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

1. Introduction
2. Volume Lower Bound
3. Factorization Upper Bounds
4. Conclusion
Discrepancy

\[
\begin{pmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
\end{pmatrix}
\begin{pmatrix}
-1 \\
1 \\
-1 \\
1 \\
\end{pmatrix}
= 
\begin{pmatrix}
1 \\
0 \\
0 \\
-1 \\
\end{pmatrix}
\]

\[
disc(U, \| \cdot \|_\infty) = \min_{\varepsilon \in \{\pm 1\}^N} \| U \varepsilon \|_\infty
\]
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 $U \in \{0, 1\}^{n \times N}$,
  \[ \text{disc}(U) \lesssim \sqrt{n} \]

- [Beck and Fiala, 1981]: For any matrix $U \in \{0, 1\}^{n \times N}$ with at most $t$ ones per column,
  \[ \text{disc}(U) \leq 2t - 1 \]

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

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

- Implied by: For any $u_1, \ldots, u_N \in B^n_\infty = [-1, 1]^n$, there exist $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, +1\}$ s.t. $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty \lesssim \sqrt{n}$. Most combinatorial discrepancy bounds are implied by geometric vector balancing arguments.
Basic Bounds

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

- **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, \( \text{disc}(U) \leq 2t - 1 \)
Basic Bounds

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

- **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, $\text{disc}(U) \leq 2t - 1$

- **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. $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty < 2$. 

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

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

  **Implied by:** For any $u_1, \ldots, u_N \in B^n_\infty = [-1, 1]^n$, there exist 
  $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, +1\}$ such that 
  \[ \|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty \lesssim \sqrt{n}. \]

- [Beck and Fiala, 1981]: For any matrix $U \in \{0, 1\}^{n \times N}$ with at most $t$ 
  ones per column, $\text{disc}(U) \leq 2t - 1$

  **Implied by:** For any $u_1, \ldots, u_N \in B^n_1$, there exist 
  $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, +1\}$ such that 
  \[ \|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_\infty < 2. \]

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

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

![Diagram showing vector balancing problem](image)
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:** $\|x\|_K = \inf\{t \geq 0 : x \in tK\}; \ t = \text{disc}((u_i)_{i=1}^N, \|\cdot\|_K)$. 
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:** \( \|x\|_K = \inf\{t \geq 0 : x \in tK\} \); \( t = \text{disc}((u_i)_{i=1}^N, \|\cdot\|_K) \).

**Vector Balancing Constant:** worst case over sequences in \( C \)

\[ \text{vb}(C, K) = \sup \left\{ \text{disc}(U, \|\cdot\|_K) : N \in \mathbb{N}, u_1, \ldots, u_N \in C, U = (u_i)_{i=1}^N \right\} \]
Questions and Prior Results

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

Banaszczyk's theorem implies $\text{vb}(B_n^2, B_n^\infty) \lesssim \sqrt{\log 2} n$. 

Komlós Problem: Prove or disprove $\text{vb}(B_n^2, B_n^\infty) \lesssim 1$. 

Sasho Nikolov (U of T) Balancing Vectors
Questions and Prior Results

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

- [Dvoretzky, 1963] “What can be said” about $\text{vb}(K, K)$?
- [Bárány and Grinberg, 1981] $\text{vb}(K, K) \leq n$ for all $K$.
- [Spencer, 1985; Gluskin, 1989] $\text{vb}(B_n^\infty, B_n^\infty) \preceq \sqrt{n}$
- [Beck and Fiala, 1981] $\text{vb}(B_1^n, B_\infty^n) < 2$
- [Banaszczyk, 1998] $\text{vb}(B_2^n, K) \leq 5$ if $K$ has Gaussian measure $\gamma_n(K) \geq \frac{1}{2}$

Komlós Problem: Prove or disprove $\text{vb}(B_2^n, B_\infty^n) \preceq 1$.

- Banaszczyk’s theorem implies $\text{vb}(B_2^n, B_\infty^n) \preceq \sqrt{\log 2n}$. 
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

\[
\| Ux - Uw \|_K \leq \text{vb}(C, K).
\]
Our Results

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

A natural volumetric lower bound on $\text{vb}(C, K)$ 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 $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_K \lesssim (1 + \log n) \text{ vb}(C, K)$. Also rounding version.

An efficiently computable upper bound on $\text{vb}(C, K)$ is tight up to factors polynomial in $\log n$. Based on an optimal application of Banaszczyk's theorem. Implies an efficient approximation algorithm for $\text{vb}(C, K)$. 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 $K$. 

Sasho Nikolov (U of T) Balancing Vectors
Our Results

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

- A natural volumetric lower bound on \( \text{vb}(C, K) \) 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 \( \|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_K \lesssim (1 + \log n) \text{vb}(C, K) \).

- Also rounding version.

Also rounding version.

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

The results extend to hereditary discrepancy with respect to arbitrary norms.
Our Results

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

- A natural volumetric lower bound on $\text{vb}(C, K)$ 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 $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_K \lesssim (1 + \log n) \text{vb}(C, K)$.
  - Also rounding version.

- An efficiently computable upper bound on $\text{vb}(C, K)$ is tight up to factors polynomial in $\log n$.
  - Based on an optimal application of Banaszczyks’ theorem.
  - Implies an efficient approximation algorithm for $\text{vb}(C, K)$.
Our Results

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

- A natural volumetric lower bound on \( \text{vb}(C, K) \) is tight up to a \( O(\log n) \) factor.
  - The proof implies an efficient algorithm to compute \( \epsilon \in \{-1, 1\}^N \) given \( u_1, \ldots, u_N \in C \), so that \( \|\epsilon_1 u_1 + \ldots + \epsilon_N u_N\|_K \lesssim (1 + \log n) \text{vb}(C, K) \).
  - Also rounding version.

- An efficiently computable upper bound on \( \text{vb}(C, K) \) is tight up to factors polynomial in \( \log n \).
  - Based on an optimal application of Banaszczyks’ theorem.
  - Implies an efficient approximation algorithm for \( \text{vb}(C, K) \).

- The results extend to hereditary discrepancy with respect to arbitrary norms.
Our Results

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

- A natural volumetric lower bound on $\text{vb}(C, K)$ 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 $\|\varepsilon_1 u_1 + \ldots + \varepsilon_N u_N\|_K \lesssim (1 + \log n) \text{vb}(C, K)$.
  - Also rounding version.

- An efficiently computable upper bound on $\text{vb}(C, K)$ is tight up to factors polynomial in $\log n$.
  - Based on an optimal application of Banaszczyks’ theorem.
  - Implies an efficient approximation algorithm for $\text{vb}(C, K)$.

- 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

1. Introduction

2. Volume Lower Bound

3. Factorization Upper Bounds

4. Conclusion
**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]
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 $U$. 

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

\[ \text{disc}(U, K) = \min \{ t : tL \cap \{-1, 1\}^N \neq \emptyset \} . \]
The Volume Lower Bound

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

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

[Lovász, Spencer, and Vesztergombi, 1986]:
If $t = \text{hd}(U, K)$, 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 \( x \)”.

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

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

[Banaszczyk, 1993]:

\[
1 = \text{vol}([0, 1]^N) \geq \text{vol}(tL) = t^N \text{vol}(L)
\]
The Volume Lower Bound

Define \( L = \{ x \in \mathbb{R}^N : Ux \in K \} \): the set of “good \( x \)”.

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

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

[Banaszczyk, 1993]:

\[ 1 = \text{vol}([0, 1]^N) \geq \text{vol}(tL) = t^N \text{vol}(L) \iff \text{hd}(U, K) \geq \text{vol}(L)^{-1/N}. \]
A Hereditary Volume Lower Bound

A simple strengthening:

$$\text{hd}(U, K) \geq \text{volLB}(U, K) = \max_{S \subseteq [N]} \frac{\text{vol}(\{ x \in \mathbb{R}^S : U_S x \in K \})}{|S|}.$$
A Hereditary Volume Lower Bound

A simple strengthening:

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

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

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

A simple strengthening:

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

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

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

Theorem

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

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

$$\operatorname{volLB}(C, K) \leq \operatorname{vb}(C, K) \lesssim (1 + \log n) \cdot \operatorname{volLB}(C, K)$$
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 $\alpha$.
($X$ 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 $\partial K$ 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 $\alpha$.

(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 $\partial K$ and is more likely to be hit by $X$.

[Rothvoß, 2014] For any small enough $\alpha$ there is a $\delta$ 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 $\alpha$.

(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 $\partial K$ and is more likely to be hit by $X$.

[Rothvoß, 2014] For any small enough $\alpha$ there is a $\delta$ 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 $K \subset \mathbb{R}^n$

$$\text{hd}(U, K) \lesssim (1 + \log n) \cdot \text{volLB}(U, K).$$
Tightness of the Volume Lower Bound

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

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

Proof by an algorithm:
Find a partial coloring with discrepancy $\lesssim \text{volLB}(U, K)$ and recurse.
Tightness of the Volume Lower Bound

Need to show: for any $U \in \mathbb{R}^{n \times N}$ and symmetric convex $K \subset \mathbb{R}^n$
$$\text{hd}(U, K) \lesssim (1 + \log n) \cdot \text{volLB}(U, K).$$

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

1. Preprocess so that $N = n$, $U = I_n$;
2. Apply Rothvoß’s algorithm to $tK$, $t \asymp \text{volLB}(I_n, K)$;
   - 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 $K \subset \mathbb{R}^n$

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

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

1. Preprocess so that $N = n$, $U = I_n$;
2. Apply Rothvoß’s algorithm to $tK$, $t \approx \text{volLB}(I_n, K)$;
   - 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

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

$$\|X^1 + \ldots + X^k\|_K \leq kt \lesssim (1 + \log n) \text{volLB}(I_n, K).$$
Tightness of the Volume Lower Bound

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

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

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

1. Preprocess so that $N = n$, $U = I_n$;
2. Apply Rothvoss’s algorithm to $tK$, $t \approx \text{volLB}(I_n, K)$;
   - 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

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

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

**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 $tK$, $t \propto \text{volLB}(I_n, K)$, has large Gaussian measure.
From Volume To Gaussian Measure

For Rothvoß’s algorithm, we need that on some subspace of large dimension, the body $tK$, $t \approx \text{volLB}(I_n, K)$, has large Gaussian measure. From the definition of $\text{volLB}(I_n, K)$:

$$\forall S \subseteq [n] : \text{vol}((\text{volLB}(I_n, K) \cdot K) \cap \mathbb{R}^S) \geq 1.$$
From Volume To Gaussian Measure

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

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

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

**Theorem (Structural result)**

For any $\delta$ there exists a $m = m(\delta)$ so that the following holds.

Let $L$ be a symmetric convex body s.t. $\text{vol}(L \cap \mathbb{R}^S) \geq 1$ 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 $L = \text{volLB}(I_n, K) \cdot K$ 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.
The bound $\text{hd}(U, K) \lesssim (1 + \log n) \text{volLB}(U, K)$ is in general tight.
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 $\approx \text{volLB}(U, K)$?
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 $\lesssim \text{volLB}(U, K)$.
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 $\simeq \text{volLB}(U, K)$?

- 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 $\leq 1$. Then there exists a $S \subseteq [n]$ s.t. $\text{diam}_{\ell_2}(K \cap \mathbb{R}^S) \lesssim \sqrt{|S|}$.***
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 $\lesssim \text{volLB}(U, K)$?

- 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 $\leq 1$. Then there exists a $S \subseteq [n]$ s.t. $\text{diam}_{\ell_2}(K \cap \mathbb{R}^S) \lesssim \sqrt{|S|}$.*

- 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

1. Introduction
2. Volume Lower Bound
3. Factorization Upper Bounds
4. Conclusion
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 \( \text{vb}(C, K) \) or \( \text{hd}(U, K) \)?

We do not know how to efficiently compute \( \text{volLB}(C, K) \).

We need a natural upper bound on \( \text{vb}(C, K) \).

Recall [Banaszczyk, 1998]:

For any convex \( K \subset \mathbb{R}^n \) such that \( \gamma_n(K) \geq \frac{1}{2} \), \( \text{vb}(B_{\ell_2^n}, K) \leq 5 \).

Observations:

If \( E \|G\|_K \leq 1 \) for \( G \sim \mathcal{N}(0, I_n) \), then \( \gamma_n(2K) \geq \frac{1}{2} \).

\( \text{vb}(B_{\ell_2^n}, K) \lesssim E \|G\|_K \).

\( \text{vb}(C, K) \lesssim (E \|G\|_K) \cdot \text{diam}_\ell_2(C) \).

The last bound can be very loose! Can we do better?
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 $vb(C, K)$ or $hd(U, K)$?

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

Recall [Banaszczyk, 1998]:

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

Observations:

- If $E \|G\|_K \leq 1$ for $G \sim N(0, I_n)$, then $\gamma_n(2K) \geq \frac{1}{2}$.
- $vb(B_n^2, K) \asymp E \|G\|_K$.
- $vb(C, K) \asymp (E \|G\|_K) \cdot diameter_\ell_2(C)$. 

Last bound can be very loose! Can we do better?
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 $\text{vb}(C, K)$ or $\text{hd}(U, K)$?

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

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

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

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

Observations:
- If \( \mathbb{E}\|G\|_K \leq 1 \) for \( G \sim N(0, I_n) \), then \( \gamma_n(2K) \geq \frac{1}{2} \).
- \( \text{vb}(B_2^n, K) \lesssim \mathbb{E}\|G\|_K \).
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 $\text{vb}(C, K)$ or $\text{hd}(U, K)$?

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

Recall [Banaszczyk, 1998]:

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

**Observations:**

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

Last bound can be very loose! Can we do better?
A Better Upper Bound

**Idea:** Map $C$ into $B^n_2$ using a linear map.

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

**Claim:** $\nu_b(C, K) \lesssim \lambda(C, K)$. 
A Better Upper Bound

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

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

**Claim:** $vb(C, K) \lesssim \lambda(C, K)$.

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

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

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

**Claim:** $\nu b(C, K) \lesssim \lambda(C, K)$.

- Take a linear map $T$ achieving $\lambda(C, K)$;
  - Can assume $\text{diam}_{\ell_2}(T(C)) = 1$, so $E\|G\|_{T(K)} = \lambda(C, K)$;
  - $\nu b(C, K) = \nu b(T(C), T(K))$ and apply Banaszczyk’s theorem.
Tightness of the Upper Bound

**Theorem**

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

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

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 $C = \text{conv}\{\pm u_1, \ldots, \pm u_N\}$, and then $\lambda(C, K)$ approximates $\text{hd}(U, K)$. 

**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.
Tightness of the Upper Bound

**Theorem**

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

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

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 $C = \text{conv}\{\pm u_1, \ldots, \pm u_N\}$, and then $\lambda(C, K)$ approximates $\text{hd}(U, K)$.

**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:** \( \inf \{ \mathbb{E}\| T^{-1}G \|_K : \text{diam}_{\ell_2}(T(C)) \leq 1 \} \)

- *Not convex:* the objective is \( \infty \) 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:** \( \inf \{ E\| T^{-1}G\|_K : \text{diam}_{\ell_2}(T(C)) \leq 1 \} \)

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

**Observation:** \( E\| T^{-1}G\|_K \) 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:** \( \inf \{ \mathbb{E} \|T^{-1}G\|_K : \text{diam}_{\ell_2}(T(C)) \leq 1 \} \)

- *Not convex:* the objective is \( \infty \) 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:**

\[ \lambda(C, K) = \inf f(A) \]

\[ \text{s.t. } \langle x, Ax \rangle \leq 1 \quad \forall x \in C \]

\[ A \succ 0. \]

- \( f(A) = \mathbb{E} \|T^{-1}G\|_K \) 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:** \( \inf \{ \mathbb{E}\| T^{-1}G \|_K : \text{diam}_2( T(C)) \leq 1 \} \)

- *Not convex:* the objective is \( \infty \) 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:**

\[ \lambda(C, K) = \inf f(A) \]

s.t. \[ \langle x, Ax \rangle \leq 1 \quad \forall x \in C \]

\[ A \succ 0. \]

- \( f(A) = \mathbb{E}\| T^{-1}G \|_K \) for any \( T \) such that \( T^* T = A \);
- \( f \) is well defined over positive definite \( A \);
- The first constraint encodes \( \text{diam}_2( T(C)) \leq 1: \)
  \[ \langle x, Ax \rangle = \langle x, T^* Tx \rangle = \langle Tx, Tx \rangle = \| Tx \|_2^2. \]
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 $\text{volLB}(C, K)$, and, therefore, on $\text{vb}(C, K)$;
  - Tools: $K$-convexity, and Sudakov minoration;
- $\Rightarrow \lambda(C, K)$ gives a lower bound on $\text{vb}(C, K)$. 

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

Sasho Nikolov (U of T)
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 $\text{volLB}(C, K)$, and, therefore, on $\text{vb}(C, K)$;
  - Tools: $K$-convexity, and Sudakov minoration;
- $\implies \lambda(C, K)$ gives a lower bound on $\text{vb}(C, K)$.

**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

1. Introduction
2. Volume Lower Bound
3. Factorization Upper Bounds
4. Conclusion
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 $\text{vb}(C, K)$, 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 \( \text{vb}(C, K) \), and hereditary discrepancy with respect to any norm.
- Our results strongly use the geometry of the underlying discrepancy problem.

Open questions:

- Does \( \text{volLB}(C, K) \) give lower bounds on partial colorings?
- \( \text{vb}(K, K) \preceq \text{volLB}(K, K) ? \) (True for \( \ell_p \).
- Can the bounds for \( \lambda(C, K) \) be improved?


Thomas Rothvoss. Constructive discrepancy minimization for convex sets.