ELSEVIER

Operations Research Letters 24 (1999) 105–114

operations research letters

www.elsevier.com/locate/orms

On dependent randomized rounding algorithms

Dimitris Bertsimas[a,*], Chungpiaw Teo[b], Rakesh Vohra[c]

[a] *Sloan School of Management and Operations Research Center, Massachusetts Institute of Technology, Room E53-363, 50 Memorial Drive, Cambridge, MA 02139-1347, USA*

[b] *Department of Decision Sciences, Faculty of Business Administration, National University of Singapore, Singapore*

[c] *Department of Managerial Economics and Management Sciences, Northwestern University, Evanston, Illinois, USA*

Received 1 January 1996; received in revised form 1 September 1998

Abstract

In recent years, approximation algorithms based on randomized rounding of fractional optimal solutions have been applied to several classes of discrete optimization problems. In this paper, we describe a class of rounding methods that exploits the structure and geometry of the underlying problem to round fractional solution to 0–1 solution. This is achieved by introducing dependencies in the rounding process. We show that this technique can be used to establish the integrality of several classical polyhedra (min cut, uncapacitated lot-sizing, Boolean optimization, k-median on cycle) and produces an improved approximation bound for the min-k-sat problem.

Keywords: Linear programming; Randomized rounding; Approximation algorithm

1. Introduction

The idea of using randomized rounding in the study of approximation algorithms was introduced by Raghavan and Thompson [17]. The generic randomized rounding technique can be described as follows:

- Formulate and solve a continuous relaxation (in polynomial time) for a 0–1 integer programming problem to obtain an optimal (possibly fractional) solution $\bar{x}$.
- Devise a randomization scheme to decide whether to round each variable x_i to 1 or 0.

The heart of the rounding procedure, given a relaxation, is in the design of the randomization scheme. In a recent survey on combinatorial optimization Grötschel and Lovász [9] write:

> ... we can obtain a heuristic primal solution by fixing those variables that are integral in the optimum solution of the linear relaxation, and rounding the remaining variables "appropriately". It seems that this natural and widely used scheme for a heuristic is not sufficiently analyzed ...

* Corresponding author.

E-mail address: dbertsim@mit.edu (D. Bertsimas).

0167-6377/99/$ - see front matter
PII: S0167-6377(99)00010-3

Raghavan and Thompson [17] derive several approximation bounds for multi-commodity routing problems by rounding *independently* each variable x_i to 1 with probability $\bar{x}_i$. Goemans and Williamson [7] introduce the idea to round each variable x_i independently but with probability $f(\bar{x}_i)$, for some particular nonlinear function $f(x)$. The algorithm (for the maximum-satisfiability problem) they obtain matches the best-known guarantee for the problem (originally obtained by Yannakakis [20].) Bertsimas and Vohra [2] use a nonlinear rounding function to obtain a randomized rounding heuristic for the set covering problem. Their method matches the best-known guarantee (originally obtained by Chvátal [4]). Brönnimann and Goodrich [3] show further that the set covering bound can be improved if the Vapnik–Cervonenkis (VC) dimension of the constraint matrix can be suitably bounded. Randomized rounding can also be seen as a generalization of deterministic rounding that exploits structural results of the fractional optimum solution to devise deterministic rounding heuristics. This technique has been used in the analysis of the bin-packing problem (see for instance, [5]), machine scheduling (see [19]) and set covering (see [11]).

In all the above applications of randomized rounding, each variable x_i was rounded independently. Few structural results of the fractional optimum solution have been used in the design of the rounding heuristics. Goemans and Williamson [8], in their study of the maximum-cut problem, show that the geometry of the fractional solution can be suitably utilized to obtain a rounding heuristic with very strong performance bounds. In their rounding process, the variables x_i are rounded in a dependent manner.

Our objective in the present paper is to describe a class of randomized rounding techniques that seems to work well on several classes of problems that are variants of the min-cut type. For another application of this technique see Bertsimas et al. [1]. The key advantage of this approach is that by using dependencies in the rounding process, the analysis of the performance of the rounding heuristic becomes extremely simple.

In the next section, we use dependent randomized rounding to establish integrality results for several basic combinatorial optimization problems. These include the min $s-t$ cut, boolean optimization, uncapacitated lotsizing and k-median problem on a cycle. In Section 3, we describe several approximation results using the rounding technique. For the feasible-cut problem studied by Yu and Cheriyan [21], our technique obtains a worst case bound of 2, matching that obtained in [21] (and independently by Ravi [18]). For the minimum satisfiability problem studied by Kohli et al. [13], our technique gives a $2(1-1/2^k)$ bound for the min k-SAT problem. For the min-2-sat problem, this result improves the bound from 2 to $\frac{3}{2}$. Marathe and Ravi [15] have obtained a bound of 2 independently using different methods. They have also shown that the minimum satisfiability problem is closely related to the node covering problem.

In this paper, we restrict the discussion only to new randomized rounding ideas and its applications. We will not discuss, for instance, running time analysis or de-randomization techniques. Furthermore, for ease of exposition, we let Z_{IP} and Z_{LP} denote the optimal integral and optimal fractional solution value, respectively. Z_H denotes the value returned by a heuristic H. All cost functions are assumed to be nonnegative. Graphs are assumed to be undirected unless stated otherwise.

2. Dependent rounding and integrality proofs

In this section, we study the connection of dependent randomized rounding and some basic combinatorial optimization problems. In particular, with the right randomization scheme, we show that the rounding argument leads to direct integrality proofs of several well-known polyhedra.

2.1. $s-t$ cut

In this section we give a direct probabilistic proof that the polyhedron defined by the min $s-t$ cut problem defined on the graph $G=(V,E)$ is integral (originally established in [6]). The $s-t$ mincut problem can be

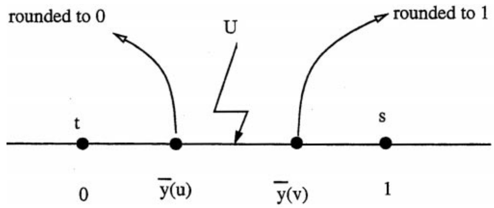

Fig. 1. Geometry of the randomization algorithm.

formulated as follows:

$$\begin{aligned}
\text{minimize} \quad & \sum_{(u,v)\in E} c(u,v)x(u,v) \\
\text{subject to} \quad & x(u,v) \geqslant y(u) - y(v), \quad (u,v) \in E, \\
& x(u,v) \geqslant y(v) - y(u), \quad (u,v) \in E, \\
& y(s) = 1, \;\; y(t) = 0, \\
& y(u), \; x(u,v) \in \{0,1\}.
\end{aligned}$$

Theorem 1. $Z_{\text{IP}} = Z_{\text{LP}}$.

Proof. Given a fractional optimal solution $(\bar{x}, \bar{y})$ of the LP relaxation, we consider the following randomization scheme:

- Generate a single random variable U uniformly distributed in $[0,1]$. Round all nodes u with $\bar{y}(u) < U$ to $y(u) = 0$, and all nodes u with $\bar{y}(u) > U$ to $y(u) = 1$ (cf. Fig. 1).

The process clearly produces a feasible cut.
Note that

$$\begin{aligned}
E(Z_H) &= E\left(\sum_e c_e x_e\right) \\
&= \sum_{e=(u,v)\in E} c_e P(\min(\bar{y}(u), \bar{y}(v)) \leqslant U \leqslant \max(\bar{y}(u), \bar{y}(v))) \\
&= \sum_{e=(u,v)\in E} c_e |\bar{y}(u) - \bar{y}(v)| \\
&= Z_{\text{LP}}.
\end{aligned}$$

Since $Z_{\text{LP}} \leqslant Z_{\text{IP}} \leqslant E(Z_H) = Z_{\text{LP}}$, we obtain $Z_{\text{LP}} = Z_{\text{IP}}$ and the relaxation is exact. $\square$

The same method also works for the directed $s-t$ cut problem. Since the LP relaxation of the directed $s-t$ cut problem is the dual of the max-flow problem, and since $Z_{\text{LP}} = Z_{\text{IP}}$ for the directed case, we can use this technique to obtain an easy probabilistic proof of the classical max-flow–min-cut theorem. Furthermore, since the randomization scheme does not violate precedence constraints of the type $y(u) \leqslant y(v)$, we see that adding this set of constraints to the formulation does not violate the integrality property.

Remark. Linial et al. [14] proved a more general result for the multi-commodity flow problem. As pointed out by one referee, their technique, when restricted to the single-commodity case, gives rise to the same proof presented in this paper.

2.2. Boolean optimization

The quadratic optimization problem is to

$$\begin{aligned} &\text{minimize} \quad \sum_{i,j} Q_{ij}x_ix_j + \sum_i c_ix_i \\ &\text{subject to} \quad x_i \in \{0,1\}. \end{aligned}$$

When the Q_{ij} are arbitrary, the problem is NP-hard. Several researchers have thus focused on identifying properties of Q_{ij} that allow the quadratic optimization problem to be solved in polynomial time.

Sign-balanced graph: Construct a graph G that has an edge between i and j if and only if $Q_{ij} \neq 0$. The edges which correspond to positive (resp. negative) Q_{ij} are called positive edges (resp. negative edges). G is called a *sign-balance graph* if it *does not* contain any cycle with an odd number of positive edges.

The notion of sign-balancedness essentially ensures that the graph G can be decomposed into $G_1 \cup G_2$, where $G_1 \cap G_2 = \emptyset, G_1 \subset G, G_2 \subset G$, and $\delta(G_1, G_2)$ contains the set of positive edges. Hansen and Simeone [10] show that the sign-balanced graph problem (i.e., a restricted version of the quadratic optimization problem where the coefficients Q_{ij} give rise to a sign-balance graph) is solvable in polynomial time. Note that this problem contains the maximum independent set problem on bipartite graphs as a special case.

Consider the following LP formulation for the problem:

$$\begin{aligned} &\text{minimize} \quad \sum_{i,j} Q_{ij}z_{ij} + \sum_i c_ix_i \\ &\text{subject to} \quad \begin{array}{ll} z_{ij} \leqslant x_i, & \text{if } Q_{ij} < 0, \\ z_{ij} \leqslant x_j, & \text{if } Q_{ij} < 0, \\ z_{ij} \geqslant x_i + x_j - 1, & \text{if } Q_{ij} > 0, \\ z_{ij}, x_i \geqslant 0, & \forall\, i,j, \\ z_{ij}, x_i \leqslant 1, & \forall\, i,j. \end{array} \end{aligned}$$

We show next that the integrality result of the above LP relaxation can be obtained in a direct manner. We round the fractional solution as follows:

- Generate a single random number U uniformly in $[0,1]$.
- Starting from an optimal solution of the LP relaxation $\bar{x}, \bar{z}$, round x_i to 1 if (i) $i \in G_1$ and $\bar{x}_i \geqslant U$, or (ii) $i \in G_2$ and $\bar{x}_i \geqslant 1 - U$.

Theorem 2. $Z_{\mathrm{LP}} = Z_{\mathrm{IP}}$ *if G is sign-balanced.*

Proof. Since $1 - U$ is also uniformly distributed in $[0,1]$, $P(x_i = 1) = \bar{x}_i$. For i,j both in G_1 or both in G_2, $P(x_ix_j = 1) = \min\{\bar{x}_i, \bar{x}_j\}$. For $i \in G_1$ and $j \in G_2$,

$$P(x_ix_j = 1) = P(U \leqslant \bar{x}_i, 1 - U \leqslant \bar{x}_j) = \max(0, \bar{x}_i + \bar{x}_j - 1).$$

At optimality, $\bar{z}_{ij} = \min(\bar{x}_i, \bar{x}_j)$ if $Q_{ij} < 0$, and $\bar{z}_{ij} = \max(0, \bar{x}_i + \bar{x}_j - 1)$ if $Q_{ij} > 0$. Then $Z_{\mathrm{LP}} \leqslant Z_{\mathrm{IP}} \leqslant E(Z_H) = \sum_{i,j} Q(i,j)E[x_ix_j] + \sum_i c_ix_i = \sum_{i,j} Q(i,j)\bar{z}_{ij} + \sum_i c_i\bar{x}_i = Z_{\mathrm{LP}}$. $\square$

2.3. Uncapacitated lot-sizing

Given a time horizon T, setup costs d_i $(i=1,\ldots,T)$ and production-inventory costs c_{ij} (indicating the cost of producing a unit in period i to satisfy a unit of demand in period j), the goal of the uncapacitated lot-sizing problem is to find a production schedule to minimize the total setup and production-inventory cost, and to satisfy the demand (denoted by f_i, $i=1,2,\ldots,T$) at all time periods. We assume further that back-ordering is not allowed in the model, and that production lead time is zero.

Let y_i be a $0-1$ decision variable that indicates whether we produce during period i. Let $w_{i,j}$ be the fraction of the demand f_j in period j that is met from production in period $i \leqslant j$. One formulation of the uncapacitated lot-sizing problem (see [16]) is as follows:

$$\begin{aligned} \text{minimize} \quad & \sum_{i,j} f_j c_{ij} w_{i,j} + \sum_i d_i y_i \\ \text{subject to} \quad & \sum_{k=1}^{j} w_{k,j} = 1, \qquad \forall j, \\ & w_{ij} \leqslant y_i, \qquad \forall i,j, \\ & y_i \in \{0,1\}, \qquad \forall i. \end{aligned}$$

It is well known that the resulting LP is integral (see [16]) when $c_{i,j} \geqslant c_{l,j}$ for all $i \leqslant l < j$. This condition is satisfied, for instance, when the unit production cost is constant throughout all time periods. Here we prove this result using randomized rounding.

For ease of exposition, we prove the result only for the case $c_{i,j} > c_{l,l}$ for all $i < l < j$. The argument can be adapted to prove the result for the general case.

In the optimal LP solution $(\bar{w}, \bar{y})$, we must have

$$\bar{w}_{i,j} \leqslant \bar{w}_{i,k}, \quad i \leqslant k < j.$$

Otherwise, from the constraints $\sum_{l=1}^{j} \bar{w}_{l,j} = 1$ and $\sum_{l=1}^{k} \bar{w}_{l,k} = 1$, there exists some time period i' such that

$$\bar{w}_{i',j} < \bar{w}_{i',k} \quad \text{whereas } \bar{w}_{i,j} > \bar{w}_{i,k},\ i,i' \leqslant j < k.$$

If $i' > i$, then transfering an ε $(\varepsilon > 0)$ amount of flow from $\bar{w}_{i,j}$ to $\bar{w}_{i',j}$ leads to feasible solution with smaller cost (due to savings in inventory holding). Similarly, if $i' < i$, then transfering an ε amount of flow from $\bar{w}_{i',k}$ to $\bar{w}_{i,k}$ leads to a feasible solution with smaller cost. Hence, without loss of generality, we can augment the LP relaxations with inequalities of the type

$$w_{i,j} \leqslant w_{i,k} \quad \text{if } j > k.$$

Let Z_{LP} denote the value of this augmented LP relaxation.

Theorem 3. $Z_{\mathrm{LP}} = Z_{\mathrm{IP}}$.

Proof. Let $(\bar{w}, \bar{y})$ be an optimal LP solution. Consider the following rounding method:

- Set $r = 1$.
- Set $y_r = 1$. Generate a random number U_r uniformly in $[0, \bar{y}_r]$.
  Let i be the index such that $\bar{w}_{r,i} \geqslant U > \bar{w}_{r,i+1}$. Set $w_{r,l}$ to 1, for all $l = r, \ldots, i$.
- Repeat step 2 with $r \leftarrow i+1$ until $r > T$.

We prove by induction that $P(y_i = 1) = \bar{y}_i$. Clearly $P(y_1 = 1) = \bar{y}_1 = 1$. Moreover,

$$\begin{aligned}
P(y_i = 1) &= \sum_{k<i} P(y_i = 1, y_{i-1} = 0, \dots, y_{k+1} = 0, y_k = 1) \\
&= \sum_{k<i} P(y_i = 1, y_{i-1} = 0, \dots, y_{k+1} = 0 | y_k = 1)\, P(y_k = 1) \\
&= \sum_{k<i} P(\bar{w}_{k,i-1} \geqslant U_k > \bar{w}_{k,i})\bar{y}_k \\
&= \sum_{k<i} (\bar{w}_{k,i-1} - \bar{w}_{k,i}) = \bar{w}_{i,i} = \bar{y}_i.
\end{aligned}$$

In addition, $P(w_{i,j} = 1) = P(w_{i,j} = 1, y_i = 1) = \bar{w}_{i,j}$. Hence the randomization scheme gives rise to an optimal integral solution with cost Z_{LP}, i.e., $Z_{\mathrm{IP}} = Z_{\mathrm{LP}}$. $\square$

2.4. *k*-median on cycle

Consider a cycle $C = (V, E)$ with $V = \{v_1, v_2, \dots, v_n\}$ and $E = \{(v_1, v_2), \dots, (v_{n-1}, v_n), (v_n, v_1)\}$. Let c be a nonnegative weight function on E. Suppose each node gives rise to a unit demand. The demand at v_i can be served by a facility located at v_j (say $j > i$) at a cost proportional to the distance between the two nodes defined as follows:

$$\min\{c(v_i, v_{i+1}) + \cdots + c(v_{j-1}, v_j), c(v_i, v_{i-1}) + \cdots + c(v_2, v_1) + c(v_1, v_n) + \cdots + c(v_{j+1}, v_j)\}.$$

The k-median problem is to locate k facilities at the nodes of the graph C in order to minimize the total cost of serving all the demands. A natural LP relaxation is as follows:

$$\begin{aligned}
&\text{(KMED)} \quad \min \textstyle\sum_{i,j} c_{i,j} x_{i,j} \\
&\text{subject to} \quad \textstyle\sum_{j=1}^{n} x_{i,j} = 1, \quad i \in \{1, \dots, n\}, \\
&\qquad x_{i,j} \leqslant y_j, \quad \forall i, j, \\
&\qquad \textstyle\sum_j y_j = k, \\
&\qquad 0 \leqslant y_j \leqslant 1, \quad \forall j.
\end{aligned}$$

Theorem 4. $Z_{\mathrm{LP}} = Z_{\mathrm{IP}}$.

Proof. Let $(\bar{x}, \bar{y})$ be an optimal fractional solution to the LP relaxation. Consider the rounding heuristic:

- Cover the interval $[0, k]$ with n nonoverlapping intervals, each of length $\overline{y}_j$, $j = 1, \dots, n$, *in that order*.
- Generate U uniformly in $[0, 1]$. Set y_j to 1 if one of the points in the set $S = \{U, 1 + U, \dots, (k-1) + U\}$ falls in the jth subinterval (of length $\overline{y}_j$). Let I_j denote the jth subinterval.
- Solve the assignment problem for the variables $(x_{i,j})$ with y_j fixed, by assigning the demand at i to the nearest location with y_j fixed at 1.

It is clear that $E(y_j) = \overline{y}_j$. We consider the assignment problem for a fixed i. The neighbors of node i are sorted in increasing distance from i. Without loss of generality, we assume that the order is $i = i_1, i_2, \dots, i_n$. A simple consequence of the geometry of the cycle C is that the union of the subintervals spanned by $I_{i_1}, I_{i_2}, \dots, I_{i_j}$ is of the type

$$\textstyle\bigcup_{l=1}^{j} I_{i_l} = [a, b] \text{ or } [b, k] \cup [0, a].$$

Note that the optimal solution $\bar{x}$ can be computed from $\bar{y}$:

$$\bar{x}_{i,i_j} = \begin{cases} \min\left\{\bar{y}_{i_j}, 1 - \sum_{l:l<j} \bar{y}_{i_l}\right\} & \text{if } \sum_{l:l<j} \bar{y}_{i_l} < 1, \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand,

$$P(x_{i,i_j} = 1) = P(S \cap I_{i_j} \neq \emptyset, S \cap I_{i_l} = \emptyset \ \forall l < j)$$

$$= \begin{cases} \min\left\{\bar{y}_{i_j}, 1 - \sum_{l:l<j} \bar{y}_{i_l}\right\} & \text{if } \sum_{l:l<j} \bar{y}_{i_l} < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $E(x_{i,i_j}) = \bar{x}_{i,i_j}$. □

3. Dependent rounding and approximation algorithms

In this section, we use the technique proposed in the previous section to obtain approximation results for two classes of problems: the feasible cut and the min-k-sat problem. Our analysis shows that the LP relaxations are within 2 times of the optimum for both problems. Hochbaum [12] has recently obtained an refinement of these results by showing that the LP relaxations are half-integral, and hence the 2-approximation results follow immediately.

3.1. Feasible cut

The feasible cut problem on a graph $G = (V, E)$ was introduced in Yu and Cheriyan [21]. Let M be a set of pairs of nodes in G. The problem asks for a cut of minimum weight, which contains a designated vertex s, but not any node pair $(u, v) \in M$. Yu and Cheriyan showed that the node covering problem can be reduced to this problem. Furthermore, the reduction preserves the approximation bound. Hence, any 2-ε approximation algorithm for the feasible cut problem would imply the same improvement for the node covering problem. Yu and Cheriyan (and independently Ravi [18]) proposed a 2-approximation algorithm for this problem. We show next how to obtain a similar bound using the rounding idea of Section 2.1.

Consider the following formulation of the feasible cut problem.

$$\begin{array}{lll} \text{minimize} & \sum_{(u,v)\in E} c(u,v)x(u,v) & \\ \text{subject to} & x(u,v) \geqslant y(u) - y(v), & (u,v) \in E, \\ & x(u,v) \geqslant y(v) - y(u), & (u,v) \in E, \\ & y(u) + y(v) \leqslant 1, & (u,v) \in M, \\ & y(s) = 1, & \\ & y(u), x(u,v) \in \{0,1\} \ . & \end{array}$$

The randomized rounding algorithm is as follows:

- Starting with an optimal solution of the LP relaxation $(\bar{x}, \bar{y})$, position the nodes in $[0,1]$ according to the value of $\bar{y}(u)$.
- Generate a single random variable U uniformly distributed in $[\frac{1}{2}, 1]$. Round all nodes u with $\bar{y}(u) < U$ to $y(u) = 0$, and all nodes u with $\bar{y}(u) > U$ to $y(u) = 1$.

Since for $(u,v) \in M$ at least one from $\bar{y}(u)$ and $\bar{y}(v)$ is larger than $\frac{1}{2}$, the rounding process produces a feasible cut.

Theorem 5. $Z_{\mathrm{IP}} \leqslant E(Z_H) \leqslant 2Z_{\mathrm{LP}}$.

Proof. If $\max(\bar{y}(u), \bar{y}(v)) \leqslant \frac{1}{2}$, then

$$E(x(u,v)) = 0.$$

If $\min(\bar{y}(u), \bar{y}(v)) \leqslant \frac{1}{2} \leqslant \max(\bar{y}(u), \bar{y}(v))$, then

$$\begin{aligned} E(x(u,v)) &= P(U \in [\tfrac{1}{2}, \max(\bar{y}(u), \bar{y}(v))]) \\ &= 2(\max(\bar{y}(u), \bar{y}(v) - \tfrac{1}{2}) \leqslant 2|\bar{y}(u) - \bar{y}(v)|. \end{aligned}$$

If $\frac{1}{2} \leqslant \min(\bar{y}(u), \bar{y}(v))$, then

$$E(x(u,v)) = 2|\bar{y}(u) - \bar{y}(v)|.$$

In all cases $E(x(u,v)) \leqslant 2|\bar{y}(u) - \bar{y}(v)| \leqslant 2\bar{x}(u,v)$. $\square$

As before, the bound of 2 holds even in the presence of precedence constraints $y(u) \leqslant y(v)$.

3.2. Minimum satisfiability

Kohli et al. [13] introduced the minimum satisfiability problem as an analog of the maximum satisfiability problem. They proved that this version of the satisfiability problem remains NP-hard, even when each clause contains at most two literals (min-2-sat).

Given a set of literals and clauses, let x_i be a literal and C_j the jth clause. Let I_j^+ be the set of unnegated literals in clause C_j and I_j^- the set of negated literals in C_j. Each literal is assigned to be "true" or "false". The clause C_j is a satisfied clause only if one of the literals in I_j^+ is assigned to be "true" or if one of the literals in I_j^- is assigned to be false. The min-sat problem is to find an assignment of the literals to minimize a weighted sum of satisfied clauses. In the rest of this section, we provide an improvement of their result in the case when the number of literals in each clause is bounded. Let k denote an upper bound on the number of literals in each clause.

The problem can be formulated as follows:

$$\begin{aligned} \text{minimize} \quad & \sum_j w_j z_j \\ \text{subject to} \quad & z_j \geqslant x_i, && \forall i \in I_j^+, \\ & z_j \geqslant 1 - x_i, && \forall i \in I_j^-, \\ & x_i, z_j \in \{0,1\}\ . \end{aligned}$$

We remark that we can assume, without loss of generality, that for every j the sets I_j^+ and I_j^- are disjoint, because if there is an $i \in I_j^+ \cap I_j^-$, then we will have

$$z_j \geqslant x_i, \qquad z_j \geqslant 1 - x_i.$$

These inequalities force $z_j = 1$ in all feasible integer solutions, and hence, in a preprocessing step we can identify all those clauses for which $I_j^+ \cap I_j^- \neq \emptyset$, set the corresponding variables z_j equal to 1 and hence reduce the problem to one for which $I_j^+ \cap I_j^- = \emptyset$.

We show that the following randomized rounding method achieves a bound of $2(1 - 1/2^k)$ for min-sat:

- Let $\bar{x}$ and $\bar{z}$ be an optimal LP solution. It is easy to see that

$$\bar{z}_j = \max\left[\max_{i \in I_j^+} \bar{x}_i, \max_{i \in I_j^-}(1 - \bar{x}_i)\right].$$

- Split the x_i's into two sets A, B randomly, i.e., each x_i is assigned to the set A (resp. B) with probability $\frac{1}{2}$.
- Generate U in $[0, 1]$ uniformly.
- For x_i in A, set $x_i = 1$ if $\bar{x}_i > U$, and 0 otherwise.
- For x_i in B, set $x_i = 1$ if $\bar{x}_i > 1 - U$, and 0 otherwise.

Theorem 6. $E(Z_H) \leqslant 2(1 - 1/2^k) Z_{\mathrm{LP}}$.

Proof. Consider clause C_j with literals $\{x_1^A, \ldots, x_a^A\}$ and $\{x_1^B, \ldots, x_b^B\}$ in A and B, respectively.

$$\begin{aligned}
P(z_j = 1) &= P\left(\left\{U \leqslant \max_{x_i^A \in I_j^+} \bar{x}_i^A\right\} \cup \left\{U \geqslant \min_{x_i^A \in I_j^-} \bar{x}_i^A\right\} \cup \left\{1 - U \leqslant \max_{x_i^B \in I_j^+} \bar{x}_i^B\right\} \cup \left\{1 - U \geqslant \min_{x_i^B \in I_j^-} \bar{x}_i^B\right\}\right) \\
&\leqslant \max\left(\max_{x_i^A \in I_j^+} \bar{x}_i^A, 1 - \min_{x_i^B \in I_j^-} \bar{x}_i^B\right) + 1 - \min\left(\min_{x_i^A \in I_j^-} \bar{x}_i^A, 1 - \max_{x_i^B \in I_j^+} \bar{x}_i^B\right) \\
&= \max\left\{\max_{x_i^A \in I_j^+} \bar{x}_i^A + \max_{x_i^A \in I_j^-}(1 - \bar{x}_i^A), \max_{x_i^A \in I_j^+} \bar{x}_i^A + \max_{x_i^B \in I_j^+} \bar{x}_i^B, \max_{x_i^B \in I_j^-}(1 - \bar{x}_i^B) + \max_{x_i^A \in I_j^-}(1 - \bar{x}_i^A),\right. \\
&\qquad \left.\max_{x_i^B \in I_j^-}(1 - \bar{x}_i^B) + \max_{x_i^B \in I_j^+} \bar{x}_i^B\right\} \\
&\leqslant 2\bar{z}_j.
\end{aligned}$$

Furthermore, with probability $1/2^k$, all elements in I_j^+ are assigned to A and all in I_j^- are assigned to B, in which case the last inequality reduces to $\bar{z}_j$. Symmetrically, with probability $1/2^k$, all elements in I_j^- are assigned to A and all in I_j^+ are assigned to B, in which case the last inequality reduces to $\bar{z}_j$. Hence,

$$E(Z_H) \leqslant \left(2\left(1 - \frac{2}{2^k}\right) + \frac{2}{2^k}\right) Z_{\mathrm{LP}} = 2\left(1 - \frac{1}{2^k}\right) Z_{\mathrm{LP}}. \qquad \square$$

The above bound is tight, as can be seen from the following example. Consider a formula with k literals $x_1, \ldots, x_k$. There are 2^k combinations of clauses, since each literal can be negated or not. Clause j corresponds to the set S of literals that are not negated. We will identify a clause by the set S of literals that appear unnegated in it. All clauses have weight equal to 1. In this case the formulation becomes

$$\begin{array}{lll} \text{minimize} & \sum_S z_S & \\ \text{subject to} & z_S \geqslant x_i, & \forall i \in S,\ \forall S, \\ & z_S \geqslant 1 - x_i, & \forall i \notin S,\ \forall S, \\ & x_i, z_S \in \{0,1\}\ . & \end{array}$$

The optimal solution of the linear programming relaxation has $x_i = \frac{1}{2}$ and therefore, $Z_{\mathrm{LP}} = 2^{k-1}$. We next find the optimal solution to the integer programming problem. Consider an arbitrary integer solution. Let T be such that $x_i = 1, i \in T$ and $x_i = 0, i \notin T$. Then, all variables z_S $(S \neq T)$ are forced to equal 1 except variable z_T. Since this is true for every T, we conclude that $Z_{\mathrm{IP}} = 2^k - 1$. Therefore, for this example,

$$Z_{\mathrm{IP}} = 2\left(1 - \frac{1}{2^k}\right) Z_{\mathrm{LP}}.$$

Acknowledgements

We would like to thank the associate editor and referee for making suggestions that improved the presentation of the current paper. We would like to thank the referee especially for pointing out several discrepencies in an earlier version of the paper.

References

[1] D. Bertsimas, C. Teo, R. Vohra, Nonlinear relaxations and improved randomized approximation algorithms for multicut problems, Proceedings of the Fourth IPCO Conference, 1995, pp. 29–39, in Networks, in press.
[2] D. Bertsimas, R. Vohra, Rounding Algorithms for covering problems, Math. Programming 80 (1998) 63–89.
[3] H. Brönnimann, M. Goodrich, Almost optimal set covers in finite VC-dimension, Proceedings of the 10th Annual Symposium Computational Geometry, 1994, pp. 293–301.
[4] V. Chvátal, A greedy heuristic for the set-covering problem, Math. Oper. Res. 4 (1979) 233–235.
[5] W. Fernandez de la Vega, G. Lueker, Bin packing solved within $1+\varepsilon$ in linear time, Combinatorica 1 (1981) 349–355.
[6] L.R. Ford, D.R. Fulkerson, Flows in Networks, Princeton University Press, Princeton, NJ, 1962.
[7] M.X. Goemans, D. Williamson, A new 3/4 approximation algorithm for MAX SAT, Proceedings of the Third IPCO Conference, 1993, pp. 313–321.
[8] M.X. Goemans, D. Williamson, .878 approximation algorithms for MAX-CUT and MAX 2SAT, Proceedings of the 26th Annual ACM STOC, 1994, pp. 422–431.
[9] M. Grötschel, L. Lovász, Combinatorial Optimization: A survey, DIMACS Technical Report 93-29, 1993.
[10] P. Hansen, B. Simeone, Unimodular functions, Discrete Appl. Math. 14 (1986) 269–281.
[11] D. Hochbaum, Approximation algorithms for set covering and vertex cover problems, SIAM J. Comput. 11 (1982) 555–556.
[12] D. Hochbaum, Instant recognition of half integrality and 2-approximation, Proceedings of the APPROX98, Lecture Notes in Computer Science, vol. 1444, Springer, Berlin, 1998, pp. 99–110.
[13] R. Kohli, R. Khrishnamurti, P. Mirchandani, The minimum satisfiability problem, SIAM J. Discrete Math. 7 (2) (1994) 275–283.
[14] N. Linial, E. London, Y. Rabinovich, The geometry of graphs and some of its algorithmic applications, Proceedings of the 35th Annual Symposium on Foundations of Computer Science, 1994, pp. 577–591.
[15] M. Marathe, S. Ravi, On approximation algorithms for the minimum satisfiability problems, preprint, 1995.
[16] G. Nemhauser, L. Wolsey, Integer and Combinatorial Optimization, Wiley, New York, 1988.
[17] P. Raghavan, C. Thompson, Randomized rounding: a technique for provably good algorithms and algorithmic proofs, Combinatorica 7 (1987) 365–374.
[18] R. Ravi, private communication.
[19] D. Shmoys, E. Tardos, An approximation algorithm for the generalized assignment problem, Math. Programming 62 (1993) 461–474.
[20] M. Yannakakis, On the approximation of maximum satisfiability, Proceedings of the Third ACM-SIAM Symposium on Discrete Algorithms, 1992, pp. 1–9.
[21] B. Yu, J. Cheriyan, Approximation algorithms for feasible cut and multicut problems, in: P. Spirakis (Ed.), Proceedings of the Algorithms – ESA'95, Third Annual European Symposium, Lecture Notes in Computer Science 979, Springer, New York, 1995, pp. 394–408.