A decorative graphic in the top-left corner consisting of several overlapping geometric shapes in teal and orange.

2016/21

## “Facet” Separation with One Linear Program

MICHELE CONFORTI AND LAURENCE A. WOLSEY

A large, stylized number '50' in a golden-yellow color, rendered with thick, rounded lines. The '5' is on the left and the '0' is on the right, with the '0' having a circular opening. The number is positioned behind the main title text.

50 YEARS OF  
**CORE**  
DISCUSSION PAPERS

## **CORE**

Voie du Roman Pays 34, L1.03.01

Tel (32 10) 47 43 04

Fax (32 10) 47 43 01

Email: [immaq-library@uclouvain.be](mailto:immaq-library@uclouvain.be)

<http://www.uclouvain.be/en-44508.html>

CORE DISCUSSION PAPER  
2016/21

## “Facet” Separation with One Linear Program

Michele Conforti<sup>1</sup> and Laurence A. Wolsey<sup>2</sup>

June 2016

### Abstract

Given polyhedron  $P$  and a point  $x^*$ , the separation problem for polyhedra asks to certify that  $x^* \in P$  and if not, to determine an inequality that is satisfied by  $P$  and violated by  $x^*$ . This problem is repeatedly solved in cutting plane methods for Integer Programming and the quality of the violated inequality is an essential feature in the performance of such methods.

In the paper we address the problem of finding efficiently an inequality that is violated by  $x^*$  and either defines an improper face or a facet of  $P$ . We provide some evidence that our method works on structured and unstructured problems.

Keywords: integer programming, separation problem, polyhedra, extended formulations, facets, cutting plane algorithm, split inequalities, Benders’ algorithm.

AMS 2010 Mathematics Subject Classification: 90C27, 90C57

---

<sup>1</sup>Dipartimento di Matematica, Università di Padova.

<sup>2</sup>CORE, Université catholique de Louvain.

This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister’s Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

# 1. Introduction

Given a polyhedron  $P$  and a point  $x^*$ , the *separation problem* asks to certify that either  $x^* \in P$  or find an inequality  $\gamma x \leq \gamma_0$  that is valid for  $P$  such that  $\gamma x^* > \gamma_0$ . The separation problem is solved at each iteration of a cutting plane phase of algorithms for pure or mixed integer programming and is therefore crucial for the performance of such methods.

A most important consequence of the Ellipsoid method for linear programming states that, assuming  $P$  satisfies some natural conditions, the separation problem can be efficiently solved if and only if one can efficiently optimize a linear function over  $P$ . In this paper we focus on the following aspect of this problem:

*If  $x^* \notin P$ , we insist on finding an inequality  $\gamma x \leq \gamma_0$  that is violated by  $x^*$  and has "strong" geometric properties: Either  $\gamma x = \gamma_0$  holds for all  $x \in P$  (i.e.  $\gamma x \leq \gamma_0$  induces an improper face of  $P$ ) or  $\gamma x \leq \gamma_0$  induces a facet of  $P$ .*

Of course, the above problem can be efficiently solved by inspection if  $P$  is described by an irredundant system of equations and inequalities and this system is small. However we investigate here the case in which either  $P$  or the cone that contains the vectors  $(\gamma, \gamma_0)$  that define inequalities that are valid for  $P$  is described in an extended space by a system of inequalities that uses additional variables. This is most often the case for instance, in cutting plane methods based on split or lift-and-project cuts and in Benders' method. There is an extensive literature on the merit and weakness of the cuts generated and the need to find the "correct" normalization (if it exists) when reducing optimization over a cone to a bounded linear program. We believe our result is relevant to such approaches.

We show that the position of  $P$  with respect to the origin is crucial to the solution of the above problem with a single LP. We show first cases depending also on the position of  $x^*$  in which a single LP suffices. Then our main result is to show that a single LP suffices when the origin is in the relative interior of  $P$  whatever the position of  $x^*$ . This fact has a simple geometric interpretation, that we explain. We also discuss how to find a point in the relative interior of  $P$ , again by solving a single LP.

We discuss the merits and issues related to the normalizations proposed in the literature and we illustrate our approach on three examples that arise in Integer Programming: split disjunctions, fixed charge networks and Benders' reformulation.

## 1.1 Relation to earlier work

Balas [2] introduced the concept of disjunctive programming in the early 70's, which studies optimization over the union of polyhedra. Split disjunctions and the lift-and-project approach developed by Balas, Ceria and Cornuéjols [3] study optimization over the union of two polyhedra, which in the 0-1 case are two faces of the same polyhedron. More generally, the theory of multi-row cuts derived from lattice-free convex sets introduced by Andersen et. al. [1] fits this framework.

Balas and Perregard [4] survey the area of disjunctive programming, and among others discuss the question of different normalizations of the cone of valid inequalities in order to obtain a "good" valid inequality by linear programming for the lift-and-project approach. This question is also addressed in Fischetti, Lodi and Tramontani [13]. They show that even when  $P \subset \mathbb{R}^2$ , the inequalities generated by using different normalizations can be very weak. They also propose a

different normalization and compare its behavior to some of the alternatives. A core point in the relative interior of  $P$  is used in Fischetti and Salvagnin [14] in an attempt to speed up the convergence of the cutting plane algorithm .

Benders' algorithm [6] can be viewed as a separation algorithm in which an extended formulation of  $P$  is given explicitly. Magnanti and Wong [18] present a modification designed to generate "pareto-optimal" cuts, see also Papadakos [21]. Their approach is also based on the use of a core point in the relative interior of  $P$ . Fischetti, Salvagnin and Zanette [15] discuss the normalization problem as it arises in Benders' algorithm.

We finally point out that the separation problem for convex sets is a central problem in Convex Analysis, see e.g. Chapter A4 in [17] and Convex Programming, see e.g. [7]. In this context, quality of separation is mostly measured in terms of maximizing a given norm.

Cornuéjols and Lemaréchal [11] study the problem of separation of the origin from a closed convex set  $Q$  from a convex analysis perspective and apply it to the polyhedral case, most notably to the disjunctive programming case. They point out the relevance of the reverse polar introduced by Balas, that contains the inequalities that are valid for  $Q$  and are violated by the origin (the point to be cut off, in this case).

Cadoux [9] takes a convex analysis/geometric approach concerning the depth/strength of a cut. He shows how a cut that maximizes a given norm can be decomposed into a conic combination of facet-inducing inequalities by solving a series of linear programs.

## 2. The separation problem for polyhedra

Given a polyhedron  $P \subseteq \mathbb{R}^n$ , in this paper we study the  $f$ -cone of  $P$ , which contains the normal vectors together with the right-hand sides of the inequalities that are valid for  $P$ . More precisely, the  $f$ -cone of  $P$  is the polyhedral cone

$$\Gamma(P) = \left\{ \begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} : \gamma x \leq \gamma_0 \ \forall x \in P \right\}.$$

(We use basic results from the theory of polyhedra, which may be found e.g. in Ch. 3 of [10]. Given a polyhedron  $P$ , we denote by  $\dim(P)$ ,  $\text{aff}(P)$ ,  $\text{lin}(P)$ ,  $\text{rec}(P)$ ,  $\text{int}(P)$ ,  $\text{relint}(P)$ ,  $\text{bd}(P)$  its dimension, affine hull, lineality space, recession cone, interior, relative interior and boundary). The following proposition, see e.g. Proposition 5.1 in [20], gives a minimal set of generators of  $\Gamma(P)$ :

**Proposition 1** *Let  $P \subset \mathbb{R}^n$  be a nonempty polyhedron and let  $\gamma^i x = \gamma_0^i$ ,  $i \in I^=$  and  $\gamma^i x \leq \gamma_0^i$ ,  $i \in I^<$  be respectively irredundant representations of  $\text{aff}(P)$  and of the set of facets of  $P$ . Then*

$$\Gamma(P) = \text{cone} \left( \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \gamma^i \\ \gamma_0^i \end{pmatrix}, i \in I^< \right) + \text{lin} \left( \begin{pmatrix} \gamma^i \\ \gamma_0^i \end{pmatrix}, i \in I^= \right). \quad (1)$$

Furthermore  $\begin{pmatrix} \gamma^i \\ \gamma_0^i \end{pmatrix}$   $i \in I^<$  and  $\begin{pmatrix} \gamma^i \\ \gamma_0^i \end{pmatrix}$   $i \in I^=$  are always necessary in the above representation.

The following proposition gives a minimal set of equations and inequalities that defines  $\Gamma(P)$ .

**Proposition 2** Given a nonempty polyhedron  $P \subset \mathbb{R}^n$ , let  $(b_k, k \in B)$  be a basis of  $\text{lin}(P)$ ,  $(v_i, i \in V)$ , and  $(r_j, j \in R)$  be the vertices and extreme rays of the projection of  $P$  on the orthogonal complement of  $\text{lin}(P)$ . Then

$$\Gamma(P) = \left\{ \begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} : \gamma v_i - \gamma_0 \leq 0, i \in V; \gamma r_j \leq 0, j \in R; \gamma b_k = 0, k \in B \right\} \quad (2)$$

Furthermore the above linear system provides an irredundant representation of  $\Gamma(P)$ .

**Remark 3** It follows from Propositions 1 and 2 that  $\Gamma(P)$  is a pointed cone if and only if  $\dim(P) = n$  and  $\dim(\Gamma(P)) = n$  if and only if  $P$  is pointed.

Whenever  $P \neq \emptyset$  is represented by a linear system,  $\Gamma(P)$  can be expressed using the Farkás multipliers that certify the validity for  $P$  of an inequality. More precisely, if  $P = \{x : \exists y \text{ s.t. } Ax + By \leq d\}$ , then

$$\Gamma(P) = \left\{ \begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} : \exists u \geq 0 \text{ s.t. } \gamma = uA, 0 = uB, \gamma_0 \geq ud \right\}. \quad (3)$$

In this paper we address the following *separation problem* for  $P$  that we (informally) state as follows:

Given a criterion and a point  $x^*$ , certify that either  $x^* \in P$  or select an inequality  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma(P)$  with  $\gamma x^* > \gamma_0$  that optimizes the given criterion.

Even if this "criterion" is a linear function, this is a challenging problem, first because  $\Gamma(P)$  is a cone and also because, in the applications we consider,  $\Gamma(P)$  is represented as in (1), (2) or (3). We now elaborate on these points and survey the existing literature.

Since  $\Gamma(P)$  is a cone, the maximum of a linear function over  $\Gamma(P)$  is either 0 or  $\infty$ . Usually this problem is overcome by *normalizing*  $\Gamma(P)$ , i.e. adding inequalities that make sure the given objective is bounded. Ideally, one would like to *truncate*  $\Gamma(P)$ : that is add a set of inequalities that transforms  $\Gamma(P)$  into a polytope. A most desirable truncation is one using a single inequality. Note that such a truncation exists if and only if  $\Gamma(P)$  is a pointed cone.

When  $\Gamma(P)$  is truncated by adding a single inequality, the set of extreme rays of  $\Gamma(P)$  is in one-to-one correspondence with the newly created vertices, and this is a desirable feature. However, the added inequality introduces a ranking of the vertices that determines which one optimizes the objective. Indeed one can choose any extreme ray  $r$  of  $\Gamma(P)$  and construct a truncation so that the vertex that corresponds to  $r$  is optimal with respect to the given criterion.

There is a further, important issue: the truncation with a single inequality is usually performed on a cone  $C$  that projects onto  $\Gamma(P)$ , such as the cone described by the inequalities in (3). Since  $\Gamma(P)$  is the projection of  $C$  in the  $(\gamma, \gamma_0)$ -space, even if  $C$  is truncated with a single inequality, the projected truncation of  $C$  does not usually result in the truncation of  $\Gamma(P)$  with a single inequality, and a ray that is extreme in  $C$ , but whose projection is not extreme in  $\Gamma(P)$ ,

may be truncated to a vertex of the projected polyhedron that optimizes the given objective function, but whose corresponding inequality is very weak.

Finally, even if  $P$  is represented by a linear system with few inequalities,  $P$  may have exponentially many vertices and extreme rays. So the size of the representation of  $\Gamma(P)$  in (2) is too large.

In this paper we address the following criterion:

*When  $x^* \notin P$ , select an inequality  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma(P)$  with  $\gamma x^* > \gamma_0$  that defines an improper face or a facet of  $P$ .*

The importance of finding violated facets has been recognized by various authors. Cadoux's approach, cited above, involves the solution of a series of linear programs and Cornuéjols and Lemaréchal present an approach due to Bonami [8] that also involves solving several linear programs. Padberg et al., in their work on 0-1 integer programs and on combinatorial optimization problems, such as the travelling salesman problem, argued strongly for the importance of violated facet-defining inequalities, see for example [12, 22].

## 2.1 Disjunctive programming, lift-and-project, split inequalities

We illustrate the above questions on a problem that is of importance in integer programming, see e.g. [10] Ch.5. Given polyhedron  $P = \{x : Ax \leq b\}$  and  $\pi \in \mathbb{Z}^n$ ,  $\pi_0 \in \mathbb{Z}$ , we define

$$P^0 := P \cap \{x : \pi x \leq \pi_0\} \quad P^1 := P \cap \{x : \pi x \geq \pi_0 + 1\}.$$

Let  $P^{(\pi, \pi_0)} := \text{conv}(P^0 \cup P^1)$ . Then  $P^{(\pi, \pi_0)} \subseteq P$  and  $P \cap \mathbb{Z}^n = P^{(\pi, \pi_0)} \cap \mathbb{Z}^n$ . Furthermore (assuming  $P$  pointed) if  $x^*$  is a vertex of  $P$  and  $\pi_0 < \pi x^* < \pi_0 + 1$ , then  $x^* \notin P^{(\pi, \pi_0)}$ . In this case  $P^{(\pi, \pi_0)}$  is a better approximation than  $P$  of  $\text{conv}(P \cap \mathbb{Z}^n)$ . In the lift-and-project method  $\pi$  is a unit vector and  $\pi_0 = 0$ .

It can be shown that  $P^{(\pi, \pi_0)}$  is a polyhedron which can be obtained as the projection in the  $x$ -space of the polyhedron defined by the following system:

$$\begin{aligned} Ax^0 &\leq b\lambda \\ \pi x^0 &\leq \pi_0\lambda \\ Ax^1 &\leq b(1-\lambda) \\ \pi x^1 &\geq \pi_0(1-\lambda) \\ x^0 + x^1 &= x \\ 0 \leq \lambda &\leq 1 \end{aligned} \tag{4}$$

An inequality is valid for  $P^{(\pi, \pi_0)}$  if and only if it is valid for both  $P^0$  and  $P^1$ , so assuming that  $P^0, P^1$  are both nonempty, let  $C$  be the cone defined by the following system:

$$\begin{aligned} uA + u_0\pi &= \gamma \\ vA - v_0\pi &= \gamma \\ ub + u_0\pi_0 &\leq \gamma_0 \\ vb - v_0(\pi_0 + 1) &\leq \gamma_0 \\ u, u_0, v, v_0 &\geq 0 \end{aligned} \tag{5}$$

Then by (3), we have that  $\Gamma(P^{(\pi, \pi_0)}) = \{(\gamma, \gamma_0) : \exists(u, u_0, v, v_0) \text{ s.t. } (u, u_0, v, v_0, \gamma, \gamma_0) \in C\}$ .

Since  $\max\{x^* \gamma - \gamma_0 : (u, u_0, v, v_0, \gamma, \gamma_0) \in C\} = \infty$  if and only if  $x^* \notin P^{(\pi, \pi_0)}$ , some normalizations of  $C$  have been introduced and studied in the literature.

- In the context of lift-and-project, Balas and Perregaard [5] mention  $\gamma_0 \in \{-1, 1\}$  as an effective normalization for covering and packing problems. The normalization  $\gamma_0 \in [-1, 1]$  is also studied and will be important in this paper.  $\sum_{i=1}^n |\gamma_i| = 1$  is a normalization that needs the  $2^n$  inequalities that define the octahedron and may introduce additional vertices. They also introduce a generic normalization with a single inequality  $a\gamma \leq b$  and mention that when  $P$  is full-dimensional and the separation LP produces a finite solution which is extreme, then a violated facet is produced.

- $u_0 + v_0 = 1$  Balas and Perregaard [5], [4] show that finding a most violated inequality (i.e. one that maximizes the objective function) can be reformulated as an LP on  $n$  variables and corresponds to a (possibly infeasible) basis of the system  $Ax \leq b$  defining  $P$ , see also [10] Ch.5.

Fischetti et al. (Theorem 2 in [13]) show that if  $x^*$  is a vertex of  $P$ , a most violated inequality can be read from a basis defining  $x^*$ .

- $\mathbf{1}u + u_0 + \mathbf{1}v + v_0 = 1$  This normalization produces a truncation of  $C$  and has been introduced by Balas and has been studied by Balas and Perregaard among others. Fischetti et al. report better computational results than for the previous normalization. However, the performance is obviously tied to the scaling of the constraints defining  $P$ .

Although a cut-generating LP is solved, the cut obtained fails to have strong geometric properties: It is neither an implicit equation nor a facet of  $P^{(\pi, \pi_0)}$ . To illustrate this point, the following 2-variable example is taken from Fischetti et al. [13].

**Example 1** *The problem is  $\max\{x_1 + 2x_2 : x \in P \cap \mathbb{Z}^2\}$  where  $P \subset [0, 1]^2$  is the polytope:*

$$\begin{array}{rcl} -4x_1 & +4x_2 & \leq 2 \\ 2x_1 & +2x_2 & \leq 3 \\ -8x_1 & +4x_2 & \leq 1 \\ x_1 & & \leq 1 \\ & x_2 & \leq 1 \\ -x_1 & & \leq 0 \\ & -x_2 & \leq 0. \end{array}$$

*The unique optimal LP solution is  $x^* = (0.5, 1)$ . Taking  $(\pi, \pi_0) = ((1, 0), 0)$ , namely the disjunction  $x_1 \leq 0$  or  $x_1 \geq 1$ , one obtains  $P_0 = \{x : x_1 = 0, 0 \leq x_2 \leq 0.25\}$  and  $P_1 = \{x : x_1 = 1, 0 \leq x_2 \leq 0.5\}$  and  $P^{(\pi, \pi_0)} = \{x : 0 \leq x_1 \leq 1, x_2 \geq 0, -x_1 + 4x_2 \leq 1\}$ .*

*Solving the cut generating LP  $\max\{\gamma x^* - \gamma_0 : (5)\}$  with either the normalization  $\mathbf{1}u + u_0 + \mathbf{1}v + v_0 = 1$  or  $u_0 + v_0 = 1$  leads to the violated inequality  $x_2 \leq 0.5$ . If the constraint  $x_2 \leq 1$  is rescaled as  $kx_2 \leq k$  with  $k \geq 8$  both normalizations give the violated inequality  $-x_1 + 2x_2 \leq 1$ . Neither of these inequalities defines a facet of  $P^{(\pi, \pi_0)}$ .*

*On the other hand we will show below why, given the relative positions of  $P$ ,  $0$  and the point  $x^*$  to be cut off, using the normalization  $-1 \leq \gamma_0 \leq 1$ , the cut generation LP produces the only facet  $-x_1 + 4x_2 \leq 1$  of  $P^{(\pi, \pi_0)}$  that is violated by  $x^*$ .*



### 3. A Separation LP

Given polyhedron  $P$ , we let  $\Gamma^{\langle \rangle}(P) = \left\{ \begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma(P), -1 \leq \gamma_0 \leq 1 \right\}$  and given  $x^*$ , we consider the LP

$$\zeta = \max \quad x^* \gamma - \gamma_0$$

$$\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma^{\langle \rangle}(P) \quad (6)$$

By (2), any linear system that defines  $\Gamma^{\langle \rangle}(P)$  (in the original space or in an extended space) is equivalent to the constraints of the LP below, which is equivalent to (6):

$$\zeta = \max \quad x^* \gamma - \gamma_0$$

$$\begin{aligned} V^T \gamma - \mathbf{1} \gamma_0 &\leq 0 \\ R^T \gamma &\leq 0 \\ B^T \gamma &= 0 \\ -1 \leq \gamma_0 &\leq 1 \end{aligned} \quad (7)$$

where  $B$  is a matrix whose columns are the vectors in a basis of  $\text{lin}(P)$ ,  $V$  and  $R$  are matrices whose columns are vertices and extreme rays of the projection of  $P$  on the orthogonal complement of  $\text{lin}(P)$ .

The dual of (7) is

$$\min |\psi - 1|$$

$$\begin{aligned} V\lambda + R\mu + B\nu &= x^* \\ \lambda, \mu &\geq 0 \\ \mathbf{1}\lambda &= \psi \end{aligned} \quad (8)$$

Note that the LP (7) admits a finite optimum if and only if (8) is feasible. We characterize when this happens. We assume  $P$  is nonempty, we let  $Ax \leq b$  be a system of inequalities defining  $P$  and consider the following cones.

- $\text{Hom}^{\text{ext}}(P) := \{(\lambda, \mu, \nu, x, \psi) : x = V\lambda + R\mu + B\nu, \mathbf{1}\lambda = \psi, \lambda, \mu \geq 0\}$
- $\text{Hom}^\psi(P) := \{(x, \psi) : Ax \leq b\psi, \psi \geq 0\}$ .
- $\text{cone}^\psi(P) := \{(x, \psi) : x = \psi z, z \in P, \psi \geq 0\}$ .

Note that if  $P \neq \emptyset$ ,  $\text{Hom}^\psi(P)$  does not depend on the specific system of inequalities representing  $P$ . When the projection of  $P$  on the orthogonal complement of  $\text{lin}(P)$  contains a vertex, inequality  $\psi \geq 0$  is implied by the system defining  $\text{Hom}^{\text{ext}}(P)$ . Therefore the representation theorem of polyhedra (see e.g. Theorem 3.37 in [10]) implies the following

**Remark 4** *Given nonempty polyhedron  $P$  let  $\text{Hom}^{\text{Ext}}(P)$ ,  $\text{Hom}^\psi(P)$  be the cones defined above. Then  $\text{Hom}^\psi(P) = \text{proj}_{(x, \psi)}(\text{Hom}^{\text{ext}}(P))$  and the LP (8) is equivalent to*

$$\min\{|\psi - 1| : (x^*, \psi) \in \text{Hom}^\psi(P)\}. \quad (9)$$

**Proposition 5** Given a nonempty polyhedron  $P$ , let  $\text{Hom}^\psi(P)$ ,  $\text{cone}^\psi(P)$  be the cones defined above:

1.  $\text{cone}^\psi(P) \subseteq \text{Hom}^\psi(P)$  and  $\text{Hom}^\psi(P) \setminus \text{cone}^\psi(P) = (\text{rec}(P), 0) \setminus (0, 0)$
2.  $(\text{rec}(P), 0)$  is the face of  $\text{Hom}^\psi(P)$  induced by  $\psi = 0$
3.  $\text{Hom}^\psi(P) = \text{cl}(\text{cone}^\psi(P))$ .

*Proof.* Let  $(z, \psi) \in \text{cone}^\psi(P)$ . Then  $z = \psi x$ , where  $\psi \geq 0$  and  $Ax \leq b$ . Therefore  $Az \leq b\psi$  and this shows  $\text{cone}^\psi(P) \subseteq \text{Hom}^\psi(P)$ . Let now  $(z, \psi) \in \text{Hom}^\psi(P) \setminus \text{cone}^\psi(P)$ . Then  $\psi = 0$ , since otherwise  $x := \frac{z}{\psi} \in P$  and  $z = \psi x$ . Furthermore  $z \neq 0$ , since otherwise  $P \neq \emptyset$ ,  $z = 0x$  for any  $x \in P$ . This shows that  $(z, 0) \in (\text{rec}(P), 0) \setminus (0, 0)$  and this proves 1..

Since  $P \neq \emptyset$ ,  $(\text{rec}(P), 0) = (\{r : Ar \leq 0\}, 0)$  and this proves 2..

By 1. and 2., we have that  $\text{cone}^\psi(P) \subseteq \text{Hom}^\psi(P)$  and  $\text{Hom}^\psi(P) \setminus \text{cone}^\psi(P) \subseteq \text{bd}(\text{Hom}^\psi(P))$ . Since  $\text{Hom}^\psi(P)$  is a closed set, this proves 3..  $\square$

**Remark 6** Given a nonempty polyhedron  $P$ , let  $\text{Hom}(P) := \{x : \exists \psi \text{ s.t. } (x, \psi) \in \text{Hom}^\psi(P)\}$ ,  $\text{cone}(P) := \{x : \exists \psi \text{ s.t. } (x, \psi) \in \text{cone}^\psi(P)\}$ . Then  $\text{cone}(P) \subseteq \text{Hom}(P)$  and  $\text{Hom}(P) \setminus \text{cone}(P) = \text{rec}(P) \setminus \text{cone}(P)$ .

**Proposition 7** Given nonempty polyhedron  $P$ , the LP (8) is feasible if and only if  $x^* \in \text{cone}(P) \cup \text{rec}(P)$ . Let  $\psi^*$  be an optimal solution of (8). Then  $\psi^* \geq 0$  and

1.  $\psi^* = 1$  if and only if  $x^* \in P$
2.  $\psi^* = 0$  if and only if  $x^* \notin P$  and  $x^* \in (\text{rec}(P) \setminus \text{cone}(P)) \cup \{0\}$
3.  $\psi^* \neq 0, 1$  if and only if  $x^* \in \text{cone}(P) \setminus (P \cup \{0\})$  and:
  - If  $\{\frac{x^*}{\psi}, \psi > 1\} \cap P \neq \emptyset$ , then  $1 < \psi^* = \min\{\psi : \frac{x^*}{\psi} \in P\}$
  - If  $\{\frac{x^*}{\psi}, \psi < 1\} \cap P \neq \emptyset$ , then  $1 > \psi^* = \max\{\psi : \frac{x^*}{\psi} \in P\}$ .

*Proof.* By Remark 4, (8) is feasible if and only if (9) is feasible and by Remark 6, this happens if and only if  $x^* \in \text{cone}(P) \cup \text{rec}(P)$ .

Now 1. is immediate, so we assume  $x^* \notin P$ . If  $x^* \in \text{rec}(P) \setminus \text{cone}(P) \cup \{0\}$ , since  $x^* \notin P$ , by Proposition 7,  $\psi = 0$  is the only value for which  $(x^*, \psi) \in \text{Hom}^\psi(P)$  and this proves 2..

Assume finally  $x^* \in \text{cone}(P) \setminus (P \cup \{0\})$ . Then  $\psi \neq 0, 1$ . Therefore exactly one of  $\{\frac{x^*}{\psi}, \psi < 1\} \cap P$ ,  $\{\frac{x^*}{\psi}, \psi > 1\} \cap P$  is nonempty and we have 3..  $\square$

**Proposition 8** Assume  $x^* \in \text{cone}(P) \setminus (P \cup \{0\})$  and let  $\psi^*$  be an optimal solution of (8). Then

$y^* := \frac{x^*}{\psi^*} \in \text{bd}(P)$  and  $F := \left\{ \begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma^{\langle \cdot \rangle}(P) : \gamma y^* = \gamma_0 \right\}$  is the optimal face of the LP(7).

Furthermore  $\gamma_0 = 1$  for every  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in F$  if  $\psi^* > 1$  and  $\gamma_0 = -1$  for every  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in F$  if  $\psi^* < 1$ .

*Proof.* Since  $x^* \in \text{cone}(P) \setminus (P \cup \{0\})$ , case 3. of Proposition 7 applies. Let  $\psi^*$  be the optimal value of  $\psi$  in (8) and let

$$F := \left\{ \begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma^{\langle \rangle}(P) : \gamma x^* = \gamma_0 + |\psi^* - 1| \right\}$$

be the optimal face of the LP (7). Since  $\max\{0, \psi^* - 1\}$  and  $\max\{0, -(\psi^* - 1)\}$  are the optimal values taken by the dual variables associated with constraints  $\gamma_0 \leq 1$  and  $\gamma_0 \geq -1$  in (7) respectively, then complementary slackness shows that  $\gamma_0 = 1$  for every  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in F$  when

$\psi^* > 1$  and  $\gamma_0 = -1$  for every  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in F$  when  $\psi^* < 1$ .

Assume  $\psi^* > 1$ . Then  $\gamma x^* = \gamma_0 + \psi^* - 1$  for every  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in F$  and  $\gamma x^* < \gamma_0 + \psi^* - 1$  for every  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma^{\langle \rangle}(P) \setminus F$ . Since  $\gamma_0 = 1$  for every  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in F$  and  $x^* = \psi^* y^*$ , then  $F := \left\{ \begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma^{\langle \rangle}(P) : \gamma y^* = \gamma_0 \right\}$ . When  $\psi^* < 1$ , the proof is the same.  $\square$

**Theorem 9** *If  $0 \in \text{relint}(P)$ , then:*

1. *The LP (8) is infeasible if and only if  $x^* \notin \text{aff}(P)$ . In this case, let  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix}$  be an unbounded solution of (7). Then  $\gamma_0 = 0$ ,  $\gamma x^* > 0$  and  $\gamma x = 0$  for all  $x \in P$ .*
2. *If the LP (8) is feasible, let  $\psi^*$  be a optimal solution of (8).  $\psi^* = 1$  if and only if  $x^* \in P$  and  $\psi^* > 1$  if and only if  $x^* \in \text{aff}(P) \setminus P$ . In this latter case, let  $F := \left\{ \begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma^{\langle \rangle}(P) : \gamma x^* = \gamma_0 + \psi^* - 1 \right\}$  be the optimal face of the LP (7) and let  $y^* = \frac{x^*}{\psi^*}$ . Then  $y^* \in \text{bd}(P)$  and  $F := \left\{ \begin{pmatrix} \gamma \\ 1 \end{pmatrix} \in \Gamma^{\langle \rangle}(P) : \gamma y^* = 1 \right\}$*

*Proof.* Since  $0 \in \text{relint}(P)$  we have that  $\text{aff}(P) = \text{Hom}(P)$ . Since the LPs (8) and (9) are equivalent, this shows that (8) is infeasible (and (7) is unbounded) if and only if  $x^* \notin \text{aff}(P)$ .

Assume this is the case and let  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix}$  an unbounded ray of  $P$ . Then  $\gamma x^* > \gamma_0$  and  $\begin{pmatrix} \lambda \gamma \\ \lambda \gamma_0 \end{pmatrix} \in \Gamma^{\langle \rangle}(P)$  for every  $\lambda > 0$ . Therefore  $\gamma_0 = 0$ . Since  $0 \in \text{relint}(P)$  and  $\gamma 0 = 0$ , then  $\gamma x = 0$  for all  $x \in P$ . This proves 1..

Assume now that (8) is feasible, i.e.  $x^* \in \text{aff}(P)$ . By Proposition 7  $\psi^* = 1$  iff  $x^* \in P$ . Assume  $x^* \in \text{aff}(P) \setminus P$ . Since  $0 \in \text{relint}(P)$ , then  $\text{rec}(P) \subseteq P$  and since  $x^* \notin P$ , case 3. of Proposition 7 applies and  $\psi^* > 1$ . Therefore  $|\psi^* - 1| = \psi^* - 1$ . Now 2. follows from Proposition 8.  $\square$

### 3.1 Computing $\text{aff}(P)$ and $x \in \text{relint}(P)$

Theorem 9 requires that  $0 \in \text{relint}(P)$ . This poses the following question:

*Let  $Ax \leq b$  be a system of inequalities that defines polyhedron  $P$ . Determine whether  $P = \emptyset$  and if  $P \neq \emptyset$ , compute  $\text{aff}(P)$  and  $\hat{x} \in \text{relint}(P)$ .*

When  $\hat{x} \in \text{relint}(P)$  is found, we can apply a linear transformation that maps  $\hat{x}$  into 0. Freund, Roundy and Todd [16] show that this question can be answered by solving a single LP:

**Proposition 10** *Given  $P = \{x \in R^n : Ax \leq b\}$ , consider the following LP:*

$$\begin{aligned} \max \quad & \mathbf{1}t \\ & Ay + t \leq b\lambda \\ 0 \leq \quad & t \leq \mathbf{1} \\ & \lambda \geq 1 \end{aligned} \tag{10}$$

1. The LP (10) is bounded and is feasible if and only if  $P \neq \emptyset$ .

2. Let  $\begin{pmatrix} \hat{y} \\ \hat{t} \\ \hat{\lambda} \end{pmatrix}$  be an optimal solution to (10). Then  $\hat{x} := \frac{\hat{y}}{\hat{\lambda}} \in \text{relint}(P)$ , and  $\hat{t} \in \{0, 1\}^m$ .

Furthermore the  $i^{\text{th}}$  constraint in the system  $Ax \leq b$  defines an improper face of  $P$  if and only if  $\hat{t}_i = 0$ .

*Proof.* 1. is straightforward. We prove 2.. Let  $\begin{pmatrix} \hat{y} \\ \hat{t} \\ \hat{\lambda} \end{pmatrix}$  be an optimal solution to (10). Since  $\hat{t} \geq 0$ , we have that  $\hat{x} := \frac{\hat{y}}{\hat{\lambda}} \in P$  and since  $\mathbf{1}t$  is maximized, then  $\hat{t}_i = 0$  if and only if  $\hat{x}$  satisfies the  $i^{\text{th}}$  constraint in  $Ax \leq b$  at equality. Furthermore, since  $\lambda$  is unbounded from above and  $0 \leq t \leq \mathbf{1}$ , then  $\hat{t}_i = 1$  if and only if  $\hat{x}$  satisfies the  $i^{\text{th}}$  constraint with strict inequality. Therefore the LP (10) finds a point  $\hat{x} \in P$  that satisfies strictly the largest number of inequalities in  $Ax \leq b$ , namely a point in  $\text{relint}(P)$ .  $\square$

There are alternative linear programming methods that solve the above problem. These methods typically find a strictly complementary solution or need the computation of a tolerance  $\varepsilon$  that depends on the system defining  $P$ , see e.g. [19]. The LP (10) avoids these issues.

**Remark 11** *If  $P := \text{proj}_x(Q)$  and  $Q := \{(x, y) : Ax + By \leq d\}$ , then applying the LP (10) to the system defining  $Q$ , one gets a point  $(\hat{x}, \hat{y}) \in \text{relint}(Q)$  and a system  $A^=x + B^=y = d^=$  of equations defining  $\text{aff}(Q)$ . Then  $\hat{x} \in \text{relint}(P)$  and  $\text{aff}(P) = \{x : uA^=x = ud^= \forall u \in N(B)\}$ , where  $N(B)$  is a basis of the space  $\{u : uB = 0\}$ .*

Furthermore given  $x^*$ , one can decide if  $x^* \in \text{aff}(P)$  or find an improper face of  $P$  that is violated by  $x^*$  by checking  $uA^=x^* = ud^= \forall u \in N(B)$ . Therefore by solving a single LP to determine  $\text{aff}(Q)$  and eventually finding a basis of a linear space (when  $Q \neq P$ ) one can check whether  $x^* \in \text{aff}(P)$ .

### 3.2 Obtaining a facet-inducing inequality almost surely

**Proposition 12** *When the LP (8) admits a finite optimum  $\psi^* \neq 0, 1$ , then an optimal solution of the LP (7) almost surely defines a facet-inducing inequality when  $x^* \in \text{aff}(P) \setminus P$  and almost surely defines an improper face of  $P$  when  $x^* \notin \text{aff}(P)$ .*

*Proof.* As  $\psi^* \neq 0, 1$ ,  $x^* \in \text{cone}(P) \setminus (P \cup \{0\})$ , so case 3. of Proposition 7 applies. Let  $y^* := \frac{x^*}{\psi^*}$ . Since for the Lebesgue measure restricted to the boundary of  $P$  the set of all faces with dimension at most  $\dim(P) - 2$  is negligible when  $P$  is full-dimensional, and the set of all faces with dimension at most  $\dim(P) - 1$  is negligible when  $P$  is not full-dimensional, a random point on the boundary of  $P$  almost surely lies in the relative interior of a facet of  $P$  when  $P$  is full-dimensional and lies in the relative interior of  $P$  when  $P$  is not full-dimensional.

By Proposition 8  $\left\{ \begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix} \in \Gamma^{\langle \rangle}(P) : \gamma y^* = \gamma_0 \right\}$  is the optimal face of the LP (7). Therefore when  $x^* \notin \text{aff}(P) \cup \{0\}$ , almost surely  $y^* \in \text{relint}(P)$  and  $F$  contains vectors  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix}$  that represent improper faces of  $P$  that are violated by  $x^*$ .

When  $x^* \in \text{aff}(P) \setminus (P \cup \{0\})$ , we may assume  $P$  full-dimensional. Almost surely  $y^* \in \text{relint}(f)$ , where  $f$  is a facet of  $P$  and in this case  $F$  contains all vectors  $\begin{pmatrix} \gamma \\ \gamma_0 \end{pmatrix}$  that represent  $f$ . □

Given polyhedron  $P$  and  $x^*$ , the *penumbra* of  $P$  with respect to  $x^*$  is the set

$$\text{Sh}(P, x^*) = \{x^* + \lambda(P - x^*), \lambda \geq 1\}. \quad (11)$$

The *reverse cone* of  $P$  with respect to  $x^*$  is the set

$$\text{Rcone}(P, x^*) = \{x^* - \lambda(P - x^*), \lambda \geq 0\}. \quad (12)$$

**Proposition 13** *Given a nonempty polyhedron  $P$  and  $x^* \notin P$ , if  $0 \in \text{relint}(\text{Sh}(P, x^*)) \cup \text{relint}(\text{Rcone}(P, x^*))$ , then the LP (8) is feasible and  $\psi^* \neq 0, 1$ .*

*Proof.* Assume  $0 \in \text{relint}(\text{Sh}(P, x^*))$ . Then  $0 = x^* + \lambda(y - x^*)$  where  $y \in P$  and  $\lambda > 1$ . Let  $\psi := \frac{\lambda}{\lambda - 1}$ . Then  $\psi > 1$  and  $y = \frac{x^*}{\psi} \in P$ . Therefore  $x^* \in \text{cone}(P)$  and, since  $x^* \notin P \cup \{0\}$ , case 3. of Proposition 7 applies. When  $0 \in \text{relint}(\text{Rcone}(P, x^*))$ , the proof is similar. □

In Figure 1, we illustrate the different regions arising in Proposition 13 and the behavior of the LP (6) as a function of the position of  $0$ .

Note that the definition of penumbra and reverse cone depend on the relative position of  $x^*$  with respect to  $P$  and Proposition 13 requires  $0 \in \text{relint}(\text{Sh}(P, x^*)) \cup \text{relint}(\text{Rcone}(P, x^*))$ . On the other hand Theorem 9 shows that when  $0 \in \text{relint}(P)$  and  $x^* \notin P$ , the separation LP(6) almost surely returns an inequality that is violated by  $x^*$  and it either induces a facet or an improper face of  $P$ .

Returning to Example 1, remark that  $0 \in \text{relint}(\text{Sh}(P, x^*))$ . The line from  $0$  to  $x^*$  traverses the boundary of  $P$  in the interior of the facet  $-x_1 + 4x_2 \leq 1$  of  $P^{(\pi, \pi_0)}$ , so by Proposition 13,  $\psi^* > 1$  and the optimal solution of the LP (6) generates this facet.

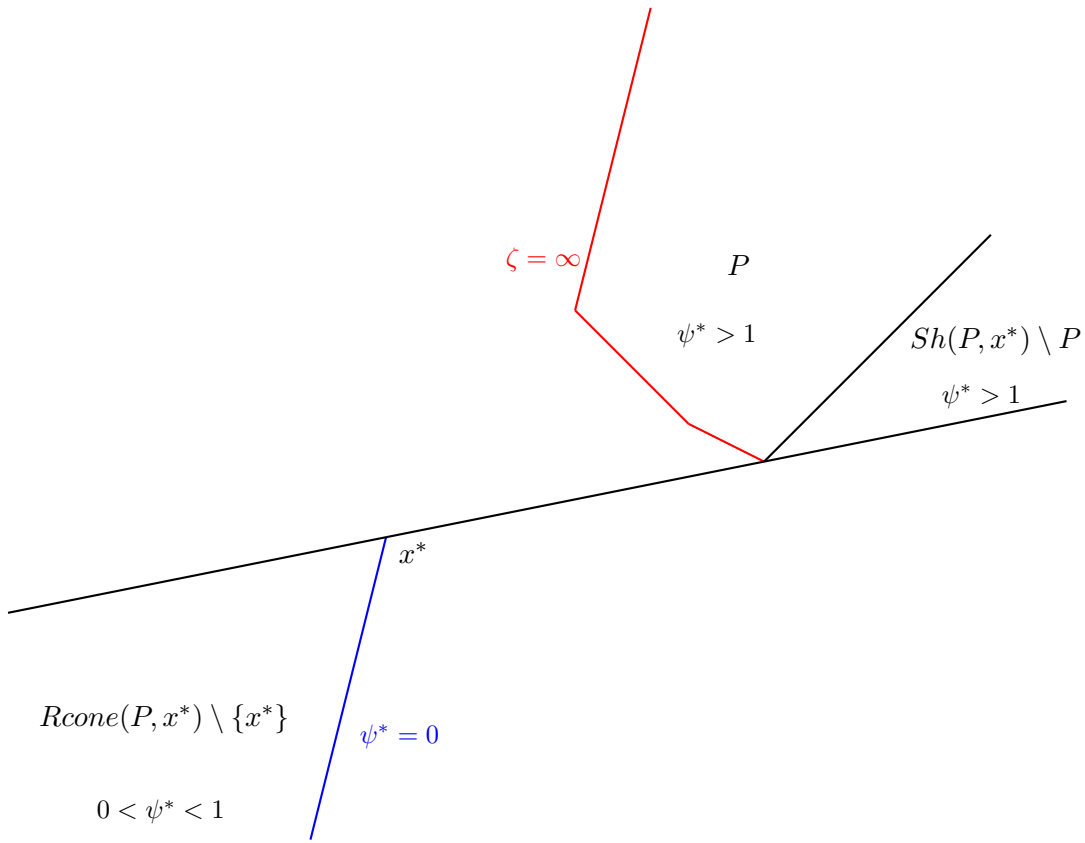


Figure 1: Placement of the Origin:  $x^* \notin P$

## 4. Applications in integer programming

Here we present the potential interest of our approach in cutting plane methods for solving structured and unstructured integer programs.

- For 0-1 integer or mixed integer programs, one approach is to use cutting planes based on split disjunctions (or more generally unions of polyhedra) of which an example was presented above.
- For uncapacitated fixed charge network flow problems, there is multi-commodity extended formulation that can be used to generate cutting planes in the space of the original arc variables.
- Mixed integer programs with a certain structure can be tackled by Benders' method which can be viewed as a cutting plane algorithm that exploits an extended formulation of the problem.

For each application we consider the form taken by the separation LP(6). For the first two we explicitly present a small example, along with some observations from solving a few small instances.

### 4.1 Split disjunctions

Given polyhedron  $P = \{x : Ax \leq b\}$  and  $(\pi, \pi_0) \in \mathbb{Z}^{n+1}$ , let  $P^0$ ,  $P^1$  and  $P^{(\pi, \pi_0)}$  be defined as in Section 2.1. Given  $x^* \notin P^{(\pi, \pi_0)}$ , when  $0 \in \text{relint}(P^{(\pi, \pi_0)})$  by Theorem 9, the LP (6) returns an inequality that almost surely either induces an improper face or a facet of  $P^{(\pi, \pi_0)}$ .

As this problem is repeatedly solved in a cutting plane algorithm, and at each iteration  $P$  and  $(\pi, \pi_0)$  change, one wants to find a point  $\hat{x}$  that can be mapped into 0 and does not depend on the current  $P$  and  $(\pi, \pi_0)$ , so that the LP (10) is not solved many times.

When  $\text{conv}(P \cap \mathbb{Z}^n)$  is not a face of  $P$ , the ideal choice is  $\hat{x} \in \text{relint}(\text{conv}(P \cap \mathbb{Z}^n))$  and for some structured problems such a point can be found by inspection, without solving the LP (10). For instance, this is easily achieved for a polytope  $P$  of the submissive type, such as the knapsack polytope. As 0 and the  $n$  unit vectors typically belong to  $P \cap \mathbb{Z}^n$ , we have that  $\hat{x} := (\frac{1}{2}, \dots, \frac{1}{2}) \in \text{relint}(\text{conv}(P \cap \mathbb{Z}^n))$ .

When such a point  $\hat{x}$  is found and mapped into 0, the separation LP (6) for  $P^{(\pi, \pi_0)}$  becomes

$$\begin{aligned} \zeta = \max \gamma(x^* - \hat{x}) & - \gamma_0 \\ uA + u_0\pi & = \gamma \\ vA - v_0\pi & = \gamma \\ u(b - A\hat{x}) + u_0(\pi_0 - \pi\hat{x}) & \leq \gamma_0 \\ v(b - A\hat{x}) - v_0(\pi_0 + 1 - \pi\hat{x}) & \leq \gamma_0 \\ -1 & \leq \gamma_0 \leq 1 \\ u, u_0, v, v_0 & \geq 0 \end{aligned}$$

This LP is used to treat the following example.

**Example 2** (*Example 2 from Fischetti et al. [13]*).

$\max\{x_1 + 2x_2 : x \in P \cap \mathbb{Z}^2\}$  where  $P \subset [0, 1]^2$  is the polytope:

$$\begin{aligned} -2x_1 + 2x_2 & \leq 1 \\ 2x_1 + 2x_2 & \leq 3 \\ -4x_1 - 4x_2 & \leq -3 \\ x_1 & \leq 1 \\ x_2 & \leq 1 \\ -x_1 & \leq 0 \\ -x_2 & \leq 0 \end{aligned}$$

The unique optimal solution of the LP relaxation is  $x^* = (0.5, 1)$ . Let  $(\pi, \pi_0) = ((1, 0), 0)$ . In this case  $P_0$  and  $P_1$  are the faces of  $P$  defined by the inequalities  $x_1 \geq 0$  and  $x_1 \leq 1$  respectively. Since  $P_0 = \emptyset$  and  $P_1 = \{x : x_1 = 1, 0 \leq x_2 \leq 1/2\}$  we have that  $P^{(\pi, \pi_0)} = \text{conv}(P_0 \cup P_1) = P_1$ .

Consider the transformation that maps  $\hat{x} = (1, 0.25) \in \text{relint}(P^{(\pi, \pi_0)})$  into 0. Since  $(0.5, 1) \notin \text{aff}(P^{(\pi, \pi_0)})$ , by Theorem 9, the separation LP (6) is unbounded and the unique unbounded ray produces the inequality  $-x_1 \leq -1$  that defines  $\text{aff}(P^{(\pi, \pi_0)})$ .

Consider the penumbra  $Sh(P^{(\pi, \pi_0)}, x^*)$  and the linear transformation that maps  $\hat{x} = (1.01, 0.25) \in \text{relint}(Sh(P^{(\pi, \pi_0)}, x^*))$  into 0. By Proposition 13, the LP (6) has a finite optimal solution and the optimal face produces the inequality  $-x_1 \leq -1$ . The same happens when  $\hat{x} = (0, 1.75) \in \text{relint}(R\text{cone}(P, x^*))$  is mapped into 0.

On the other hand, if 0 is not changed, the LP (6) is unbounded and the unbounded ray produces the inequality  $-x_1 + 0.5x_2 \leq 0$  that is valid, but not facet defining for  $P^{(\pi, \pi_0)}$ .

We carried out tests on small instances of 0–1 packing problems with up to 40 variables and compared our strategy with the normalizations proposed in the literature (mainly  $u_0 + v_0 = 1$  and  $\mathbf{1}u + u_0 + \mathbf{1}v + v_0 = 1$ ). The number of iterations is considerably reduced and one can run many iterations without encountering numerical problems. The normalizations proposed in

the literature typically fail to solve these instances with a pure cutting plane method, while our strategy does.

## 4.2 Multi-commodity reformulation of fixed charge network flow

We consider a single source uncapacitated network flow problem for which there is a multicommodity extended formulation, see Rardin and Choe [24]. We present the original and extended formulations, state a known result concerning the structure of the facets, demonstrate in detail how to derive the separation LP (6) and then present a small example indicating the behavior of the facet separation algorithm.

Given a network  $D = (V, A)$  where  $V = \{0, 1, \dots, n\}$  and  $A = \{1, \dots, m\}$ , source node 0 and demands  $b_i \geq 0$  for  $i \in V \setminus \{0\}$  where  $b_0 = -\sum_{i \in V \setminus \{0\}} b_i$ , and for each  $e \in A$  we are given a unit flow cost  $p_e$  and a fixed set-up cost  $f_e$ . The Fixed Charge Network Flow Problem is to find a feasible flow of minimum cost (i.e. flow cost plus set-up cost).

The following is a mixed integer programming formulation. We define  $x_e$  to be the flow value in arc  $e$ ,  $y_e$  to be the binary variable indicating if arc  $e$  carries positive flow.

$$\begin{aligned} \min \quad & \sum_{e \in A} p_e x_e + \sum_{e \in A} f_e y_e \\ & \sum_{e: h(e)=i} x_e - \sum_{e: t(e)=i} x_e = b_i \quad i = 1, \dots, n \\ & x_e \leq -b_0 y_e \quad e = 1, \dots, m \\ & x \in \mathbb{R}_+^m, y \in \{0, 1\}^m \end{aligned} \tag{13}$$

To obtain the multi-commodity reformulation, let  $K = \{i \in V \setminus \{0\} : b_i > 0\}$  be the set of terminals. We represent the flow as the sum of  $|K|$  distinct flows from 0 to  $k$  for  $k \in K$ . We define the variable  $w_e^k$  to be the flow in arc  $e$  with destination node  $k$ . Assuming  $p_e, f_e \geq 0$  for all  $e \in A$ , it can be shown that the following is an equivalent MIP formulation:

$$\begin{aligned} \min \quad & \sum_{e \in A} p_e x_e + \sum_{e \in A} f_e y_e \\ & \sum_{e: h(e)=i} w_e^k - \sum_{e: t(e)=i} w_e^k = 0 \quad i = 1, \dots, n, k \in K, k \neq i \\ & \sum_{e: h(e)=k} w_e^k - \sum_{e: t(e)=k} w_e^k = b_k \quad k \in K \\ & w_e^k \leq b_k y_e \quad e = 1, \dots, m, k \in K \\ & \sum_k w_e^k \leq x_e \quad e = 1, \dots, m, \\ & w \in \mathbb{R}_+^{m(|K|)}, y \in \mathbb{Z}^m. \end{aligned} \tag{14}$$

Let  $Q$  be the polyhedron defined by the constraints of (14) by dropping the integrality requirement on  $y$  and let  $P = \{(x, y) : \exists w \text{ s.t. } (x, y, w) \in Q\}$ .

It is known that the linear programming relaxation of (13) is in general considerably weaker than that obtained from solving the linear program over  $P$ . In particular for uncapacitated lot-sizing with or without backlogging [23], that is a special case of the fixed charge network flow problem,  $P$  is known to provide the convex hull of the feasible solutions of (13). Though a complete description of  $P$  in the original space is not known in general, the following result of Rardin and Wolsey [25] provides information on the structure of the facet-defining inequalities of  $P$ .

**Proposition 14** *Every nontrivial facet defining inequality for  $P$  can be written in the form:  $\mu^x x + \mu^y y \geq \mu_0$  where*

- $\mu_0 = \sum_{k \in K} \alpha_k b_k > 0$  with  $\alpha \in \mathbb{Z}_+^{|K|} \setminus \{0\}$



- for all  $k \in K$  and  $i = 1 \dots \alpha_k$ , there a directed cut (dicut)  $C^{i,k} \subset A$  separating node 0 and node  $k$  that is partitioned into two sets  $C_x^{i,k}, C_y^{i,k}$ , so that either the variable  $x_e$  or the variable  $y_e$  is associated to each arc  $e$  of  $C^{i,k}$
- for the  $y$  variables, the coefficient  $\mu_e^y = \sum_{k \in K} b_k |\{i : C_y(i,k) \ni e\}| \geq 0$
- for the  $x$  variables, the coefficient  $\mu_e^x = \max_k |\{i : C_x(i,k) \ni e\}| \geq 0$ .

### The separation LP for fixed charge network flow

Typically a point  $(\hat{x}, \hat{y}, \hat{w})$  in the relative interior of  $Q$  is easy to construct. Let  $\hat{w}$  be a feasible flow in which  $b_k > \hat{w}_e^k > 0$  unless  $w_e^k = 0$  or  $w_e^k = b_k$  for all feasible solutions in which case  $\hat{w}_e^k = 0$  or  $\hat{w}_e^k = b_k$  respectively. Set  $\hat{x}_e > \sum_k \hat{w}_e^k$  for all  $e$ ,  $\hat{y}_e > 1$  if  $w_e^k = b^k$  for some  $k$  and  $\hat{y}_e = 1$  otherwise.

By applying a linear transformation that maps  $(\hat{x}, \hat{y}, \hat{w})$  into 0, the constraints of  $Q$  become:

$$\begin{aligned}
\sum_{e:h(e)=i} \tilde{w}_e^k - \sum_{e:t(e)=i} \tilde{w}_e^k &= - \sum_{e:h(e)=i} \hat{w}_e^k + \sum_{e:t(e)=i} \hat{w}_e^k \quad i = 1, \dots, n, k \in K, & \mu_i^k \\
\sum_{e:h(e)=k} \tilde{w}_e^k - \sum_{e:t(e)=k} \tilde{w}_e^k &= b_k - \sum_{e:h(e)=k} \hat{w}_e^k + \sum_{e:t(e)=k} \hat{w}_e^k \quad k \in K & \mu_k^k \\
\tilde{w}_e^k - b_k \tilde{y}_e &\leq b_k \hat{y}_e - \hat{w}_e^k \quad e = 1, \dots, m, k \in K & v_e^k \\
\sum_k \tilde{w}_e^k - \tilde{x}_e &\leq - \sum_k \hat{w}_e^k + \hat{x}_e \quad e = 1, \dots, m, & \pi_e \\
-\tilde{w}_e^k &\leq \hat{w}_e^k \quad e = 1, \dots, m, k \in K & \phi_e^k \\
\tilde{w} &\in \mathbb{R}^{m(|K|)}, \tilde{y} \in \mathbb{R}^m, \tilde{x} \in \mathbb{R}^m
\end{aligned}$$

where  $x_e = \tilde{x}_e + \hat{x}_e$ ,  $w_e^k = \tilde{w}_e^k + \hat{w}_e^k$ ,  $y_e = \tilde{y}_e + \hat{y}_e$ .

The separation LP (6) now takes the form:

$$\begin{aligned}
\max \sum_e (\gamma_e^x (x_e^* - \hat{x}_e) + \gamma_e^y (y_e^* - \hat{y}_e)) & \quad -\gamma_0 \\
\gamma_e^x + \pi_e &= 0 \quad e \in 1 \dots, m \\
\gamma_e^y + \sum_{k \in K} b_k v_e^k &= 0 \quad e \in 1 \dots, m \\
\pi_e + v_e^k - \phi_e^k + \sum_{i:h(e)=i} \mu_i^k - \sum_{i:t(e)=i} \mu_i^k &= 0 \quad e \in 1 \dots, m, k \in K \quad (15) \\
\gamma_0 - \sum_k \mu_k^k b_k - \sum_{i,k} \mu_i^k (- \sum_{e:h(e)=i} \hat{w}_e^k + \sum_{e:t(e)=i} \hat{w}_e^k) \\
\sum_e \pi_e (- \sum_k \hat{w}_e^k + \hat{x}_e) - \sum_{e,k} v_e^k (b_k \hat{y}_e - \hat{w}_e^k) - \sum_{k,e} \phi_e^k \hat{w}_e^k &\geq 0 \\
-1 \leq \gamma_0 &\leq 1 \\
\gamma^x, \gamma^y \in \mathbb{R}^m, \gamma_0 \in \mathbb{R}^1, \mu \in \mathbb{R}^{nK}, v, \pi, \phi &\geq 0
\end{aligned}$$

As we always work with the same polyhedron, the set of constraints of the LP (15) does not change at every iteration. The only change is in the objective function that is defined by the current point to be separated. So one only needs to compute a point in  $\text{relint}(P)$  once.

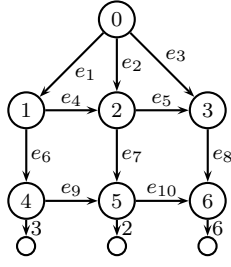


Figure 2: The Fixed Charge Network

### A Small Example in Detail

**Example 3** Here we consider an instance with  $n = 6$  nodes other than the root node 0 and  $m = 10$  arcs. The network is shown in Figure 2 along with the demands at nodes 4, 5, 6, corresponding to commodities 1, 2, 3. The variable costs are  $p = (4, 2, 2, 1, 1, 2, 2, 2, 1, 1)$  and the fixed costs  $f = (25, 42, 17, 36, 18, 34, 25, 48, 37, 46)$ .

Consider the point

$e$	1	2	3	4	5	6	7	8	9	10
$\hat{y}_e$	1.1	1	1	1	1	1.1	1	1	1	1
$\hat{w}_e^1$	3	0	0	0	0	3	0	0	0	0
$\hat{w}_e^2$	3/2	1/2	0	1/2	0	1	1	0	1	0
$\hat{w}_e^3$	3	3/2	3/2	3/2	3/2	3/2	3/2	3	3/2	3

with  $\hat{x}_e = \sum_k \hat{w}_e^k + 0.1$ .

Since  $(\hat{x}, \hat{y}) \in \text{relint}(P)$ , we apply the linear transformation that maps this point into 0. Since the separation LP is solved with  $(x^*, y^*) \in \text{aff}(P)$ , it follows from Theorem 9 that the LP(6) is always bounded.

Below we show the results obtained using this facet separation LP. The optimal value of the LP relaxation of (13) is 113.727. The cuts added at each iteration and the resulting LP values are shown below. The algorithm terminates after generating eight inequalities that are facet inducing for  $P$  (and therefore satisfy the conditions of Proposition 14).

Cut	Inequality		LPval
1	$y_6$	$\geq 1$	138.455
2	$y_1$	$\geq 1$	156.636
3	$6y_8 + x_{10}$	$\geq 6$	178.455
4	$y_7 + y_9$	$\geq 1$	195.636
5	$x_8 + 6y_{10}$	$\geq 6$	198.909
6	$6y_3 + x_5 + x_{10}$	$\geq 6$	206.636
7	$y_2 + y_4 + y_9$	$\geq 1$	216.455
8	$2y_7 + x_9$	$\geq 2$	217.

We have solved other small instances of fixed charge network flow problem (16 nodes and 27 arcs). Here Proposition 14 provides a necessity check on whether the cuts generated cuts are

facet-defining. Nearly all the cuts generated were active at the optimum, so that there was no sign of a tailing-off effect. Comparing with cut generation using three well-known normalizations and without moving the origin, the number of iterations was typically reduced by a half or more.

### 4.3 Benders' Algorithm

Benders' algorithm applies to problems of the form

$$\zeta = \max\{cx + hy : Ax + By \leq d, x \in X\} \quad (16)$$

where typically in mixed integer programming  $X \subseteq \mathbb{Z}^n$ . Let  $Q := \{(\eta, x, y) : \eta - cx - hy \leq 0, Ax + By \leq d\}$  and  $P := \{(\eta, x) : \exists y \text{ s.t. } (\eta, x, y) \in Q\}$ . Then (16) is equivalent to

$$\max\{\eta : (\eta, x) \in P, x \in X\}.$$

Given  $(\eta^*, x^*)$ , Benders' approach consists in checking whether  $(\eta^*, x^*) \in P$  and if not, finding a violated inequality.

Given a point  $(\hat{\eta}, \hat{x}, \hat{y}) \in \text{relint}(Q)$ , then  $(\hat{\eta}, \hat{x}) \in \text{relint}(P)$  and we can apply the linear transformation that maps this point into 0. The separation LP (6) becomes

$$\begin{aligned} \max \quad & \gamma^\eta(\eta^* - \hat{\eta}) + \gamma^x(x^* - \hat{x}) - \gamma_0 \\ & \gamma^\eta = u_0 \\ & \gamma^x = uA - u_0c \\ & uB - u_0h = 0 \\ \gamma_0 \geq & u_0(-\hat{\eta} + c\hat{x} + h\hat{y}) + u(d - A\hat{x} - B\hat{y}) \\ & -1 \leq \gamma_0 \leq 1 \\ & u_0, u \geq 0 \end{aligned} \quad (17)$$

Since  $(\hat{\eta}, \hat{x}) \in \text{relint}(P)$ , given  $(\eta^*, x^*) \notin P$ , by Theorem 9, we have the following possibilities:

- The LP (17) is bounded if and only if  $(\eta^*, x^*) \in \text{aff}(P)$ . In this case  $\gamma_0 = 1$  and an optimal solution almost surely induces a violated inequality that is facet-inducing for  $P$ . Furthermore  $u_0(-\hat{\eta} + c\hat{x} + h\hat{y}) + u(d - A\hat{x} - B\hat{y}) = 1$ . There are two cases:
  - $-u_0 = 0$ . Then  $x^* \in \text{aff}(P') \setminus P'$ , where  $P' := \{x : \exists y \text{ s.t. } Ax + By \leq d\}$  and one obtains the inequality  $uA(x - \hat{x}) \leq 1$ . As  $u(d - A\hat{x} - B\hat{y}) = 1$ , this yields a feasibility cut  $uAx \leq u(d - B\hat{y})$ .
  - $-u_0 > 0$ . In this case the inequality is  $u_0(\eta - \hat{\eta}) + (uA - u_0c)(x - \hat{x}) \leq 1$ , and again as  $u_0(-\hat{\eta} + c\hat{x} + h\hat{y}) + u(d - A\hat{x} - B\hat{y}) = 1$ , this is an optimality cut  $\eta \leq cx + h\hat{y} + \frac{u}{u_0}(d - Ax - B\hat{y})$ .
- LP (17) is unbounded if and only if  $(\eta^*, x^*) \notin \text{aff}(P)$ . In this case  $\gamma_0 = 0$  and an optimal solution produces a violated inequality that induces an improper face of  $P$ . As  $\text{aff}(P) = \{(\eta, x) : x \in \text{aff}(P')\}$ , this inequality is of the form  $uA(x - \hat{x}) = 0$ .

It is well known that if one applies Benders' algorithm to  $\max\{cx + hy : Ax + By \leq d\}$  as described above, the resulting linear programming relaxation at each iteration is the the dual

of the restricted Master problem obtained when applying Dantzig-Wolfe decomposition to the linear program

$$\min\{ud : uA = c, uB = h, u \geq 0\}.$$

This suggests that the separation LP(17) may be a good option for the column generation subproblem in Dantzig-Wolfe.

## 5. Final Remarks

It seems clear that the normalization  $-1 \leq \gamma_0 \leq 1$  together with a linear transformation that maps 0 into a point in the relative interior of the feasible region is a good strategy. This raises some questions:

- The separation LP (6) provides a violated inequality that induces a facet or an improper face when the dual is feasible and  $\psi^* > 0$ . The dual can be reformulated as in (9), which is a one dimensional LP. Can one devise efficient methods that take advantage of this?
- As solving a large separation LP at each iteration is costly, several methods, such as variable fixing and cut-lifting, may allow one to reduce the size of such LPs
- Our examples indicate that our strategy reduces the tailing-off effect and the numerical stability that typically affect pure cutting-plane algorithms. However this clearly has to be substantiated by further testing.

**Acknowledgement** The authors gratefully acknowledge that this work started at the University of Bonn when they were participants in the Hausdorff Trimester Program on "Combinatorial Optimization".

## References

- [1] K. Andersen, Q. Louveaux, R. Weismantel and L.A. Wolsey, Cutting Planes from two rows of a simplex tableau, *Proceedings of IPCO XII*, Ithaca, New York, *Lecture Notes in Computer Science 4513* 1-15 (2007).
- [2] E. Balas, Disjunctive programming: properties of the convex hull of feasible points, *GSIA Management Science Research Report MSRR 348*, Carnegie Mellon University (1974), published as invited paper in *Discrete Applied Mathematics 89* 1-44 (1998).
- [3] E. Balas, S. Ceria and G. Cornuéjols, A lift-and-project cutting plane algorithm for mixed 0-1 programs, *Mathematical Programming 58* 295-324 (1993).
- [4] E. Balas and M. Perregaard, Lift-and-project for mixed 0-1 programming: recent progress, *Discrete Applied Mathematics 123* 129-154 (2002).
- [5] E. Balas and M. Perregaard, A precise correspondence between lift-and-project cuts, simple disjunctive cuts and mixed integer Gomory cuts for 0-1 programming, *Mathematical Programming B 94* 221-245 (2003).

- [6] J.F. Benders, Partitioning procedures for solving mixed variables programming problems, *Numerische Mathematik* 4 238-252 (1962).
- [7] A. Ben-Tal, A. Nemirovski, *Lectures on Modern Convex Optimization*, MPS-SIAM Series on Optimization, SIAM, Philadelphia (2001).
- [8] P. Bonami, Etude et mise en oeuvre d'approches polyédriques pour la résolution de programmes en nombres entiers ou mixtes généraux, *PhD thesis, Université de Paris 6* (2003).
- [9] F. Cadoux, Computing deep facet-defining disjunctive cuts for mixed-integer programming *Mathematical Programming A* 122 197-223 (2010).
- [10] M. Conforti, G. Cornuéjols and G. Zambelli, *Integer Programming*, Springer, (2014).
- [11] G. Cornuéjols and C. Lemaréchal, A convex-analysis perspective of disjunctive cuts, *Mathematical Programming A* 106 567-586 (2006).
- [12] H. Crowder, E.L. Johnson and M.W. Padberg, Solving large-scale zero-one linear programming problems, *Operations Research* 31 803-834 (1983).
- [13] M. Fischetti, A. Lodi, A. Tramontani, On the separation of disjunctive cuts, *Mathematical Programming* 128 205-230 (2011).
- [14] M. Fischetti and D. Salvagnin, An in-out approach to disjunctive optimization, *CPAIOR, Lecture Notes in Computer Science, Vol. 6140* 136-140 (2010).
- [15] M. Fischetti, D. Salvagnin and A. Zanette, A note on the selection of Benders' cuts, *Mathematical Programming B* 124 175-182 (2010).
- [16] R. M. Freund, R. Roundy and M. J. Todd, Identifying the Set of Always-Active Constraints in a System of Linear Inequalities by a Single Linear Program *Sloan W.P. No. 1674-85* (1985).
- [17] J. B. Hiriart-Urruty and C. Lemaréchal *Fundamentals of Convex Analysis* Springer Verlag, Grundlehren Text Editions, (2001).
- [18] T. L. Magnanti and R. T. Wong, Accelerating Benders Decomposition: Algorithmic Enhancement and Model Selection Criteria, *Operations Research* 29 464-484 (1981).
- [19] M. Mehdiloozad, K. Tone, R. Askarpour and M.B. Ahmadi, Finding a maximal element of a non-negative convex set through its characteristic cone: an application to finding a strictly complementary solution *Computational and Applied Mathematics, to appear* (2016).
- [20] G.L. Nemhauser and L.A. Wolsey, *Integer and Combinatorial Optimization*, Wiley, New York (1988).
- [21] N. Papadakos, Practical enhancements to the Magnanti-Wong method, *Operations Research Letters* 36 444-449 (2008).
- [22] M.W. Padberg and M. Grötschel, *Polyhedral Computations, Ch. 9* in: *The Traveling Salesman Problem*, E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan and D.B. Shmoys eds, Wiley and Sons Ltd. (1985).

- [23] Y. Pochet and L.A. Wolsey, Lot-sizing models with backlogging: Strong formulations and cutting planes, *Mathematical Programming* 40 317-335 (1988).
- [24] R.L. Rardin and U. Choe, Tighter relaxations of fixed charge network flow problems, *Report J-79-18, Industrial and Systems Engineering, Georgia Institute of Technology* (1979).
- [25] R. Rardin and L.A. Wolsey, Valid inequalities and projecting the multicommodity extended formulation for uncapacitated fixed charge network flow problems, *European J. of Operations Research* 71 95-109 (1993).