RC23169 (W0404-002) April 1, 2004
Mathematics

IBM Research Report

Identifying Important Facets of the Master Polyhedra

Sanjeeb Dash, Oktay Günlük
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Identifying important facets of the master polyhedra

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)

March 30, 2004

Abstract

We analyze and extend the shooting experiment described in Gomory, Johnson and Evans (2003), which is an empirical approach to identifying the "important" facets of the master cyclic group polyhedra. Evans (2002) and Gomory, Johnson and Evans (2003) suggest using these facets in the interpolation procedure of Gomory and Johnson (1972). The interpolation procedure is a method to generate cutting planes for general integer programs from facets of master cyclic group polyhedra.

1 Introduction

In a sequence of papers, Gomory [10], and later Gomory and Johnson [11, 12], studied the polyhedral structure of the master cyclic group polyhedron

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

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 $S \subseteq R^n$, $conv(S)$ denotes the convex hull of vectors in S. The following characterization of the nontrivial facets (i.e., excluding the non-negativity inequalities) of $P(n,r)$ is presented in [10].

Theorem 1 (Gomory [10]) *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

$$\eta_i + \eta_j \geq \eta_{(i+j)\, mod\, n} \quad \forall i,j \in \{1,\ldots,n-1\}, \tag{2}$$
$$\eta_i + \eta_j = \eta_r \quad \forall i,j \textit{ such that } r = (i+j)\, mod\, n, \tag{3}$$
$$\eta_j \geq 0 \quad \forall j \in \{1,\ldots,n-1\}, \tag{4}$$
$$\eta_r = 1. \tag{5}$$

Let $\sum_{i=1}^{n-1} \eta_i w_i \geq 1$ be a facet of $P(n,r)$. Let $h : R \rightarrow [0,1]$ 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 $h(v) = h(v+1)$ for all $v \in R$. We will call $h(v)$ a *facet interpolated template function*, abbreviated as a *template function*.

Gomory and Johnson [11] derived the following *sub-additivity* property of template functions from the sub-additivity property of facet coefficients in inequality (2):

$$h \text{ is a template function of } P(n,r) \Rightarrow h(x) + h(y) \geq h(x+y),\ \forall x,y \in R. \tag{6}$$

Based on this, Gomory and Johnson [11] also showed that template functions can be used to derive valid inequalities for general integer programs. In particular, let

$$Y = \left\{ x \in Z^{|J|}, y \in Z \ :\ \sum_{j \in J} a_j\, x_j + y = b, \quad x \geq 0 \right\} \tag{7}$$

where $\sum_{j \in J} a_j x_j + y = b$ could be derived from a row of the simplex tableau.

Proposition 2 (Gomory and Johnson [11]) *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 $\bar{n}$. Notice that Y can be viewed as a lower-dimensional face of an associated master polyhedron $P(\bar{n}, \bar{r})$ [10]. Therefore all facets of $conv(Y)$ can be obtained using template functions. However, $\bar{n}$ would be too large n general.

In a recent paper, Gomory, Johnson and Evans [13] 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. To identify the "important" facets of $P(n, r)$, they use a computational approach called the *shooting experiment*. We next give an overview of this procedure and then describe it in detail in Section 2 .

1.1 The shooting experiment

The *shooting experiment* is a randomized procedure, where, at each step, a random direction $d \geq 0$ is chosen and the facet f first encountered along the ray $\{\lambda d \,|\, \lambda \geq 0\}$ as λ increases from 0 is identified. One can view the facet f as being *hit* by the "shot" d. In [13], Gomory, Johnson and Evans 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" above 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 $P(n, r)$. 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 [13] and [7] the authors present results of their computational experiments performed on master polyhedra with $n \leq 30$. Their main observations 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 [5]. Evans [7] 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 [5]. We describe t-scaled MIR facets and two-step MIR facets in detail below.

1.2 MIR-based facets

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 Gomory mixed-integer cut (GMIC) [17]. Using this simple idea, in [5], we define t-scaled MIR inequalities. Let $\hat{c}$ denote $c - \lfloor c \rfloor$ for $c \in R$.

Definition 3 *For $b \in R$, the MIR function with parameter b is defined as:*

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

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

$$\sum_{j \in J} f^{tb}(ta_j)x_j \geq 1 \tag{8}$$

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.

Using similar ideas [5], we define the two-step MIR inequality based on the following function.

Definition 4 *For $b, \alpha \in R$ satisfying $\hat{b} > \alpha > 0$, and $1/\alpha \geq \lceil \hat{b}/\alpha \rceil > \hat{b}/\alpha$, the* two-step MIR function *with parameters b and α is defined as:*

$$g^{b,\alpha}(v) = \begin{cases} \dfrac{\hat{v}(1-\rho\tau) - k(v)(\alpha-\rho)}{\rho\tau(1-\hat{b})} & \text{if } \hat{v} - k(v)\alpha < \rho \\[2ex] \dfrac{k(v) + 1 - \tau\hat{v}}{\tau(1-\hat{b})} & \text{if } \hat{v} - k(v)\alpha \geq \rho, \end{cases}$$

where $\rho = \hat{b} - \alpha\lfloor \hat{b}/\alpha \rfloor$, $\tau = \lceil \hat{b}/\alpha \rceil$ 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

$$\sum_{j \in J} g^{tb,\alpha}(ta_j)x_j \geq 1 \tag{9}$$

is valid for Y.

When applied to $P(n,r)$ these inequalities take the form

$$\sum_{i=1}^{n-1} f^{tr/n}(ti/n)\, w_i \geq 1 \tag{10}$$

and

$$\sum_{i=1}^{n-1} g^{tr/n,\Delta/n}(ti/n)\, w_i \geq 1 \tag{11}$$

for integers t and Δ. They are valid and facet defining under mild conditions (see [5]). We call facets derived from inequality (10) *t-scaled MIR* facets, and facets from inequality (11) *t-scaled two-step MIR* facets. In particular, the inequality in (10) defines a facet of $P(n,r)$ if tr is not a multiple of n. For t equal to 1, inequality (11) defines a facet of $P(n,r)$ if $n > \Delta \lceil r/\Delta \rceil > r$ and $r > \Delta$.

In this paper we analyze the shooting experiment formally and extend the experimental results presented in [13] and [7]. In particular, we investigate if their observations on the importance of scaled MIR facets and 2slope facets extend to $P(n,r)$ for larger values of n. In addition, we measure the importance of two-step MIR facets which are not 2slope facets.

The structure of the paper is as follows: In Section 2, we describe the shooting experiment in detail. In Section 3, we discuss how many shots are needed to make reliable conclusions from the experimental observations. In Sections 4 and 5, we present our computational results.

2 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 (2)-(5). Therefore, for any $d \in R_+^{n-1}$, a basic optimal solution of the linear program

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

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 [8], 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 $\phi(i) = ki \,\mathrm{mod}\, n$; 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[10]. 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.

Furthermore, 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 $\phi(i) = ki \,\mathrm{mod}\, n$. 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 (10), the t-scaled of $P(n,r)$ is isomorphic to the following facet of $P(n,\phi(r))$:

$$\sum_{i=1}^{n-1} f^{tr/n}(ti/n)\, w_{\phi(i)} \geq 1. \tag{12}$$

Let s be the unique integer such that $sk \bmod n = t$. We will show that (12) 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 (12) 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 (11) 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 $P(10,3)$.

3 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. 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]):

$$P(p > \bar{p}_t - \Phi^{-1}(\alpha)\sqrt{\bar{p}_t(1-\bar{p}_t)/t}) \geq \alpha \tag{13}$$

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 (13) 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 1 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 5). As seen in the Figure 1, $\bar{p}_t$ becomes quite stable after 25,000 shots and almost does not change after 50,000 shots.

4 Computational results

We use two approaches to obtain our experimental results. For smaller n, we solve the linear program described in Section 2 to optimality for each direction d. Here we use the LP solver QSopt 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.

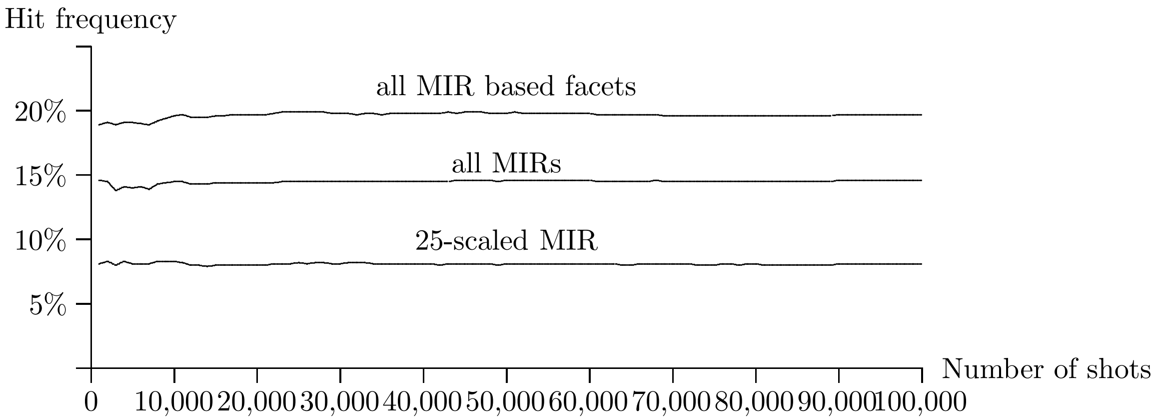

Figure 1: Shooting results for $P(50, 1)$

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, 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 [9] 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 2, 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 course a single choice for r.

Our observations are similar to those in [13], especially regarding the importance of scaled MIR facets. We however also examine the importance of two-step MIR

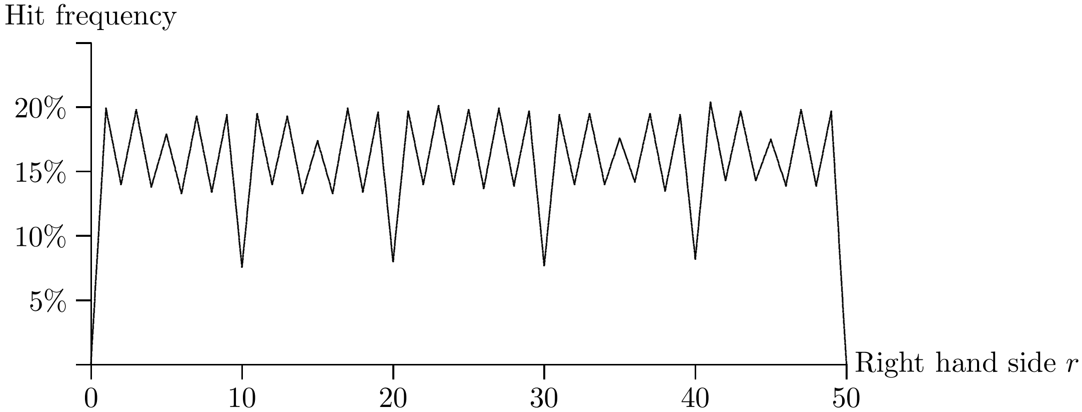

Figure 2: % of shots absorbed by all MIR based facets for $P(50, r)$

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 (which is much higher). 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–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, 7].

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.

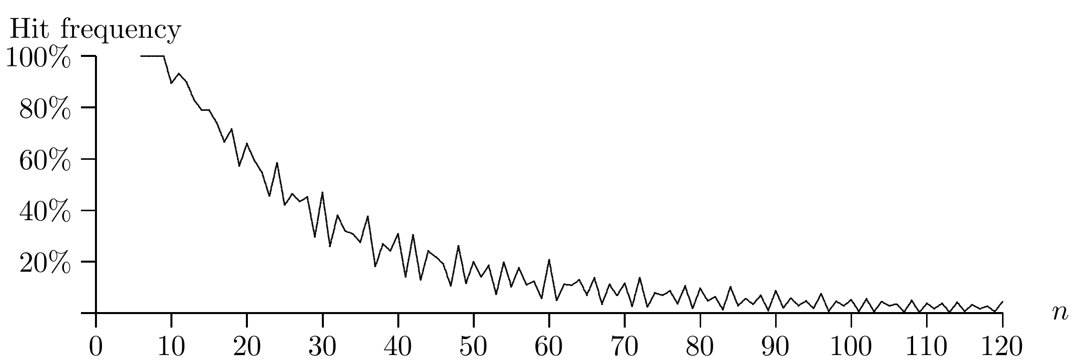

Figure 3: % of shots absorbed by all MIR based facets for $P(n,1)$

Finally in Figure 2, we consider $P(n,1)$ for $n = 6,\ldots,120$. This figure illustrates how the 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 [12]. Also note that even for

	Number of facets		% of shots hit				
The group	MIRs	two-step MIRs	top MIR	MIRs	2slopes	two-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 $P(n, r)$

larger n the frequency of hits absorbed by MIR based facets is non-negligible.

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

As we discussed earlier, our experiments suggest that the most important facets are t-MIR facets where t is a divisor of n, especially the largest divisor. We believe that *if t is the largest divisor of n, then the t-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 largest MIR based facet is the 10-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

$$\bar{\eta}^T d \;=\; \min_{\eta \in \Gamma(n,r)} \{\eta^T d\} \tag{14}$$

where $\Gamma(n, r)$ is the set of extreme points of (2)-(5). 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 $P(n, r)$ with $r \neq 0$. Then $P(Z(\eta) > Z(\beta)) = P(Z(\eta) < Z(\beta))$.*

Proof. We will show that $P(Z(\eta) - Z(\beta) > 0) = P(Z(\eta) - Z(\beta) < 0)$. 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 (14). We believe that a similar property holds for t-scaled MIRs, with t the largest divisor of n.

6 Concluding Remarks

In this paper we analyzed the shooting experiment and extended the experimental results presented in [13] and [7]. Our computational results confirm the importance of the MIR-based facets described in our earlier paper [5].

Our main observations are

1. Scaled MIR and two-step MIR facets are important facet classes and neither is conistently more important than the other,

2. The two-step MIR facets that are not 2slope facets are also important,

3. The most important facets of $P(n, r)$ are the scaled MIR facets where the scaling factor is a large divisor of n.

Combined with the fact that [6] MIR-based inequalities dominate the template inequalities based on the MIR-based facets of the master polyhedra, we conclude that template inequalities based on the important facets of small master polyhedra should not be used as cutting planes for mixed integer programs.

Acknowledgments: We would like to thank John Forrest for his help with using COIN-Clp for our computational 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, To appear in *INFORMS Journal on Computing*.

[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] S. Dash and O. Gunluk, Comparing valid inequalities for Cyclic Group Polyhedra, IBM Research Report RC22989, T. J. Watson Research Center, Yorktown Heights, New York, 2003.

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

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

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

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

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

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

[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).