Valid inequalities based on the interpolation procedure

Sanjeeb Dash and Oktay Günlük

Mathematical Sciences Department,
IBM T. J. Watson Research Center,Yorktown Heights, NY 10598
(sanjeebd@us.ibm.com, oktay@watson.ibm.com)

April 7, 2004

Abstract

We study the interpolation procedure of Gomory and Johnson (1972), which generates cutting planes for general integer programs from facets of cyclic group polyhedra. This idea has recently been re-considered by Evans (2002) and Gomory, Johnson and Evans (2003). We compare inequalities generated by this procedure with mixed-integer rounding (MIR) based inequalities discussed in Dash and Gunluk (2003). We first analyze and extend the shooting experiment described in Gomory, Johnson and Evans. We show that MIR based inequalities dominate inequalities generated by the interpolation procedure in some important cases. We also show that the Gomory mixed-integer cut is likely to dominate any inequality generated by the interpolation procedure in a certain probabilistic sense. We also generalize a result of Cornuéjols, Li and Vandenbussche (2003) on comparing the strength of the Gomory mixed-integer cut with related inequalities.

1 Introduction

In a sequence of papers, Gomory [9], and later Gomory and Johnson [10, 11], studied the polyhedral structure of the master cyclic group polyhedron

$$P(n,r) = conv\Big\{w \in Z^{n-1} \ : \ \sum_{i=1}^{n-1}(i/n)\, w_i \equiv r/n \,(\text{mod } 1), \quad w \geq 0\Big\} \tag{1}$$

where $n, r \in Z$ and $n > r > 0$, and $a \equiv b\,(\text{mod } 1)$ means that $a - b$ is an integer. For a set $S \subseteq R^n$, we use $conv(S)$ to denote the convex hull of vectors in S. As discussed in [9, 10], facets of $P(n,r)$ can easily be translated into sub-additive functions using an *interpolation procedure*. These functions yield cutting planes for general integer programs. In particular, let

$$Y \ = \ \Big\{x \in Z^{|J|}, y \in Z \ : \ \sum_{j \in J} a_j\, x_j + y = b, \quad x \geq 0\Big\} \tag{2}$$

($\sum_{j\in J} a_j x_j + y = b$ could be derived from a row of the simplex tableau) and let $h(v)$ be a sub-additive function derived via interpolation from a facet of some $P(n,r)$. Then it is possible to show that

$$\sum_{i=1}^{|J|} h(a_j)x_j \geq h(b) \tag{3}$$

is a valid inequality for Y. Recently, there has been renewed interest in $P(n,r)$, see [13], [3], [12] and [5]. In particular, Evans [6] and Gomory, Johnson and Evans [13] present an empirical approach to identify "important" facets of $P(n,r)$ for small n and propose using sub-additive functions based on these important facets as cutting planes for general integer programs. The most important classes of facets identified in these empirical studies are the facets based on the Gomory mixed-integer cut (GMIC), and the so-called "2slope facets".

In a recent paper [5], we discussed how valid inequalities for Y can be derived from the mixed-integer rounding (MIR) principle, and a related two-step MIR principle. The GMIC inequalities can also be derived as MIR inequalities [17] and the 2slope inequalities form a subclass of two-step MIR inequalities [5].

In this paper, we are primarily interested in comparing these MIR based inequalities with cuts derived via the interpolation procedure. Another motivation of this study is to increase our understanding of the interpolation procedure. Our main result is that in a number of important cases, given an inequality based on the interpolation procedure, one can generate a stronger MIR based inequality. Combined with our computational results, this suggests that interpolation procedure described in [13] is not likely to be an effective way to generate cutting planes. Our analysis also yields a generalization of a result of Cornuéjols, Li and Vandenbussche [4] on comparing the strength of the GMIC with related inequalities.

The structure of the paper is as follows: in the rest of this section, we review earlier results and motivate our study. We also describe the interpolation procedure of Gomory and Johnson [10]. In Section 2, we discuss the *shooting experiment* of Gomory (see [13]) and present some computational results that identify the important facets of $P(n,r)$. In Section 3, we study properties of inequalities generated by the interpolation procedure and compare their strength with MIR based inequalities.

1.1 Valid inequalities based on simple sets

It is well-known that the basic MIR principle, based on a simple mixed-integer set with two variables, can be used to derive the GMIC. In particular, if

$$Q^1 = \Big\{ x \in R,\ z \in Z \ :\ x + z \ \geq\ \beta,\ \ x \geq 0 \Big\}$$

where $\beta \notin Z$, then the MIR inequality $x \ \geq\ (\beta - \lfloor \beta \rfloor)(\lceil \beta \rceil - z)$ is valid and facet defining for Q^1 (see [18]). The above inequality can be combined with an appropriate relaxation of Y to obtain the GMIC for Y [17]. Based on similar relaxations of Y, one can define t-scaled MIR inequalities, see [5]. Let $\hat{c}$ denote $c - \lfloor c \rfloor$ for $c \in R$.

Definition 1 *For $c \in R$, the MIR function with parameter c is defined as:*

$$f^c(v) = \begin{cases} \hat{v}/\hat{c} & \text{if } \hat{v} < \hat{c}, \\ (1-\hat{v})/(1-\hat{c}) & \text{if } \hat{v} \geq \hat{c}, \end{cases}$$

For any integer t such that $tb \notin Z$, the t-scaled MIR inequality

$$\sum_{j \in J} f^{tb}(ta_j)x_j \geq 1 \tag{4}$$

is valid for Y. These inequalities are called k-cuts in [4], and can be viewed as the GMIC (or, MIR) applied to $\sum_{j \in J} a_j x_j + y = b$ after scaling the equation by an integer t.

Let

$$Q^2 = \Big\{ x \in R,\ y, z \in Z\ :\ x + \alpha y + z\ \geq\ \beta,\ \ x, y \geq 0 \Big\},$$

where $1 > \beta > \alpha > 0$. In [5], we show that the *two-step MIR inequality*, defined as $x \geq (\beta - \alpha \lfloor \beta/\alpha \rfloor)(\lceil \beta/\alpha \rceil - \lceil \beta/\alpha \rceil z - y)$, is valid and facet defining for Q^2 provided that $1/\alpha \geq \lceil \beta/\alpha \rceil$. We then use the two-step MIR inequalities to derive valid inequalities for Y.

Definition 2 *For $c, \alpha \in R$ satisfying $\hat{c} > \alpha > 0$, and $1/\alpha \geq \lceil \hat{c}/\alpha \rceil > \hat{c}/\alpha$, the* two-step MIR function *with parameters c and α is defined as:*

$$g^{c,\alpha}(v) = \begin{cases} \dfrac{\hat{v}(1-\rho\tau) - k(v)(\alpha - \rho)}{\rho\tau(1-\hat{c})} & \text{if } \hat{v} - k(v)\alpha < \rho \\ \\ \dfrac{k(v) + 1 - \tau\hat{v}}{\tau(1-\hat{c})} & \text{if } \hat{v} - k(v)\alpha \geq \rho, \end{cases}$$

where $\rho = \hat{c} - \alpha\lfloor \hat{c}/\alpha \rfloor$, $\tau = \lceil \hat{c}/\alpha \rceil$ and $k(v) = \min\{\lceil \hat{v}/\alpha \rceil, \tau\} - 1$.

For any $t \in Z$ and $\alpha \in R$ such that tb and α are valid parameters for the two-step MIR function, the t-scaled two-step MIR inequality

$$\sum_{j \in J} g^{tb,\alpha}(ta_j)x_j \geq 1 \tag{5}$$

is valid for Y.

When applied to $P(n,r)$ (which has the same form as Y) these inequalities take the form

$$\sum_{i=1}^{n-1} f^{tr/n}(ti/n)\, w_i \geq 1 \tag{6}$$

and

$$\sum_{i=1}^{n-1} g^{tr/n, \Delta/n}(ti/n)\, w_i \geq 1 \tag{7}$$

for integers t and Δ. They are valid and facet defining under mild conditions (see [5]). We call facets derived from inequality (6) *t-scaled MIR* facets, and facets from inequality (7) *t-scaled two-step MIR* facets. In particular, the inequality in (6) defines a facet of $P(n,r)$ if tr is not a multiple of n. For t equal to 1, inequality (7) defines a facet of $P(n,r)$ if $n > \Delta \lceil r/\Delta \rceil > r$ and $r > \Delta$.

1.2 Valid inequalities based on interpolation

We first present a result of Gomory [9] that gives a complete characterization of the nontrivial facets (i.e., excluding the non-negativity inequalities) of $P(n, r)$.

Theorem 3 (Gomory [9]) *If $r \neq 0$, then $\sum_{i=1}^{n-1} \eta_i w_i \geq 1$ is a non-trivial facet of $P(n, r)$ if and only if $\eta = (\eta_j)$ is an extreme point of the inequality system*

$$\eta_i + \eta_j \geq \eta_{(i+j)\, mod\, n} \quad \forall i, j \in \{1, \ldots, n-1\}, \tag{8}$$

$$\eta_i + \eta_j = \eta_r \quad \forall i, j \text{ such that } r = (i+j)\, mod\, n, \tag{9}$$

$$\eta_j \geq 0 \quad \forall j \in \{1, \ldots, n-1\}, \tag{10}$$

$$\eta_r = 1. \tag{11}$$

The property (8) is called *sub-additivity*; we call the property (9) *r-additivity*. We will only be interested in non-trivial facets of $P(n, r)$.

Let $\sum_{i=1}^{n-1} \eta_i w_i \geq 1$ be a facet of $P(n, r)$. Let $h : R \to [0, 1]$ be an associated piecewise-linear function defined by:

$$h(v) = \begin{cases} 0, & \text{if } v = \lfloor v \rfloor, \\ \eta_i & \text{if } v - \lfloor v \rfloor = \frac{i}{n} \text{ for } i \in \{1, \ldots, n-1\}, \\ \delta h(\frac{i}{n}) + (1-\delta) h(\frac{i+1}{n}) & \text{if } v - \lfloor v \rfloor = \frac{i+\delta}{n}, \text{ for } i \in \{1, \ldots, n-1\},\ 0 < \delta < 1. \end{cases}$$

Note that $0 \leq h(v) \leq 1$ as $0 \leq \eta_i \leq \eta_r = 1$, and $h(v) = h(\hat{v})$ for all $v \in R$. We will call $h(v)$ a *facet interpolated template function*, abbreviated as a *template function*. In Figure 1 we present a template function based on the 6-scaled MIR facet for $P(13, 12)$. Observe that it is enough to describe $h(v)$ over the interval $[0, 1]$, as $h(v) = h(\hat{v})$. The functions f^c and $g^{c,\alpha}$ defined earlier also have this property; they only depend on the fractional part of their argument.

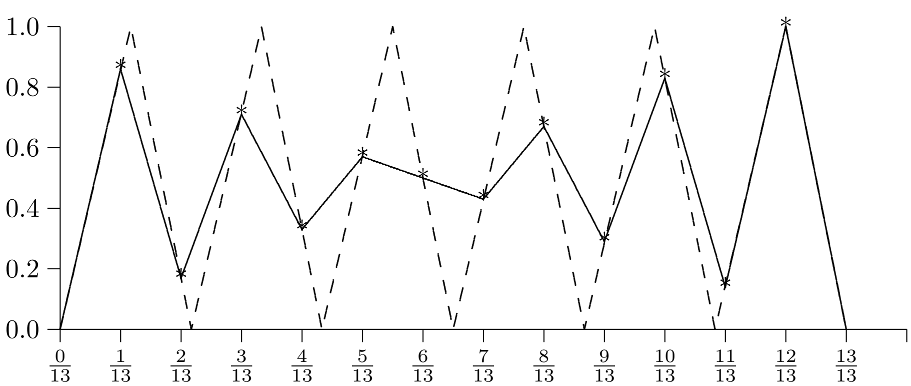

Figure 1: 6-scaled MIR for $P(13, 12)$ and template function

Gomory and Johnson [10] derived the following *sub-additivity* property of template functions from the sub-additivity property of facet coefficients η_i:

$$h \text{ is a template function of } P(n, r) \Rightarrow h(x) + h(y) \geq h(x + y),\ \forall x, y \in R. \qquad (12)$$

Note that the functions f^c and $g^{c,\alpha}$ are both sub-additive functions. From (12) they show that template functions can be used to derive cutting planes for Y.

Proposition 4 (Gomory and Johnson [10]) *If h is a template function, then*

$$\sum_{i \in J} h(a_j)x_j \geq h(b)$$

is a valid inequality for Y.

An important question is: which template function based cutting planes are useful for Y ?

Some template functions for Y can be viewed as being special in the following sense. Let $\bar{n}$ be the smallest positive integer such that all coefficients of $\sum_{j \in J} a_j x_j + y = b$ become integral when multiplied by $\bar{n}$. Define $\bar{r} = \bar{n}\hat{b}$, and consider $P(\bar{n}, \bar{r})$, which we call the *canonical master polyhedron* for Y. Notice that Y can be viewed as a lower-dimensional face of $P(\bar{n}, \bar{r})$, see Gomory [9]. Therefore, every facet of $conv(Y)$ is of the form $\sum_{j \in J} h(a_j)x_j \geq h(b)$ where h is some template function derived from $P(\bar{n}, \bar{r})$. In general, it is impractical to work directly with the canonical $P(\bar{n}, \bar{r})$ and the associated inequality system (8)-(11), as $\bar{n}$ is too large. In a recent paper, Gomory, Johnson and Evans[13] instead propose using template functions of important facets of $P(n, r)$ with small values of n, to get cutting planes for Y. They show that small master polyhedra (with $n \leq 30$) are amenable to computational analysis; we examine their approach in detail in the next section.

2 Important facets of master polyhedra

Gomory, Johnson and Evans [13] propose a computational approach to identify the "important" facets of small master cyclic group polyhedra $P(n, r)$. More precisely, they suggest a randomized procedure, which they call the *shooting experiment*, where they repeatedly choose a random direction $d \geq 0$ and identify the facet f first encountered along the ray $\{\lambda d \,|\, \lambda \geq 0\}$ as λ increases from 0. One can view the facet f as being *hit* by the "shot" d. After a number of shots, they determine the frequency with which different facets are hit. They propose that a facet should be considered important if it is hit frequently. Similar shooting experiments were first performed by Kuhn in the 1950s in the context of the TSP, see [15, 16].

In this experiment a "random direction" means a random vector v such that $v/||v||_2$ is uniformly distributed over the surface of the unit-sphere. Therefore, the probability of a given facet being hit is proportional to the solid angle subtended by the facet at the origin, or equivalently to the area of the "projection" of the facet on the unit-sphere. The definition of "importance" for a facet has two inherent assumptions: (i) a facet is important if it occupies a large area and (*ii*) large facets will have a large projection on the unit-sphere. One justification for (*i*) is that if a facet with small area is removed, $P(n,r)$ does not change much. A partial justification for (*ii*) is given in [13]. Clearly, there are some drawbacks to this measure of importance. For example, the projection of a non-negativity constraint, which induces an unbounded facet of $P(n,r)$, is negligible on the surface of the unit-sphere and therefore it is considered unimportant. Nonetheless, we still agree that this is a reasonable, and more importantly, measurable definition of importance of the *non-trivial* facets of $P(n,r)$.

The main observations in [13] are: (a) a relatively small number of facets of $P(n,r)$ absorb most of the hits; (b) the most important facets of $P(n,r)$ are what we call t-scaled MIR facets (6). A similar study by Evans [6] reports that the so-called "2slope facets" [3] constitute the second most important class of facets . The 2slope facets form a sub-class of the two-step MIR facets (7). The experiments reported in [13] and [6] are performed on master polyhedra with $n \leq 30$. Related work on the shooting experiment of Gomory can be found in Hunsaker's thesis [14], where he examines different measures of quality of facets, including the notion of importance defined above.

In this section we analyze the shooting experiment more formally and extend the experimental results in [13] and [6] in two ways. First, we investigate if observations (a) and (b) extend to $P(n,r)$ for larger values of n, up to 200. Secondly, we measure the importance of additional facet classes.

2.1 Experimental framework

Given a direction $d \geq 0$, the non-trivial facet $\eta^T x \geq 1$ of $P(n,r)$ hit by d is the facet which minimizes $d^T\eta$ over all facets [13]. The non-trivial facets are extreme points of the system of inequalities (8)-(11). Therefore, for any $d \in R_+^{n-1}$, a basic optimal solution of the linear program

$$
\begin{array}{lrcl}
\mathit{min} & d^T\eta & & \\
\mathit{subject\ to} & & & \\
 & \eta_i + \eta_j & \geq & \eta_{(i+j) \bmod n} \quad \forall i,j \in \{1,\ldots,n-1\}, \\
 & \eta_i + \eta_j & = & 1 \quad \forall i,j \text{ such that } r = (i+j) \bmod n, \\
 & \eta_j & \geq & 0 \quad \forall j \in \{1,\ldots,n-1\}, \\
 & \eta_r & = & 1.
\end{array}
$$

gives the facet hit by d. For a random direction d, the probability that the LP above has multiple optimal solutions is negligible. This linear program has $n-1$ variables and less than $n^2/2$ constraints. About half of the variables can be substituted using the equality constraints.

To generate the random vector d, we first generate n Gaussian random variables X_i. The distribution of the vector $d/||d||_2$ where $d = [X_1, X_2, \ldots, X_n]$ is known to be uniform over the surface of the unit-sphere (see [7], pp. 234). To restrict the vector d to the non-negative quadrant, we first generate the vector without this restriction, and then simply take the absolute value of its components. Finally, we use the code of Acklam [1] to generate the Gaussian distribution (also called the standard normal distribution). In our computations we use d instead of $d/||d||_2$.

Gomory, Johnson and Evans perform experiments on $P(n, r)$ for all $n \leq 30$ and for selected values of r for each n. Though not explicitly stated, the values of r they select actually cover the $n-1$ possible choices of r. Many of these $n-1$ choices yield essentially identical master polyhedra. More precisely, let k be an integer co-prime with n, i.e., k and n have no common divisors. Let $\phi(i) = ki \bmod n$; here ϕ defines a permutation of $\{1, 2, \ldots, n-1\}$. Then $P(n, r)$ and $P(n, \phi(r))$ are *isomorphic* in the following sense. If $\sum_i \alpha_i w_i \geq 1$ is a non-trivial facet of $P(n, r)$, then $\sum_i \alpha_i w_{\phi(i)} \geq 1$ is a facet of $P(n, \phi(r))$, and vice-versa[9]. We say that the first facet above is isomorphic to the second. Thus, the non-trivial facets of $P(n, r)$ and $P(n, \phi(r))$ are identical after permuting variables via ϕ. One can show that non-isomorphic $P(n, r)$ correspond to distinct divisors of n. For example, $P(10, 1), P(10, 3), P(10, 7)$ and $P(10, 9)$ are isomorphic to one another, whereas $P(10, 1), P(10, 2)$ and $P(10, 5)$ are non-isomorphic.

We can show that the MIR based facets of $P(n, r)$ are isomorphic to the MIR based facets of $P(n, \phi(r))$. Then two isomorphic master polyhedra will have their MIR based facets hit equally frequently in shooting experiments. We state our result on isomorphism between MIR based facets below.

Theorem 5 *Let $P(n, r)$ be some master cyclic polyhedron. Let k be an integer co-prime with n, and let $\phi(i) = ki \bmod n$. Then the scaled MIR facets of $P(n, r)$ and the scaled two-step MIR facets of $P(n, r)$ are isomorphic, respectively, to the scaled MIR facets and scaled two-step MIR facets of $P(n, \phi(r))$. Further, if t is the largest divisor of n and $tr \bmod n \neq 0$, then the t-mir facet of $P(n, r)$ is isomorphic to the t-mir facet of $P(n, \phi(r))$.*

Proof. The values of the functions $f^b(v)$ and $g^{b,\alpha}(v)$ depend only on $\hat{b}$ and $\hat{v}$, the fractional parts of b and v, respectively. Thus if b' and v' are numbers such that $b - b'$ and $v - v'$ are integral, then

$$f^b(v) = f^{b'}(v') \text{ and } g^{b,\alpha}(v) = g^{b',\alpha}(v').$$

Consider the t-scaled MIR of $P(n, r)$ for some integer t. Let k and ϕ be defined as in the theorem. From (6), the t-scaled of $P(n, r)$ is isomorphic to the following facet of $P(n, \phi(r))$:

$$\sum_{i=1}^{n-1} f^{tr/n}(ti/n)\, w_{\phi(i)} \geq 1. \tag{13}$$

Let s be the unique integer such that $sk \bmod n = t$. We will show that (13) is the s-scaled MIR for $P(n, \phi(r))$. If j is any integer between 1 and $n-1$, then $s\phi(j) \bmod n = skj \bmod n = tj \bmod n \Rightarrow s\phi(j)/n - tj/n$ is integral. Therefore (13) is the same as

$$\sum_{i=1}^{n-1} f^{s\phi(r)/n}(s\phi(i)/n) w_{\phi(i)} \geq 1,$$

which is the s-scaled MIR of $P(n, \phi(r))$. Similarly, the t-scaled two-step MIR facet of $P(n, r)$ with parameter Δ in (7) is isomorphic to the s-scaled two-step MIR facet of $P(n, \phi(r))$ with parameter Δ. Finally, if t is the largest divisor of n, and $sk \bmod n = t$, then $s = t$, and the theorem follows. ■

For example, the scaled MIR facets of $P(10, 1)$ are isomorphic to the scaled MIR facets of $P(10, 3)$.

2.2 Statistical relevance of experiments

For probabilistic experiments, an important issue is to decide on the number of trials that would give reliable estimates of true probability values. In their experiments, Gomory, Johnson and Evans select 10,000 random directions. (For his experiments in 1953, Kuhn [16] used ten shots, with random numbers "obtained by sticking a pin at random in the Los Angeles telephone book"; such experiments have certainly become easier to perform). Let p denote the probability that a random direction in the shooting experiment hits a member of the family of facets we are interested in. With each shot we can associate a Benoulli random variable X_i, $i \geq 1$, that takes value 1 if the shot hits an interesting facet, or takes value 0 otherwise. Using the Central Limit Theorem for sums of Bernoulli variables with success probability p, $Y_t = (1/t)\sum_1^t X_i$ is approximately normally distributed with mean p and variance $p(1-p)/t$. Let $\bar{p}_t$ be the success frequency observed after t shots. For any fixed $\alpha \in (0, 1)$, it is possible to derive a reliable lower bound on p based on the estimate $\bar{p}_t$ as follows (see [19]):

$$P(p > \bar{p}_t - \Phi^{-1}(\alpha)\sqrt{\bar{p}_t(1-\bar{p}_t)/t}) \geq \alpha \tag{14}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution $N(0, 1)$. Notice that this analysis does not depend on n in $P(n, r)$. In Table 1 we use (14) with $\alpha = 1 - 10^{-4}$ to derive the lower bounds on p based on $\bar{p}_t$ for $t = 10,000$, $25,000$, $100,000$ and $1,000,000$.

Observed value of $\bar{p}_t$		1%	3%	5%	10%	20%	30%	50%
Lower bound on p with t=	10,000	0.0%	1.0%	2.5%	6.6%	15.4%	24.8%	44.3%
Lower bound on p with t=	25,000	0.8%	2.6%	4.5%	9.3%	19.1%	29.0%	48.9%
Lower bound on p with t=	100,000	0.9%	2.8%	4.8%	9.7%	19.5%	29.5%	49.4%
Lower bound on p with t=	1,000,000	1.0%	2.9%	4.9%	9.9%	19.9%	29.8%	49.8%

Table 1: Lower bounds on p for $\alpha = .9999$ and $\Phi^{-1}(\alpha) = 3.62$

As seen in Table 1, only 10,000 shots may not lead to accurate conclusions, especially when $\bar{p}_t$ is small. Given the tradeoff between computation time and accuracy of the results, we decided to use 100,000 shots in our experiments.

In Figure 2 we show how $\bar{p}_t$ changes as t increases for three groups of facets for $P(50, 1)$. The first group consists of all scaled MIR and scaled two-step MIR facets, and the second group contains only the scaled MIR facets. The last group consists of only the 25-scaled MIR facet (we discuss this facet further in Section 2.4). As seen in the Figure 2, $\bar{p}_t$ becomes quite stable after 25,000 shots and almost does not change after 50,000 shots.

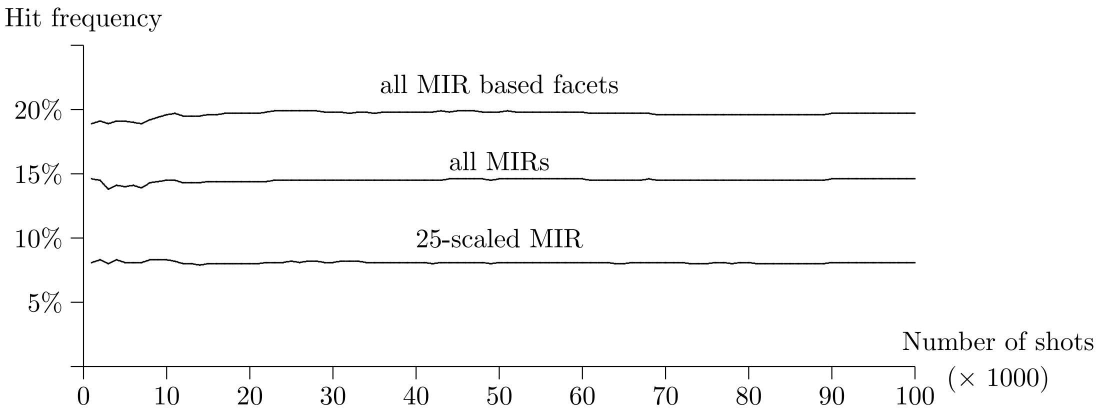

Figure 2: Shooting results for $P(50, 1)$

2.3 Computational results

We use two approaches to obtain our experimental results. For smaller n, we solve the linear program described in Section 2.1 to optimality for each direction d. Here we use the LP solver QSopt version 1.0 [2]; any other solver could be used for this purpose. We compute the frequency with which different facets are hit, including non-MIR based facets. We call this *complete shooting*, and use this approach for $n \leq 90$. For $P(90, 1)$ for example, 100,000 shots took about 96,700 seconds on a 375 MHz PowerPC running AIX Version 5, release 1.

For larger n, we adopt a different approach and only compute the frequency with which MIR based facets are hit. This approach yields less information but is faster. For a random direction $d \geq 0$, we check if d hits a MIR-based facet. To do this, we enumerate the $O(n^2)$ MIR based facets, and find the facet $\bar{\eta}$ from this class which minimizes $d^T\eta$. We then test if $\bar{\eta}$ is an optimal solution of the LP in Section 2.1, by giving $\bar{\eta}$ to a simplex based LP solver and terminating before optimality if there is a simplex step that improves the objective function. We call this *partial shooting*, and we go up to $n = 200$ with this approach. We use COIN-Clp version 0.93 [8] for partial shooting. Specific features of COIN-Clp make it easier to perform the optimality test mentioned above, and influenced our decision to use it in this context. On an 800 MHz Itanium2 processor running HP TrueUnix, 100,000 shots take about 18,000 seconds when $n = 100$, and about 120,000 seconds when $n = 200$.

In Figure 3, we consider $P(n, r)$ for fixed n – here n is 50 – and let r run from 1 to $n - 1$. We plot the frequency with which MIR based facets are hit versus the right-hand side r (on the x-axis). This figure illustrates Theorem 5 in the case $n = 50$. For example, Theorem 5 implies that the MIR based facets of $P(50, 10)$ will be hit with the same frequency as those of $P(50, 20), P(50, 30)$ and $P(50, 40)$. This frequency is 7.6% for each master polyhedron above.

In Tables 2 and 3, we present our results with selected master polyhedra. For each value of n, we choose up to three different factors of n; letting r stand for these different factors, we get non-isomorphic master polyhedra $P(n, r)$. We always choose 1, the largest factor of n other than n, and some other factor of n as possible choices for r. If n is prime, there is of

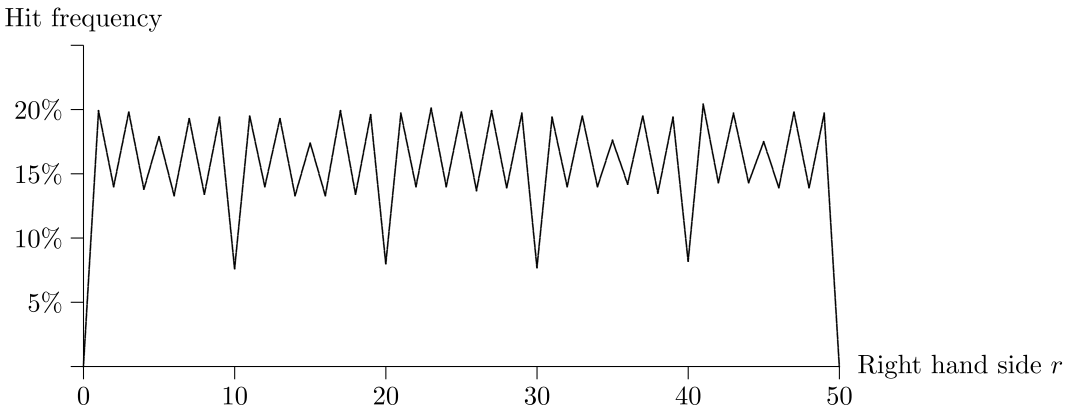

Figure 3: % of shots absorbed by all MIR based facets for $P(50, r)$

course a single choice for r.

Our observations are similar to those in [13], especially regarding the importance of scaled MIR facets. We also report on the importance of two-step MIR facets. In Table 2, we give our shooting results for three different values of n. The second column gives the number of all MIR based facets. The third column gives the number of *distinct* facets hit in 100,000 shots; this is a lower bound on the total number of facets. The fourth column gives the frequency (in percentage terms) with which the MIR based facets are hit. The fifth and sixth columns give, respectively, the hit frequencies for the ten most frequently hit facets, and for the ten most frequently hit MIR based facets. Our results clearly suggest that MIR based facets are very important for these polyhedra. Looking at the row for $P(50, 1)$ for example, we see that 378 MIR based facets (out of 65,346 facets) absorb almost 20% of all hits. The ten most frequently hit facets absorb 13.7% of all hits. Also, the ten most frequently hit MIR based facets absorb 13.7% of all hits. It turns out that for each master polyhedron in Table 2, at least nine out of the ten most frequently hit facets are MIR based facets. Thus, a very few facets absorb a large fraction of all hits, and are mostly MIR based facets.

the group	total MIRs	total facets hit	% all mirs	% top 10	% top 10 MIRs
P(31,1)	194	29,617	26.0	8.6	8.6
P(42,1)	207	46,407	30.4	21.9	21.8
P(42,7)	216	47,918	28.4	19.3	19.2
P(42,21)	247	53,754	27.2	15.5	15.5
P(50,1)	378	65,346	19.7	13.7	13.7
P(50,10)	370	74,736	7.6	1.4	1.3
P(50,25)	465	68,202	20.0	11.2	11.2

Table 2: Shots absorbed by MIR based facets for different $P(n, r)$

In Table 3, we look separately at the MIRs and the two-step MIRs. Columns 2 and 3 give the number of MIRs and two-step MIRs, respectively. Columns 4,5,6,7 and 8 give, respectively, the hit frequency for the most important MIR based facet, the scaled MIRs, 2slope facets, two–step MIRs and all MIR based facets. The most important MIR based facet is defined to be the one hit most frequently. Looking at columns 5,7 and 8, it is clear that the MIR facets and the two-step MIR facets are important facets of $P(n,r)$. Also, neither class is uniformly more important than the other. For example, for $P(42,1)$ the MIRs are more important, whereas the two–step MIRs are more important for $P(42,21)$. Comparing columns 6 and 7, we see that the two-step mir facets which are not 2slopes are also important facets of $P(n,r)$. Thus, we establish the importance of another class of facets of $P(n,r)$ in addition to the scaled MIRs and 2slopes already discussed in [3, 6].

In column 4, except for the numbers indicated by a '*', the most important MIR based facet is a t-scaled MIR facet with t a divisor of n. In particular, t is the largest divisor of n for which the t-scaled MIR inequality is valid and facet-defining for $P(n,r)$. For example, for $P(50,1)$ and $P(50,25)$, the largest MIR based facet is the 25-scaled MIR facet. For $P(50,1)$, the percentage of hits absorbed by the 25-scaled MIR facet is 8.1% whereas all remaining scaled MIR facets combined absorb only 6.5% of the shots. The total fraction of shots absorbed by all scaled two-step MIR facets is 5.1% . However, for $P(72,18)$, the 36-scaled MIR does not define a facet as 36×18 is a multiple of 72. Instead, the most important facet is the 18-scaled MIR.

We have verified for all cases in Table 2 other than $P(50,10)$, that the most important MIR facet is in fact the most important facet. We discuss this further in the next section.

	Number of facets		% of shots hit				
The group	MIRs	2-step MIRs	top MIR	MIRs	2-slopes	2-step MIRs	all MIRs
P(31,1)	15	179	1.5	8.9	11.6	17.1	26.0
P(42,1)	21	186	9.4	23.5	4.1	6.9	30.4
P(42,7)	18	198	9.3	18.4	6.8	10.0	28.4
P(42,21)	11	236	9.6	11.8	13.5	15.4	27.2
P(50,1)	25	353	8.1	14.6	2.9	5.1	19.7
P(50,10)	20	350	0.2*	1.9	3.6	5.8	7.6
P(50,25)	13	452	8.1	8.9	9.8	11.1	20.0
P(72,1)	36	673	5.0	11.6	1.3	2.2	13.8
P(72,18)	27	616	0.9	2.3	2.0	2.2	4.5
P(72,36)	18	682	0.2*	0.6	3.5	3.6	4.2
P(83,1)	41	1814	0.1	0.4	0.7	1.0	1.5
P(90,1)	45	995	4.0	7.7	0.6	0.9	8.7
P(90,6)	42	740	0.6	1.8	1.0	1.1	3.0
P(90,45)	23	1074	4.0	4.2	2.3	2.4	6.7
P(100,1)	50	1534	3.2	4.8	0.3	0.5	5.4
P(100,4)	48	1325	0.4	1.1	0.5	0.6	1.7
P(100,50)	25	1784	0.4	0.5	1.0	1.0	1.5
P(150,1)	75	3065	1.2	1.8	0.0	0.0	1.8
P(150,5)	73	2891	1.2	1.5	0.1	0.1	1.5
P(150,75)	38	3302	1.2	1.3	0.2	0.2	1.4
P(200,1)	100	6584	0.4	0.5	0.0	0.0	0.5

Table 3: Shots absorbed by MIR based facets for different $P(n,r)$

Finally in Figure 3, we consider $P(n,1)$ for $n = 6, \ldots, 120$. This figure illustrates how the

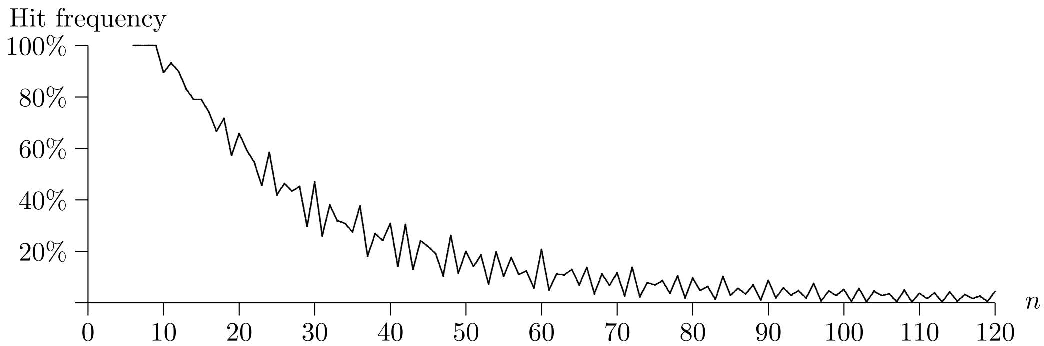

Figure 4: % of shots absorbed by all MIR based facets for $P(n, 1)$

frequency of hits for MIR based facets decreases as n increases. We would like to emphasize that the number of scaled MIR facets increases linearly, and the number of two-step MIR facets increases quadratically with n, whereas the total number of facets of $P(n, 1)$ increases exponentially [11]. Also note that even for larger n the frequency of hits absorbed by MIR based facets is non-negligible.

2.4 On the most important facet of $P(n, r)$

As we discussed earlier, our experiments suggest that the most important facets are t-scaled MIR facets where t is a divisor of n, especially the largest divisor. This fact will be relevant later on in Section 3.2 where we compare template functions based on important facets of $P(n, r)$ with MIR based inequalities. We believe that *if t is the largest divisor of n, then the t-scaled MIR facet (when it exists) is the most important facet of $P(n, r)$.* However, we do not expect such a role for other divisors, independent of r, as Theorem 5 only guarantees the invariance of the t-scaled MIR – where t is the largest divisor of n – over isomorphic master polyhedra. For example, for $P(100, 4)$, the most important MIR based facet is the 10-scaled MIR facet (neither 25 nor 50 are valid scaling parameters). This facet of $P(100, 4)$ is isomorphic to the 30-scaled MIR of $P(100, 28)$, which is therefore the most important MIR based facet of $P(100, 28)$.

Hunsaker [14] presents some sufficient conditions for a facet of $P(n, r)$ to be at least as important as another. His results yield a partial order on the facets of $P(n, r)$ consistent with the shooting experiment notion of importance. In every instance in [14] where this partial order rates a single facet as being at least as important as the rest, that facet is the t-scaled facet of $P(n, r)$ with t the largest valid divisor of n.

In a shooting experiment, a vector d is said to hit the facet of $P(n, r)$ defined by $\bar{\eta}^T w \geq 1$ if

$$\bar{\eta}^T d \;=\; \min_{\eta \in \Gamma(n,r)} \{\eta^T d\} \tag{15}$$

where $\Gamma(n, r)$ is the set of extreme points of (8)-(11). Note that in our experiments we use random vectors $d = [X_1, X_2, \ldots, X_{n-1}]$ where X_is are independent and identically distributed (i.i.d.) random variables. For each $\eta \in \Gamma(n, r)$, we can define a random variable

$Z(\eta)$ where $Z(\eta) = \sum_i \eta_i X_i$. The probability that a shot d hits a facet $\bar{\eta}^T w \geq 1$ is then equal to the probability that $Z(\bar{\eta}) \leq Z(\eta)$ for all $\eta \in \Gamma(n,r)$. Clearly the random variables $Z(\eta)$ for $\eta \in \Gamma(n,r)$ are not independent as the same vector d is used in their definition. We next show a basic property of $Z(\eta)$.

Proposition 6 *Let $X_1, \ldots, X_{n-1}$ be i.i.d. random variables and let $\eta, \beta \in \Gamma(n,r)$ for $P(n,r)$. Then $P(Z(\eta) > Z(\beta)) = P(Z(\eta) < Z(\beta))$.*

Proof. We will show that $P(Z(\eta) - Z(\beta) > 0) = P(Z(\eta) - Z(\beta) < 0)$. For coefficient indices of η and β, consider addition and subtraction to be modulo n. Using the fact that $\eta_i + \eta_{r-i} = \beta_i + \beta_{r-i} = 1$ for $i = 1, \ldots, n-1$ with $i \neq r$, we can write $Z(\eta) - Z(\beta)$ as $Y_1 - Y_2$, where

$$\begin{aligned} Y_1 &= \sum_{\eta_i - \beta_i > 0} (\eta_i - \beta_i) X_i, \quad \text{and} \\ Y_2 &= \sum_{\eta_i - \beta_i < 0} -(\eta_i - \beta_i) X_i = \sum_{\eta_i - \beta_i > 0} (\eta_i - \beta_i) X_{r-i}. \end{aligned}$$

Therefore Y_1 and Y_2 are identically distributed and $P(Y_1 - Y_2 > 0) = P(Y_1 - Y_2 < 0)$. ■

Proposition 6 says that if any two facets from $\Gamma(n,r)$ are compared, while ignoring the other facets, then they will be hit equally frequently in the shooting experiment. Geometrically this means that their projections on the unit-sphere, after removing the remaining facets in $\Gamma(n,r)$, have the same area. It is also possible to show that the expected value of $Z(\eta)$ is the same for all $\eta \in \Gamma(n,r)$. This is interesting because our shooting experiments suggest that some scaled MIR facets are more important than the rest of the facets. One explanation for the difference in importance is the difference in the variance of $Z(\eta)$ for $\eta \in \Gamma(n,r)$. The variance of $Z(\eta)$ can be written as

$$var(Z(\eta)) = \sum_{i=1}^{n-1} var(\eta_i X_i) = \sum_{i=1}^{n-1} \eta_i^2 var(X_i) \quad = \quad \sigma^2 \sum_{i=1}^{n-1} \eta_i^2$$

where σ^2 is the variance of X_i. Using the r-additivity of the facet coefficients (i.e., $\eta_i = 1 - \eta_j$ for $i + j = r \bmod n$), and defining η_0 to be 0, we can write

$$\sum_{i=0}^{n-1} \eta_i^2 \;=\; \frac{1}{2} \sum_{i=0}^{n-1} (\eta_i^2 + \eta_{r-i}^2) \;=\; \frac{1}{2} \sum_{i=0}^{n-1} (\eta_i^2 + (1-\eta_i)^2) \leq \frac{n}{2}$$

for all $\eta \in \Gamma(n,r)$. For n even and r odd, let $\bar{\eta}^T w \geq 1$ denote the $n/2$-scaled MIR facet. Notice that $\sum_{i=1}^{n-1} \bar{\eta}_i^2 = n/2$ and therefore $Z(\bar{\eta})$ has the largest variance among all $Z(\eta)$. In other words, $Z(\bar{\eta})$ is more likely to produce extreme values, and we believe this makes it significantly more likely to give the term that attains the minimum in (15). We believe that a similar property holds for t-scaled MIRs, with t the largest divisor of n.

3 Comparing template inequalities

In this section we compare different template inequalities for the set Y. If $\eta^T x \geq 1$ and $\beta^T x \geq 1$ are valid inequalities for Y, and $\eta \leq \beta$, then the first inequality is obviously preferable to the second. How does one compare valid inequalities when such obvious criteria do not apply? We adopt the approach proposed in Cornuéjols, Li and Vandenbussche [4], where the authors compare t-scaled MIR inequalities, for different t, by comparing their coefficients component-wise.

Definition 7 *Let $h_1(x)$ and $h_2(x)$ be two functions defined over $[0,1]$, and let $\mathcal{X}$ be a random variable distributed uniformly over $[0,1]$.*
(i) h_1 dominates *h_2 if $h_1(v) \leq h_2(v)$ for all $v \in [0,1]$, and if there is some interval $(a,b) \subset [0,1]$ such that $h_1(v) < h_2(v)$ for all $v \in (a,b)$.*
(ii) h_1 dominates *h_2 in a* probabilistic sense *if $\mathcal{P}(h_1(\mathcal{X}) < h_2(\mathcal{X})) > \mathcal{P}(h_1(\mathcal{X}) > h_2(\mathcal{X}))$.*

Based on Proposition 4 we know that the *template inequality*.

$$\sum_{j \in J} \frac{h(a_j)}{h(b)} x_j \geq 1, \tag{16}$$

is valid for Y (we assume $h(b) \neq 0$ throughout). Let h_1 and h_2 be two template functions such that $h_1(b) = h_2(b)$. If h_1 dominates h_2 then the template inequality generated by h_1 implies the inequality generated by h_2. Notice that if h_1 and h_2 arise from the canonical master polyhedron for Y, then $h_1(b) = h_2(b) = 1$. Further, if the coefficients of Y are chosen randomly from $[n_1, n_2]$, where n_1 and n_2 are integers, then the fractional parts of the coefficients of Y are uniformly distributed over $[0,1]$. If h_1 dominates h_2 in a probabilistic sense, then it is more likely that $h_1(a_j) < h_2(a_j)$ than $h_2(a_j) < h_1(a_j)$ for an index $j \in J$.

3.1 Properties of template functions

We next study some basic properties of template functions and show that there is no dominance relationship between them. This generalizes a result of Cornuéjols, Li and Vandenbussche [4] on the relative strength of scaled MIR inequalities. We start with a property of facets of $P(n,r)$ that follows from r-additivity.

Proposition 8 *Let $\sum_{i=1}^{n-1} \eta_i w_i \geq 1$ and $\sum_{i=1}^{n-1} \beta_i w_i \geq 1$ be non-trivial facets of $P(n,r)$.*
(i) The average value of the facet coefficients η_i is $n/(2(n-1))$, and
(ii) $|\{i : \eta_i > \beta_i\}| = |\{i : \beta_i > \eta_i\}|$.

Proof. For coefficient indices of (η_i) and (β_i), consider addition and subtraction to be modulo n. Define η_0 to be zero. Then

$$\sum_{i=1}^{n-1} \eta_i = \sum_{i=0}^{n-1} \eta_i = 1/2 \sum_{i=0}^{n-1} (\eta_i + \eta_{r-i}) = n/2.$$

The last equality is easily derived from the r-additivity of facet coefficients, and part (i) follows. Part (ii) follows from the fact that $\eta_i > \beta_i \Leftrightarrow \beta_{r-i} > \eta_{r-i}$ for $i = 1, \ldots, n-1$. ∎

Proposition 8 shows that the coefficients of all non-trivial facets of $P(n, r)$ have the same average value. Also, given two facet-defining inequalities, either inequality has the same number of coefficients which exceed the corresponding coefficients in the other. This, in some sense, shows that there is no probabilistic dominance between the facets of $P(n, r)$. To show an analogous statement for template functions, we next derive a property of template functions similar to the r-additivity of facet coefficients.

Lemma 9 *If $h(x)$ is a template function derived from a facet of $P(n, r)$, then $h(x) + h((r/n) - x) = h(r/n)$, $\forall x \in [0, 1]$.*

Proof. By the definition of template functions and because of (9), this is true if $x = i/n$ for any integer i between 0 and n. Let $x = (i + \delta)/n$ for some $\delta \in (0, 1)$ and some integer i between 0 and $n - 1$. Now,

$$h(x) = \delta h(\frac{i}{n}) + (1-\delta)h(\frac{i+1}{n}), \quad \text{and} \quad h(\frac{r}{n} - x) = \delta h(\frac{r-i}{n}) + (1-\delta)h(\frac{r-i-1}{n}).$$

Adding up the right-hand sides of the equations above we get $\delta(r/n) + (1 - \delta)(r/n) = r/n$ and the lemma follows. ∎

We now show that for any two template functions derived from non-trivial facets of a common master polyhedron, neither dominates the other in a probabilistic sense.

Theorem 10 *Let $h_1(x)$ and $h_2(x)$ be two template functions associated with some $P(n, r)$. Let $\mathcal{X}$ be a random variable which is uniformly distributed over $[0, 1]$.*
(i) The expected value of $h_1(\mathcal{X})$ is $1/2$.
(ii) $\mathcal{P}(h_1(\mathcal{X}) > h_2(\mathcal{X})) = \mathcal{P}(h_1(\mathcal{X}) < h_2(\mathcal{X}))$.

Proof. Define $\beta = r/n$. Observe that h_1 and h_2 are piece-wise linear functions, with $h_i(x) + h_i(\beta - x) = h_i(\beta) = 1$ and $h_i(0) = h_i(1) = 0$ for $i = 1, 2$. Hence

$$\int_0^\beta h_1(x)dx = \int_0^{\beta/2} (h_1(x) + h_1(\beta - x))dx = \beta/2.$$

Similarly $\int_\beta^1 h_1(x)dx = (1 - \beta)/2$, and part (i) of the proposition follows.

As h_1 and h_2 are piece-wise linear and have n linear segments in $[0, 1]$, there are fewer than n points at which they cross in $[0, 1]$. Define a finite ordered set $S = (s_1, s_2, \ldots, s_m)$, containing the x-coordinates of these crossing points and the x-coordinates of the end points of the linear segments of h_1 andh_2 in $[0, 1]$. Let the members of S be arranged by increasing magnitude. Then S has the following properties:
(a) $\{0, \frac{\beta}{2}, \beta, \frac{1+\beta}{2}, 1\} \subseteq S \subseteq [0, 1]$;
(b) In the interval (s_i, s_{i+1}), for $1 \le i \le m - 1$, if one of the functions h_1 and h_2 is strictly

greater than the other at one point, it is strictly greater throughout the interval;
(c) If $x \in S$, then so does $\beta - x$; here subtraction is taken modulo 1.
Now consider an interval $(s_i, s_{i+1}) \subseteq (0, \beta/2)$. If $h_1(x) > h_2(x)$ for all x in (s_i, s_{i+1}), then because of Lemma 9, $h_2(x) > h_1(x)$ for all x in $(\beta - s_{i+1}, \beta - s_i) \subseteq (\beta/2, \beta)$. Arguing similarly for intervals in $(\beta/2, \beta), (\beta, \frac{1+\beta}{2})$ and $(\frac{1+\beta}{2}, 1)$, it follows that $\mathcal{P}(h_1(\mathcal{X}) > h_2(\mathcal{X})) \leq \mathcal{P}(h_1(\mathcal{X}) < h_2(\mathcal{X}))$. Reversing the roles of h_1 and h_2, the result follows. ∎

Suppose we choose Y to be a row of a simplex tableau associated with the linear relaxation of an integer program. Over many different integer programs, it is reasonable to assume that the fractional parts of the coefficients in Y will be uniformly distributed over $[0, 1]$. The above theorem indicates that for two template functions h_1 and h_2 defined as in Theorem 10 with $h_1(b) = h_2(b)$, neither is likely to yield stronger inequalities than the other, when relative strength is measured by counting the number of coefficients of one inequality that exceed the corresponding coefficients in the other inequality. Also, the expected value of any coefficient in the template inequality generated by h_1 is $1/(2h_1(b))$. If $h_1(b) \neq h_2(b)$, then one can argue that h_1 is preferable to h_2 if $h_1(b) > h_2(b)$. We discuss and use this idea in Section 3.2.

We next restrict our attention to template functions arising from scaled MIR facets of $P(n, r)$. We define the *t-scaled MIR function with parameter c*, where t is an integer, in terms of the MIR function described earlier:

$$f^{t,c}(v) = f^{tc}(tv).$$

Then the t-scaled MIR inequality for Y can be written as $\sum_{j \in J} f^{t,b}(a_j)x_j \geq 1$. In Figure 1 the template function of the 6-scaled MIR cut of $P(13, 12)$ does not coincide with the 6-scaled MIR function. This is because the scaling coefficient t does not satisfy the following property.

Proposition 11 *If t is a divisor of n, then the template function $h(v)$ of the t-scaled MIR facet of $P(n, r)$ coincides with the t-scaled MIR function $f^{t,r/n}(v)$.*

Proof. It is sufficient to observe that $h(x) = f^{t,r/n}(x)$ whenever $f^{t,r/n}(x) = 0$ *or* 1. ∎

In a recent paper [4], Cornuéjols, Li, and Vandenbussche compare the GMIC (1-scaled MIR inequality) with t-scaled MIR inequalities for different values of $t \in Z$. For a uniformly distributed random variable $\mathcal{X}$, they compute $\mathcal{P}(f^{t,b}(\mathcal{X}) > f^{1,b}(\mathcal{X}))$ and $\mathcal{P}(f^{t,b}(\mathcal{X}) < f^{1,b}(\mathcal{X}))$. They also show that these probabilities are equal [Theorem 1(ii),Theorem 3]. We can derive this result by observing that for any two positive integers t_1,t_2, the t_1-scaled and t_2-scaled MIR functions with a rational parameter b coincide with appropriate template functions for some common $P(n, r)$. In particular, let $\bar{n}, \bar{r} \in Z$ be such that $\hat{b} = \bar{r}/\bar{n}$, and define $n = t_1 t_2 \bar{n}$ and $r = t_1 t_2 \bar{r}$. Applying Proposition 11 and Theorem 10 we obtain the following corollary.

Corollary 12 *Let b be a rational number and let t_1 and t_2 be positive integers. If $\mathcal{X}$ is uniformly distributed over $[0, 1]$, then $\mathcal{P}(f^{t_1,b}(\mathcal{X}) > f^{t_2,b}(\mathcal{X})) = \mathcal{P}(f^{t_1,b}(\mathcal{X}) < f^{t_2,b}(\mathcal{X}))$.*

We know that the expected value of template functions, and therefore of scaled MIR functions, in the interval $[0, 1]$ is $1/2$. We close this section by showing another property of scaled MIR functions.

Proposition 13 *Let $t \in Z$ and $0 < b < 1$ be such that $tb \notin Z$. If $\mathcal{X}$ is distributed uniformly in $[n_1, n_2)$, for $n_1, n_2 \in Z$, then $\mathcal{Y} = f^{t,b}(\mathcal{X})$ is distributed uniformly in $[0, 1)$.*

Proof. We will show that $\mathcal{P}(\mathcal{Y} \leq \delta) = \delta$ for $0 \leq \delta \leq 1$. First note that if $\mathcal{X}$ is distributed uniformly in $[n_1, n_2)$, then $\hat{\mathcal{X}} = t\mathcal{X} - \lfloor t\mathcal{X} \rfloor$ is distributed uniformly in $[0, 1)$. Define $\beta = tb$. Therefore given $\delta \in [0, 1)$, we can write

$$\begin{aligned} \mathcal{P}(\mathcal{Y} \leq \delta) &= \mathcal{P}(\mathcal{Y} \leq \delta, \hat{\mathcal{X}} < \hat{\beta}) + \mathcal{P}(\mathcal{Y} \leq \delta, \hat{\mathcal{X}} \geq \hat{\beta}) \\ &= \mathcal{P}(\hat{\mathcal{X}}/\hat{\beta} \leq \delta) \; + \; \mathcal{P}((1 - \hat{\mathcal{X}})/(1 - \hat{\beta}) \leq \delta)) \\ &= \mathcal{P}(\hat{\mathcal{X}} \leq \hat{\beta}\delta) \; + \; \mathcal{P}(\hat{\mathcal{X}} \geq 1 - (1 - \hat{\beta})\delta)) \\ &= \delta\,\hat{\beta} + \delta(1 - \hat{\beta}) = \delta \end{aligned}$$ ∎

This is an interesting result especially when combined with the fact that coefficients of scaled MIR facets of $P(n, r)$ do not necessarily have the same variance. In particular, as discussed in Section 2.4 , the coefficients of the $n/2$-scaled MIR facet (when it exists) give the largest variance among all facets. Proposition 13, on the other hand, shows that for a template function h the variance (or, distribution) of $h(\mathcal{X})$ does not necessarily depend on the variance (or, distribution) of the facet coefficients used to derive h. In particular, the variance and distribution of $f^{n/2,b}(\mathcal{X})$ are the same as, respectively, the variance and distribution of $f^{1,b}(\mathcal{X})$.

3.2 Comparing template and MIR inequalities

As discussed in Section 2, Gomory, Johnson and Evans [13] propose using template inequalities based on important facets of $P(n, r)$ with small n. Based on our computational experiments, this approach favors using template inequalities based on scaled MIR facets, and two-step MIR facets.

Using Definition 7, we next compare the above classes of template inequalities with scaled MIR inequalities and two-step MIR inequalities for Y We first show that template inequalities are almost always dominated by the 1-scaled MIR inequality (GMIC) in a probabilistic sense. We then show that template inequalities based on the "most important" facets of $P(n, r)$ are dominated by appropriate MIR based inequalities.

Theorem 14 *Let $h(x)$ denote the template function of a facet of $P(n, r)$, and let $b \in R$. If $nb \notin Z$, then the MIR function $f^b(x)$ dominates $h(x)/h(b)$ in a probabilistic sense.*

Proof. As $f^b(x) = f^{\hat{b}}(x)$ for all $x \in R$, and $h(b) = h(\hat{b})$, we can assume that $b \in (0, 1)$. Define $h'(x) = h(x)/h(b)$ and observe that the sub-additivity of h implies that h' is sub-additive. Therefore, for all $x \in (0, 1)$

$$h'(x) + h'(b - x) \geq h'(b) = 1 = f^b(x) + f^b(b - x).$$

If $f^b(x) > h'(x)$ for some $x \in (0,1)$, then $f^b(b-x) < h'(b-x)$. Therefore, if $f^b(x) > h'(x)$ for all x in some interval $(u_1, u_2) \subseteq (0,1)$, then $f^b(x) < h'(x)$ for all x in $(b-u_2, b-u_1)$. In other words, for every interval in which $f^b(x) > h'(x)$, there is a corresponding interval of the same length where $f^b(x) < h'(x)$. This implies that

$$\mathcal{P}(f^b(\mathcal{X}) < h'(\mathcal{X})) \geq \mathcal{P}(f^b(\mathcal{X}) > h'(\mathcal{X})) \tag{17}$$

where $\mathcal{X}$ is uniformly distributed over $[0,1]$. To show that the inequality in (17) is strict, we next identify an interval $(\bar{u}_1, \bar{u}_2)$ such that $f^b(x) < h'(x)$ for all x in $(\bar{u}_1, \bar{u}_2)$ and $f^b(x) \leq h'(x)$ for all x in $(b-\bar{u}_2, b-\bar{u}_1)$. Let $t = \lfloor nb \rfloor$ so that $t/n < b < (t+1)/n$. Using the definition of h (i.e., interpolation), $h'(b) = \alpha h'(t/n) + (1-\alpha)h'((t+1)/n)$ for some $\alpha \in (0,1)$, and therefore,

$$\max\left\{h'(t/n), h'((t+1)/n)\right\} \geq h'(b) = 1.$$

Assume $h'(t/n) \geq 1$. Since $f^b(t/n) < 1$ and both f^b and h' are linear in $[t/n, b]$, we have $f^b(x) < h'(x)$ for all x in $(t/n, b)$. In addition, sub-additivity of h' implies that $nh'(b/n) \geq h'(b) = 1$. That is, $h'(b/n) \geq 1/n$. Furthermore, as h' is linear in $[0, 1/n]$ and $h'(0) = 0$, it follows that

$$h'(x) \geq x/b = f^b(x) \text{ for } x \in (0, 1/n).$$

As $b - t/n < 1/n$, the above inequality holds for all x in $(0, b-t/n)$. This proves the claim with $\bar{u}_1 = t/n$ and $\bar{u}_2 = b$. As $\mathcal{P}(\bar{u}_1 > \mathcal{X} > \bar{u}_2) = b - t/n$, we have

$$\mathcal{P}(f^b(\mathcal{X}) < h'(\mathcal{X})) - \mathcal{P}(f^b(\mathcal{X}) > h'(\mathcal{X})) \geq (b - t/n) > 0.$$

If $h'(t/n) < 1$, then $h'((t+1)/n) \geq 1$ and the claim is proved in a similar fashion. ∎

Theorem 14 is a relatively strong result in that a simple inequality almost always dominates template inequalities, in a probabilistic sense. In practice, if Y is constructed from a row of a simplex tableau, it would be impractical to work with the canonical master polyhedron. Let n' be the smallest integer such that $n'b$ is integral. Template inequalities from $P(n,r)$ with $n < n'$ would then be dominated by the GMIC in a probabilistic sense, as the condition $nb \notin Z$ would be true. For template functions arising from scaled MIR facets, we can obtain stronger dominance results. We use the following definition in the proof of our next result: for $c \in R$ and $t \in Z$, define $\hat{c}^t = tc - \lfloor tc \rfloor$.

Theorem 15 *Let $h(x)$ be the template function of the t-scaled MIR facet of $P(n,r)$ and let $b \in R$. Let $\mathcal{X}$ be uniformly distributed over $[0,1]$. If t is a divisor of n and $nb \notin Z$, then*
(i) the t-scaled MIR function $f^{t,b}(x)$ dominates $h(v)/h(b)$;
(ii) $\mathcal{P}(f^{t,b}(\mathcal{X}) < h(\mathcal{X})/h(b)) \geq \min\{\hat{b}^t, 1 - \hat{b}^t\}$.

Proof. As in the previous theorem, we can assume that $b \in (0,1)$. Let ρ stand for r/n, and f for $f^{t,b}$, and h' for $h/h(b)$. As t is a divisor of n, Lemma 11 implies that $h = f^{t,\rho}$. Both f and h' are periodic functions with period $1/t$ and $f(0) = h'(0) = 0$. Part (i) of the theorem will follow if we show that $f \leq h'$ in the interval $[0, 1/t]$, and that this inequality is strict for some non-empty sub-interval of $[0, 1/t]$. Now, f and h' are piece-wise linear functions, with

f being linear in the intervals $[0, \hat{b}^t/t]$ and $[\hat{b}^t/t, 1/t]$, and h' being linear in the intervals $[0, \hat{\rho}^t/t]$ and $[\hat{\rho}^t/t, 1/t]$. As $h(\hat{b}^t/t) = h(b)$, we know that $f(\hat{b}^t/t) = h'(\hat{b}^t/t) = 1$. Therefore, showing that

$$f(\hat{\rho}^t/t) < h'(\hat{\rho}^t/t) \qquad (18)$$

is enough to prove part (i). As $nb \notin Z$, we know that $h(b) < 1$. Because $h(\hat{\rho}^t/t) = 1$, it follows that $h'(\hat{\rho}^t/t) > 1$. As $f(x)$ is never greater than 1, (18) is true and part (i) of the theorem follows. For part (ii), we know that $\hat{b}^t \neq \hat{\rho}^t$, and we will consider the following two cases.

Case 1: $\hat{b}^t < \hat{\rho}^t$. In this case, $f(x) < h'(x)$ if $x \in (\hat{b}^t/t, 1/t)$. Because of the periodicity of f and h', it follows that $f(\mathcal{X}) < h'(\mathcal{X})$ if $\hat{\mathcal{X}}^t > \hat{b}^t$, and the probability of this happening is $1 - \hat{b}^t$.

Case 2: $\hat{b}^t > \hat{\rho}^t$. An easy calculation shows that $f(\mathcal{X}) < h'(\mathcal{X})$ if $\hat{\mathcal{X}}^t < \hat{b}^t$. The probability of the latter being true is $\hat{b}^t$ and the theorem follows. ∎

As discussed in Section 2, our computational results indicate that the most important facets of $P(n, r)$ are scaled MIR facets with the scaling factor being a divisor of n. We also note that 9 of the 13 important facets presented in [13] are scaled MIR facets with this property. Theorem 15 implies that template inequalities based on such facets will be dominated by appropriate scaled MIR inequalities. In Figure 3.2 we give an application of Theorem 15, where $t = 2, b = 0.75, n = 10$ and $r = 7$. The solid line stands for $f^{2,0.75}$, and the dashed line for $h(x)/h(0.75)$, where h is the template function of the 2-scaled MIR facet of $P(10, 7)$. Because of Lemma 11, $h = f^{2,0.7}$.

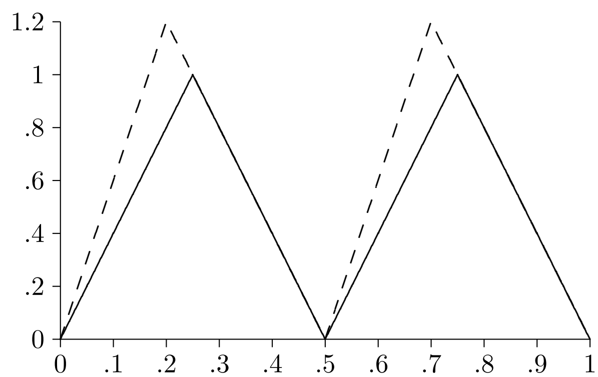

Figure 5: Domination of a template inequality by a scaled-MIR inequality

We will now prove a result analogous to Theorem 15 about template functions arising from 1-scaled two-step MIR facets, described in Section 1.1. Such facets are obtained by setting t to 1 in (7). In particular, we will consider a sub-class of these facets obtained by restricting the parameter Δ such that $n/2 > \Delta > r/2$ and $r > \Delta$. Under these assumptions on Δ, and letting $b = r/n$ and $\alpha = \Delta/n$, we have $1/\alpha > \lceil b/\alpha \rceil = 2$, and $\rho = b - \alpha = (r - \Delta)/n$. The two-step MIR function $g^{b,\alpha}$ becomes

$$g^{b,\alpha}(v) = \begin{cases} \dfrac{v(1-2\rho)}{2\rho(1-b)} & \text{if } 0 \le v \le b-\alpha, \\[2ex] \dfrac{1/2-v}{1-b} & \text{if } b-\alpha < v \le \alpha, \\[2ex] \dfrac{v(1-2\rho)-(b-2\rho)}{2\rho(1-b)} & \text{if } \alpha < v \le b, \\[2ex] \dfrac{1-v}{1-b} & \text{if } b < v \le 1. \end{cases} \tag{19}$$

Thus $g^{b,\alpha}$ is piece-wise linear, and has four linear segments. Also, the slope of $g^{b,\alpha}$ in the first and third intervals is the same, and its slope in the second and fourth intervals is the same. This is why this function is referred to as a "2slope" function in [10, 11]. Observe that if α is chosen to be $b/2$ in (19), then $b-\alpha=\alpha$, and the second interval in (19) disappears, and $g^{b,\alpha}(v) = f^b(v)$ for all $v \in [0,1]$.

Consider the two-step MIR facet of $P(n,r)$ generated with Δ chosen as above, and let $h(v)$ be the corresponding template function. As the underlying two-step MIR function $g^{r/n,\Delta/n}(v)$ changes slope only at $(r-\Delta)/n$, Δ/n and r/n, which are multiples of $1/n$, the template function $h(v)$ coincides with $g^{r/n,\Delta/n}(v)$. Under fairly general conditions, we next show how to find valid inequalities for Y that dominate inequalities derived from such template functions. As before, we use $\hat{c}$ to denote $c - \lfloor c \rfloor$ for $c \in R$.

Theorem 16 *Let $h(v)$ be the template function of a 1-scaled two-step MIR facet of $P(n,r)$, with parameter Δ such that $n/2 > \Delta > r/2$ and $r > \Delta$. Let $b \in R$ such that $\hat{b} \neq r/n$. If $\Delta/n < \hat{b}$ and $\hat{b}/2 < r/n$, then for some $\alpha \in R$ the two-step MIR function $g^{b,\alpha}(v)$ dominates $h(v)/h(b)$.*

Proof. As $g^{b,\alpha} = g^{\hat{b},\alpha}$, without loss of generality, we assume that $b \in (0,1)$ and $b = \hat{b}$. Let b' and α' stand for r/n and Δ/n, respectively. Then $h(v)$ is the 1-scaled two-step MIR function $g^{b',\alpha'}(v)$ and is given by (19).

We will find an α such that

$$g^{b,\alpha}(v) \le h(v)/h(b) \text{ for all } v \in [0,1], \tag{20}$$

and the inequality is strict for some sub-interval of $(0,1)$. For convenience, let g stand for $g^{b,\alpha}$. Also let $\rho = b-\alpha$, and $\rho' = b'-\alpha'$. Once α is chosen, (20) will follow if we show that the inequality in (20) holds at the points in

$$S = \{b'-\alpha',\ b-\alpha,\ \alpha',\ \alpha,\ b',\ b\}.$$

This is because g and h are piece-wise linear functions in the interval $[0,1]$, with value 0 at $v=0$ and $v=1$, and change slope only at points in S. Further, if we show that the inequality in (20) is strict for some point in S, then the theorem will follow. Now, $h(b')=1$ and $h(b)<1$. Therefore, $h(v)/h(b)$ equals 1 at $v=b$, and is at greater than 1 at $v=b'$.

At both these points, $g(v) \leq 1$ for any choice of α. We will therefore focus on the first four points in S.

Case 1: Assume $b' < b$. Define $\alpha = \max\{\alpha', b/2\}$. It trivially follows that $\alpha \in [b/2, 1/2)$, and $\alpha < b$. Therefore $1/\alpha > \lceil b/\alpha \rceil = 2$, and $g^{b,\alpha}$ is well-defined and has the form in (19). By assumption $b/2 < b'$ and $\alpha' < b'$, and therefore $\alpha' \leq \alpha < b'$. Also, $b' - \alpha' < b - \alpha$; if $\alpha = b/2$, then $b' - \alpha' < b'/2 < b/2 = b - \alpha$, and the other case is trivial.

As $b' < b$, $h(b) = (1-b)/(1-b')$. Therefore in the interval $[b'-\alpha', \alpha']$, $h(v)/h(b)$ becomes $(1/2 - v)/(1-b)$, which is precisely the form of $g(v)$ in the interval $[b-\alpha, \alpha]$. Also, $h(v)/h(b)$ increases linearly as v increases from α' to b', and $g(v)$ decreases linearly as v decreases from $b - \alpha$ to 0. As $b' - \alpha' < b - \alpha$, and $\alpha' \leq \alpha < b'$, we can conclude that (20) holds at the first four points in S, which proves the theorem.

See Figure 6 for illustration. In the the first example, we let $b = .7, b' = .6, \alpha' = .4$ and $h = g^{.6,.4}$. From the proof, $\alpha = \max\{\alpha', b/2\} = .4$. The solid line represents $g^{.7,.4}$ which dominates $h(v)/h(b)$, represented by the dashed line. In the second example, $b = .7, b' = .6, \alpha' = .32$. We choose $\alpha = \max\{\alpha', b/2\} = .35$. In this case, $g^{b,\alpha}$ is just $f^{.7}$.

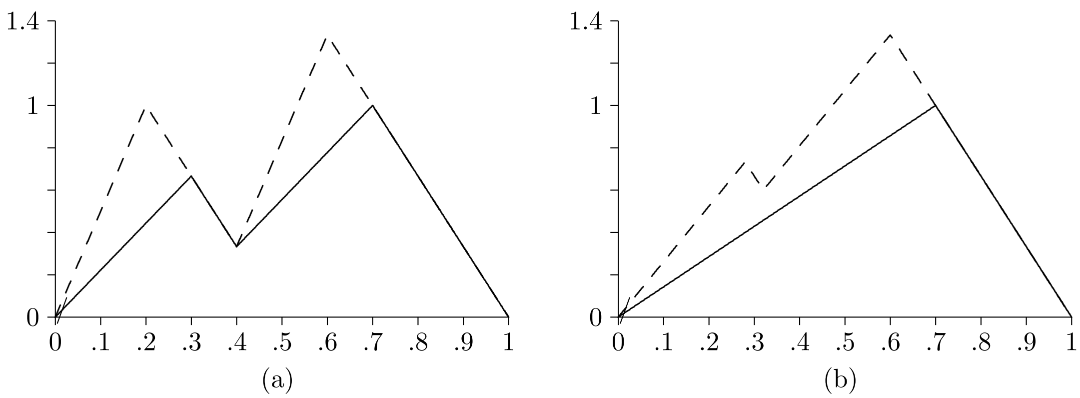

Figure 6: Dominating two-step MIR template inequalities when $r/n < b$.

Case 2: Let $b < b'$. We will find an α such that the functions $g^{b,\alpha}(v)$ and $h(v)/h(b)$ have the same (one-sided) slope at $v = 0$. Let $\rho' = b' - \alpha'$ and define δ' by

$$\delta' = \frac{b' - 2\rho'}{1 - 2\rho'}.$$

Because $\alpha' < 1/2$, and from the definition of δ' it follows that

$$0 < \delta' < 1/2 \text{ and } \frac{b' - \delta'}{2(1-\delta')} = \rho' = b' - \alpha'.$$

We now define α by

$$\rho = b - \alpha = \frac{b - \delta'}{2(1-\delta')}.$$

We need to verify that α chosen in this manner is valid. From the definitions of α' and α above, it follows that

$$\alpha' - \alpha = (b' - b) - \frac{b' - b}{2(1 - \delta')}.$$

As $b' - b > 0$ and $\delta' < 1/2 \Rightarrow 2(1 - \delta') > 1$, the right-hand side of the above equation is positive. Therefore $\alpha < \alpha' < 1/2$. By assumption, $\alpha' < b$, and therefore $b - \alpha$ is positive. As $(b - \delta')/(1 - \delta') < b$, we see from the above that

$$b - \alpha < b/2 \Rightarrow \alpha > b/2 \Rightarrow \lceil b/\alpha \rceil = 2.$$

Therefore $g^{b,\alpha}$ is well-defined and has the form in (19). Because $b \in [\alpha', b']$, therefore

$$h(b) = \frac{b(1 - 2\rho') - (b' - 2\rho')}{2\rho'(1 - b')}.$$

One can verify that the slope of $h(v)/h(b)$ in the intervals $[0, b' - \alpha']$ and $[\alpha', b']$ is equal to the slope of $g(v)$ in the interval $[0, b - \alpha]$ and the interval $[\alpha, b]$. In fact, we chose α to ensure this condition. Comparing the expressions for ρ and ρ' above, we see that $b - \alpha < b' - \alpha'$. Observe that $g(v)$ decreases linearly as v increases from $b - \alpha$ to α, and decreases linearly as v decreases from b to α. The above observations, combined with the fact that $h(b)/h(b) = g(b) = 1$, and $h(b - \alpha)/h(b) = g(b - \alpha)$, imply (20).

See Figure 7 for illustration. Here $h = g^{.8,.45}$ and $b = .7$. Therefore, $b' = .8, \alpha' = .45$. We then choose $\alpha = .425$ and $g^{.7,.425}$, denoted by the solid line, dominates $h(v)/h(b)$ which is denoted by the dashed line.

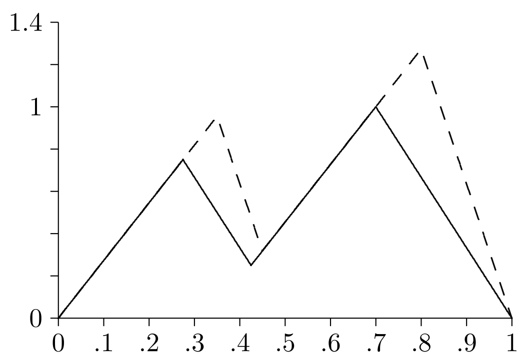

Figure 7: Dominating two-step MIR template inequalities when $r/n > b$.

■

It is shown in [5] that the "2slope facets" in [3] form a sub-class of scaled two-step MIR facets. A simple transformation can be used to show that Theorem 16 is also true for these 2slope facets, that is, associated template inequalities for Y are dominated by appropriately chosen scaled two-step MIR inequalities.

4 Concluding Remarks

In this paper, we analyzed the shooting experiment of Gomory, and extended the experimental results in [13, 6]. We verified the importance of MIR-based facets for higher-dimensional master polyhedra. We also examined the strength of template inequalities derived from MIR based facets. Any valid inequality for Y is implied by a convex combination of facet-defining inequalities for the canonical master polyhedron of Y, though such a decomposition is typically hard to find. However, our results show that this decomposition can easily be derived for template inequalities associated with MIR based facets. This suggests that the interpolation procedure is not an effective way to generate valid inequalities for Y.

In addition, viewing scaled MIR inequalities as template functions of master cyclic group polyhedra, we show that some results of Cornuéjols, Li, and Vandenbussche [4] on scaled MIR inequalities follow from general properties of template functions .

We would like to emphasize that master polyhedra are important structures to get cutting planes for Y, as all facets of Y come from the canonical master polyhedron of Y. We feel that analyzing the polyhedral structure of small master polyhedra can yield families of inequalities which are facet-defining for all master polyhedra, in particular the canonical master polyhedron of Y. This approach has been studied in [3] and [5]. These inequalities could then be used as cutting planes for Y.

Acknowledgments: We would like to thank John Forrest for his help with using COIN-Clp for our shooting experiments.

References

[1] P. J. Acklam, An algorithm for computing the inverse normal cumulative distribution function, `http://home.online.no/~pjacklam/notes/invnorm/index.html`.

[2] D. Applegate, W. Cook, S. Dash, and M. Mevenkamp, QSopt Linear Programming Solver, `http://www.isye.gatech.edu/~wcook/qsopt`.

[3] J. Araoz, R.E. Gomory, E.L. Johnson, and L. Evans, Cyclic group and knapsack facets, *Mathematical Programming* **96** 377–408 (2003).

[4] G. Cornuejols, Y. Li and D. Vandenbussche, K-Cuts: A Variation of Gomory Mixed Integer Cuts from the LP Tableau, *INFORMS Journal on Computing* **15** 385–396 (2003).

[5] S. Dash and O. Gunluk, Valid inequalities based on simple mixed-integer sets, IBM Research Report RC22922, T. J. Watson Research Center, Yorktown Heights, New York, 2003.

[6] L. Evans, Cyclic Groups and Knapsack Facets with Applications to Cutting Planes, Ph.D. Thesis, Georgia Institute of Technology, Atlanta, Georgia, 2002.

[7] G. Fishman, *Monte Carlo*, Springer, New York (1995).

[8] J. J. Forrest, CLP: Coin LP, a native simplex solver, `http://www.ibm.com/developerworks/opensource/coin`.

[9] R.E. Gomory, Some Polyhedra Related to Combinatorial Problems, Journal of Linear Algebra and its Applications, **2**, 451–558 (1969).

[10] R.E. Gomory and E. Johnson, Some Continuous Functions Related to Corner Polyhedra I, *Mathematical Programming* **3** 23–85 (1972).

[11] R.E. Gomory and E. Johnson, Some Continuous Functions Related to Corner Polyhedra II, *Mathematical Programming* **3** 359–389 (1972).

[12] R.E. Gomory and E. Johnson, T-space and cutting planes, *Mathematical Programming* **96** 341–375 (2003).

[13] R.E. Gomory, E.L. Johnson, and L. Evans, Corner Polyhedra and their connection with cutting planes, *Mathematical Programming* **96** 321–339 (2003).

[14] B. Hunsaker, Measuring facets of polyhedra to predict usefulness in branch-and-cut algorithms, Ph.D. Thesis, Georgia Institute of Technology, Atlanta, Georgia, 2003.

[15] H.W. Kuhn, Discussion, in *Proceedings of the IBM Scientific Computing Symposium on Combinatorial Problems: March 16–18, 1964*, L. Robinson, ed., IBM, White Plains, New York (1966), pp. 118–121.

[16] H.W. Kuhn, On the origin of the Hungarian Method, in *History of mathematical programming: a collection of personal reminiscences* J.K. Lenstra, A.H.G. Rinnooy Kan, and A. Schrijver, eds., Elsevier Science Publishers B.V., The Netherlands (1991), pp. 77–81.

[17] H. Marchand and L. Wolsey, Aggregation and mixed integer rounding to solve MIPs, *Oper. Res.* **49**, 363–371 (2001).

[18] G.L. Nemhauser and L.A. Wolsey, *Integer and Combinatorial Optimization*, Wiley, New York (1988).

[19] S. Ross, *Introduction to Probability and Statistics*, John Wiley and Sons, New York (1987).