.5 The triangle inequalities

**Theorem 4.7.** *If  $P_{n,K}$  is full-dimensional, then the inequality  $x_{ik} + x_{jk} - x_{ij} \leq 1$  for vertices  $i, j, k$  distinct in  $V$  is facet-defining if and only if the following conditions are satisfied*

- (i)  $k < i$  or  $k < j$ ;
- (ii)  $\{i, j, k\} \neq \{1, 2, 3\}$ .

## 4.6 The strengthened triangle inequalities

Theorem 4.7 states that inequalities (1) are not facet-defining if  $k$  is greater than both  $i$  and  $j$ . However, they can be strengthened by adding the term  $x_k$  to the left side of the inequality whenever  $k$  is greater than three (otherwise  $x_k$  is an artificial variable):

$$x_{ik} + x_{jk} - x_{ij} + x_k \leq 1. \quad (2')$$

For three distinct vertices  $i, j$  and  $k$ , let  $P_{i,j,k}$  be the face of  $P_{n,K}$  defined by Equation (2').

**Theorem 4.8.** *Let  $i, j$  and  $k$  be three vertices in  $V$  such that  $i < j < k$  and  $k > 3$ . When  $P_{n,K}$  is full-dimensional, the inequality  $x_{ik} + x_{jk} - x_{ij} + x_k \leq 1$  is facet-defining if and only if  $(j > 3)$  or  $(K \leq n - 3)$ .*

*Proof.* Assume that  $j \leq 3$  and  $K = n - 2$ . Let  $\pi$  be a  $K$ -partition such that the three first vertices are in the same cluster. As  $K$  is equal to  $n - 2$  the  $K - 1$  other clusters are necessarily reduced to one vertex. Hence the left part of equation (2') is equal to zero and  $\pi$  is not in  $P_{i,j,k}$ . Since 1, 2 and 3 cannot be together the equation

$$\sum_{i=4}^n x_i - x_{1,2} - x_{1,3} - x_{2,3} = K - 3. \quad (10)$$

is always satisfied. Thus,  $P_{i,j,k}$  is included in the hyperplane induced by (10).

Let  $U = \{u_1, u_2, \dots, u_{|U|}\}$  be  $V \setminus \{i, j, k\}$  such that  $u_1 < u_2 < \dots < u_{|U|}$  and let  $H = \{x \in \mathbb{R}^{|E|+|V|-3} \mid \alpha^T x = \alpha_0\}$  be a hyperplane which includes  $P_{i,j,k}$ . Let  $\pi$  be a  $K$ -partition such that vertex  $k$  is either a representative of its cluster or it is in the same cluster than vertices  $i$  or  $j$ . The vector  $x^\pi$  associated to such a  $K$ -partition is in  $P_{i,j,k}$ . This remark ensures that each transformation considered throughout this proof is valid for  $P_{i,j,k}$ .

We first consider two cases (*i.e.*,  $j > 3$  and  $j \leq 3$ ) and prove that in both of them  $\alpha_{ib}$  and  $\alpha_{jb}$  are equal to 0 for all vertices  $b$  in  $U$ . Let  $m$  and  $a$  be two vertices respectively in  $\{i, j\}$  and  $U \setminus \{b\}$ . If  $j$  is greater than 3 we additionally assume that  $a \leq 3$ . In this case, the transformations  $\mathcal{T}(\{m, k\}, a, m)$  and  $\mathcal{T}(\{m, k\}, \{a, b\}, m)$  give the result by respectively leading to:  $\alpha_m + \alpha_{km} = \alpha_k + \alpha_{am}$  and  $\alpha_m + \alpha_{km} = \alpha_k + \alpha_{am} + \alpha_{bm}$ . In the other case, the theorem hypothesis ensures that  $K$  is strictly lower than  $n - 2$  and the same result is obtained using the two following transformations:  $\mathcal{T}(\{i, j, k\}, a, m)$  and  $\mathcal{T}(\{b, i, j, k\}, a, m)$ .

The transformations  $\mathcal{T}(\{j, k\}, \{a, i\}, a)$  and  $\mathcal{T}(\{i, k\}, \{a, j\}, a)$  give  $\alpha_i = \alpha_j$ . If  $i$  is strictly lower than 4,  $x_i$  is an artificial variable,  $\alpha_i$  and  $\alpha_j$  are then equal to 0. If  $j \geq 4$  the same can be proved through  $\mathcal{T}(\{u_1, u_3\}, u_2, u_3)$ ,  $\mathcal{T}(\{i, u_1, u_2\}, u_3, \{u_1, u_2\})$  and  $\mathcal{T}(\{u_1, u_2\}, i, u_2)$ .

The value of the remaining coefficients of  $H$  is obtained using the transformations presented in table 1.

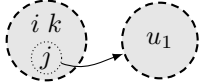
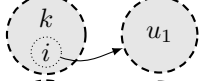
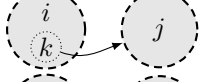
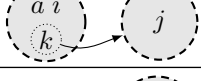



Conditions	Transformation	Results
-		$\alpha_{ik} = -\alpha_{ij}$
-		$\alpha_{ik} = \alpha_k$
-		$\alpha_{ik} = \alpha_{jk}$
-		$\alpha_{ak} = 0$
$\forall b \in U \setminus u_1$ $\forall a \in V \setminus \{b, k\} \ a < b$		$\alpha_{u_1 b} = \alpha_b$
-		$\alpha_{ab} = \alpha_b$
-		$\alpha_b = 0$

Table 1: Transformations used in Theorem 4.8. Each line presents a step of the proof. The last column corresponds to the result.

□

## 5 Valid inequalities

In this section we consider several classes of inequalities, looking for facets. Except for the *paw inequalities*, which contain our representative variables, the other classes of inequalities have already been studied for other

variants of graph partitioning problems. Due to technicality, we skip the proofs in the two first subsections but the reader may refer to [20].

## 5.1 The 2–chorded cycle inequalities

In this subsection, we address the 2–chorded cycle class of inequalities, first introduced in [6]. Let  $C = \{e_1, \dots, e_{|C|}\}$  be a cycle in  $E$  so that  $e_i = c_i c_{i+1}$  for all vertices  $i$  in  $\{1, 2, \dots, |C| - 1\}$  and  $e_{|C|} = c_1 c_{|C|}$ . Let  $V_C \stackrel{\text{def}}{=} \{c_1, c_2, \dots, c_{|C|}\}$  and  $U \stackrel{\text{def}}{=} V \setminus V_C$ . The set of 2–chords of  $C$  is defined as  $\overline{C} = \{c_i c_{i+2} \in E \mid i = 1, \dots, |C| - 2\} \cup \{c_1 c_{|C|-1}, c_2 c_{|C|}\}$ . The 2–chorded cycle inequality induced by a given cycle  $C$  of length at least 5 and its corresponding  $\overline{C}$  is defined as

$$x(E(C)) - x(E(\overline{C})) \leq \lfloor \frac{1}{2} |C| \rfloor. \quad (11)$$

We skip the proof of the following lemma. The reader can refer to [6] for further details.

**Lemma 5.1.** *The 2–chorded cycle inequality (11) induced by a cycle  $C$  of length at least 5 is valid for  $P_{n,K}$ . The corresponding face  $F_C$  is not facet-defining if  $|C|$  is even.*

**Theorem 5.2.** *The face  $F_C$  induced by an odd cycle  $C$  of size  $2p+1$  is facet-defining if the following conditions are satisfied:*

- (i)  $P_{n,K}$  is full-dimensional;
- (ii)  $|U \cap \{1, 2, 3\}| \geq 2$ ;
- (iii)  $2 \leq p \leq n - K - |U \cap \{1, 2, 3\}|$ ;
- (iv)  $K \geq 4$ .

## 5.2 The 2-Partition inequalities

This section is dedicated to the study of the 2-partition inequalities, first introduced in [6] for the general clique partitioning problem. For two disjoint nonempty subsets  $S$  and  $T$  of  $V$  are defined as

$$x(E(S), E(T)) - x(E(S)) - x(E(T)) \leq \min(|S|, |T|). \quad (12)$$

Let  $F_{S,T}$  be the face of  $P_{n,K}$  defined by Equation 12.

The proof of the following lemma is skipped, for further details the reader may refer to [6].

**Lemma 5.3.** *Inequality (12) is valid for  $P_{n,K}$ .*

**Theorem 5.4.** *If  $P_{n,K}$  is full-dimensional, the 2-partition inequality (12) is facet-defining for two non empty disjoint subsets  $S$  and  $T$  of  $V$  if and only if the following conditions are satisfied:*

- (i)  $|T| - |S| \in \{1, 2, \dots, K - 1\}$ ;
- (ii)  $|S| \leq n - (K + 2)$ ;
- (iii)  $\forall s \in S \exists t \in T, t > s$ ;
- (iv) if  $|S| = 1 \exists u \in U \cap \{1, 2, 3\}$ .



### 5.3 The general clique inequalities

The *clique inequalities* have been introduced by Chopra and Rao [1] and correspond to the fact that for any  $m$ -partition  $\pi$  ( $m \leq K$ ) and any set  $Z \subset V$  of size  $K + 1$ , at least two vertices of  $Z$  are necessarily in the same cluster. The clique inequality induced by a given set  $Z$  is:

$$x(E(Z)) \geq 1. \quad (13)$$

The general clique inequalities are obtained by increasing both the size of  $Z$  and the value of the right-hand side of Equation 13. Thus, the general clique inequalities induced by a set  $Z \subset V$  of size  $qK + r$  with  $q \in \mathbb{N}$  and  $r \in \{0, 1, \dots, K - 1\}$  is defined by Chopra and Rao as:

$$x(E(Z)) \geq \binom{q+1}{2} r + \binom{q}{2} (K - r). \quad (14)$$

Let  $P_Z$  be the face of  $P_{n,K}$  defined by Equation (14). As represented in Figure 6, the lower bound of inequality (14) corresponds to the minimal value of  $x(E(Z))$  in the incidence vector of a  $K$ -partition – thus ensuring the validity of this inequality. It is obtained by setting  $q + 1$  vertices of  $Z$  in each of the  $r$  first clusters and  $q$  vertices in each of the  $K - r$  remaining clusters.

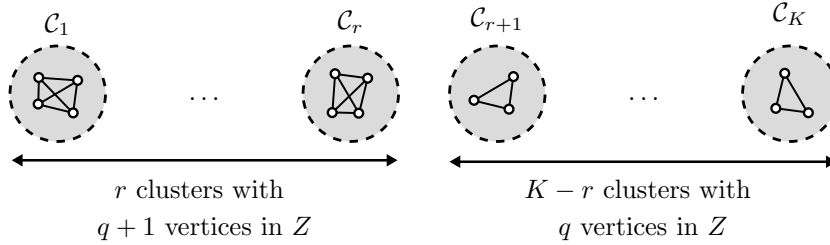


Figure 6: Distribution of  $Z$  vertices in a  $K$ -partition included in  $F_Z$  (case where  $q$  is equal to three).

These inequalities have also been studied by Labbé and Öszozy [19] in the case where the clusters must contain at least  $F_L$  vertices. In this context, the size of  $Z$  must be greater than or equal to  $\lfloor \frac{n}{F_L} \rfloor$ . Finally, Ji and Mitchell also studied these inequalities that they called the *pigeon inequalities* [23].

In the following  $U = \{u_1, u_2, \dots, u_{|U|}\}$  is used to denote  $V \setminus Z$  such that  $u_1 < u_2 < \dots < u_{|U|}$ . The vertices in  $Z = \{z_1, \dots, z_{K+1}\}$  are similarly sorted.

**Theorem 5.5.** *If  $P_{n,K}$  is full-dimensional, for a given  $Z \subset V$  of size  $K+1$ , then inequality (14) is facet-defining if and only if:*

- (i)  $|U| \geq 1$  and  $u_1 \leq 3$ ;
- (ii)  $z_{|Z|} = n$
- (iii)  $|Z| \in \{K + 1, \dots, 2K - 1\}$ .

**Lemma 5.6.** *Let  $V_1$  and  $V_2$  be two disjoint subsets of  $V$  and let  $Z$  be a subset of  $V$  which satisfies the conditions of Theorem 5.5. Then, there exists a  $K$ -partition in  $F_Z$  which includes  $V_1$  and  $V_2$  if  $\{|V_1 \cap Z|, |V_2 \cap Z|\}$  is equal to  $\{1, 2\}$ .*

*Proof.* Given the bounds on the size of  $Z$  in theorem 5.5,  $q$  is necessarily equal to one and  $r \in \{1, \dots, K - 1\}$ . Consequently, each  $K$ -partition included in  $F_Z$  contains at least one cluster with exactly one vertex in  $Z$  and at least one cluster with exactly two vertices in  $Z$ .

Let  $\pi = \{C_1, \dots, C_K\}$  be the  $K$ -partition such that:

- $C_1 = V_1$  and  $C_2 = V_2$ .
- Cluster  $C_i$ ,  $i \in \{3, \dots, r + 1\}$ , contains  $q + 1$  vertices from  $Z$ .
- Cluster  $C_j$ ,  $j \in \{r + 2, \dots, K\}$ , contains  $q$  vertices from  $Z$ .

- The vertices in  $U$  which are not included in  $V_1$  or  $V_2$  are in  $C_K$ .

This construction is always possible since  $|Z|$  is equal to  $qK + r$ . It can be easily checked – by computing  $x^\pi(Z)$  – that  $\pi$  is in  $F_Z$ .  $\square$

Each transformation  $\mathcal{T}(C_1, C_2, R) \mapsto \{C'_1, C'_2\}$ , considered in the proof of Theorem 5.5 is such that the couples  $(C_1, C_2)$  and  $(C'_1, C'_2)$  satisfy the conditions imposed on  $V_1$  and  $V_2$  in Lemma 5.6. This ensures the validity of the transformations. We now present the proof of Theorem 5.5.

*Proof.* If the first condition of the theorem is not satisfied, the three first vertices are in  $Z$  and cannot be in the same cluster. Consequently  $P_Z$  is included in the hyperplane defined by  $\sum_{i=4}^n x_i - x_{1,2} - x_{1,3} - x_{2,3} = K - 3$ . If (ii) is false, the vertices  $u$  which are greater than  $z_{K+1}$  cannot be representative since each cluster contains at least one element of  $Z$ . Thus,  $P_Z$  is included in the hyperplanes:  $x_u = 0 \forall u > z_{|Z|}$ . Eventually, if  $Z$  contains more than  $2K - 1$  vertices, each cluster necessarily include at least two vertices from  $Z$ . Thus,  $z_{|Z|}$  cannot be a representative and  $F_Z$  is included in the hyperplane induced by  $x_{z_{|Z|}} = 0$ .

Let  $H = \{x \in \mathbb{R}^{|E|+|V|-3} \mid \alpha^T x = \alpha_0\}$  be a hyperplane which includes  $P_Z$  and let  $z_i < z_j < z_k$  be three elements of  $Z$ . The transformation  $\mathcal{T}(\{z_i, z_k\}, \{z_j\}, \{z_k\})$ , first shows:  $\alpha_{z_i z_k} = \alpha_{z_j z_k}$ . Thus, for a given vertex  $k$  and for all vertices  $j \in \{1, \dots, k-1\}$  the coefficients  $\alpha_{z_j z_k}$  are equal to a constant, referred to as  $\beta_k$ .

For all  $j$  and  $k$  greater than  $i$ ,  $\mathcal{T}(\{z_i, z_k\}, \{z_j\}, \{z_i\})$  gives

$$\beta_k - z_k = \beta_j - z_j. \quad (15)$$

Let  $z, z'$  and  $z''$  be three distinct vertices of  $Z$ . The transformation  $\mathcal{T}(\{u_1, z\}, \{z'\}, \{u_1\})$  leads to  $\alpha_{z'} + \alpha_{u_1 z} = \alpha_z + \alpha_{u_1 z'}$ . This result and  $\mathcal{T}(\{u_1, z\}, \{z', z''\}, \{u_1\})$  give for all  $h \in \{2, \dots, K+1\}$ :  $\alpha_{u_1 z_h} = 0$  and

$$\alpha_{u_1 z_1} + \alpha_{z_h} = \alpha_{z_1}. \quad (16)$$

From Equations (15) and (16), we obtain that for all  $h \in \{2, \dots, K+1\}$ , the representative coefficients of  $z_h$  are equal and that the same applies to the  $\beta_h$ .

If  $|U|$  is equal to one, the proof is over. Indeed, in that case,  $z_2$  is lower than four and thus  $\alpha_{z_2}$  is equal to zero, which gives via Equation (16)  $\alpha_{u_1 z_1} = 0$ .

If  $|U|$  is greater than two, we then prove that  $\alpha_{u z}$  is equal to zero for all  $u \in U \setminus \{u_1\}$  and all  $z \in Z$ . This is obtained due to  $\mathcal{T}(\{u_1, u, z_1\}, \{z\}, \{u_1, u\})$ ,  $\mathcal{T}(\{u_1, u, z_1\}, \{z, z'\}, \{u_1, u\})$  and Equation (16).

We show that  $\alpha_{z_1}$  is equal to  $\alpha_{z_2}$ , thanks to  $\mathcal{T}(\{u_2, z_2\}, \{z_1\}, \{u_2\})$  which leads through Equation (16) to  $\alpha_{u_1 z_1} = 0$ .

If  $|U|$  is equal to two, then  $\alpha_z$  is equal to zero, and it remains to prove that  $\alpha_{u_1 u_2}$  is equal to zero, which can be done by  $\mathcal{T}(\{u_1, u_2, z_2\}, \{z_1\}, \{u_2\})$ .

Otherwise, for a given vertex  $u$  in  $U$ , let  $U'$  be a subset of  $U \setminus \{u\}$  which contains vertices  $u_1$  or  $u_2$ . The transformation  $\mathcal{T}(\{u, z, U'\}, \{n\}, \{u\})$  gives

$$\sum_{u' \in U'} \alpha_{uu'} + \alpha_z = \alpha_u \quad \forall z \in Z. \quad (17)$$

This equation shows that the sum of  $\alpha_{uu'}$  is equal to a constant for any possible  $U'$ . Let  $u'$  and  $u''$  be two vertices in  $U \setminus \{u\}$ . By successively choosing  $U'$  equal to  $\{u''\}$  and  $\{u', u''\}$  we obtain:  $\alpha_{uu'} = 0$ .

Eventually, Equation (17) gives:  $\alpha_z = \alpha_u \quad \forall (u, z) \in U \times Z$ . Since  $\alpha_{u_1}$  is equal to zero, the same applies to the other representative variables.  $\square$

## 5.4 The paw inequalities

Given a subset  $W = \{a, b, c, d\}$  of  $V$ , we define the *paw inequality* associated to  $W$  by:

$$x_{ab} + x_{bc} - x_{ac} + x_{cd} + x_b + x_c \leq 2. \quad (18)$$

Figure 7 represents the variables in this inequality.

**Lemma 5.7.** *Let  $K \in \{3, \dots, n-2\}$ . Inequality (18) is valid for  $P_{n,K}$  if and only if*

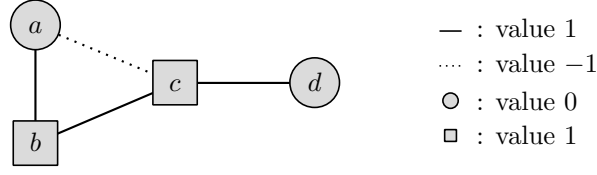


Figure 7: Representation of the coefficients of the paw inequality associated to a subset  $\{a, b, c, d\}$  of  $V$ .

1.  $a < b$ ;
2.  $\min(b, c, d) = d$ .

*Proof.* If  $\min(b, c, d)$  is not  $d$ , the left-hand side of Equation (18) is equal to 3 for any  $K$ -partition with a cluster equal to  $\{b, c, d\}$ . If  $b$  is lower than  $a$ , then Equation (18) is not satisfied for any  $K$ -partition  $\pi = \{C_1, \dots, C_K\}$  such that  $\{a, b\} \subset C_1$  and  $C_2 = \{c\}$ .

The addition of the triangle inequality (1)

$$x_{ab} + x_{bc} - x_{ac} \leq 1 \quad (19)$$

and the lower representative inequality (3)

$$x_c + x_{cd} \leq 1 \quad (20)$$

ensures that the paw inequality is valid if  $x_b$  is equal to zero.

If  $x_b$  is equal to one, we show that (18) is still valid since Equation (19) and Equation (20) cannot both be tight. In that case, vertices  $a$  and  $d$  cannot be in the same cluster as vertex  $b$  since their indices are lower than  $b$ . The only way for (19) to be tight under these conditions is for  $b$  and  $c$  to be together. Equation (20) is tight if vertex  $c$  is representative or if vertices  $c$  and  $d$  are together. In both cases  $x_b$  cannot be equal to one if vertices  $b$  and  $c$  are together.  $\square$

Let  $F_P$  be the face of  $P_{n,K}$  associated to inequality (18).

**Lemma 5.8.** *Under the conditions of Lemma 5.7, the face  $F_P$  is not a facet if  $c < b$  or  $K = n - 2$ .*

*Proof.* If  $c$  is lower than  $b$  we prove that  $F_P$  is included in the hyperplane induced by  $x_c + x_{cd} = 1$ . The expression

$$x_c + x_{cd} \quad (21)$$

can be equal to 0, 1 or 2.

If expression (21) equals 0, the solutions in  $F_P$  satisfy :  $x_{ab} + x_{bc} - x_{ac} + x_b = 2$ . This equation cannot be true since  $b$  has to be greater than both  $a$  and  $c$  according to Lemma 5.7 and the condition of the current lemma. The expression (21) cannot be equal to two either since  $d$  is lower than  $c$ . As a result, expression (21) is necessarily equal to one.

If  $K$  is equal to  $n - 2$ , no  $K$ -partition can contain both vertices  $a$  and  $c$  and thus,  $F_P$  is included in the hyperplane induced by  $x_{ac}$ .  $\square$

**Theorem 5.9.** *Let  $K \in \{3, n - 3\}$  and  $b \in \{4, \dots, n\}$ ,  $F_P$  is facet defining of  $P_{n,K}$  if and only if*

1.  $d < b < c$ ;
2.  $a < b$ .

*Proof.* A  $K$ -partition containing a cluster equal to  $\{b, c\}$  satisfies the paw inequality. Thus, by setting  $C_3$  equal to  $\{b, c\}$ , one can use the same reasoning as in the proof of Theorem (3.1) to obtain the following relations on the coefficients of an equation  $\alpha^T x = \alpha_0$  satisfied by all the points in  $F_P$ :

- $\alpha_{ij} = 0 \ \forall i \in V \setminus \{b, c, 1, 2, 3\} \forall j \in V \setminus \{b, c, i\}$ ;
- $\alpha_{1,2} = \alpha_{1,3} = \alpha_{2,3} \stackrel{def}{=} \beta$ ;

- $\alpha_i = -2\beta \forall i \in V \setminus \{b, c, 1, 2, 3\}$ .

The value of the remaining  $\alpha$  coefficients can be obtained through the transformations represented in Table 2. □

Conditions	Transformation	Results
-		$\alpha_{bc} = \alpha_{cd} \stackrel{def}{=} \gamma$
$\forall i \in V \setminus \{a, b, c, d\}$		$\alpha_{ci} = 0$
		$\alpha_{bi} = 0$
if $d \geq 4 \forall e, f \in \{1, 2, 3\}$ $\{a, b\} \subset C_3$		$\beta = 0$
if $d \leq 3$ $\forall e, f \in \{1, 2, 3\} \setminus \{d\}$ $C_2 \subset V \setminus \{b, c, d, e, f\}$ $C_3 = \{b, c, d\}$		$\beta = 0$
$\forall i \in V \setminus \{a, b, c, d\}$		$\alpha_c + \alpha_{ab} = \gamma + \alpha_b$
		$\alpha_{bd} = 0$
		$\alpha_{ac} = -\gamma$
		$\alpha_{ab} = \gamma, \alpha_b = \alpha_c$
		$\alpha_b = \gamma$

Table 2: Transformations used in Theorem 5.9. Each line presents a step of the proof. The last column corresponds to the result.

**Theorem 5.10.** Let  $K \in \{3, n - 3\}$ . The face  $F_P$  associated to the inequality  $x_{ab} + x_{bc} - x_{ac} + x_{cd} + x_c + x_{1,2} - \sum_{i=4}^n x_i \leq 4 - K$  - which corresponds to the paw inequality (18) for  $b$  equal to three - is facet for  $P_{n,K}$  if and only if

1.  $d < 3 < c$ ;
2.  $a < 3$ .

## 6 Numerical experiments

In this section we study the strength of our formulation and of the reinforcements with facets of the previous sections. We consider three data sets generated randomly, and we believe that the instances are quite difficult

since there are no preexisting classes to detect. Each data set  $D_1, D_2, D_3$  contains 100 instances formed from complete graphs. In  $D_1, D_2$  and  $D_3$ , the edge weights are respectively in the intervals  $[0, 500]$ ,  $[-250, 250]$  and  $[-500, 0]$ .

We first compare the value of the linear relaxation from our formulation to that of Chopra and Rao's formulation [1] (also in [17, 13]) adapted to obtain exactly  $K$  clusters:

$$(P_{cr}) \left\{ \begin{array}{l} \min \sum_{ij \in E} w_{i,j} x_{i,j} \\ -y_{it} + y_{jt} + x_{ij} \leq 1 \quad \forall ij \in E \forall t \in \{1, \dots, K\} \quad (22) \\ y_{it} - y_{jt} + x_{ij} \leq 1 \quad \forall ij \in E \forall t \in \{1, \dots, K\} \quad (23) \\ y_{it} + y_{jt} - x_{ij} \leq 1 \quad \forall ij \in E \forall t \in \{1, \dots, K\} \quad (24) \\ \sum_{t \in \{1, \dots, K\}} y_{it} = 1 \quad \forall i \in V \quad (25) \\ \sum_{i \in V} y_{it} \geq 1 \quad \forall t \in \{1, \dots, K\} \quad (26) \\ y_{it} \in \{0, 1\} \quad \forall i \in V \forall t \in \{1, \dots, K\} \quad (27) \\ x_{ij} \in \{0, 1\} \quad \forall ij \in E \quad (28) \end{array} \right.$$

For each vertex  $i \in V$  and each cluster  $k \in \{1, \dots, K\}$ , variable  $y_{i,k}$  is equal to 1 if vertex  $i$  is assigned to the cluster number  $K$  and 0 otherwise. Similarly to the triangle inequalities (1), Constraints (22) to (24) ensure the coherence of the obtained partition. Constraints (25) guarantee that each vertex is assigned to exactly one cluster and Constraints (26) ensure that the  $K$  clusters are non empty.

The comparison of the results of formulations  $(P_{er})$  and  $(P_{cr})$  over the three data sets are displayed in tables 3, 4 and 5. In each table and for each couple  $(n, K)$ , the value corresponding to formulation  $(P_{er})$  is the second one. Our formulation gives better relaxation values in all cases except in the case of data set  $D_1$  (*i.e.*, positive weights) when  $K$  is equal to 2.

We now only focus on formulation  $(P_{er})$ . Tables 6 and 7 give the number of instances for which the linear relaxation of  $(P_{er})$  gives an optimal solution for data set  $D_1$  and  $D_2$ . No optimal solution is obtained for the instances of  $D_3$ .  $D_2$  instances have weights of both signs, like in [10], and  $D_3$  corresponds to a variant which it considered to be easier (minimizing a cut with  $K$  parts and positive weights).

To evaluate the efficiency of a family of inequalities, we use a separation algorithm to add some of them to the formulation and observe the percentage of improvement of the value of the linear relaxation. This requires the definition of separation algorithms for each of the considered family.

The values of  $n$  considered in our experiments are low enough to allow an exhaustive enumeration of all the valid paw inequalities. Separating the 2-partition inequalities is NP-hard [7] and we are not able to enumerate them all. Instead we use a heuristic inspired from the well-known Kernighan-Lin algorithm [24]. A similar procedure is used for the separation of the general clique inequalities.

Separating the 2-chorded cycle inequalities is a bit more technical. In [25], Müller adapted an approach, introduced by Barahona and Mahjoub [26], to separate in polynomial time *odd closed walk inequalities* in directed graphs. Müller showed that the same algorithm can be applied to undirected graphs to allow the separation of a class of inequalities which includes the 2-chorded cycle inequalities. We adapted this approach to separate 2-chorded cycle inequalities from cycles which may contain repetitions.

We define a graph  $H = (V_H, A_H)$  such that for each edge  $ij \in E$ ,  $A_H$  contains (see example Figure 8):

- eight vertices:  $u_1^{ij}, u_2^{ij}, v_1^{ij}, v_2^{ij}, u_1^{ji}, u_2^{ji}, v_1^{ji}$  and  $v_2^{ji}$ ;
- four arcs:  $(u_1^{ij}, u_2^{ij}), (v_1^{ij}, v_2^{ij}), (u_1^{ji}, u_2^{ji}), (v_1^{ji}, v_2^{ji})$  of weight  $x_{ij}$ .

Moreover, to each pair of edges  $ij, ik \in E$  with a common endnode, we associate four additional arcs in  $A_H$ :  $(u_2^{ji}, v_1^{ik}), (v_2^{ji}, u_1^{ik}), (u_2^{ki}, v_1^{ij})$  and  $(v_2^{ki}, u_1^{ij})$  of weight  $-x_{jk} - \frac{1}{2}$ .

Let  $C = \{c_1, \dots, c_{2p+1}\}$  be an odd cycle of  $G$ . By construction,  $C$  induces a walk in  $H$  from  $u_1^{c_1, c_2}$  to  $v_1^{c_1, c_2}$  (see example Figure 9) of weight

$$\begin{aligned} & x_{c_1, c_2} - \frac{1}{2} - x_{c_1, c_3} + \dots + x_{c_{2p+1}, c_1} - \frac{1}{2} - x_{c_{2p+1}, c_2} \\ & = x(E(C)) - x(E(\overline{C})) - \frac{2p+1}{2} \\ & = x(E(C)) - x(E(\overline{C})) - \lfloor \frac{|C|}{2} \rfloor - \frac{1}{2}. \end{aligned}$$

Thus, there exists a cycle  $C$  which violates inequality (11) if and only if there exists a path from  $u_1^{c_1, c_2}$  to  $v_1^{c_1, c_2}$  in  $H$  whose length is greater than  $-\frac{1}{2}$ .

n	K								
	2	3	4	5	6	7	8	9	10
10	1102	114	40	16	7	3	2	1	0
	978	679	462	295	172	88	39	11	0
11	1268	123	45	19	8	4	2	1	0
	1033	737	522	356	227	135	71	33	10
12	1408	140	49	20	9	4	2	1	0
	1124	816	593	422	288	187	112	59	27
13	1529	125	46	20	8	4	2	1	0
	1177	874	649	476	341	231	149	90	47
14	1607	124	46	17	8	4	2	1	0
	1207	904	687	521	387	278	191	123	72
15	1733	125	48	20	8	4	2	1	0
	1275	971	749	578	440	327	234	159	100
16	1883	127	44	19	8	4	2	1	0
	1284	993	784	623	490	378	284	206	142
17	2010	118	43	19	8	4	2	1	0
	1375	1079	858	684	542	426	329	246	177
18	2055	117	41	18	9	4	2	1	0
	1391	1105	893	726	589	474	377	293	221
19	2270	133	49	21	9	4	2	1	1
	1463	1164	941	765	621	503	404	317	243
20	2359	123	43	17	8	4	2	1	0
	1439	1146	936	774	642	530	433	348	273

Table 3: Mean value of the linear relaxation from formulation ( $P_{er}$ ) and formulation ( $P_{cr}$ ) over the data set  $D_1$ . For each couple  $(n, K)$ , second line corresponds to formulation ( $P_{er}$ ).

Müller’s approach for undirected graphs only considers four vertices per edge ( $u_1^{ij}$ ,  $u_2^{ij}$ ,  $v_1^{ij}$  and  $v_2^{ij}$ ). As a consequence, a path in  $H$  between vertices  $u_1^{ij}$  and  $v_1^{ij}$  corresponds to a sequence of edges in  $G$  such that each edge has a common endnode with its neighbors. Such a sequence may not be a cycle (*e.g.*:  $\{ij, ik, il\}$ ). Four additional vertices per edge enable to give an orientation to the edge in the obtained sequence and thus ensure that it is a cycle (possibly with vertex repetitions).



Figure 8: Vertices and arcs of  $H$  associated to edges  $(ij)$  and  $(ik)$  in  $E$ .

After creating  $H$ , we obtain for all  $ij \in E$  the shortest path between vertices  $u_1^{ij}$  and  $v_1^{ij}$  using the Floyd-Warshall shortest path algorithm [27] and deduce the corresponding cycle in  $G$  and its associated 2-chorded cycle inequality. Eventually, the violated inequalities are added to the problem and the root relaxation is updated. This process is repeated until no more violated inequality is found.

For each family of inequalities and each couple  $(n, K)$  such that  $n \in \{10, 11, \dots, 20\}$  and  $K \in \{2, 3, \dots, 10\}$ , the average gain obtained over the instances of each data set is computed. The results are summed up in Table 8.

In our experiments, the paw inequalities are useless when all the weights are non negative (instances of  $D_1$ ) but they improve the value of the relaxation in the two other data sets. In the case of  $D_3$  (all negative weights)

n	K								
	2	3	4	5	6	7	8	9	10
10	-2246	-2251	-2213	-2146	-2064	-1940	-1754	-1409	0
	-1566	-1620	-1570	-1431	-1236	-990	-709	-385	0
11	-2784	-2778	-2740	-2672	-2591	-2484	-2341	-2136	-1731
	-1861	-1919	-1869	-1746	-1563	-1330	-1062	-752	-407
12	-3374	-3352	-3314	-3245	-3159	-3057	-2931	-2775	-2532
	-2115	-2171	-2149	-2050	-1891	-1679	-1422	-1131	-804
13	-4159	-4140	-4091	-4017	-3933	-3833	-3717	-3577	-3389
	-2557	-2606	-2576	-2480	-2326	-2119	-1867	-1575	-1250
14	-4936	-4901	-4846	-4768	-4674	-4566	-4452	-4318	-4157
	-2938	-2997	-2978	-2891	-2747	-2556	-2318	-2037	-1719
15	-5732	-5693	-5634	-5552	-5467	-5361	-5243	-5115	-4965
	-3332	-3399	-3388	-3314	-3180	-2994	-2759	-2483	-2171
16	-6683	-6641	-6582	-6496	-6410	-6306	-6193	-6067	-5919
	-3803	-3861	-3850	-3779	-3648	-3467	-3242	-2970	-2662
17	-7510	-7484	-7428	-7347	-7257	-7150	-7038	-6919	-6783
	-4250	-4308	-4304	-4240	-4117	-3945	-3726	-3463	-3161
18	-8539	-8510	-8449	-8364	-8272	-8163	-8050	-7925	-7784
	-4788	-4839	-4829	-4768	-4657	-4492	-4277	-4014	-3711
19	-9606	-9559	-9501	-9412	-9319	-9209	-9093	-8960	-8828
	-5300	-5361	-5357	-5306	-5199	-5041	-4839	-4588	-4300
20	-10770	-10725	-10666	-10576	-10483	-10374	-10259	-10135	-10000
	-5936	-5991	-5979	-5915	-5797	-5628	-5410	-5153	-4860

Table 4: Mean value of the linear relaxation from formulation ( $P_{er}$ ) and formulation ( $P_{cr}$ ) over the data set  $D_2$ . For each couple  $(n, K)$ , the second line corresponds to formulation ( $P_{er}$ ).

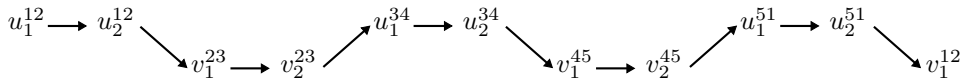


Figure 9: Path in  $H$  which corresponds to the cycle  $C = \{1, 2, 3, 4, 5\}$  in  $G$ .

we fail at finding any violated inequality except for the paw inequalities. These instances are however easier in practice. Regarding  $D_2$  the paw inequalities give the best improvement in average, but they are actually complementary to the generalized clique inequalities in the sense that they help more for the highest values of  $K$  while the generalized clique inequalities are more efficient for the smallest values of  $K$ .

For instances in  $D_1$  and  $D_2$ , the 2-chorded cycle inequalities are much less efficient than the 2-partition inequalities or the generalized clique inequalities. In particular, the general clique inequalities lead to a spectacular improvement for small values of  $K$  in the case of  $D_1$  instances.

### Conclusion

We have introduced a new formulation for the  $K$ -partitioning problem. By the addition of representative variables, we are able to break the symmetry in the edge variable formulation and fix the number of clusters. The resulting formulation seems to be stronger than the formulation with node-cluster variables and edge variables used by several authors ([1, 17, 13]) when  $K$  is greater than 2, at least on complete graphs. We have proved in this paper facet-defining results for several classical families of inequalities, and for a new family of inequalities that seems to be useful when there are negative weights.

The computing time for the 20-vertex instances takes only a few minutes using CPLEX 12.5 on a desktop computer. To actually solve problems to optimality for higher values of  $n$  will need to find a compromise between the separation and the solving of the linear programs at the nodes of a branch and bound procedure.

n	K								
	2	3	4	5	6	7	8	9	10
10	-10929	-10438	-9935	-9294	-8622	-7848	-6925	-5549	0
	-9959	-8724	-7488	-6253	-5017	-3780	-2540	-1293	0
11	-13490	-13001	-12498	-11865	-11191	-10453	-9589	-8557	-6986
	-12375	-11009	-9643	-8277	-6911	-5545	-4174	-2801	-1422
12	-16107	-15611	-15111	-14477	-13802	-13068	-12246	-11317	-10185
	-14876	-13397	-11918	-10439	-8960	-7480	-5999	-4517	-3028
13	-19402	-18912	-18414	-17773	-17102	-16372	-15574	-14680	-13651
	-18019	-16390	-14761	-13131	-11502	-9872	-8243	-6611	-4977
14	-22638	-22135	-21626	-20988	-20311	-19567	-18784	-17912	-16963
	-21141	-19390	-17639	-15888	-14137	-12385	-10634	-8883	-7130
15	-26015	-25517	-25003	-24366	-23716	-22983	-22206	-21357	-20428
	-24402	-22537	-20673	-18808	-16943	-15078	-13214	-11347	-9481
16	-29871	-29372	-28864	-28226	-27573	-26841	-26078	-25250	-24336
	-28121	-26122	-24124	-22125	-20126	-18128	-16129	-14130	-12132
17	-33803	-33308	-32805	-32174	-31515	-30786	-30034	-29226	-28346
	-31935	-29818	-27702	-25585	-23469	-21352	-19236	-17118	-15001
18	-38126	-37624	-37113	-36473	-35816	-35081	-34329	-33520	-32650
	-36139	-33893	-31648	-29402	-27157	-24911	-22666	-20420	-18174
19	-42640	-42143	-41643	-41001	-40352	-39621	-38871	-38047	-37207
	-40515	-38143	-35771	-33399	-31027	-28655	-26283	-23911	-21539
20	-47565	-47065	-46560	-45916	-45262	-44534	-43788	-42981	-42148
	-45309	-42801	-40294	-37786	-35279	-32771	-30263	-27755	-25247

Table 5: Mean value of the linear relaxation from formulation ( $P_{er}$ ) and formulation ( $P_{cr}$ ) over the data set  $D_3$ . For each couple  $(n, K)$ , the second line corresponds to formulation ( $P_{er}$ ).

n	K																		
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
10	0	0	0	12	47	75	95	100											
11	0	0	0	0	15	50	77	94	100										
12	0	0	0	0	8	29	70	91	99	100									
13	0	0	0	0	0	8	37	60	87	99	100								
14	0	0	0	0	0	0	20	45	72	91	98	100							
15	0	0	0	0	0	0	6	28	61	84	94	99	100						
16	0	0	0	0	0	0	3	13	31	66	80	91	98	100					
17	0	0	0	0	0	0	0	3	18	46	69	84	90	96	100				
18	0	0	0	0	0	0	0	0	6	19	51	70	91	97	99	100			
19	0	0	0	0	0	0	0	0	1	13	34	56	76	92	96	99	100		
20	0	0	0	0	0	0	0	0	0	4	12	37	60	78	94	99	100	100	

Table 6: Number of instances of  $D_1$  for which the linear relaxation of ( $P_{er}$ ) gives an optimal solution.

Still the results of this work are promising and show the interest of the polyhedral approach.

Further work will concentrate on improving the separation procedures and developing a branch and cut



n	K																		
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
10	20	53	31	4	3	0	0	0											
11	16	37	16	4	1	1	0	0	0										
12	8	21	16	4	0	0	0	0	0	0									
13	12	17	11	2	1	0	0	0	0	0	0								
14	6	14	9	3	0	0	0	0	0	0	0	0							
15	2	10	9	3	0	0	0	0	0	0	0	0	0						
16	2	5	5	1	0	0	0	0	0	0	0	0	0	0					
17	2	10	4	2	0	0	0	0	0	0	0	0	0	0	0				
18	2	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
19	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
20	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Table 7: Number of instances of  $D_2$  for which the linear relaxation of  $(P_{er})$  gives an optimal solution.

Data sets	Inequalities	Minimum	Maximum	Mean	Standard deviation
$D_1$	2-chorded cycle	0.0	4.7	1.6	1.3
	2-partition	0.0	30.1	13.6	8.2
	general clique	0.0	1186.6	198.3	253.4
	paw	0.0	0.0	0.0	0.0
$D_2$	2-chorded cycle	0.0	5.5	2.9	1.8
	2-partition	1.2	14.4	7.7	3.5
	general clique	0.0	24.9	3.1	6.5
	paw	0.0	37.3	8.2	10.3
$D_3$	2-chorded cycle	0.0	0.0	0.0	0.0
	2-partition	0.0	0.0	0.0	0.0
	general clique	0.0	0.0	0.0	0.0
	paw	0.6	33.3	5.9	6.5

Table 8: Statistical indicators related to the average gain in percentage obtained for values of  $n \in \{10, 11, \dots, 20\}$  and values of  $K \in \{2, 3, \dots, 10\}$  over the three data sets for each family of inequalities.

framework for the application that motivated this study [12].

## 7 References

### References

- [1] S. Chopra, M. Rao, The partition problem, *Mathematical Programming* 59 (1) (1993) 87–115.
- [2] C. Bichot, P. Siarry, *Graph Partitioning*, ISTE, John Wiley & Sons, London, 2013.
- [3] C. E. Ferreira, A. Martin, C. C. de Souza, R. Weismantel, L. A. Wolsey, The node capacitated graph partitioning problem: A computational study, *Mathematical Programming* 81 (2) (1998) 229–256.
- [4] E. Johnson, A. Mehrotra, G. Nemhauser, Min-cut clustering, *Mathematical Programming* 62 (1) (1993) 133–151.
- [5] M. R. Garey, D. S. Johnson, L. Stockmeyer, Some simplified NP-complete graph problems, *Theoretical computer science* 1 (3) (1976) 237–267.

- [6] M. Grötschel, Y. Wakabayashi, Facets of the clique partitioning polytope, *Mathematical Programming* 47 (1) (1990) 367–387.
- [7] M. Oosten, J. Rutten, F. Spieksma, The clique partitioning problem: Facets and patching facets, *Networks* 38 (4) (2001) 209–226.
- [8] J. E. Mitchell, Realignment in the national football league: Did they do it right?, *Naval Research Logistics (NRL)* 50 (7) (2003) 683–701.
- [9] M. Conforti, M. Rao, A. Sassano, The equipartition polytope. i: formulations, dimension and basic facets, *Mathematical Programming* 49 (1) (1990) 49–70.
- [10] M. Grötschel, Y. Wakabayashi, A cutting plane algorithm for a clustering problem, *Mathematical Programming* 45 (1) (1989) 59–96.
- [11] H.-J. Bandelt, M. Oosten, J. H. Rutten, F. C. Spieksma, Lifting theorems and facet characterization for a class of clique partitioning inequalities, *Operations research letters* 24 (5) (1999) 235–243.
- [12] Z. Ales, G. D. Duplessis, O. Serban, A. Pauchet, A methodology to design human-like embodied conversational agents based on dialogue analysis, in: *Proceedings of the 11th international conference on Autonomous Agents and Multiagent Systems (AAMAS) at Human-Agent Interaction Design and Models (HAIDM)*, 2012, pp. 34–49.
- [13] N. Fan, P. M. Pardalos, Linear and quadratic programming approaches for the general graph partitioning problem, *Journal of Global Optimization* 48 (1) (2010) 57–71.
- [14] W. W. Hager, D. T. Phan, H. Zhang, An exact algorithm for graph partitioning, *Mathematical Programming* 137 (1-2) (2013) 531–556.
- [15] P. Bonami, V. H. Nguyen, M. Klein, M. Minoux, On the solution of a graph partitioning problem under capacity constraints, in: *Proceedings of the 2nd International Symposium on Combinatorial Optimization (ISCO)*, LNCS, Vol. 7422, Springer, 2012, pp. 285–296.
- [16] C. Ferreira, A. Martin, C. de Souza, R. Weismantel, L. Wolsey, Formulations and valid inequalities for the node capacitated graph partitioning problem, *Mathematical Programming* 74 (3) (1996) 247–266.
- [17] V. Kaibel, M. Peinhardt, M. E. Pfetsch, Orbitopal fixing, *Discrete Optimization* 8 (4) (2011) 595–610.
- [18] M. Campêlo, V. A. Campos, R. C. Corrêa, On the asymmetric representatives formulation for the vertex coloring problem, *Discrete Applied Mathematics* 156 (7) (2008) 1097–1111.
- [19] M. Labbé, F. Özsoy, Size-constrained graph partitioning polytopes, *Discrete Mathematics* 310 (24) (2010) 3473–3493.
- [20] Z. Ales, A. Knippel, A. Pauchet, On the polyhedron of the  $k$ -partitioning problem with representative variables, Tech. rep., LMI/LITIS, INSA de Rouen (2014). arXiv:0902.0885.
- [21] Z. Ales, Extraction et partitionnement pour la recherche de régularités : application à l’analyse de dialogues, Ph.D. thesis, INSA de Rouen (2014).
- [22] G. L. Nemhauser, L. A. Wolsey, *Integer and Combinatorial Optimization*, Wiley-Interscience, New York, NY, USA, 1988.
- [23] X. Ji, J. Mitchell, The clique partitioning problem with minimum size requirement, Tech. rep., Rensselaer Polytechnic Institute (2005).
- [24] B. W. Kernighan, S. Lin, An efficient heuristic procedure for partitioning graphs, *Bell System Technical Journal* 49 (2) (1970) 291–307.
- [25] R. Müller, On the partial order polytope of a digraph, *Mathematical Programming* 73 (1) (1996) 31–49.
- [26] F. Barahona, A. R. Mahjoub, On the cut polytope, *Mathematical Programming* 36 (2) (1986) 157–173.
- [27] R. K. Ahuja, T. L. Magnanti, J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1993.