On the Integrality Gap of Binary Integer Programs with Gaussian Data

Sander Borst*, Daniel Dadush*,**, Sophie Huiberts**, and Samarth Tiwari*

Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands
{sander.borst,dadush,s.huiberts,samarth.tiwari}@cwi.nl

Abstract. For a binary integer program (IP) $\max c^\mathsf{T} x, Ax \leq b, x \in \{0,1\}^n$, where $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^n$ have independent Gaussian entries and the right-hand side $b \in \mathbb{R}^m$ satisfies that its negative coordinates have ℓ_2 norm at most $n/10$, we prove that the gap between the value of the linear programming relaxation and the IP is upper bounded by $\operatorname{poly}(m)(\log n)^2/n$ with probability at least $1 - 1/n^7 - 2^{-\operatorname{poly}(m)}$. Our results give a Gaussian analogue of the classical integrality gap result of Dyer and Frieze (Math. of O.R., 1989) in the case of random packing IPs. In constrast to the packing case, our integrality gap depends only polynomially on m instead of exponentially. By recent breakthrough work of Dey, Dubey and Molinaro (SODA, 2021), the bound on the integrality gap immediately implies that branch and bound requires $n^{\operatorname{poly}(m)}$ time on random Gaussian IPs with good probability, which is polynomial when the number of constraints m is fixed.

Keywords: Integer Programming · Integrality Gap · Branch and Bound

1 Introduction

Consider the following linear program with n variables and m constraints

$$\begin{aligned} \mathrm{val}_{\mathsf{LP}} = \max_x \ & \mathrm{val}(x) = c^\mathsf{T} x \\ \text{s.t. } & Ax \leq b \\ & x \in [0,1]^n \end{aligned} \tag{Primal LP}$$

Let $\mathrm{val}_{\mathsf{IP}}$ be the value of the same optimization problem with the additional restriction that x is integral, i.e., $x \in \{0,1\}^n$. We denote the integrality gap as $\mathsf{IPGAP} := \mathrm{val}_{\mathsf{LP}} - \mathrm{val}_{\mathsf{IP}}$. The integrality gap of integer linear programs forms an important measure for the complexity of solving said problem in a number of works on the average-case complexity of integer programming [1,3,4,5,7,11].

* This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement QIP–805241)

** This work was done while the author was participating in a program at the Simons Institute for the Theory of Computing.

So far, probabilistic analyses of the integrality gap have focussed on 0–1 packing IPs and the generalized assignment problem. In particular, the entries of $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$ in these problems are all non-negative, and the entries of b were assumed to scale linearly with n.

In this paper, we analyze the integrality gap of (Primal LP) under the assumption that the entries of A and c are all independent Gaussian $\mathcal{N}(0,1)$ distributed, and that the negative part of b is small: $\|b^-\|_2 \leq n/10$.

We prove that, with high probability, the integrality gap IPGAP is small, i.e., (Primal LP) admits a solution $x \in \{0,1\}^n$ with value close to the optimum.

Theorem 1. *There exists an absolute constant $C \geq 1$, such that, for $m \geq 1$, $n \geq Cm$, $b \in \mathbb{R}^m$ with $\|b^-\|_2 \leq n/10$, if* (Primal LP) *is sampled with independent $\mathcal{N}(0,1)$ entries in c and A, we have that*

$$\Pr\left(\mathsf{IPGAP} \geq 10^{15} \cdot t \cdot \frac{m^{2.5}(m+\log n)^2}{n}\right) \leq 4 \cdot \left(1 - \frac{1}{25}\right)^t + n^{-7},$$

for all $1 \leq t \leq \frac{n}{Cm^{2.5}(m+\log n)^2}$.

In the previous probabilistic analyses by [4,5,11], it is assumed that $b_i = \beta_i n$ for fixed $\beta_1, \ldots, \beta_m \in (0, 1/2)$ and the entries of (A, c) are independently distributed uniformly in the interval $[0,1]$. Those works prove a similar bound as above, except that in their results the dependence on m is exponential instead of polynomial. Namely, they require $n \geq 2^{O(m)}$ and the integrality gap scales like $2^{O(m)} \log^2 n/n$. Furthermore, the integrality gap in [5] also had a $O(1/\beta^m)$ dependence, where $\beta := \min_{i\in[m]} \beta_i$, whereas the integrality gap in Theorem 1 does not depend on the "shape" of b (other than requiring $\|b^-\|_2 \leq n/10$). Due a recent breakthrough [3], the integrality gap above also implies that branch and bound applied to the above IP produces a branching tree of size at most $n^{\mathrm{poly}(m)}$ with good probability.

In the rest of the introduction, we begin with an overview of the main techniques we use to prove Theorem 1, describing the similarities and differences with the analysis of Dyer and Frieze [5], and highlight several open problems. We continue by explaining the relation between the integrality gap and the complexity of branch and bound (subsection 1.2 below), and conclude with a discussion of related work.

1.1 Techniques

Our proof strategy follows along similar lines to that of Dyer and Frieze [5], which we now describe. In their strategy, one first solves an auxiliary LP $\max c^\mathsf{T} x, Ax \leq b - \epsilon 1_m$, for $\epsilon > 0$ small, to get its optimal solution x^*, which is both feasible and nearly optimal for the starting LP (proved by a simple scaling argument), together with its optimal dual solution $u^* \geq 0$ (see subsection 2.2 for the formulation of the dual). From here, they round down the fractional components of x^* to get a feasible IP solution $x' := \lfloor x^* \rfloor$. We note that the feasibility of

x' depends crucially on the packing structure of the LPs they work with, i.e., that A has non-negative entries (which does not hold in the Gaussian setting). Lastly, they construct a nearly optimal integer solution x'', by carefully choosing a subset of coordinates $T \subset \{i \in [n] : x'_i = 0\}$ of size $O(\mathrm{poly}(m) \log n)$, where they flip the coordinates of x' in T from 0 to 1 to get x''. The coordinates of T are chosen accordingly the following criteria. Firstly, the coordinates should be *very cheap* to flip, which is measured by the absolute value of their *reduced costs*. Namely, they enforce that $|c_i - A_{\cdot,i}^\mathsf{T} u^*| = O(\log n/n)$, $\forall i \in T$. Secondly, T is chosen to make the *excess slack* $\|A(x^* - x'')\|_\infty = 1/\mathrm{poly}(n)$, i.e., negligible. We note that guaranteeing the existence of T is highly non-trivial. Crucial to the analysis is that after conditioning on the exact value of x^* and u^*, the columns of $W := \begin{bmatrix} c^\mathsf{T} \\ A \end{bmatrix} \in \mathbb{R}^{(m+1)\times n}$ (the objective extended constraint matrix), indexed by $N_0 := \{i \in [n] : x^*_i = 0\}$, are independently distributed subject to having negative reduced cost, i.e., subject to $c_i - A_{\cdot,i}^\mathsf{T} u^* < 0$ for $i \in N_0$ (see Lemma 5). It is the large amount of left-over randomness in these columns that allowed Dyer and Frieze to show the existence of the subset T via a discrepancy argument (more on this below). Finally, given a suitable T, a simple sensitivity analysis is used to show the bound on the gap between $c^\mathsf{T} x''$ and the (Primal LP) value. This analysis uses the basic formula for the optimality gap between primal and dual solutions (see (Gap Formula) in subsection 2.2), and relies upon bounds on the size of the reduced costs of the flipped variables, the total excess slack and the norm of the dual optimal solution u^*.

As a first difference with the above strategy, we are able to work directly with the optimal solution x^* of the original LP without having to replace b by $b' := b - \epsilon 1_m$. The necessity of working with this more conservative feasible region in [5] is that flipping 0 coordinates of x' to 1 can only *decrease* $b - Ax'$. In particular, if the coordinates of $b - Ax' \geq 0$ are too small, it becomes difficult to find a set T that doesn't force x'' to be infeasible. By working with b' instead of b, they can insure that $b - Ax' \geq \epsilon 1_m$, which avoids this problem. In the Gaussian setting, it turns out that we have equal power to both increase and decrease the slack of $b - Ax'$, due to the fact that the Gaussian distribution is symmetric about 0. We are in fact able to simultaneously fix both the feasibility and optimality error of x', which gives us more flexibility. In particular, we will be able to use randomized rounding when we move from x^* to x', which will allow us to start with a smaller initial slack error than is achievable by simply rounding x^* down.

Our main quantitative improvement – the reduction from an exponential to a polynomial dependence in m – arises from two main sources. The first source of improvement is a substantially improved version of a discrepancy lemma of Dyer and Frieze [5, Lemma 3.4]. This lemma posits that for any large enough set of "suitably random" columns in $\mathbb{R}^m$ and any not too big target vector $t \in \mathbb{R}^m$, then with non-negligible probability there exists a set containing half the columns whose sum is very close to t. This is the main lemma used to show the existence of the subset T, chosen from a suitably filtered subset of the columns of A in N_0, used to reduce the excess slack. The non-negligible probability in their lemma was of order $2^{-O(m)}$, which implied that one had to try $2^{O(m)}$ disjoint subsets of the filtered columns before having a constant probability of success of finding a suitable T. In our improved variant of the discrepancy lemma, given in Lemma 8, we show that by sub-selecting a $1/(2\sqrt{m})$-fraction of the columns instead of 1/2-fraction, we can increase the success probability to constant, with the caveat of requiring a slightly larger set of initial columns.

The second source of improvement is the use of a much milder filtering step mentioned above. In both the uniform and Gaussian case, the subset T is chosen from a subset of N_0 associated with columns of A having reduced costs of absolute value at most some parameter $\Delta > 0$. The probability of finding a suitable T increases as Δ grows larger, since we have more columns to choose from, and the target integrality gap scales linearly with Δ, as the columns we choose from become more expensive as Δ grows. Depending on the distribution of c and A, the reduced cost filtering induces non-trivial correlations between the entries of the corresponding columns of A, which makes it difficult to use them within the context of the discrepancy lemma. To deal with this problem in the uniform setting, Dyer and Frieze filtered much more aggressively, by additionally restricting to columns of A lying in a sub-cube $[\alpha, \beta]^m$, where $\alpha = \Omega(\log^3 n/n)$ and $\beta := \min_{i\in[m]} \beta_i$ as above. By doing the reduced cost filtering more carefully, this allowed them to ensure that the distribution of the filtered columns in A is in fact uniform in $[\alpha, \beta]^m$, thereby removing the unwanted correlations. With this aggressive filtering, the columns in N_0 only pass the filtering step with probability $O(\beta^m \Delta)$, which is the source of the $(1/\beta)^m$ dependence in their integrality gap. In the Gaussian context, we show how to work directly with the columns of A with only reduced cost filtering, which increases the success probability of the filtering test to $\Theta(\Delta)$. While the entries of the filtered columns of A do indeed correlate, using the rotational symmetry of the Gaussian distribution, we show that after applying a suitable rotation R, the coordinates of filtered columns of RA are all independent. This allows us to apply the discrepancy lemma in a "rotated space", thereby completely avoiding the correlation issues in the uniform setting.

As already mentioned, we are also able to substantially relax the rigid requirements on the right hand side b and to remove any stringent "shape-dependence" of the integrality gap on b. Specifically, for $b_i = \beta_i n$, $\beta_i \in (0, 1/2)$, $\forall i \in [m]$, the shape parameter $\beta := \min_{i\in m} \beta_i$, is used to both lower bound $|N_0|$ by roughly $\Omega((1-2\beta)n)$, the number of zeros in x^*, as well as upper bound the ℓ_1 norm of the optimal dual solution u^* by $O(1/\beta)$ (this a main reason for the choice of the $[\alpha, \beta]^m$ sub-cube above). These bounds are both crucial for determining the existence of T. In the Gaussian setting, we are able to establish $|N_0| = \Omega(n)$ and $\|u^*\|_2 = O(1)$, using only that $\|b^-\|_2 \leq n/10$. Due to the different nature of the distributions we work with, our arguments to establish these bounds are completely different from those used by Dyer and Frieze. Firstly, the lower bound on $|N_0|$, which is strongly based on the packing structure of the IP in [5], is replaced by a sub-optimality argument. Namely, we show that the objective value of any

LP basic solution with too few zero coordinates must be sub-optimal, using the concentration properties of the Gaussian distribution. The upper bound on the ℓ_1 norm of u^* in [5] is deterministic and based on packing structure; namely, that the objective value of a Primal LP of packing-type is at most $\sum_{i=1}^n c_i \leq n$ (since $c_i \in [0,1], \forall i \in [m]$). In the Gaussian setting, we prove our bound on the norm of u^* by first establishing simple upper and lower bounds on the dual objective function, which hold with overwhelming probability, and optimizing over these simple approximations (see Lemma 4).

We note that we expect the techniques we develop here to yield improvements to the analysis of random packing IPs as well. The main technical difficulty at present is showing that a milder filtering step (e.g., just based on the reduced costs) is also sufficient in the packing case. This in essence reduces to understanding whether Lemma 8 can be generalized to handle random columns whose entries are allowed to have non-trivial correlations. Another possible approach to obtain improvements, in particular with respect to further relaxing the restrictions on b, is to try and flip both 0s to 1 and 1s to 0 in the rounding of x' to x''. The columns of W associated with the one coordinates of x^* are no longer independent however. A final open question is whether these techniques can be extended to handle discrete distributions on A and c.

1.2 Relation to Branch and Bound

In a recent breakthrough work, Dey, Dubey and Molinaro [3] proved that, if the entries of A, c are independently distributed in $[0,1]$ and for all $0 < \alpha \min\{30m, \frac{\log n}{a_2}$ one has

$$\Pr\left(\mathsf{IPGAP} \geq \alpha a_1 \frac{\log^2 n}{n}\right) \leq 4 \cdot 2^{-\alpha a_2} + \frac{1}{n}$$

for some $a_1, a_2 > 0$, then with probability at least $1 - \frac{2}{n} - 4 \cdot 2^{-\alpha a_2}$, a branch-and-bound algorithm that branches on variables, always selecting the node with the largest LP value, will produce a tree of size at most

$$n^{O(m a_1 \log a_1 + \alpha a_1 \log m)}$$

for all $\alpha \leq \min\{30m, \frac{\log n}{a_2}\}$. The values of a_1 and a_2, they get from [5].

In fact, their analysis goes through for the entries of A and c independently Gaussian $\mathcal{N}(0,1)$ distributed as well with minor modifications. Specifically, one needs to condition on the columns of $\begin{bmatrix} c^\mathsf{T} \\ A \end{bmatrix}$ having bounded norm, after which the final net argument needs to be slightly adapted. The details of this adaptation are give in the full version.

Taking $a_1 = 2 \cdot 10^{15} m^{4.5}$, $a_2 = 1/30$ and $\alpha = 30 \cdot \min\{m, \log n\}$, this result together with Theorem 1, proves that branch and bound can find the best integer solution to (Primal LP) with a branching tree of size $n^{\mathrm{poly}(m)}$ with probability $1 - \frac{6}{n} - 4 \cdot 2^{-m}$. Here, the fact that Theorem 1 depends only polynomially on m results in a much better upper bound than the exponential dependence on m from [5].

1.3 Related Work

The worst-case complexity of solving $\max\{c^\mathsf{T} x : Ax = b, x \geq 0, x \in \mathbb{Z}^n\}$ scales as $n^{O(n)}$ times a polynomial factor in the bit complexity of the problem. This is a classical result due to Lenstra [10] and Kannan [9].

If we restrict to IPs with integer data, a dynamic programming algorithm can solve $\max\{c^\mathsf{T} x : Ax = b, x \geq 0, x \in \mathbb{Z}^n\}$ in time $O(\sqrt{m}\Delta)^{2m} \log(\|b\|_\infty) + O(nm)$, where Δ is the largest absolute value of entries in the input matrix A [6,8,12]. Integer programs of the form $\max\{c^\mathsf{T} x : Ax = b, 0 \leq x \leq u, x \in \mathbb{Z}^n\}$ can similarly be solved in time

$$n \cdot O(m)^{(m+1)^2} \cdot O(\Delta)^{m\cdot(m+1)} \log^2(m \cdot \Delta),$$

which was proved in [6]. Note that integer programs of the form $\max\{c^\mathsf{T} x : Ax \leq b, x \in \{0,1\}^n\}$ can be rewritten in this latter form by adding m slack variables.

In terms of random inputs, we mention the work of Röglin and Vöcking [13]. They prove that a class of IPs satisfying some minor conditions has polynomial smoothed complexity if and only if that class admits a pseudopolynomial time algorithm. An algorithm has polynomial smoothed complexity if its running time is polynomial with high probability when its input has been perturbed by adding random noise, where the polynomial may depend on the magnitude φ^{-1} of the noise as well as the dimensions n, m of the problem. An algorithm runs in pseudopolynomial time if the running time is polynomial when the numbers are written in unary, i.e., when the input data consists of integers of absolute value at most Δ and the running time is bounded by a polynomial $p(n, m, \Delta)$. In particular, they prove that solving the randomly perturbed problem requires only polynomially many calls to the pseudopolynomial time algorithm with numbers of size $(nm\varphi)^{O(1)}$ and considering only the first $O(\log(nm\varphi))$ bits of each of the perturbed entries.

If, for the sake of comparison, we choose $b \in \mathbb{R}^m$ in (Primal LP) from a $\mathcal{N}(0,1)$ Gaussian distribution independently of c and A, then the result of [13] proves that, with high probability, it is sufficient to solve polynomially many problems with integer entries of size $(nm)^{O(1)}$. Since $\Delta = (nm)^{O(m)}$ in this setting (by Hadamard's inequality), the result of [6] tells us that this problem can be solved in time $(nm)^{O(m^3)}$.

1.4 Organization

In Section 2, we give preliminaries on probability theory, linear programming and integer rounding. In Section 3, we prove properties of the LP optimal solution x^*, and in Section 4 we look at the distribution of the columns of $\begin{bmatrix} c^\mathsf{T} \\ A \end{bmatrix}$ corresponding to indices $i \in [n]$ with $x_i^* = 0$. Then in Section 5, we prove Theorem 1, using a discrepancy result that we prove in the full version of this paper [2]. All other proofs can be found in the full version as well.

2 Preliminaries

2.1 Basic Notation

We denote the reals and non-negative reals by $\mathbb{R}, \mathbb{R}^+$ respectively, and the integers and positive integers by $\mathbb{Z}, \mathbb{N}$ respectively. For $k \geq 1$ an integer, we let $[k] := \{1, \ldots, k\}$. For $s \in \mathbb{R}$, we let $s^+ := \max\{s, 0\}$ and $s^- := \min\{s, 0\}$ denote the positive and negative part of s. We extend this to a vector $x \in \mathbb{R}^n$ by letting $x^{+(-)}$ correspond to applying the positive (negative) part operator coordinate-wise. We let $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ and $\|x\|_1 = \sum_{i=1}^n |x_i|$ denote the ℓ_2 and ℓ_1 norm respectively. We use $\log x$ to denote the base e natural logarithm. We use $0_m, 1_m \in \mathbb{R}^m$ to denote the all zeros and all ones vector respectively, and $e_1, \ldots, e_m \in \mathbb{R}^m$ denote the standard coordinate basis. We write $\mathbb{R}^m_+ := [0, \infty)^m$. For a random variable $X \in \mathbb{R}$, we let $\mathbb{E}[X]$ denote its expectation and $\mathrm{Var}[X] := \mathbb{E}[X^2] - \mathbb{E}[X]^2$ denote its variance.

2.2 The Dual Program, Gap Formula and the Optimal Solutions

A convenient formulation of the dual of (Primal LP) is given by

$$\begin{aligned} &\min\ \mathrm{val}^\star(u) := b^\mathsf{T} u + \sum_{i=1}^n (c - A^\mathsf{T} u)_i^+ \qquad \text{(Dual LP)} \\ &\text{s.t. } u \geq 0. \end{aligned}$$

To keep the notation concise, we will often use the identity $\|(c - A^t u)^+\|_1 = \sum_{i=1}^n (c - A^\mathsf{T} u)_i^+$.

For any primal solution x and dual solution u to the above pair of programs, we have the following standard formula for the primal-dual gap:

$$\begin{aligned} \mathrm{val}^\star(u) - \mathrm{val}(x) &:= b^\mathsf{T} u + \sum_{i=1}^n (c - A^\mathsf{T} u)_i^+ - c^\mathsf{T} x \qquad \text{(Gap Formula)} \\ &= (b - Ax)^\mathsf{T} u + \left(\sum_{i=1}^n x_i (A^\mathsf{T} u - c)_i^+ + (1 - x_i)(c - A^\mathsf{T} u)_i^+ \right). \end{aligned}$$

Throughout the rest of the paper, we let x^* and u^* denote primal and dual optimal basic feasible solutions for (Primal LP) and (Dual LP) respectively, which we note are unique with probability 1. We use the notation

$$W := \begin{bmatrix} c^\mathsf{T} \\ A \end{bmatrix} \in \mathbb{R}^{(m+1) \times n}, \tag{1}$$

to denote the objective extended constraint matrix. We will frequently make use of the sets $N_b := \{i \in [n] : x_i^* = b\}$, $b \in \{0, 1\}$, the 0 and 1 coordinates of x^*, and $S := \{i \in [n] : x_i^* \in (0, 1)\}$, the fractional coordinates of x^*. We will also use the fact that $|S| \leq m$, which follows since x^* is a basic solution to (Primal LP) and A has m rows.

2.3 Gaussian and Sub-Gaussian Random Variables

The standard, mean zero and variance 1, Gaussian $\mathcal{N}(0,1)$ has density function $\varphi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. A standard Gaussian vector in $\mathbb{R}^d$, denoted $\mathcal{N}(0, I_d)$, has probability density $\prod_{i=1}^d \varphi(x_i) = \frac{1}{\sqrt{2\pi}^d} e^{-\|x\|^2/2}$ for $x \in \mathbb{R}^d$. A random variable $Y \in \mathbb{R}$ is σ-sub-Gaussian if for all $\lambda \in \mathbb{R}$, we have

$$\mathbb{E}[e^{\lambda Y}] \le e^{\sigma^2 \lambda^2 / 2}. \tag{2}$$

A standard normal random variable $X \sim \mathcal{N}(0,1)$ is 1-sub-Gaussian. If variables $Y_1, \dots, Y_k \in \mathbb{R}$ are independent and respectively σ_i-sub-Gaussian, $i \in [k]$, then $\sum_{i=1}^k Y_i$ is $\sqrt{\sum_{i=1}^k \sigma_i^2}$-sub-Gaussian.
For a σ-sub-Gaussian random variable $Y \in \mathbb{R}$ we have the following standard tailbound:

$$\max\{\Pr[Y \le -\sigma s], \Pr[Y \ge \sigma s]\} \le e^{-\frac{s^2}{2}}, s \ge 0. \tag{3}$$

For $X \sim \mathcal{N}(0, I_d)$, we will use the following higher dimensional analogue:

$$\Pr[\|X\|_2 \ge s\sqrt{d}] \le e^{-\frac{d}{2}(s^2 - 2\log s - 1)} \le e^{-\frac{d}{2}(s-1)^2}, s \ge 1. \tag{4}$$

We will use this bound to show that the columns of A corresponding to the fractional coordinates in the (almost surely unique) optimal solution x^* are bounded.

Lemma 1. *Letting $S := \{i \in [n] : x_i^* \in (0,1)\}$, we have that*

$$\Pr[\exists i \in S : \|A_{\cdot,i}\|_2 \ge (4\sqrt{\log(n)} + \sqrt{m})] \le n^{-7}.$$

2.4 A Local Limit Theorem

Before we can state the discrepancy lemma, we introduce the concept of Gaussian convergence.

Definition 1. *Suppose $X_1, X_2, \dots$ is a sequence of i.i.d copies of a random variable X with density f. X is said to be (γ, k)-Gaussian convergent if the density f_n of $\sum_{i=1}^n X_i/\sqrt{n}$ satisfies:*

$$|f_n(x) - \varphi(x)| \le \frac{\gamma}{n} \quad \forall x \in \mathbb{R}, n \ge k,$$

where $\varphi := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ is the Gaussian probability density function.

The above definition quantifies the speed of convergence in the context of the central limit theorem. The rounding strategy used to obtain the main result utilizes random variables that are the weighted sum of a uniform and an independent normal variable. Crucially, the given convergence estimate will hold for these random variables:

Lemma 2. *Let U be uniform on $[-\sqrt{3}, \sqrt{3}]$ and let $Z \sim \mathcal{N}(0,1)$. Then there exists a universal constant $k_0 \ge 1$ such that $\forall \epsilon \in [0,1]$, the random variable $\sqrt{\epsilon} U + \sqrt{1-\epsilon} Z$ is $(1/10, k_0)$-Gaussian convergent and has maximum density at most 1.*

2.5 Rounding to Binary Solutions

In the proof of Theorem 1, we will take our optimal solution x^* and round it to an integer solution x', by changing the fractional coordinates. Note that as x^* is a basic solution, it has at most m fractional coordinates. One could round to a integral solution by setting all of them to 0, i.e., $x' = \lfloor x^* \rfloor$. If we assume that the Euclidean norm of every column of A is bounded by C, then we have $\|A(x^* - x')\|_2 \le mC$, since x^* has at most m fractional variables. However, by using randomized rounding we can make this bound smaller, as stated in the next lemma. We use this to obtain smaller polynomial dependence in Theorem 1.

Lemma 3. *Consider an $m \times n$ matrix A with $\|A_{\cdot,i}\|_2 \le C$ for all $i \in [m]$ and $y \in [0,1]^n$. Let $S = \{i \in [n] : y_i \in (0,1)\}$. There exists a vector $y' \in \{0,1\}^n$ with $\|A(y - y')\|_2 \le C\sqrt{|S|}/2$ and $y'_i = y_i$ for all $i \notin S$.*

3 Properties of the Optimal Solutions

The following lemma is the main result of this section, which gives principal properties we will need of the optimal primal and dual LP solutions. Namely, we prove an upper bound on the norm of the optimal dual solution u^* and a lower bound on the number of zero coordinates of the optimal primal solution x^*.

Lemma 4. *Given $\delta := \frac{\sqrt{2\pi}}{n}\|b^-\|_2 \in [0, 1/2)$, $\epsilon \in (0, 1/5)$, let x^*, u^* denote the optimal primal and dual LP solutions, and let $\alpha := \frac{1}{\sqrt{2\pi}}\sqrt{\left(\frac{1-3\epsilon}{1-\epsilon}\right)^2 - \delta^2}$ and choose $\beta \in [1/2, 1]$ with $H(\beta) = \frac{\alpha^2}{4}$. Then, with probability at least $1 - 2\left(1 + \frac{2}{\epsilon}\right)^{m+1} e^{-\frac{\epsilon^2 n}{8\pi}} - e^{-\frac{\alpha^2 n}{4}}$, the following holds:*

1. $c^\mathsf{T} x^* \ge \alpha n$.
2. $\|u^*\|_2 \le \frac{1+\epsilon}{1-3\epsilon-(1-\epsilon)\delta}$.
3. $|\{i \in [n] : x^*_i = 0\}| \ge (1-\beta)n - m$.

4 Properties of the 0 Columns

For $Y := (c, a_1, \ldots, a_m) \sim \mathcal{N}(0, I_{m+1})$ and $u \in \mathbb{R}^m_+$, let Y^u denote the random variable Y conditioned the event $c - \sum_{i=1}^m u_i a_i \le 0$. We will crucially use the following lemma directly adapted from Dyer and Frieze [5, Lemma 2.1], which shows that the columns of W associated with the 0 coordinates of x^* are independent subject to having negative reduced cost.

Recall that $N_b = \{i \in [n] : x^*_i = b\}$, that $S = \{i \in [n] : x^*_i \in (0,1)\}$ and that $W := \begin{bmatrix} c^\mathsf{T} \\ A \end{bmatrix} \in \mathbb{R}^{(m+1)\times n}$ is the objective extended constraint matrix.

Lemma 5. *Let $N'_0 \subseteq [n]$. Conditioning on $N_0 = N'_0$, the submatrix $W_{\cdot,[n]\setminus N'_0}$ uniquely determines x^* and u^* almost surely. If we further condition on the exact value of $W_{\cdot,[n]\setminus N'_0}$, assuming x^* and u^* are uniquely defined, then any column $W_{\cdot,i}$ with $i \in N'_0$ is distributed according to Y^{u^*} and independent of $W_{\cdot,[n]\setminus\{i\}}$.*

To make the distribution of the columns $A_{\cdot,i}$ easier to analyze we rotate them.

Lemma 6. *Let R be a rotation that sends the vector u to the vector $\|u\|_2 e_m$. Suppose $(c, a) \sim Y^u$. Define $a' := Ra$. Then $(c, Ra) \sim (c', a')$, where (c', a') is the value of $(\bar{c}', \bar{a}') \sim \mathcal{N}(0, I_{m+1})$ conditioned on $\|u\|_2 \bar{a}'_m - \bar{c}' \geq 0$.*

We will slightly change the distribution of the (c', a'_m) above using rejection sampling, as stated in the next lemma. This will make it easier to apply the discrepancy result of Lemma 8, which is used to round x^* to an integer solution of nearby value. In what follows, we denote the probability density function of a random variable X by f_X. In the following lemma, we use $\mathrm{unif}(0, \nu)$ to denote the uniform distribution on the interval $[0, \nu]$, for $\nu \geq 0$.

Lemma 7. *For any $\omega \geq 0$, $\nu > 0$, let $X, Y \sim \mathcal{N}(0, 1)$ be independent random variables and let $Z = \omega Y - X$. Let X', Y', Z' be these variables conditioned on $Z \geq 0$. We apply rejection sampling on (X', Y', Z') with acceptance probability*

$$\Pr[accept | Z' = z] = \frac{2\varphi(\nu/\sqrt{1+\omega^2}) \mathbf{1}_{z \in [0,\nu]}}{2\varphi(z/\sqrt{1+\omega^2})}.$$

Let X'', Y'', Z'' be the variables X', Y', Z' conditioned on acceptance. Then:

1. $\Pr[accept] = 2\nu\varphi(\nu/\sqrt{1+\omega^2})/\sqrt{1+\omega^2}$.
2. *$Y'' \sim W + V$ where $W \sim \mathcal{N}(0, \frac{1}{1+\omega^2})$, $V \sim \mathrm{unif}(0, \frac{\nu\omega}{1+\omega^2})$ and W, V are independent.*

5 Proof of Theorem 1

Recall that $S = \{i \in [n] : x^*_i \in (0, 1)\}$ and $N_0 = \{i \in [n] : x^*_i = 0\}$. To prove Theorem 1, we will assume that the following three conditions hold:

1. $\|A_{\cdot,i}\|_2 \leq 4\sqrt{\log(n)} + \sqrt{m}$, $\forall i \in S$.
2. $\|u^*\| \leq 3$.
3. $|N_0| \geq n/500$.

Using Lemmas 1 and 4 we can show that these events hold with probability $1 - n^{-\Omega(1)}$. Now we take our optimal basic solution x^* and round it to an integral vector x' using Lemma 3. Then we can generate a new solution x'' from x' by flipping the values at indices $T \subseteq N_0$ to one. In Lemma 9 we show that with high probability there is such a set T, such that x'' is a feasible solution to our primal problem and that $\mathrm{val}(x^*) - \mathrm{val}(x'')$ is small.

We do this by looking at t disjoint subsets of N_0 with small reduced costs. Then we show for each of these sets that with constant probability it contains a subset T such that for x'' obtained from T, x'' is feasible and all constraints that are tight for x^* are close to being tight for x''. This argument relies on the following improved discrepancy lemma.

Lemma 8. *For $k, m \in \mathbb{N}$, let $a = \lceil 2\sqrt{m} \rceil$ and $\theta > 0$ satisfy $\left(\frac{2\theta}{\sqrt{2\pi k}}\right)^m \binom{ak}{k} = 1$. Let $Y_1, \dots, Y_{ak} \in \mathbb{R}^m$ be i.i.d. random vectors whose coordinates are independent random variables. For $k_0 \in \mathbb{N}, \gamma \geq 0, M > 0$, assume that $Y_{1,i}$, $i \in [m]$, is (γ, k_0)-Gaussian convergent and admits a probability density $g_i : \mathbb{R} \to \mathbb{R}_+$ satisfying $\max_{x \in \mathbb{R}} g_i(x) \leq M$. Then, if*

$$k \geq \max\{(4\sqrt{m}+2)k_0, 144m^{\frac{3}{2}}(\log M + 3), 150\,000(\gamma+1)m^{\frac{7}{4}}\},$$

then for any vector $A \in \mathbb{R}^m$ with $\|A\|_2 \leq \sqrt{k}$ the following holds:

$$\Pr\left[\exists K \subset [ak] : |K| = k, \|(\sum_{j \in K} Y_j) - A\|_\infty \leq \theta\right] \geq \frac{1}{25}. \tag{5}$$

If a suitable T exists, then using the gap formula we show that $\mathrm{val}(x^*) - \mathrm{val}(x'')$ is small. Because the t sets independent the probability of failure decreases exponentially with t. Hence, we can make the probability of failure arbitrarily small by increasing t. We know $\mathrm{val}(x^*) = \mathrm{val}_{\mathsf{LP}}$ and because $x'' \in \{0,1\}^n$ we have $\mathrm{val}_{\mathsf{IP}} \geq \mathrm{val}(x'')$, so $\mathsf{IPGAP} = \mathrm{val}_{\mathsf{LP}} - \mathrm{val}_{\mathsf{IP}} \leq \mathrm{val}(x^*) - \mathrm{val}(x'')$, which is small with high probability.

Lemma 9. *If $n \geq \exp(k_0)$ for k_0 from Lemma 2 and conditions 1, 2 and 3 above hold, then*

$$\Pr\left[\mathsf{IPGAP} > 10^{15} t \cdot \frac{m^{2.5}(\log n + m)^2}{n}.\right] \leq 2 \cdot \left(1 - \frac{1}{25}\right)^t \tag{6}$$

for $1 \leq t \leq \frac{n}{20\,000\sqrt{m}k^2}$, where $k := \lceil 165\,000 m(\log(n) + m) \rceil$.

Proof. It suffices to condition on N_0 and $W_{\cdot,[n]\setminus N_0}$, subject to the conditions 1-3. Now let R be a rotation that sends the vector u^* to the vector $\|u^*\|_2 e_m$. Define:

$$\begin{aligned} \Delta &:= 10\,000\sqrt{m}k/n, \\ B_i &:= RA_{\cdot,i}, && \text{for } i \in N_0, \\ Z_t &:= \{i \in N_0 : \|u^*\|_2 (B_i)_m - c_i \in [0, t\Delta]\}, && \text{for } 1 \leq t \leq \frac{1}{2\Delta k}. \end{aligned}$$

We consider a (possibly infeasible) integral solution x' to the LP, generated by rounding the fractional coordinates from x^*. By Lemma 3 we can find such a solution with $\|A(x^* - x')\|_2 \leq (4\sqrt{\log n} + \sqrt{m})\sqrt{|S|}/2 \leq (4\sqrt{\log n} + \sqrt{m})\sqrt{m}/2$. We will select a subset $T \subseteq Z_t$ of size $|T| = k$ of coordinates to flip from 0 to 1 to obtain $x'' \in \{0,1\}^n$ from x', so $x'' := x' + \sum_{i \in T} e_i$. By complementary slackness, we know for $i \in [n]$ that $x_i^*(A^\mathsf{T} u^* - c)_i^+ = (1 - x_i^*)(c - A^\mathsf{T} u^*)^+ = 0$ and that $x_i^* \notin \{0,1\}$ implies $(c - A^\mathsf{T} u^*)_i = 0$, and for $j \in [m]$ that $u_j^* > 0$ implies $b_j = (Ax^*)_j$. This observation allows us to prove the following key bound for the integrality gap of (Primal LP)

$$\mathrm{val}(x^*) - \mathrm{val}(x'') = \mathrm{val}^\star(u^*) - \mathrm{val}(x'')$$

$$
\begin{aligned}
&= (b - Ax'')^\top u^* \\
&+ \left(\sum_{i=1}^{n} x''_i (A^\top u^* - c)_i^+ + (1 - x''_i)(c - A^\top u^*)_i^+ \right) \qquad \text{(by Gap Formula)} \\
&= (x^* - x'')^\top A^\top u^* + \sum_{i \in T} (A^\top u^* - c)_i \qquad \text{(by complementary slackness)} \\
&\leq \sqrt{m} \|u^*\|_2 \, \|A(x'' - x^*)\|_\infty + t\Delta k \qquad \text{(since } T \subseteq Z_t).
\end{aligned}
$$

Condition 2 tells us that $\|u^*\|_2 \leq 3$, and by definition we have

$$
t\Delta k \leq 27226 \cdot 10^{10} t \cdot \frac{m^{2.5}(\log(n) + m)^2}{n},
$$

so the rest of this proof is dedicated to showing the existence of a set $T \subseteq Z_t$ such that $\|A(x'' - x^*)\|_\infty \leq O(1/n)$ and $Ax'' \leq b$.

By applying Lemma 5, we see that $\{(c_i, A_{\cdot,i})\}_{i \in N_0}$ are independent vectors, distributed as $\mathcal{N}(0, I_{m+1})$ conditioned on $c_i - A_{\cdot,i}^\top u^* \leq 0$. This implies that the vectors $\{(c_i, B_i)\}_{i \in N_0}$, are also independent. By Lemma 6, it follows that $(c_i, B_i) \sim \mathcal{N}(0, I_{m+1}) \mid \|u^*\|_2 (B_i)_m - c_i \geq 0$. Note that the coordinates of B_i are therefore independent and $(B_i)_j \sim \mathcal{N}(0,1)$ for $j \in [m-1]$.

To simplify the upcoming calculations, we apply rejection sampling as specified in Lemma 7 with $\nu = \Delta t$ on $(c_i, (B_i)_m)$, for each $i \in N_0$. Let $Z'_t \subseteq N_0$ denote the indices which are accepted by the rejection sampling procedure. By the guarantees of Lemma 7, we have that $Z'_t \subseteq Z_t$ and

$$
\Pr[i \in Z'_t \mid i \in N_0] = \frac{2\Delta t \varphi(\Delta t / \sqrt{1 + \|u^*\|_2^2})}{\sqrt{1 + \|u^*\|_2^2}} \geq \frac{2\Delta t \varphi(1/2)}{\sqrt{10}} \geq \Delta t / 5.
$$

Furthermore, for $i \in Z'_t$ we know that $(B_i)_m$ is distributed as a sum of independent $N(0, \frac{1}{1 + \|u^*\|_2^2})$ and $\mathrm{unif}(0, t\Delta)$ random variables, and thus $(B_i)_m$ has mean and variance

$$
\begin{aligned}
\mu_t &:= \mathbb{E}[(B_i)_m | i \in Z'_t] = \Delta t / 2, \\
\sigma_t^2 &:= \mathrm{Var}[(B_i)_m | i \in Z'_t] = \frac{1}{1 + \|u^*\|_2^2} + \frac{1}{12} \left(\frac{\|u^*\|_2 \Delta t}{1 + \|u^*\|_2^2} \right)^2 \in [1/10, 2].
\end{aligned}
$$

Now define $\Sigma^{(t)}$ to be the diagonal matrix with $\Sigma^{(t)}_{j,j} = 1$, $j \in [m-1]$, and $\Sigma^{(t)}_{m,m} = \sigma_t$. Conditional on $i \in Z'_t$, define $B_i^{(t)}$ as the random variable

$$
B_i^{(t)} := (\Sigma^{(t)})^{-1} (B_i - \mu_t e_m) \mid i \in Z'_t.
$$

This ensures that all coordinates of $B^{(t)}$ are independent, mean zero and have variance one.

We have assumed that $|N_0| \geq n/500$ and we know $\Pr[i \in Z'_t | i \in N_0] \geq \Delta t / 5$. Now, using the Chernoff bound we find that:

$$
\Pr[|Z'_t| < 2t\sqrt{m}k] \leq \Pr\left[|Z'_t| < \frac{1}{5} t\Delta |N_0| / 2 \right]
$$

$$\leq \exp\left(-\frac{1}{8}\cdot\frac{1}{5}t\Delta|N_0|\right)$$
$$\leq \left(1-\frac{1}{25}\right)^t. \tag{7}$$

Now we define:

$$\theta := \frac{\sqrt{2\pi k}}{2}\binom{\lceil 2\sqrt{m}k\rceil}{k}^{-1/m}, \qquad d := A(x^*-x').$$
$$\theta' := 2\sqrt{m}\theta. \qquad d' := d - 1_m\theta'.$$

Observe that

$$\theta = \frac{\sqrt{2\pi k}}{2}\binom{\lceil 2\sqrt{m}k\rceil}{k}^{-1/m} \leq \frac{\sqrt{2\pi k}}{2}\left(2\sqrt{m}\right)^{-k/m} \leq \frac{1}{32m^2n}.$$

So $\theta' \leq 1/8$.

If $|Z_t'| \geq \lceil 2\sqrt{m}\rceil kt$, then we can take t disjoint subsets $Z_t^{(1)}, \ldots Z_t^{(k)}$ of Z_t' of size $\lceil 2\sqrt{m}\rceil k$. Conditioning on this event, we wish to apply Lemma 8 to each set $\{B_i^{(t)}\}_{i\in Z_t^{(l)}}$, $l \in [t]$, to help us find a candidate rounding of x' to a "good" integer solution x''.

Now we check that all conditions of Lemma 8 are satisfied. By definition we have $\left(\frac{2\theta}{\sqrt{2\pi k}}\right)^m \binom{ak}{k} = 1$, and we can bound

$$\begin{aligned}\left\|(\Sigma^{(t)})^{-1}(Rd' - e_m k\mu_t)\right\|_2 &\leq \max(1, 1/\sigma_t)(\|Rd\|_2 + \theta' + k\mu_t)\\ &\leq \sqrt{10}(\|RA(x^*-x')\|_2 + \theta' + k\Delta t/2)\\ &\leq \sqrt{10}\left(\sqrt{m}(4\sqrt{\log n} + \sqrt{m})/2 + \frac{1}{8} + \frac{1}{4}\right)\\ &\leq 4\sqrt{10m(\log n + m)} \leq \sqrt{k}.\end{aligned}$$

We now show that the conditions of Lemma 8 for $M = 1$,$\gamma = 1/10$, and k_0 specified below, are satisfied by $\{B_i^{(t)}\}_{i\in Z_t^{(l)}}$, $\forall l \in [t]$.

First, we observe that the $B_i^{(t)}$ are distributed as $(B_i^{(t)})_m \sim \sqrt{\epsilon}V + \sqrt{1-\epsilon}U$ for $\epsilon = \frac{1}{(1+\|u^*\|_2^2)\sigma_t^2}$, where U is uniform on $[-\sqrt{3}, \sqrt{3}]$ and $V \sim \mathcal{N}(0,1)$. By Lemma 2, $(B_i^{(t)})_m$ is $(1/10, k_0)$-Gaussian convergent for some k_0 and has maximum density at most 1. Recalling that the coordinates of $B_i^{(t)}$, $i \in Z_t'$, are independent and $(B_i^{(t)})_j \sim \mathcal{N}(0,1)$, $\forall j \in [m-1]$, we see that $B_i^{(t)}$ has independent $(1/10, k_0)$-Gaussian convergent entries of maximum density at most 1. Lastly, we note that

$$\begin{aligned}k &= 165\,000m(\log(n) + m) \geq 165\,000(m^2 + k_0 m)\\ &\geq \max\{(4\sqrt{m}+2)k_0, 144m^{\frac{3}{2}}(\log 1 + 3), 150\,000(\gamma+1)m^{\frac{7}{4}}\}\end{aligned}$$

as needed to apply Lemma 8, using that $n \geq \exp(k_0)$.

Therefore, applying Lemma 8, for each $l \in [t]$, with probability at least $1 - 1/25$, there exists a set $T_l \subseteq Z_t^{(l)}$ of size k such that:

$$\left\| \sum_{i \in T_l} B_i^{(t)} - (\Sigma^{(t)})^{-1}(Rd' - e_m k \mu_t) \right\|_\infty \leq \theta. \tag{8}$$

Call the event that (8) is valid for any of the t sets E_t. Because the success probabilities for each of the t sets are independent, we get:

$$\Pr[\neg E_t \mid \ |Z_t'| \geq \lceil 2\sqrt{m} \rceil tk] \leq \left(1 - \frac{1}{25}\right)^t.$$

Combining this with Equation (7), we see that $\Pr[\neg E_t] \leq 2 \cdot (1 - \frac{1}{25})^t$. If E_t occurs, we choose $T \subseteq Z_t'$, $|T| = k$, satisfying (8). Then,

$$\begin{aligned} \left\| \sum_{i \in T} A_{\cdot,i} - d' \right\|_\infty &\leq \left\| \sum_{i \in T} A_{\cdot,i} - d' \right\|_2 = \left\| \sum_{i \in T} B_{\cdot,i} - Rd' \right\|_2 \\ &= \left\| \sum_{i \in T} (\Sigma^{(t)}) B_{\cdot,i}^{(t)} + k\mu_t e_m - Rd' \right\|_2 \\ &\leq \max(1, \sigma_t)\sqrt{m} \left\| \sum_{i \in T} B_{\cdot,i}^{(t)} - (\Sigma^{(t)})^{-1}(Rd' - e_m k \mu_t) \right\|_\infty \\ &\leq 2\sqrt{m}\theta = \theta'. \end{aligned}$$

Now we will show that when E_t occurs, x'' is feasible and $\|A(x'' - x^*)\|_\infty = O(1/n)$. First we check feasibility:

$$\begin{aligned} \sum_{i=1}^m x_i'' a_{ji} &= (Ax')_j + \sum_{i \in T} a_{ji} \leq (Ax')_j + d_j' + \theta' \\ &= (Ax')_j + (A(x^* - x'))_j = (Ax^*)_j \leq b_j. \end{aligned}$$

Hence the solution is feasible for our LP. We also have

$$\begin{aligned} \|A(x'' - x^*)\|_\infty &= \|Ax'' - Ax' - d\|_\infty \\ &= \| \sum_{i \in T} A_{\cdot,i} - d' \|_\infty \leq \| \sum_{i \in T} A_{\cdot,i} - d \|_\infty + \theta' \leq 2\theta'. \end{aligned}$$

Now we can finalize our initial computation:

$$\begin{aligned} \mathrm{val}(x^*) - \mathrm{val}(x'') &\leq \sqrt{m} \|u^*\|_2 \, \|A(x'' - x^*)\|_\infty + t\Delta k \\ &\leq 6\sqrt{m}\theta' + 10\,000 \cdot \frac{\sqrt{m} \cdot t \cdot k^2}{n} \\ &\leq \frac{12}{32mn} + 27226 \cdot 10^{10} t \cdot \frac{m^{2.5}(\log n + m)^2}{n} \end{aligned}$$

$$\leq 10^{15} t \cdot \frac{m^{2.5}(\log n + m)^2}{n}.$$

□

References

1. Beier, R., Vöcking, B.: Probabilistic analysis of knapsack core algorithms. In: Munro, J.I. (ed.) Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2004, New Orleans, Louisiana, USA, January 11-14, 2004. pp. 468–477. SIAM (2004)
2. Borst, S., Dadush, D., Huiberts, S., Tiwari, S.: On the Integrality Gap of Binary Integer Programs with Gaussian Data. arXiv:2012.08346 [cs, math] (Dec 2020)
3. Dey, S.S., Dubey, Y., Molinaro, M.: Branch-and-bound solves random binary IPs in polytime. In: Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 579–591. Society for Industrial and Applied Mathematics (Jan 2021). https://doi.org/10.1137/1.9781611976465.35
4. Dyer, M., Frieze, A.: Probabilistic analysis of the generalised assignment problem. Math. Program. **55**(1-3), 169–181 (Apr 1992). https://doi.org/10.1007/bf01581197
5. Dyer, M., Frieze, A.: Probabilistic analysis of the multidimensional knapsack problem. Mathematics of OR **14**(1), 162–176 (Feb 1989). https://doi.org/10.1287/moor.14.1.162
6. Eisenbrand, F., Weismantel, R.: Proximity results and faster algorithms for integer programming using the steinitz lemma. ACM Trans. Algorithms **16**(1), 1–14 (Jan 2020). https://doi.org/10.1145/3340322
7. Goldberg, A., Marchetti-Spaccamela, A.: On finding the exact solution of a zero-one knapsack problem. In: Proceedings of the sixteenth annual ACM symposium on Theory of computing - STOC '84. ACM Press (1984). https://doi.org/10.1145/800057.808701
8. Jansen, K., Rohwedder, L.: Integer programming (Oct 2019). https://doi.org/10.1002/9781119454816.ch10
9. Kannan, R.: Minkowski's convex body theorem and integer programming. Mathematics of OR **12**(3), 415–440 (Aug 1987). https://doi.org/10.1287/moor.12.3.415
10. Lenstra, H.: Integer programming with a fixed number of variables. Mathematics of OR **8**(4), 538–548 (Nov 1983). https://doi.org/10.1287/moor.8.4.538
11. Lueker, G.S.: On the average difference between the solutions to linear and integer knapsack problems. In: Applied Probability-Computer Science: The Interface Volume 1, pp. 489–504. Birkhäuser Boston (1982). https://doi.org/10.1007/978-1-4612-5791-2
12. Papadimitriou, C.H.: On the complexity of integer programming. J. ACM **28**(4), 765–768 (Oct 1981). https://doi.org/10.1145/322276.322287
13. Röglin, H., Vöcking, B.: Smoothed analysis of integer programming. Math. Program. **110**(1), 21–56 (Jan 2007). https://doi.org/10.1007/s10107-006-0055-7