A Probabilistic Analysis of the Strength of the Split and Triangle Closures

Amitabh Basu[1], Gérard Cornuéjols[2⋆], and Marco Molinaro[2⋆⋆]

[1] University of California, Davis
[2] Carnegie Mellon University

Abstract. In this paper we consider the 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:

$$\min ay$$
$$Ay = b \qquad\qquad \text{(IP)}$$
$$y \geq 0, \quad y \in \mathbb{Z}^d.$$

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

$$\min \bar{a}_N y_N$$
$$y_B = \bar{b} - \bar{N} y_N \qquad \text{(IP')}$$
$$y \geq 0, \quad y \in \mathbb{Z}^d$$

where $\bar{a}_N \geq 0$ 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 [12] 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

$$\min cs$$
$$x = f + \sum_{j=1}^{n} r^j s_j \qquad \text{(RCP)}$$
$$x \in \mathbb{Z}^m, \quad s \geq 0$$

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 convex hull of feasible solutions of $RCP(\mathcal{E}, c)$ onto the s-space. In other words, $P(\mathcal{E}) = \text{conv}(\{s \geq 0 \,|\, f + \sum_{j=1}^{n} r^j s_j \in \mathbb{Z}^m\})$.

Intersection cuts. Let X be a closed, full-dimensional 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

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

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 geometric 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 ith 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 [6] 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 [11, 3, 5, 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 it 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 $P(\mathcal{E})$, $G(\mathcal{E})$, $S(\mathcal{E})$ and $T(\mathcal{E})$. Roughly, in the 'worst-cost' measure we look at the direction where the relaxation 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 [6] and shows that they are not as pathological as one could think. On the other hand, with respect to the average-cost measure we show that, with high probability, the simple split closure $G(\mathcal{E})$ is a very good approximation of $P(\mathcal{E})$. Since $P(\mathcal{E}) \subseteq T(\mathcal{E}) \subseteq S(\mathcal{E}) \subseteq G(\mathcal{E})$, this shows in particular that the split and the triangle closure are very similar under the average-cost measure and also approximate the integer hull very well. This gives a partial justification why triangle cuts seem to have a similar performance to split cuts in practice.

We remark that, due to space constraints, the proof of some claims are omitted; they can be found in [7].

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. [13] 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. [14] 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 [6]) 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:

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

Notice that this value is greater than 1 if A is stronger than B, i.e. $A \subseteq B$. 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:

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

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

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.

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 $\mathbb{R}^m_+$ and let $c \sim \mathcal{C}$ indicate that c is a random vector with distribution $\mathcal{C}$. Then, the *average-cost* measure between A and B is defined by

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

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 $\mathcal{E}'$ all of whose rays are unit vectors and a cost vector c', such that the optimal value of $\text{RCP}(\mathcal{E}, c)$ equals to the optimal value of $\text{RCP}(\mathcal{E}', c')$. Moreover, there exists an invertible affine transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ such that $P(\mathcal{E}) = A(P(\mathcal{E}'))$. Hence, in our model we assume that 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*

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

where

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

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 α.

Remark 1. Since $P(\mathcal{E}) \subseteq T(\mathcal{E})$, Theorem 1 implies that the same statement holds when $T(\mathcal{E})$ is replaced by $P(\mathcal{E})$, showing that the split closure is a bad approximation of the integer hull with the same probability.

Table 1. Values of the bound of Theorem 1 for different values of n and approximation factor α. The value of β in every entry was chosen empirically and attempts to optimize the bound.

n	α	β	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 %

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.

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 $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \mathrm{wc}(T(\mathcal{E}'), S(\mathcal{E}'))$.*

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 $(s_1, s_2) = (1/M, 1/M)$ 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 $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq M/2$. The following lemma formalizes the observation that if r^1 and r^2 are spread out then the split closure is weak.

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*

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

Now we are ready to establish the main lemma of this section, which exhibits bad ensembles for the split closure.

Lemma 3. *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 Lemma 2 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 γ 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 $\mathrm{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq (1-2\epsilon)/2f_1$.

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 Δ denote the triangle in $\mathbb{R}^2$ with vertices $(0,0)$, $(0,1)$, $(1/2, 1/2)$.

Lemma 4. *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*

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

where

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

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 3 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 θ is the angle between the vectors $(0, 1-\epsilon) - \bar{f}$ and $(0,1) - \bar{f}$. Expressing θ using the function $\arctan(.)$ and then using the concavity of the later, one can prove that $\theta \geq 2\pi g(\bar{f}_1)$ (we refer to the full version of the paper for a proof). 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 Δ_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 4 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:

$$\Pr\left(\text{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \frac{1 - 2\epsilon}{2t} \Big| f \in \Delta\right)$$

$$\geq \Pr\left(\text{wc}(T(\mathcal{E}), S(\mathcal{E})) \geq \frac{1 - 2\epsilon}{2t} \Big| f \in \Delta_t \setminus \Delta_{\beta t}\right) \Pr\left(f \in \Delta_t \setminus \Delta_{\beta t} \Big| f \in \Delta\right)$$

$$\geq [1 - 2(1 - g(\beta t))^n] \cdot 4 \cdot [(1 - t)t - (1 - \beta t)\beta t],$$

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

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

4   Average-Cost 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,*

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

Remark 2. Since $P(\mathcal{E}) \subseteq T(\mathcal{E}) \subseteq S(\mathcal{E}) \subseteq G(\mathcal{E})$, Theorem 2 implies that $S(\mathcal{E})$ is a good approximation of $T(\mathcal{E})$ and $P(\mathcal{E})$ under the average-cost measure. In fact, the same statement as Theorem 2 holds for any pair of bodies from $P(\mathcal{E}), T(\mathcal{E}), S(\mathcal{E}), G(\mathcal{E})$ with the appropriate containments. 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 ϵ. 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 (β, 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 (β, 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 $t = \hat{f} - f$ and let ρ 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 β with e^m. We also define H_i^+ as the halfspace given by $\{x \in \mathbb{R}^m : x_i \geq 0\}$ and $H_i^- = \{x \in \mathbb{R}^m : x_i \leq 0\}$. 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 ρ^{-1}.

Using these structures, we say that $\mathcal{E}$ is (β, 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 5. *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 $\sum_{i=1}^n s_i \leq \frac{|t|}{\beta}$.*

Proof. Without loss of generality assume that $R \cap C(\beta) = \{r^1, r^2, \ldots, r^{n'}\}$. First, one can use Farkas' Lemma and the hypothesis $R \cap C_I(\beta) \neq \emptyset$ for all $I \subseteq [m-1]$ to show that $t \in cone(R \cap C(\beta))$ (see full version of the paper for more details).

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 6. *Fix $\beta, \epsilon > 0$ and an integer $k \geq 0$. Consider a (β, 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 ϵ/β. Then employing Lemma 5 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 ϵ/β 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 (β, k)-Good Ensemble

In this section we estimate the probability that a random ensemble in $\mathcal{D}_n^m$ is (β, k)-good. Let

$$\bar{k} \doteq n \frac{area(\bar{C}_\emptyset(\beta))}{area(S^{m-1})} - \sqrt{\frac{n(\ln n + m - 1)}{2}}. \tag{6}$$

Using some Chernoff bound arguments, we can show the following lemma.

Lemma 7. *Consider a random ensemble $\mathcal{E} \sim \mathcal{D}_n^m$ and let $\bar{k}$ be defined as in (6). 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 8. *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 ϵ. This concludes the proof of Lemma 8.

4.4   Proof of Theorem 2

Recall that ϵ, α and m are fixed. Let β 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 (6). 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 6 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 $1 - p \le \epsilon\beta\alpha/2$ and hence $\mathbb{E}_{c \sim \mathcal{P}_\epsilon}[z(c)] \le |t|\epsilon\alpha$. If in addition we use the lower bound from Lemma 8, we obtain that $\mathrm{avg}(P(\mathcal{E}), G(\mathcal{E}), \mathcal{P}_\epsilon) \le \alpha$. 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 7.

References

1. Kent Andersen, Quentin Louveaux, Robert Weismantel, and Laurence A. Wolsey. Inequalities from two rows of a simplex tableau. In *IPCO '07, Lecture Notes in Computer Science*, volume 4513, pages 1–15. Springer-Verlag, 2007.
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. *Mathematical Programming, Series A*, 113(2):219–240, 2008.
5. Amitabh Basu, Pierre Bonami, Gérard Cornuéjols, and François Margot. Experiments with two-row cuts from degenerate tableaux. To appear in *INFORMS Journal on Computing, DOI: 10.1287/ijoc.1100.0437* .
6. Amitabh Basu, Pierre Bonami, Gérard Cornuéjols, and François Margot. On the relative strength of split, triangle and quadrilateral cuts. *Mathematical Programming, Series A*, 126:281–314, 2011.
7. Amitabh Basu, Gérard Cornuéjols, and Marco Molinaro. A probabilistic comparison of split, triangle and quadrilateral cuts for two row mixed-integer programs (extended version). http://www.optimization-online.org/DB_HTML/2010/10/2770.html.
8. Valentin Borozan and Gérard Cornuéjols. Minimal valid inequalities for integer constraints. *Mathematics of Operations Research*, 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 '10, Lecture Notes in Computer Science*, volume 6080, pages 424–437. Springer-Verlag, 2010.
11. Daniel G. Espinoza. Computing with multi-row gomory cuts. *Operations Research Letters*, 38(2):115–120, 2010.
12. Ralph E. Gomory. Some polyhedra related to combinatorial problems. *Linear Algebra and its Applications*, 2(4):451–558, 1969.
13. Qie He, Shabbir Ahmed, and George Nemhauser. A probabilistic comparison of split and type 1 triangle cuts for two row mixed-integer programs. Working paper, School of OR&IE, Goergia Tech, 2009, http://www.optimization-online.org/DB_HTML/2010/06/2636.html.

14. 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, http://arxiv.org/abs/1009.5253.