Algorithmica (1997) 18: 44–66

Algorithmica

Randomized Approximation of Bounded Multicovering Problems[1]

D. Peleg,[2] G. Schechtman,[3] and A. Wool[4]

Abstract. The problem of finding approximate solutions for a subclass of multicovering problems denoted by *ILP*(k, b) is considered. The problem involves finding $x \in \{0, 1\}^n$ that minimizes $\sum_j x_j$ subject to the constraint $Ax \geq \boldsymbol{b}$, where A is a 0–1 $m \times n$ matrix with at most k ones per row, $\boldsymbol{b}$ is an integer vector, and b is the smallest entry in $\boldsymbol{b}$. This subclass includes, for example, the Bounded Set Cover problem when $b = 1$, and the Vertex Cover problem when $k = 2$ and $b = 1$.

An approximation ratio of $k - b + 1$ is achievable by known deterministic algorithms. A new randomized approximation algorithm is presented, with an approximation ratio of $(k - b + 1)(1 - (c/m)^{1/(k-b+1)})$ for a small constant $c > 0$. The analysis of this algorithm relies on the use of a new bound on the sum of independent Bernoulli random variables, that is of interest in its own right.

Key Words. Integer linear programs, Randomized rounding, Set cover, Approximation algorithms.

1. Introduction

1.1. *The Problem.* The problem dealt with in this paper is the *ILP*(k, b) problem, formally defined as

$$ILP(k, b): \quad \begin{cases} \min_{\boldsymbol{x}} \sum_{j=1}^{n} x_j, \\ A\boldsymbol{x} \geq \boldsymbol{b}, \\ \boldsymbol{x} \in \{0, 1\}^n, \end{cases}$$

where A is an $m \times n$ matrix, $a_{ij} \in \{0, 1\}$, such that, for all i, $\sum_{j=1}^{n} a_{ij} \leq k$, and $\boldsymbol{b} = (b_1, \ldots, b_m)$ is an integer vector such that $b_i \geq b$ for all i. Thus problems in the *ILP*(k, b) class consist of a system of m *inequalities* of n *variables*, where each inequality uses at most k variables, its right-hand side is "$\geq b_i$" such that the smallest b_i is b, and the goal is to minimize the cost function $\sum_j x_j$.

We demonstrate the richness of the class *ILP*(k, b) by exhibiting a number of NP-hard problems, taken from [GJ], which can all be formulated in a natural way as special cases of the *ILP*(k, b) problem. They are all presented in their standard version, i.e., with $b_i = b = 1$ in our terminology.

---

[1] The first author was supported in part by a Walter and Elise Haas Career Development Award and by a grant from the Israeli Science Foundation. This work was done while the third author was at the Department of Applied Mathematics and Computer Science, The Weizmann Institute, Rehovot 76100, Israel.

[2] Department of Applied Mathematics and Computer Science, The Weizmann Institute, Rehovot 76100, Israel. peleg@wisdom.weizmann.ac.il.
[3] Department of Theoretical Mathematics, The Weizmann Institute, Rehovot 76100, Israel.
[4] Bell Laboratories, 700 Mountain Avenue, Murray Hill, NJ 07974, USA. yash@research.bell-labs.com.

Received June 6, 1994; revised July 7, 1995. Communicated by M. X. Goemans.

VERTEX COVER. Given a graph $G = (V, E)$, find a subset $V' \subseteq V$ with minimal cardinality, such that for each edge $(u, v) \in E$ at least one of u and v belongs to V'.

ILP(2, 1) formulation: Define a variable x_v for each $v \in V$. The inequalities are $x_u + x_v \geq 1, \forall (u, v) \in E$.

BOUNDED DEGREE DOMINATING SET. Given a graph $G = (V, E)$, with maximum degree Δ, find a subset $V' \subseteq V$ with minimal cardinality, such that for each $u \in V \backslash V'$ there is a $v \in V'$ such that $(u, v) \in E$.

ILP($\Delta + 1$, 1) formulation: Define a variable x_v for each $v \in V$. Denote the *neighborhood* of a vertex v by $\Gamma(v) = \{u \in V | (u, v) \in E\} \cup \{v\}$. The inequalities are $\sum_{u \in \Gamma(v)} x_u \geq 1, \forall v \in V$.

PARTIAL FEEDBACK EDGE SET. Given a graph $G = (V, E)$ find a subset $E' \subseteq E$ with minimal cardinality, such that E' contains at least one edge from each circuit of length L or less for some fixed integer $L \geq 3$.

ILP(L, 1) formulation: Define a variable x_e for each $e \in E$. The inequalities are $\sum_{e \in C} x_e \geq 1, \forall$ circuit $C, |C| \leq L$.

BOUNDED SET COVER. Given a collection $C = \{C_1, \dots, C_n\}$ of subsets of a finite set S, denote the *rank* of an element $i \in S$ by $\rho_i = |\{j | i \in C_j\}|$, and let $\max_i \rho_i \leq k$ for some constant k. Find a subcollection $C' \subseteq C$ with minimal cardinality, such that every element of S belongs to at least one member of C'.

ILP(k, 1) formulation: Define a variable x_j for each set C_j, $j = 1, \dots, n$. The inequalities are $\sum_{C_j \ni i} x_j \geq 1, \forall i \in S$.

BOUNDED HITTING SET. Given a collection $C = \{C_1, \dots, C_m\}$ of subsets of a finite set S, such that $\max_i |C_i| \leq k$ for some constant k, find a subset $S' \subseteq S$ with minimal cardinality, such that S' contains at least one element from each subset in C.

ILP(k, 1) formulation: Define a variable x_j for each element $j \in S$. The inequalities are $\sum_{j \in C_i} x_j \geq 1$, for $i = 1, \dots, m$.

BOUNDED CHOICE TEST SET. Given a collection $C = \{C_1, \dots, C_n\}$ of subsets of a finite set S, let D_{uv} be the set of indices of possible test sets for each $u, v \in S$, $u \neq v$: $D_{uv} = \{j | (u \in C_j \wedge v \notin C_j) \vee (u \notin C_j \wedge v \in C_j)\}$, such that $\max_{u,v} |D_{uv}| \leq k$. Find a subcollection $C' \subseteq C$ with minimal cardinality, such that, for each pair of distinct elements $u, v \in S$, there is some set $C_r \in C'$ that contains exactly one of u and v.

ILP(k, 1) formulation: Define a variable x_j for each set $C_j \in C$. The inequalities are $\sum_{j \in D_{uv}} x_j \geq 1, \forall u, v \in S, u \neq v$.

The *ILP*(k, b) problem is NP-hard. The case when $b = 1$ and $k = n$ is the Minimum Set Cover problem, and when $k = 2$ and $b = 1$ it is the Vertex Cover problem, see

[GJ]. This also is a MAX-SNP-hard problem [F], i.e., does not have a polynomial-time approximation scheme unless P = NP [ALM$^+$]. In contrast, the relaxed fractional problem, $LP(k, b)$, in which the solution vector is not required to be integral, i.e., $x \in [0, 1]^n$, appears to be easier; the optimal solution for the relaxed problem can be found in polynomial time using Linear Programming algorithms [Kha], [Kar].

The best known problems in the $ILP(k, b)$ class are those with $b = 1$. In this work we concentrate on the effect of the b parameter, when it is larger than 1. One typical situation in which $b > 1$ may naturally appear is in network design problems, when a fault tolerance requirement is imposed on the design. As an example, consider the following center selection problem [HS], [BKP2]. The network is given as a graph $G = (V, E)$ with maximum degree Δ. We wish to select nodes of the network as centers, in which copies of some resource or service are to be placed. Each node should either be a center, or have a center as an immediate neighbor. For increased crash resilience, we also require that every node will still have at least one functioning center in its neighborhood even after b nodes have crashed. To meet this requirement we must select the centers so that at least $b + 1$ of them appear in each node's neighborhood. For cost efficiency, we need to select the minimal number of centers possible. Therefore, if we define a binary variable x_v for each network node v, with $x_v = 1$ meaning "v is selected," then writing an inequality per neighborhood and minimizing $\sum_v x_v$ as our cost function will give us an instance of $ILP(\Delta + 1, b + 1)$.

1.2. *Related Work.* The earliest published approximation algorithms for the Minimum Set Cover problem [J], [L], [Chv], or $ILP(n, 1)$, use the greedy heuristic. This heuristic chooses variables one by one, according to the number of inequalities they satisfy. The approximation ratio achieved by these algorithms is $R_{\text{greedy}} \leq 1 + \ln m$.

In [D], [Wol], and [BKP1] the greedy heuristic is examined in a more general setting. In particular, it is shown to be applicable to the $ILP(k, b)$ problem for all values of b and k. The analysis shows that the approximation ratio achieved is $R_{\text{greedy}} \leq \ln(mb)$.

There are several results on $ILP(k, b)$ problems that take advantage of the bounded number of variables per inequality. The first algorithm for the Vertex Cover problem with a constant approximation ratio of 2 is attributed to Gavril in [GJ]. This algorithm repeatedly picks an uncovered edge, and takes both of its endpoints into the output set, until all the edges are covered.

A different algorithm was described by Hochbaum [Hoc1] for the subclass $ILP(k, 1)$. The algorithm solves the relaxed fractional $LP(k, 1)$ problem, and takes all the variables with fractional values of at least $1/k$ to be the solution set. The analysis shows an approximation ratio of k (and in particular, 2 for the Vertex Cover problem).

Both Gavril's algorithm and Hochbaum's algorithm can be extended to the more general $ILP(k, b)$ setting, and both generalizations yield approximation ratios of $k - b + 1$ [Woo].

Subsequent algorithms for the Vertex Cover problem, with approximation ratios of slightly less than 2 (for any *fixed* value of m), rely on a result by Nemhauser and Trotter [NT], concerning the properties of the weighted Vertex Cover problem. Their work shows that there always exists an optimal solution to the fractional $LP(2, 1)$ problem where the variables all have values of 0, 1, or $\frac{1}{2}$. Hochbaum [Hoc2] uses this property to obtain improved approximation algorithms for the Vertex Cover problem. Unfortunately, it may be difficult to extend this approach even to *ILP*(3, 1) since there are examples in this class where the optimal solution cannot behave as multiples of $\frac{1}{3}$ [Woo].

The best currently known approximation algorithm for the weighted Vertex Cover problem is by Bar-Yehuda and Even [BE]. They use the Nemhauser–Trotter preprocessing, coupled with a local-ratio theorem to obtain several approximation algorithms, the best of which has an approximation ratio of $R \leq 2 - (\log\log m)/(2\log m)$.

The weighted *ILP*(k, b) problem, in which the objective function is $\min \sum_j w_j x_j$, is studied by Hall and Hochbaum [HH]. Their algorithm is a generalization of the approach in [BE], and has an approximation ratio of k. However, it is not hard to see that in the unweighted case the algorithm becomes a variant of Gavril's algorithm, and hence has an approximation ratio of $k - b + 1$ for *ILP*(k, b) instances. This algorithm has a time complexity of $O(n \cdot \max\{n, m\})$, which is better than that of algorithms requiring the solution of the relaxed linear program (e.g., [Hoc1]).

A randomized algorithm for the *b-matching* problem in hypergraphs has been proposed by Raghavan and Thompson [RT]. The problem is closely related to the *ILP*(k, b) problem, with the following differences: the inequalities are "$\leq b$," the goal is maximizing the cost, and there is no k bound on the number of variables per inequality. The algorithm solves the relaxed fractional problem using Linear Programming, and then uses the values of the optimal solution vector as the defining *probabilities* for independent Bernoulli random variables.

A computational study comparing the performance of many of the above-mentioned algorithms on large-scale problem instances is [GW]. It is shown that on the tested instances, the quality of the solutions found differs considerably between algorithms with the same approximation ratio.

1.3. *New Results.* We first show some interesting structural properties of the *ILP*(k, b) problem. We present approximation-preserving reductions to and from the *ILP*$(k, 1)$ problem, thus showing their equivalence in the sense of [PY]. We also present an example showing that the gap between the optima of the integral problem and its fractional relaxation may approach a factor of $k - b + 1$.

Our randomized algorithm RND for *ILP*(k, b) has an approximation ratio of

$$R_{\mathrm{RND}} \leq (k - b + 1)\left(1 - \left(\frac{c}{m}\right)^{1/(k-b+1)}\right)$$

for a small constant $c > 0$. This is better than the ratios of the deterministic algorithms for any fixed value of m, although asymptotically it is the same. The analysis of this algorithm relies on the use of a new bound that we prove on the sum of independent Bernoulli random variables.

An extended abstract of this paper can be found in [PSW].

2. Preliminaries

2.1. *Definitions and Notation.* We refer to a solution to *ILP*(k, b) either as a vector, e.g., $\boldsymbol{x} \in \{0, 1\}^n$, or as a set of chosen variables $T = \{j \mid x_j = 1\}$. We usually identify the variables with their indices, i.e., we may interchangeably use either x_j or j to denote the jth variable.

We use the following notation. The *cost* of a solution $z \in [0, 1]^n$ is defined to be $C(z) = \sum_{j=1}^{n} z_j$. We use $x^{\rm opt}$ to denote the vector of an optimal (integral) solution to $ILP(k, b)$ and $C^{\rm opt} = C(x^{\rm opt})$ to denote its cost.

We use $LP(k, b)$ to denote the corresponding relaxed fractional problem, x^* to denote the vector of an optimal solution to $LP(k, b)$, and C^* to denote the cost of an optimal feasible solution to $LP(k, b)$, namely, $C^* = C(x^*)$.

We are interested in several measures of the algorithm's quality. For some algorithm, B, let $C_B(A)$ denote the cost of the solution found by B on the problem defined by the matrix A. Then the *approximation ratio* of B is

$$R_B(m) = \sup_A \left\{ \frac{C_B(A)}{C^{\rm opt}(A)} | A \text{ has } m \text{ rows} \right\},$$

and its asymptotic approximation ratio is $R_B = \lim_{m\to\infty} R_B(m)$. We may also consider the *fractional approximation ratio* of B, $R_B^*(m) = \sup_A\{(C_B(A)/C^*(A)) | A \text{ has } m \text{ rows}\}$, and the asymptotic fractional approximation ratio, $R_B^* = \lim_{m\to\infty} R_B^*(m)$.

2.2. *Technical Lemmas.* In the analysis of our algorithm, we deal with sets of variables in the range $[0, 1]$, whose sum is bounded from below. We now present two technical lemmas regarding the properties of such variables.

The first lemma characterizes the distribution of "large" values of variables in the range $[0, 1]$ whose sum is bounded from below.

LEMMA 2.1. *Let ℓ, b, t be integers, $t \le b \le \ell - t$. Let $0 \le x_j \le 1$, for $j = 1, \ldots, \ell$, be such that $\sum_{j=1}^{\ell} x_j \ge b$. Then at least $b-t+1$ of the values x_j satisfy $x_j \ge 1/(\ell-b-t+2)$.*

PROOF. Assume that $\sum_{j=1}^{\ell} x_j \ge b$, but there are at most $b - t$ values x_j satisfying $x_j \ge 1/(\ell - b - t + 2)$. Assume without loss of generality that $x_1 \ge x_2 \ge \cdots \ge x_\ell$. Then

$$\begin{aligned}\sum_{j=1}^{\ell} x_j &= \sum_{j=1}^{b-t} x_j + \sum_{j=b-t+1}^{\ell} x_j \le (b-t) + \sum_{j=b-t+1}^{\ell} x_j \\ &< (b-t) + (\ell - b + t) \cdot \frac{1}{\ell - b - t + 2} \\ &= (b-t+1) + \frac{2t-2}{\ell - b - t + 2} \le b,\end{aligned}$$

contradiction. The last step of the derivation uses the fact that $b \le \ell - t$, and therefore $(2t-2)/(\ell - b - t + 2) \le t - 1$. □

The useful cases of this lemma, which are applied in later parts of this work, are:

- $t = 1$. If $1 \le b \le \ell - 1$, then at least b of the values x_j satisfy $x_j \ge 1/(\ell - b + 1)$.
- $t = 2$. If $2 \le b \le \ell - 2$, then at least $b - 1$ of the values x_j satisfy $x_j \ge 1/(\ell - b)$.

Our second lemma deals with the maximum of a product of linear terms. In the analysis of our randomized algorithm (Section 5) we need to bound the maximal value of a specific multilinear function. We need to maximize this function constrained to the domain where the variables are in the range $[0, 1]$ and their sum is bounded from below. The following lemma gives us the desired result.

LEMMA 2.2. *For $a \geq 0$ let $\langle a \rangle = \min\{a, 1\}$. For any $\delta > 1$ and integer d let $f(x) = \prod_{j=1}^{d}(1 - \langle \delta x_j \rangle)$, $x \in [0, 1]^d$ and $\sum_{j=1}^{d} x_j \geq 1$. If $d > \delta$, then $f(x)$ attains its maximum when $x_j = 1/d$ for all j, and $f(x) = (1 - \delta/d)^d$. If $d \leq \delta$, then $f \equiv 0$.*

PROOF. The condition $\sum_{j=1}^{d} x_j \geq 1$ implies that there exists j such that $x_j \geq 1/d$. If $d \leq \delta$, then, for this j, $\langle \delta x_j \rangle = 1$, and $f = 0$.

Assume now that $d > \delta$. If x has some $x_j \geq 1/\delta$, then $f(x) = 0$. Therefore since $1/\delta < 1$, we need to search for the maximum of the function only inside the domain

$$H_d(\delta) = \left[0, \frac{1}{\delta}\right)^d \cap \left\{x \,|\, \sum_{j=1}^{d} x_j \geq 1\right\}.$$

In this domain we can drop the $\langle \cdot \rangle$, and consider $f(x) = \prod_{j=1}^{d}(1 - \delta x_j)$.

By the fact that the geometric mean is always less than or equal to the arithmetic mean (with equality only if the numbers are all equal), we have that

$$\prod_{j=1}^{d}(1 - \delta x_j) \leq \left(\frac{\sum_{j=1}^{d}(1 - \delta x_j)}{d}\right)^d = \left(1 - \frac{\delta \sum_{j=1}^{d} x_j}{d}\right)^d \leq \left(1 - \frac{\delta}{d}\right)^d,$$

with equality everywhere only if all the x_j's are equal and $\sum_j x_j = 1$. □

3. Structural Properties

3.1. *Reductions to and from the $b = 1$ Case.* The following propositions show that the approximability of *ILP*(k, b) problems with $b > 1$ is closely related to that of *ILP*$(k, 1)$.

PROPOSITION 3.1. *Any instance of ILP(k, b) with m inequalities can be reduced to an instance of ILP$(k - b + 1, 1)$ with at most $\sum_{i=1}^{m} \binom{k}{b_i - 1}$ inequalities, and if $b_i = b$ for all i, then there are at most $m\binom{k}{b-1}$ inequalities.*

PROOF. Consider the inequality

$$\sum_{j \in T} x_j \geq b_i, \qquad |T| = \ell \leq k. \tag{1}$$

We replace it with $\binom{\ell}{b_i - 1}$ inequalities:

$$\sum_{j \in S} x_j \geq 1, \qquad \forall S \subseteq T, \quad |S| = \ell - b_i + 1. \tag{2}$$

It suffices to show that, for any vector $y \in \{0, 1\}^\ell$,

$$y \text{ satisfies (1)} \quad \Leftrightarrow \quad y \text{ satisfies (2)}.$$

1. Assume y does not satisfy (2). That is, assume that there exists $S \subseteq T, |S| = \ell - b_i + 1$, such that $\sum_{j \in S} y_j = 0$. Then

$$\sum_{j \in T} y_j = \sum_{j \in T \setminus S} y_j \leq b_i - 1,$$

hence y does not satisfy (1) either.
2. Assume y does not satisfy (1). That is, assume that $\sum_{j \in T} y_j < b_i$. Let $L = \{j \mid y_j = 1\}$ (or equivalently, $T \setminus L = \{j \mid y_j = 0\}$). Then

$$|L| \leq b_i - 1 \quad \Rightarrow \quad |T \setminus L| \geq \ell - b_i + 1.$$

It follows that there exists an $S \subseteq T \setminus L \subseteq T$, $|S| = \ell - b_i + 1$ such that $\sum_{j \in S} y_j = 0$, hence y does not satisfy (2) either. □

REMARKS.

• The transformation does not apply to the *relaxed* problem *LP*(k, b), i.e., the resulting *LP*$(k - b + 1, 1)$ instance will not necessarily be equivalent to the original instance.

• In general it may be impractical to use the transformation, since the output set of inequalities is considerably larger than the input one. Both the algorithm of [HH] and our randomized algorithm have the same approximation ratios on *ILP*(k, b) and on the corresponding Set Cover problem, so using the transformation only increases the time-complexity.

• The increase in the number of inequalities could be by a nonpolynomial factor, even when $k - b + 1 < \ln(mb)$, i.e., when the algorithm of [HH] has a better approximation ratio than the greedy algorithm [BKP1].

• When b_i or $k - b_i$ is bounded for all i the transformation is polynomial, since

$$b_i \leq M \ \ or \ \ k - b_i \leq M \quad \Rightarrow \quad \binom{k}{b_i - 1} = \binom{k}{k - b_i + 1} \approx k^{M+1} \leq n^{M+1}.$$

Therefore it is an L-reduction, using the terminology from [PY].

PROPOSITION 3.2. *Any instance of ILP*$(k, 1)$ *with n variables and optimal cost C^{opt} can be reduced to an instance of ILP*$(k + b - 1, b)$ *with $n + b - 1$ variables and optimal cost* $C^{\text{opt}} + b - 1$.

PROOF. Consider an instance of *ILP*$(k, 1)$ with n variables $x_1, \ldots, x_n$. We add $b - 1$ new variables, x_s for $s = n + 1, \ldots, n + b - 1$. We replace the inequalities

$$(3) \qquad \sum_{j \in S_i} x_j \geq 1, \qquad \text{for } i = 1, \ldots, m,$$

with

$$(4) \qquad \sum_{j\in S_i} x_j + \sum_{s=n+1}^{n+b-1} x_s \geq b, \qquad \text{for } i = 1, \dots, m.$$

The inequalities of (4) contain at most $k + b - 1$ variables each, since $|S_i| \leq k$ for all i. We need to show that

$$\exists y,\ \ y \text{ satisfies (3)},\ \ C(y) = t \quad \Leftrightarrow \quad \exists y',\ \ y' \text{ satisfies (4)},\ \ C(y') = t + b - 1.$$

1. Assume $y \in \{0, 1\}^n$ satisfies (3). Then clearly $y' = (y_1, \dots, y_n, 1, \dots, 1) \in \{0, 1\}^{n+b-1}$ satisfies (4), and $C(y') = C(y) + b - 1$ as required.
2. Assume $y' \in \{0, 1\}^{n+b-1}$ satisfies (4). We first construct $y'' \in \{0, 1\}^{n+b-1}$ that also satisfies (4), such that $y''_s = 1$ for $s = n+1, \dots, n+b-1$, and $C(y'') = C(y')$. The vector y'' is constructed by the following procedure.

   $y'' \leftarrow y'$
   while $\exists s,\ n + 1 \leq s \leq n + b - 1,\ y''_s = 0$ **do**
   Find $1 \leq j \leq n,\ y''_j = 1$
   $y''_s \leftarrow 1;\ \ y''_j \leftarrow 0$
   end-while

   Note that the "Find" always succeeds whenever there exists $y''_s = 0$, since $C(y') \geq b$. Note also that $C(y'') = C(y')$ throughout the procedure.

   We prove that y'' satisfies (4) by induction on the loop. Initially $y'' = y'$ satisfies (4) by assumption. For the inductive step, consider a specific modification of y'', at indices j and s. By construction the variable x_s appears in all the inequalities, thus it appears in any inequality that x_j appears in. Therefore swapping the values of y''_s and y''_j does not violate any inequality.

   The vector y is obtained by setting $y_j = y''_j$ for $j = 1, \dots, n$. Clearly, this y satisfies (3), and $C(y) = C(y') - b + 1$ as required. □

Proposition 3.2 leads us to the following conjecture, which extends conjectures from [Hoc2] and [BE].

CONJECTURE 3.3. *Unless $P = NP$, there is no polynomial-time approximation algorithm for ILP(k, b) which, for fixed k and b, has an approximation ratio less than $k - b + 1$.*

3.2. *The Ratio Between the Integral and Fractional Optima.* In this section we present a family of *ILP*(k, b) problem instances on which the gap between the optimal integral and fractional solutions is "large." We show that on this family of instances the ratio C^{opt}/C^* gets arbitrarily close to $k - b + 1$. This is an extension of an example given in [Hoc3] for the case $b = 1$.

Consider the following set of inequalities, denoted by A. For some t, the number of variables is $n = b - 1 + t$, and there are $m = \binom{t}{k-b+1}$ inequalities. Variables

$1, \dots, b-1$ appear in all m inequalities. Each inequality contains one of the possible choices of $k-b+1$ variables out of variables $b, \dots, n$, and the right-hand side is "$\geq b$."

The optimal cost of $LP(k, b)$ is $C^*(A) = b - 1 + t/(k-b+1)$. A possible optimal solution is

$$x_j^* = \begin{cases} 1, & j = 1, \dots, b-1, \\ \dfrac{1}{k-b+1}, & j = b, \dots, n. \end{cases}$$

There exists an optimal solution x^{opt} that contains variables $1, \dots, b-1$ (see the proof of Proposition 3.2). After removing these variables, the problem becomes $ILP(k-b+1, 1)$ with $m = \binom{t}{k-b+1}$ inequalities. A simple counting argument shows that the minimal number of variables that cover all the inequalities is $t-(k-b+1)+1 = t-k+b$, and that *any* choice of $t-k+b$ variables covers all the inequalities. Therefore the optimal integral cost is $C^{\text{opt}} = (b-1)+t-k+b = t-k+2b-1$, so

$$\frac{C^{\text{opt}}(A)}{C^*(A)} = \frac{t-k+2b-1}{b-1+t/(k-b+1)},$$

and the ratio tends to $k-b+1$ with t.

This example shows that any approximation algorithm B for the $ILP(k, b)$ problem has a worst-case *fractional* approximation ratio $R_B^* \geq k-b+1$.

4. An Inequality Concerning Sums of Independent Bernoulli Random Variables. The following theorem is a key tool in the analysis of our randomized algorithm, which appears in Section 5.

Let $\{x\} = x - [x] = x(\text{mod}\, 1)$, let $\mathbb{E}$ denote expectation, and let $X \sim B(p)$ denote a Bernoulli random variable X with a distribution rule of

$$\mathbb{P}(X = 1) = p, \qquad \mathbb{P}(X = 0) = 1 - p.$$

THEOREM 4.1. *There exists a constant $q > 0$ such that if $X_i \sim B(p_i)$, $i = 1, \dots, n$, are independent Bernoulli random variables, then putting $E = \sum_{i=1}^n p_i = \mathbb{E}(\sum_{i=1}^n X_i)$,*

$$\mathbb{P}\left(\sum_{i=1}^n X_i \leq E\right) \geq q(1 - \{E\}).$$

Note that for every E and $n \geq E$ there are p_i's with $\sum_{i=1}^n p_i = E$ and $\mathbb{P}(\sum_{i=1}^n X_i \leq E) = 1 - \{E\}$. Indeed, take $p_1 = \cdots = p_{[E]} = 1$, $p_{[E]+1} = \{E\}$, and $p_{[E]+2} = \cdots = p_n = 0$.

PROOF. Let X_i, $i = 1, \dots, n$, be independent Bernoulli random variables with $\mathbb{P}(X_i = 1) = p_i$, $\mathbb{P}(X_i = 0) = 1 - p_i$. If $p_i \in \{0, 1\}$ for all i, then the claim is trivial, so assume otherwise. Put

$$Z_i = \frac{X_i - p_i}{(\sum_{i=1}^n p_i(1-p_i))^{1/2}}, \qquad i = 1, \dots, n.$$

Then

$$\mathbb{E}Z_i = 0, \qquad \sum_{i=1}^{n} \sigma^2(Z_i) = 1,$$

and

$$\Gamma = \sum_{i=1}^{n} \mathbb{E}|Z_i|^3 = \frac{\sum_{i=1}^{n} p_i(1-p_i)((1-p_i)^2 + p_i^2)}{(\sum_{i=1}^{n} p_i(1-p_i))^{3/2}} \le \frac{1}{(\sum_{i=1}^{n} p_i(1-p_i))^{1/2}}.$$

By the Berry–Esseen theorem (a quantitative version of the Central Limit Theorem, see p. 225 of [Chu]),

$$\sup_x \left| \mathbb{P}\left(\sum_{i=1}^{n} Z_i \le x\right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \right| \le A\Gamma$$

for some absolute constant A (≥ 1). In particular, for $x = 0$,

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \le \sum_{i=1}^{n} p_i\right) = \mathbb{P}\left(\sum_{i=1}^{n} Z_i \le 0\right) \ge \tfrac{1}{2} - A\Gamma$$

and it follows that, as long as $\sum_{i=1}^{n} p_i(1-p_i)$ is larger than $16A^2$,

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \le \sum_{i=1}^{n} p_i\right) > \tfrac{1}{4}.$$

So we may assume $\sum_{i=1}^{n} p_i(1-p_i) \le 16A^2$ and, in particular,

$$\sum_{p_i \ge 1/2} (1 - p_i) \le 32A^2 \tag{5}$$

and

$$\sum_{p_i < 1/2} p_i \le 32A^2.$$

Put $t = 1 - \{\sum_{i=1}^{n} p_i\}$. By increasing some of the p_i's without changing $[\sum_{i=1}^{n} p_i]$ we may assume $t < \frac{1}{2}$. We may also assume that $p_1 \ge p_2 \ge \cdots \ge p_n$. Let $k_0 = 0$ and let k_1 be the first index (if it exists) such that $p_{k_1} \ge \frac{1}{2}$ and

$$k_1 - 1 < \sum_{i=1}^{k_1} p_i < k_1 - 1 + \frac{1-t}{64A^2}.$$

In a similar manner define indices k_j inductively. Let $k_j > k_{j-1}$ be the first index, if it exists, with $p_{k_j} \ge \frac{1}{2}$ and

$$k_j - k_{j-1} - 1 < \sum_{i=k_{j-1}+1}^{k_j} p_i < k_j - k_{j-1} - 1 + \frac{1-t}{64A^2}. \tag{6}$$

Let k_m be the last such index to exist. Then

$$k_m - m < \sum_{i=1}^{m} \sum_{i=k_{j-1}+1}^{k_j} p_i < k_m - m + \frac{m(1-t)}{64A^2} \tag{7}$$

and in particular, by negating the last inequality,

$$\sum_{j=1}^{m} \sum_{i=k_{j-1}+1}^{k_j} (1 - p_i) > m - \frac{m(1-t)}{64A^2}\,.$$

Using (5) it follows from the last inequality that

$$32A^2 + 1 > m.$$

Going back to (7) we get

$$k_m - m \le \sum_{j=1}^{m} \sum_{i=k_{j-1}+1}^{k_j} p_i < k_m - m + \frac{3(1-t)}{4},$$

so that

$$\frac{1-t}{4} < \left\{ \sum_{i=k_m+1}^{n} p_i \right\} \le 1 - t. \tag{8}$$

There can be three reasons why k_{m+1} does not exist:

(i) $k_m = n$ (e.g., all p_i's are $\ge \frac{1}{2}$),
(ii) $\sum_{i=k_m+1}^{n} p_i \ge n - k_m - 1 + (1-t)/64A^2$ (and then $\sum_{i=k_m+1}^{n} p_i \le n - k_m - t$),
(iii) for some $n > \ell > k_m$, both

$$\sum_{i=k_m+1}^{\ell} p_i \ge \ell - k_m - 1 + \frac{1-t}{64A^2}$$

and either $p_{\ell+1} \le \frac{1}{2}$ or $\frac{1}{2} < p_{\ell+1} \le \ell - k_m - \sum_{i=k_m+1}^{\ell} p_i \le 1 - (1-t)/64A^2$.

We examine each of the three cases. For each $1 \le j \le m$, using (6),

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=k_{j-1}+1}^{k_j} X_i \le \sum_{i=k_{j-1}+1}^{k_j} p_i \right) &= \mathbb{P}\left(\sum_{i=k_{j-1}+1}^{k_j} (1 - X_i) \ge 1 \right) \\
&= 1 - \mathbb{P}\left(\sum_{i=k_{j-1}+1}^{k_j} (1 - X_i) = 0 \right) \\
&= 1 - \mathbb{P}(X_i = 1, \forall k_{j-1} + 1 \le i \le k_j) \\
&= 1 - \prod_{i=k_{j-1}+1}^{k_j} p_i \ge 1 - \exp\left(- \sum_{i=k_{j-1}+1}^{k_j} (1 - p_i) \right) \\
&\ge 1 - \exp\left(- \left(1 - \frac{1-t}{64A^2} \right) \right) \ge 1 - e^{-1/2}
\end{aligned}$$

and

$$\mathbb{P}\left(\sum_{i=1}^{k_m} X_i \le \sum_{i=1}^{k_m} p_i\right) \ge (1 - e^{-1/2})^m \ge (1 - e^{-1/2})^{32A^2+1},$$

which proves the theorem if case (i) occurs. If case (ii) occurs, then in a similar manner

$$\mathbb{P}\left(\sum_{i=k_m+1}^{n} X_i \le \sum_{i=k_m+1}^{n} p_i\right) \ge 1 - \exp\left(-\sum_{i=k_m+1}^{n} (1 - p_i)\right) \ge 1 - e^{-t}$$

and

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \le \sum_{i=1}^{n} p_i\right) \ge (1 - e^{-1/2})^{32A^2+1}(1 - e^{-t}).$$

If case (iii) occurs, then $p_i < 1 - 1/128A^2$ for $i > \ell$ and thus

$$\sum_{i=\ell+1}^{n} p_i \le 16A^2 \cdot 128A^2 = 2048A^2.$$

So either

$$\ell - k_m - \sum_{i=k_m+1}^{\ell} p_i \ge t,$$

in which case

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} X_i \le \sum_{i=1}^{n} p_i\right) &\ge (1 - e^{-1/2})^{32A^2+1}\mathbb{P}\left(\sum_{i=k_m+1}^{\ell} X_i \le \sum_{i=k_m+1}^{\ell} p_i\right) \\
&\quad \cdot \mathbb{P}\left(\sum_{i=\ell+1}^{n} X_i \le \sum_{i=\ell+1}^{n} p_i\right) \\
&\ge (1 - e^{-1/2})^{32A^2+1}(1 - e^{-t})\mathbb{P}\left(\sum_{i=\ell+1}^{n} X_i = 0\right) \\
&= (1 - e^{-1/2})^{32A^2+1}(1 - e^{-t}) \prod_{i=\ell+1}^{n} (1 - p_i) \\
&\ge (1 - e^{-1/2})^{32A^2+1}(1 - e^{-t}) \exp\left(-B \sum_{i=\ell+1}^{n} p_i\right) \\
&\ge (1 - e^{-1/2})^{32A^2+1}(1 - e^{-t})e^{-2048A^4 B}
\end{aligned}$$

(here $B = 128A^2 \ln(128A^2)$ so that $1 - x \ge e^{-Bx}$ for $0 < x < 1 - 1/128A^2$)

or

$$\ell - k_m - \sum_{i=k_m+1}^{\ell} p_i < t,$$

in which case necessarily

$$\sum_{i=k_m+1}^{n} p_i > \ell - k_m$$

and

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} X_i \le \sum_{i=1}^{n} p_i\right) &\ge (1-e^{-1/2})^{32A^2+1}\mathbb{P}\left(\sum_{i=k_m+1}^{n} X_i \le \sum_{i=k_m+1}^{n} p_i\right)\\
&\ge (1-e^{-1/2})^{32A^2+1}\mathbb{P}\left(\sum_{i=k_m+1}^{n} X_i = \ell - k_m\right)\\
&\ge (1-e^{-1/2})^{32A^2+1}\mathbb{P}(X_i = 1,\ i = k_m+1,\dots,\ell\\
&\qquad\qquad \&\, X_i = 0,\ i = \ell+1,\dots,n)\\
&\ge (1-e^{-1/2})^{32A^2+1}\prod_{i=k_m+1}^{\ell} p_i \prod_{i=\ell+1}^{n}(1-p_i)\\
&\ge (1-e^{-1/2})^{32A^2+1}\exp\left(-B\sum_{i=k_m+1}^{\ell}(1-p_i)\right)\\
&\quad\cdot \exp\left(-B\sum_{i=\ell+1}^{n} p_i\right)\\
&\ge (1-e^{-1/2})^{32A^2+1}e^{-B(2048A^4+1)}. \qquad \square
\end{aligned}$$

We would now like to state two corollaries; the first is stated to emphasize the amazement that we find in the theorem above, the second one, suggested to us by Uri Feige, may prove to be a more useful form of the theorem.

COROLLARY 4.2. *For $0 < \alpha < 1$, let*

$$f(\alpha) = \inf \mathbb{P}\left(\sum_{i=1}^{n} X_i \le \alpha n\right),$$

where the X_i's are as in the statement of Theorem 4.1 and the inf *is taken over all n and all $p_1, \dots, p_n$ with $\sum_{i=1}^{n} p_i = \alpha n$. Then $f(\alpha) > 0$ if and only if α is rational.*

PROOF. ($\Leftarrow$) Assume α is rational. Then the sequence $\{\alpha n\}$, $n = 1, 2, \dots$, has a finite set of values, so let $\varepsilon = \max_n(\{\alpha n\}) < 1$. Let n and $p_1, \dots, p_n$ be as before. By Theorem 4.1 we get

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \le \alpha n\right) \ge q(1 - \{\alpha n\}) \ge q(1-\varepsilon) > 0.$$

($\Rightarrow$) Assume α is irrational. For any n let $k = \lfloor \alpha n \rfloor$ and $\beta = \{\alpha n\}$. Let the probabilities be $p_1 = \cdots = p_k = 1$, $p_{k+1} = \beta$, and $p_{k+2} = \cdots = p_n = 0$. Then

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \leq \alpha n\right) = \mathbb{P}(X_{k+1} = 0) = 1 - \beta = 1 - \{\alpha n\}.$$

Since α is irrational the infimum of the last expression is 0. □

COROLLARY 4.3. *There exists a constant $q > 0$ such that for all $0 < \varepsilon \leq 1$ if $X_i \sim B(p_i)$, $i = 1, \ldots, n$ are independent Bernoulli random variables, then putting $E = \sum_{i=1}^{n} p_i = \mathbb{E}(\sum_{i=1}^{n} X_i)$,*

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \leq E + \varepsilon\right) \geq q\varepsilon.$$

PROOF. We separate the discussion into two cases:

1. $\{E\} \leq 1 - \varepsilon$. Then using Theorem 4.1,

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \leq E + \varepsilon\right) \geq \mathbb{P}\left(\sum_{i=1}^{n} X_i \leq E\right) \geq q(1 - \{E\}) \geq q\varepsilon.$$

2. $\{E\} > 1 - \varepsilon$. Then $E = \lceil E \rceil - 1 + \{E\} > \lceil E \rceil - \varepsilon$, so $\lceil E \rceil < E + \varepsilon$. Define an auxiliary random variable $Y \sim B(\lceil E \rceil - E)$, so $\mathbb{E}(\sum_{i=1}^{n} X_i + Y) = \lceil E \rceil$. We get

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{n} X_i \leq E + \varepsilon\right) &\geq \mathbb{P}\left(\sum_{i=1}^{n} X_i + Y \leq E + \varepsilon\right)\\ &\geq \mathbb{P}\left(\sum_{i=1}^{n} X_i + Y \leq \lceil E \rceil\right) \geq q(1 - \{\lceil E \rceil\}) = q \geq q\varepsilon. \quad \square\end{aligned}$$

5. The Randomized Algorithm

5.1. *Introduction.* In this section we present a randomized approximation algorithm for the *ILP*(k, b) problem. Our goal is to obtain an algorithm with an approximation ratio which is lower than the ratio of $k - b + 1$ we already have from deterministic algorithms [HH].

A randomized algorithm for the closely related general b-matching problem in hypergraphs has been proposed by Raghavan and Thompson [RT]. Their technique, called "randomized rounding with scaling," is based on first solving the fractional b-matching problem, and then using randomization to obtain an approximate solution for the integral problem. The analysis of their algorithm relies in two essential ways on the use of Chernoff bounds on sums of Bernoulli random variables [Che].

In our algorithm for *ILP*(k, b) we also use randomized rounding with scaling, but with a new analysis which avoids the use of Chernoff bounds. This is both necessary and advantageous for the following reasons. First, the analogue of one of the two applications of Chernoff bounds in [RT] is not possible for $ILP(k, b)$ (the source of the difficulty is mentioned in what follows). Second, the other application of a Chernoff bound yields an inferior approximation ratio in our case. Instead of the Chernoff bounds, our analysis hinges on the utilization of the k bound on edge cardinalities, and on the use of the new bound on the sum of Bernoulli random variables, Theorem 4.1.

5.2. *Randomized Rounding with Scaling.* The fundamental idea behind randomized rounding is quite simple. Let $x^* = (x_1^*, \ldots, x_n^*)$ be an optimal solution to an $LP(k, b)$ instance. Since $x_j^* \in [0, 1]$ we can define independent Bernoulli random variables:

$$Y_j \sim B(x_j^*) \qquad \text{for} \quad j = 1, \ldots, n.$$

The Y_j's are referred to as the rounded version of x^*. A natural idea is to take a random assignment for these variables as a candidate for an integral solution for the corresponding $ILP(k, b)$ instance. Indeed, consider a specific inequality

$$\sum_{j \in S} X_j \geq b_i \qquad \text{for} \quad |S| \leq k, \quad \text{and} \quad b_i \geq b. \tag{9}$$

Then the expected sum of the Y_j variables satisfies the constraint, since by linearity of expectation

$$\mathbb{E}\left(\sum_{j \in S} Y_j\right) = \sum_{j \in S} \mathbb{E}(Y_j) = \sum_{j \in S} x_j^* \geq b_i$$

and similarly the expected cost is optimal,

$$\mathbb{E}\left(\sum_{j=1}^{n} Y_j\right) \leq C^*.$$

The difficulty in utilizing this idea lies in the fact that there is a significant probability that the rounded Y_j values will not satisfy (9). Furthermore, it is not clear whether there is a nonzero probability that the randomized rounding will yield a solution in which *none* of the m constraints is violated, *and* which has a low cost *simultaneously*.

More precisely, we require that for a (random) integral solution to be called *satisfactory*, with respect to some cost μ', two events must occur:

$$\mathcal{E}_1 = \{\text{no constraint is violated}\}$$

and

$$\mathcal{E}_2(\mu') = \left\{\sum_{j=1}^{n} Y_j \leq \mu'\right\}.$$

In other words, we define the event of finding a satisfactory solution by

$$\mathcal{E} = \mathcal{E}_1 \wedge \mathcal{E}_2(\mu'). \tag{10}$$

Our task is to find a method of choosing the random variables Y_j, and a suitable cost bound μ', such that for probabilities γ and ε we have

- $\mathbb{P}(\mathcal{E}_1) \geq 1 - \gamma$,
- $\mathbb{P}(\mathcal{E}_2(\mu')) \geq \varepsilon$.

Since the events $\mathcal{E}_1$ and $\mathcal{E}_2(\mu')$ are not independent, we cannot preclude the possibility that these two events are largely disjoint. We are only guaranteed that $\mathbb{P}(\mathcal{E}) \geq \mathbb{P}(\mathcal{E}_1) + \mathbb{P}(\mathcal{E}_2(\mu')) - 1$. Therefore we should ensure that

$$(1 - \gamma) + \varepsilon \geq 1 + \theta > 1. \tag{11}$$

Then θ is a lower bound on the probability of finding a satisfactory solution. Therefore the expected number of trials until a satisfactory solution is actually found is $1/\theta$. To claim that we have a *randomized approximation algorithm*, we need to show that $1/\theta$ is polynomial, and then the approximation ratio is determined by the cost μ'.

The first step we take is to ensure that $\mathcal{E}_1$ occurs with probability $1 - \gamma$. To achieve this, the random variables need to be scaled. By scaling, we mean that the x_j^* probabilities are multiplied by some factor $\delta > 1$. The resultant values are then used to define the Y_j random variables. The factor δ is chosen so the probability of event $\mathcal{E}_1$ is at least $1 - \gamma$ for some "safety probability" γ.

REMARK. Multiplying a probability x_j^* with $\delta > 1$ may yield a value larger than 1, in which case we take 1 as the result. Therefore, we cannot claim that the expected sum of scaled variables in a specific inequality is at least δb_i (consider the extreme case of $x_j^* \in \{0, 1\}$ for all j; then the scaled variables have the same probabilities as the unscaled ones). This is the cause of our inability to apply a Chernoff bound at this point in the analysis. This difficulty does not arise in Raghavan and Thompson's analysis for the b-matching problem, since the problem there is a maximization problem, and their scaling is done with a factor $\delta < 1$.

Trivially, the expected cost of the scaled variables is at most δC^*. However, it turns out that for the δ that we use the expected cost is in fact at most $\delta C^* - 1$. This observation enables us to claim that the algorithm can (with a high probability) find a solution with cost of at most δC^*, i.e., one higher than the bound on the expected cost. In other words, our analysis shows that, for some $\varepsilon > 0$,

$$\mathbb{P}(\mathcal{E}_2(\delta C^*)) \geq \varepsilon.$$

We need to show that for these γ and ε, (11) holds, and that $1/\theta$ is indeed polynomial (in fact it will be constant). We can then claim that we have a randomized algorithm RND that finds a satisfactory solution, i.e., a feasible solution with a cost of at most δC^*. The approximation ratio would then be

$$R_{\text{RND}} \leq R^*_{\text{RND}} \leq \delta.$$

5.3. *The Main Results*

THEOREM 5.1. *There exists a polynomial-time randomized approximation algorithm* RND *for the ILP(k, b) problem with $4 \le b \le k_{\min} - 2$, where $k_{\min}$ is the minimal number of variables in any inequality. Algorithm* RND *has the following approximation ratio, using q, the universal constant guaranteed by Theorem* 4.1, *and $M = (q/2)(k - b + 1)^{k-b+1}$*:

$$R^*_{\mathrm{RND}} \le \begin{cases} k - b, & m \le M, \\ (k - b + 1)\left(1 - \left(\dfrac{q}{2m}\right)^{1/(k-b+1)}\right), & m > M. \end{cases}$$

REMARK. This does not contradict Hochbaum's conjecture (3.3) since for $k - b$ fixed the bound on the approximation ratio tends to $k - b + 1$ with m.

The proof of this result is presented in two stages:

1. Satisfying the Constraints (Section 5.4). For any probability γ we obtain a lower bound on the scaling factor δ that will guarantee that the probability of the event $\mathcal{E}_1$ is at least $1 - \gamma$. This is achieved by direct analysis of the event. We then show that δ also guarantees that the expected cost is at most $\delta C^* - 1$.
2. The Algorithm (Section 5.5). We find the choice of probabilities (γ and ε) that will yield a cost of at most δC^*, while satisfying (11). We conclude by showing that θ, the probability of success, ensures a polynomial-time algorithm.

5.4. *Satisfying the Constraints.* Let $x^* = (x_1^*, \ldots, x_n^*)$ be an optimal solution to $LP(k, b)$. Consider an inequality on $\ell \le k$ variables. Assume without loss of generality that it is

$$\sum_{j=1}^{\ell} x_j \ge b_i. \tag{12}$$

NOTATION. For $a \ge 0$ let $\langle a \rangle = \min\{a, 1\}$. For $\delta > 1$ define scaled Bernoulli random variables

$$Y_j \sim B(\langle \delta x_j^* \rangle)$$

and their sum

$$Y = \sum_{j=1}^{\ell} Y_j.$$

Let $\alpha(\delta) = \mathbb{P}(Y < b_i)$, the probability that (12) is violated.

PROPOSITION 5.2. *For $1 \le b_i \le k_{\min} - 2$, and a safety probability γ, if*

$$\delta \ge \max\left\{k - b_i, (k - b_i + 1)\left(1 - \left(\frac{\gamma}{m}\right)^{1/(k-b_i+1)}\right)\right\},$$

then $\alpha(\delta) \le \gamma/m$.

PROOF. Define A_r for $r = 0, \ldots, b_i - 1$ as the event that exactly r variables from $I = \{1, \ldots, \ell\}$ are chosen to be 1, i.e., $A_r = \{Y = r\}$. From the definition of Y_j,

$$\mathbb{P}(A_r) = \sum_{\substack{S \subseteq I \\ |S|=r}} \prod_{j \in S} \langle \delta x_j^* \rangle \prod_{j \in I \setminus S} (1 - \langle \delta x_j^* \rangle). \tag{13}$$

The events are disjoint, therefore

$$\alpha(\delta) = \sum_{r=0}^{b_i - 1} \mathbb{P}(A_r).$$

By assumption, $b_i \le k_{\min} - 2 \le \ell - 2$. Assume also that $b_i \ge 2$. Define the set of "large values" in x^* as $B^* = \{j | x_j^* \ge 1/(\ell - b_i)\}$. Then by the fact that x^* obeys (12), and by Lemma 2.1 applied with $t = 2$, $|B^*| \ge b_i - 1$. For each of the $j \in B^*$, $\langle \delta x_j^* \rangle = 1$ (since by assumption $\delta \ge k - b_i \ge \ell - b_i$).

Consider the events A_r for $r = 0, \ldots, b_i - 2$. Since $r < b_i - 1$, from the pigeonhole principle we get that in every term of (13) there is at least one $j \in I \setminus S$ such that $(1 - \langle \delta x_j^* \rangle) = 0$. Therefore $\mathbb{P}(A_r) = 0$ for $r = 0, \ldots, b_i - 2$.

We now consider the event $A_{b_i - 1}$. If $|B^*| \ge b_i$, then by a similar argument $\delta = k - b_i$ suffices to get $\mathbb{P}(A_{b_i - 1}) = 0$ and we are done. Otherwise, $|B^*| = b_i - 1$. Then the only nonzero term in (13) is the one based on the set $S = B^*$. Therefore

$$\alpha(\delta) = \mathbb{P}(A_{b_i - 1}) = \prod_{j \in B^*} \langle \delta x_j^* \rangle \prod_{j \in I \setminus B^*} (1 - \langle \delta x_j^* \rangle) = \prod_{j \in I \setminus B^*} (1 - \langle \delta x_j^* \rangle). \tag{14}$$

Note the following facts:

- $|I \setminus B^*| = \ell - b_i + 1$,
- $\sum_{j \in I \setminus B^*} x_j^* \ge b_i - \sum_{j \in B^*} x_j^* \ge 1$.

Under these conditions, by Lemma 2.2, if $\alpha(\delta) > 0$ then it gets its maximal value as a function of x^* when all x_j^*'s are equal, and $x_j^* = 1/(\ell - b_i + 1)$ for all $j \in I \setminus B^*$. Therefore

$$\alpha(\delta) \le \left(1 - \frac{\delta}{\ell - b_i + 1}\right)^{\ell - b_i + 1} \le \left(1 - \frac{\delta}{k - b_i + 1}\right)^{k - b_i + 1}, \tag{15}$$

since, for $a > 0$ and $t \ge a$, the function $(1 - a/t)^t$ is monotonously increasing with t. In our case $\ell - b_i + 1 \ge \delta$, otherwise $\alpha(\delta) = 0$. Therefore taking δ as specified in the proposition ensures that $\alpha(\delta) \le \gamma/m$.

The only remaining case is $b_i = 1$. In this case the only event is A_0, and then we do not need to assume that $\delta \ge k - b_i$. The analysis can start directly at (14), with $B^* = \varnothing$. This completes the proof of Proposition 5.2. □

REMARKS.

- For the special case $b_i = 2$ we can simplify the proof. Specifically, we can weaken the requirement that $\delta \ge k - b_i$, and replace it by

$$\delta \ge \frac{k}{2}.$$

Instead of using Lemma 2.1, we can then use the observation that there exists at least one j with $x_j^* \geq 2/\ell$, and the analysis holds without the premise $b_i \leq k_{\min} - 2$.

• If $k - b_i + 1$ is not fixed, then we can take the limit in (15) and get $\alpha(\delta) < e^{-\delta}$. To ensure $\alpha(\delta) \leq \gamma/m$ we need to take $\delta \geq \ln(m/\gamma)$, which leads to a logarithmic approximation ratio.

COROLLARY 5.3. *For $1 \leq b \leq k_{\min} - 2$, and a safety probability γ, if*

$$\delta \geq \max\left\{k - b, (k - b + 1)\left(1 - \left(\frac{\gamma}{m}\right)^{1/(k-b+1)}\right)\right\},$$

then $\mathbb{P}(\mathcal{E}_1) \geq 1 - \gamma$.

PROOF. We need δ to satisfy the requirements of Proposition 5.2 for every inequality i, i.e., we must take the value from the inequality on which the expression in Proposition 5.2 is largest. This maximum is obtained for inequalities where $b_i = b$, the smallest right-hand side value. To prove this, it suffices to show that, for every $0 \leq \tau \leq 1$, if $b < b_i$, then

$$(k - b + 1)(1 - \tau^{1/(k-b+1)}) \geq (k - b_i + 1)(1 - \tau^{1/(k-b_i+1)}),$$

and after rearranging,

$$b_i - b \geq (k - b + 1)\tau^{1/(k-b+1)} - (k - b_i + 1)\tau^{1/(k-b_i+1)}.$$

The last inequality holds since the right-hand side is a monotonous increasing function of τ in the domain $[0, 1]$ which has a value of $b_i - b$ when $\tau = 1$. □

PROPOSITION 5.4. *If $4 \leq b \leq k - 2$, and δ is as in Corollary 5.3, then $\mathbb{E}(Y) \leq \delta C^* - 1$.*

PROOF. Let S_i be the set of variables in inequality i, $|S_i| = \ell \leq k$. Let $D_i \subset S_i$ be the set of $b - 1$ variables $j \in S_i$ with largest fractional values x_j^*. Since $\sum_{j \in S_i} x_j^* \geq b_i \geq b$, the average value for x_j^* is at least $b/\ell \geq b/k$. Therefore summing the $b - 1$ largest values we get

$$\sum_{j \in D_i} x_j^* \geq (b - 1)\frac{b}{k}.$$

Now let D be the set of $b - 1$ variables $j \in \{1, \ldots, n\}$ with largest fractional values x_j^*. Then for any i, $\sum_{j \in D} x_j^* \geq \sum_{j \in D_i} x_j^*$, so

$$\sum_{j \in D} x_j^* \geq (b - 1)\frac{b}{k}. \tag{16}$$

Using the definitions of Y and $\langle\cdot\rangle$ we get

$$\mathbb{E}(Y) = \sum_{j=1}^{n} \langle \delta x_j^* \rangle = \delta C^* - \sum_{j=1}^{n} \max\{\delta x_j^* - 1, 0\}.$$

Since by assumption $\delta \geq k - b$, Lemma 2.1 guarantees that, for all $j \in D$, $\langle \delta x_j^* \rangle = 1$. Therefore

$$\mathbb{E}(Y) \leq \delta C^* - \sum_{j \in D} (\delta x_j^* - 1) = \delta C^* - \left(\delta \sum_{j \in D} x_j^* - |D| \right).$$

By (16) and using the assumptions that $b \geq 4$ and $k - b \geq 2$ we get

$$\mathbb{E}(Y) \leq \delta C^* - \left((k-b)(b-1)\frac{b}{k} - (b-1) \right) \leq \delta C^* - 1. \qquad \square$$

5.5. *The Algorithm.* Now the groundwork is prepared for proceeding to prove our main result, Theorem 5.1. We need to combine the results of Corollary 5.3 and Proposition 5.4 regarding the properties of the scaling factor δ, and then to use Corollary 4.3. With these tools we can now build the randomized algorithm RND and prove its claimed approximation ratio, and polynomial-time complexity.

Using the universal constant q from Theorem 4.1, we define

$$\varepsilon = q, \tag{17}$$

$$\gamma = \frac{q}{2}. \tag{18}$$

Hence

$$\theta = (1 - \gamma) + \varepsilon - 1 = \frac{\varepsilon}{2}. \tag{19}$$

We use this γ to calculate δ according to Corollary 5.3, i.e.,

$$\delta = \max\left\{ k - b, (k - b + 1)\left(1 - \left(\frac{\gamma}{m}\right)^{1/(k-b+1)} \right) \right\}. \tag{20}$$

With this δ denote the scaled probabilities and their sum by

$$z_j = \langle \delta x_j^* \rangle, \qquad \text{for} \quad j = 1, \ldots, n,$$

$$\mu = \sum_{j=1}^{n} z_j.$$

Note that, by Proposition 5.4, $\mu \leq \delta C^* - 1$.

Now define the random variables using the modified probabilities:

$$Y_j \sim B(z_j) \qquad \text{for} \quad j = 1, \ldots, n.$$

The algorithm RND flips n independent coins for the Y_j random variables, until event $\mathcal{E}$ occurs (for a cost of δC^*).

LEMMA 5.5. *The algorithm* RND *enjoys the following properties*:

1. $\mathbb{P}(\mathcal{E}_1) \geq 1 - \gamma$,

1. Calculate γ by (18) and δ by (20) for k, b and m.
2. Linear Programming : find the optimal fractional solution $x^* = (x_1^*, \ldots, x_n^*)$.
3. Calculate the scaled probability vector $z = (z_1, \ldots, z_n)$ by $z_j = \langle \delta x_j^* \rangle$.
4. Randomize : flip n independent coins with probabilities z_j until the event $\mathcal{E}$ (10) occurs.

Fig. 1. Algorithm RND.

2. $\mathbb{P}(\mathcal{E}_2(\delta C^*)) \geq \varepsilon$,
3. $\mathbb{P}(\mathcal{E}) \geq \theta = \varepsilon/2$,
4. $R^*_{\mathrm{RND}} \leq \delta$.

PROOF. By the definition of δ, Corollary 5.3 holds, thus

$$\mathbb{P}(\mathcal{E}_1) \geq 1 - \gamma.$$

Using Corollary 4.3, along with Proposition 5.4 we get

$$\mathbb{P}(\mathcal{E}_2(\delta C^*)) = \mathbb{P}(Y \leq \delta C^*) \geq \mathbb{P}(Y \leq \mu + 1) \geq q \cdot 1 = \varepsilon.$$

We conclude by (19) that the probability of finding a satisfactory solution is

$$\mathbb{P}(\mathcal{E}) \geq \theta = \frac{\varepsilon}{2}.$$

Since $\mathbb{P}(\mathcal{E}) \geq \theta > 0$, algorithm RND will find a solution $\bar{y} \in \{0, 1\}^n$ such that $C_{\mathrm{RND}} = \sum_{j=1}^n \bar{y}_j \leq \delta C^*$. Therefore the approximation ratio is

$$R^*_{\mathrm{RND}} \leq \frac{C_{\mathrm{RND}}}{C^*} \leq \delta. \qquad \square$$

We need to show that θ is polynomial, which is trivial since $\theta = q/2$, a constant. Therefore the time complexity is dominated by the Linear Programming phase, which is polynomial.

By a straightforward computation we obtain that, for $M = (q/2)(k - b + 1)^{k-b+1}$, the value of δ depends on m by

$$\delta = \begin{cases} k - b, & m \leq M, \\ (k - b + 1)\left(1 - \left(\dfrac{q}{2m}\right)^{1/(k-b+1)}\right), & m > M. \end{cases}$$

This concludes the proof of Theorem 5.1.

5.6. *Final Remarks.* 1. We have not computed the exact value of the constant q. It is bounded very roughly from below by

$$q \geq 2^{-2^{50}}.$$

A tighter bound can probably be found by careful analysis. Nevertheless, it seems that q is too small to be of practical value.

2. It is not beneficial to reduce $ILP(k, b)$ to $ILP(k - b + 1, 1)$ using Proposition 3.1. The reduction increases the number of equations by a (potentially) nonpolynomial factor, causing the time complexity to increase significantly in the Linear Programming phase. Moreover, the increased number of inequalities increases the bound on the approximation ratio R^*_{RND}.

An important exception is the case $b = k - 1$, which is not covered by Theorem 5.1. Using Proposition 3.1 in this case increases the number of equations by a factor of only $O(k^2)$, and reduces the problem into an instance of $ILP(2, 1)$. However, this is the Vertex Cover problem, for which better deterministic algorithms exist (e.g., [BE]).

3. The requirement $b \geq 4$ is not a real limitation. We can transform an $ILP(k, b)$ instance with $b \leq 3$ into an instance of $ILP(k+4-b, 4)$ by adding $4-b$ new variables that appear in all the inequalities, and increasing the right-hand side to 4, as in Proposition 3.2. This has no adverse effect on the analysis since in all the expressions containing b, it appears only in the difference $k - b$, which is invariant under the transformation.

4. As presented, the algorithm has a small (though constant) probability of finding a satisfactory solution. Therefore derandomizing it would be a significant improvement in terms of its time complexity. Since the Raghavan–Thompson algorithm was successfully derandomized [R] (see also [BV]), we could hope to derandomize our algorithm as well. So far we have not succeeded in doing this, and it remains an open problem. The main difficulty seems to be derandomizing Theorem 4.1.

Acknowledgments. We are very grateful to Uri Feige for his helpful comments and suggestions. We thank Michel Goemans for providing a shorter proof for Lemma 2.2, and the anonymous referees, whose remarks helped us improve our presentation.

References

[ALM+] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and hardness of approximating problems. In *Proc. 33rd IEEE Symp. on Foundations of Computer Science*, pages 14–23, 1992.

[BE] R. Bar-Yehuda and S. Even. A local-ratio theorem for approximating the weighted vertex cover problem. *Ann. Discrete Math.*, 25:27–46, 1985.

[BKP1] J. Bar-Ilan, G. Kortsarz, and D. Peleg. Greedy approximation algorithms. Technical Report CS92–28, The Weizmann Institute of Science, Rehovot, 1992.

[BKP2] J. Bar-Ilan, G. Kortsarz, and D. Peleg. How to allocate network centers. *J. Algorithms*, 15:385–415, 1993.

[BV] D. Bertsimas and R. Vohra. Linear programming relaxations, approximation algorithms and randomization: a unified view of covering problems, Working Paper, Operations Research Center, MIT, Cambridge, MA, 1994.

[Che] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.*, 23:493–509, 1952.

[Chu] K. L. Chung. *A Course in Probability Theory*. Academic Press, New York, 1974.

[Chv] V. Chvatal. A greedy heuristic for the set-covering problem. *Math. Oper. Res.*, 4(3):233–235, 1979.

[D] G. Dobson. Worst case analysis of greedy heuristics for integer programming with nonnegative data. *Math. Oper. Res.*, 7:515–531, 1982.

[F] U. Feige. Personal communication, 1993.

[GJ] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, San Francisco, 1979.

[GW] T. Grossman and A. Wool. Computational experience with approximation algorithms for the set covering problem. *Euro. J. Operational Research*, To appear. See also Technical Report CS94-25, The Weizmann Institute of Science, Rehovot.

[HH] N. G. Hall and D. S. Hochbaum. A fast approximation algorithm for the multicovering problem. *Discrete Appl. Math.*, 15:35–40, 1986.

[Hoc1] D. S. Hochbaum. Approximation algorithms for the set covering and vertex cover problems. *SIAM J. Comput.*, 11(3):555–556, 1982.

[Hoc2] D. S. Hochbaum. Efficient bounds for the stable set, vertex cover and set packing problems. *Discrete Appl. Math.*, 6:243–254, 1983.

[Hoc3] D. S. Hochbaum. On the fractional solution to the set covering problem. *SIAM J. Algebraic Discrete Methods*, 4(2):221–222, 1983.

[HS] D. S. Hochbaum and D. B. Shmoys. A unified approach to approximation algorithms for bottleneck problems. *J. Assoc. Comput. Mach.*, 33(3):533–550, July 1986.

[J] D. S. Johnson. Approximation algorithms for combinatorial problems. *J. Comput. System Sci.*, 9:256–278, 1974.

[Kar] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[Kha] L. G. Khachiyan. A polynomial algorithm in linear programming. *Dokl. Akad. Nauk SSSR*, 244(S):1093–1096, 1979. Translated in *Soviet Math. Dokl.*, 20(1):191–194, 1979.

[L] L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Math.*, 13:383–390, 1975.

[NT] G. L. Nemhauser and L. E. Trotter. Vertex packing: Structural properties and algorithms. *Math. Programming*, 8:232–248, 1975.

[PSW] D. Peleg, G. Schechtman, and A. Wool. Approximating bounded 0–1 integer linear programs. In *Proc. 2nd Israel Symp. on Theory of Computing Systems*, pages 69–77, Netanya, 1993.

[PY] C. H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *J. Comput. System Sci.*, 43:425–440, 1991.

[R] P. Raghavan. Probabilistic construction of deterministic algorithms: approximating packing integer programs. *J. Comput. System Sci.*, 37:130–143, 1988.

[RT] P. Raghavan and C. D. Thompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7(4):365–374, 1987.

[Wol] L. A. Wolsey. An analysis of the greedy algorithm for the submodular set covering problem. *Combinatorica*, 2:385–393, 1982.

[Woo] A. Wool. Approximating bounded 0–1 integer linear programs. Master's thesis, Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot, 1992.