# A probabilistic analysis of the strength of the split and triangle closures

Amitabh Basu     Gérard Cornuéjols     Marco Molinaro

October 21, 2010

**Abstract**

In this paper we consider a relaxation of the corner polyhedron introduced by Andersen et al., which we denote by RCP. We study the relative strength of the split and triangle cuts of RCP's. Basu et al. showed examples where the split closure can be arbitrarily worse than the triangle closure under a 'worst-cost' type of measure. However, despite experiments carried out by several authors, the usefulness of triangle cuts in practice has fallen short of its theoretical strength.

In order to understand this issue, we consider two types of measures between the closures: the 'worst-cost' one mentioned above, where we look at the weakest direction of the split closure, and the 'average-cost' measure which takes an average over all directions. Moreover, we consider a natural model for generating random RCP's. Our first result is that, under the worst-cost measure, a random RCP has a weak split closure with reasonable probability. This shows that the bad examples given by Basu et al. are not pathological cases. However, when we consider the average-cost measure, with high probability both split and triangle closures obtain a very good approximation of the RCP. The above result holds even if we replace split cuts by the simple split or Gomory cuts. This gives an indication that split/Gomory cuts are indeed as useful as triangle cuts.

## 1 Introduction

Consider an IP in standard form:

$$\begin{aligned} &\min cy \\ &Ay = b \\ &y \geq 0 \\ &y \in \mathbb{Z}^d\ . \end{aligned} \tag{IP}$$

Suppose that $B$ is an optimal basis for the LP relaxation of (IP). Rewriting IP in tableaux form with respect to $B$ (i.e., pre-multiplying the system by $B^{-1}$) we obtain the equivalent system

$$\begin{aligned} &\min \bar{c}_N y_N \\ &y_B = \bar{b} - \bar{N} y_N \\ &y \geq 0 \\ &y \in \mathbb{Z}^d \end{aligned} \tag{IP'}$$

where $\bar{c} \geq 0$ due to the optimality of $B$.

In [1], Andersen et al. introduced a relaxation of (IP') which we call Relaxed Corner Polyhedron (RCP). This is a further weakening of the Corner relaxation [13] where: (i) the integrality constraints are dropped for the nonbasic variables and (ii) the non-negativity constraints are dropped for the basic variables. Rewriting in a different way, an RCP is a MIP of the form

$$\begin{aligned} &\min cs \\ &x = f + \sum_{j=1}^{n} r^j s_j \\ &s \geq 0 \\ &x \in \mathbb{Z}^m \end{aligned} \tag{RCP}$$

with $c \geq 0$.

An RCP is defined by the vectors $f, r^1, \ldots, r^n$ and the cost vector $c$. We call a tuple $\langle f, r^1, r^2, \ldots, r^n \rangle$ an *ensemble*. Given an ensemble $\mathcal{E}$ and cost vector $c$, we use $RCP(\mathcal{E}, c)$ to denote the corresponding RCP.

In this work, we are interested in comparing the cost of an optimal solution to $RCP(\mathcal{E}, c)$ against the cost of an optimal solution to some of its relaxations. In order to simplify things, we work over the projection onto the $s$-space; the crucial property given by the structure of RCP's is that the projection of any solution onto the $s$-space has the same cost as the original solution. We define $P(\mathcal{E})$ as the projection of the feasible region of $RCP(\mathcal{E}, c)$ onto the $s$-space. We also use $P_L(\mathcal{E})$ to denote the linear relaxation of $P(\mathcal{E})$.

**Intersection cuts.** Let $X$ be a closed convex set in $\mathbb{R}^m$ which: (i) contains $f$ in its interior and (ii) does not contain any integer point in its interior. The functional $\psi_X : \mathbb{R}^m \to \mathbb{R}$ is defined as

$$\psi_X(r) = \inf\{\lambda > 0 : f + \frac{r}{\lambda} \in X\}. \tag{1}$$

The inequality $\sum_{j=1}^{n} \psi_X(r^j) s_j \geq 1$ is valid for $P(\mathcal{E})$, since it is an intersection cut [2]. Moreover, it was shown in [8] that all minimal inequalities of $P(\mathcal{E})$ are of this form (for some $X$). Since these inequalities are constructed based on convex sets in $\mathbb{R}^m$, this gives a geometrical way of analyzing $P(\mathcal{E})$.

One important family of inequalities is derived from sets $X$ which are 'splits', that is, $X$ is of the form $\{x : b_1 \leq ax \leq b_2\}$. We call them split cuts. The split closure of $P(\mathcal{E})$, denoted by $S(\mathcal{E})$, is the intersection of all split cuts. We also consider cuts from *simple* splits, that is, splits of the form $\{x : \lfloor b \rfloor \leq e^i x \leq \lceil b \rceil\}$ where $b \in \mathbb{R} \setminus \mathbb{Z}$ and $e^i$ is the $i$th canonical vector in $\mathbb{R}^m$. We denote the closure of these cuts by $G(\mathcal{E})$. Finally, for the case $m = 2$, another important family is when $X$ is a triangle. The triangle closure is denoted by $T(\mathcal{E})$.

**Strength of relaxations.** In [7] the authors showed a family of RCP's whose split closure gets arbitrarily weaker than the triangle closure. In their work, a 'worst-case' type of measure is used to compare these two closures. Interestingly, split cuts for (IP) derived by taking into account all integrality constraints actually perform very well in practice [4]. The apparent strength of triangle cuts suggests that even stronger cuts for (IP) can be obtained from them.

Motivated by this observation, several authors recently started to test experimentally the effectiveness of the triangle cuts [12, 3, 6, 10]. Despite some positive results obtained by Balas and Qualizza in [3], the usefulness of the triangle cuts has been very far from those indicated by theory.

However, it is still not clear if the disappointing performance of triangle cuts is due to insufficient separation heuristics or is really an innate property of these cuts, although the latter seems more likely. Our goal in this paper is to further explore the relative strength of the split and triangle closure from a theoretical perspective.

**Our results.** Informally, the results presented in this paper can be described as follows. We consider two ways of comparing the closures $P(\mathcal{E})$, $G(\mathcal{E})$ and $T(\mathcal{E})$. Roughly, in the 'worst-cost' measure we look at the direction where $G(\mathcal{E})$ is weakest, while in the 'average-cost' measure we take an average over all directions. We show that the strength of the split closure is very dependent on which measure is used to perform the comparison.

First we consider the worst-cost measure and $m = 2$. We show that with reasonable probability the split closure of a random RCP is very weak when compared to the triangle closure. This generalizes the bad examples from [7] and shows that they are not as pathological as one could think. On the other hand, with respect to the average-case measure we show that, with high probability, both the split and the triangle closure are very good approximations of $P(\mathcal{E})$. In particular, this shows that the split and the triangle closure are very similar under the average-case measure. This gives a partial justification why triangle cuts seem to have a similar performance to split cuts in practice.

**Related work.** Two recent papers address the fundamental question of comparing the strengths of triangle and split cuts from a probabilistic point of view.

He et al. [14] use the same random model for generating RCP's, but a different measure to compare the strength of cuts, comparing the random coefficients of the inequalities induced by the randomness of the rays. Their analysis does not consider the important triangles of Type 3. Although the results cannot be directly compared, their paper also indicates that split cuts perform at least as well as some classes of triangles.

Del Pia et al. [15] base their analysis on the lattice width of the underlying convex set. They show that the importance of triangle cuts generated from Type 2 triangles (the same family which was considered in [7]) decreases with decreasing lattice width, on average. They also have results for triangles of Type 3 and for quadrilaterals.

Our approach is very different from these two papers.

## 2 Preliminaries

**Measures of strength.** Let $A$ and $B$ be convex relaxations of $P(\mathcal{E})$ such that $A, B \subseteq \mathbb{R}^n_+$. A closed, convex set $X \subseteq \mathbb{R}^n_+$ is said to be of blocking type if $y \geq x \in X$ implies $y \in X$. It is well-known that the recession cone of $P(\mathcal{E})$ is $\mathbb{R}^n_+$(see [9]) and hence $P(\mathcal{E})$, $A$ and $B$ are convex sets of blocking type. A traditional measure of strength for integer programs is the *integrality gap*, which compares the ratio of the minimization over the IP and its linear relaxation. More generally, we define the *gap* between $A$ and $B$ with respect to the cost vector $c$ as:

$$\mathrm{gap}(A, B, c) = \frac{\inf\{cs : s \in A\}}{\inf\{cs : s \in B\}}. \tag{2}$$

Notice that this value is greater than 1 if $A$ is stronger than $B$. We define the gap to be $+\infty$ if $A$ is empty or $\inf\{cs : s \in B\} = 0$.

Based on this idea, we can define the *worst-cost* measure between the two relaxations $A$ and $B$ as the worst possible gap over all non-negative cost vectors:

$$\mathrm{wc}(A, B) = \sup_{c \in \mathbb{R}^m_+} \{\mathrm{gap}(A, B, c)\} \tag{3}$$

$$= \sup_{c \in [0,1]^m} \{\mathrm{gap}(A, B, c)\}, \tag{4}$$

where the second equation follows from the fact that the ratios are preserved under positive scaling of the cost vectors. Note that for convex sets of blocking type, only non-negative cost vectors have bounded optimum, hence we will restrict ourselves to this case.

For any convex set $X$ of blocking type, define $\alpha X = \{\frac{x}{\alpha} : x \in X\}$. An equivalent definition for the worst-case measure used in [7] is the following: $\mathrm{wc}(A, B)$ is the amount that $A$ has to be "blown up" in order to contain $B$, namely $\mathrm{wc}(A, B) = \inf\{\alpha : \alpha A \supseteq B\}$. Therefore, if there is any direction/cost where $B$ is far away from $A$, the value of this measure becomes large. We prove this equivalence in the Appendix in Lemma 10.

Now we define another (more robust) measure of strength which tries to capture the average strength with respect to different costs. Consider a distribution $\mathcal{C}$ over vectors in $\mathbb{R}^m_+$. Then, the *average-cost* measure between $A$ and $B$ is defined by

$$\mathrm{avg}(A, B, \mathcal{C}) = \mathbb{E}_{c\sim\mathcal{C}}\left[\mathrm{gap}(A, B, c)\right]. \tag{5}$$

In the sequel we study the worst-cost and average-cost strength of the split and triangle closures for random RCP's. We define our model for random RCP's next.

**Random model.** Let $\mathcal{D}^m_n$ denote the distribution of ensembles $\langle f, r^1, \ldots, r^n\rangle$ where $f$ is picked uniformly from $[0, 1]^m$ and each of $r^1, \ldots, r^n$ is picked independently and uniformly at random from the set of rational unit vectors in $\mathbb{R}^m$. We make a note here that the rays in RCP can be assumed to be unit vectors, by suitably scaling the cost coefficients. In other words, given an ensemble $\mathcal{E}$ and a cost vectors $c$, there exists an ensemble $\mathcal{E}'$ and a cost vector $c'$ such that the optimal value of $\mathrm{RCP}(\mathcal{E}, c)$ equals to the optimal value of $\mathrm{RCP}(\mathcal{E}', c')$. Moreover, there exists an affine invertible affine transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ such that $P(\mathcal{E}) = A(P(\mathcal{E}'))$. Appendix states these observations in the form we need in the paper and provides rigorous proofs. Hence, in our model, we assume the rays are sampled from the set of rational unit vectors. When the dimension is 2 we write $\mathcal{D}_n$ for the distribution, omitting the superscript.

## 3 Worst-cost measure in $\mathbb{R}^2$

The main result of this section is that, for a significant fraction of the RCP's in the plane, $S(\mathcal{E})$ is significantly worse than $T(\mathcal{E})$ based on the worst-cost measure.

**Theorem 1.** *For any $\alpha \geq 1$ and $\beta \in [0, 1]$, a random ensemble $\mathcal{E} \sim \mathcal{D}_n$ satisfies*

$$\Pr\left(\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \alpha\right) \geq \left[1 - 2\left(1 - g(\frac{\beta}{4\alpha})\right)^n\right]\left[\frac{1-\beta}{\alpha} - \frac{1-\beta^2}{4\alpha^2}\right],$$

*where*

$$g(x) = \left(\frac{x}{0.75 - (2-\sqrt{2})x} - \frac{x}{1 - (2-\sqrt{2})x}\right).$$

Notice that this bound increases as $n$ grows. In the limit $n \to \infty$, and using the optimal choice $\beta \to 0$, the bound becomes $1/\alpha - 1/4\alpha^2$. To obtain an idea about the probabilities in the above theorem, Table 1 presents the bound obtained for different values of $n$ and $\alpha$.

The way to prove this result is to consider a particular (deterministic) ensemble $\langle f, r^1, r^2\rangle$ which is 'bad' for the split closure and show that it appears with significant probability in a random ensemble. We employ the following monotonicity property to transfer the 'badness' to the whole RCP. The proof appears at the end of Appendix A.

**Lemma 1.** *Consider an ensemble $\mathcal{E} = \langle f, r^1, \ldots, r^n\rangle$ and let $\mathcal{E}' = \langle f, r^{i_1}, r^{i_2}, \ldots, r^{i_k}\rangle$ be a subensemble of it. Then* $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \mathrm{wc}(T(\mathcal{E}'), S(\mathcal{E}'))$.

| $n$ | $\alpha$ | $\beta$ | Pr |
|---|---|---|---|
| 100 | 1.5 | 0.37 | 25.7% |
| 100 | 2 | 0.43 | 16.7% |
| 500 | 2 | 0.16 | 33.6 % |
| 500 | 3 | 0.22 | 21.3% |
| 1000 | 2 | 0.01 | 37.7% |
| 1000 | 3 | 0.14 | 25.0 % |
| 1000 | 4 | 0.17 | 18.2 % |
| $+\infty$ | 2 | 0 | 43.75 % |
| $+\infty$ | 4 | 0 | 30.56 % |

Table 1: Values of the bound of Theorem 1 for different values of $n$ and approximation factor $\alpha$. The value of $\beta$ in every entry was chosen empirically and attempts to optimize the bound.

### 3.1 A bad ensemble for the split closure

First, we introduce the following notation: Given a point $f$ and a ray $r$, we say that $f + r$ *crosses* a region $R \subseteq \mathbb{R}^n$ if there is $\lambda \geq 0$ such that $f + \lambda r \in R$.

In this part we will focus on ensembles $\mathcal{E} = \langle f, r^1, r^2 \rangle$ where $f \in (0,1)^2$, and $f+r^1$ and $f+r^2$ cross the open segment connecting $(0,0)$ to $(0,1)$. The high-level idea is the following. Suppose that $r^1$ an $r^2$ have $x_1$-value equal to -1 and consider a lattice-free triangle $T$ containing the points $f+r^1$ and $f+r^2$, and also containing $f$ in its interior. This triangle gives an inequality which is at least as strong as $s_1+s_2 \geq 1$, hence we have a lower bound of 1 for minimizing $s_1+s_2$ over the triangle closure $T(\mathcal{E})$. However, further assume that the angle between rays $r^1$ and $r^2$ is large. Then we can see that any split that contains $f$ in its interior will have a very large coefficient for either $s_1$ or $s_2$. More specifically, suppose that there is a large $M$ such that, for every inequality $\psi(r^1)s_1 + \psi(r^2)s_2 \geq 1$ coming from a split, we have $\max\{\psi(r^1), \psi(r^2)\} \geq M$. Then the point $(s_1, s_2) = (1/2M, 1/2M)$ satisfies every such inequality and hence is feasible for the split closure $S(\mathcal{E})$; this gives an upper bound of $2/M$ for minimizing $s_1 + s_2$ over the split closure. Then using the choice of $c = [1,1]$ in the maximization in (3) gives $\text{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq M/2$.

The following lemma is presented in Section 5.6.2 of [5] and formalizes the observation that if $r^1$ and $r^2$ are spread out then the split closure is weak. The proof is included in Appendix B.

**Lemma 2.** *Consider an ensemble $\mathcal{E} = \langle f, r^1, r^2 \rangle$ where $f = (f_1, f_2) \in (0,1)^2$, $r^1 = c_1(-1, t_1)$ and $r^2 = c_2(-1, t_2)$ with $c_1, c_2 \geq 0$ and $t_1 \geq t_2$. Moreover, assume that both $f + r^1$ and $f + r^2$ cross the left facet of the unit square. Then*

$$\min\{c_1 s_1 + c_2 s_2 : (s_1, s_2) \in S(\mathcal{E})\} \leq \frac{f_1(t_1 - t_2) + 1}{t_1 - t_2}.$$

**Corollary 1.** *Let $\mathcal{E}$ as in the previous lemma. Then*

$$\min\{c_1 s_1 + c_2 s_2 : (s_1, s_2) \in S(\mathcal{E})\} \leq \frac{2}{t_1 - t_2}.$$

This corollary follows by applying the next simple geometric fact to Lemma 2.

**Lemma 3.** *Consider a point $f = (f_1, f_2) \in (0,1)^2$ and rays $r^1 = (-1, t_1)$ and $r^2 = (-1, t_2)$, with $t_1 \geq t_2$. If $f + r^1$ and $f + r^2$ cross the left facet of the unit square then $f_1(t_1 - t_2) \leq 1$. Moreover, $f_1(t_1 - t_2) = 1$ iff $f + r^1$ crosses $(0,1)$ and $f + r^2$ crosses $(0,0)$.*

*Proof.* First notice that $f + f_1 r^1 = (0, f_2 + f_1 t_1)$ and using the crossing property we get that $f_2 + f_1 t_1 \leq 1$. Similarly, $f + f_1 r^2 = (0, f_2 + f_1 t_2)$, hence $f_2 + f_1 t_2 \geq 0$. Isolating $f_2$ in both inequalities and chaining them we obtain $f_1(t_1 - t_2) \leq 1$, obtaining the first part of the lemma. The second part follows from the fact that all the previous inequalities hold with equality iff $f + r^1$ crosses $(0,1)$ and $f + r^2$ crosses $(0,0)$. $\square$

Using Corollary 1 we establish the main lemma of this section, which exhibits bad ensembles for the split closure.

**Lemma 4.** *Consider an ensemble $\mathcal{E} = \langle f, r^1, r^2 \rangle$ where $f = (f_1, f_2) \in (0,1)^2$. Suppose that $f + r^1$ crosses the open segment connecting $(0, 1-\epsilon)$ and $(0,1)$ and that $f + r^2$ crosses the open segment connecting $(0,0)$ and $(0,\epsilon)$, for some $0 < \epsilon < 1$. Then $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq (1-2\epsilon)/2f_1$.*

*Proof.* Let $v^1 = (-1, t_1)$, $v^2 = (-1, t_2)$ and let $c_1, c_2 \geq 0$ be such that $r^1 = c_1 v^1$ and $r^2 = c_2 v^2$. By the assumptions on the rays, we have $t_1 \geq t_2$.

Consider the rays $\underline{v}^1 = (-1, \underline{t}_1)$ and $\underline{v}^2 = (-1, \underline{t}_2)$ such that $f + \underline{v}^1$ crosses $(0, 1-\epsilon)$ and $f + \underline{v}^2$ crosses $(0, \epsilon)$.

Notice that $t_1 \geq \underline{t}_1 \geq \underline{t}_2 \geq t_2$, implying that $t_1 - t_2 \geq \underline{t}_1 - \underline{t}_2$. Moreover, using similarity of triangles we obtain that $\underline{t}_1 - \underline{t}_2 = \frac{1-2\epsilon}{f_1}$. Therefore, $t_1 - t_2 \geq (1-2\epsilon)/f_1$.

Employing Corollary 1 over $\mathcal{E}'$ gives $\min\{c_1 s_1 + c_2 s_2 : (s_1, s_2) \in S(\mathcal{E}')\} \leq 2f_1/(1-2\epsilon)$. In contrast, $\min\{c_1 s_1 + c_2 s_2 : (s_1, s_2) \in T(\mathcal{E}')\} \geq 1$, because of the inequality $c_1 s_1 + c_2 s_2 \geq 1$ derived from the lattice-free triangle with vertices $f + v^1$, $f + v^2$ and $f - (\gamma, 0)$ for some $\gamma > 0$. Notice that such $\gamma$ exists because $f + v^1$ and $f + v^2$ do not cross the points $(0,1)$ and $(0,0)$ respectively. Using the cost vector $c = [c_1, c_2]$, we obtain the desired bound $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq (1-2\epsilon)/2f_1$. $\square$

### 3.2 Probability of bad ensembles

Using the ensemble constructed in the previous section and the monotonicity property from Lemma 1, we now analyze the probability that a random ensemble $\mathcal{E} \sim \mathcal{D}_n$ is bad for the split closure. Let $\Delta$ denote the triangle in $\mathbb{R}^2$ with vertices $(0,0)$, $(0,1)$, $(1/2, 1/2)$.

**Lemma 5.** *Let $\mathcal{E} = \langle f, r^1, \ldots, r^n \rangle$ be a random ensemble from $\mathcal{D}_n$, where $f = (f_1, f_2)$. Then for all $\bar{f} = (\bar{f}_1, \bar{f}_2) \in \Delta$ and all $\epsilon \in (0, 1/2)$, we have*

$$\Pr\left(\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \frac{1-2\epsilon}{\bar{f}_1} \middle| f = \bar{f}\right) \geq 1 - 2\left(1 - g(\bar{f}_1)\right)^n,$$

*where*

$$g(x) = \left(\frac{x}{1-\epsilon-(2-\sqrt{2})x} - \frac{x}{1-(2-\sqrt{2})x}\right).$$

*Proof.* Let us call *portals* the open segment connecting $(0, 1-\epsilon)$ and $(0,1)$ and the open segment connecting $(0,\epsilon)$ and $(0,0)$. Due to Lemmas 1 and 4 it suffices to lower bound the probability that a random ensemble has rays $r^i$ and $r^j$ such that $f + r^i$ crosses one portal and $f + r^j$ crosses the other portal.

Consider a ray $r^i$; the probability that $f + r^i$ crosses the open segment connecting $(0, 1-\epsilon)$ and $(0,1)$ equals to $\theta/2\pi$, where $\theta$ is the angle between the vectors $(0, 1-\epsilon) - \bar{f}$ and $(0,1) - \bar{f}$. So we have

$$\theta = \arctan\left(\frac{1-\bar{f}_2}{\bar{f}_1}\right) - \arctan\left(\frac{1-\epsilon-\bar{f}_2}{\bar{f}_1}\right). \tag{6}$$

Recall that $\arctan(.)$ is concave in $\mathbb{R}_+$. This implies that (6) is minimized when $\bar{f}_2$ is minimum. Since $\bar{f} \in \Delta$, $\bar{f}_2 \geq \bar{f}_1$ and hence we have

$$\theta \geq \arctan\left(\frac{1-\bar{f}_1}{\bar{f}_1}\right) - \arctan\left(\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}\right). \tag{7}$$

In order to simplify the previous bound we integrate arctan and notice that its derivative can be bounded as $1/(x^2+1) \geq 1/(x+\sqrt{2}-1)^2$ for all $x \in [1,\infty)$. Thus:

$$\theta \geq \arctan\left(\frac{1-\bar{f}_1}{\bar{f}_1}\right) - \arctan\left(\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}\right) = \int_{\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}}^{\frac{1-\bar{f}_1}{\bar{f}_1}} \frac{1}{x^2+1} \geq \int_{\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}}^{\frac{1-\bar{f}_1}{\bar{f}_1}} \frac{1}{(x+\sqrt{2}-1)^2}$$

$$= -\frac{1}{x+\sqrt{2}-1}\Bigg|_{\frac{1-\epsilon-\bar{f}_1}{\bar{f}_1}}^{\frac{1-\bar{f}_1}{\bar{f}_1}} = \left(\frac{\bar{f}_1}{1-\epsilon-(2-\sqrt{2})\bar{f}_1} - \frac{\bar{f}_1}{1-(2-\sqrt{2})\bar{f}_1}\right) = g(\bar{f}_1).$$

Therefore, the probability that $\bar{f}+r^i$ crosses the open segment connecting $(0,1-\epsilon)$ and $(0,1)$ is at least $g(\bar{f}_1)$. By symmetry, we can also prove that the probability that $\bar{f}+r^i$ crosses the open segment connecting $(0,\epsilon)$ and $(0,0)$ is also at least $g(\bar{f}_1)$; this bounds also holds for this case because it is independent of $\bar{f}_2$.

Let $B_1$ denote the event that no ray of $\mathcal{E}$ crosses the open segment connecting $(0,1-\epsilon)$ and $(0,1)$ and let $B_2$ denote the even that no ray of $\mathcal{E}$ crosses the open segment connecting $(0,\epsilon)$ and $(0,0)$. Using our previous bound we obtain that $\Pr(B_1) \leq (1-g(\bar{f}_1))^n$, and the same lower bound holds for $\Pr(B_2)$. Notice that the probability that $\mathcal{E}$ has rays $r^i$ and $r^j$ such that $f+r^i$ and $f+r^j$ cross distinct portals is $1-\Pr(B_1 \vee B_2)$; from union bound we get that this probability is at least $1-2(1-g(\bar{f}_1))^n$. This concludes the proof of the lemma. □

## 3.3 Proof of Theorem 1

In order to conclude the proof of Theorem 1 we need to remove the conditioning in the previous lemma. To make progress towards this goal, for $t \in [0,1/2]$ let $\Delta_t = \Delta \cap \{(x_1,x_2) : x_1 \leq t\}$. It is easy to see that the area of $\Delta_t$ equals $(1-t)t$. Now it is useful to focus on the set $\Delta_t \setminus \Delta_{\beta t}$, for some $\beta \in [0,1]$, since we can bound the probability that a uniform point lies in it and Lemma 5 is still meaningful. Using the independence properties of the distribution $\mathcal{D}_n$ we get that for every $\beta \in [0,1]$ and $\epsilon \in (0,1/2)$ a random ensemble $\mathcal{E} = \langle f, r^1, \ldots, r^n\rangle \sim \mathcal{D}_n$ satisfies:

$$\begin{aligned}
&\Pr\left(\mathrm{wc}(T(\mathcal{E}),S(\mathcal{E})) \geq \frac{1-2\epsilon}{2t}\Big| f \in \Delta\right)\\
&\geq \Pr\left(\mathrm{wc}(T(\mathcal{E}),S(\mathcal{E})) \geq \frac{1-2\epsilon}{2t}\Big| f \in \Delta_t \setminus \Delta_{\beta t}\right) \Pr\left(f \in \Delta_t \setminus \Delta_{\beta t}\Big| f \in \Delta\right)\\
&\geq \left[1-2\left(1-g\left(\beta t\right)\right)^n\right] \cdot 4 \cdot \left[(1-t)t - (1-\beta t)\beta t\right],
\end{aligned}$$

where the first inequality follows from the fact that $\Delta_t \setminus \Delta_{\beta t} \subseteq \Delta$ and the second inequality follows from the fact that $\beta t \leq f_1 \leq t$ and that the function $g(x)$ is increasing in $x$.

Finally, notice that this bound holds for all four 90-degree rotations of $\Delta$ around the point $(1/2,1/2)$; this is because of the symmetries of $\mathcal{D}_n$. Thus, by law of total probability we can remove the last conditioning. Using $\epsilon = 1/4$ and $\alpha = 1/4t$ we then obtain Theorem 1. We remark that we fixed the value of $\epsilon$ in order to simplify the expression in the theorem and that the value 1/4 was chosen experimentally in order to obtain good bounds specially for reasonably small values of $n$.

Since $T(\mathcal{E})$ is a relaxation of $P(\mathcal{E})$, as a corollary of the theorem we obtain a bound on the probability that the split closure is bad for random RCP's.

**Corollary 2.** *For any $\alpha \geq 1$ and $\beta \in [0,1]$, a random ensemble $\mathcal{E} \sim \mathcal{D}_n$ satisfies*

$$\Pr\left(\mathrm{wc}(P(\mathcal{E}),S(\mathcal{E})) \geq \alpha\right) \geq \left[1-2\left(1-g(\frac{\beta}{4\alpha})\right)^n\right]\left[\frac{1-\beta}{\alpha} - \frac{1-\beta^2}{4\alpha^2}\right],$$

*where*

$$g(x) = \left( \frac{x}{0.75 - (2-\sqrt{2})x} - \frac{x}{1-(2-\sqrt{2})x} \right).$$

# 4 Average-case measure

For $\epsilon > 0$ we define the product distribution $\mathcal{P}_\epsilon$ over $[\epsilon, 1]^n$ where a vector is obtained by taking each of its $n$ coefficients independently uniformly in $[\epsilon, 1]$. In this section we show that $avg(P(\mathcal{E}), G(\mathcal{E}), \mathcal{P}_\epsilon)$ is small for most ensembles $\mathcal{E}$ in $\mathcal{D}_n^m$.

**Theorem 2.** *Fix reals $\epsilon > 0$ and $\alpha > 1$ and an integer $m > 0$. Then for large enough $n$,*

$$\Pr_{\mathcal{E} \sim \mathcal{D}_n^m} \left( \text{avg}(P(\mathcal{E}), G(\mathcal{E}), \mathcal{P}_\epsilon) \le \alpha \right) \ge 1 - \frac{1}{n}.$$

We remark that the property that the cost vector is bounded away from zero in every coordinate is crucial in our analysis. This is needed because the ratio in (2) can become ill-defined in the presence of rays of zero cost.

The high level idea for proving the theorem is the following. Consider an ensemble $\mathcal{E} = \langle f, r^1, \ldots, r^n \rangle$. Define $\hat{f}$ as the integral point closest to $f$ in $l_2$ norm. It is not difficult to see that for every $c \in \mathcal{P}_\epsilon$, $\min\{cs : s \in P(\mathcal{E})\}$ is lower bounded by $\epsilon|\hat{f} - f|$, and this is achieved when the ensemble has the ray $(\hat{f} - f)/|\hat{f} - f|$ with cost $\epsilon$. We prove that this lower bound also holds for minimizing over $G(\mathcal{E})$ instead of $P(\mathcal{E})$. In addition, we show that for most ensembles $\mathcal{E}$, there are enough rays similar to $\hat{f} - f$ that have small cost. This allows us to upper bound $\min\{cs : s \in P(\mathcal{E})\}$ by roughly $\epsilon|\hat{f} - f|$ for most of the ensembles, which gives the desired result.

We start by proving the upper bound. For that, we need to study a specific subset of the ensembles in $\mathcal{D}_n^m$. We remark that the bounds presented are not optimized and were simplified in order to allow a clearer presentation.

## 4.1 $(\beta, k)$-good ensembles

Consider an ensemble $\mathcal{E} = \langle f, r^1, \ldots, r^n \rangle$. At a high level, we consider special regions in $\mathbb{R}^m$ 'around' $f - \hat{f}$, whose size depends on a parameter $\beta > 0$; then an ensemble is $(\beta, k)$-good if it has at least $k$ rays in each of these regions.

To make this precise, let $S^{m-1}$ denote the $(m-1)$-dimensional unit sphere in $\mathbb{R}^m$. Define $t \doteq \hat{f} - f$ and let $\rho$ be a rotation of $\mathbb{R}^m$ which maps $t/|t|$ into $e^m$. Let $\bar{C}(\beta)$ be the cap of the hypersphere $S^{m-1}$ consisting of all unit vectors with dot product at least $\beta$ with $e^m$. We also define $H_i^+$ as the halfspace given by $\{x \in \mathbb{R}^m : x_i \ge 0\}$ and $H_i^- = \{x \in \mathbb{R}^m : x_i \le 0\}$. We use the halfspaces $H_i^+$ and $H_i^-$ to partition $\bar{C}(\beta)$ into $2^{m-1}$ parts. That is, for $I \subseteq [m-1]$, let $\bar{C}_I(\beta) = \bar{C}(\beta) \cap (\bigcap_{i \in I} H_i^+) \cap (\bigcap_{i \in [m-1] \setminus I} H_i^-)$. Finally, let $C(\beta) = \rho^{-1}\bar{C}(\beta)$ and $C_I(\beta) = \rho^{-1}\bar{C}_I(\beta)$, that is, the sets obtained by applying the inverse rotation $\rho^{-1}$.

Using these structures, we say that $\mathcal{E}$ is $(\beta, k)$-good if for every $I \subseteq [m-1]$ there are at least $k$ rays $r^i$ in $C_I(\beta)$. The main property of such ensembles is that they allow us to use the following lemma.

**Lemma 6.** *Let $R$ be a subset of the rays of $\mathcal{E}$ such that $R \cap C_I(\beta) \neq \emptyset$ for all $I \subseteq [m-1]$. Then there is a solution $s \in P(\mathcal{E})$ supported in $R$ such that $\sum_{i=1}^n s_i \le \frac{|t|}{\beta}$.*

*Proof.* Without loss of generality assume that $R \cap C(\beta) = \{r^1, r^2, \ldots, r^{n'}\}$. First we show that $t \in cone(R \cap C(\beta))$. This follows from Farkas' Lemma and the hypothesis $R \cap C_I(\beta) \neq \emptyset$ for all $I \subseteq [m-1]$; the proof is deferred to the appendix.

**Claim 1.** *$t \in cone(R \cap C(\beta))$.*

So consider $s_1, s_2, \ldots, s_{n'} \geq 0$ with $\sum_{i=1}^{n'} s_i r^i = t$. We claim that $\sum_{i=1}^{n'} s_i \leq |t|/\beta$. To see this, first notice that by definition of $C(\beta)$ we have $r(t/|t|) \geq \beta$ for all $r \in C(\beta)$. Then multiplying the equation $\sum_{i=1}^{n'} s_i r^i = t$ by $t$ gives $\sum_{i=1}^{n'} s_i \beta |t| \leq \sum_{i=1}^{n'} s_i r^i t = tt = |t|^2$ and the claim follows.

Since $f + t = \hat{f}$ is integral we obtain that $s$ is a feasible solution for $P(\mathcal{E})$. This concludes the proof of the lemma. ☐

Using this lemma we can prove an upper bound on optimizing a cost vector in $\mathcal{P}_\epsilon$ over $P(\mathcal{E})$.

**Lemma 7.** *Fix $\beta, \epsilon > 0$ and an integer $k \geq 0$. Consider a $(\beta, k)$-good ensemble $\mathcal{E}$ and let $z(c) = \min\{cs : s \in P(\mathcal{E})\}$. Then*

$$\mathbb{E}_{c \sim \mathcal{P}_\epsilon}[z(c)] \leq |t| \left( p \frac{\epsilon}{\beta^2} + (1-p)\frac{1}{\beta} \right),$$

*where*

$$p = 1 - 2^{m-1} \left( \frac{1 - \epsilon/\beta}{1 - \epsilon} \right)^k.$$

*Proof.* Consider a vector $c$ which satisfies the following property: (*) for each $I \subseteq [m-1]$ there is a ray in $C_I(\beta)$ which has cost w.r.t $c$ at most $\epsilon/\beta$. Then employing Lemma 6 we obtain that $z(c) \leq |t|\epsilon/\beta^2$. Similarly, for a general vector $c \in [\epsilon, 1]^m$ we have the bound $z(c) \leq |t|/\beta$.

Now consider a vector $c \sim \mathcal{P}_\epsilon$. For a fixed $I$, the probability that every ray in $\mathcal{E} \cap C_I(\beta)$ has cost greater than $\epsilon/\beta$ is at most $((1-\epsilon/\beta)/(1-\epsilon))^k$. By union bound, $c$ satisfies property (*) with probability at least

$$1 - 2^{m-1} \left( \frac{1 - \epsilon/\beta}{1 - \epsilon} \right)^k.$$

The lemma then follows by employing the bounds on $z(c)$. ☐

### 4.2 Probability of obtaining a $(\beta, k)$-good ensemble

In this section we estimate the probability that a random ensemble in $\mathcal{D}_n^m$ is $(\beta, k)$-good. In order to simplify the presentation, we actually present a bound only for the specific value of $k$ chosen in hindsight to be equal to

$$\bar{k} \doteq n \frac{area(\bar{C}_\emptyset(\beta))}{area(S^{m-1})} - \sqrt{\frac{n(\ln n + m - 1)}{2}}. \tag{8}$$

Furthermore, we assume that $\beta, n$ and $m$ are such that $\bar{k} \geq 0$.

Consider a random ensemble $\mathcal{E} = \langle f, r^1, \ldots, r^n \rangle$ from $\mathcal{D}_n^m$ and let $R$ denote the set of rays of $\mathcal{E}$. We have that

$$\Pr\left(\mathcal{E} \text{ is } (\beta, \bar{k})\text{-good }\right) = \Pr\left( \bigwedge_{I \subseteq [m-1]} |R \cap C_I(\beta)| \geq \bar{k} \right) \geq 1 - 2^{m-1} \Pr\left(|R \cap \bar{C}_\emptyset(\beta)| < \bar{k}\right), \tag{9}$$

where the last inequality follows from the union bound and the fact that, by symmetry, $\Pr\left(|R \cap C_I(\beta)| < \bar{k}\right)$ is the same as $\Pr\left(|R \cap \bar{C}_\emptyset(\beta)| < \bar{k}\right)$ for every $I \subseteq [m-1]$.

Due to the independence of the rays, $|R \cap \bar{C}_\emptyset(\beta)|$ behaves as a sum of $n$ 0/1 random variables which take value 1 with probability $area(\bar{C}_\emptyset(\beta))/area(S^{m-1})$. At this point we recall the additive Chernoff bound on the tail of such distributions.

**Theorem 3** (Theorem 1.1 of [11]). *Let $X = \sum_{i=1}^n X_i$, where $X_i$ are random variables independently distributed in $[0,1]$. Then for all $t > 0$*

$$\Pr(X < \mathbb{E}[X] - t) \leq e^{-2t^2/n}.$$

By linearity of expectation we obtain that $\mathbb{E}[|R \cap \bar{C}_\emptyset(\beta)|] = n(area(\bar{C}_\emptyset(\beta))/area(S^{m-1}))$, hence employing the previous bound with $t = \sqrt{n(\ln n + m - 1)/2}$ we obtain that

$$\Pr\left(|R \cap \bar{C}_\emptyset(\beta)| < \bar{k}\right) \leq \frac{1}{ne^{m-1}}.$$

This upper bound together with inequality (9) gives that $\mathcal{E}$ is $(\beta, \bar{k})$-good with high probability.

**Lemma 8.** *Consider a random ensemble $\mathcal{E} \sim \mathcal{D}_n^m$ and let $\bar{k}$ be defined as in (8). If $\bar{k} \geq 0$, then*

$$\Pr\left(\mathcal{E} \text{ is } (\beta, \bar{k})\text{-good}\right) \geq 1 - \frac{1}{n}.$$

### 4.3 Lower bound for simple splits

In this section we show that $\epsilon|t|$ is also a lower bound for optimizing any vector in $[\epsilon, 1]^n$ over $G(\mathcal{E})$.

**Lemma 9.** *Fix $\epsilon > 0$ and consider an ensemble $\mathcal{E}$ in $\mathcal{D}_n^m$ and a vector $c \in [\epsilon, 1]^n$. For $t$ defined as before, we have*

$$\min\{cs : s \in G(\mathcal{E})\} \geq \epsilon|t|.$$

*Proof.* To prove this lemma, let $S_i \equiv \sum_{j=1}^n \psi^i(r^j)s^j \geq 1$ be the inequality for $P(\mathcal{E})$ obtained from the simple split $\{x : 0 \leq x_i \leq 1\}$. Clearly $S_i$ is valid for $G(\mathcal{E})$. Using the definition of Minkowski's functional, it is not difficult to see that

$$\psi^i(r^j) = \frac{r_i^j}{[r_i^j \geq 0] - f_i},$$

where $[r_i^j \geq 0]$ is the function that is equal to 1 if $r_i^j \geq 0$ and equal to 0 otherwise.

Now consider the inequality $\sum_{j=1}^n \psi(r^j)s_j \geq 1$ where

$$\psi(r^j) = \frac{\sum_{i=1}^m (\hat{f}_i - f_i)^2 \psi^i(r^j)}{\sum_{i=1}^m (\hat{f}_i - f_i)^2}.$$

This inequality is a non-negative combination of the inequalities $S_i$ and therefore is valid for $G(\mathcal{E})$. We claim that for any $c \in [\epsilon, 1]^m$, $\min\{cs : \sum_{j=1}^n \psi(r^j)s_j \geq 1\} \geq \epsilon|t|$, which will give the desired lower bound on optimizing $c$ over $G(\mathcal{E})$.

To prove the claim recall that $\sum_{i=1}^m (\hat{f}_i - f_i)^2 = |t|^2$ and notice that

$$\psi(r^j) = \frac{1}{|t|^2} \sum_{i=1}^m (\hat{f}_i - f_i)^2 \psi^i(r^j) = \frac{1}{|t|^2} \sum_{i=1}^m \frac{(\hat{f}_i - f_i)^2 r_i^j}{[r_i^j \geq 0] - f_i}.$$

Employing the Cauchy-Schwarz inequality and using the fact that $|r^j| = 1$, we get

$$\psi(r^j) \leq \frac{1}{|t|^2} |r^j| \sqrt{\sum_{i=1}^m \left(\frac{(\hat{f}_i - f_i)^2}{[r_i^j \geq 0] - f_i}\right)^2} \leq \frac{1}{|t|^2} \sqrt{\sum_{i=1}^m \frac{(\hat{f}_i - f_i)^4}{([r_i^j \geq 0] - f_i)^2}}.$$

However, since $\hat{f}$ is the integral point closest to $f$, for all $i$ it holds that $(\hat{f}_i - f_i)^2 \leq ([r_i^j \geq 0] - f_i)^2$. Employing this observation on the previous displayed inequality gives $\psi(r^j) \leq 1/|t|$. Therefore, any $s$ satisfying $\sum_{j=1}^n \psi(r^j)s_j \geq 1$ also satisfies $\sum_{j=1}^n s_j \geq |t|$. The claim then follows from the fact that every coordinate of $c$ is lower bounded by $\epsilon$. This concludes the proof of Lemma 9. □

### 4.4 Proof of Theorem 2

Recall that $\epsilon, \alpha$ and $m$ are fixed. Let $\beta$ be the minimum between $\sqrt{2/\alpha}$ and a positive constant strictly less than 1; this guarantees that $\bar{C}_\emptyset(\beta) > 0$. Consider a large enough positive integer $n$. Let $\mathcal{E}$ be a $(\beta, \bar{k})$-good ensemble in $\mathcal{D}_n^m$, where $\bar{k}$ is defined as in (8). Notice that $\bar{k}$, as a function of $n$, has asymptotic behavior $\Omega(n)$. We assume that $n$ is large enough so that $\bar{k} > 0$.

Now let us consider Lemma 7 with $k = \bar{k}$. The value $p$ defined in this lemma is also function of $n$, now with asymptotic behavior $1 - o(1)$. Thus, if $n$ is chosen sufficiently large we get $1 - p \leq \epsilon\beta\alpha/2$ and hence $\mathbb{E}_{c \sim \mathcal{P}_\epsilon}[z(c)] \leq |t|\epsilon\alpha$. If in addition we use the lower bound from Lemma 9, we obtain that $\text{avg}(P(\mathcal{E}), G(\mathcal{E}), \mathcal{P}_\epsilon) \leq \alpha$. The theorem then follows from the fact that an ensemble in $\mathcal{D}_n^m$ is $(\beta, \bar{k})$-good with probability at least $1 - 1/n$, according to Lemma 8.

## 5 Implications for the mixed integer case

In this section we consider the mixed integer model obtained from (RCP) by introducing integral 'non-basic' variables. That is we consider IP's of the form

$$\begin{aligned} &\min c^1 s + c^2 y \\ &x = f + \sum_{j=1}^{n} r^j s_j + \sum_{j=1}^{p} q^j y_j \\ &x \in \mathbb{Z}^m \\ &s \geq 0 \\ &y \in \mathbb{Z}_+^p \end{aligned} \tag{MG}$$

As RCP's arise as relaxations for IP's in tableaux form, the above IP also appears in the same context and offers a possibly much tighter relaxation since it only relaxes possible non-negativity constraints of basic variables. Our goal is to understand how the results from the previous section carry over to this model.

The IP (MG) is completely defined by a cost vector $c$ and a tuple $\mathcal{E} = \langle f, r^1, \ldots, r^n, y^1, \ldots, y^q \rangle$. With some overload in the notation we also call such tuple an ensemble. Given an ensemble $\mathcal{E}$ and a cost vector $c$, we use $MG(\mathcal{E}, c)$ to denote the associated mixed-integer program. As before, we work on the space of the 'non-basic' variables, hence we define $\tilde{P}(\mathcal{E})$ as the projection of the feasible region of $MG(\mathcal{E}, c)$ onto the $s, y$-space.

The random model for RCP's can be extended naturally for these mixed integer programs. That is, define the distribution $\tilde{\mathcal{D}}_{n,p}^m$ over ensembles where $f$ is picked uniformly from $[0,1]^m$ and each of the rays $r^1, \ldots, r^n, y^1, \ldots, y^q$ is picked independently and uniformly at random from the set of rational unit vectors in $\mathbb{R}^m$. Similarly, we define the cost distribution $\tilde{\mathcal{P}}_\epsilon$ where a vector in $[\epsilon, 1]^{n+p}$ is obtained by selecting each coefficient independently uniformly in $[\epsilon, 1]$.

Again our goal is to study the strength of simple split cuts for these random mixed integer programs. The (unstrengthened) extension of split cuts to MG's is also direct: given a split $X$ in $\mathbb{R}^m$, its associated split cuts is $\sum_{j=1}^n \psi_X(r^j)s_j + \sum_{j=1}^p \psi_X(q^j)y_j \geq 1$, where $\psi_X$ is defined in (1). It is easy to see that such inequality is valid for $\tilde{P}(\mathcal{E})$, since it is in fact valid for the set of solutions when the integrality constraints on the 'non-basic' variables are dropped. The simple split cut of $MG(\mathcal{E})$, denoted by $\tilde{G}(\mathcal{E})$, is defined as the intersection of all simple split cuts.

With definitions at hand, we extend Theorem 2 for MG's. The proof is a fairly direct modification of the proof of Theorem 2. We give it in the appendix.

**Theorem 4.** *Fix reals $\epsilon > 0$ and $\alpha > 1$ and positive integers $m$ and $p$. Then for $n$ sufficiently larger than $m$,*

$$\Pr_{\mathcal{E}\sim\tilde{\mathcal{D}}^m_{n,p}}\left(\text{avg}(\tilde{P}(\mathcal{E}), \tilde{G}(\mathcal{E}), \tilde{\mathcal{P}}_\epsilon) \leq \alpha\right) \geq 1 - \frac{1}{n}.$$

# References


[1] Kent Andersen, Quentin Louveaux, Robert Weismantel, and Laurence A. Wolsey. Inequalities from two rows of a simplex tableau. In *IPCO '07: Proceedings of the 12th international conference on Integer Programming and Combinatorial Optimization*, pages 1–15, Berlin, Heidelberg, 2007. Springer-Verlag.

[2] Egon Balas. Intersection cuts-a new type of cutting planes for integer programming. *Operations Research*, 19(1):19–39, 1971.

[3] Egon Balas and Andrea Qualizza. Personal communication.

[4] Egon Balas and Anureet Saxena. Optimizing over the split closure. *Math. Program.*, 113(2):219–240, 2008.

[5] Amitabh Basu. Corner polyhedra and maximal lattice-free convex sets : A geometric approach to cutting planes. PhD Thesis.

[6] Amitabh Basu, Pierre Bonami, Gérard Cornuéjols, and François Margot. Experiments with two-row cuts from degenerate tableaux. To appear in *INFORMS Journal on Computing*.

[7] Amitabh Basu, Pierre Bonami, Gérard Cornuéjols, and François Margot. On the relative strength of split, triangle and quadrilateral cuts. *Mathematical Programming*, 2009. DOI:s10107-009-0281-x.

[8] Valentin Borozan and Gérard Cornuéjols. Minimal valid inequalities for integer constraints. *Math. Oper. Res.*, 34(3):538–546, 2009.

[9] Gérard Cornuéjols and François Margot. On the facets of mixed integer programs with two integer variables and two constraints. *Mathematical Programming, Series A*, 120:429 – 456, 2009.

[10] Santanu S. Dey, Andrea Lodi, Andrea Tramontani, and Laurence A. Wolsey. Experiments with two row tableau cuts. In *IPCO*, pages 424–437, 2010.

[11] Devdatt Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms.* Cambridge University Press, New York, NY, USA, 2009.

[12] Daniel G. Espinoza. Computing with multi-row gomory cuts. *Operations Research Letters*, 38(2):115 – 120, 2010.

[13] Ralph E. Gomory. Some polyhedra related to combinatorial problems. *Linear Algebra and its Applications*, 2(4):451 – 558, 1969.

[14] Qie He, Shabbir Ahmed, and George Nemhauser. A probabilistic comparison of split, triangle and quadrilateral cuts for two row mixed-integer programs. Working paper, School of OR&IE, Goergia Tech, 2009.

[15] Alberto Del Pia, Christian Wagner, and Robert Weismantel. A probabilistic comparison of the strength of split, triangle, and quadrilateral cuts. Working paper, ETH Zürich, September 2010.

# Appendix

## A Some properties of the worst-cost

Consider a closed convex set $P = \{x \in \mathbb{R}^n_+ : a^i x \geq b_i \ \ \forall i \in I\}$ (we include sets where the index set $I$ may not be finite). The definition of $\alpha P$ implies that $\alpha P = \{x \in \mathbb{R}^n_+ : \alpha a^i x \geq b_i \ \ \forall i \in I\}$; we also define $\infty P = \mathcal{R}^n_+$. Moreover, $P$ is of *blocking type* if and only if $a^i \geq 0$ and $b_i > 0$ for all $i \in I$.

The next lemma establishes the connection between the above definitions and the worst-cost measure.

**Lemma 10.** *Consider two sets $P \subseteq Q$ of blocking type. Then*

$$\mathrm{wc}(P, Q) = \inf\{\alpha : Q \subseteq \alpha P\}.$$

*Proof.* Let $\alpha^* = \mathrm{wc}(P, Q)$.

($\geq$) If $\alpha^* = \infty$ then the inequality is obvious, so suppose $\alpha^* < \infty$. Let $ax \geq b$ be a valid inequality for $P$ with $a \geq 0$ and $b > 0$. Then

$$\frac{\inf\{ax : x \in P\}}{\inf\{ax : x \in Q\}} \leq \alpha^*.$$

Notice that both terms in the ratio of the LHS are nonzero, thus the last expression implies that for any $q \in Q$

$$aq \geq \inf\{ax : x \in Q\} \geq \frac{\inf\{ax : x \in P\}}{\alpha^*} \geq \frac{b}{\alpha^*},$$

where the last inequality follows from the fact $ax \geq b$ is valid for $P$. Consequently, $\alpha^* aq \geq b$ and since this holds for every supporting inequality of $P$ we get $Q \subseteq \alpha^* P$. This further implies that $\alpha^* = \mathrm{wc}(P, Q) \geq \inf\{\alpha : Q \subseteq \alpha P\}$.

($\leq$) Notice that $\mathrm{gap}(P, Q, c)$ as a function of $c$ from $[0,1]^n$ to $\mathbb{R} \cup \infty$ is upper semicontinuous, that is, for every $c \in [0,1]^n$ and every $\epsilon > 0$ there is a neighborhood $U$ of $c$ such that $\mathrm{gap}(P, Q, c') \leq \mathrm{gap}(P, Q, c) + \epsilon$ for every $c' \in U$. To see this, notice that the required condition holds trivially for points $c \in [0,1]^n$ where $\mathrm{gap}(P, Q, c) = \infty$ and it also holds for all $c \in [0,1]^n$ where $\mathrm{gap}(P, Q, c) < \infty$ due to continuity at those points. Therefore, since $[0,1]^n$ is a compact set, by Weierstrass' theorem there exists $a \in [0,1]^n$ such that $\alpha^* = \mathrm{gap}(P, Q, a)$.

If $\inf\{\alpha : Q \subseteq \alpha P\} = \infty$ then again the inequality holds trivially, so we assume that this quantity is finite. In particular, this implies that $P$ is non-empty. Using these two observations, let $\alpha' < \infty$ be such that $Q \subseteq \alpha' P$ and let $b = \inf\{ax : x \in P\}$. Since $ax \geq b$ is a valid inequality for $P$, $\alpha' ax \geq b$ is valid for $\alpha' P$ and by the definition of $\alpha'$ it is satisfied by every $x \in Q$. Consequently:

$$\big[\inf\{\alpha' ax : x \in Q\} \geq b\big] \equiv \big[\alpha'\{\inf ax : x \in Q\} \geq b\big] \Rightarrow \alpha' \geq \frac{b}{\inf\{ax : x \in Q\}} = \alpha^*,$$

where the last equation follows from $\alpha^* = \mathrm{gap}(P, Q, a)$. Taking the infimum over all $\alpha'$'s gives the desired result. □

Given a set $P = \{(x, y) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \bar{a}^i x + \hat{a}^i y \geq b_i \ \ \forall i \in I\}$, the *truncation* of $P$ to its first $k$ coordinates is defined as $P_k = \{x : \bar{a}^i x \geq b_i \ \ \forall i \in I\}$; alternatively, $P_k = \{x : (x, 0) \in P\}$. The next lemma shows that the worst-cost measure cannot increase when truncating.

**Lemma 11.** *Consider two sets $P \supseteq Q$ of blocking type. Then for any $k \leq n$*

$$\mathrm{wc}(P, Q) \geq \mathrm{wc}(P_k, Q_k).$$

*Proof.* To simplify the notation, let $\mathrm{wc}(P, Q) = \alpha$. Also let $P = \{(x, y) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \bar{a}^i x + \hat{a}^i y \geq b_i \quad \forall i \in I\}$. Notice that $P_k$ and $Q_k$ are of blocking type.

Lemma 10 gives that $Q \subseteq \alpha P$. We claim that this implies that $Q_k \subseteq \alpha P_k$. To see this, consider $q \in Q_k$. Since $(q, 0) \in Q \subseteq \alpha P$, for every $i \in I$ we have $\alpha(\bar{a}^i q + \hat{a}^i 0) \geq b_i$, or equivalently $\alpha \bar{a}^i q \geq b_i$; the claim then holds. Again employing Lemma 10, we have that $\mathrm{wc}(P_k, Q_k) \leq \alpha$ and the result follows. □

*Proof of Lemma 1.* In order to simplify the notation, we assume without loss of generality that $\mathcal{E}' = \langle f, r^1, r^2, \ldots, r^k \rangle$.

Consider $T(\mathcal{E})$ and recall that it is defined as $\{s : \sum_{i=1}^n \psi_T(r^i) \geq 1 \quad \forall T \in \mathcal{T}\}$, where $\mathcal{T}$ is the set of all triangles in $\mathbb{R}^2$ that contain $f$ but no integral point in their interior. Since $\psi_T \geq 0$, it follows that $T(\mathcal{E})$ is of blocking type. Using again its definition, we have that $T(\mathcal{E}') = \{s : \sum_{i=1}^k \psi_T(r^i) \geq 1 \quad \forall T \in \mathcal{T}\}$, and therefore $T(\mathcal{E}') = T(\mathcal{E})_k$. The same argument can be used to show that $S(\mathcal{E}') = S(\mathcal{E})_k$ as well. Then employing Lemma 11 we obtain the desired result $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \mathrm{wc}(T(\mathcal{E}'), S(\mathcal{E}'))$. □

# B Proof of Lemma 2

A key step in the proof of Lemma 2 is a method for constructing a polyhedron contained in the split closure. We minimize the function $c_1 s_1 + c_2 s_2$ over this strengthening of the split closure. The resulting LP implies an upper bound on the objective value when minimizing the function over the split closure.

To obtain this polyhderon, we define some inequalities which dominate the split closure $S(\mathcal{E})$. A *pseudo-split* is the convex set between two distinct parallel lines passing through $(0, 0)$ and $(0, 1)$ respectively. The direction of the lines, called *direction* of the pseudo-split, is a parameter.The *pseudo-split inequality* is derived from a pseudo-split exactly in the same way as from any maximal lattice-free convex set using formula (1). Note that pseudo-splits are in general not lattice-free and hence do not generate valid inequalities for $RCP(\mathcal{E}, c)$. However, we can dominate any split inequality cutting $f$ by an inequality derived from these convex sets. Indeed, consider any split $S$ containing the fractional point $f$ in its interior and passing through the segment joining $(0, 0)$ and $(0, 1)$. The pseudo-split with direction identical to the direction of $S$ generates an inequality that dominates the split inequality derived from $S$, as the coefficient for any ray is smaller in the pseudo-split inequality. The condition imposed on the rays to cross the left facet of the unit square implies the following. Any split which contains $f$, but does not pass through the segment $(0, 0), (0, 1)$, is dominated by any pseudo-split passing through the segment joining $(0, 0)$ and $(0, 1)$. So to dominate the split closure in this case, we only need to consider the inequalities derived from the pseudo-splits.

The next lemma states that we can dominate the split closure by using only the inequalities generated by the pseudo-splits with direction parallel to the rays $r^1, r^2$.

**Lemma 12.** *Consider an ensemble $\mathcal{E} = \langle f, r^1, r^2 \rangle$ where $f = (f_1, f_2) \in (0, 1)^2$, $r^1 = c_1(-1, t_1)$ and $r^2 = c_2(-1, t_2)$ with $c_1, c_2 \geq 0$ and $t_1 \geq t_2$, such that both $f + r^1$ and $f + r^2$ cross the segment joining $(0, 0)$ and $(0, 1)$. Then any pseudo-split inequality is dominated by the convex combination of the two pseudo-splits parallel to $r^1, r^2$.*

*Proof.* Let the pseudo-split parallel to $r^1$ be denoted by $S_1$ and similarly the pseudo-split parallel to $r^2$ be $S_2$. Consider any other pseudo split $S'$. Consider the point $\bar{f}$ on the segment joining

$(0,0)$ and $(0,1)$ be such that the segment joining $f$ and $\bar{f}$ is parallel to the direction of $S'$. Let $\bar{\mathcal{E}}$ be the ensemble $\langle \bar{f}, r^1, r^2 \rangle$. We compare the inequalities generated by the convex set $S'$ using the formula (1) for $P(\mathcal{E})$ and $P(\bar{\mathcal{E}})$. Let $\psi_X(r^i)$ be the coefficeint for $r^i$ in $P(\mathcal{E})$ and $\bar{\psi}_X(r^i)$ be the coefficient for $r^i$ in $P(\bar{\mathcal{E}})$ with respect to the convex set $X$.

**Observation 1.** *$\psi_{S'}(r^i) = \bar{\psi}_{S'}(r^i)$ for $i = 1, 2$ since the distance cut by $S'$ on the rays $r^1, r^2$ does not change in the two ensembles.*

**Observation 2.** *$\psi_{S_1}(r^i) \geq \bar{\psi}_{s_1}(r^i)$. This is because the coefficient for $r^1$ remains 0 and the distance cut by $S_1$ on $r^2$ is more in ensemble $\mathcal{E}$ as compared to in ensemble $\bar{\mathcal{E}}$. By a similar argument, $\psi_{S_2}(r^i) \geq \bar{\psi}_{s_2}(r^i)$ : the coefficient for $r^2$ remains 0 and the distance cut off on $r^1$ is more in $\mathcal{E}$ compared to $\bar{\mathcal{E}}$.*

We now make the following claim.

**Claim 2.** *There exists $0 \leq \lambda \leq 1$ such that $\bar{\psi}_{S'}(r^i) = \lambda \bar{\psi}_{S^1}(r^i) + (1-\lambda)\bar{\psi}_{S^2}(r^i)$ for $i = 1, 2$.*

*Proof.* We first note that $\bar{\psi}_{S_i}(r^i) = 0$ for $i = 1, 2$. This implies it suffices to show that $\frac{\bar{\psi}_{S'}(r^2)}{\bar{\psi}_{S^1}(r^2)} + \frac{\bar{\psi}_{S'}(r^1)}{\bar{\psi}_{S^2}(r^1)} = 1$. Indeed, we can then pick $\lambda = \frac{\bar{\psi}_{S'}(r^2)}{\bar{\psi}_{S^1}(r^2)}$. We use similarity of triangles to establish that $\frac{\bar{\psi}_{S'}(r^2)}{\bar{\psi}_{S^1}(r^2)} + \frac{\bar{\psi}_{S'}(r^1)}{\bar{\psi}_{S^2}(r^1)} = 1$. Refer to Figure 1 for the following notation. In the figure, ray $r^2$ is extended back to intersect $S_1$ at $D$ and $S'$ at $E$. Note that $\frac{\bar{\psi}_{S'}(r^2)}{\bar{\psi}_{S^1}(r^2)}$ is equal to $Ff/Gf$. By similarity of triangles, $Ff/Gf = Df/Ef = AC/AE$. Also, $\frac{\bar{\psi}_{S'}(r^1)}{\bar{\psi}_{S^2}(r^1)}$ is equal to $Bf/Af = CD/Af$ and by similarity of triangles, $CD/Af = CE/AE$. Since, $AC/AE + CE/AE = 1$, we have our identity. ∎

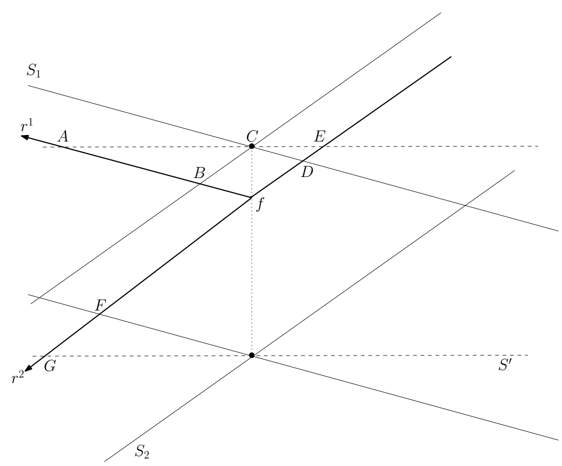


Figure 1: Figure for proof of Claim 2

Now combining Claim 2 and Observations 1 and 2, we obtain that the inequality $\psi_{S'}(r^1)s_1 + \psi_{S'}(r^2) \leq 1$ is dominated by the convex combination of the two inequalities $\psi_{S_1}(r^1)s_1 + \psi_{S_1}(r^2) \leq 1$ and $\psi_{S_2}(r^1)s_1 + \psi_{S_2}(r^2) \leq 1$ defined by $\lambda$ from Claim 2.

∎

We have thus shown that we need to consider the following LP to bound $\min\{c_1 s_1 + c_2 s_2 : (s_1, s_2) \in S(\mathcal{E})\}$ from above.

$$\begin{array}{lll} \min & c_1 s_1 + c_2 s_2 & \\ & \psi_{S_1}(r^1)s_1 + \psi_{S_1}(r^2)s_2 & \geq 1 \\ & \psi_{S_2}(r^1)s_1 + \psi_{S_2}(r^2)s_2 & \geq 1 \\ & s \in \mathbb{R}^2_+ . & \end{array} \tag{10}$$

We can derive the constraints corresponding to $S_1, S_2$. We have then to find an upper bound on the value of the following LP.

$$\begin{array}{lll} \min & s_1 + s_2 & \\ & 0 \cdot s_1 + \dfrac{c_2(t_1 - t_2)}{f_2 + f_1 t_1} s_2 & \geq 1 \\ & \dfrac{c_1(t_1 - t_2)}{1 - f_2 - f_1 t_2} s_1 + 0 \cdot s_2 & \geq 1 \\ & s \in \mathbb{R}^2_+ \ . & \end{array} \tag{11}$$

The upper bound can be obtained by exhibiting a feasible solution :

$$s_1 = \frac{1 - f_2 - f_1 t_2}{c_1(t_1 - t_2)} \quad \text{and} \quad s_2 = \frac{f_2 + f_1 t_1}{c_2(t_1 - t_2)} \ .$$

The value of this feasible solution is

$$c_1 s_1 + c_2 s_2 = \frac{1 + f_1(t_1 - t_2)}{t_1 - t_2}.$$

# C Proof of Claim 1 in Lemma 6

First we need a preliminary lemma.

**Lemma 13.** *Let $R' \subseteq \bar{C}(\beta)$ be such that $R' \cap \bar{C}_I(\beta) \neq \emptyset$ for all $I \subseteq [m-1]$. Then $e^m \in cone(R')$.*

*Proof.* Consider a vector $a \in \mathbb{R}^m$ such that $ar \geq 0$ for all $r \in R'$; we claim that $a_m \geq 0$. To see this, consider the set of indices $I = \{i \in [m-1] : a_i < 0\}$. Making use of our hypothesis, there is $r' \in R' \cap \bar{C}_I(\beta)$, which then satisfies $\sum_{i \in I} a_i r'_i + \sum_{i \in [m-1] \setminus I} a_i r'_i \leq 0$. Since $ar' \geq 0$, this implies that $a_m r'_m \geq 0$. Finally, since $r' e^m \geq \beta > 0$, we obtain that $r'_m > 0$ and hence $a_m \geq 0$.

From Farkas' Lemma $e^m \in cone(R')$ iff there is no vector with $a_m r \geq 0$ for all $r \in R'$ and $a_m < 0$, so the result follows from the previous claim. □

In order to prove Claim 1 we can proceed as follows. Letting $R' \doteq \rho R$, the definition of $R$ and the fact that $\bar{C}_I(\beta) = \rho C_I(\beta)$ implies that $R' \cap \bar{C}_I(\beta) \neq \emptyset$ for all $I \subseteq [m-1]$. Then Lemma 13 implies that $e^m \in cone(R')$. Since $\rho^{-1}$ is a linear transformation, we have $t = \rho^{-1} e^m \in \rho^{-1}(cone(R')) = cone(R)$.

# D Proof of Theorem 4

As in the proof of Theorem 2, we need an upper bound on $z(c) = \min\{cs : s \in \tilde{P}(\mathcal{E})\}$ and a lower bound on $\min\{cs : s \in \tilde{G}(\mathcal{E})\}$.

For the upper bound, we consider solutions of (MG) with $y = 0$. We are now back to a problem of the form (RCP). We say that a tuple $\mathcal{E} = \langle f, r^1, \ldots, r^n, y^1, \ldots, y^q \rangle$ is $(\beta, k)$-good if the ensemble $\langle f, r^1, \ldots, r^n \rangle$ is $(\beta, k)$-good as defined in Section 4.1. We can apply Lemmas 6-8 to $(\beta, k)$-good tuples $\mathcal{E}$.

For the lower bound, we relax the integrality constraint on the $y$ variables of (MG). We are now back to a problem of the form (RCP) with $n + p$ continuous variables. Applying Lemma 9, we get

**Lemma 14.** *Fix* $\epsilon > 0$ *and consider an ensemble* $\mathcal{E}$ *in* $\tilde{\mathcal{D}}^m_{n,p}$ *and a vector* $(c^1, c^2) \in [\epsilon, 1]^{n+p}$. *For* $t = \hat{f} - f$, *we have*

$$\min\{c^1 s + c^2 y : (s, y) \in \tilde{G}(\mathcal{E})\} \geq \epsilon |t|.$$

The proof of Theorem 4 now follows the proof of Theorem 2:

Let $\beta$ be the minimum between $\sqrt{2/\alpha}$ and a positive constant strictly less than 1; this guarantees that $\bar{C}_{\emptyset}(\beta) > 0$. Consider a large enough positive integer $n$. Let $\mathcal{E}$ be a $(\beta, \bar{k})$-good tuple in $\tilde{\mathcal{D}}^m_{n+p}$, where $\bar{k}$ is defined as in (8). Notice that $\bar{k}$, as a function of $n$, has asymptotic behavior $\Omega(n)$. We assume that $n$ is large enough so that $\bar{k} > 0$.

Now let us consider Lemma 7 with $k = \bar{k}$. The value $p$ defined in this lemma is also function of $n$, now with asymptotic behavior $1 - o(1)$. Thus, if $n$ is chosen sufficiently large we get $1 - p \leq \epsilon\beta\alpha/2$ and hence $\mathbb{E}_{c \sim \tilde{\mathcal{P}}_\epsilon}[z(c)] \leq |t|\epsilon\alpha$. If in addition we use the lower bound from Lemma 14, we obtain that $\text{avg}(\tilde{P}(\mathcal{E}), \tilde{G}(\mathcal{E}), \tilde{\mathcal{P}}_\epsilon) \leq \alpha$. The theorem then follows from the fact that a tuple in $\tilde{\mathcal{D}}^m_{n+p}$ is $(\beta, \bar{k})$-good with probability at least $1 - 1/n$, according to Lemma 8.