Balancing Vectors in Any Norm

Aleksandar (Sasho) Nikolov

University of Toronto

Based on joint work with

Daniel Dadush, Kunal Talwar, and Nicole Tomczak-Jaegermann

Outline

Discrepancy

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

$$\mathrm{disc}(U, \|\cdot\|_\infty) = \min_{\varepsilon \in \{\pm 1\}^N} \|U\varepsilon\|_\infty$$

Discrepancy

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

$$\text{disc}(U, \|\cdot\|_\infty) = \min_{\varepsilon \in \{\pm 1\}^N} \|U\varepsilon\|_\infty$$

Natural to consider arbitrary norms: any norm can be written as $\|U \cdot\|_\infty$.

Basic Bounds

- [Spencer, 1985; Gluskin, 1989]: For any matrix $U \in \{0,1\}^{n \times N}$, $\mathrm{disc}(U) \lesssim \sqrt{n}$

Basic Bounds

- [Spencer, 1985; Gluskin, 1989]: For any matrix $U \in \{0,1\}^{n \times N}$, $\mathrm{disc}(U) \lesssim \sqrt{n}$
- *Implied by*: For any $u_1, \ldots, u_N \in B_\infty^n = [-1,1]^n$, there exist $\varepsilon_1, \ldots, \varepsilon_N \in \{-1,+1\}$ s.t. $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty \lesssim \sqrt{n}$.

Basic Bounds

- [Spencer, 1985; Gluskin, 1989]: For any matrix $U \in \{0,1\}^{n\times N}$, $\mathrm{disc}(U) \lesssim \sqrt{n}$
- *Implied by*: For any $u_1, \ldots, u_N \in B_\infty^n = [-1,1]^n$, there exist $\varepsilon_1, \ldots, \varepsilon_N \in \{-1,+1\}$ s.t. $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty \lesssim \sqrt{n}$.
- [Beck and Fiala, 1981]: For any matrix $U \in \{0,1\}^{n\times N}$ with at most t ones per column, $\mathrm{disc}(U) \le 2t - 1$

Basic Bounds

- [Spencer, 1985; Gluskin, 1989]: For any matrix $U \in \{0,1\}^{n \times N}$, $\mathrm{disc}(U) \lesssim \sqrt{n}$
- *Implied by*: For any $u_1, \ldots, u_N \in B_\infty^n = [-1,1]^n$, there exist $\varepsilon_1, \ldots, \varepsilon_N \in \{-1,+1\}$ s.t. $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty \lesssim \sqrt{n}$.
- [Beck and Fiala, 1981]: For any matrix $U \in \{0,1\}^{n \times N}$ with at most t ones per column, $\mathrm{disc}(U) \leq 2t - 1$
- *Implied by*: For any $u_1, \ldots, u_N \in B_1^n$, there exist $\varepsilon_1, \ldots, \varepsilon_N \in \{-1,+1\}$ s.t. $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty < 2$.

Basic Bounds

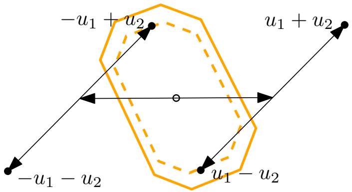
- [Spencer, 1985; Gluskin, 1989]: For any matrix $U \in \{0,1\}^{n \times N}$, $\mathrm{disc}(U) \lesssim \sqrt{n}$
- *Implied by*: For any $u_1, \ldots, u_N \in B_\infty^n = [-1,1]^n$, there exist $\varepsilon_1, \ldots, \varepsilon_N \in \{-1,+1\}$ s.t. $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty \lesssim \sqrt{n}$.
- [Beck and Fiala, 1981]: For any matrix $U \in \{0,1\}^{n \times N}$ with at most t ones per column, $\mathrm{disc}(U) \le 2t - 1$
- *Implied by*: For any $u_1, \ldots, u_N \in B_1^n$, there exist $\varepsilon_1, \ldots, \varepsilon_N \in \{-1,+1\}$ s.t. $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty < 2$.

Most combinatorial discrepancy bounds are implied by geometric vector balancing arguments.

The Vector Balancing Problem

Given $u_1, \ldots, u_N \in \mathbb{R}^n$, and symmetric convex body $K \subset \mathbb{R}^n$ ($K = -K$), find the smallest t such that

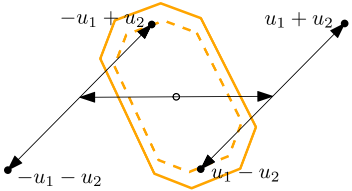
$$\exists\ \varepsilon_1, \ldots, \varepsilon_N \in \{-1, +1\} : \varepsilon_1 u_1 + \ldots + \varepsilon_N u_N \in tK$$



The Vector Balancing Problem

Given $u_1, \ldots, u_N \in \mathbb{R}^n$, and symmetric convex body $K \subset \mathbb{R}^n$ $(K = -K)$, find the smallest t such that

$$\exists\ \varepsilon_1, \ldots, \varepsilon_N \in \{-1, +1\} : \varepsilon_1 u_1 + \ldots + \varepsilon_N u_N \in tK$$

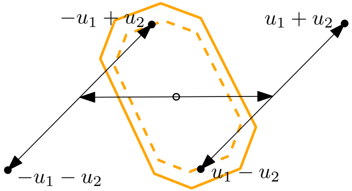


Minkowski Norm: $\|x\|_K = \inf\{t \geq 0 : x \in tK\}$; $t = \mathrm{disc}((u_i)_{i=1}^N, \|\cdot\|_K)$.

The Vector Balancing Problem

Given $u_1, \dots, u_N \in \mathbb{R}^n$, and symmetric convex body $K \subset \mathbb{R}^n$ ($K = -K$), find the smallest t such that

$$\exists\ \varepsilon_1, \dots, \varepsilon_N \in \{-1, +1\} : \varepsilon_1 u_1 + \dots + \varepsilon_N u_N \in tK$$



Minkowski Norm: $\|x\|_K = \inf\{t \geq 0 : x \in tK\}$; $t = \text{disc}((u_i)_{i=1}^N, \|\cdot\|_K)$.

Vector Balancing Constant: worst case over sequences in C

$$\text{vb}(C, K) = \sup \left\{ \text{disc}(U, \|\cdot\|_K) : N \in \mathbb{N}, u_1, \dots, u_N \in C, U = (u_i)_{i=1}^N \right\}$$

Questions and Prior Results

- [Dvoretzky, 1963] "What can be said" about $\mathrm{vb}(K, K)$?
- [Bárány and Grinberg, 1981] $\mathrm{vb}(K, K) \le n$ for all K.

Questions and Prior Results

- [Dvoretzky, 1963] "What can be said" about $\mathrm{vb}(K, K)$?
- [Bárány and Grinberg, 1981] $\mathrm{vb}(K, K) \leq n$ for all K.
- [Spencer, 1985; Gluskin, 1989] $\mathrm{vb}(B_\infty^n, B_\infty^n) \lesssim \sqrt{n}$
- [Beck and Fiala, 1981] $\mathrm{vb}(B_1^n, B_\infty^n) < 2$

Questions and Prior Results

- [Dvoretzky, 1963] "What can be said" about $\mathrm{vb}(K, K)$?
- [Bárány and Grinberg, 1981] $\mathrm{vb}(K, K) \leq n$ for all K.
- [Spencer, 1985; Gluskin, 1989] $\mathrm{vb}(B_\infty^n, B_\infty^n) \lesssim \sqrt{n}$
- [Beck and Fiala, 1981] $\mathrm{vb}(B_1^n, B_\infty^n) < 2$
- [Banaszczyk, 1998] $vb(B_2^n, K) \leq 5$ if K has Gaussian measure $\gamma_n(K) \geq \frac{1}{2}$
- *Komlós Problem*: Prove or disprove $\mathrm{vb}(B_2^n, B_\infty^n) \lesssim 1$.
    - Banaszczyk's theorem implies $\mathrm{vb}(B_2^n, B_\infty^n) \lesssim \sqrt{\log 2n}$.

Vector Balancing and Rounding

For any $w \in [0,1]^N$, any $U = (u_i)_{i=1}^N$, $u_i \in C$, and any symmetric convex K, there exists a $x \in \{0,1\}^N$ such that

$$\|Ux - Uw\|_K \leq \mathrm{vb}(C, K).$$

Our Results

We initiate a systematic study of *upper* and *lower bounds* on $\mathrm{vb}(C, K)$ and its computational complexity:

Our Results

We initiate a systematic study of *upper* and *lower bounds* on $\mathrm{vb}(C, K)$ and its computational complexity:

- A natural volumetric lower bound on $\mathrm{vb}(C, K)$ is tight up to a $O(\log n)$ factor.
  - The proof implies an efficient algorithm to compute $\varepsilon \in \{-1, 1\}^N$ given $u_1, \ldots, u_N \in C$, so that $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_K \lesssim (1 + \log n)\, \mathrm{vb}(C, K)$.
  - Also rounding version.

Our Results

We initiate a systematic study of *upper* and *lower bounds* on $\mathrm{vb}(C,K)$ and its computational complexity:

- A natural volumetric lower bound on $\mathrm{vb}(C,K)$ is tight up to a $O(\log n)$ factor.
  - The proof implies an efficient algorithm to compute $\varepsilon \in \{-1,1\}^N$ given $u_1,\ldots,u_N \in C$, so that $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_K \lesssim (1+\log n)\,\mathrm{vb}(C,K)$.
  - Also rounding version.
- An efficiently computable upper bound on $\mathrm{vb}(C,K)$ is tight up to factors polynomial in $\log n$.
  - Based on an optimal application of Banaszczyks' theorem.
  - Implies an efficient approximation algorithm for $\mathrm{vb}(C,K)$.

Our Results

We initiate a systematic study of *upper* and *lower bounds* on $\mathrm{vb}(C, K)$ and its computational complexity:

- A natural volumetric lower bound on $\mathrm{vb}(C, K)$ is tight up to a $O(\log n)$ factor.
  - The proof implies an efficient algorithm to compute $\varepsilon \in \{-1, 1\}^N$ given $u_1, \ldots, u_N \in C$, so that $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_K \lesssim (1 + \log n)\,\mathrm{vb}(C, K)$.
  - Also rounding version.
- An efficiently computable upper bound on $\mathrm{vb}(C, K)$ is tight up to factors polynomial in $\log n$.
  - Based on an optimal application of Banaszczyks' theorem.
  - Implies an efficient approximation algorithm for $\mathrm{vb}(C, K)$.
- The results extend to hereditary discrepancy with respect to arbitrary norms.

Our Results

We initiate a systematic study of *upper* and *lower bounds* on $\mathrm{vb}(C, K)$ and its computational complexity:

- A natural volumetric lower bound on $\mathrm{vb}(C, K)$ is tight up to a $O(\log n)$ factor.
    - The proof implies an efficient algorithm to compute $\varepsilon \in \{-1, 1\}^N$ given $u_1, \ldots, u_N \in C$, so that $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_K \lesssim (1 + \log n)\,\mathrm{vb}(C, K)$.
    - Also rounding version.
- An efficiently computable upper bound on $\mathrm{vb}(C, K)$ is tight up to factors polynomial in $\log n$.
    - Based on an optimal application of Banaszczyks' theorem.
    - Implies an efficient approximation algorithm for $\mathrm{vb}(C, K)$.
- The results extend to hereditary discrepancy with respect to arbitrary norms.

Prior work [Bansal, 2010; Nikolov and Talwar, 2015] implies bounds which deteriorate with the number of facets of K.

Outline

Hereditary Discrepancy

Issue: $\mathrm{disc}(U, K) = \mathrm{disc}(U, \|\cdot\|_K)$ is

- not robust to slight changes in U (e.g. repeat each column)
- hard to approximate [Charikar, Newman, and Nikolov, 2011]

Hereditary Discrepancy

Issue: $\mathrm{disc}(U, K) = \mathrm{disc}(U, \|\cdot\|_K)$ is

- not robust to slight changes in U (e.g. repeat each column)
- hard to approximate [Charikar, Newman, and Nikolov, 2011]

$\mathrm{vb}(C, K)$ is more robust, but not about a specific matrix U.

Hereditary Discrepancy

Issue: $\mathrm{disc}(U, K) = \mathrm{disc}(U, \|\cdot\|_K)$ is

- not robust to slight changes in U (e.g. repeat each column)
- hard to approximate [Charikar, Newman, and Nikolov, 2011]

$\mathrm{vb}(C, K)$ is more robust, but not about a specific matrix U.

Hereditary discrepancy is a robust analog of discrepancy:

$$\mathrm{hd}(U, K) = \max_{S \subseteq [N]} \mathrm{disc}(U_S, K),$$

where $U_S = (u_i)_{i \in S}$ is the submatrix of U indexed by S.

Hereditary Discrepancy

Issue: $\text{disc}(U, K) = \text{disc}(U, \|\cdot\|_K)$ is

- not robust to slight changes in U (e.g. repeat each column)
- hard to approximate [Charikar, Newman, and Nikolov, 2011]

$\text{vb}(C, K)$ is more robust, but not about a specific matrix U.

Hereditary discrepancy is a robust analog of discrepancy:

$$\text{hd}(U, K) = \max_{S \subseteq [N]} \text{disc}(U_S, K),$$

where $U_S = (u_i)_{i \in S}$ is the submatrix of U indexed by S.

Observation:

$$\text{vb}(C, K) = \sup \left\{ \text{hd}(U, K) : N \in \mathbb{N}, u_1, \ldots, u_N \in C, U = (u_i)_{i=1}^N \right\}.$$

The Volume Lower Bound

Define $L = \{x \in \mathbb{R}^N : Ux \in K\}$: the set of "good x".

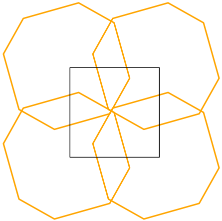
- $\mathrm{disc}(U, K) = \min\{t : tL \cap \{-1, 1\}^N \neq \emptyset\}$.

The Volume Lower Bound

Define $L = \{x \in \mathbb{R}^N : Ux \in K\}$: the set of "good x".

- $\mathrm{disc}(U, K) = \min\{t : tL \cap \{-1, 1\}^N \neq \emptyset\}$.

[Lovász, Spencer, and Vesztergombi, 1986]:
If $t = \mathrm{hd}(U, K)$, then $[0, 1]^N \subseteq \bigcup_{x \in \{0,1\}^N} (x + tL)$.

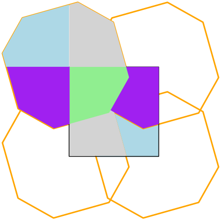


The Volume Lower Bound

Define $L = \{x \in \mathbb{R}^N : Ux \in K\}$: the set of "good x".

- $\mathrm{disc}(U, K) = \min\{t : tL \cap \{-1, 1\}^N \neq \emptyset\}$.

[Lovász, Spencer, and Vesztergombi, 1986]:
If $t = \mathrm{hd}(U, K)$, then $[0, 1]^N \subseteq \bigcup_{x \in \{0,1\}^N} (x + tL)$.



[Banaszczyk, 1993]:

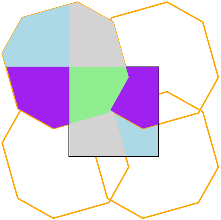
$$1 = \mathrm{vol}([0, 1]^N) \geq \mathrm{vol}(tL) = t^N \mathrm{vol}(L)$$

The Volume Lower Bound

Define $L = \{x \in \mathbb{R}^N : Ux \in K\}$: the set of "good x".

- $\mathrm{disc}(U, K) = \min\{t : tL \cap \{-1, 1\}^N \neq \emptyset\}$.

[Lovász, Spencer, and Vesztergombi, 1986]:
If $t = \mathrm{hd}(U, K)$, then $[0, 1]^N \subseteq \bigcup_{x \in \{0,1\}^N} (x + tL)$.



[Banaszczyk, 1993]:

$$1 = \mathrm{vol}([0, 1]^N) \geq \mathrm{vol}(tL) = t^N \mathrm{vol}(L) \iff \mathrm{hd}(U, K) \geq \mathrm{vol}(L)^{-1/N}.$$

A Hereditary Volume Lower Bound

A simple strengthening:

$$\mathrm{hd}(U, K) \geq \mathrm{volLB}(U, K) = \max_{S \subseteq [N]} \mathrm{vol}(\{x \in \mathbb{R}^S : U_S x \in K\})^{-1/|S|}.$$

A Hereditary Volume Lower Bound

A simple strengthening:

$$\mathrm{hd}(U, K) \geq \mathrm{volLB}(U, K) = \max_{S \subseteq [N]} \mathrm{vol}(\{x \in \mathbb{R}^S : U_S x \in K\})^{-1/|S|}.$$

Lower Bound on $\mathrm{vb}(C, K)$:

$$\mathrm{vb}(C, K) \geq \mathrm{volLB}(C, K) = \sup\left\{\mathrm{volLB}((u_i)_{i=1}^N, K) : u_1, \ldots, u_N \in C\right\}.$$

A Hereditary Volume Lower Bound

A simple strengthening:

$$\mathrm{hd}(U,K) \geq \mathrm{volLB}(U,K) = \max_{S\subseteq[N]} \mathrm{vol}(\{x \in \mathbb{R}^S : U_S x \in K\})^{-1/|S|}.$$

Lower Bound on $\mathrm{vb}(C,K)$:

$$\mathrm{vb}(C,K) \geq \mathrm{volLB}(C,K) = \sup\left\{\mathrm{volLB}((u_i)_{i=1}^N, K) : u_1,\dots,u_N \in C\right\}.$$

Theorem

For any $n \times N$ matrix U, and any symmetric convex $C, K \subset \mathbb{R}^n$,

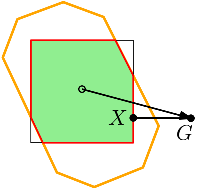
$$\mathrm{volLB}(U,K) \leq \mathrm{hd}(U,K) \lesssim (1+\log n)\cdot \mathrm{volLB}(U,K)$$

$$\mathrm{volLB}(C,K) \leq \mathrm{vb}(C,K) \lesssim (1+\log n)\cdot \mathrm{volLB}(C,K)$$

Rothvoß's Algorithm

Algorithm [Rothvoß, 2014]: given $K \subset \mathbb{R}^n$,

1. Sample a standard Gaussian $G \sim N(0, I_n)$;
2. Output
   $X = \arg\min\{\|x - G\|_2^2 : x \in K \cap [-1,1]^n\}$.



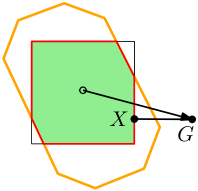
Goal: $|\{i : X_i \in \{-1,+1\}\}| \geq \alpha n$ for a constant α.
(X is a *partial coloring*.)

Intuition: If K is "big enough," then in an average direction $\partial[-1,1]^n$ is closer to the origin than ∂K and is more likely to be hit by X.

Rothvoß's Algorithm

Algorithm [Rothvoß, 2014]: given $K \subset \mathbb{R}^n$,

1. Sample a standard Gaussian $G \sim N(0, I_n)$;
2. Output
   $X = \arg\min\{\|x - G\|_2^2 : x \in K \cap [-1,1]^n\}$.



Goal: $|\{i : X_i \in \{-1,+1\}\}| \geq \alpha n$ for a constant α.
(X is a *partial coloring*.)

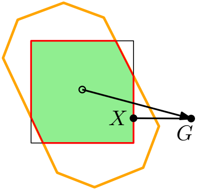
Intuition: If K is "big enough," then in an average direction $\partial[-1,1]^n$ is closer to the origin than ∂K and is more likely to be hit by X.

[Rothvoß, 2014] For any small enough α there is a δ so that
if K has Gaussian measure $\gamma_n(K) \geq e^{-\delta n}$, then
with high probability $|\{i : X_i \in \{-1,+1\}| \geq \alpha n$.

Rothvoß's Algorithm

Algorithm [Rothvoß, 2014]: given $K \subset \mathbb{R}^n$,

1. Sample a standard Gaussian $G \sim N(0, I_n)$;
2. Output
   $X = \arg\min\{\|x - G\|_2^2 : x \in K \cap [-1,1]^n\}$.



Goal: $|\{i : X_i \in \{-1,+1\}\}| \geq \alpha n$ for a constant α.
(X is a *partial coloring*.)

Intuition: If K is "big enough," then in an average direction $\partial[-1,1]^n$ is closer to the origin than ∂K and is more likely to be hit by X.

[Rothvoß, 2014] For any small enough α there is a δ so that *if there exists a dimension* $(1-\delta)n$ *subspace* W *for which* $K \cap W$ has Gaussian measure $\gamma_W(K \cap W) \geq e^{-\delta n}$, then with high probability $|\{i : X_i \in \{-1,+1\}\}| \geq \alpha n$.

Tightness of the Volume Lower Bound

Need to show: for any $U \in \mathbb{R}^{n\times N}$ and symmetric convex $K \subset \mathbb{R}^n$

$$\mathrm{hd}(U, K) \lesssim (1 + \log n) \cdot \mathrm{volLB}(U, K).$$

Tightness of the Volume Lower Bound

Need to show: for any $U \in \mathbb{R}^{n \times N}$ and symmetric convex $K \subset \mathbb{R}^n$

$$\mathrm{hd}(U, K) \lesssim (1 + \log n) \cdot \mathrm{volLB}(U, K).$$

Proof by an algorithm:
Find a partial coloring with discrepancy $\lesssim \mathrm{volLB}(U, K)$ and recurse.

Tightness of the Volume Lower Bound

Need to show: for any $U \in \mathbb{R}^{n\times N}$ and symmetric convex $K \subset \mathbb{R}^n$

$$\mathrm{hd}(U, K) \lesssim (1 + \log n) \cdot \mathrm{volLB}(U, K).$$

Proof by an algorithm:
Find a partial coloring with discrepancy $\lesssim \mathrm{volLB}(U, K)$ and recurse.

1. Preprocess so that $N = n$, $U = I_n$;
2. Apply Rothvoß's algorithm to tK, $t \asymp \mathrm{volLB}(I_n, K)$;
   - *If conditions hold*, gives a partial coloring $X \in tK$;
3. $S = \{i : -1 < X_i < 1\}$; Project K on $\mathbb{R}^S$ and recurse.
   - Need a "recentered" variant of Rothvoß's algorithm.

Tightness of the Volume Lower Bound

Need to show: for any $U \in \mathbb{R}^{n \times N}$ and symmetric convex $K \subset \mathbb{R}^n$

$$\mathrm{hd}(U, K) \lesssim (1 + \log n) \cdot \mathrm{volLB}(U, K).$$

Proof by an algorithm:
Find a partial coloring with discrepancy $\lesssim \mathrm{volLB}(U, K)$ and recurse.

1. Preprocess so that $N = n$, $U = I_n$;
2. Apply Rothvoß's algorithm to tK, $t \asymp \mathrm{volLB}(I_n, K)$;
   - *If conditions hold*, gives a partial coloring $X \in tK$;
3. $S = \{i : -1 < X_i < 1\}$; Project K on $\mathbb{R}^S$ and recurse.
   - Need a "recentered" variant of Rothvoß's algorithm.

After $k \lesssim 1 + \log n$ iterations, we have $X^1, \ldots X^k$ so that

$$X^1 + \ldots + X^k \in \{-1, 1\}^n;$$
$$\|X^1 + \ldots + X^k\|_K \leq kt \lesssim (1 + \log n)\,\mathrm{volLB}(I_n, K).$$

Tightness of the Volume Lower Bound

Need to show: for any $U \in \mathbb{R}^{n \times N}$ and symmetric convex $K \subset \mathbb{R}^n$

$$\mathrm{hd}(U, K) \lesssim (1 + \log n) \cdot \mathrm{volLB}(U, K).$$

Proof by an algorithm:
Find a partial coloring with discrepancy $\lesssim \mathrm{volLB}(U, K)$ and recurse.

1. Preprocess so that $N = n$, $U = I_n$;
2. Apply Rothvoß's algorithm to tK, $t \asymp \mathrm{volLB}(I_n, K)$;
   - *If conditions hold*, gives a partial coloring $X \in tK$;
3. $S = \{i : -1 < X_i < 1\}$; Project K on $\mathbb{R}^S$ and recurse.
   - Need a "recentered" variant of Rothvoß's algorithm.

After $k \lesssim 1 + \log n$ iterations, we have $X^1, \ldots X^k$ so that

$$X^1 + \ldots + X^k \in \{-1, 1\}^n;$$

$$\|X^1 + \ldots + X^k\|_K \leq kt \lesssim (1 + \log n)\,\mathrm{volLB}(I_n, K).$$

Main Challenge: Show that the conditions of Rothvoß's algorithm are satisfied.

From Volume To Gaussian Measure

For Rothvoß's algorithm, we need that on some subspace of large dimension, the body tK, $t \asymp \mathrm{volLB}(I_n, K)$, has large Gaussian measure.

From Volume To Gaussian Measure

For Rothvoß's algorithm, we need that on some subspace of large dimension, the body tK, $t \asymp \mathrm{volLB}(I_n, K)$, has large Gaussian measure.

From the definition of $\mathrm{volLB}(I_n, K)$:

$$\forall S \subseteq [n] : \mathrm{vol}((\mathrm{volLB}(I_n, K) \cdot K) \cap \mathbb{R}^S) \geq 1.$$

From Volume To Gaussian Measure

For Rothvoß's algorithm, we need that on some subspace of large dimension, the body tK, $t \asymp \mathrm{volLB}(I_n, K)$, has large Gaussian measure.

From the definition of $\mathrm{volLB}(I_n, K)$:

$$\forall S \subseteq [n] : \mathrm{vol}((\mathrm{volLB}(I_n, K) \cdot K) \cap \mathbb{R}^S) \geq 1.$$

Theorem (Structural result)

For any δ there exists a $m = m(\delta)$ so that the following holds. Let L be a symmetric convex body s.t. $\mathrm{vol}(L \cap \mathbb{R}^S) \geq 1$ for all $S \subseteq [n]$. There exists a subspace W of dimension $(1 - \delta)n$ for which

$$\gamma_W((mL) \cap W) \geq e^{-\delta n}.$$

Apply to $L = \mathrm{volLB}(I_n, K) \cdot K$ to get that the conditions of Rothvoß's algorithm are satisfied.

Proof Ideas

Generally applicable strategy:

1. Prove the theorem for an ellipsoid $E = T(B_2^n)$.
    - Reduces to linear algebra!

Proof Ideas

Generally applicable strategy:

1. Prove the theorem for an ellipsoid $E = T(B_2^n)$.
   - Reduces to linear algebra!
2. Approximate a general convex body L by an appropriate ellipsoid.

Theorem (Regular M-ellipsoid, [Milman, 1986; Pisier, 1989])

For any symmetric convex $L \subseteq \mathbb{R}^n$ there exists an ellipsoid E such that for any $t \geq 1$

$$\max\{N(L, tE), N(E, tL)\} \leq e^{cn/t},$$

where c is a constant.

$N(K, L)$ = number of translates of L needed to cover K.
E preserves "large scale" information about L.

Proof Ideas

Generally applicable strategy:

1. Prove the theorem for an ellipsoid $E = T(B_2^n)$.
   - Reduces to linear algebra!
2. Approximate a general convex body L by an appropriate ellipsoid.

Theorem (Regular M-ellipsoid, [Milman, 1986; Pisier, 1989])

For any symmetric convex $L \subseteq \mathbb{R}^n$ there exists an ellipsoid E such that for any $t \geq 1$

$$\max\{N(L, tE), N(E, tL)\} \leq e^{cn/t},$$

where c is a constant.

$N(K, L) =$ number of translates of L needed to cover K.
E preserves "large scale" information about L.

- $L \cap \mathbb{R}^S$ has large volume $\implies$ $E \cap \mathbb{R}^S$ has large volume.
- $E \cap W$ has large Gaussian measure $\implies$ $L \cap W$ has large Gaussian measure.

Partial Colorings

The bound $\mathrm{hd}(U, K) \lesssim (1 + \log n)\,\mathrm{volLB}(U, K)$ is in general tight.

Partial Colorings

The bound $\mathrm{hd}(U,K) \lesssim (1+\log n)\,\mathrm{volLB}(U,K)$ is in general tight.

Is the hereditary discrepancy of partial colorings $\asymp \mathrm{volLB}(U,K)$?

Partial Colorings

The bound $\mathrm{hd}(U, K) \lesssim (1 + \log n)\,\mathrm{volLB}(U, K)$ is in general tight.

Is the hereditary discrepancy of partial colorings $\asymp \mathrm{volLB}(U, K)$?

- The hereditary discrepancy of partial colorings is $\lesssim \mathrm{volLB}(U, K)$.

Partial Colorings

The bound $\mathrm{hd}(U, K) \lesssim (1 + \log n)\,\mathrm{volLB}(U, K)$ is in general tight.

Is the hereditary discrepancy of partial colorings $\asymp \mathrm{volLB}(U, K)$?

- The hereditary discrepancy of partial colorings is $\lesssim \mathrm{volLB}(U, K)$.
- A lower bound would follow from

Conjecture

Suppose $K \subset \mathbb{R}^n$ is a symmetric convex body of volume ≤ 1. Then there exists a $S \subseteq [n]$ s.t. $\mathrm{diam}_{\ell_2}(K \cap \mathbb{R}^S) \lesssim \sqrt{|S|}$.

Partial Colorings

The bound $\text{hd}(U, K) \lesssim (1 + \log n)\,\text{volLB}(U, K)$ is in general tight.

Is the hereditary discrepancy of partial colorings $\asymp \text{volLB}(U, K)$?

- The hereditary discrepancy of partial colorings is $\lesssim \text{volLB}(U, K)$.
- A lower bound would follow from

Conjecture

Suppose $K \subset \mathbb{R}^n$ is a symmetric convex body of volume ≤ 1. Then there exists a $S \subseteq [n]$ s.t. $\mathrm{diam}_{\ell_2}(K \cap \mathbb{R}^S) \lesssim \sqrt{|S|}$.

- True for ellipsoids and reduces to the Restricted Invertibility Principle.
- True for general bodies K if we replace $\mathbb{R}^S$ with an arbitrary subspace W and $|S|$ with $\dim W$.

Outline

Upper Bounds from Banaszczyk's Theorem

We showed how to efficiently compute near optimal signs $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, 1\}$ for any $u_1, \ldots, u_N$.

But what if we want to compute $\mathrm{vb}(C, K)$ or $\mathrm{hd}(U, K)$?

Upper Bounds from Banaszczyk's Theorem

We showed how to efficiently compute near optimal signs $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, 1\}$ for any $u_1, \ldots, u_N$.

But what if we want to compute $\mathrm{vb}(C, K)$ or $\mathrm{hd}(U, K)$?

- We do not know how to efficiently compute $\mathrm{volLB}(C, K)$.
- We need a natural *upper bound* on $\mathrm{vb}(C, K)$.

Upper Bounds from Banaszczyk's Theorem

We showed how to efficiently compute near optimal signs $\varepsilon_1, \dots, \varepsilon_N \in \{-1, 1\}$ for any $u_1, \dots, u_N$.

But what if we want to compute $\mathrm{vb}(C, K)$ or $\mathrm{hd}(U, K)$?

- We do not know how to efficiently compute $\mathrm{volLB}(C, K)$.
- We need a natural *upper bound* on $\mathrm{vb}(C, K)$.

Recall [Banaszczyk, 1998]:

For any convex $K \subset \mathbb{R}^n$ such that $\gamma_n(K) \geq \frac{1}{2}$, $\mathrm{vb}(B_2^n, K) \leq 5$.

Upper Bounds from Banaszczyk's Theorem

We showed how to efficiently compute near optimal signs $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, 1\}$ for any $u_1, \ldots, u_N$.

But what if we want to compute $\mathrm{vb}(C, K)$ or $\mathrm{hd}(U, K)$?

- We do not know how to efficiently compute $\mathrm{volLB}(C, K)$.
- We need a natural *upper bound* on $\mathrm{vb}(C, K)$.

Recall [Banaszczyk, 1998]:

For any convex $K \subset \mathbb{R}^n$ such that $\gamma_n(K) \geq \frac{1}{2}$, $\mathrm{vb}(B_2^n, K) \leq 5$.

Observations:

- If $\mathbb{E}\|G\|_K \leq 1$ for $G \sim N(0, I_n)$, then $\gamma_n(2K) \geq \frac{1}{2}$.
- $\mathrm{vb}(B_2^n, K) \lesssim \mathbb{E}\|G\|_K$.

Upper Bounds from Banaszczyk's Theorem

We showed how to efficiently compute near optimal signs $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, 1\}$ for any $u_1, \ldots, u_N$.

But what if we want to compute $\mathrm{vb}(C, K)$ or $\mathrm{hd}(U, K)$?

- We do not know how to efficiently compute $\mathrm{volLB}(C, K)$.
- We need a natural *upper bound* on $\mathrm{vb}(C, K)$.

Recall [Banaszczyk, 1998]:
For any convex $K \subset \mathbb{R}^n$ such that $\gamma_n(K) \geq \frac{1}{2}$, $\mathrm{vb}(B_2^n, K) \leq 5$.

Observations:

- If $\mathbb{E}\|G\|_K \leq 1$ for $G \sim N(0, I_n)$, then $\gamma_n(2K) \geq \frac{1}{2}$.
- $\mathrm{vb}(B_2^n, K) \lesssim \mathbb{E}\|G\|_K$.
- $\mathrm{vb}(C, K) \lesssim (\mathbb{E}\|G\|_K) \cdot \mathrm{diam}_{\ell_2}(C)$.

Last bound can be very loose! Can we do better?

A Better Upper Bound

Idea: Map C into B_2^n using a linear map.

$$\lambda(C,K) = \inf\{(\mathbb{E}\|G\|_{T(K)}) \cdot \mathrm{diam}_{\ell_2}(T(C)) : T \text{ a linear map}\}.$$

Claim: $\mathrm{vb}(C,K) \lesssim \lambda(C,K)$.

A Better Upper Bound

Idea: Map C into B_2^n using a linear map.

$$\lambda(C, K) = \inf\{(\mathbb{E}\|G\|_{T(K)}) \cdot \mathrm{diam}_{\ell_2}(T(C)) : T \text{ a linear map}\}.$$

Claim: $\mathrm{vb}(C, K) \lesssim \lambda(C, K)$.

- Take a linear map T achieving $\lambda(C, K)$;
  - Can assume $\mathrm{diam}_{\ell_2}(T(C)) = 1$, so $\mathbb{E}\|G\|_{T(K)} = \lambda(C, K)$;

A Better Upper Bound

Idea: Map C into B_2^n using a linear map.

$$\lambda(C, K) = \inf\{(\mathbb{E}\|G\|_{T(K)}) \cdot \mathrm{diam}_{\ell_2}(T(C)) : T \text{ a linear map}\}.$$

Claim: $\mathrm{vb}(C, K) \lesssim \lambda(C, K)$.

- Take a linear map T achieving $\lambda(C, K)$;
  - Can assume $\mathrm{diam}_{\ell_2}(T(C)) = 1$, so $\mathbb{E}\|G\|_{T(K)} = \lambda(C, K)$;
- $\mathrm{vb}(C, K) = \mathrm{vb}(T(C), T(K))$ and apply Banaszczyk's theorem.

Tightness of the Upper Bound

Theorem

For any symmetric convex $C, K \subset \mathbb{R}^n$,

$$\frac{\lambda(C,K)}{(1+\log n)^{5/2}} \lesssim \mathrm{vb}(C,K) \lesssim \lambda(C,K).$$

Moreover, given membership oracle access to K *and a vertex representation of* C*, we can efficiently compute* $\lambda(C,K)$*.*

For a matrix $U \in \mathbb{R}^{n\times N}$, we can take $C = \mathrm{conv}\{\pm u_1, \dots, \pm u_N\}$, and then $\lambda(C,K)$ approximates $\mathrm{hd}(U,K)$.

Tightness of the Upper Bound

Theorem

For any symmetric convex $C, K \subset \mathbb{R}^n$,

$$\frac{\lambda(C,K)}{(1+\log n)^{5/2}} \lesssim \mathrm{vb}(C,K) \lesssim \lambda(C,K).$$

Moreover, given membership oracle access to K *and a vertex representation of* C*, we can efficiently compute* $\lambda(C,K)$*.*

For a matrix $U \in \mathbb{R}^{n\times N}$, we can take $C = \mathrm{conv}\{\pm u_1, \dots, \pm u_N\}$, and then $\lambda(C,K)$ approximates $\mathrm{hd}(U,K)$.

Proof outline:

1. Formulate $\lambda(C,K)$ as a convex minimization problem;
2. Derive the Lagrange dual: an equivalent maximization problem;
3. Relate dual solutions to the volume lower bound.

Convex Formulation

$\|x\|_{T(K)} = \|T^{-1}x\|_K$

First attempt: $\inf\{\mathbb{E}\|T^{-1}G\|_K : \mathrm{diam}_{\ell_2}(T(C)) \le 1\}$

- *Not convex*: the objective is ∞ for $T = 0$ and finite for any invertible T, but $0 = \frac{1}{2}(T + (-T))$.

Convex Formulation

$\|x\|_{T(K)} = \|T^{-1}x\|_K$

First attempt: $\inf\{\mathbb{E}\|T^{-1}G\|_K : \mathrm{diam}_{\ell_2}(T(C)) \leq 1\}$

- *Not convex*: the objective is ∞ for $T = 0$ and finite for any invertible T, but $0 = \frac{1}{2}(T + (-T))$.

Observation: $\mathbb{E}\|T^{-1}G\|_K$ is defined entirely by $A = T^*T$, because the covariance of $T^{-1}G$ is given by A^{-1}.

Convex Formulation

$\|x\|_{T(K)} = \|T^{-1}x\|_K$

First attempt: $\inf\{\mathbb{E}\|T^{-1}G\|_K : \mathrm{diam}_{\ell_2}(T(C)) \le 1\}$

- *Not convex*: the objective is ∞ for $T = 0$ and finite for any invertible T, but $0 = \frac{1}{2}(T + (-T))$.

Observation: $\mathbb{E}\|T^{-1}G\|_K$ is defined entirely by $A = T^*T$, because the covariance of $T^{-1}G$ is given by A^{-1}.

Formulation:

$$\begin{aligned} \lambda(C,K) = &\inf f(A) \\ \text{s.t.} \quad & \langle x, Ax\rangle \le 1 \quad \forall x \in C \\ & A \succ 0. \end{aligned}$$

- $f(A) = \mathbb{E}\|T^{-1}G\|_K$ for any T such that $T^*T = A$;
  - f is well defined over positive definite A;

Convex Formulation

$\|x\|_{T(K)} = \|T^{-1}x\|_K$

First attempt: $\inf\{\mathbb{E}\|T^{-1}G\|_K : \mathrm{diam}_{\ell_2}(T(C)) \le 1\}$

- *Not convex*: the objective is ∞ for $T = 0$ and finite for any invertible T, but $0 = \frac{1}{2}(T + (-T))$.

Observation: $\mathbb{E}\|T^{-1}G\|_K$ is defined entirely by $A = T^*T$, because the covariance of $T^{-1}G$ is given by A^{-1}.

Formulation:

$$\begin{aligned} \lambda(C,K) = &\inf f(A) \\ \text{s.t.} \ \ &\langle x, Ax\rangle \le 1 \quad \forall x \in C \\ &A \succ 0. \end{aligned}$$

- $f(A) = \mathbb{E}\|T^{-1}G\|_K$ for any T such that $T^*T = A$;
  - f is well defined over positive definite A;
- The first constraint encodes $\mathrm{diam}_{\ell_2}(T(C)) \le 1$: $\langle x, Ax\rangle = \langle x, T^*Tx\rangle = \langle Tx, Tx\rangle = \|Tx\|_2^2$.

Properties of the Formulation

- The function $f(A)$ is convex in A, and the constraints are also convex;
- **Lagrange Duality**: there exists an *equivalent* dual maximization problem, whose value also equals $\lambda(U, C)$;

Properties of the Formulation

- The function $f(A)$ is convex in A, and the constraints are also convex;
- **Lagrange Duality**: there exists an *equivalent* dual maximization problem, whose value also equals $\lambda(U, C)$;
- Each dual solution gives a lower bound on $\mathrm{volLB}(C, K)$, and, therefore, on $\mathrm{vb}(C, K)$;
  - Tools: K-convexity, and Sudakov minoration;
- $\implies \lambda(C, K)$ gives a lower bound on $\mathrm{vb}(C, K)$.

Properties of the Formulation

- The function $f(A)$ is convex in A, and the constraints are also convex;
- **Lagrange Duality**: there exists an *equivalent* dual maximization problem, whose value also equals $\lambda(U, C)$;
- Each dual solution gives a lower bound on $\mathrm{volLB}(C, K)$, and, therefore, on $\mathrm{vb}(C, K)$;
    - Tools: K-convexity, and Sudakov minoration;
- $\implies \lambda(C, K)$ gives a lower bound on $\mathrm{vb}(C, K)$.

Computation: The convex optimization problem can be solved using the ellipsoid method, given a membership oracle for K and a vertex representation of C.

Outline

Conclusion

In this work:

- Tightness of natural upper and lower bounds for vector balancing.
- Efficient algorithms to find nearly optimal vector balancing signs, and to compute $\mathrm{vb}(C, K)$, and hereditary discrepancy with respect to any norm.
- Our results strongly use the geometry of the underlying discrepancy problem.

Conclusion

In this work:

- Tightness of natural upper and lower bounds for vector balancing.
- Efficient algorithms to find nearly optimal vector balancing signs, and to compute $\mathrm{vb}(C, K)$, and hereditary discrepancy with respect to any norm.
- Our results strongly use the geometry of the underlying discrepancy problem.

Open questions:

- Does $\mathrm{volLB}(C, K)$ give lower bounds on partial colorings?
- $\mathrm{vb}(K, K) \asymp \mathrm{volLB}(K, K)$? (True for ℓ_p.)
- Can the bounds for $\lambda(C, K)$ be improved?

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

Wojciech Banaszczyk. Balancing vectors and convex bodies. *Studia Math.*, 106(1):93–100, 1993. ISSN 0039-3223.

Nikhil Bansal. Constructive algorithms for discrepancy minimization. In *51st Annual IEEE Symposium on Foundations of Computer Science (FOCS), 2010*, pages 3–10. IEEE, 2010.

Nikhil Bansal, Moses Charikar, Ravishankar Krishnaswamy, and Shi Li. Better algorithms and hardness for broadcast scheduling via a discrepancy approach. In *SODA*, pages 55–71, 2014.

I. Bárány and VS Grinberg. On some combinatorial questions in finite-dimensional spaces. *Linear Algebra and its Applications*, 41:1–9, 1981.

J. Beck and T. Fiala. Integer-making theorems. *Discrete Applied Mathematics*, 3(1):1–8, 1981.

József Beck. Balanced two-colorings of finite sets in the square i. *Combinatorica*, 1(4):327–335, 1981.

Moses Charikar, Alantha Newman, and Aleksandar Nikolov. Tight hardness results for minimizing discrepancy. In *SODA*, pages 1607–1614, 2011.

Aryeh Dvoretzky. Problem. In *Proc. Sympos. Pure Math., Vol. VII*. Amer. Math. Soc., Providence, R.I., 1963.

Efim Davydovich Gluskin. Extremal properties of orthogonal parallelepipeds and their applications to the geometry of banach spaces. *Mathematics of the USSR-Sbornik*, 64(1):85, 1989.

Rebecca Hoberg and Thomas Rothvoss. A logarithmic additive integrality gap for bin packing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2616–2625. SIAM, Philadelphia, PA, 2017. doi: 10.1137/1.9781611974782.172. URL https://doi.org/10.1137/1.9781611974782.172.

Kasper Green Larsen. On range searching in the group model and combinatorial discrepancy. *SIAM J. Comput.*, 43(2):673–686, 2014. doi: 10.1137/120865240. URL http://dx.doi.org/10.1137/120865240.

L. Lovász, J. Spencer, and K. Vesztergombi. Discrepancy of set-systems and matrices. *European Journal of Combinatorics*, 7(2):151–160, 1986.

Jiri Matousek. Approximations and optimal geometric divide-and-conquer. *Journal of Computer and System Sciences*, 50(2):203–208, 1995.

Vitali D. Milman. Inégalité de Brunn-Minkowski inverse et applications à la théorie locale des espaces normés. *C. R. Acad. Sci. Paris Sér. I Math.*, 302(1):25–28, 1986. ISSN 0249-6291.

Alantha Newman, Ofer Neiman, and Aleksandar Nikolov. Beck's three permutations conjecture: a counterexample and some consequences. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science—FOCS 2012*, pages 253–262. IEEE Computer Soc., Los Alamitos, CA, 2012.

Aleksandar Nikolov. An improved private mechanism for small databases. In Magnús M. Halldórsson, Kazuo Iwama, Naoki Kobayashi, and Bettina Speckmann, editors, *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part I*, volume 9134 of *Lecture Notes in Computer Science*, pages 1010–1021. Springer, 2015. doi: 10.1007/978-3-662-47672-7_82. URL http://dx.doi.org/10.1007/978-3-662-47672-7_82.

Aleksandar Nikolov and Kunal Talwar. Approximating hereditary discrepancy via small width ellipsoids. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 324–336. SIAM, Philadelphia, PA, 2015. doi: 10.1137/1.9781611973730.24. URL https://doi.org/10.1137/1.9781611973730.24.

Aleksandar Nikolov, Kunal Talwar, and Li Zhang. The geometry of differential privacy: The small database and approximate cases. *SIAM J. Comput.*, 45(2):575–616, 2016. doi: 10.1137/130938943. URL http://dx.doi.org/10.1137/130938943.

Gilles Pisier. A new approach to several results of V. Milman. *J. Reine Angew. Math.*, 393:115–131, 1989. ISSN 0075-4102. doi: 10.1515/crll.1989.393.115. URL https://doi.org/10.1515/crll.1989.393.115.

Thomas Rothvoss. The entropy rounding method in approximation algorithms. In *Symposium on Discrete Algorithms (SODA)*, pages 356–372, 2012.

Thomas Rothvoß. Constructive discrepancy minimization for convex sets.

In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, pages 140–145. IEEE Computer Society, 2014. doi: 10.1109/FOCS.2014.23. URL http://dx.doi.org/10.1109/FOCS.2014.23.

Joel Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289:679–706, 1985.

Zhewei Wei and Ke Yi. The space complexity of 2-dimensional approximate range counting. In Sanjeev Khanna, editor, *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013, New Orleans, Louisiana, USA, January 6-8, 2013*, pages 252–264. SIAM, 2013. ISBN 978-1-61197-251-1. doi: 10.1137/1.9781611973105.19. URL http://dx.doi.org/10.1137/1.9781611973105.19.