Towards a Constructive Version of Banaszczyk’s Vector Balancing Theorem

Daniel Dadush \(^1\)  
Shashwat Garg \(^2\)  
Shachar Lovett \(^3\)  
Sasho Nikolov \(^4\)

\(^1\)CWI  
\(^2\)TU Eindhoven  
\(^3\)UCSD  
\(^4\)U of Toronto
Discrepancy of Set Systems

Given: System of \( m \) subsets \( S = \{ S_1, \ldots, S_m \} \) of \([n] = \{1, \ldots, n\}\).
Color each element of \( P \) red or blue, so that each set is as balanced as possible.

Discrepancy of a coloring: maximum imbalance (above: 1).
Discrepancy of \( S \): discrepancy of the best coloring.

\[
\text{disc } S := \min_{\chi:[n] \rightarrow \{-1,1\}} \max_i \left| \sum_{j \in S_i} \chi(j) \right|
\]
Discrepancy of Set Systems

Given: System of $m$ subsets $S = \{S_1, \ldots, S_m\}$ of $[n] = \{1, \ldots, n\}$.
Color each element of $P$ red or blue, so that each set is as balanced as possible.

Discrepancy of a coloring: maximum imbalance (above: 1).
Discrepancy of $S$: discrepancy of the best coloring.

$$\text{disc } S := \min_{\chi: [n] \to \{-1, 1\}} \max_i \left| \sum_{j \in S_i} \chi(j) \right|$$
Beck-Fiala

**Theorem ([Beck and Fiala, 1981])**

Suppose each $i \in [n]$ appears in at most $t$ sets of $S$. Then $\text{disc } S \leq 2t - 1$.

**Beck-Fiala Conjecture.** $\text{disc } S = O(\sqrt{t})$.
Beck-Fiala

Theorem ([Beck and Fiala, 1981])

*Suppose each* $i \in [n]$ *appears in at most* $t$ *sets of* $S$. *Then* $\text{disc } S \leq 2t - 1$.

Beck-Fiala Conjecture. $\text{disc } S = O(\sqrt{t})$.

- Recently improved to $2t - \log^* t$ [Bukh, 2013]
- No better bound known in terms of $t$ only!
- The proof of the theorem is *an (efficient) algorithm*!
Komlòs Conjecture

Komlòs Conjecture. For any vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ with $\max_i \|u_i\|_2 \leq 1$, there exist signs $\varepsilon_1, \ldots, \varepsilon_n$ for which

$$\left\| \sum_i \varepsilon_i u_i \right\|_\infty = O(1).$$
Komlós Conjecture

Komlós Conjecture. For any vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ with $\max_i \|u_i\|_2 \leq 1$, there exist signs $\varepsilon_1, \ldots, \varepsilon_n$ for which

$$\left\| \sum_i \varepsilon_i u_i \right\|_\infty = O(1).$$

- $O(1)$ is independent of $m$ and $n$.
- Implies the Beck-Fiala Conjecture: Take $u_j$ to be the $j$-th column of the incidence matrix of $S$, scaled by $t^{-1/2}$.
  - $j$-th column of incidence matrix: indicator vector of $\{i : j \in S_i\}$.
  - $\sqrt{t}\left\| \sum_j \varepsilon_j u_j \right\|_\infty$ is the discrepancy of the coloring $\chi(j) = \varepsilon_j$. 
Banaszczyk’s Theorem

Theorem ([Banaszczyk, 1998])

Let $X$ be a standard $m$-dimensional Gaussian, and let $K$ be a convex body in $\mathbb{R}^m$ such that $\Pr[X \in K] \geq 1/2$.
Banaszczyk’s Theorem

**Theorem ([Banaszczyk, 1998])**

Let $X$ be a standard $m$-dimensional Gaussian, and let $K$ be a convex body in $\mathbb{R}^m$ such that $\Pr[X \in K] \geq 1/2$.

For any vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ with $\max_i \|u_i\|_2 \leq 1/5$, there exist signs $\varepsilon_1, \ldots, \varepsilon_n$ for which

$$\sum_i \varepsilon_i u_i \in K.$$
Banaszczyk’s Theorem

Theorem ([Banaszczyk, 1998])

Let $X$ be a standard $m$-dimensional Gaussian, and let $K$ be a convex body in $\mathbb{R}^m$ such that $\Pr[X \in K] \geq 1/2$.

For any vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ with $\max_i \|u_i\|_2 \leq 1/5$, there exist signs $\varepsilon_1, \ldots, \varepsilon_n$ for which

$$\sum_i \varepsilon_i u_i \in K.$$

- The proof is not an efficient algorithm!

- By taking $K = O(\sqrt{\log m}) \cdot [-1, 1]^m$, we get a bound of $O(\sqrt{\log m})$ for Komlós and $O(\sqrt{t \log m})$ for Beck-Fiala.
  
  - Recent algorithmic proof of these bounds, but not the full theorem, in [Bansal, Dadush, and Garg, 2016].
Banaszczyk’s Theorem

Theorem ([Banaszczyk, 1998])

Let $X$ be a standard $m$-dimensional Gaussian, and let $K$ be a convex body in $\mathbb{R}^m$ such that $\Pr[X \in K] \geq 1/2$.

For any vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ with $\max_i \|u_i\|_2 \leq 1/5$, there exist signs $\varepsilon_1, \ldots, \varepsilon_n$ for which

$$
\sum_i \varepsilon_i u_i \in K.
$$

- The proof is not an efficient algorithm!
- By taking $K = O(\sqrt{\log m}) \cdot [-1, 1]^m$, we get a bound of $O(\sqrt{\log m})$ for Komlós and $O(\sqrt{t \log m})$ for Beck-Fiala.
  - Recent algorithmic proof of these bounds, but not the full theorem, in [Bansal, Dadush, and Garg, 2016].
- Also used in approximation algorithm for hereditary discrepancy, bounds on discrepancy of boxes, vector-rearrangement problems.
Interlude: Subgaussian Random Variables

**Definition**

A real-valued random variable \( X \) is \( s \)-subgaussian if

\[
\Pr[|X| \geq t] \leq 2 \exp \left( -\frac{t^2}{2s^2} \right).
\]

A random variable \( Y \in \mathbb{R}^m \) is \( s \)-subgaussian if for every unit vector \( \theta \in \mathbb{S}^{m-1} \) the marginal \( \langle \theta, Y \rangle \) is \( s \)-subgaussian.

I.e., an \( s \)-subgaussian random variable shrinks about as fast as a Gaussian with variance \( s^2 \) in every direction.
Interlude: Subgaussian Random Variables

Definition

A real-valued random variable $X$ is $s$-subgaussian if

$$\Pr[|X| \geq t] \leq 2 \exp\left(-\frac{t^2}{2s^2}\right).$$

A random variable $Y \in \mathbb{R}^m$ is $s$-subgaussian if for every unit vector $\theta \in S^{m-1}$ the marginal $\langle \theta, Y \rangle$ is $s$-subgaussian.
Interlude: Subgaussian Random Variables

Definition

A real-valued random variable $X$ is $s$-subgaussian if

$$\Pr[|X| \geq t] \leq 2 \exp \left( -\frac{t^2}{2s^2} \right).$$

A random variable $Y \in \mathbb{R}^m$ is $s$-subgaussian if for every unit vector $\theta \in S^{m-1}$ the marginal $\langle \theta, Y \rangle$ is $s$-subgaussian.

I.e., an $s$-subgaussian random variable shrinks about as fast as a Gaussian with variance $s^2$ in every direction.
The Main Equivalence

Theorem

Let \( T = \{ \sum_i \pm u_i \} \) where the vectors \( u_1, \ldots, u_n \) satisfy \( \max_i \|u_i\|_2 \leq 1/5 \).

The following two are equivalent:

1. Banaszczyk’s theorem restricted to convex bodies \( K \) symmetric around 0.
2. There exists an \( O(1) \)-subgaussian \( Y \) supported on \( T \), where \( O(1) \) is independent of \( m, n \), or the vectors.
The Main Equivalence

**Theorem**

Let \( T = \{ \sum_i \pm u_i \} \) where the vectors \( u_1, \ldots, u_n \) satisfy \( \max_i \| u_i \|_2 \leq 1/5 \). The following two are equivalent:

1. Banaszczyk’s theorem restricted to convex bodies \( K \) symmetric around 0.
2. There exists an \( O(1) \)-subgaussian \( Y \) supported on \( T \), where \( O(1) \) is independent of \( m, n \), or the vectors.

- 2. was not known before, and we know no direct proof.
- If we can sample \( Y \) efficiently, we would have an algorithmic version of Banaszczyk’s theorem!
- Using a random walk, we can sample an \( O(\sqrt{\log m}) \)-subgaussian \( Y \): recovers Banaszczyk algorithmically for symmetric \( K \), up to a factor of \( O(\sqrt{\log m}) \).
Theorem
Let $X$ be a standard Gaussian in $\mathbb{R}^m$, and $K \subset \mathbb{R}^m$ be a symmetric convex body such that $\Pr[X \in K] \geq 1/2$. Then, for any $s$-subgaussian $Y$,

$$\Pr[Y \in O(s) \cdot K] \geq 1/2.$$  

- **Universal sampler**: there is a single distribution on $\sum_i \pm u_i$ which works for all $K$.  

Proof of Theorem

Need: $\mathbb{E}\|Y\|_K = O(s)$. Then done by Markov.
Proof of Theorem

\[ \|x\|_K = \min \{ t : x \in tK \} \]

Need: \( \mathbb{E}\|Y\|_K = O(s) \). Then done by Markov.

i. [Borell, 1975] For any symmetric convex body \( K \), and a standard Gaussian \( X \), \( \Pr[X \in K] \geq 1/2 \Rightarrow \mathbb{E}\|X\|_K = O(1) \).

ii. [Talagrand, 1987] For any \( s \)-subgaussian \( Y \), and any symmetric convex body \( K \), \( \mathbb{E}\|Y\|_K = O(s) \cdot \mathbb{E}\|X\|_K \).

From i. and ii., we get \( \mathbb{E}\|Y\|_K = O(s) \).
Define a zero-sum game:

- **Min** has strategies $T = \{\sum_i \pm u_i\}$.
- **Max** player has strategies $\{v \in \mathbb{R}^m\}$.
- The payoff of $y \in T$ and $v \in \mathbb{R}^m$ is $(e^{\langle y, v \rangle} + e^{-\langle y, v \rangle})/e\|v\|^2/2$.

Using Banaszczyk's theorem and the von Neumann min-max principle, we can bound the value of the game:

$$\min_{Y \text{ r.v. supp. on } T} \max_{v \in \mathbb{R}^m} E[e^{\langle Y, v \rangle} + e^{-\langle Y, v \rangle}/e\|v\|^2/2] \leq 2.$$

Implies $E[e^{\lvert \langle Y, v \rangle \rvert}]/e\|v\|^2/2 \leq 2$ by Chernoff trick, $Y$ is $O(1)$-subgaussian.
Define a zero-sum game:

- **Min** has strategies \( T = \{ \sum_i \pm u_i \} \).
- **Max** player has strategies \( \{ v \in \mathbb{R}^m \} \).
- The payoff of \( y \in T \) and \( v \in \mathbb{R}^m \) is \( (e^{\langle y, v \rangle} + e^{-\langle y, v \rangle})/e\|v\|_2^2/2 \).

Using Banaszczyk’s theorem, and the von Neumann min-max principle, we can bound the value of the game:

\[
\min_{Y \text{ r.v. supp. on } T} \max_{v \in \mathbb{R}^m} \mathbb{E} \left[ \frac{e^{\langle Y, v \rangle} + e^{-\langle Y, v \rangle}}{e\|v\|_2^2/2} \right] \leq 2.
\]

Implies \( \mathbb{E}[e^{\langle Y, v \rangle}] \leq 2e\|v\|_2^2/2 \). By Chernoff trick, \( Y \) is \( O(1) \)-subgaussian.
Asymmetric Bodies

*Does an efficient sampler for $O(1)$-subgaussian $Y$ imply algorithmic Banaszczyk for asymmetric $K$?*

---

**Bad News:** Take $K = \{ x \in \mathbb{R}^m : x_1 \leq 0 \}$ and $Y = e_1$. Then:

- $Y$ is $O(1)$-subgaussian
- $\Pr[X \in K] = 1/2$ for standard Gaussian $X$.

**Good news:** If $K$'s barycenter $b(K) = E[X \cdot 1_{\{X \in K\}}]$ is at the origin, then $\Pr[Y \in O(1) \cdot (K \cap -K)] \geq 1/2$.

We design a recentering procedure that:

- Either finds signs $\varepsilon_1, \ldots, \varepsilon_n$ such that $\sum_i \varepsilon_i u_i \in K$,
- Or reduces to the case when $b(K) = 0$. 

---
Asymmetric Bodies

Does an efficient sampler for $O(1)$-subgaussian $Y$ imply algorithmic Banaszczyk for asymmetric $K$?

**Bad News:** Take $K = \{ x \in \mathbb{R}^m : x_1 \leq 0 \}$ and $Y = e_1$. Then:

- $Y$ is $O(1)$-subgaussian
- $\Pr[X \in K] = 1/2$ for standard Gaussian $X$.
- For any $t > 0$, $Y \notin tK = K$. 

---

DGLN  Vector Balancing  11 / 12
Asymmetric Bodies

Does an efficient sampler for $O(1)$-subgaussian $Y$ imply algorithmic Banaszczyk for asymmetric $K$?

**Bad News:** Take $K = \{x \in \mathbb{R}^m : x_1 \leq 0\}$ and $Y = e_1$. Then:
- $Y$ is $O(1)$-subgaussian
- $\Pr[X \in K] = 1/2$ for standard Gaussian $X$.
- For any $t > 0$, $Y \notin tK = K$.

**Good news:** If $K$’s barycenter $b(K) = \mathbb{E}[X \cdot 1\{X \in K\}]$ is at the origin, then $\Pr[Y \in O(1) \cdot (K \cap -K)] \geq 1/2$. 

We design a recentering procedure that

* Either finds signs $\varepsilon_1, \ldots, \varepsilon_n$ such that $\sum_i \varepsilon_i u_i \in K$,
* Or reduces to the case when $b(K) = 0$. 

DGLN Vector Balancing
Asymmetric Bodies

Does an efficient sampler for $O(1)$-subgaussian $Y$ imply algorithmic Banaszczyk for asymmetric $K$?

**Bad News**: Take $K = \{ x \in \mathbb{R}^m : x_1 \leq 0 \}$ and $Y = e_1$. Then:
- $Y$ is $O(1)$-subgaussian
- $\Pr[X \in K] = 1/2$ for standard Gaussian $X$.
- For any $t > 0$, $Y \not\in tK = K$.

**Good news**: If $K$’s barycenter $b(K) = \mathbb{E}[X \cdot 1\{X \in K\}]$ is at the origin, then $\Pr[Y \in O(1) \cdot (K \cap -K)] \geq 1/2$.

We design a recentering procedure that
- Either finds signs $\varepsilon_1, \ldots, \varepsilon_n$ such that $\sum_i \varepsilon_i u_i \in K$,
- Or reduces to the case when $b(K) = 0$. 
Open Problems

- Find a direct proof that there exists an $O(1)$-subgaussian $Y$ supported on $\{\sum_i \pm u_i\}$.
- Find an efficient algorithm to sample $Y$.

Thank you!


