

# Bin Packing via Discrepancy of Permutations<sup>1</sup>

Friedrich Eisenbrand<sup>2</sup>

EPFL, Lausanne, Switzerland

and

Dömötör Pálvölgyi<sup>3</sup>

Eötvös Loránd University (ELTE), Budapest, Hungary

and

Thomas Rothvoß<sup>4</sup>

M.I.T., Cambridge, USA

---

A well studied special case of *bin packing* is the *3-partition problem*, where  $n$  items of size  $> \frac{1}{4}$  have to be packed in a minimum number of bins of capacity one. The famous *Karmarkar-Karp algorithm* transforms a fractional solution of a suitable LP relaxation for this problem into an integral solution that requires at most  $O(\log n)$  additional bins.

The *three-permutations-problem* of Beck is the following. Given any 3 permutations on  $n$  symbols, color the symbols red and blue, such that in any interval of any of those permutations, the number of red and blue symbols is roughly the same. The necessary difference is called the *discrepancy*.

We establish a surprising connection between bin packing and Beck's problem: The additive integrality gap of the 3-partition linear programming relaxation can be bounded by the discrepancy of 3 permutations.

This connection yields an alternative method to establish an  $O(\log n)$  bound on the additive integrality gap of the 3-partition. Reversely, making use of a recent example of 3 permutations, for which a discrepancy of  $\Omega(\log n)$  is necessary, we prove the following: The  $O(\log^2 n)$  upper bound on the additive gap for bin packing with arbitrary item sizes cannot be improved by any technique that is based on rounding up items. This lower bound holds for a large class of algorithms including the Karmarkar-Karp procedure.

Categories and Subject Descriptors: F.2.2 [Computations on discrete structures]: Non-numerical Algorithms and Problems

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Bin packing, linear programming relaxations, discrepancy theory

---

## 1. INTRODUCTION

The *bin packing* problem is the following. Given  $n$  items of size  $s_1, \dots, s_n \in [0, 1]$  respectively, the goal is to pack these items in as few bins of capacity one as possible. Bin packing is a fundamental problem in Computer Science with numerous

---

<sup>1</sup>A preliminary version of this paper appeared in SODA'11 [10].

<sup>2</sup>Email: [friedrich.eisenbrand@epfl.ch](mailto:friedrich.eisenbrand@epfl.ch). Supported by the Swiss National Science Foundation (SNF).

<sup>3</sup>Email: [dom@cs.elte.hu](mailto:dom@cs.elte.hu)

<sup>4</sup>Email: [rothvoss@math.mit.edu](mailto:rothvoss@math.mit.edu). Supported by the German Research Foundation (DFG) within the Priority Program 1307 "Algorithm Engineering", by the Alexander von Humboldt Foundation within the Feodor Lynen program, by ONR grant N00014-11-1-0053 and by NSF contract CCF-0829878.

applications in theory and practice.

The development of heuristics for bin packing with better and better performance guarantee is an important success story in the field of *Approximation Algorithms*. Johnson [16; 17] has shown that the *First Fit* algorithm requires at most  $1.7 \cdot OPT + 1$  bins and that *First Fit Decreasing* yields a solution with  $\frac{11}{9}OPT + 4$  bins (see [8] for a tight bound of  $\frac{11}{9}OPT + \frac{6}{9}$ ). An important step forward was made by Fernandez de la Vega and Luecker [11] who provided an asymptotic polynomial time approximation scheme for bin packing. The *rounding technique* that is introduced in their paper has been very influential in the design of PTAS's for many other difficult combinatorial optimization problems.

In 1982, Karmarkar and Karp [18] proposed an approximation algorithm for bin packing that can be analyzed to yield a solution using at most  $OPT + O(\log^2 n)$  bins. This seminal procedure is based on the *Gilmore Gomory LP relaxation* [13; 9]:

$$\begin{aligned} \min \sum_{p \in \mathcal{P}} x_p \\ \sum_{p \in \mathcal{P}} p \cdot x_p &\geq \mathbf{1} \\ x_p &\geq 0 \quad \forall p \in \mathcal{P} \end{aligned} \quad (\text{LP})$$

Here  $\mathbf{1} = (1, \dots, 1)^T$  denotes the all ones vector and  $\mathcal{P} = \{p \in \{0, 1\}^n : s^T p \leq 1\}$  is the set of all feasible *patterns*, i.e. every vector in  $\mathcal{P}$  denotes a feasible way to pack one bin. Let  $OPT$  and  $OPT_f$  be the value of the best integer and fractional solution respectively. The linear program (LP) has an exponential number of variables but still one can compute a basic solution  $x$  with  $\mathbf{1}^T x \leq OPT_f + \delta$  in time polynomial in  $n$  and  $1/\delta$  [18] using the Grötschel-Lovász-Schrijver variant of the Ellipsoid method [14].

The procedure of Karmarkar and Karp [18] yields an *additive integrality gap* of  $O(\log^2 n)$ , i.e.  $OPT \leq OPT_f + O(\log^2 n)$ , see also [27]. This corresponds to an asymptotic FPTAS<sup>5</sup> for bin packing. The authors in [22] conjecture that even  $OPT \leq \lceil OPT_f \rceil + 1$  holds and this even if one replaces the right-hand-side  $\mathbf{1}$  by any other positive integral vector  $b$ . This *Modified Integer Round-up Conjecture* was proven by Sebő and Shmonin [23] if the number of different item sizes is at most 7. We would like to mention that Jansen and Solis-Oba [15] recently provided an  $OPT + 1$  approximation-algorithm for bin packing if the number of item sizes is fixed.

Much of the hardness of bin packing seems to appear already in the special case of *3-partition*, where  $3n$  items of size  $\frac{1}{4} < s_i < \frac{1}{2}$  with  $\sum_{i=1}^{3n} s_i = n$  have to be packed. It is strongly **NP**-hard to distinguish between  $OPT \leq n$  and  $OPT \geq n + 1$  [12]. No stronger hardness result is known for general bin packing. A closer look into [18] reveals that, with the restriction  $s_i > \frac{1}{4}$ , the Karmarkar-Karp algorithm uses  $OPT_f + O(\log n)$  bins<sup>6</sup>.

<sup>5</sup>An asymptotic fully polynomial time approximation scheme (AFPTAS) is an approximation algorithm that produces solutions of cost at most  $(1 + \varepsilon)OPT + p(1/\varepsilon)$  in time polynomial in  $n$  and  $1/\varepsilon$ , where also  $p$  must be a polynomial.

<sup>6</sup>The *geometric grouping* procedure (Lemma 5 in [18]) discards items of size  $O(\log \frac{1}{s_{\min}})$ , where  $s_{\min}$  denotes the size of the smallest item. The geometric grouping is applied  $O(\log n)$  times in the Karmarkar-Karp algorithm. The claim follows by using that  $s_{\min} > \frac{1}{4}$  for 3-partition.

### Discrepancy theory.

Let  $[n] := \{1, \dots, n\}$  and consider a set system  $\mathcal{S} \subseteq 2^{[n]}$  over the ground set  $[n]$ . A *coloring* is a mapping  $\chi : [n] \rightarrow \{\pm 1\}$ . In *discrepancy theory*, one aims at finding colorings for which the difference of “red” and “blue” elements in all sets is as small as possible. Formally, the *discrepancy* of a set system  $\mathcal{S}$  is defined as

$$\text{disc}(\mathcal{S}) = \min_{\chi: [n] \rightarrow \{\pm 1\}} \max_{S \in \mathcal{S}} |\chi(S)|.$$

where  $\chi(S) = \sum_{i \in S} \chi(i)$ . A random coloring provides an easy bound of  $\text{disc}(\mathcal{S}) \leq O(\sqrt{n \log |\mathcal{S}|})$  [20]. The famous “*Six Standard Deviations suffice*” result of Spencer [24] improves this to  $\text{disc}(\mathcal{S}) \leq O(\sqrt{n \log(2|\mathcal{S}|/n)})$ .

If every element appears in at most  $t$  sets, then the *Beck-Fiala Theorem* [3] yields  $\text{disc}(\mathcal{S}) < 2t$ . The same authors conjecture that in fact  $\text{disc}(\mathcal{S}) = O(\sqrt{t})$ . Srinivasan [26] gave a  $O(\sqrt{t} \log n)$  bound, which was improved by Banaszczyk [1] to  $O(\sqrt{t} \log n)$ . Many such discrepancy proofs are purely existential, for instance due to the use of the pigeonhole principle. In a very recent breakthrough Bansal [2] showed how to obtain the desired colorings for the Spencer [24] and Srinivasan [26] bounds by considering a random walk, guided by the solution of a semidefinite program.

For several decades, the following *three-permutations-conjecture* or simply *Beck’s conjecture* (see Problem 1.9 in [4]) was open:

Given any 3 permutations on  $n$  symbols, one can color the symbols with red and blue, such that in every interval of every of those permutations, the number of red and blue symbols differs by  $O(1)$ .

Formally, a set of permutations  $\pi_1, \dots, \pi_k : [n] \rightarrow [n]$  induces a set-system<sup>7</sup>

$$\mathcal{S} = \{\{\pi_i(1), \dots, \pi_i(j)\} : j = 1, \dots, n; i = 1, \dots, k\}.$$

We denote the maximum discrepancy of such a set-system induced by  $k$  permutations over  $n$  symbols as  $D_k^{\text{perm}}(n)$ , then Beck’s conjecture can be rephrased as  $D_3^{\text{perm}}(n) = O(1)$ . One can provably upper bound  $D_3^{\text{perm}}(n)$  by  $O(\log n)$  and more generally  $D_k^{\text{perm}}(n)$  can be bounded by  $O(k \log n)$  [5] and by  $O(\sqrt{k} \log n)$  [26; 25] using the so-called *entropy method*.

But very recently a counterexample to Beck’s conjecture was found by Newman and Nikolov [21] (earning a prize of 100 USD offered by Joel Spencer)<sup>8</sup>. In fact, they fully settle the question by proving that  $D_3^{\text{perm}}(n) = \Theta(\log n)$ .

### Our contribution.

The first result of this paper is the following theorem.

**THEOREM 1.** *The additive integrality gap of the linear program (LP) restricted to 3-partition instances is bounded by  $6 \cdot D_3^{\text{perm}}(n)$ .*

<sup>7</sup>We only consider intervals of permutations that start from the first element. Since any interval is the difference of two such prefixes, this changes the discrepancy by a factor of at most 2.

<sup>8</sup>The counterexample was announced few months after SODA’11. As a small anecdote, both authors of [21] had a joint paper [6] on a related topic, which was presented in the same session of SODA’11 as the conference version of this paper.

This result is constructive in the following sense. If one can find a  $\alpha$  discrepancy coloring for any three permutations in polynomial time, then there is an  $OPT + O(\alpha)$  approximation algorithm for 3-partition.

The proof of Theorem 1 itself is via two steps.

- i) We show that the additive integrality gap of (LP) is at most twice the maximum *linear discrepancy* of a  $k$ -monotone matrix if all item sizes are larger than  $1/(k+1)$  (Section 3). This step is based on matching techniques and *Hall's theorem*.
- ii) We then show that the linear discrepancy of a  $k$ -monotone matrix is at most  $k$  times the discrepancy of  $k$  permutations (Section 4). This result uses a theorem of Lovász, Spencer and Vesztergombi.

The theorem then follows by setting  $k$  equal to 3 in the above steps.

Furthermore, we show that the discrepancy of  $k$  permutations is at most 4 times the linear discrepancy of a  $k$ -monotone matrix. Moreover in Section 5, we provide a  $5k \cdot \log_2(2 \min\{m, n\})$  upper bound on the linear discrepancy of a  $k$ -monotone  $n \times m$ -matrix.

Recall that most approximation algorithms for bin packing or corresponding generalizations rely on “rounding up items”, i.e. they select some patterns from the support of a fractional solution which form a valid solution to a dominating instance. Reversing the above connection, we can show that no algorithm that is only based on this principle can obtain an additive integrality of  $o(\log n)$  for item sizes  $> \frac{1}{4}$  and  $o(\log^2 n)$  for arbitrary item sizes (see Section 6). This still holds if we allow to discard and greedily pack items. More precisely:

**THEOREM 2.** *For infinitely many  $n$ , there is a bin packing instance  $s_1 \geq \dots \geq s_n > 0$  with a feasible fractional (LP) solution  $y \in [0, 1]^P$  such that the following holds: Let  $x \in \mathbb{Z}_{\geq 0}^P$  be an integral solution and  $D \subseteq [n]$  be those items that are not covered by  $x$  with the properties:*

- Use only patterns from fractional solution:  $\text{supp}(x) \subseteq \text{supp}(y)$ .
- Feasibility:  $\exists \sigma : [n] \setminus D \rightarrow [n]$  with  $\sigma(i) \leq i$  and  $\sum_{p:i \in p} x_p \geq |\sigma^{-1}(i)|$  for all  $i \in \{1, \dots, n\}$ .

*Then one has  $\mathbf{1}^T x + 2 \sum_{i \in D} s_i \geq \mathbf{1}^T y + \Omega(\log^2 n)$ .*

Improving the Karmarkar-Karp algorithm has been a longstanding open problem for many decades now. Our result shows that the recursive rounding procedure of the algorithm is optimal. In order to break the  $O(\log^2 n)$  barrier it does not suffice to consider only the patterns that are contained in an initial fractional solution as it is the case for the Karmarkar-Karp algorithm.

## 2. PRELIMINARIES

We first review some further necessary preliminaries on discrepancy theory. We refer to [20] for further details.

If  $A$  is a matrix, then we denote the  $i$ th row of  $A$  by  $A_i$  and the  $j$ th entry in the  $i$ th row by  $A_{ij}$ . The notation of discrepancy can be naturally extended to real

matrices  $A \in \mathbb{R}^{m \times n}$  as

$$\text{disc}(A) := \min_{x \in \{0,1\}^n} \|A(x - 1/2 \cdot \mathbf{1})\|_\infty,$$

see, e.g. [20]. Note that if  $A$  is the incidence matrix of a set system  $\mathcal{S}$  (i.e. each row of  $A$  corresponds to the characteristic vector of a set  $S \in \mathcal{S}$ ), then  $\text{disc}(A) = \frac{1}{2} \text{disc}(\mathcal{S})$ , hence this notation is consistent — apart from the  $\frac{1}{2}$  factor.

The *linear discrepancy* of a matrix  $A \in \mathbb{R}^{m \times n}$  is defined as

$$\text{lindisc}(A) := \max_{y \in [0,1]^n} \min_{x \in \{0,1\}^n} \|Ax - Ay\|_\infty.$$

This value can be also described by a two player game. The first player chooses a fractional vector  $y$ , then the second player chooses a 0/1 vector  $x$ . The goal of the first player is to maximize, of the second to minimize  $\|Ax - Ay\|_\infty$ . The inequality  $\text{disc}(A) \leq \text{lindisc}(A)$  holds by choosing  $y := (1/2, \dots, 1/2)$ . One more notion of defining the “complexity” of a set system or a matrix is that of the *hereditary discrepancy*:

$$\text{herdisc}(A) := \max_{B \text{ submatrix of } A} \text{disc}(B).$$

Notice that one can assume that  $B$  is formed by choosing a subset of the columns of  $A$ . This parameter is obviously at least  $\text{disc}(A)$  since we can choose  $B := A$  and in [19] even an upper bound for  $\text{lindisc}(A)$  is proved (see again [20] for a recent description).

**THEOREM 3 (LOVÁSZ, SPENCER, VESZTERGOMBI).** *For  $A \in \mathbb{R}^{m \times n}$  one has*

$$\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A).$$

### 3. BOUNDING THE GAP VIA THE DISCREPANCY OF MONOTONE MATRICES

A matrix  $A$  is called *k-monotone* if all its column vectors have non-decreasing entries from  $0, \dots, k$ . In other words  $A \in \{0, \dots, k\}^{m \times n}$  and  $A_{1j} \leq \dots \leq A_{mj}$  for any column  $j$ . We denote the maximum linear discrepancy of such matrices by

$$D_k^{\text{mon}}(n) := \max_{\substack{A \in \mathbb{Z}^{m \times n} \\ k\text{-monotone}}} \text{lindisc}(A).$$

The next theorem establishes step i) mentioned in the introduction.

**THEOREM 4.** *Consider the linear program (LP) and suppose that the item sizes satisfy  $s_1, \dots, s_n > \frac{1}{k+1}$ . Then*

$$\text{OPT} \leq \text{OPT}_f + \left(1 + \frac{1}{k}\right) D_k^{\text{mon}}(n).$$

**PROOF.** Assume that the item sizes are sorted such that  $s_1 \geq \dots \geq s_n$ . Let  $y$  be any optimum basic solution of (LP) and let  $p_1, \dots, p_m$  be the list of patterns. Since  $y$  is a basic solution, its support satisfies  $|\{i : y_i > 0\}| \leq n$ . Hence by deleting unused patterns, we may assume<sup>9</sup> that  $m = n$ .

<sup>9</sup>In case that there are less than  $n$  patterns, we add empty patterns.

We define  $B = (p_1, \dots, p_n) \in \{0, 1\}^{n \times n}$  as the matrix composed of the patterns as column vectors. Clearly  $By = \mathbf{1}$ . Let  $A$  be the matrix that is defined by  $A_i := \sum_{j=1}^i B_j$ , again  $A_i$  denotes the  $i$ th row of  $A$ . In other words,  $A_{ij}$  denotes the number of items of types  $1, \dots, i$  in pattern  $p_j$ . Since  $By = \mathbf{1}$  we have  $Ay = (1, 2, 3, \dots, n)^T$ . Each column of  $A$  is monotone. Furthermore, since no pattern contains more than  $k$  items one has  $A_{ij} \in \{0, \dots, k\}$ , thus  $A$  is  $k$ -monotone.

We attach a row  $A_{n+1} := (k, \dots, k)$  as the new last row of  $A$ . Clearly  $A$  remains  $k$ -monotone. There exists a vector  $x \in \{0, 1\}^n$  with

$$\|Ax - Ay\|_\infty \leq \text{lindisc}(A) \leq D_k^{\text{mon}}(n).$$

We buy  $x_i$  times pattern  $p_i$  and  $D_k^{\text{mon}}(n)$  times the pattern that only contains the largest item of size  $s_1$ .

It remains to show: (1) this yields a feasible solution; (2) the number of patterns does not exceed the claimed bound of  $OPT_f + (1 + \frac{1}{k}) \cdot D_k^{\text{mon}}(n)$ .

For the latter claim, recall that the constraint emerging from row  $A_{n+1} = (k, \dots, k)$  together with  $\sum_{i=1}^n y_i = OPT_f$  provides

$$k \sum_{i=1}^n x_i \leq k \cdot \sum_{i=1}^n y_i + D_k^{\text{mon}}(n) = k \cdot OPT_f + D_k^{\text{mon}}(n).$$

We use this to upper bound the number of opened bins by

$$\sum_{i=1}^n x_i + D_k^{\text{mon}}(n) \leq OPT_f + \left(1 + \frac{1}{k}\right) \cdot D_k^{\text{mon}}(n).$$

It remains to prove that our integral solution is feasible. To be more precise, we need to show that any item  $i$  can be assigned to a space reserved for an item of size  $s_i$  or larger.

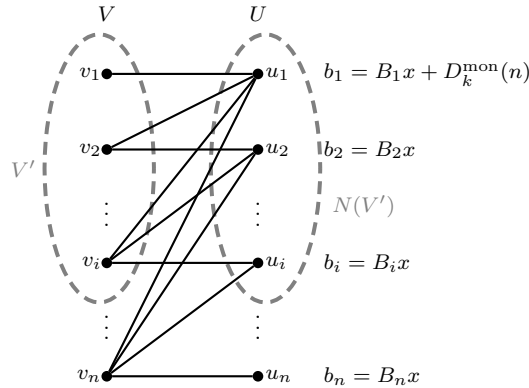


Fig. 1. The bipartite graph in the proof of Theorem 4

To this end, consider a bipartite graph with nodes  $V = \{v_1, \dots, v_n\}$  on the left, representing the items. The nodes on the right are the set  $U = \{u_1, \dots, u_n\}$ , where

each  $u_i$  is attributed with a multiplicity  $b_i$  representing the number of times that we reserve space for items of size  $s_i$  in our solution, see Figure 1. Recall that

$$b_i = \begin{cases} B_i x + D_k^{\text{mon}}(n) & \text{if } i = 1 \\ B_i x & \text{otherwise} \end{cases}.$$

We insert an edge  $(v_i, u_j)$  for all  $i \geq j$ . The meaning of this edge is the following. One can assign item  $i$  into the space which is reserved for item  $j$  since  $s_i \leq s_j$ . We claim that there exists a  $V$ -perfect matching, respecting the multiplicities of  $U$ . By *Hall's Theorem*, see, e.g. [7], it suffices to show for any subset  $V' \subseteq V$  that the multiplicities of the nodes in  $N(V')$  (the neighborhood of  $V'$ ) are at least  $|V'|$ . Observe that  $N(v_i) \subseteq N(v_{i+1})$ , hence it suffices to prove the claim for sets of the form  $V' = \{1, \dots, i\}$ . For such a  $V'$  one has

$$\sum_{u_j \in N(V')} b_j = D_k^{\text{mon}}(n) + \sum_{j=1}^i B_j x = D_k^{\text{mon}}(n) + A_i x \geq A_i y = i$$

and the claim follows.  $\square$

#### 4. BOUNDING THE DISCREPANCY OF MONOTONE MATRICES BY THE DISCREPANCY OF PERMUTATIONS

In this section, we show that the linear discrepancy of  $k$ -monotone matrices is essentially bounded by the discrepancy of  $k$  permutations. This corresponds to step ii) in the proof of the main theorem. By Theorem 3 it suffices to bound the discrepancy of  $k$ -monotone matrices by the discrepancy of  $k$  permutations times a suitable factor.

We first explain how one can associate a permutation to a 1-monotone matrix. Suppose that  $B \in \{0, 1\}^{m \times n}$  is a 1-monotone matrix. If  $B^j$  denotes the  $j$ -th column of  $B$ , then the permutation  $\pi$  that we associate with  $B$  is the (not necessarily unique) permutation that satisfies  $B^{\pi(1)} \geq B^{\pi(2)} \geq \dots \geq B^{\pi(n)}$  where  $u \geq v$  for vectors  $u, v \in \mathbb{R}^m$  if  $u_i \geq v_i$  for all  $1 \leq i \leq m$ . On the other hand the matrix  $B$  (potentially plus some extra rows and after merging identical rows) gives the incidence matrix of the set-system induced by  $\pi$ .

A  $k$ -monotone matrix  $B$  can be decomposed into a sum of 1-monotone matrices  $B^1, \dots, B^k$ . Then any  $B^\ell$  naturally corresponds to a permutation  $\pi_\ell$  of the columns as we explained above. A low-discrepancy coloring of these permutations yields a coloring that has low discrepancy for any  $B^\ell$  and hence also for  $B$ , as we show now in detail.

**THEOREM 5.** *For any  $k, n \in \mathbb{N}$ , one has  $D_k^{\text{mon}}(n) \leq k \cdot D_k^{\text{perm}}(n)$ .*

**PROOF.** Consider any  $k$ -monotone matrix  $A \in \mathbb{Z}^{m \times n}$ . By virtue of Theorem 3, there is a  $m \times n'$  submatrix,  $B$ , of  $A$  such that  $\text{lindisc}(A) \leq 2 \cdot \text{disc}(B)$ , thus it suffices to show that  $\text{disc}(B) \leq \frac{k}{2} \cdot D_k^{\text{perm}}(n)$ . Of course,  $B$  itself is again  $k$ -monotone.

Let  $B^\ell$  also be a  $m \times n'$  matrix, defined by

$$B_{ij}^\ell := \begin{cases} 1 & \text{if } B_{ij} \geq \ell \\ 0, & \text{otherwise.} \end{cases}$$

The matrices  $B^\ell$  are 1-monotone, and the matrix  $B$  decomposes into  $B = B^1 + \dots + B^k$ . As mentioned above, for any  $\ell$ , there is a (not necessarily unique) permutation  $\pi_\ell$  on  $[n]$  such that  $B^{\ell, \pi_\ell(1)} \geq B^{\ell, \pi_\ell(2)} \geq \dots \geq B^{\ell, \pi_\ell(n)}$ , where  $B^{\ell, j}$  denotes the  $j$ th column of  $B^\ell$ . Observe that the row vector  $B_i^\ell$  is the characteristic vector of the set  $\{\pi_\ell(1), \dots, \pi_\ell(j)\}$ , where  $j$  denotes the number of ones in  $B_i^\ell$ .

Let  $\chi : [n] \rightarrow \{\pm 1\}$  be the coloring that has discrepancy at most  $D_k^{\text{perm}}(n)$  with respect to all permutations  $\pi_1, \dots, \pi_k$ . In particular  $|B_i^\ell \chi| \leq D_k^{\text{perm}}(n)$ , when interpreting  $\chi$  as a  $\pm 1$  vector. Then by the triangle inequality

$$\text{disc}(B) \leq \frac{1}{2} \|B\chi\|_\infty \leq \frac{1}{2} \sum_{\ell=1}^k \|B^\ell \chi\|_\infty \leq \frac{k}{2} D_k^{\text{perm}}(n).$$

□

Combining Theorem 4 and Theorem 5, we conclude

**COROLLARY 6.** *Given any bin packing instance with  $n$  items of size bigger than  $\frac{1}{k+1}$  one has*

$$OPT \leq OPT_f + 2k \cdot D_k^{\text{perm}}(n).$$

In particular, this proves Theorem 1, our main result.

**Bounding the discrepancy of permutations in terms of the discrepancy of monotone matrices**

In addition we would like to note that the discrepancy of permutations can be also bounded by the discrepancy of  $k$ -monotone matrices as follows.

**THEOREM 7.** *For any  $k, n \in \mathbb{N}$ , one has  $D_k^{\text{perm}}(n) \leq 4 \cdot D_k^{\text{mon}}(n)$ .*

**PROOF.** We will show that for any permutations  $\pi_1, \dots, \pi_k$  on  $[n]$ , there is a  $kn \times n$   $k$ -monotone matrix  $C$  with  $\text{disc}(\pi_1, \dots, \pi_k) \leq 4 \cdot \text{disc}(C)$ . Let  $\Sigma \in \{1, \dots, n\}^{kn}$  be the string which we obtain by concatenating the  $k$  permutations. That means  $\Sigma = (\pi_1(1), \dots, \pi_1(n), \dots, \pi_k(1), \dots, \pi_k(n))$ . Let  $C$  the matrix where  $C_{ij}$  is the number of appearances of  $j \in \{1, \dots, n\}$  among the first  $i \in \{1, \dots, kn\}$  entries of  $\Sigma$ . By definition,  $C$  is  $k$ -monotone, in fact it is the “same”  $k$ -monotone matrix as in the previous proof.

Choose  $y := (\frac{1}{2}, \dots, \frac{1}{2})$  to have  $Cy = (\frac{1}{2}, 1, \dots, \frac{kn}{2})$ . Let  $x \in \{0, 1\}^n$  be a vector with  $\|Cx - Cy\|_\infty \leq \text{disc}(C)$ . Consider the coloring  $\chi : [n] \rightarrow \{\pm 1\}$  with  $\chi(j) := 1$  if  $x_j = 1$  and  $\chi(j) := -1$  if  $x_j = 0$ . We claim that the discrepancy of this coloring is bounded by  $4 \cdot \text{disc}(C)$  for all  $k$  permutations. Consider any prefix  $S := \{\pi_i(1), \dots, \pi_i(\ell)\}$ . Let  $r = C_{(i-1)n+\ell} \in \{i-1, i\}^n$  be the row of  $C$  that corresponds to this prefix. With these notations we have

$$|\chi(S)| \leq |(r - (i-1)\mathbf{1}) \cdot (2x - 2y)| \leq 2 \cdot \left( \underbrace{|r(x-y)|}_{\leq \text{disc}(C)} + \underbrace{|k \cdot \mathbf{1}(x-y)|}_{\leq \text{disc}(C)} \right) \leq 4 \cdot \text{disc}(C).$$

Here the inequality  $|(k \cdot \mathbf{1}) \cdot (x-y)| \leq \text{disc}(C)$  comes from the fact that  $k \cdot \mathbf{1} = (k, \dots, k)$  is the last row of  $C$ . □



## 5. A BOUND ON THE DISCREPANCY OF MONOTONE MATRICES

Finally, we want to provide a non-trivial upper bound on the linear discrepancy of  $k$ -monotone matrices. The result of Spencer, Srinivasan and Tetali [25; 26] together with Theorem 5 yields a bound of  $D_k^{\text{mon}}(n) = O(k^{3/2} \log n)$ . This bound can be reduced by a direct proof that shares some similarities with that of Bohus [5]. Note that  $D_k^{\text{mon}}(n) \geq k/2$ , as the  $k$ -monotone  $1 \times 1$  matrix  $A = (k)$  together with target vector  $y = (1/2)$  witnesses.

**THEOREM 8.** *Consider any  $k$ -monotone matrix  $A \in \mathbb{Z}^{n \times m}$ . Then*

$$\text{lindisc}(A) \leq 5k \cdot \log_2(2 \min\{n, m\}).$$

**PROOF.** If  $n = m = 1$ ,  $\text{lindisc}(A) < \frac{k}{2}$ , hence the claim is true. Let  $y \in [0, 1]^m$  by any vector. We can remove all columns  $i$  with  $y_i = 0$  or  $y_i = 1$  and then apply induction (on the size of the matrix). Next, if  $m > n$ , i.e. the number of columns is bigger than the number of constraints, then  $y$  is not a basic solution of the system

$$\begin{aligned} Ay &= b \\ 0 \leq y_i &\leq 1 \quad \forall i = 1, \dots, m. \end{aligned}$$

We replace  $y$  by a basic solution  $y'$  and apply induction (since  $y'$  has some integer entries and  $Ay = Ay'$ ).

Finally it remains to consider the case  $m \leq n$ . Let  $a_1, \dots, a_n$  be the rows of  $A$  and let  $d(j) := \|a_{j+1} - a_j\|_1$  for  $j = 1, \dots, n-1$ , i.e.  $d(j)$  gives the cumulated differences between the  $j$ th and the  $(j+1)$ th row. Since the columns are  $k$ -monotone, each column contributes at most  $k$  to the sum  $\sum_{j=1}^{n-1} d(j)$ . Thus

$$\sum_{j=1}^{n-1} d(j) \leq mk \leq nk.$$

By the pigeonhole principle at least  $n/2$  many rows  $j$  have  $d(j) \leq 2k$ . Take any second of these rows and we obtain a set  $J \subseteq \{1, \dots, n-1\}$  of size  $|J| \geq n/4$  such that for every  $j \in J$  one has  $d(j) \leq 2k$  and  $(j+1) \notin J$ . Let  $A'y = b'$  be the subsystem of  $n' \leq \frac{3}{4}n$  many equations, which we obtain by deleting the rows in  $J$  from  $Ay = b$ . We apply induction to this system and obtain an  $x \in \{0, 1\}^m$  with

$$\begin{aligned} \|A'x - A'y\|_\infty &\leq 5k \cdot \log_2(2n') \\ &\leq 5k \log_2\left(2 \cdot \frac{3}{4}n\right) \\ &\leq 5k \log_2(2n) - 5k \log_2\left(\frac{4}{3}\right) \\ &\leq 5k \log_2(2n) - 2k. \end{aligned}$$

Now consider any  $j \in \{1, \dots, n\}$ . If  $j \notin J$ , then row  $j$  still appeared in  $A'y = b'$ , hence  $|a_j^T x - a_j^T y| \leq 5k \log_2(2n) - 2k$ . Now suppose  $j \in J$ . We remember that  $j+1 \notin J$ , thus  $|a_{j+1}^T(x-y)| \leq 5k \log_2(2n) - 2k$ . But then using the triangle inequality

$$|a_j^T x - a_j^T y| \leq \underbrace{|(a_{j+1} - a_j)^T(x-y)|}_{\leq d(j) \leq 2k} + \underbrace{|a_{j+1}^T(x-y)|}_{\leq 5k \log_2(2n) - 2k} \leq 5k \cdot \log(2n).$$

□

## 6. LOWER BOUNDS FOR ALGORITHMS BASED ON ROUNDING UP ITEMS

Let us remind ourselves, how the classical approximation algorithms for bin packing work. For example in the algorithm of de la Vega and Lueker [11] one first *groups* the items, i.e. the item sizes  $s_i$  are rounded up to some  $s'_i \geq s_i$  such that (1) the number of different item sizes in  $s'$  is at most  $O(1/\varepsilon^2)$  (for some proper choice of  $\varepsilon$ ) and (2) the optimum number of bins increases only by a  $(1 + \varepsilon)$  factor. Note that any solution for the new instance with bigger item sizes induces a solution with the same value for the original instance. Then one computes a basic solution<sup>10</sup>  $y \in \mathbb{Q}_{\geq 0}^{\mathcal{P}'}$  to (LP) with  $|\text{supp}(y)| \leq O(1/\varepsilon^2)$  and uses  $(\lceil y_p \rceil)_{p \in \mathcal{P}'}$  as approximate solution (here  $\mathcal{P}'$  are the feasible patterns induced by sizes  $s'$ ).

In contrast, the algorithm of Karmarkar and Karp [18] uses an iterative procedure, where in each of the  $O(\log n)$  iterations, the item sizes are suitably rounded and the integral parts  $\lfloor y_p \rfloor$  from a basic solution  $y$  are bought. Nevertheless, both algorithms rely only on the following properties of bin-packing:

- *Replacement property*: If  $p$  is a feasible pattern (i.e.  $\sum_{i \in p} s_i \leq 1$ ) with  $j \in p$  and  $s_i \leq s_j$ , then  $(p \setminus \{j\}) \cup \{i\}$  is also feasible.
- *Discarding items*: Any subset  $D \subseteq [n]$  of items can be greedily assigned to at most  $2s(D) + 1$  many bins ( $s(D) := \sum_{i \in D} s_i$ ).

For a vector  $x \in \mathbb{Z}_{\geq 0}^{\mathcal{P}}$ , we say that  $x$  *buys*  $\sum_{p \in \mathcal{P}: i \in p} x_p$  many slots for item  $i$ . The replacement property implies that e.g. for two items  $s_1 \geq s_2$ ;  $x$  induces a feasible solution already if it buys no slot for item 2, but 2 slots for the larger item 1.

In the following we always assume that  $s_1 \geq \dots \geq s_n$ . We say that an integral vector  $x$  *covers* the non-discarded items  $[n] \setminus D$ , if there is a map  $\sigma : [n] \setminus D \rightarrow [n]$  with  $\sigma(i) \leq i$  and  $\sum_{p \in \mathcal{P}: i \in p} x_p \geq |\sigma^{-1}(i)|$ . Here the map  $\sigma$  assigns items  $i$  to a slot that  $x$  reserves for an item of size  $s_{\sigma(i)} \geq s_i$ . In other words, a tuple  $(x, D)$  corresponds to a feasible solution if  $x$  covers the items in  $[n] \setminus D$  and the cost of this solution can be bounded by  $\mathbf{1}^T x + 2s(D) + 1$ .

It is not difficult to see<sup>11</sup> that for the existence of such a mapping  $\sigma$  it is necessary (though i.g. not sufficient) that

$$\sum_{p \in \mathcal{P}} x_p \cdot |p \cap \{1, \dots, i\}| \geq i - |D| \quad \forall i \in [n]. \quad (1)$$

The algorithm of Karmarkar and Karp starts from a fractional solution  $y$  and obtains a pair  $(x, D)$  with  $\mathbf{1}^T x \leq \mathbf{1}^T y$  and  $\sum_{i \in D} s_i = O(\log^2 n)$  such that  $x$  covers  $[n] \setminus D$ . Moreover, it has the property<sup>12</sup> that  $\text{supp}(x) \subseteq \text{supp}(y)$ , which means that

<sup>10</sup>Alternatively one can compute an optimum solution for the rounded instance by dynamic programming in time  $n^{(1/\varepsilon)^{O(1/\varepsilon)}}$ , but using the LP reduces the running time to  $f(\varepsilon) \cdot n$ .

<sup>11</sup>Proof sketch: Assign input items  $i$  iteratively in increasing order (starting with the largest one) to the smallest available slot. If there is none left for item  $i$ , then there are less than  $i$  slots for items  $1, \dots, i$ .

<sup>12</sup>The Karmarkar-Karp method solves the (LP)  $O(\log n)$  many times for smaller and smaller instances. This can either be done by reoptimizing the previous fractional solution or by starting from scratch. We assume here that the first option is chosen.

it only uses patterns that are already contained in the support of the fractional solution  $y$ . Hence this method falls into an abstract class of algorithms that can be characterized as follows:

**DEFINITION 1.** *We call an approximation algorithm for bin packing based on rounding up items, if for given item sizes  $s_1, \dots, s_n$  and a given fractional solution  $y \in [0, 1]^P$  to (LP) it performs as follows: The algorithm produces a tuple  $(x, D)$  such that (1)  $x \in \mathbb{Z}_{\geq 0}^P$ , (2)  $\text{supp}(x) \subseteq \text{supp}(y)$  and (3)  $x$  covers  $[n] \setminus D$ . We define the additive integrality gap for a tuple  $(x, D)$  as*

$$\mathbf{1}^T x + 2 \sum_{i \in D} s_i - \mathbf{1}^T y.$$

We can now argue that the method of Karmarkar and Karp is optimal for all algorithms that are based on rounding up items. The crucial ingredient is the recent result of Newman and Nikolov [21] that there are 3 permutations of discrepancy  $\Omega(\log n)$ . For a permutation  $\pi$  we let  $\pi([i]) = \{\pi(1), \dots, \pi(i)\}$  be the prefix consisting of the first  $i$  symbols. In the following, let  $\mathbb{O} = \{\dots, -5, -3, -1, 1, 3, 5, \dots\}$  be the set of odd integers.

**THEOREM 9.** [21] *For every  $k \in \mathbb{N}$  and  $n = 3^k$ , there are permutations  $\pi_1, \pi_2, \pi_3 : [n] \rightarrow [n]$  such that  $\text{disc}(\pi_1, \dots, \pi_3) \geq k/3$ . Additionally, for every coloring  $\chi : [n] \rightarrow \mathbb{O}$  one has:*

- If  $\chi([n]) \geq 1$ , then there are  $i, j$  such that  $\chi(\pi_j([i])) \geq (k+2)/3$
- If  $\chi([n]) \leq -1$ , then there are  $i, j$  such that  $\chi(\pi_j([i])) \leq -(k+2)/3$ .

Note that the result of [21] was only stated for  $\{\pm 1\}$  colorings. But the proof uses only the fact that the colors  $\chi(i)$  are odd integers<sup>13</sup>. This theorem does not just yield a  $\Omega(\log n)$  discrepancy, but also the stronger claim that any coloring  $\chi$  which is balanced (i.e.  $|\chi([n])|$  is small) yields a prefix of one of the permutations which has a “surplus” of  $\Omega(\log n)$  and another prefix that has a “deficit” of  $\Omega(\log n)$ .

We begin with slightly reformulating the result. Here we make no attempt to optimize any constant. A *string*  $\Sigma = (\Sigma(1), \dots, \Sigma(q))$  is an ordered sequence;  $\Sigma(\ell)$  denotes the symbol at the  $\ell$ th position and  $\Sigma[\ell] = (\Sigma(1), \dots, \Sigma(\ell))$  denotes the prefix string consisting of the first  $\ell$  symbols. We write  $\chi(\Sigma[\ell]) = \sum_{i=1}^{\ell} \chi(\Sigma(i))$  and  $\mathbb{O}_{\geq -1} = \{-1, 1, 3, 5, \dots\}$ .

**COROLLARY 10.** *For infinitely many even  $n$ , there is a string  $\Sigma \in [n]^{3n}$ , each of the  $n$  symbols appearing exactly 3 times, such that: for all  $\chi : [n] \rightarrow \mathbb{O}_{\geq -1}$  with  $\chi([n]) \leq \frac{\log n}{40}$ , there is an even  $\ell \in \{1, \dots, 3n\}$  with  $\chi(\Sigma[\ell]) \leq -\frac{\log n}{20}$ .*

Note that this statement is in fact true for every large enough  $n$  using a similar argument but we omit the proof as for us this weaker version suffices.

<sup>13</sup>The only point where [21] uses that  $\chi(i) \in \{\pm 1\}$  is the base case  $k = 1$  of the induction in the proof of Lemma 2. In fact, the case  $\chi(\{3\}) \geq 1$  with a single positive symbol  $i \in \{1, 2, 3\}$  becomes possible if one considers colorings with odd numbers. However, also this case can easily be seen to be true. Interestingly, coloring all multiples of 3 with +2 and all other numbers with -1 would yield a constant discrepancy.

PROOF. For some  $k \in \mathbb{N}$ , let  $\pi_1, \pi_2, \pi_3$  be the permutations on  $[3^k]$  according to Theorem 9. We append the permutations together to a string  $\Sigma$  of length  $3 \cdot 3^k$ . Additionally, for  $n := 3^k + 1$ , we append 3 times the symbol  $n$  to  $\Sigma$ . Thus

$$\Sigma = (\pi_1(1), \dots, \pi_1(3^k), \pi_2(1), \dots, \pi_2(3^k), \pi_3(1), \dots, \pi_3(3^k), n, n, n)$$

and  $\Sigma$  has even length.

Next, let  $\chi : [n] \rightarrow \mathbb{O}_{\geq -1}$  be any coloring with  $|\chi([n])| \leq \frac{\log n}{40}$ . Reducing the values of at most  $\frac{1}{2}(\frac{\log n}{40} + 1)$  colors by 2, we obtain a coloring  $\chi' : [n] \rightarrow \mathbb{O}_{\geq -1}$  with  $\chi'([3^k]) \leq -1$ . Then by Theorem 9 there are  $j \in \{1, \dots, 3\}$  and  $i \in \{1, \dots, 3^k\}$  such that  $\chi'(\pi_j([i])) \leq -(k+2)/3$ . For  $\ell := (j-1) \cdot 3^k + i$  one has

$$\chi(\Sigma[\ell]) \leq \chi'(\Sigma[\ell]) + 3\left(\frac{\log n}{40} + 2\right) \leq (j-1) \cdot \underbrace{\chi'([3^k])}_{<0} + \underbrace{\chi'(\pi_j[i])}_{\leq -(k+2)/3} + 3\left(\frac{\log n}{20} + 2\right) \leq -\frac{\log n}{20}$$

for  $n$  large enough. If  $\ell$  is not even, we can increment it by 1 — the discrepancy is changed by at most 2 (since we may assume that the last symbol  $\Sigma(\ell)$  is negative, thus  $-1$ ), which can be absorbed into the slack that we still have.  $\square$

### 6.1 A $\Omega(\log n)$ lower bound for the case of item sizes $> 1/4$

In the following, for an even  $n$ , let  $\Sigma$  be the string from Cor. 10. We define a matrix  $A \in \{0, 1\}^{3n \times n}$  such that

$$A_{ij} := \begin{cases} 1 & \Sigma(i) = j \\ 0 & \text{otherwise.} \end{cases}$$

Note that  $A$  has a single one entry per row and 3 one entries per column.

Next, we add up pairs of consecutive rows to obtain a matrix  $B \in \{0, 1, 2\}^{(3/2)n \times n}$ . Formally  $B_i := A_{2i-1} + A_{2i}$ . We define a bin packing instance by choosing item sizes  $s_i := \frac{1}{3} - \varepsilon i$  for items  $i = 1, \dots, \frac{3}{2}n$  with  $\varepsilon := \frac{1}{20n}$ . Then  $\frac{1}{3} > s_1 > s_2 > \dots > s_{(3/2)n} > \frac{1}{4}$ . Furthermore we consider  $B$  as our pattern matrix and  $y := (\frac{1}{2}, \dots, \frac{1}{2})$  a corresponding feasible fractional solution. Note that  $By = \mathbf{1}$ .

In the following theorem we will assume for the sake of contradiction that this instance admits a solution  $(x, D)$  respecting Def. 1 with additive gap  $o(\log n)$ . It is not difficult to see, that then  $|D| = o(\log n)$  and  $|\mathbf{1}^T x - \mathbf{1}^T y| = o(\log n)$ . The integral vector  $x$  defines a coloring  $\chi : [n] \rightarrow \mathbb{O}_{\geq -1}$  via the equation  $x_i = y_i + \frac{1}{2}\chi(i)$ . This coloring is balanced, i.e.  $|\chi([n])| = o(\log n)$ . Thus there is a prefix string  $\Sigma[\ell]$  with a deficit of  $\chi(\Sigma[\ell]) \leq -\Omega(\log n)$ . This corresponds to  $x$  having  $\ell/2 - \Omega(\log n)$  slots for the largest  $\ell/2$  items, which implies that  $x$  cannot be feasible. Now the proof in detail:

**THEOREM 11.** *There is no algorithm for bin packing, based on rounding up items which achieves an additive integrality gap of  $o(\log n)$  for all instances with  $s_1, \dots, s_n > 1/4$ .*

PROOF. Let  $(x, D)$  be a solution to the constructed instance with  $\text{supp}(x) \subseteq \text{supp}(y)$  such that  $x$  is integral and covers the non-discarded items  $[\frac{3}{2}n] \setminus D$ . For the sake of contradiction assume that

$$\mathbf{1}^T x + 2 \sum_{i \in D} s_i \leq \mathbf{1}^T y + o(\log n).$$

Clearly we may assume that  $\mathbf{1}^T x \leq \mathbf{1}^T y + \frac{1}{600} \log n$ , otherwise there is nothing to show. Note that  $\mathbf{1}^T x + 2s(D) \geq \frac{(3/2)^n - |D|}{3} + 2 \cdot \frac{|D|}{4} = \mathbf{1}^T y + \frac{|D|}{6}$  (since  $\frac{1}{3} > s_i > \frac{1}{4}$ ) and thus  $|D| \leq \frac{1}{100} \log n$ . Furthermore  $\mathbf{1}^T x \geq \frac{(3/2)^n - |D|}{3} \geq \mathbf{1}^T y - \frac{1}{300} \log n$ . We can summarize:

$$|\mathbf{1}^T x - \mathbf{1}^T y| \leq \frac{\log n}{300} \quad \text{and} \quad \sum_{i'=1}^i B_{i'} x \geq i - \frac{\log n}{100} \quad \forall i \in [\frac{3}{2}n] \quad (2)$$

We will now lead this to a contradiction. Recall that every symbol  $i \in \{1, \dots, n\}$  corresponds to a column of matrix  $B$ . Define a coloring  $\chi : [n] \rightarrow \mathbb{O}_{\geq -1}$  such that  $x_i = \frac{1}{2} + \frac{1}{2}\chi(i)$ . Note that indeed the integrality of  $x_i$  implies that  $\chi(i)$  is an odd integer. Furthermore  $|\chi([n])| = 2 \cdot |\mathbf{1}^T x - \mathbf{1}^T y| \leq \frac{1}{150} \log n$ . Using Cor. 10 there is a  $2q \in \{1, \dots, 3n\}$  such that  $\chi(\Sigma[2q]) \leq -\frac{\log n}{20}$ . The crucial observation is that by construction  $\chi(\Sigma[2q]) = \sum_{i=1}^q B_i \chi$ . Then the number of slots that  $x$  reserves for the largest  $q$  items is

$$\sum_{i=1}^q B_i x = \underbrace{\sum_{i=1}^q B_i y}_{=q} + \frac{1}{2} \sum_{i=1}^q B_i \chi = q + \frac{1}{2} \underbrace{\chi(\Sigma[2q])}_{\leq -\frac{\log n}{20}} \leq q - \frac{\log n}{40}.$$

Thus  $x$  cannot cover items  $[\frac{3}{2}n] \setminus D$ .  $\square$

## 6.2 A $\Omega(\log^2 n)$ lower bound for the general case

Starting from the pattern matrix  $B$  defined above, we will construct another pattern matrix  $C$  and a vector  $b$  of item multiplicities such that for the emerging instance even a  $o(\log^2 n)$  additive integrality gap is not achievable by just rounding up items.

Let  $\ell := \log n$  be a parameter. We will define *groups* of items for every  $j = 1, \dots, \ell$ , where group  $j \in \{1, \dots, \ell\}$  contains  $\frac{3}{2}n$  many different item types; each one with multiplicity  $2^{j-1}$ . Define

$$C := \begin{pmatrix} 2^0 \cdot B & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & 2^1 \cdot B & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 2^2 \cdot B & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & 2^{\ell-1} \cdot B \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 2^0 \cdot \mathbf{1} \\ 2^1 \cdot \mathbf{1} \\ 2^2 \cdot \mathbf{1} \\ \vdots \\ 2^{\ell-1} \cdot \mathbf{1} \end{pmatrix},$$

thus  $C$  is an  $\frac{3}{2}n\ell \times n\ell$  matrix and  $b$  is a  $\frac{3}{2}n\ell$ -dimensional vector. In other words, each group is a scaled clone of the instance in the previous section. Choosing again  $y := (1/2, \dots, 1/2) \in \mathbb{R}^{n\ell}$  as fractional solution, we have  $Cy = b$ . Note that allowing multiplicities is just for notational convenience and does not make the problem setting more general. Since the total number of items is still bounded by a polynomial in  $n$  (more precisely  $\mathbf{1}^T b \leq O(n^2)$ ), each item  $i$  could still be replaced by  $b_i$  items of multiplicity 1. Let  $s_i^j := \frac{1}{3} \cdot (\frac{1}{2})^{j-1} - i \cdot \varepsilon$  the size of the  $i$ th item in group  $j$  for  $\varepsilon := \frac{1}{12n^3}$ . Note that the size contribution of each item type is  $2^{j-1} \cdot s_i^j \in [\frac{1}{3}, \frac{1}{3} - \frac{1}{n}]$ . Abbreviate the number of different item types by  $m := \ell \cdot \frac{3}{2}n$ .

**THEOREM 12.** *There is no algorithm for bin packing which is based on rounding up items and achieves an additive integrality gap of  $o(\log^2 n)$ .*

PROOF. Let  $(x, D)$  be arbitrary with  $\text{supp}(x) \subseteq \text{supp}(y)$  such that  $x$  is integral and covers  $[m] \setminus D$  (considering  $D$  now as a multiset). Assume for the sake of contradiction that

$$\mathbf{1}^T x + 2 \sum_{i \in D} s_i \geq \mathbf{1}^T y + o(\log^2 n).$$

As in Theorem 11, we can assume that  $|\mathbf{1}^T x - \mathbf{1}^T y| \leq \frac{1}{10000} \log^2 n$ . First, observe that the bins in  $y$  are packed pretty tight, i.e.  $|\mathbf{1}^T y - s([m])| \leq 1$ . If an item  $i$  is covered by a slot for a larger item  $i'$ , then this causes a *waste* of  $s_{i'} - s_i$ , which is not anymore available for any other item. The additive gap is defined as  $\mathbf{1}^T x + 2s(D) - \mathbf{1}^T y \geq s([m]/D) + \text{waste} + 2s(D) - s([m]) - 1 = \text{waste} + s(D) - 1$ . Thus both, the waste and the size of the discarded items  $s(D)$  must be bounded by  $\frac{1}{10000} \log^2 n$ .

Observe that the items in group  $j - 1$  are at least a factor  $3/2$  larger than the items in group  $j$ . In other words, every item  $i$  from group  $j$  which is mapped to group  $1, \dots, j - 1$  generates a waste of at least  $\frac{1}{2}s_i^j$ . Thus the total size of items which are mapped to the slot of an item in a larger group is bounded by  $\frac{1}{5000} \log^2 n$ . Thus there lies no harm in discarding these items as well — let  $D'$  be the union of such items and  $D$ . Then  $s(D') \leq s(D) + \frac{1}{5000} \log^2 n \leq \frac{1}{3000} \log^2 n$ .

For group  $j$ , let  $D'_j \subseteq D'$  be the discarded items in the  $j$ th group and let  $x^j$  ( $y^j$ , resp.) be the vector  $x$  ( $y$ , resp.), restricted to the patterns corresponding to group  $j$ . In other words,  $x = (x^1, \dots, x^\ell)$  and  $y = (y^1, \dots, y^\ell)$ . By  $x_i^j \in \mathbb{Z}_{\geq 0}$  we denote the entry belonging to column  $(\mathbf{0}, \dots, \mathbf{0}, 2^{j-1}B^i, \mathbf{0}, \dots, \mathbf{0})$  in  $C$ . Pick  $j \in \{1, \dots, \ell\}$  uniformly at random, then  $E[|\mathbf{1}^T x^j - \mathbf{1}^T y^j|] \leq \frac{1}{10000} \log n$  and  $E[s(D'_j)] \leq \frac{1}{3000} \log n$ . By Markov's inequality, there must be an index  $j$ , such that  $|\mathbf{1}^T x^j - \mathbf{1}^T y^j| \leq \frac{1}{1000} \log n$  and  $s(D'_j) \leq \frac{\log n}{2000}$ . Recall that  $|D'_j| \leq 4 \cdot 2^{j-1} s(D'_j) \leq 2^{j-1} \frac{\log n}{500}$ . Since  $x^j$  covers all items in group  $j$  (without  $D'_j$ ), we obtain

$$\sum_{i'=1}^i 2^{j-1} B_{i'} x^j \geq i \cdot 2^{j-1} - 2^{j-1} \frac{\log n}{500} \quad \forall i = 1, \dots, \frac{3}{2}n$$

After division by  $2^{j-1}$ , this implies Condition (2), which leads to a contradiction.

The claim follows since the number of items counted with multiplicity is bounded by  $O(n^2)$ , thus  $\log^2(\mathbf{1}^T b) = \Theta(\log^2 n)$ .  $\square$

REMARK 1. Note that the additive integrality gap for the constructed instance is still small, once arbitrary patterns may be used. For example a First Fit Decreasing assignment will produce a solution of cost exactly  $OPT_f$ . This can be partly fixed by slightly increasing the item sizes. For the sake of simplicity consider the construction in Section 6.1 and observe that the used patterns are still feasible if the items corresponding to the first permutation have sizes in the range  $[\frac{1}{3} + 10\delta, \frac{1}{3} + 11\delta]$  and the items corresponding to the 2nd and 3rd permutation have item sizes in  $[\frac{1}{3} - 7\delta, \frac{1}{3} - 6\delta]$  (for a small constant  $\delta > 0$ ). Then a First Fit Decreasing approach will produce a  $\Omega(n)$  additive gap.

## REFERENCES

W. Banaszczyk. Balancing vectors and Gaussian measures of  $n$ -dimensional convex bodies. *Random Structures Algorithms*, 12(4):351–360, 1998.

- N. Bansal. Constructive algorithms for discrepancy minimization. *CoRR*, abs/1002.2259, 2010. informal publication.
- J. Beck and T. Fiala. “Integer-making” theorems. *Discrete Appl. Math.*, 3(1):1–8, 1981.
- J. Beck and V. Sós. Discrepancy theory. In *Handbook of combinatorics, Vol. 1, 2*, pages 1405–1446. Elsevier, Amsterdam, 1995.
- G. Bohus. On the discrepancy of 3 permutations. *Random Structures Algorithms*, 1(2):215–220, 1990.
- Moses Charikar, Alantha Newman, and Aleksandar Nikolov. Tight hardness results for minimizing discrepancy. In *SODA*, pages 1607–1614, 2011.
- R. Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, 1997.
- G. Dósa. The tight bound of first fit decreasing bin-packing algorithm is  $\text{FFD}(I) \leq 11/9\text{OPT}(I) + 6/9$ . In Bo Chen, Mike Paterson, and Guochuan Zhang, editors, *Combinatorics, Algorithms, Probabilistic and Experimental Methodologies, First International Symposium, ESCAPE 2007, Hangzhou, China, April 7-9, 2007, Revised Selected Papers*, volume 4614 of *Lecture Notes in Computer Science*, pages 1–11. Springer, 2007.
- K. Eisemann. The trim problem. *Management Science*, 3(3):279–284, 1957.
- Friedrich Eisenbrand, Dömötör Pálvölgyi, and Thomas Rothvoß. Bin packing via discrepancy of permutations. In *SODA*, pages 476–481, 2011.
- W. Fernandez de la Vega and G. S. Lueker. Bin packing can be solved within  $1 + \epsilon$  in linear time. *Combinatorica*, 1(4):349–355, 1981.
- M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, New York, New York, 1979.
- P. C. Gilmore and R. E. Gomory. A linear programming approach to the cutting-stock problem. *Operations Research*, 9:849–859, 1961.
- M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981.
- K. Jansen and R. Solis-Oba. An  $\text{opt} + 1$  algorithm for the cutting stock problem with constant number of object lengths. In Friedrich Eisenbrand and F. Bruce Shepherd, editors, *Proceedings of the 14th International Conference on Integer Programming and Combinatorial Optimization, IPCO 2010*, pages 438–449, 2010.
- D. S. Johnson. *Near-optimal bin packing algorithms*. PhD thesis, MIT, Cambridge, MA, 1973.
- D. S. Johnson, A. Demers, J. D. Ullman, M. R. Garey, and R. L. Graham. Worst-case performance bounds for simple one-dimensional packing algorithms. *SIAM Journal on Computing*, 3(4):299–325, 1974.
- N. Karmarkar and R. M. Karp. An efficient approximation scheme for the one-dimensional bin-packing problem. In *23rd annual symposium on foundations of computer science (Chicago, Ill., 1982)*, pages 312–320. IEEE, New York, 1982.
- L. Lovász, J. Spencer, and K. Vesztegombi. Discrepancy of set-systems and matrices. *European J. Combin.*, 7(2):151–160, 1986.
- J. Matoušek. *Geometric discrepancy*, volume 18 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1999. An illustrated guide.
- A. Newman and A. Nikolov. A counterexample to beck’s conjecture on the discrepancy of three permutations. *CoRR*, abs/1104.2922, 2011.
- G. Scheithauer and J. Terno. Theoretical investigations on the modified integer round-up property for the one-dimensional cutting stock problem. *Operations Research Letters*, 20(2):93–100, 1997.
- A. Sebő and G. Shmonin. Proof of the modified integer round-up conjecture for bin packing in dimension 7. Personal communication, 2009.
- J. Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–706, 1985.
- J. H. Spencer, A. Srinivasan, and P. Tetali. The discrepancy of permutation families. Unpublished manuscript.

A. Srinivasan. Improving the discrepancy bound for sparse matrices: Better approximations for sparse lattice approximation problems. In *Proceedings of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA'97 (New Orleans, Louisiana, January 5-7, 1997)*, pages 692–701, Philadelphia, PA, 1997. ACM SIGACT, SIAM, Society for Industrial and Applied Mathematics.

D. Williamson. Lecture notes on approximation algorithms, Fall 1998. IBM Research Report, 1998. <http://legacy.orie.cornell.edu/~dpw/cornell.ps> .