Bin Packing via Discrepancy of Permutations

Friedrich Eisenbrand*, Dömötör Pálvölgyi†, Thomas Rothvoß‡

Abstract
A well-studied special case of bin packing is the 3-partition problem, where $n$ items of size $s_1, \ldots, s_n \in [0,1]$ 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 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 [14, 15] has shown that the First Fit algorithm requires at most $1.7 \cdot \text{OPT} + 1$ bins and that First Fit Decreasing yields a solution with $1.5 \cdot \text{OPT} + 4$ bins (see [7] for a tight bound of $1.5 \cdot \text{OPT} + 2$). An important step forward was made by Fernandez de la Vega and Luecker [9] 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 [16] proposed an approximation algorithm for bin packing that can be analyzed to yield a solution using at most $\text{OPT} + O(\log^2 n)$ bins. This seminal procedure is based on the Gilmore-Gomory LP relaxation [11, 8]:

$$\min \sum_{p \in P} x_p$$

$$\sum_{p \in P} p^T x_p \geq 1$$

$$\geq 0 \ \forall p \in P$$

Here $1 = (1, \ldots, 1)^T$ denotes the all ones vector and $P = \{p \in \{0,1\}^n : s^T p \leq 1\}$ is the set of all feasible patterns, i.e. every vector in $P$ denotes a feasible way to pack one bin. Let $\text{OPT}$ and $\text{OPT}_f$ 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 basis solution $x$ with $1^T x \leq \text{OPT}_f + \delta$ in time polynomial in $n$ and $1/\delta$ [16] using the Grötschel-Lovász-Schrijver variant of the Ellipsoid method [12].

The procedure of Karmarkar and Karp [16] yields an additive integrality gap of $O(\log^2 n)$, i.e. $\text{OPT} \leq \text{OPT}_f + O(\log^2 n)$, see also [24]. This corresponds to an asymptotic FPTAS for bin packing. The authors in [19] conjecture that even $\text{OPT} \leq \text{OPT}_f + 1$ holds and this even if one replaces the right-hand-side $1$ by any other positive integral vector $b$. This Modified Integer Round-up Conjecture was proven by Sebő and Shmonin [20] if the number of different item sizes is at most 7. We would like to mention that Jansen and Solis-Oba [13] recently provided an $\text{OPT} + 1$ 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 $n$ items of size $s_1, s_2, s_3$ with $\sum_{i=1}^3 s_i = n$ have to be packed. It is strongly NP-hard to distinguish between $\text{OPT} \leq n$ and $\text{OPT} \geq n + 1$ [10]. No stronger hardness result is known for general bin packing. A closer look into [16] reveals that, with the restriction $s_1 > \frac{1}{3}$, the Karmarkar-Karp algorithm uses $\text{OPT}_f + O(\log n)$ bins$^2$.

$^*$EPFL, Lausanne, Switzerland, friedrich.eisenbrand@epfl.ch. Supported by the Swiss National Science Foundation (SNF).

†Eötvös Loránd University (ELTE), Budapest, Hungary, dom@cs.elte.hu

‡EPFL, Lausanne, Switzerland, thomas.rothvoss@epfl.ch. Supported by the German Research Foundation (DFG) within the Priority Program 1307 “Algorithm Engineering”.

$^2$An asymptotic fully polynomial time approximation scheme (AFPTAS) is an approximation algorithm that produces solutions of cost at most $(1 + \epsilon) \text{OPT} + p(1/\epsilon)$ in time polynomial in $n$ and $1/\epsilon$, where also $p$ must be a polynomial.

$^2$The geometric grouping procedure (Lemma 5 in [16]) discards
Discrepancy theory. Let $[n] := \{1, \ldots, n\}$ and consider a set system $\mathcal{S} \subseteq 2^{[n]}$ over the ground set $[n]$. A coloring is a mapping $\chi : [n] \to \{\pm 1\}$. 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
\[
\text{disc}(\mathcal{S}) = \min_{\chi : [n] \to \{\pm 1\}} \max_{S \in \mathcal{S}} |\chi(S)|.
\]
where $\chi(S) = \sum_{i \in S} \chi(i)$. A random coloring provides an easy bound of $\text{disc}(\mathcal{S}) \leq O(\sqrt{n \log |\mathcal{S}|})$ [18]. The famous “Six Standard Deviations suffice” result of Spencer [21] 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 [23] 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 improved by Banaszczyk [1] to $\text{disc}(\mathcal{S}) \leq O(\sqrt{n \log (2 |\mathcal{S}|/n)})$.

The following conjecture is coined three-permutations-conjecture or simply Beck’s conjecture (see Problem 1.9 in [4]):

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 $O(1)$.

A set of permutations $\pi_1, \ldots, \pi_k : [n] \to [n]$ induces a set-system
\[
\mathcal{S} = \{\{\pi_1(i), \ldots, \pi_k(j)\} : j = 1, \ldots, n; \ i = 1, \ldots, k\}.
\]
If we denote the maximum discrepancy of such a set-system induced by $k$ permutations over $n$ symbols as $D^\permn{\text{perm}}(n)$, then Beck’s conjecture can be rephrased as $D^\permn{\text{perm}}(n) = O(1)$.

So far the best known bound on $D^\permn{\text{perm}}(n)$ is $O(\log n)$ and more generally $D^\permn{\text{perm}}(n)$ can be bounded by $O(k \log n)$ [5] and by $O(\sqrt{k \log n})$ [23, 22] using the so-called entropy method.

The first player chooses a fractional vector $x$, and the second player non-adversely picks another fractional vector $y$. The integrality gap of the linear program (LP) restricted to 3-partition instances is bounded by an additive constant $c$.

Our contribution. The main result of this paper is a proof of the following theorem.

**Theorem 1.1.** If Beck’s conjecture holds, then the integrality gap of the linear program (LP) restricted to 3-partition instances is bounded by an additive constant $c$.

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

The proof of Theorem 1.1 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 4.1). 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. And finally, 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.

2 Preliminaries

We first review some further necessary preliminaries on discrepancy theory. We refer to [18] 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
\[
\text{disc}(A) := \min_{x \in [0, 1]^n} \|A(x - 1/2 \cdot 1)\|_\infty,
\]
see, e.g. [18]. 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 $\text{disc}(A) = 1/2\text{disc}(\mathcal{S})$, hence this notation is consistent — apart from the $1/2$ factor.

The linear discrepancy of a matrix $A \in \mathbb{R}^{m \times n}$ is defined as
\[
\text{lin}\text{disc}(A) := \max_{y \in [0, 1]^n} \min_{x \in [0, 1]^n} \|Ax - Ay\|_\infty.
\]
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 $y := (1/2, \ldots, 1/2)$. 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 [17] even an upper bound for $\text{herdisc}(A)$ is proved (see again [18] for a recent description).

**Theorem 2.1. (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 $0, \ldots, k$. In other words $A \in \{0, \ldots, k\}^{m \times n}$ and $A_{ij} \leq \ldots \leq A_{mj}$ for any column $j$. We denote the maximum linear discrepancy of such matrices by

$$D_k^{\text{mon}}(n) := \max_{A \in \{0, \ldots, k\}^{n \times n}} \text{lindisc}(A).$$

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

**Theorem 3.1.** Consider the linear program (LP) and suppose that the item sizes satisfy $s_1, \ldots, s_n > \frac{1}{k+1}$. 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 $s_1 \geq \ldots \geq s_n$. Let $y$ be any optimum basic solution of (LP) and let $p_1, \ldots, p_m$ 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 $4$ that $m = n$.

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

We attach a row $A_{n+1} := (k, \ldots, k)$ 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 $s_i$.

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

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

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

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

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

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.

![Figure 1: The bipartite graph in the proof of Theorem 3.1](image)

To this end, consider a bipartite graph with nodes $V = \{v_1, \ldots, v_n\}$ on the left, representing the items. The nodes on the right are the set $U = \{u_1, \ldots, u_n\}$, 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. [6], 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 $V' = \{1, \ldots, i\}$. For such a $V'$ one has
$$\sum_{u_j \in N(V')} b_j = D_k^{\text{mon}}(n) + \sum_{j=1}^i B_j x = D_k^{\text{mon}}(n) + A_i x \geq A_{ij} y = i$$
and the claim follows.

## 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 2.1 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 $B_{\pi(1)} \geq B_{\pi(2)} \geq \ldots \geq B_{\pi(n)}$ 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 $B^1, \ldots, B^k$. 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 4.1.** 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 2.1, 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 $\text{disc}(B) \leq \frac{1}{2} \cdot D_k^{\text{perm}}(n)$. Of course, $B$ itself is again $k$-monotone.

Let $B^\ell$ also be a $m \times n'$ matrix, defined by
$$B^\ell_{ij} := \begin{cases} 1 & \text{if } B_{ij} \geq \ell \\ 0, & \text{otherwise.} \end{cases}$$

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

Let $\chi : [n'] \to \{\pm 1\}$ be the coloring that has discrepancy at most $D_k^{\text{perm}}(n)$ with respect to all permutations $\pi_1, \ldots, \pi_k$. In particular $|B^\ell_i \chi| \leq D_k^{\text{perm}}(n)$, 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 3.1 and Theorem 4.1, we conclude

**Corollary 4.1.** Given any bin packing instance with $n$ items of size bigger than $\frac{1}{\sqrt{n}}$, one has $\text{OPT} \leq \text{OPT} + 2k \cdot D_k^{\text{perm}}(n)$.

In particular, this proves Theorem 1.1, 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 4.2.** 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 $\pi_1, \ldots, \pi_k$ on $[n]$, there is a $kn \times k$-monotone matrix $C$ with $\text{disc}(\pi_1, \ldots, \pi_k) \leq 4 \cdot \text{disc}(C)$. Let $\Sigma \in \{1, \ldots, n\}^{kn}$ be the string which we obtain by concatenating the $k$ permutations. That means $\Sigma = (\pi_1(1), \ldots, \pi_1(n), \ldots, \pi_k(1), \ldots, \pi_k(n))$. Let $C$ the matrix where $C_{ij}$ is the number of appearances of $j \in \{1, \ldots, n\}$ among the first $i \in \{1, \ldots, kn\}$ entries of $\Sigma$. By definition, $C$ is $k$-monotone, in fact it is the “same” $k$-monotone matrix as in the previous proof.
Choose $y := (\frac{1}{2}, \ldots, \frac{1}{2})$ to have $Cy = (\frac{1}{2}, 1, \ldots, \frac{k}{2})$. Let $x \in \{0,1\}^n$ be a vector with $\|Cx-Cy\|_\infty \leq \text{disc}(C)$. Consider the coloring $\chi: [n] \to \{\pm 1\}$ 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 $S := \{\pi(1), \ldots, \pi(l)\}$. Let $r = C_{(i-1)n+i} \in \{1, i\}^n$ be the row of $C$ that corresponds to this prefix. With these notations we have

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

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

From Theorem 4.1 and Theorem 4.2, we obtain that the following conjecture is equivalent to the Three-Permutations-Conjecture.

**Conjecture 4.1.** There exists a constant $c > 0$, such that for any $3$-monotone matrix $A \in \mathbb{Z}^{m \times n}$ one has

$$\text{lindisc}(A) \leq c.$$

### 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 [22, 23] together with Theorem 4.1 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 5.1.** 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 $n = m = 1$, $\text{lindisc}(A) \leq \frac{k}{2}$, 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 then the number of constraints, then $y$ is not a basic solution of the system

$$Ay = b, \quad 0 \leq y_i \leq 1 \quad \forall i = 1, \ldots, m.$$

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 $a_1, \ldots, a_n$ be the rows of $A$ and let $d(j) := \|a_{j+1} - a_j\|_1$ for $j = 1, \ldots, n-1$, 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

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

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 $J \subseteq \{1, \ldots, n-1\}$ 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 $A' = b'$ be the subsystem of $n' = \frac{n}{2}$ 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

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

Now consider any $j \in \{1, \ldots, n\}$. If $j \notin J$, then row $j$ still appeared in $A' = b'$, 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

$$|a_{j+1}^T x - a_{j+1}^T y| \leq \left|a_{j+1}^T (x-y)\right| \leq \left(|a_{j+1}^T (x-y)| + |a_{j+1}^T (x-y)|\right) \leq 5k \cdot \log_2(2n).$$

**References**


