RC22838 (W0307-029) July 2, 2003
Mathematics

IBM Research Report

Parsimonious Binary-Encoding In Integer Programming

Jon Lee, Don Coppersmith
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

PARSIMONIOUS BINARY-ENCODING IN INTEGER PROGRAMMING

DON COPPERSMITH AND JON LEE

Abstract. We describe an effective method for doing binary-encoded modeling, in the context of 0/1 linear programming, when the number of feasible configurations is not a power of two. Our motivation comes from modeling all-different restrictions.

Introduction

We assume some familiarity with the basics of polytopes (see [7]) and integer programming (see [6], for example).

Our motivation follows that of [2] (also see [3, 4]). In the context of integer programming, we are expressing "colors" $0, 1, \ldots, \kappa - 1$ in binary. The number of bits that we need is $n := \lceil \log_2 \kappa \rceil$. Lee [2] studied, in some detail, the *all-different polytope*: Namely, the convex hull of $m \times n$ 0/1 matrices with all-different rows — so $m \leq 2^n$. We can think of each such 0/1 matrix X as applying different colors, from a set of 2^n colors, to m objects. Our goal is to find an efficient way to handle the case where κ is not a power of 2. Lee [2] provided one simple technique, but it is not very effective from a polyhedral point of view. Lee's technique is simply to append the inequality $\sum_{i=0}^{n-1} 2^i x_i \leq \kappa - 1$, for each row $(x_{n-1}, x_{n-2}, \ldots, x_1, x_0)$ of X. But already for $m = 1$, $n = 2$, $\kappa = 2$, we have the fractional extreme point $(x_1, x_0) = (1/2, 0)$.

Let $k := 2^n - \kappa$ be the number of n-bit strings that will not describe colors. Clearly, $k < 2^{n-1}$. We are free to choose which of the k n-bit strings will not describe colors. Our goal is to choose them conveniently, from a polyhedral combinatorics point of view. In particular, we seek to cut off these points from the standard n-cube $H_n := [0, 1]^n$ using a standard set of so-called "cropping" inequalities (see [1]), so that the resulting polytope has only integer vertices (corresponding to the κ valid colors). But our goal is to accomplish this "parsimoniously"; for examples, we may seek to minimize: (i) the number of cropping inequalities used, or (ii) the volume of the resulting polytope.

Before continuing, it is worth remarking that although our motivation came from coloring, the problem that we address is more fundamental than that. Generally, we may wish to model κ "feasible configurations" as vertices of a lowest-dimensional H_n. The general issue is how to inject the feasible configurations into the vertices of H_n so that we can easily and efficiently describe the convex hull of the image by linear inequalities.

Date: June 2003.
Key words and phrases. polytope, integer programming, coloring, all different,

1. The Problem

Let n and k be positive integers with $k < 2^{n-1}$. We seek to find μ tri-partitions of $N := \{1, 2, \ldots, n\}$ as $(S_i, T_i, U_i)_{i=1}^{\mu}$, so that:

(*i*) $k = \sum_{i=1}^{\mu} 2^{|U_i|}$;
(*ii*) $|S_i \cap T_j| + |T_i \cap S_j| \geq 2$, for all distinct i, j;
(*iii*) $\sum_{i=1}^{\mu} w(S_i, T_i, U_i)$ is minimized,

where w is an arbitrary function from tri-partitions to $\Re$. Soon, we will restrict the class of w that we will analyze and give some examples. But for now, take $w(S_i, T_i, U_i) := 1$, so that in (*iii*) we are just minimizing μ.

First, we discuss the geometry of the problem. Associated with the partition (S_i, T_i, U_i) is the *cropping inequality*

$$\sum_{j \in S_i} (1 - x_j) + \sum_{j \in T_i} x_j \geq 1. \tag{1.1}$$

In $\Re^n$, the inequality (1.1) cuts, from H_n, all vertices of the form

$$(\underbrace{1, 1, \ldots, 1, 1}_{S_i}, \underbrace{0, 0, \ldots, 0, 0}_{T_i}, \underbrace{*, *, \ldots, *, *}_{U_i}). \tag{1.2}$$

These generating points (1.2) are the vertices of a $|U^i|$-dimensional face of H_n. Of course, the inequality (1.1) removes much more from H_n than this face. The n-cube H_n is cut at the $2^{|U_i|}(|S_i| + |T_i|)$ vertices of H_n that are adjacent to the set of points (1.2) and satisfy (1.1) as an equation. Moreover, the volume cut off is $1/|S_i \cup T_i|!$.

In Figures 1–2, we see two possible choices of sets of tri-partitions when $n = 3$ and $k = 2$.

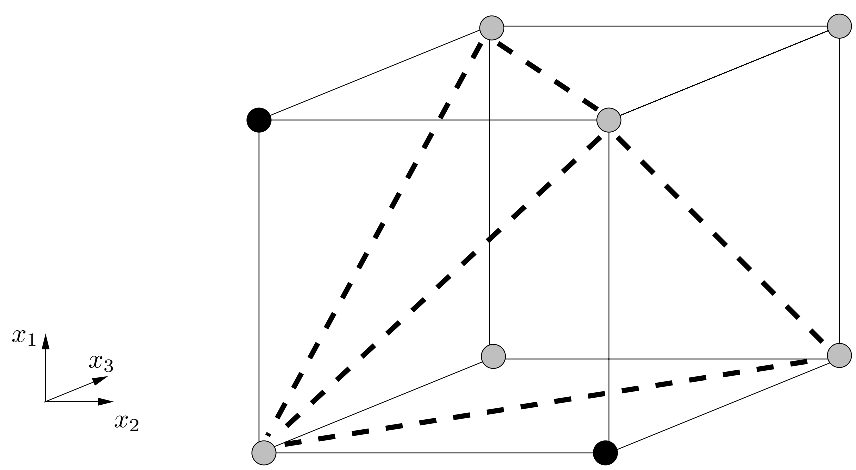

Figure 1. $\mu = 2$: $S_1=\{1\}$, $T_1=\{2,3\}$, $U_1=\emptyset$; $S_2=\{2\}$, $T_2=\{1,3\}$, $U_2=\emptyset$

Each inequality (1.1), individually, creates no fractional vertices when applied to H_n. Condition (*i*) specifies that we cut off the desired number k of points. Condition (*ii*) ensures that the parts of H_n cut off by each of the μ inequalities are disjoint (see [1]). Condition (*iii*) seeks to minimize some criterion.

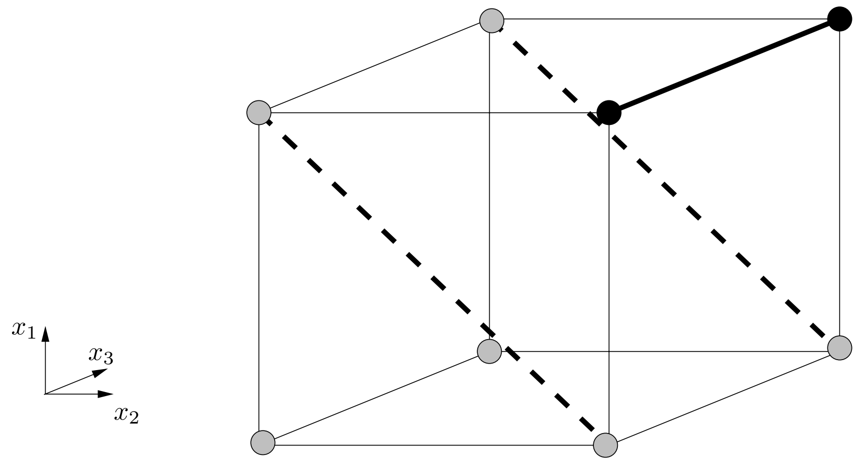

Figure 2. $\mu = 1$: $S_1 = \{1, 2\}$, $T_1 = \emptyset$, $U_1 = \{3\}$

We refer to any set of tri-partitions satisfying (i) as an *n-breakup* of k. If (ii) is also satisfied, then we have a *valid n-breakup*. If (iii) is also satisfied, then we have an *optimal n-breakup* (*with respect to w*). We may omit the "n-" and/or "valid" when it is clear from context.

With $k \leq 2^{n-1}$, a valid n-breakup always exists since we can just take any k of the 2^{n-1} tri-partitions having $|S_i|$ even, $T_i := N \setminus S_i$, and $U_i := \emptyset$. But this is the most inefficient possible choice of a set of tri-partitions.

A *binary breakup* is a breakup with all $|U_i|$ different, so that μ is the Hamming weight of the binary representation of k. Obviously, when a binary breakup is valid, it is optimal with respect to minimizing μ, so this Hamming weight is always a lower bound on μ.

We provide further motivation by considering the case of dimension $n = 4$ and $k = 7$ points to cut off. The interested reader can check that $\mu = 3$ (with $2^{|U_i|} = 4, 2, 1$) is not possible. So μ cannot always be achieved as the Hamming weight of the binary representation of k. But $\mu = 4$ with $2^{|U_i|} = 2, 2, 2, 1$ is possible (and optimal with respect to minimizing μ):

$$S_1 = \{2, 3\},\ T_1 = \{1\},\ U_1 = \{4\};$$
$$S_2 = \{1, 3\},\ T_2 = \{2\},\ U_2 = \{4\};$$
$$S_3 = \{1, 2\},\ T_3 = \{3\},\ U_3 = \{4\};$$
$$S_4 = \emptyset,\ T_4 = \{1, 2, 3, 4\},\ U_4 = \emptyset.$$

In the next section, we characterize the optimal breakups for certain functions w. In the remainder of this section, we fix some useful notation and introduce the functions w that we are able to analyze.

As was already noted, the set of points of the form (1.2) is the set of vertices of a $|U_i|$-dimensional face of H_n. That face is, itself, a $|U_i|$-cube. To keep notation less cluttered, we will denote one of these standard subcubes of H_n by (permutations of) the n-string

$$a_1 a_2 \ldots a_h \underbrace{* * \ldots * *}_{n-h},$$

where the a_i are in $\{0,1\}$. For example, the optimal 4-breakup (above) of $k=7$ as

$$011*$$
$$101*$$
$$110*$$
$$0000\ .$$

For convenience, when all of the "free" components are at the end, we may write

$$a_1 \dots a_h *^{n-h}\quad .$$

In this case, we refer to the substring $a_1 \dots a_h$ as the *address* of the subcube.

Notation: The *size* of a cube or subcube is the number of its vertices: $|H_n| = 2^n$. The *relative size* of a subcube A is $|A|/2^n$. A subcube is also called a *block*; a subcube filled with unused colors is a *piece*. A set K of unused colors has a *number* $k = |K|$ and a *mass* $\nu = \nu(K) = k/2^n$. If A is a subcube of H_n, we set $m(A) = m_K(A) = |K \cap A|/2^n$, and denote the *density* within A as $\rho(A) = \rho_K(A) = |K \cap A|/|A|$.

We are given a (strictly) subadditive *cost function* $\tilde{c}$ on sizes of subcubes. So the cost function satisfies $\tilde{c}(2^h) < 2\tilde{c}(2^{h-1})$. The functions w that we are interested in are of the form

$$w(S_i, T_i, U_i) := \tilde{c}(2^{|U_i|}),$$

where $\tilde{c}$ is a strictly subadditive cost function. It will be more convenient to use relative size in many of our arguments; fixing the ambient space H_n, we will set $c(2^{-h}) := \tilde{c}(2^{n-h})$ as the cost of a subcube of relative size 2^{-h}. Again, we have $c(2^{-h}) < 2c(2^{-h-1})$. So

$$w(S_i, T_i, U_i) = c(2^{-|S_i \cup T_i|}).$$

Let $\rho := k/2^n$ denote the density of colors that we will not use. When ρ is a density, we will speak interchangeably of an n-breakup of ρ or an n-breakup of $k = 2^n\rho$. The cost of a valid n-breakup of ρ is the sum of the costs of its pieces. We let $C(\rho)$ denote the minimum cost of a valid breakup of ρ. If $\rho = \sum \epsilon_h 2^{-h}$, $\epsilon_h \in \{0,1\}$ then $\mathrm{Bin}(\rho) := \sum \epsilon_h c(2^{-h})$, that is, the cost of a binary breakup of ρ, whether or not such a breakup is valid.

We give two example cost functions. First, in the case where we count subcubes, we have $c(2^{-h}) = 1$. Certainly this choice of c is strictly subadditive. Second, if we set $c(2^{-h}) = -1/h!$, then 1 plus the total cost of a valid break up is just the volume of H_n that satisfies the cropping inequalities associated with the breakup. It is easy to check that this choice of c is subadditive — and strictly so, except between $h=1$ and $h=2$ (which is of no concern to us). This latter choice of c is motivated by [5].

2. The Solution

Certainly, if $\rho = 2^{-h}$ for some nonnegative integer h, then the trivial binary breakup is possible (and optimal). Also, if $\rho < 1/4$ then a binary breakup will be possible (and optimal); this is described in Lemma 2.1 below.

Otherwise it will turn out that any optimal breakup can be obtained as follows. Select a largest block relative size, say relative size 2^{-h}, recalling that the normalization is such so that H_n has relative size 1. Break H_n into 2^h parallel blocks, and pay attention to the parity of the h-bit address of each block. We will leave empty the 2^{h-1} blocks of (say) odd parity. Among the 2^{h-1} blocks with even parity, we will fill all but one, two or three of them completely, and fill the remaining few only partially. The choices of block size 2^{-h}, the number of complete blocks, and the layout of the partial blocks depends on the exact value of ρ.

Our theorem will show that this gives the only optimal breakups. The most troublesome case to rule out, occurring when $11/32 < \rho < 12/32 = 3/8$, is the possibility of completely filling one block (address 000) of size 1/8, leaving three blocks empty (address 001, 010, 100), and partially filling four blocks (addresses 110, 101, 011, 111), each to density at most 1/2. Pieces could be shared between block 111 and any of the three neighboring blocks. In this case we need to examine the largest pieces within the partially filled blocks, in order to conclude that the main procedure is still optimal.

Lemma 2.1. *If $\rho = 1$ or $\rho = 1/2$ or $\rho \leq 1/4$, then k admits a binary breakup.*

Proof. If $\rho \in \{1, 1/2, 1/4\}$, the result is immediate. If $\rho < 1/4$, we proceed inductively. The base cases of $n = 0, 1, 2$ are trivial. So, when $\rho < 1/4$, we use the four cubes of dimension $n-3$, whose leading 3-tuples are the four strings of length 3 having even weight. One will accommodate a piece of relative size 1/8 (if the binary representation of ρ has a 1 in the 1/8 place); the second, 1/16; the third, 1/32; and the fourth will accommodate the rest of ρ. Since this remainder is smaller than 1/32, its density within the cube of dimension $n-3$ is less than 1/4, and, by induction, it admits a binary breakup. □

We say two vertices of H_n are *adjacent* if they share an edge, and two pairs of adjacent vertices are *parallel* if their respective shared edges are parallel.

The following technical lemma will be helpful in our analysis.

Lemma 2.2. *Assign to each vertex $v \in H_n$ a value $\tau = \tau_v \in [0, 1/2] \cup \{1\}$ subject to the following conditions:*

1. *At least k vertices (with $0 \leq k \leq n-1$) have $\tau_v \in (0, 1/2]$;*
2. *at least one vertex has $\tau_v = 1$;*
3. *if v, w are adjacent vertices then $\tau_v + \tau_w \leq 1$.*

Define the total value to be $R = R(H_n) = \sum_{v \in H_n} \tau_v$. Then we conclude:

$$R(H) \leq 2^{n-1} - \frac{k}{2}$$

Proof. If $n = 1$ then we must have $k = 0$ and $R = 1$, so the conclusion is easily seen to be true. So assume $n > 1$.

If exactly one vertex v has $\tau_v = 1$, then each of its n neighbors w has $\tau_w = 0$, and each other vertex w has $\tau_w \leq 1/2$, so that

$$R(H_n) \leq 1 + \frac{2^n - (n+1)}{2} = 2^{n-1} - \frac{n-1}{2} \leq 2^{n-1} - \frac{k}{2}$$

as desired.

So assume that at least two vertices each have $\tau = 1$. Select a coordinate in which they differ, and divide H_n into H^0_{n-1} and H^1_{n-1} along this coordinate, where each H^j_{n-1} is an H_{n-1}, and each H^j_{n-1} has at least one vertex with $\tau = 1$.

For each j, suppose that H^j_{n-1} has exactly h_j vertices with $\tau \in (0, 1/2]$, so that $h_0 + h_1 \geq k$. If some $h_j = 0$, then among the 2^{n-1} edges between H^0_{n-1} and H^1_{n-1}, at least k satisfy $\tau_s+\tau_t \leq 0+\frac{1}{2}$ and the other $2^{n-1}-k$ satisfy $\tau_s+\tau_t \leq 1$. Summing, we find

$$R(H_n) \leq \frac{k}{2} + (2^{n-1}-k)(1) = 2^{n-1} - \frac{k}{2}$$

as desired.

In the remaining case, $h_0, h_1 \geq 1$, and we can define $k_j = \min(h_j, n-2)$, and check that $k_0+k_1 \geq k$. Applying the lemma inductively to both halves, we conclude that

$$R(H_n) = R(H^0_{n-1}) + R(H^1_{n-1}) \leq 2^{(n-1)-1} - \frac{k_0}{2} + 2^{(n-1)-1} - \frac{k_1}{2} \leq 2^{n-1} - \frac{k}{2}.$$

□

We outline a procedure for finding a breakup for an arbitrary density ρ. It leaves a few choices open, which will depend on the cost function $c(\cdot)$ and the structure of ρ. Then, in Theorem 2.4, we will show that, for each subadditive cost function $c(\cdot)$ and each density ρ, PROCEDURE 1 will produce an optimal breakup, for some setting of the choices.

PROCEDURE 1

Given a density $\rho \in [0, 1/2] \cup \{1\}$ and a subadditive cost function $c(\cdot)$, we find a breakup ρ. Its form depends on ρ as follows:

Case 1: $\rho = 1$, $\rho = 1/2$ or $\rho \leq 1/4$. (Binary representations are 1.0, 0.1, 0.01, or 0.00xxx for some unspecified continuation "xxx").

Lemma 2.1 provides a binary breakup; and when a binary breakup exists, it is optimal.

Case 2: $\rho = 1/2 - 1/2^h$. (Binary representation is 0.0111 (if $h = 4$, as it will be in the remaining examples)).

Lemma 2.2 shows we cannot possibly use pieces larger than $1/2^{h-1}$. Use $2^{h-2}-1$ pieces of size $1/2^{h-1}$ (corresponding to even weight words of length $h-1$); use the remaining block to accommodate the remaining piece of size $1/2^h$.

Case 3: $1/2 - 1/2^h < \rho < 1/2 - 1/2^h + 1/2^{h+2}$. (Binary representation is 0.011100xxx).

Lemma 2.2 again shows we cannot use pieces larger than $1/2^h$. Use $2^{h-1}-1$ pieces of size $1/2^h$, leaving one block of size $1/2^h$ and an unused mass of $\rho-(1/2-1/2^h) = \rho' \times 1/2^h$ where $0 < \rho' < 1/4$. This unused mass can be represented within that block with a binary breakup, since its density will be less than 1/4.

Case 4: $1/2 - 1/2^h + 1/2^{h+2} \leq \rho < 1/2 - 1/2^h + 1/2^{h+2} + 1/2^{h+3}$. (Binary representation is 0.0111010xxx).

By Lemma 2.2, we cannot use pieces smaller than $1/2^h$. There are two possibilities; evaluate both and use the cheaper.

Case 4A: We can use $2^{h-1}-1$ pieces of size $1/2^h$, leaving one block unused and residual mass $\rho - (1/2 - 1/2^h) = \rho' \times 1/2^h$ where ρ' is between 1/4 and 3/8. Use PROCEDURE 1 inductively on ρ' to solve that problem.

Case 4B: Or we can use 2^h-2 pieces of size $1/2^{h+1}$, leaving two blocks. Use one to handle the piece of size $1/2^{h+2}$. The remaining mass is $\rho-(1/2-2/2^{h+1}+1/2^{h+2}) = \rho' \times 1/2^{h+1}$ where $\rho' < 1/4$, so it admits a binary breakup.

Remark: If our cost measure is "number of pieces" ($c(1/2^h) = 1$) then Case 4A is always preferable over 4B. But under a more general measure either could be preferable: If $\rho = .0101011$, $c(1/4) + 2c(1/32) + c(1/64) + c(1/128)$ and $2c(1/8) + c(1/16) + c(1/64) + c(1/128)$ are *a priori* incomparable.

Case 5: $1/2 - 1/2^h + 1/2^{h+2} + 1/2^{h+3} \leq \rho < 1/2 - 1/2^{h+1}$**.** (Binary representation is 0.0111011xxx).

We cannot use pieces smaller than $1/2^h$. There are three possibilities; evaluate all three and use the cheapest.

Case 5A: We can use $2^{h-1} - 1$ pieces of size $1/2^h$, leaving the last block with a density ρ' between 3/8 and 1/2, to handled by induction. Here $\rho - (1/2 - 1/2^h) = \rho' \times 1/2^h$.

Case 5B: We can use $2^h - 2$ pieces of size $1/2^{h+1}$, leaving one block to fill halfway with the piece of size $1/2^{h+2}$, and leaving a second block with density ρ'' between 1/4 and 1/2, to be handled by induction. Here $\rho - (1/2 - 2/2^{h+1} + 1/2^{h+2}) = \rho'' \times 1/2^{h+1}$. This piece is different from the one in the first case ($\rho'' \neq \rho'$) and could conceivably be handled more efficiently.

Case 5C: We can to use $2^{h+1} - 3$ pieces of size $1/2^{h+2}$. This leaves three blocks: one accommodates the piece of size $1/2^{h+3}$ (relative density 1/2); a second, the piece of size $1/2^{h+4}$ (density 1/4); and the third accommodates the remainder, whose density is less than 1/4, so a binary breakup is possible. $\rho - (1/2 - 3/2^{h+2} + 1/2^{h+3} + 1/2^{h+4}) = \rho''' \times 1/2^{h+2}$.

Note that each of these cases (5A, 5B, 5C) is optimal in some region. Suppose that our objective is to minimize the number of pieces. If $\rho = 1/2 - 1/2^5 - 1/2^9 = .011101111$ then Case 5A is optimal with 15 pieces; Case 5B uses 19 and Case 5C uses 32. If $\rho = 1/2 - 1/2^5 - 1/2^{11} = .01110111111$ then Case 5B is optimal with 31 pieces; Case 5A uses 39 and Case 5C uses 34. If $\rho = 1/2 - 1/2^5 - 1/2^{12} = .011101111111$ then Case 5C is optimal with 35 pieces; Case 5A uses 71 and Case 5B uses 47.

This ends the description of Procedure 1; all ranges of ρ have been treated. It remains to prove that in each case the optimal treatment is one of the possibilities allowed by Procedure 1.

Lemma 2.3. $\mathrm{Bin}(\rho) \leq C(\rho)$.

Proof. This follow from subadditivity of $c(\cdot)$. □

Theorem 2.4.

(1) Procedure 1 *achieves optimality; it represents* ρ *using cost* $C(\rho)$*.*
(2) *If* $0 \leq \alpha, \beta, \gamma \leq 1/2$ *and* $\alpha + \beta = \gamma + 1/2$*, then* $C(\alpha) + C(\beta) \geq C(\gamma) + C(1/2)$*.*

Proof. We need to prove the two clauses simultaneously by induction; the second is only useful for maintaining the induction.

In Cases 1, 2 and 3, optimality follows from the fact that we use the largest possible pieces (from Lemma 2.2), as many of them as possible, and use binary breakup for the remainder.

In Case 4, Lemma 2.2 demands that no piece be larger than $1/2^h$.

Suppose first that there is a piece of size $1/2^h$, and $h \geq 3$ (this is Case 4A). Then by Lemma 2.2, using $d = h \geq 3$ and $k = 2$, because $R(H) = 2^h \times \rho > 2^{d-1} - \frac{k}{2}$, we conclude that there are not two "partial" blocks of size $1/2^h$ (that is, each with mass strictly between 0 and $1/2^h$); there is only one such block, and $2^{h-1} - 1$ "full" blocks. Optimality of our treatment of this partial block follows by induction.

If $h = 2$, the fact that one piece is of size $1/4$ forces the rest to be confined to the opposite block of size $1/4$, so that Case 4A is the only possible treatment.

Otherwise the largest piece is at most $1/2^{h+1}$, in which case the treatment in Case 4B is optimal (by the same argument as Cases 1,2,3).

In Case 5, Lemma 2.2 demands that no piece be larger than $1/2^h$.

Case 5A is when we do use a piece of size $1/2^h$. The treatment is similar to that of Case 4A, and again requires breaking into subcases, depending whether $h = 2$ or $h \geq 3$.

Optimality of Case 5C (in the event that no piece is larger than $1/2^{h+2}$) is similar to that of Case 4B.

In Case 5B, if $h = 2$ we repeat the argument from 4A. If $h \geq 4$, use Lemma 2.2 with $d = h \geq 4$ and $k = 3$ to show that we cannot have ≥ 3 partial blocks. Nor can we have one partial block (since its relative density would be strictly between $1/2$ and 1). We have two partial blocks, with relative densities α, β with $0 < \alpha, \beta \leq 1/2$. By induction, it is cheaper to use two blocks $\gamma, 1/2$ with $\alpha + \beta = \gamma + 1/2$.

If $h = 3$, then Lemma 2.2 is insufficient, and indeed we can have one full block and four partial blocks. The 3-bit address of the full block is 000; the addresses of the four partial blocks are 011, 101, 110, 111. Notice that partial block 111 is adjacent to the other three, and pieces can be shared between two neighboring partial blocks. In this case we need to rely on the following technical lemma:

Lemma 2.5. *If* $11/32 \leq \rho \leq 3/8$, *the outlined procedure gives a better cost than the possibility of completely filling one subcube (address 000) of size 1/8, leaving three subcubes empty (address 001, 010, 100), and partially filling four subcubes (addresses 110, 101, 011, 111), each to density at most 1/2, and possibly sharing pieces between subcube 111 and its neighbors.*

Proof. Name the partially filled subcubes $A = 110$, $B = 101$, $C = 011$, and $D = 111$. We will break into cases, according to which of these subcubes are filled with density exactly $1/2$ (so contributing $1/16$ to ρ) and which have a "deficit" (density less than $1/2$). In each case we will identify a block size h, related to the largest piece within certain subcubes or the largest piece shared between two subcubes. Let $q = (1/16)/h$ be the number of such pieces required to fill a subcube to density $1/2$.

In some cases we will replace the given configuration by one where four subcubes of size $1/8$ (000, 110, 101, 011) are filled with masses $1/8$, $1/8$, $1/16$, $\rho - 5/16$ respectively, the latter being covered with at most $q - 3$ pieces of size h and an optimal covering of the remaining mass. We will show that the pieces in the original covering can be combined to form the pieces in the new covering; in particular, that the pieces of size at least $2h$ (within $A \cup B \cup C \cup D$, that is, outside of block 000) accounted for mass at most $1/8 + 1/16 = 3/16$. Thus the cost of the new arrangement will be less than that of the original arrangement, and the new arrangement is one of them produced by our procedure.

In other cases we will use only two subcubes of size $1/4$ (00 and 11), again giving a smaller cost arrangement that could be produced by our procedure.

Notation: we will let $b_{2h}(A)$ denote the total mass of pieces wholly within subcube A whose individual sizes are at least $2h$; $b_{2h}(AD)$ is the total mass of pieces which straddle A, D (so half of the piece is in each subcube) and with sizes at least $2h$; $b_{2h}(A, BD, CD) = b_{2h}(A) + b_{2h}(BD) + b_{2h}(CD)$; and when the threshold $2h$ is understood we may write $b(A), b(BD)$, and so on (b means "big").

By symmetry among A, B, C, we need only consider eight cases, labelled L1-L8 to distinguish them from the cases in PROCEDURE 1.

Case L1: No subcubes have deficits.

Then $\rho = 1/8 + 4(1/16) = 3/8$, so the binary breakup $\rho = 1/4 + 1/8$ is possible and thus optimal.

Case L2: Only subcube A has a deficit.

If no pieces straddle AD, then we can replace the given configuration by one in which $B = 101$ is a single piece of mass $1/8$, $C = 011$ contains a single piece of mass $1/16$, $D = 111$ is empty, and $A = 110$ remains as is.

Otherwise at least one piece straddles AD; let its size be $2h$. Consider the covering of D by pieces, some of which may be wholly contained within D, while others are the restriction to D of pieces straddling AD or BD or CD. Since D has no deficit, all these pieces (intersected with D) have the same size h. This implies that all pieces straddling AD have identical size $2h$. Since BD has no deficit, and one piece intersected with BD has size h, all the pieces intersected with BD have size h, so that $b(B) = b(BD) = b(C) = b(CD) = b(D) = 0$. The only pieces of size at least $2h$ are contained with A or AD, so their total mass is less than 1/16+1/16=1/8. We will sever the pieces straddling AD and, as above, replace the given configuration by one in which $B' = 101$ is a single piece of mass $1/8$, $C' = 011$ contains a single piece of mass $1/16$, $D' = 111$ is empty, and $A' = 110$ remains as is (the pieces straddling AD are broken in half). Because $b(A, AD) < 1/8 < 3/16 = 1/8 + 1/16 = m(B' \cup C')$, the new configuration has a smaller cost than the old one.

Case L3: Subcubes A, B have deficits.

If no pieces straddle AD or BD, replace C and D by a full piece of mass $1/8$ and an empty piece, respectively. Then use the second inductive clause of Theorem 2.4 to say $c(1/8) + c(1/8) + c(A) + c(B) \leq c(1/8) + c(1/8) + c(1/16) + c(\rho - 5/16)$, so that case 5A is better than the original configuration.

If there are pieces straddling AD or BD, let $2h$ be the largest size of such a piece. D is covered by (intersected) pieces of size h. All pieces straddling AD or BD have identical size $2h$. Because at least one piece intersected with CD has size h, CD is covered by (intersected) pieces of size h. So the pieces of size at least $2h$ are in A, AD, B, BD. Again we replace C and D by a full piece C' of mass $1/8$ and an empty block D', respectively. This breaks pieces of size $2h$ of total mass $b(AD, BD) \leq 2m(D) = 2(1/16) = 1/8 = m(C' \cup D')$, so the new configuration is no more expensive than the old one. As above, $c(1/8) + c(1/8) + c(A) + c(B) \leq c(1/8) + c(1/8) + c(1/16) + c(\rho - 5/16)$, so that case 5A is better than the original configuration.

Case L4: Subcubes A, B, C have deficits.

D has no deficit so it is covered by q pieces of size h. No piece of A, B, C can be larger than h, but pieces of size exactly $2h$ can straddle AD, BD or CD. Their total mass is at most twice the mass of D, or $2/16 = 1/8$. First we isolate D from A, B, C, incurring a temporary cost of at most $q \times (2c(h) - c(2h))$. Each of A, B, C has $2q$ blocks of size h, of which q must be empty. The remaining q (in each subcube) can be empty, full, or partially full, and are not adjacent. Of these $3q$ blocks of size h, if fewer than q are completely full (so the remaining $2q$ are at most half full), then the total density will be less than $1/8+1/16+q(h)+2q(h/2) = 1/8 + 1/16 + 1/16 + 1/16 = 5/16$, so that a binary breakup will be possible and thus optimal (Case 3). So we assume at least q are completely full. Trade these blocks around so that C ends up with q full blocks. Replace C and D with a single piece C' (at location C) of mass $1/8$, with D' becoming empty. We have regained a cost advantage of $2q \times c(h) - c(1/8)$, outweighing our initial loss. Finish as before: $c(1/8) + c(1/8) + c(A) + c(B) \leq c(1/8) + c(1/8) + c(1/16) + c(\rho - 5/16)$, so that Case 5A is better than the original configuration.

Case L5: Only subcube D has a deficit.

Let h be the largest piece size intersected with D. Any pieces straddling AD must have size exactly $2h$; if larger than $2h$ then the part intersected with D would exceed the maximum, and if smaller than $2h$ then A would be covered by (intersected) pieces of half that size, strictly smaller than h, and would have a deficit due to empty space opposite the large piece in D.

Suppose A contains pieces of size at least $2h$. Then it completely covered with such blocks (all the same size; half of them empty and half of them full), and no pieces can straddle AD. Further, in this case, since D has large empty blocks opposite the full pieces of A, and piece size bounded by h, the total mass of D cannot exceed $1/32$. So if $b(A) > 0$ then $b(AD) = 0$ and $b(BD)+b(CD) \leq 1/16$. Denote the total mass of large pieces by $M = b(A) + b(B) + b(C) + b(AD) + b(BD) + b(CD)$. If $b(A) = b(B) = b(C) = 0$ then $M \leq 3 \times 0 + 3 \times 1/16 = 3/16$. If $b(A) > 0$ and $b(B) = b(C) = 0$ then M is bounded by $M = b(A) + [b(BD) + b(CD)] \leq 1/16 + 1/16 = 1/8$. If $b(A) > 0$, $b(B) > 0$ and $b(C) = 0$ then M is bounded by $M = b(A) + b(B) + b(CD) \leq 1/16 + 1/16 + 1/16 = 3/16$. If $b(A), b(B), b(C)$ are all nonzero, then $b(AD) = b(BD) = b(CD) = 0$ and $M = 3/16$. In either case $M \leq 3/16$, so that we can only profit by replacing the current configuration by a piece of size $1/8$ at A, a piece of size $1/16$ within B, a copy of the current D at the new C', and nothing at D'; the two pieces of sizes $1/8 + 1/16 = 3/16 \geq M$ accommodate all the large pieces from the old configuration.

Case L6: Subcubes A, D have deficits.

Let h be the largest piece intersected with D. As above, any piece straddling BD or CD has size exactly $2h$. Consider $M' = b(B) + b(C) + b(BD) + b(CD)$ (ignoring large pieces in A or AD). If $b(B) = b(C) = 0$ then $M' = b(BD)+b(CD) \leq 2m(D) < 1/8$. If $b(B) = 0$ and $b(C) > 0$ then $M' = b(BD) + b(C) \leq 1/16 + 1/16 = 1/8$ by arguments similar to Case 5. If both $b(B), b(C) > 0$, then $M' = b(B) + b(C) \leq 1/16 + 1/16 = 1/8$. In any case $M' \leq 1/8$. We will combine B, C and the piece of size $1/8$ at 000 into a single piece of size $1/4$, and let AD occupy the opposite subcube of size $1/4$. The large pieces in the old configuration that are lost in transition to the new one are $M' + 1/8 \leq 1/4$, so that we have only gained.

Case L7: Subcubes A, B, D have deficits.

Let h_A, h_B, h_D denote the largest (possibly intersected) piece in A, B or D respectively, and set $h = \max\{h_D, \min\{h_A, h_B\}\}$.

We claim that the total deficit

$$3/8 - \rho = (1/16 - m(A)) + (1/16 - m(B)) + (1/16 - m(D))$$

is at least $3h/2$, and in fact each of A, B, D contributes at least $h/2$ to the deficit. If $h = h_D$ then D has deficit at least $h/2$; and A either has a piece of size at least h (and so a deficit of at least $h/2$) or its largest piece is at most $h/2$ but it has an empty block of size h opposite the full one in D (again necessitating a deficit at least $h/2$). If $h = h_A \leq h_B$ then A, B each contributes a deficit of at least $h/2$, and repeating the above argument (reversing the roles of A, D) shows that D also has such a deficit.

Assume $h_A \leq h_B$, implying $h_A \leq h$. Pieces of size at least $2h$ can occur in either C or straddling CD (but not both, since C has no deficit); straddling AD but not A; both in B and straddling BD; and not in D. We wish to show that these large pieces have total mass M at most 3/16, setting $M = b(B)+b(C)+b(AD)+b(BD)+b(CD)$ and recalling $b(A) = b(D) = 0$. If $b(CD) > 0$ then $b(C) = 0$, so that $M \leq b(B) + b(AD) + b(BD) + b(CD) \leq b(B) + 2m(D) < 1/16 + 1/8 = 3/16$. If $b(C) > 0$ then C is covered with large pieces, $b(CD) = 0$. D cannot have large pieces (since $h_D \leq h$), so its mass is at most 1/32, and $b(AD)+b(BD) \leq 2m(D) \leq 1/16$, whence $M = b(B) + b(C) + b(AD) + b(BD) \leq 1/16 + 1/16 + 1/16 = 3/16$. If $b(CD) = b(C) = 0$ then $M = b(B)+b(AD)+b(BD) \leq m(B)+2m(D) = 1/16+2/16 = 3/16$. In any case the large pieces have total mass at most 3/16, and we can profitably combine them into a piece of size 1/8 at 110 and a piece of size 1/16 at 101. The subcube at 111 will be empty, and the subcube at 011 will have total mass $\rho - 5/16 \leq 1/16 - 3h/2$, so that we can use at most $(1/16)/h - 3$ pieces of size h and optimally cover the rest with binary breakup. This is equivalent to Case 5B (with pieces of size 1/8) followed by 5C (pieces of size h).

Case L8: Subcubes A, B, C, D have deficits.

Let $h_A \leq h_B \leq h_C$ and define $h = \max\{h_D, h_B\}$. The large pieces ($\geq 2h$) are bounded by $M = b(C)+b(AD)+b(BD)+b(CD) \leq m(B)+2m(C) < 1/16+1/8 = 3/16$. We proceed as in the previous case.

□

(Returning to proof of Theorem 2.4). We still need to prove $C(\alpha) + C(\beta) \geq C(\gamma) + C(1/2)$. If α allows a binary breakup, then either $\alpha = 1/2$ (so that $\beta = \gamma$ and the result is trivial), or $\alpha \leq 1/4$, in which case $\gamma \leq 1/4$ so that $C(\gamma) = \mathrm{Bin}(\gamma)$, and

$$\begin{aligned} C(\gamma) + C(1/2) &= \mathrm{Bin}(\gamma) + \mathrm{Bin}(1/2) = \mathrm{Bin}(\gamma + 1/2) = \mathrm{Bin}(\alpha + \beta) \\ &\leq \mathrm{Bin}(\alpha) + \mathrm{Bin}(\beta) \leq C(\alpha) + C(\beta). \end{aligned}$$

So assume neither α nor β falls into Case 1. Let an optimal breakup of α use largest piece $1/2^r$, and an optimal breakup of β use largest piece $1/2^s$, with $r \geq s$.

Assume first that $r > s$. We have

$$\alpha \leq \frac{1}{2} - \frac{1}{2 \times 2^r}$$

$$\beta \leq \frac{1}{2} - \frac{1}{2 \times 2^s} \leq \frac{1}{2} - \frac{1}{2^r}$$

$$\gamma \leq \frac{1}{2} - \frac{3}{2} \times \frac{1}{2^r}$$

If

$$\gamma = \frac{k}{2^r} + \frac{\epsilon_1}{2^{r+1}} + \frac{\epsilon_2}{2^{r+2}} + \frac{1}{2^r} \times \gamma',\ \epsilon_j \in \{0,1\},\ 0 \leq \gamma' < \frac{1}{4},$$

then $k + \epsilon_1 + \epsilon_2 \leq 2^{r-1} - 1$, so that we can represent γ within 2^{r-1} blocks of size $1/2^r$: k pieces of size $1/2^r$, ϵ_j pieces of size $1/2^{r+j}$ $(j = 1, 2)$ within blocks of size $1/2^r$, and one remaining block to represent $\gamma'/2^r$ with a binary breakup. So:

$$C(\gamma) \leq kc(1/2^r)+\epsilon_1 c(1/2^{r+1})+\epsilon_2 c(1/2^{r+2})+\mathrm{Bin}(\gamma'/2^r) = kc(1/2^r)+\mathrm{Bin}(\gamma-k/2^r).$$

Set $\alpha = (k_1 + \alpha')/2^r$ with $k_1 = 2^{r-1} - 1$ and $0 \le \alpha' \le 1/2$, and $\beta = k_2/2^s + (\ell + \beta')/2^n$ with $0 \le \ell < 2^{r-s} - 1$ and $0 \le \beta' < 1$. We find:

$$C(\alpha) \ge k_1 c(1/2^r) + \mathrm{Bin}(\alpha'/2^r)$$

$$C(\beta) \ge k_2 c(1/2^s) + \mathrm{Bin}(\ell/2^r) + \mathrm{Bin}(\beta'/2^r)$$

$$\mathrm{Bin}(\alpha'/2^r) + \mathrm{Bin}(\beta'/2^r) \ge \mathrm{Bin}((\alpha' + \beta')/2^r)$$

Collect some larger pieces together to equal 1/2:

$$C(1/2) \ge k_2 c(1/2^s) + \mathrm{Bin}(\ell/2^r) + \left[\frac{\frac{1}{2} - \frac{k_2}{2^s} - \frac{\ell}{2^r}}{1/2^r} \right] c(1/2^r)$$

Combining, we find

$$C(1/2) + C(\gamma) \le C(\alpha) + C(\beta)$$

in this case, as desired.

This leaves the case $r = s$. If $\gamma \le \frac{1}{2} - \frac{3}{2} \times \frac{1}{2^r}$, then we can just mimic the previous proof. So assume

$$\frac{1}{2} - \frac{3}{2} \times \frac{1}{2^r} < \gamma \le \frac{1}{2} - \frac{1}{2^r}$$

Further, since $\alpha, \beta \le \frac{1}{2} - \frac{1}{2 \times 2^r}$, we know $\alpha, \beta \ge \frac{1}{2} - \frac{1}{2^r}$. We can write

$$\alpha = (2^{r-1} - 1)\frac{1}{2^r} + \frac{\alpha'}{2^r},\ 0 \le \alpha' \le \frac{1}{2}$$

$$\beta = (2^{r-1} - 1)\frac{1}{2^r} + \frac{\beta'}{2^r},\ 0 \le \beta' \le \frac{1}{2}$$

$$C(\alpha) \ge (2^{r-1} - 1)c(\frac{1}{2^r}) + C'(\alpha')$$

$$C(\beta) \ge (2^{r-1} - 1)c(\frac{1}{2^r}) + C'(\beta')$$

where $C'(\alpha')$ is the cost of representing $\alpha'/2^r$ within a block of size $1/2^r$. Setting $\gamma' + 1/2 = \alpha' + \beta'$, we have (by induction)

$$C'(\gamma') + C'(1/2) \le C'(\alpha') + C'(\beta')$$

This is enough to show:

$$\begin{aligned} C(\alpha) + C(\beta) &\ge (2^{r-1} - 1)c(\tfrac{1}{2^r}) + C'(\alpha') + (2^{r-1} - 1)c(\tfrac{1}{2^r}) + C'(\beta') \\ &\ge (2^r - 2)c(\tfrac{1}{2^r}) + c(\tfrac{1}{2^{r+1}}) + C'(\gamma') \\ &\ge c(\tfrac{1}{2}) + (2^{r-1} - 2)c(\tfrac{1}{2^r}) + c(\tfrac{1}{2^{r+1}}) + C'(\gamma') \\ &\ge c(\tfrac{1}{2}) + C(\gamma) \end{aligned}$$

The last inequality follows because the indicated breakup is valid but not necessarily optimal.

□

References

1. Jon Lee, *Cropped cubes*, IBM Research Report RC21830 (2000), to appear in: Journal of Combinatorial Optimization.
2. ______, *All-different polytopes*, Journal of Combinatorial Optimization **6** (2002), 335–352.
3. Jon Lee, Janny Leung, and Sven de Vries, *Separating type-I odd-cycle inequalities for a binary encoded edge-coloring formulation*, IBM Research Report RC22303 (2002), to appear in: Journal of Combinatorial Optimization.
4. Jon Lee and François Margot, *Manuscript in preparation*, (2003).
5. Jon Lee and Walter D. Morris, Jr., *Geometric comparison of combinatorial polytopes*, Discrete Applied Mathematics **55** (1994), no. 2, 163–182.
6. George L. Nemhauser and Laurence A. Wolsey, *Integer and combinatorial optimization*, John Wiley & Sons Inc., New York, 1999, Reprint of the 1988 original, A Wiley-Interscience Publication.
7. Günter M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics, vol. 152, Springer-Verlag, New York, 1995, Revised edition, 1998.

IBM T.J. Watson Research Center, PO Box 218, Yorktown Heights, NY 10598 USA
E-mail address: {dcopper,jonlee}@us.ibm.com