Banach-Mazur distance to the cube

Apostolos Giannopoulos

National and Kapodistrian University of Athens

September 6, 2018
If $X$ and $Y$ are two $n$-dimensional normed spaces then their *Banach-Mazur distance* $d(X, Y)$ is defined by

$$d(X, Y) = \min \{ \| T \| \| T^{-1} \| \mid T : X \to Y \text{ is an isomorphism} \}.$$
Banach-Mazur distance

If $X$ and $Y$ are two $n$-dimensional normed spaces then their *Banach-Mazur distance* $d(X, Y)$ is defined by

$$d(X, Y) = \min \{ \| T\| \| T^{-1}\| \mid T : X \to Y \text{ is an isomorphism} \}.$$  

**Geometric interpretation**

Let $B_X$ and $B_Y$ denote the unit balls of $X$ and $Y$. Then, $d(X, Y)$ is the smallest possible $r \geqslant 1$ for which there exists an isomorphism $T : X \to Y$ such that

$$B_Y \subseteq T(B_X) \subseteq rB_Y.$$
Banach-Mazur distance

- If $X$ and $Y$ are two $n$-dimensional normed spaces then their Banach-Mazur distance $d(X, Y)$ is defined by

$$d(X, Y) = \min\{ \| T \| \| T^{-1} \| \mid T : X \to Y \text{ is an isomorphism} \}.$$ 

**Geometric interpretation**

Let $B_X$ and $B_Y$ denote the unit balls of $X$ and $Y$. Then, $d(X, Y)$ is the smallest possible $r \geq 1$ for which there exists an isomorphism $T : X \to Y$ such that

$$B_Y \subseteq T(B_X) \subseteq rB_Y.$$ 

**Basic properties**

- $d(X, Y) \geq 1$ with equality if and only if $X$ is isometrically isomorphic to $Y$.
- $d(X, Y) = d(Y, X)$.
- $d(X, Z) \leq d(X, Y)d(Y, Z)$.
- $d(X^*, Y^*) = d(X, Y)$. 
The $n$-th Banach-Mazur (or Minkowski) compactum is the set $\mathcal{B}_n$ of all equivalence classes of isometrically isomorphic $n$-dimensional normed spaces.

$\mathcal{B}_n$ becomes a compact metric space with the metric $\log d$.

Usually, instead of $\log d$, we consider $d$ as a “multiplicative” distance on $\mathcal{B}_n$. 

Diameter of the compactum

Upper bound:

\[
\operatorname{diam}(\mathcal{B}_n) \leq n.
\]

This is a consequence of John's theorem which can be stated as follows: for any $n$-dimensional normed space $X$, 

$$d(X,\ell_n^2) \leq \sqrt{n}.$$ 

Then, for any $X$ and $Y$, 

$$d(X,Y) \leq d(X,\ell_n^2)d(\ell_n^2,Y) \leq \sqrt{n} \cdot \sqrt{n} = n.$$ 

Notation:

$$\ell_n^p = (\mathbb{R}^n, \|\cdot\|_p),$$ where 

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$$ if $1 \leq p < \infty$ and 

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$
The $n$-th Banach-Mazur (or Minkowski) compactum is the set $\mathcal{B}_n$ of all equivalence classes of isometrically isomorphic $n$-dimensional normed spaces.

$\mathcal{B}_n$ becomes a compact metric space with the metric $\log d$.

Usually, instead of $\log d$, we consider $d$ as a “multiplicative” distance on $\mathcal{B}_n$.

### Diameter of the compactum

$$\text{diam}(\mathcal{B}_n) \leq n.$$  

This is a consequence of John's theorem which can be stated as follows: for any $n$-dimensional normed space $X$,

$$d(X, \ell^2_n) \leq \sqrt{n}.$$  

Then, for any $X$ and $Y$,

$$d(X, Y) \leq d(X, \ell^2_n) d(\ell^2_n, Y) \leq \sqrt{n} \cdot \sqrt{n} = n.$$
The \(n\)-th Banach-Mazur (or Minkowski) compactum is the set \(\mathcal{B}_n\) of all equivalence classes of isometrically isomorphic \(n\)-dimensional normed spaces.

\(\mathcal{B}_n\) becomes a compact metric space with the metric \(\log d\).

Usually, instead of \(\log d\), we consider \(d\) as a “multiplicative” distance on \(\mathcal{B}_n\).

**Diameter of the compactum**

Upper bound: \(\text{diam}(\mathcal{B}_n) \leq n\).
The $n$-th Banach-Mazur (or Minkowski) compactum is the set $\mathcal{B}_n$ of all equivalence classes of isometrically isomorphic $n$-dimensional normed spaces.

$\mathcal{B}_n$ becomes a compact metric space with the metric $\log d$.

Usually, instead of $\log d$, we consider $d$ as a “multiplicative” distance on $\mathcal{B}_n$.

**Diameter of the compactum**

Upper bound: $\text{diam}(\mathcal{B}_n) \leq n$.

This is a consequence of John’s theorem which can be stated as follows: for any $n$-dimensional normed space $X$,

$$d(X, \ell_2^n) \leq \sqrt{n}.$$  

Then, for any $X$ and $Y$,

$$d(X, Y) \leq d(X, \ell_2^n)d(\ell_2^n, Y) \leq \sqrt{n} \cdot \sqrt{n} = n.$$  

**Notation: $\ell_p^n$**

$$\ell_p^n = (\mathbb{R}^n, \| \cdot \|_p), \text{ where } \|x\|_p = \left( \sum_{i=1}^{n} |x_i|^p \right)^{1/p} \text{ if } 1 \leq p < \infty \text{ and } \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$
There exists an absolute constant $c > 0$ with the following property: for any $n \in \mathbb{N}$ one may find two $n$-dimensional normed spaces $X_n, Y_n$ with $d(X_n, Y_n) \geq cn$. Consequently, $\text{diam}(B_n) \geq cn$. 

The proof introduces a class of random spaces, sometimes called Gluskin spaces. Let $x_1, \ldots, x_m$ be random vectors which are independently and uniformly distributed in the Euclidean unit sphere $S^{n-1}$. We consider the symmetric random polytope $B_m := B_m(x_1, \ldots, x_m) = \text{conv}\{\pm e_1, \pm e_2, \ldots, \pm e_n, \pm x_1, \ldots, \pm x_m\}$, where $\{e_i\}_{i \leq n}$ is the standard orthonormal basis of $\mathbb{R}^n$. The space whose unit ball is $B_m$ is denoted by $X_{B_m}$. We write $A_m$ for the set of all these spaces equipped with the probability measure $\mu \equiv \otimes_{i=1}^m \sigma$.

Gluskin proves that if $m = 2n$ and $B'_m$ is an independent copy of $B_m$ then $d(X_m, X_m') \geq cn$ with probability greater than $1 - 2^{-n^2}$. 

Apostolos Giannopoulos (University of Athens)
Diameter of the Banach-Mazur compactum

**Gluskin’s theorem**

There exists an absolute constant $c > 0$ with the following property: for any $n \in \mathbb{N}$ one may find two $n$-dimensional normed spaces $X_n, Y_n$ with $d(X_n, Y_n) \geq cn$. Consequently, $\text{diam}(B_n) \geq cn$.

- The proof introduces a class of random spaces, sometimes called *Gluskin spaces*. Let $x_1, \ldots, x_m$ be random vectors which are independently and uniformly distributed in the Euclidean unit sphere $S^{n-1}$. We consider the symmetric random polytope

  $$B_m := B_m(x_1, \ldots, x_m) = \text{conv}\{\pm e_1, \pm e_2, \ldots, \pm e_n, \pm x_1, \ldots, \pm x_m\},$$

  where $\{e_i\}_{i \leq n}$ is the standard orthonormal basis of $\mathbb{R}^n$. The space whose unit ball is $B_m$ is denoted by $X_{B_m}$. We write $\mathcal{A}_m$ for the set of all these spaces equipped with the probability measure $\mu \equiv \otimes_{i=1}^{m} \sigma$. 

Apostolos Giannopoulos (University of Athens) 

Banach-Mazur distance to the cube 

Castro Urdiales, September 2018
Gluskin’s theorem

There exists an absolute constant $c > 0$ with the following property: for any $n \in \mathbb{N}$ one may find two $n$-dimensional normed spaces $X_n, Y_n$ with $d(X_n, Y_n) \geq cn$. Consequently, $\text{diam}(B_n) \geq cn$.

The proof introduces a class of random spaces, sometimes called Gluskin spaces. Let $x_1, \ldots, x_m$ be random vectors which are independently and uniformly distributed in the Euclidean unit sphere $S^{n-1}$. We consider the symmetric random polytope

$$B_m := B_m(x_1, \ldots, x_m) = \text{conv}\{\pm e_1, \pm e_2, \ldots, \pm e_n, \pm x_1, \ldots, \pm x_m\},$$

where $\{e_i\}_{i \leq n}$ is the standard orthonormal basis of $\mathbb{R}^n$. The space whose unit ball is $B_m$ is denoted by $X_{B_m}$. We write $\mathcal{A}_m$ for the set of all these spaces equipped with the probability measure $\mu \equiv \otimes_{i=1}^m \sigma$.

Gluskin proves that if $m = 2n$ and $B'_m$ is an independent copy of $B_m$ then

$$d(X_{B_m}, X_{B'_m}) \geq cn$$

with probability greater than $1 - 2^{-n^2}$. 
Let $X_0 \in \mathcal{B}_n$. We denote by $R(X_0)$ the “radius” of the Banach-Mazur compactum $\mathcal{B}_n$ with respect to $X_0$, defined by

$$R(X_0) = \max\{d(X, X_0) : X \in \mathcal{B}_n\}.$$
Let $X_0 \in B_n$. We denote by $\mathcal{R}(X_0)$ the “radius” of the Banach-Mazur compactum $B_n$ with respect to $X_0$, defined by

$$\mathcal{R}(X_0) = \max\{d(X, X_0) \mid X \in B_n\}.$$  

John’s theorem implies that $\mathcal{R}(\ell_n^2) = \sqrt{n}$ because one can show that $d(\ell_\infty^n, \ell_2^n) = d(\ell_1^n, \ell_2^n) = \sqrt{n}$. 

We shall discuss the radius of $B_n$ with respect to $\ell_n^\infty$, defined by

$$\mathcal{R}_n^\infty = \max\{d(X, \ell_n^\infty) \mid X \in B_n\}.$$ 

Pełczynski

What is the asymptotic behavior of $\mathcal{R}_n^\infty$ as $n$ tends to infinity?

One clearly has $\mathcal{R}_n^\infty \leq \text{diam}(B_n) \leq n$ and the fact that $d(\ell_\infty^n, \ell_2^n) = \sqrt{n}$ shows that $\sqrt{n} \leq \mathcal{R}_n^\infty \leq n$. 

Pełczynski
Let $X_0 \in B_n$. We denote by $\mathcal{R}(X_0)$ the “radius” of the Banach-Mazur compactum $B_n$ with respect to $X_0$, defined by

$$\mathcal{R}(X_0) = \max\{d(X, X_0) : X \in B_n\}.$$ 

John’s theorem implies that $\mathcal{R}(\ell^n_2) = \sqrt{n}$ because one can show that

$$d(\ell^n_\infty, \ell^n_2) = d(\ell^n_1, \ell^n_2) = \sqrt{n}.$$ 

We shall discuss the radius of $B_n$ with respect to $\ell^n_\infty$, defined by

$$\mathcal{R}^n_\infty = \max\{d(X, \ell^n_\infty) : X \in B_n\}.$$
Let $X_0 \in B_n$. We denote by $\mathcal{R}(X_0)$ the “radius” of the Banach-Mazur compactum $B_n$ with respect to $X_0$, defined by

$$\mathcal{R}(X_0) = \max\{d(X, X_0) : X \in B_n\}.$$ 

John’s theorem implies that $\mathcal{R}(\ell_2^n) = \sqrt{n}$ because one can show that

$$d(\ell_\infty^n, \ell_2^n) = d(\ell_1^n, \ell_2^n) = \sqrt{n}.$$ 

We shall discuss the radius of $B_n$ with respect to $\ell_\infty^n$, defined by

$$\mathcal{R}_\infty^n = \max\{d(X, \ell_\infty^n) : X \in B_n\}.$$ 

Pełczynski

What is the asymptotic behavior of $\mathcal{R}_\infty^n$ as $n$ tends to infinity?
Let $X_0 \in \mathcal{B}_n$. We denote by $R(X_0)$ the “radius” of the Banach-Mazur compactum $\mathcal{B}_n$ with respect to $X_0$, defined by

$$R(X_0) = \max \{ d(X, X_0) : X \in \mathcal{B}_n \}.$$ 

John’s theorem implies that $R(\ell_2^n) = \sqrt{n}$ because one can show that

$$d(\ell_\infty^n, \ell_2^n) = d(\ell_1^n, \ell_2^n) = \sqrt{n}.$$ 

We shall discuss the radius of $\mathcal{B}_n$ with respect to $\ell_\infty^n$, defined by

$$R_\infty^n = \max \{ d(X, \ell_\infty^n) : X \in \mathcal{B}_n \}.$$ 

Pełczyński

What is the asymptotic behavior of $R_\infty^n$ as $n$ tends to infinity?

One clearly has $R_\infty^n \leq \text{diam}(\mathcal{B}_n) \leq n$ and the fact that $d(\ell_\infty^n, \ell_2^n) = \sqrt{n}$ shows that

$$\sqrt{n} \leq R_\infty^n \leq n.$$
**Upper bounds** were obtained by:

- Bourgain-Szarek: $R_n^\infty \leq n \cdot \exp(-c\sqrt{\log n})$.

There exists an absolute constant $c > 0$ such that, for any $n \geq 2$, $R_n^\infty \leq cn^{5/6}$. Lower bounds: Szarek, using random spaces of Gluskin type, proved that $R_n^\infty \geq c\sqrt{n \log n}$. Tikhomirov, 2018, showed that there exist absolute constants $c, b > 0$ such that, for any $n \geq 2$, $R_n^\infty \geq cn^{5/9} \log^{-b} n$. This means that $R_n^\infty$ has order of growth much larger than $\sqrt{n}$; in other words, $\ell_n^\infty$ is not an asymptotic center of the Banach-Mazur compactum, in a very strong sense.
**Upper bounds** were obtained by:

- Bourgain-Szarek: $R_n^\infty \leq n \cdot \exp(-c\sqrt{\log n})$.
- Szarek-Talagrand: $R_n^\infty \leq cn^{7/8}$.
Upper bounds were obtained by:

- Bourgain-Szarek: $\mathcal{R}_\infty^n \leq n \cdot \exp(-c\sqrt{\log n})$.
- Szarek-Talagrand: $\mathcal{R}_\infty^n \leq cn^{7/8}$.

G., 1993

There exists an absolute constant $c > 0$ such that, for any $n \geq 2$,

$$\mathcal{R}_\infty^n \leq cn^{5/6}.$$
Upper bounds were obtained by:

- Bourgain-Szarek: $\mathcal{R}_\infty^n \leq n \cdot \exp(-c\sqrt{\log n})$.
- Szarek-Talagrand: $\mathcal{R}_\infty^n \leq cn^{7/8}$.

G., 1993

There exists an absolute constant $c > 0$ such that, for any $n \geq 2$,

$$\mathcal{R}_\infty^n \leq cn^{5/6}.$$  

Lower bounds: Szarek, using random spaces of Gluskin type, proved that

$$\mathcal{R}_\infty^n \geq c\sqrt{n \log n}.$$
Upper bounds were obtained by:

- Bourgain-Szarek: $R_n^\infty \leq n \cdot \exp(-c\sqrt{\log n})$.
- Szarek-Talagrand: $R_n^\infty \leq cn^{7/8}$.

G., 1993

There exists an absolute constant $c > 0$ such that, for any $n \geq 2$,

$$R_n^\infty \leq cn^{5/6}.$$ 

Lower bounds: Szarek, using random spaces of Gluskin type, proved that

$$R_n^\infty \geq c\sqrt{n \log n}.$$ 

Tikhomirov, 2018

There exist absolute constants $c, b > 0$ such that, for any $n \geq 2$,

$$R_n^\infty \geq cn^{5/9} \log^{-b} n.$$
Banach-Mazur distance to the cube

**Upper bounds** were obtained by:

- Bourgain-Szarek: \( R_n^\infty \leq n \cdot \exp(-c\sqrt{\log n}) \).
- Szarek-Talagrand: \( R_n^\infty \leq cn^{7/8} \).

**G., 1993**

There exists an absolute constant \( c > 0 \) such that, for any \( n \geq 2 \),

\[
R_n^\infty \leq cn^{5/6}.
\]

**Lower bounds**: Szarek, using random spaces of Gluskin type, proved that

\[
R_n^\infty \geq c\sqrt{n} \log n.
\]

**Tikhomirov, 2018**

There exist absolute constants \( c, b > 0 \) such that, for any \( n \geq 2 \),

\[
R_n^\infty \geq cn^{5/9} \log^{-b} n.
\]

- This means that \( R_n^\infty \) has order of growth much larger than \( \sqrt{n} \); in other words, \( \ell_\infty^n \) is not an asymptotic center of the Banach-Mazur compactum, in a very strong sense.
Upper bound for $\mathcal{R}_n^\infty$

It is more convenient to work with the dual quantity

$$\mathcal{R}_1^n = \max\{d(X, \ell_1^n) : X \in B_n\}.$$
It is more convenient to work with the dual quantity
\[ R_n^1 = \max\{d(X, \ell_1^n) : X \in B_n\}. \]

Since \( d(X^*, Y^*) = d(X, Y) \) we see that \( R_n^\infty = R_n^1 \).
Upper bound for $\mathcal{R}_\infty^n$

- It is more convenient to work with the dual quantity
  \[ \mathcal{R}_1^n = \max \{ d(X, \ell_1^n) : X \in \mathcal{B}_n \}. \]

- Since $d(X^*, Y^*) = d(X, Y)$ we see that $\mathcal{R}_\infty^n = \mathcal{R}_1^n$.

- We want an upper bound for $d(X, \ell_1^n)$ where $X = (\mathbb{R}^n, \| \cdot \|)$, and we may also assume that the minimal volume ellipsoid of the unit ball $K$ of $X$ is the Euclidean unit ball $B_2^n$. 
Upper bound for $\mathcal{R}_\infty^n$

- It is more convenient to work with the dual quantity
  \[ \mathcal{R}_1^n = \max\{ d(X, \ell_1^n) : X \in \mathcal{B}_n \}. \]

- Since $d(X^*, Y^*) = d(X, Y)$ we see that $\mathcal{R}_\infty^n = \mathcal{R}_1^n$.

- We want an upper bound for $d(X, \ell_1^n)$ where $X = (\mathbb{R}^n, \| \cdot \|)$, and we may also assume that the minimal volume ellipsoid of the unit ball $K$ of $X$ is the Euclidean unit ball $B_2^n$.

- We need to find $n$ vectors $u_1, \ldots, u_n \in \mathbb{R}^n$ such that, for all $t_1, \ldots, t_n \in \mathbb{R}$,
  \[ \frac{1}{cn^{5/6}} \sum_{i=1}^{n} |t_i| \leq \left\| \sum_{i=1}^{n} t_i u_i \right\| \leq \sum_{i=1}^{n} |t_i|. \]
Upper bound for $\mathcal{R}_\infty^n$

- It is more convenient to work with the dual quantity
  \[ \mathcal{R}_1^n = \max\{d(X, \ell_1^n) : X \in \mathcal{B}_n\} . \]

- Since $d(X^*, Y^*) = d(X, Y)$ we see that $\mathcal{R}_\infty^n = \mathcal{R}_1^n$.

- We want an upper bound for $d(X, \ell_1^n)$ where $X = (\mathbb{R}^n, \| \cdot \|)$, and we may also assume that the minimal volume ellipsoid of the unit ball $K$ of $X$ is the Euclidean unit ball $B_2^n$.

- We need to find $n$ vectors $u_1, \ldots, u_n \in \mathbb{R}^n$ such that, for all $t_1, \ldots, t_n \in \mathbb{R}$,
  \[ \frac{1}{cn^{5/6}} \sum_{i=1}^n |t_i| \leq \left\| \sum_{i=1}^n t_i u_i \right\| \leq \sum_{i=1}^n |t_i| . \]

- Then, the operator $T : \ell_1^n \to X$ defined by $T(e_i) = u_i$ satisfies $\| T \| \leq 1$ and $\| T^{-1} \| \leq cn^{5/6}$, which implies the bound
  \[ d(X, \ell_1^n) \leq \| T \| \| T^{-1} \| \leq cn^{5/6} . \]
Upper bound for $\mathcal{R}_n^\infty$

- It is more convenient to work with the dual quantity
  \[ \mathcal{R}_1^n = \max \{ d(X, \ell_1^n) : X \in \mathcal{B}_n \}. \]

- Since $d(X^*, Y^*) = d(X, Y)$ we see that $\mathcal{R}_n^\infty = \mathcal{R}_1^n$.

- We want an upper bound for $d(X, \ell_1^n)$ where $X = (\mathbb{R}^n, \| \cdot \|)$, and we may also assume that the minimal volume ellipsoid of the unit ball $K$ of $X$ is the Euclidean unit ball $B_2^n$.

- We need to find $n$ vectors $u_1, \ldots, u_n \in \mathbb{R}^n$ such that, for all $t_1, \ldots, t_n \in \mathbb{R}$,
  \[ \frac{1}{cn^{5/6}} \sum_{i=1}^n |t_i| \leq \left\| \sum_{i=1}^n t_i u_i \right\| \leq \sum_{i=1}^n |t_i|. \]

- Then, the operator $T : \ell_1^n \to X$ defined by $T(e_i) = u_i$ satisfies $\| T \| \leq 1$ and $\| T^{-1} \| \leq cn^{5/6}$, which implies the bound
  \[ d(X, \ell_1^n) \leq \| T \| \| T^{-1} \| \leq cn^{5/6}. \]

- The main ingredients for the proof are the combinatorial Sauer-Shelah lemma and a Dvoretzky-Rogers type lemma of Szarek and Talagrand on the distribution of the contact points of $K$ and $B_2^n$ when $K$ is in Löwner position.
The lemma of Szarek and Talagrand

Recall John's representation of the identity: since $B_2^n$ is the minimal volume ellipsoid of $K$, there exist contact points $x_1, \ldots, x_m$ of $K$ and $B_2^n$, and positive real numbers $c_1, \ldots, c_m$ such that

$$x = \sum_{i=1}^{m} c_i \langle x, x_i \rangle x_i$$

(1)

for all $x \in \mathbb{R}^n$. 
The lemma of Szarek and Talagrand

Recall John’s representation of the identity: since $B^n_2$ is the minimal volume ellipsoid of $K$, there exist contact points $x_1, \ldots, x_m$ of $K$ and $B^n_2$, and positive real numbers $c_1, \ldots, c_m$ such that

$$x = \sum_{i=1}^{m} c_i \langle x, x_i \rangle x_i$$

for all $x \in \mathbb{R}^n$.

Szarek-Talagrand

Let $B^n_2$ be the minimal volume ellipsoid of $K$. For every $\epsilon \in (0,1)$, we can find $k \geq (1 - \epsilon)n$ and contact points $y_1, \ldots, y_k$ of $K$ and $B^n_2$ with the following property: If $j \in \{1, \ldots, k\}$ and $F_j = \text{span}\{y_i : i \neq j\}$, then

$$|P_{F_j^\perp}(y_j)| \geq \sqrt{\epsilon} \quad \text{for all } 1 \leq j \leq k.$$
The lemma of Szarek and Talagrand

- Recall John’s representation of the identity: since $B_2^n$ is the minimal volume ellipsoid of $K$, there exist contact points $x_1, \ldots, x_m$ of $K$ and $B_2^n$, and positive real numbers $c_1, \ldots, c_m$ such that

$$x = \sum_{i=1}^{m} c_i \langle x, x_i \rangle x_i$$  \hspace{1cm} (1)$$

for all $x \in \mathbb{R}^n$.

Szarek-Talagrand

Let $B_2^n$ be the minimal volume ellipsoid of $K$. For every $\epsilon \in (0, 1)$, we can find $k \geq (1 - \epsilon)n$ and contact points $y_1, \ldots, y_k$ of $K$ and $B_2^n$ with the following property: If $j \in \{1, \ldots, k\}$ and $F_j = \text{span}\{y_i : i \neq j\}$, then

$$|P_{F_j^\perp}(y_j)| \geq \sqrt{\epsilon} \quad \text{for all } 1 \leq j \leq k.$$

- Among all $k$-sets $\{x_{i_1}, \ldots, x_{i_k}\}$ of contact points in (1) choose one, say $\{y_1, \ldots, y_k\}$, which maximizes $\text{vol}_k(\text{conv}\{\pm x_{i_1}, \ldots, \pm x_{i_k}\})$. 
The lemma of Szarek and Talagrand

- Recall John’s representation of the identity: since $B_2^n$ is the minimal volume ellipsoid of $K$, there exist contact points $x_1, \ldots, x_m$ of $K$ and $B_2^n$, and positive real numbers $c_1, \ldots, c_m$ such that

$$x = \sum_{i=1}^{m} c_i \langle x, x_i \rangle x_i$$

(1)

for all $x \in \mathbb{R}^n$.

Szarek-Talagrand

Let $B_2^n$ be the minimal volume ellipsoid of $K$. For every $\epsilon \in (0, 1)$, we can find $k \geq (1 - \epsilon)n$ and contact points $y_1, \ldots, y_k$ of $K$ and $B_2^n$ with the following property: If $j \in \{1, \ldots, k\}$ and $F_j = \text{span}\{y_i : i \neq j\}$, then

$$|P_{F_j^\perp}(y_j)| \geq \sqrt{\epsilon} \quad \text{for all } 1 \leq j \leq k.$$ 

- Among all $k$-sets $\{x_{i_1}, \ldots, x_{i_k}\}$ of contact points in (1) choose one, say $\{y_1, \ldots, y_k\}$, which maximizes $\text{vol}_k(\text{conv}\{\pm x_{i_1}, \ldots, \pm x_{i_k}\})$.

- Then, for all $1 \leq j \leq k$ and all $1 \leq i \leq m$ we have

$$|P_{F_j^\perp}(y_j)| \geq |P_{F_j^\perp}(x_i)|.$$
Let $B_n^2$ be the minimal volume ellipsoid of $K$. For every $\epsilon \in (0, 1)$, we can find $k \geq (1 - \epsilon)n$ and contact points $y_1, \ldots, y_k$ of $K$ and $B_n^2$ with the following property: If $j \in \{1, \ldots, k\}$ and $F_j = \text{span}\{y_i : i \neq j\}$, then $|P_{F_j^\perp}(y_j)| \geq \sqrt{\epsilon}$ for all $1 \leq j \leq k$. 

Apostolos Giannopoulos (University of Athens)
The lemma of Szarek and Talagrand

Szarek-Talagrand

Let $B_n^2$ be the minimal volume ellipsoid of $K$. For every $\epsilon \in (0, 1)$, we can find $k \geq (1 - \epsilon)n$ and contact points $y_1, \ldots, y_k$ of $K$ and $B_n^2$ with the following property: If $j \in \{1, \ldots, k\}$ and $F_j = \text{span}\{y_i : i \neq j\}$, then $|P_{F_j^+}(y_j)| \geq \sqrt{\epsilon}$ for all $1 \leq j \leq k$.

- Among all $k$-sets $\{x_{i_1}, \ldots, x_{i_k}\}$ of contact points in (1) choose one, say $\{y_1, \ldots, y_k\}$, which maximizes $\text{vol}_k(\text{conv}\{\pm x_{i_1}, \ldots, \pm x_{i_k}\})$. 

The lemma of Szarek and Talagrand

Szarek-Talagrand

Let $B^2_n$ be the minimal volume ellipsoid of $K$. For every $\epsilon \in (0, 1)$, we can find $k \geq (1 - \epsilon)n$ and contact points $y_1, \ldots, y_k$ of $K$ and $B^2_n$ with the following property: If $j \in \{1, \ldots, k\}$ and $F_j = \text{span}\{y_i : i \neq j\}$, then $|P_{F_j} (y_j)| \geq \sqrt{\epsilon}$ for all $1 \leq j \leq k$.

- Among all $k$-sets $\{x_{i_1}, \ldots, x_{i_k}\}$ of contact points in (1) choose one, say $\{y_1, \ldots, y_k\}$, which maximizes $\text{vol}_k(\text{conv}\{\pm x_{i_1}, \ldots, \pm x_{i_k}\})$.
- Then, for all $1 \leq j \leq k$ and all $1 \leq i \leq m$ we have $|P_{F_j} (y_j)| \geq |P_{F_j} (x_i)|$.
Let $B_2^n$ be the minimal volume ellipsoid of $K$. For every $\epsilon \in (0, 1)$, we can find $k \geq (1 - \epsilon)n$ and contact points $y_1, \ldots, y_k$ of $K$ and $B_2^n$ with the following property: If $j \in \{1, \ldots, k\}$ and $F_j = \text{span}\{y_i : i \neq j\}$, then $|P_{F_j^\perp}(y_j)| \geq \sqrt{\epsilon}$ for all $1 \leq j \leq k$.

Among all $k$-sets $\{x_{i_1}, \ldots, x_{i_k}\}$ of contact points in (1) choose one, say $\{y_1, \ldots, y_k\}$, which maximizes $\text{vol}_k(\text{conv}\{\pm x_{i_1}, \ldots, \pm x_{i_k}\})$.

Then, for all $1 \leq j \leq k$ and all $1 \leq i \leq m$ we have $|P_{F_j^\perp}(y_j)| \geq |P_{F_j^\perp}(x_i)|$.

Note that $P_{F_j^\perp}(x) = \sum_{i=1}^m c_i \langle x, x_i \rangle P_{F_j^\perp}(x_i)$. Using this, we see that

$$n - k + 1 = \text{tr}(P_{F_j^\perp}) = \sum_{i=1}^m c_i \langle x_i, P_{F_j^\perp}(x_i) \rangle = \sum_{i=1}^m c_i |P_{F_j^\perp}(x_i)|^2,$$

and since $\sum_{i=1}^m c_i = n$ there exists $x_i$ such that

$$|P_{F_j^\perp}(x_i)|^2 = \langle x_i, P_{F_j^\perp}(x_i) \rangle \geq \text{tr}(P_{F_j^\perp})/n = (n - k + 1)/n.$$
Szarek-Talagrand

Let $B^n_2$ be the minimal volume ellipsoid of $K$. For every $\epsilon \in (0, 1)$, we can find $k \geq (1 - \epsilon)n$ and contact points $y_1, \ldots, y_k$ of $K$ and $B^n_2$ with the following property: If $j \in \{1, \ldots, k\}$ and $F_j = \text{span}\{y_i : i \neq j\}$, then $|P_{F_j^\perp}(y_j)| \geq \sqrt{\epsilon}$ for all $1 \leq j \leq k$.

- Among all $k$-sets $\{x_{i_1}, \ldots, x_{i_k}\}$ of contact points in (1) choose one, say $\{y_1, \ldots, y_k\}$, which maximizes $\text{vol}_k(\text{conv}\{\pm x_{i_1}, \ldots, \pm x_{i_k}\})$.
- Then, for all $1 \leq j \leq k$ and all $1 \leq i \leq m$ we have $|P_{F_j^\perp}(y_j)| \geq |P_{F_j^\perp}(x_i)|$.
- Note that $P_{F_j^\perp}(x) = \sum_{i=1}^m c_i \langle x, x_i \rangle P_{F_j^\perp}(x_i)$. Using this, we see that

$$n - k + 1 = \text{tr}(P_{F_j^\perp}) = \sum_{i=1}^m c_i \langle x_i, P_{F_j^\perp}(x_i) \rangle = \sum_{i=1}^m c_i |P_{F_j^\perp}(x_i)|^2,$$

and since $\sum_{i=1}^m c_i = n$ there exists $x_i$ such that

$$|P_{F_j^\perp}(x_i)|^2 = \langle x_i, P_{F_j^\perp}(x_i) \rangle \geq \text{tr}(P_{F_j^\perp})/n = (n - k + 1)/n.$$

- Taking $k = \lfloor (1 - \epsilon)n \rfloor + 1$, we see that $k \geq (1 - \epsilon)n$ and, for all $1 \leq j \leq k$,

$$|P_{F_j^\perp}(y_j)| = \max_{i \leq m} |P_{F_j^\perp}(x_i)| \geq \sqrt{(n - k + 1)/n} \geq \sqrt{\epsilon}.$$
The Sauer-Shelah lemma

Let $X$ be a set with cardinality $|X| = n$ and $1 \leq k \leq n$. If $\mathcal{F}$ is a family of subsets of $X$ with

$$|\mathcal{F}| > \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k-1}$$

then we can find $A \subseteq X$ with $|A| \geq k$ and $A \cap \mathcal{F} = \mathcal{P}(A)$, where $\mathcal{P}(A)$ is the family of all subsets of $A$. 

Sauer-Shelah II

Let $A$ be a subset of $E^n_2 = \{-1, 1\}^n$ with cardinality $|A| > \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k-1}$. There exists $\sigma \subseteq \{1, \ldots, n\}$ with $|\sigma| \geq k$, such that the map $P_\sigma$ is onto. That is, $P_\sigma(A) = \{-1, 1\}^\sigma$. 

Apostolos Giannopoulos (University of Athens)
The Sauer-Shelah lemma

Sauer-Shelah

Let $X$ be a set with cardinality $|X| = n$ and $1 \leq k \leq n$. If $\mathcal{F}$ is a family of subsets of $X$ with

$$|\mathcal{F}| > \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k-1}$$

then we can find $A \subseteq X$ with $|A| \geq k$ and $A \cap \mathcal{F} = \mathcal{P}(A)$, where $\mathcal{P}(A)$ is the family of all subsets of $A$.

- Consider the discrete cube $E_2^n = \{-1, 1\}^n$. For any $\sigma \subseteq [n]$ we consider the coordinates restriction function $P_\sigma : E_2^n = \{-1, 1\}^n \rightarrow \{-1, 1\}^\sigma$ with $(\epsilon_1, \ldots, \epsilon_n) \mapsto (\epsilon_j)_{j \in \sigma}$. Since the map $\varphi : \mathcal{P}\{1, \ldots, n\} \rightarrow E_2^n$ with $\varphi(\sigma)_i = 1$ if $i \in \sigma$ and $\varphi(\sigma)_i = -1$ if $i \notin \sigma$ is a bijection, we can immediately translate the Sauer-Shelah lemma as follows:
The Sauer-Shelah lemma

Sauer-Shelah

Let $X$ be a set with cardinality $|X| = n$ and $1 \leq k \leq n$. If $\mathcal{F}$ is a family of subsets of $X$ with

$$|\mathcal{F}| > \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k-1},$$

then we can find $A \subseteq X$ with $|A| \geq k$ and $A \cap \mathcal{F} = \mathcal{P}(A)$, where $\mathcal{P}(A)$ is the family of all subsets of $A$.

- Consider the discrete cube $E_2^n = \{-1, 1\}^n$. For any $\sigma \subseteq [n]$ we consider the coordinates restriction function $P_\sigma : E_2^n = \{-1, 1\}^n \to \{-1, 1\}^\sigma$ with $(\epsilon_1, \ldots, \epsilon_n) \mapsto (\epsilon_j)_{j \in \sigma}$. Since the map $\varphi : \mathcal{P}\{1, \ldots, n\} \to E_2^n$ with $\varphi(\sigma)_i = 1$ if $i \in \sigma$ and $\varphi(\sigma)_i = -1$ if $i \notin \sigma$ is a bijection, we can immediate translate the Sauer-Shelah lemma as follows:

Sauer-Shelah II

Let $A$ be a subset of $E_2^n = \{-1, 1\}^n$ with cardinality $|A| > \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k-1}$. There exists $\sigma \subseteq \{1, \ldots, n\}$ with $|\sigma| \geq k$, such that the map $P_\sigma$ is onto. That is,

$$P_\sigma(A) = \{-1, 1\}^\sigma.$$
Sauer-Shelah II

Let $A$ be a subset of $E_2^n = \{-1, 1\}^n$ with cardinality $|A| > \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k-1}$. There exists $\sigma \subseteq \{1, \ldots, n\}$ with $|\sigma| \geq k$, such that the map $P_\sigma$ is onto. That is,

$$P_\sigma(A) = \{-1, 1\}^\sigma.$$
Let $A$ be a subset of $E_2^n = \{-1, 1\}^n$ with cardinality $|A| > \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k-1}$. There exists $\sigma \subseteq \{1, \ldots, n\}$ with $|\sigma| \geq k$, such that the map $P_\sigma$ is onto. That is,

$$P_\sigma(A) = \{-1, 1\}^\sigma.$$

- It is useful to think of the elements of $E_2^n$ as the vertices of the cube $Q_n = [-1, 1]^n$ in $\mathbb{R}^n$. 

Apostolos Giannopoulos (University of Athens)  
Banach-Mazur distance to the cube  
Castro Urdiales, September 2018
The Sauer-Shelah lemma

Sauer-Shelah II

Let $A$ be a subset of $E_2^n = \{-1, 1\}^n$ with cardinality $|A| > \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k-1}$. There exists $\sigma \subseteq \{1, \ldots, n\}$ with $|\sigma| \geq k$, such that the map $P_\sigma$ is onto. That is,

$$P_\sigma(A) = \{-1, 1\}^\sigma.$$ 

- It is useful to think of the elements of $E_2^n$ as the vertices of the cube $Q_n = [-1, 1]^n$ in $\mathbb{R}^n$.
- Then, the coordinates restriction function $P_\sigma$ is the orthogonal projection onto $\mathbb{R}^\sigma$. 

Apostolos Giannopoulos (University of Athens)  
Banach-Mazur distance to the cube  
Castro Urdiales, September 2018  
11 / 43
The Sauer-Shelah lemma

**Sauer-Shelah II**

Let $A$ be a subset of $E_2^n = \{-1, 1\}^n$ with cardinality $|A| > \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k-1}$. There exists $\sigma \subseteq \{1, \ldots, n\}$ with $|\sigma| \geq k$, such that the map $P_\sigma$ is onto. That is,

$$P_\sigma(A) = \{-1, 1\}^\sigma.$$

- It is useful to think of the elements of $E_2^n$ as the vertices of the cube $Q_n = [-1, 1]^n$ in $\mathbb{R}^n$.
- Then, the coordinates restriction function $P_\sigma$ is the orthogonal projection onto $\mathbb{R}^\sigma$.
- In this setting, the Sauer-Shelah lemma tells us the following.

**Geometric Sauer-Shelah lemma**

If $A \subseteq \{-1, 1\}^n \subseteq \mathbb{R}^n$, and $|A| > \sum_{i=0}^{k-1} \binom{n}{i}$, then there exists $\sigma \subseteq \{1, \ldots, n\}$ with $|\sigma| \geq k$ such that the orthogonal projection $P_\sigma(\text{conv}(A))$ of the convex hull of $A$ onto $\mathbb{R}^\sigma$ is the full unit cube of $\mathbb{R}^\sigma$:

$$P_\sigma(\text{conv}(A)) = Q_\sigma := [-1, 1]^\sigma.$$
Let $u_1, \ldots, u_s \in B_2^n$ and $\mathcal{E} = \{(\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \}$. Then, for every $\epsilon \in (0, 1)$ there exists $\sigma \subseteq \{1, \ldots, s\}$ with cardinality $|\sigma| \geq (1 - \epsilon)s$, such that $P_\sigma(\mathcal{E}) \supseteq c\sqrt{\epsilon} [-1, 1]^\sigma$, where $c > 0$ is an absolute constant, and $P_\sigma$ is the orthogonal projection onto $\mathbb{R}^\sigma$. 

For the proof we use an inductive scheme; first, consider all points of the form $(\delta^{(1)}_j)_{j \leq s} \in \mathbb{R}^s$, with $\delta^{(1)}_j = \pm 1$. By the parallelogram law, $E_{\delta^{(1)}_j} = \pm 1 \left| \sum_{j=1}^s \delta^{(1)}_j u_j \right|^2 \leq 2s$. Using Markov’s inequality, we find $M_1 \subseteq \{-1, 1\}^s$ with cardinality $|M_1| \geq 2^s - 1$, such that for every $(\delta^{(1)}_j) \in M_1$, $\left| \sum_{j=1}^s \delta^{(1)}_j u_j \right|^2 \leq 2s$.

Using the geometric Sauer-Shelah lemma we find $\sigma_1 \subseteq S$ with cardinality $|\sigma_1| \geq s^2$, such that $P_{\sigma_1}(M_1) = \{-1, 1\}^{\sigma_1}$. Since $M_1 \subseteq E \cap Q$ and the last set is convex, we have $Q_{\sigma_1} \subseteq P_{\sigma_1}(E \cap Q)$.
Let $u_1, \ldots, u_s \in B^n_2$ and $\mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \right\}$. Then, for every $\epsilon \in (0, 1)$ there exists $\sigma \subseteq \{1, \ldots, s\}$ with cardinality $|\sigma| \geq (1 - \epsilon)s$, such that $P_\sigma(\mathcal{E}) \supseteq c\sqrt{\epsilon} [-1, 1]^\sigma$, where $c > 0$ is an absolute constant, and $P_\sigma$ is the orthogonal projection onto $\mathbb{R}^\sigma$.

For the proof we use an inductive scheme; first, consider all points of the form $(\delta_j^{(1)})_{j \leq s} \in \mathbb{R}^s$, with $\delta_j^{(1)} = \pm 1$. By the parallelogram law,

$$
\mathbb{E}_{\delta_j^{(1)} = \pm 1} \left| \sum_{j=1}^s \delta_j^{(1)} u_j \right|^2 = \sum_{j=1}^s |u_j|^2 \leq s.
$$
Let \( u_1, \ldots, u_s \in B_2^n \) and \( \mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \right\} \). Then, for every \( \epsilon \in (0, 1) \) there exists \( \sigma \subseteq \{1, \ldots, s\} \) with cardinality \( |\sigma| \geq (1 - \epsilon)s \), such that \( P_\sigma(\mathcal{E}) \supseteq c \sqrt{\epsilon} [-1, 1]^\sigma \), where \( c > 0 \) is an absolute constant, and \( P_\sigma \) is the orthogonal projection onto \( \mathbb{R}^\sigma \).

- For the proof we use an inductive scheme; first, consider all points of the form \( (\delta_j^{(1)})_{j \leq s} \in \mathbb{R}^s \), with \( \delta_j^{(1)} = \pm 1 \). By the parallelogram law,

\[
\mathbb{E}_{\delta_j^{(1)} = \pm 1} \left| \sum_{j=1}^s \delta_j^{(1)} u_j \right|^2 = \sum_{j=1}^s |u_j|^2 \leq s.
\]

- Using Markov’s inequality, we find \( M^1 \subseteq \{-1, 1\}^s \) with cardinality \( |M^1| \geq 2^{s-1} \), such that for every \( (\delta_j^{(1)}) \in M^1 \),

\[
\left| \sum_{j=1}^s \delta_j^{(1)} u_j \right|^2 \leq 2s.
\]
Let $u_1, \ldots, u_s \in B_2^n$ and $\mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^{s} \delta_j u_j \right|^2 \leq 2s \right\}$. Then, for every $\epsilon \in (0, 1)$ there exists $\sigma \subseteq \{1, \ldots, s\}$ with cardinality $|\sigma| \geq (1 - \epsilon)s$, such that $P_\sigma(\mathcal{E}) \supseteq c\sqrt{\epsilon} [-1, 1]^\sigma$, where $c > 0$ is an absolute constant, and $P_\sigma$ is the orthogonal projection onto $\mathbb{R}^\sigma$.

- For the proof we use an inductive scheme; first, consider all points of the form 
  
  $$(\delta_j^{(1)})_{j \leq s} \in \mathbb{R}^s, \text{ with } \delta_j^{(1)} = \pm 1.$$  

  By the parallelogram law,

  $$\mathbb{E}_{\delta_j^{(1)} = \pm 1} \left| \sum_{j=1}^{s} \delta_j^{(1)} u_j \right|^2 = \sum_{j=1}^{s} |u_j|^2 \leq s.$$  

- Using Markov’s inequality, we find $M^1 \subseteq \{-1, 1\}^s$ with cardinality $|M^1| \geq 2^{s-1}$, such that for every $(\delta_j^{(1)}) \in M^1$,

  $$\left| \sum_{j=1}^{s} \delta_j^{(1)} u_j \right|^2 \leq 2s.$$  

- Using the geometric Sauer-Shelah lemma we find $\sigma_1 \subseteq S$, with cardinality $|\sigma_1| \geq \frac{s}{2}$, such that $P_{\sigma_1}(M^1) = \{-1, 1\}^{\sigma_1}$. Since $M^1 \subseteq \mathcal{E} \cap Q$ and the last set is convex, we have $Q_{\sigma_1} \subseteq P_{\sigma_1}(\mathcal{E} \cap Q)$.  

Let \( u_1, \ldots, u_s \in B^n_2 \) and \( \mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \right\} \). Then, for every \( \epsilon \in (0, 1) \) there exists \( \sigma \subseteq \{1, \ldots, s\} \) with cardinality \( |\sigma| \geq (1 - \epsilon)s \), such that \( P_{\sigma}(\mathcal{E}) \supseteq c\sqrt{\epsilon} [-1, 1]^\sigma \), where \( c > 0 \) is an absolute constant, and \( P_{\sigma} \) is the orthogonal projection onto \( \mathbb{R}^\sigma \).
Let $u_1, \ldots, u_s \in B_2^n$ and $\mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \right\}$. Then, for every $\epsilon \in (0, 1)$ there exists $\sigma \subseteq \{1, \ldots, s\}$ with cardinality $|\sigma| \geq (1 - \epsilon)s$, such that $P_\sigma(\mathcal{E}) \supseteq c \sqrt{\epsilon} [1, 1]^{\sigma}$, where $c > 0$ is an absolute constant, and $P_\sigma$ is the orthogonal projection onto $\mathbb{R}^\sigma$.

We set $S = \{1, \ldots, s\}$, $Q = [-1, 1]^s$, $Q_\tau = [-1, 1]^\tau$ for every $\tau \subseteq S$, and for every $k \geq 1$ we define $\alpha_k = \sum_{r=0}^{k-1} 2^{r/2}$ and $\beta_k = \sum_{r=0}^{k-1} 2^r = 2^k - 1$. 
Let \( u_1, \ldots, u_s \in B_2^n \) and \( \mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \right\} \). Then, for every \( \epsilon \in (0, 1) \) there exists \( \sigma \subseteq \{1, \ldots, s\} \) with cardinality \( |\sigma| \geq (1 - \epsilon)s \), such that \( P_\sigma(\mathcal{E}) \supseteq c\sqrt{\epsilon} [-1, 1]^\sigma \), where \( c > 0 \) is an absolute constant, and \( P_\sigma \) is the orthogonal projection onto \( \mathbb{R}^\sigma \).

We set \( S = \{1, \ldots, s\} \), \( Q = [-1, 1]^s \), \( Q_\tau = [-1, 1]^\tau \) for every \( \tau \subseteq S \), and for every \( k \geq 1 \) we define \( \alpha_k = \sum_{r=0}^{k-1} 2^{r/2} \) and \( \beta_k = \sum_{r=0}^{k-1} 2^r = 2^k - 1 \).

Claim (proved by induction on \( k \))

For every \( k \geq 1 \) there exists \( \sigma_k \subseteq S \) with cardinality \( |\sigma_k| \geq (1 - \frac{1}{2^k})s \), such that

\[
Q_{\sigma_k} \subseteq P_{\sigma_k}(\alpha_k \mathcal{E} \cap \beta_k Q).
\]
Isomorphic Sauer-Shelah lemma

Let \( u_1, \ldots, u_s \in B_2^n \) and \( \mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \right\} \). Then, for every \( \epsilon \in (0, 1) \) there exists \( \sigma \subseteq \{1, \ldots, s\} \) with cardinality \( |\sigma| \geq (1 - \epsilon)s \), such that \( P_\sigma(\mathcal{E}) \supseteq c \sqrt{\epsilon} \left[ -1, 1 \right]^{\sigma} \), where \( c > 0 \) is an absolute constant, and \( P_\sigma \) is the orthogonal projection onto \( \mathbb{R}^\sigma \).

We set \( S = \{1, \ldots, s\} \), \( Q = [-1, 1]^s \), \( Q_\tau = [-1, 1]^\tau \) for every \( \tau \subseteq S \), and for every \( k \geq 1 \) we define \( \alpha_k = \sum_{r=0}^{k-1} 2^{r/2} \) and \( \beta_k = \sum_{r=0}^{k-1} 2^r = 2^k - 1 \).

Claim (proved by induction on \( k \))

For every \( k \geq 1 \) there exists \( \sigma_k \subseteq S \) with cardinality \( |\sigma_k| \geq (1 - \frac{1}{2^k})s \), such that

\[
Q_{\sigma_k} \subseteq P_{\sigma_k}(\alpha_k \mathcal{E} \cap \beta_k Q).
\]

The claim shows that for every \( k = 1, 2, \ldots \), there exists \( \sigma_k \subseteq S \) with \( |\sigma_k| \geq (1 - \frac{1}{2^k})s \), such that

\[
P_{\sigma_k}(\mathcal{E}) \supseteq c \sqrt{\frac{1}{2^k}} \left[ -1, 1 \right]^{\sigma_k},
\]

where \( c = \sqrt{2} - 1 \). Then, we easily arrive at the statement of the isomorphic Sauer-Shelah lemma with a slightly worse value for the constant \( c \).
The inductive step

- Consider all points of the form $\delta_{j}^{(k+1)}$, $j \leq s$, where $\delta_{j}^{(k+1)} = 0$ if $j \in \sigma_k$ and $\delta_{j}^{(k+1)} = \pm 2^{k/2}$ if $j \notin \sigma_k$. 
The inductive step

- Consider all points of the form $\delta_j^{(k+1)}$, $j \leq s$, where $\delta_j^{(k+1)} = 0$ if $j \in \sigma_k$ and $\delta_j^{(k+1)} = \pm 2^{k/2}$ if $j \notin \sigma_k$.
- As in the first step,

$$
E \left( \delta_j^{(k+1)} \right)_{j \leq s} \left| \sum_{j=1}^{s} \delta_j^{(k+1)} u_j \right|^2 = \sum_{j \notin \sigma_k} 2^k |u_j|^2 \leq s.
$$

Observe that the cardinality of the set of points $(\delta_j^{(k+1)})_{j \leq s}$ is $2^{s-|\sigma_k|}$. From Markov’s inequality we may find $M^{k+1} \subseteq [0_{\sigma_k} \times \{-2^{k/2}, 2^{k/2}\}^{S \setminus \sigma_k}] \cap \mathcal{E}$ with $|M^{k+1}| \geq 2^{s-|\sigma_k|-1}$. 

Apostolos Giannopoulos  (University of Athens)  
Banach-Mazur distance to the cube  
Castro Urdiales, September 2018
The inductive step

Consider all points of the form $\delta_j^{(k+1)}$, $j \leq s$, where $\delta_j^{(k+1)} = 0$ if $j \in \sigma_k$ and $\delta_j^{(k+1)} = \pm 2^{k/2}$ if $j \notin \sigma_k$.

As in the first step,

$$
\mathbb{E}_{(\delta_j^{(k+1)})_{j \leq s}} \left| \sum_{j=1}^{s} \delta_j^{(k+1)} u_j \right|^2 = \sum_{j \notin \sigma_k} 2^k |u_j|^2 \leq s.
$$

Observe that the cardinality of the set of points $(\delta_j^{(k+1)})_{j \leq s}$ is $2^{s-|\sigma_k|}$. From Markov’s inequality we may find $M^{k+1} \subseteq [0_{\sigma_k} \times \{-2^{k/2}, 2^{k/2}\} \setminus \sigma_k] \cap \mathcal{E}$ with $|M^{k+1}| \geq 2^{s-|\sigma_k|-1}$.

By the Sauer-Shelah lemma there exists $\sigma_{k+1}^* \subseteq S \setminus \sigma_k$, with cardinality $|\sigma_{k+1}^*| \geq \frac{1}{2} (s - |\sigma_k|)$, such that

$$
P_{\sigma_k \cup \sigma_{k+1}^*} (M^{k+1}) = 0_{\sigma_k} \times \{-2^{k/2}, 2^{k/2}\} \sigma_{k+1}^*.$$

Apostolos Giannopoulos (University of Athens)  
Banach-Mazur distance to the cube  
Castro Urdiales, September 2018  14 / 43
The inductive step

- Consider all points of the form $\delta_j^{(k+1)}$, $j \leq s$, where $\delta_j^{(k+1)} = 0$ if $j \in \sigma_k$ and $\delta_j^{(k+1)} = \pm 2^{k/2}$ if $j \notin \sigma_k$.

- As in the first step,

$$\mathbb{E}_{(\delta_j^{(k+1)})_{j \leq s}} \left| \sum_{j=1}^{s} \delta_j^{(k+1)} u_j \right|^2 = \sum_{j \notin \sigma_k} 2^k |u_j|^2 \leq s.$$ 

Observe that the cardinality of the set of points $(\delta_j^{(k+1)})_{j \leq s}$ is $2^{s-|\sigma_k|}$. From Markov’s inequality we may find $M^{k+1} \subseteq [0_{\sigma_k} \times \{-2^{k/2}, 2^{k/2}\} S \setminus \sigma_k] \cap \mathcal{E}$ with $|M^{k+1}| \geq 2^{s-|\sigma_k|-1}$.

- By the Sauer-Shelah lemma there exists $\sigma_k^{*+1} \subseteq S \setminus \sigma_k$, with cardinality $|\sigma_k^{*+1}| \geq \frac{1}{2} (s - |\sigma_k|)$, such that

$$P_{\sigma_k \cup \sigma_k^{*+1}} (M^{k+1}) = 0_{\sigma_k} \times \{-2^{k/2}, 2^{k/2}\} \sigma_k^{*+1}.$$ 

- Since $M^{k+1} \subseteq \mathcal{E} \cap 2^{k/2} Q$ and the last set is convex, it follows that

$$0_{\sigma_k} \times 2^k Q_{\sigma_k^{*+1}} \subseteq P_{\sigma_k \cup \sigma_k^{*+1}} (2^{k/2} \mathcal{E} \cap 2^k Q).$$
The inductive step

- We know that $Q_{\sigma_k} \subseteq P_{\sigma_k}(\alpha_k E \cap \beta_k Q)$ and

$$0_{\sigma_k} \times 2^k Q_{\sigma^*_k+1} \subseteq P_{\sigma_k \cup \sigma^*_k+1}(2^{k/2} E \cap 2^k Q).$$
The inductive step

- We know that $Q_{\sigma_k} \subseteq P_{\sigma_k}(\alpha_k E \cap \beta_k Q)$ and
  
  $0_{\sigma_k} \times 2^k Q_{\sigma^*_k+1} \subseteq P_{\sigma_k \cup \sigma^*_k+1}(2^{k/2} E \cap 2^k Q)$.

- Suppose that $a \in Q_{\sigma_k}$ and $b \in Q_{\sigma^*_k+1}$. By the inductive hypothesis, we can find $w_a \in \beta_k Q_{\sigma^*_k+1}$ for which

  $$(a, w_a) \in P_{\sigma_k \cup \sigma^*_k+1}(\alpha_k E \cap \beta_k Q).$$
The inductive step

- We know that \( Q_{\sigma_k} \subseteq P_{\sigma_k}(\alpha_k \mathcal{E} \cap \beta_k Q) \) and
  \[
  0_{\sigma_k} \times 2^k Q_{\sigma_k^*} \subseteq P_{\sigma_k \cup \sigma_k^*}(2^{k/2} \mathcal{E} \cap 2^k Q).
  \]

- Suppose that \( a \in Q_{\sigma_k} \) and \( b \in Q_{\sigma_k^*} \). By the inductive hypothesis, we can find \( w_a \in \beta_k Q_{\sigma_k^*} \) for which
  \[
  (a, w_a) \in P_{\sigma_k \cup \sigma_k^*}(\alpha_k \mathcal{E} \cap \beta_k Q).
  \]

- We define \( v_{a,b} = b - w_a \). It is clear that \( v_{a,b} \in (\beta_k + 1) Q_{\sigma_k^*} = 2^k Q_{\sigma_k^*} \), and hence
  \[
  (0_{\sigma_k}, v_{a,b}) \in P_{\sigma_k \cup \sigma_k^*}(2^{k/2} \mathcal{E} \cap 2^k Q).
  \]
The inductive step

- We know that $Q_{\sigma_k} \subseteq P_{\sigma_k}(\alpha_k \mathcal{E} \cap \beta_k Q)$ and
  
  $$0_{\sigma_k} \times 2^k Q_{\sigma_{k+1}^*} \subseteq P_{\sigma_k \cup \sigma_{k+1}^*}(2^{k/2} \mathcal{E} \cap 2^k Q).$$

- Suppose that $a \in Q_{\sigma_k}$ and $b \in Q_{\sigma_{k+1}^*}$. By the inductive hypothesis, we can find $w_a \in \beta_k Q_{\sigma_{k+1}^*}$ for which
  
  $$(a, w_a) \in P_{\sigma_k \cup \sigma_{k+1}^*}(\alpha_k \mathcal{E} \cap \beta_k Q).$$

- We define $v_{a,b} = b - w_a$. It is clear that $v_{a,b} \in (\beta_k + 1) Q_{\sigma_{k+1}^*} = 2^k Q_{\sigma_{k+1}^*}$, and hence
  
  $$(0_{\sigma_k}, v_{a,b}) \in P_{\sigma_k \cup \sigma_{k+1}^*}(2^{k/2} \mathcal{E} \cap 2^k Q).$$

- Consequently,
  
  $$(a, b) = (a, w_a) + (0_{\sigma_k}, v_{a,b}) \in P_{\sigma_k \cup \sigma_{k+1}^*}(\alpha_k \mathcal{E} \cap \beta_k Q) + P_{\sigma_k \cup \sigma_{k+1}^*}(2^{k/2} \mathcal{E} \cap 2^k Q)$$
  
  $$\subseteq P_{\sigma_k \cup \sigma_{k+1}^*}(\alpha_{k+1} \mathcal{E} \cap \beta_{k+1} Q).$$
The inductive step

- We know that \( Q_{\sigma_k} \subseteq P_{\sigma_k} (\alpha_k E \cap \beta_k Q) \) and
  \[
  0_{\sigma_k} \times 2^k Q_{\sigma_k^*} \subseteq P_{\sigma_k \cup \sigma_k^*} (2^{k/2} E \cap 2^k Q).
  \]

- Suppose that \( a \in Q_{\sigma_k} \) and \( b \in Q_{\sigma_k^*} \). By the inductive hypothesis, we can find \( w_a \in \beta_k Q_{\sigma_k^*} \) for which
  \[
  (a, w_a) \in P_{\sigma_k \cup \sigma_k^*} (\alpha_k E \cap \beta_k Q).
  \]

- We define \( v_{a,b} = b - w_a \). It is clear that \( v_{a,b} \in (\beta_k + 1) Q_{\sigma_k^*} = 2^k Q_{\sigma_k^*}, \) and hence
  \[
  (0_{\sigma_k}, v_{a,b}) \in P_{\sigma_k \cup \sigma_k^*} (2^{k/2} E \cap 2^k Q).
  \]

- Consequently,
  \[
  (a, b) = (a, w_a) + (0_{\sigma_k}, v_{a,b}) \in P_{\sigma_k \cup \sigma_k^*} (\alpha_k E \cap \beta_k Q) + P_{\sigma_k \cup \sigma_k^*} (2^{k/2} E \cap 2^k Q)
  \]
  \[
  \subseteq P_{\sigma_k \cup \sigma_k^*} (\alpha_{k+1} E \cap \beta_{k+1} Q).
  \]

- We have thus proved that
  \[
  Q_{\sigma_k \cup \sigma_k^*} \subseteq P_{\sigma_k \cup \sigma_k^*} (\alpha_{k+1} E \cap \beta_{k+1} Q).
  \]

We set \( \sigma_{k+1} = \sigma_k \cup \sigma_{k+1}^* \) and observe that \( |\sigma_{k+1}| \geq (1 - \frac{1}{2^{k+1}}) s \).
The main proposition

Let $X = (\mathbb{R}^n, \| \cdot \|)$ be a normed space and let $\epsilon \in (0, 1)$. Assume that the unit ball $K$ of $X$ is in Löwner position. Then, we can find $m \geq (1 - \epsilon)n$ and vectors $z_1, \ldots, z_m$ in $X$ with $\|z_i\| = |z_i| = 1$ so that, for any choice of real numbers $t_1, \ldots, t_m$,

$$\left| \sum_{i=1}^{m} t_i z_i \right| \geq c \frac{\epsilon}{\sqrt{n}} \sum_{i=1}^{m} |t_i|,$$

where $c > 0$ is an absolute constant.
The main proposition

Let \( X = (\mathbb{R}^n, \| \cdot \|) \) be a normed space and let \( \epsilon \in (0, 1) \). Assume that the unit ball \( K \) of \( X \) is in L"owner position. Then, we can find \( m \geq (1 - \epsilon)n \) and vectors \( z_1, \ldots, z_m \) in \( X \) with \( \| z_i \| = |z_i| = 1 \) so that, for any choice of real numbers \( t_1, \ldots, t_m \),

\[
\left| \sum_{i=1}^{m} t_i z_i \right| \geq c \frac{\epsilon}{\sqrt{n}} \sum_{i=1}^{m} |t_i|,
\]

where \( c > 0 \) is an absolute constant.

Proof:

- We use the lemma of Szarek and Talagrand to choose \( x_1, \ldots, x_s \in K \) with \( s \geq (1 - \frac{\epsilon}{2})n \), such that \( \text{dist} \left( x_i, \text{span}\{x_j, j \neq i\} \right) \geq \sqrt{\epsilon}/2 \) for all \( i = 1, \ldots, s \).
The main proposition

Let $X = (\mathbb{R}^n, \| \cdot \|)$ be a normed space and let $\epsilon \in (0, 1)$. Assume that the unit ball $K$ of $X$ is in Löwner position. Then, we can find $m \geq (1 - \epsilon)n$ and vectors $z_1, \ldots, z_m$ in $X$ with $\|z_i\| = |z_i| = 1$ so that, for any choice of real numbers $t_1, \ldots, t_m$,

$$\left| \sum_{i=1}^{m} t_i z_i \right| \geq c \frac{\epsilon}{\sqrt{n}} \sum_{i=1}^{m} |t_i|,$$

where $c > 0$ is an absolute constant.

Proof:

- We use the lemma of Szarek and Talagrand to choose $x_1, \ldots, x_s \in K$ with $s \geq (1 - \frac{\epsilon}{2})n$, such that $\text{dist} \left( x_i, \text{span}\{x_j, j \neq i\} \right) \geq \sqrt{\epsilon}/2$ for all $i = 1, \ldots, s$.

- There exist $v_i \perp \text{span}\{x_j, j \neq i\}$ which form a biorthogonal system with the $x_j$’s and have length $|v_i| \leq \sqrt{2/\epsilon}$. In other words, we can find $v_1, \ldots, v_s \in \mathbb{R}^n$ such that

$$|v_i| \leq \sqrt{2/\epsilon} \quad \text{and} \quad \langle x_i, v_j \rangle = \delta_{ij} \quad i, j = 1, \ldots, s.$$
Upper bound for $\mathcal{R}_\infty^n$

Proof (continued):
- We define $u_i = \sqrt{\epsilon/2} v_i$, and applying the isomorphic Sauer-Shelah lemma for the set $E = \{(\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \}$ we find $\sigma \subseteq \{1, \ldots, s\}$ of cardinality $|\sigma| \geq (1 - \epsilon/2)s$, with

$$P_{\sigma}(E) \supseteq c\sqrt{\epsilon} [-1, 1]^\sigma.$$ 

Then, $|\sigma| \geq (1 - \epsilon)n$. 

Proof (continued):
- We define \( u_i = \sqrt{\epsilon/2} v_i \), and applying the isomorphic Sauer-Shelah lemma for the set \( \mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \right\} \) we find \( \sigma \subseteq \{1, \ldots, s\} \) of cardinality \( |\sigma| \geq (1 - \epsilon/2)s \), with
  \[
P_\sigma(\mathcal{E}) \supseteq c\sqrt{\epsilon} [-1, 1]^\sigma.
\]
  Then, \( |\sigma| \geq (1 - \epsilon)n \).
- Note that for all \((t_i)_{i \in \sigma}\) we have
  \[
  \sum_{i \in \sigma} |t_i| = \left\langle \sum_{i \in \sigma} t_i x_i, \sum_{j \in \sigma} \text{sign}(t_j)v_j \right\rangle.
  \]
Upper bound for $\mathcal{R}_\infty^n$

Proof (continued):

- We define $u_i = \sqrt{\frac{\epsilon}{2}} v_i$, and applying the isomorphic Sauer-Shelah lemma for the set $\mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right|^2 \leq 2s \right\}$ we find $\sigma \subseteq \{1, \ldots, s\}$ of cardinality $|\sigma| \geq (1 - \frac{\epsilon}{2})s$, with $P_\sigma(\mathcal{E}) \supseteq c\sqrt{\epsilon} [-1, 1]^\sigma$.

Then, $|\sigma| \geq (1 - \epsilon)n$.

- Note that for all $(t_i)_{i \in \sigma}$ we have

$$\sum_{i \in \sigma} |t_i| = \left\langle \sum_{i \in \sigma} t_i x_i, \sum_{j \in \sigma} \text{sign}(t_j) v_j \right\rangle.$$ 

- Since $(c\sqrt{\epsilon} \text{sign}(t_j))_{j \in \sigma} \in P_\sigma(\mathcal{E})$, we can find a point $(\delta_j)_{j \leq s}$ in $\mathcal{E}$, such that $\delta_j = c\sqrt{\epsilon} \text{sign}(t_j)$ if $j \in \sigma$. Note that if $i \in \sigma$ and $j \notin \sigma$ then $\langle x_i, v_j \rangle = 0$, and hence

$$\left\langle \sum_{i \in \sigma} t_i x_i, \sum_{j \in \sigma} \text{sign}(t_j) v_j \right\rangle = \frac{1}{c\sqrt{\epsilon}} \left\langle \sum_{i \in \sigma} t_i x_i, \sum_{j=1}^s \delta_j v_j \right\rangle \leq \frac{1}{c\sqrt{\epsilon}} \left| \sum_{i \in \sigma} t_i x_i \right| \sqrt{\frac{2}{\epsilon}} \left| \sum_{j=1}^s \delta_j u_j \right| \leq \frac{2 \sqrt{s}}{c \epsilon} \left| \sum_{i \in \sigma} t_i x_i \right| \leq \frac{\sqrt{n}}{c_1 \epsilon} \left| \sum_{i \in \sigma} t_i x_i \right|.$$
Upper bound for $R^n_{\infty}$

Proof (continued):

- We define $u_i = \sqrt{\epsilon/2} \nu_i$, and applying the isomorphic Sauer-Shelah lemma for the set $E = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^{s} \delta_j u_j \right|^2 \leq 2s \right\}$ we find $\sigma \subseteq \{1, \ldots, s\}$ of cardinality $|\sigma| \geq (1 - \frac{\epsilon}{2})s$, with
  
  $$P_\sigma(E) \supseteq c\sqrt{\epsilon} [-1, 1]^\sigma.$$

  Then, $|\sigma| \geq (1 - \epsilon)n$.

- Note that for all $(t_i)_{i \in \sigma}$ we have
  
  $$\sum_{i \in \sigma} |t_i| = \left\langle \sum_{i \in \sigma} t_i x_i, \sum_{j \in \sigma} \text{sign}(t_j) v_j \right\rangle.$$

- Since $(c\sqrt{\epsilon} \text{sign}(t_j))_{j \in \sigma} \in P_\sigma(E)$, we can find a point $(\delta_j)_{j \leq s}$ in $E$, such that $\delta_j = c\sqrt{\epsilon} \text{sign}(t_j)$ if $j \in \sigma$. Note that if $i \in \sigma$ and $j \notin \sigma$ then $\langle x_i, v_j \rangle = 0$, and hence
  
  $$\left\langle \sum_{i \in \sigma} t_i x_i, \sum_{j \in \sigma} \text{sign}(t_j) v_j \right\rangle \leq \frac{1}{c\sqrt{\epsilon}} \left\langle \sum_{i \in \sigma} t_i x_i, \sum_{j=1}^{s} \delta_j v_j \right\rangle \leq \frac{1}{c\sqrt{\epsilon}} \left| \sum_{i \in \sigma} t_i x_i \right| \sqrt{\frac{2}{\epsilon}} \left| \sum_{j=1}^{s} \delta_j u_j \right| \leq \frac{2\sqrt{s}}{c\epsilon} \left| \sum_{i \in \sigma} t_i x_i \right| \leq \frac{\sqrt{n}}{c_1 \epsilon} \left| \sum_{i \in \sigma} t_i x_i \right|.$$

- We choose as $z_i, i = 1, \ldots, |\sigma| = m$, the $x_j$’s for which $j \in \sigma$, and the proof is complete.
Proof of $R^n_\infty \leq cn^{5/6}$

Let $X = (\mathbb{R}^n, \| \cdot \|)$ be an $n$-dimensional normed space.
Proof of $R^n_\infty \leq cn^{5/6}$

- Let $X = (\mathbb{R}^n, \| \cdot \|)$ be an $n$-dimensional normed space.
- We may assume that the unit ball $K$ of $X$ is in Löwner position. Fix $\epsilon \in (0, 1)$. We have found $m \geq (1 - \epsilon)n$ and $z_1, \ldots, z_m$ in $X$ with $\|z_i\| = |z_i| = 1$ so that, for any choice of real numbers $t_1, \ldots, t_m$,

$$\left| \sum_{i=1}^{m} t_i z_i \right| \geq c \frac{\epsilon}{\sqrt{n}} \sum_{i=1}^{m} |t_i|.$$
Proof of $\mathcal{R}_\infty \leq cn^{5/6}$

- Let $X = (\mathbb{R}^n, \| \cdot \|)$ be an $n$-dimensional normed space.
- We may assume that the unit ball $K$ of $X$ is in Löwner position. Fix $\epsilon \in (0, 1)$. We have found $m \geq (1 - \epsilon)n$ and $z_1, \ldots, z_m$ in $X$ with $\|z_i\| = |z_i| = 1$ so that, for any choice of real numbers $t_1, \ldots, t_m$,
  \[ \left| \sum_{i=1}^{m} t_i z_i \right| \geq c \frac{\epsilon}{\sqrt{n}} \sum_{i=1}^{m} |t_i|. \]
- We define $F = \text{span}\{z_1, \ldots, z_m\}$ and choose any orthonormal basis $y_1, \ldots, y_{n-m}$ of $F^\perp$. By John’s theorem, for every $j = 1, \ldots, n - m$ we have
  \[ |y_j| \leq \|y_j\| \leq \sqrt{n}|y_j| = \sqrt{n}. \]
Proof of $R_n \leq cn^{5/6}$

- Let $X = (\mathbb{R}^n, \| \cdot \|)$ be an $n$-dimensional normed space.
- We may assume that the unit ball $K$ of $X$ is in Löwner position. Fix $\epsilon \in (0, 1)$. We have found $m \geq (1 - \epsilon)n$ and $z_1, \ldots, z_m$ in $X$ with $\|z_i\| = |z_i| = 1$ so that, for any choice of real numbers $t_1, \ldots, t_m$,

$$\left| \sum_{i=1}^{m} t_i z_i \right| \geq c \frac{\epsilon}{\sqrt{n}} \sum_{i=1}^{m} |t_i|.$$ 

- We define $F = \text{span}\{z_1, \ldots, z_m\}$ and choose any orthonormal basis $y_1, \ldots, y_{n-m}$ of $F^\perp$. By John's theorem, for every $j = 1, \ldots, n - m$ we have

$$|y_j| \leq \|y_j\| \leq \sqrt{n}|y_j| = \sqrt{n}.$$ 

- Therefore, if we set $w_j = y_j/\|y_j\|$ we have $\|w_j\| = 1$ and $|w_j| \geq 1/\sqrt{n}$, $j = 1, \ldots, n - m$. 


Proof of $\mathcal{R}_\infty \leq cn^{5/6}$

- Let $X = (\mathbb{R}^n, \| \cdot \|)$ be an $n$-dimensional normed space.
- We may assume that the unit ball $K$ of $X$ is in Löwner position. Fix $\epsilon \in (0, 1)$. We have found $m \geq (1 - \epsilon)n$ and $z_1, \ldots, z_m$ in $X$ with $\|z_i\| = |z_i| = 1$ so that, for any choice of real numbers $t_1, \ldots, t_m$,
  $$\left| \sum_{i=1}^{m} t_i z_i \right| \geq c \epsilon \sqrt{n} \sum_{i=1}^{m} |t_i|.$$  
- We define $F = \text{span}\{z_1, \ldots, z_m\}$ and choose any orthonormal basis $y_1, \ldots, y_{n-m}$ of $F^\perp$. By John's theorem, for every $j = 1, \ldots, n - m$ we have
  $$|y_j| \leq \|y_j\| \leq \sqrt{n} |y_j| = \sqrt{n}.$$  
- Therefore, if we set $w_j = y_j/\|y_j\|$ we have $\|w_j\| = 1$ and $|w_j| \geq 1/\sqrt{n}$, $j = 1, \ldots, n - m$.
- Consider the $n$-tuple of vectors $z_1, \ldots, z_m, w_1, \ldots, w_{n-m}$. Note that $n - m \leq \epsilon n$. 
Proof of $\mathcal{R}_\infty^n \leq cn^{5/6}$

Let $t_1, \ldots, t_m, s_1, \ldots, s_{n-m} \in \mathbb{R}$. Then,

$$\left| \sum_{i=1}^{m} t_i z_i + \sum_{j=1}^{n-m} s_j w_j \right| \leq \left\| \sum_{i=1}^{m} t_i z_i + \sum_{j=1}^{n-m} s_j w_j \right\| \leq \sum_{i=1}^{m} |t_i| + \sum_{j=1}^{n-m} |s_j|.$$
Proof of $R^n \leq cn^{5/6}$

- Let $t_1, \ldots, t_m, s_1, \ldots, s_{n-m} \in \mathbb{R}$. Then,

$$\left| \sum_{i=1}^{m} t_i z_i + \sum_{j=1}^{n-m} s_j w_j \right| \leq \left\| \sum_{i=1}^{m} t_i z_i + \sum_{j=1}^{n-m} s_j w_j \right\| \leq \sum_{i=1}^{m} |t_i| + \sum_{j=1}^{n-m} |s_j|.$$

- On the other hand, $\sum_i t_i z_i$ is orthogonal to $\sum_j s_j w_j$. It follows that

$$\left| \sum_{i=1}^{m} t_i z_i + \sum_{j=1}^{n-m} s_j w_j \right| = \left( \left| \sum_{i=1}^{m} t_i z_i \right|^2 + \left| \sum_{j=1}^{n-m} s_j w_j \right|^2 \right)^{1/2} \geq \frac{1}{\sqrt{2}} \left( \left| \sum_{i=1}^{m} t_i z_i \right| + \left| \sum_{j=1}^{n-m} s_j w_j \right| \right)$$

$$= \frac{1}{\sqrt{2}} \left( \left| \sum_{i=1}^{m} t_i z_i \right| + \left( \sum_{j=1}^{n-m} s_j^2 |w_j|^2 \right)^{1/2} \right) \geq \frac{1}{\sqrt{2}} \left( \frac{c \epsilon}{\sqrt{n}} \sum_{i=1}^{m} |t_i| + \frac{1}{\sqrt{n}} \frac{1}{\sqrt{n-m}} \sum_{j=1}^{n-m} |s_j| \right)$$

$$\geq \frac{1}{\sqrt{2}} \min \left\{ \frac{c \epsilon}{\sqrt{n}}, \frac{1}{\sqrt{\epsilon n}} \right\} \left( \sum_{i=1}^{m} |t_i| + \sum_{j=1}^{n-m} |s_j| \right).$$
Proof of $\mathcal{R}_\infty^n \leq cn^{5/6}$

- Let $t_1, \ldots, t_m, s_1, \ldots, s_{n-m} \in \mathbb{R}$. Then,
  \[
  \left| \sum_{i=1}^m t_i z_i + \sum_{j=1}^{n-m} s_j w_j \right| \leq \left\| \sum_{i=1}^m t_i z_i + \sum_{j=1}^{n-m} s_j w_j \right\| \leq \sum_{i=1}^m |t_i| + \sum_{j=1}^{n-m} |s_j|.
  \]

- On the other hand, $\sum_i t_i z_i$ is orthogonal to $\sum_j s_j w_j$. It follows that
  \[
  \left| \sum_{i=1}^m t_i z_i + \sum_{j=1}^{n-m} s_j w_j \right| = \left( \left| \sum_{i=1}^m t_i z_i \right| \right)^2 + \left( \left| \sum_{j=1}^{n-m} s_j w_j \right| \right)^2 \geq \frac{1}{\sqrt{2}} \left( \left| \sum_{i=1}^m t_i z_i \right| + \left| \sum_{j=1}^{n-m} s_j w_j \right| \right)^2 \\
  = \frac{1}{\sqrt{2}} \left( \left| \sum_{i=1}^m t_i z_i \right| + \left( \sum_{j=1}^{n-m} s_j^2 \right)^{1/2} \right)^2 \geq \frac{1}{\sqrt{2}} \left( \frac{c\epsilon}{\sqrt{n}} \sum_{i=1}^m |t_i| + \frac{1}{\sqrt{n}} \frac{1}{\sqrt{n-m}} \sum_{j=1}^{n-m} |s_j| \right) \\
  \geq \frac{1}{\sqrt{2}} \min \left\{ \frac{c\epsilon}{\sqrt{n}}, \frac{1}{\sqrt{\epsilon n}} \right\} \left( \sum_{i=1}^m |t_i| + \sum_{j=1}^{n-m} |s_j| \right).
  \]

- We have thus proved that
  \[
  d(X, \ell_1^n) \leq \sqrt{2} \max \left\{ \sqrt{n}/c\epsilon, \sqrt{\epsilon n} \right\}
  \]
  for every $\epsilon \in (0, 1)$. The optimal choice of $\epsilon$ is $\epsilon \approx 1/n^{1/3}$. For a value of $\epsilon$ of this order we have $d(X, \ell_1^n) \leq cn^{5/6}$.
In their study of the radius $\mathcal{R}_\infty^n$, Bourgain and Szarek obtained a proportional Dvoretzky-Rogers factorization theorem.

**Bourgain-Szarek**

Assume that $B^n_2$ is the minimal volume ellipsoid of $K$, For every $\epsilon \in (0, 1)$ one can find $m \geq (1 - \epsilon)n$ and $x_1, \ldots, x_m$ among the contact points of $K$ and $B^n_2$, so that for every choice of scalars $(t_i)_{i \leq m}$

$$f(\epsilon) \left( \sum_{i=1}^{m} t_i^2 \right)^{1/2} \leq \left| \sum_{i=1}^{m} t_i x_i \right| \leq \left\| \sum_{i=1}^{m} t_i x_i \right\|_K \leq \sum_{i=1}^{m} |t_i|.$$
Proportional Dvoretzky-Rogers factorization

In their study of the radius $R_{\infty}^n$, Bourgain and Szarek obtained a proportional Dvoretzky-Rogers factorization theorem.

Bourgain-Szarek

Assume that $B_2^n$ is the minimal volume ellipsoid of $K$, For every $\epsilon \in (0, 1)$ one can find $m \geq (1 - \epsilon)n$ and $x_1, \ldots, x_m$ among the contact points of $K$ and $B_2^n$, so that for every choice of scalars $(t_i)_{i \leq m}$

$$f(\epsilon)\left(\sum_{i=1}^{m} t_i^2\right)^{1/2} \leq \left|\sum_{i=1}^{m} t_i x_i\right| \leq \left\|\sum_{i=1}^{m} t_i x_i\right\|_K \leq \sum_{i=1}^{m} |t_i|.$$  

- The important part in this string of inequalities is the first one; it provides a much-stronger version of the classical Dvoretzky–Rogers Lemma which implied a similar inequality only for $m \leq \sqrt{n}$. 

Apostolos Giannopoulos (University of Athens)  
Banach-Mazur distance to the cube  
Castro Urdiales, September 2018  20 / 43
Proportional Dvoretzky-Rogers factorization

In their study of the radius $\mathcal{R}_\infty^n$, Bourgain and Szarek obtained a proportional Dvoretzky-Rogers factorization theorem.

**Bourgain-Szarek**

Assume that $B_2^n$ is the minimal volume ellipsoid of $K$, For every $\epsilon \in (0, 1)$ one can find $m \geq (1 - \epsilon)n$ and $x_1, \ldots, x_m$ among the contact points of $K$ and $B_2^n$, so that for every choice of scalars $(t_i)_{i \leq m}$

$$f(\epsilon) \left( \sum_{i=1}^{m} t_i^2 \right)^{1/2} \leq \left\| \sum_{i=1}^{m} t_i x_i \right\| \leq \left\| \sum_{i=1}^{m} t_i x_i \right\|_K \leq \sum_{i=1}^{m} |t_i|.$$

- The important part in this string of inequalities is the first one; it provides a much-stronger version of the classical Dvoretzky–Rogers Lemma which implied a similar inequality only for $m \leq \sqrt{n}$.
- Equivalently, it can be stated in the form of a “proportional factorization result”:

**Proportional Dvoretzky-Rogers factorization**

Let $X$ be an $n$-dimensional normed space. For any $\epsilon > 0$ there exists $k \geq (1 - \epsilon)^2 n$ such that the identity operator $i_{2,\infty} : l_2^k \to l_\infty^k$ can be written in the form $i_{2,\infty} = \alpha \circ \beta$, where $\beta : l_2^k \to X$, $\alpha : X \to l_\infty^k$ and $\|\alpha\| \cdot \|\beta\| \leq \frac{1}{\epsilon}$.
The first proof by Bourgain and Szarek gave a weaker dependence on $\epsilon$. The work of Szarek and Talagrand improved the dependence on $\epsilon$ to $\epsilon^{-2}$.
Proportional Dvoretzky-Rogers factorization

- The first proof by Bourgain and Szarek gave a weaker dependence on \( \epsilon \). The work of Szarek and Talagrand improved the dependence on \( \epsilon \) to \( \epsilon^{-2} \).
- The best known dependence on \( \epsilon \) is \( c(\epsilon) = \frac{c}{\epsilon} \). The tools that are used are factorization arguments related to Grothendieck’s inequality and the following stronger version of the isomorphic Sauer-Shelah lemma.

G., 1993

Let \( u_1, \ldots, u_s \in B_2^n \) and define \( \mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^{s} \delta_j u_j \right| \leq 1 \right\} \). For every \( \epsilon \in (0, 1) \) we can find \( \sigma \subseteq \{1, \ldots, s\} \) with \( |\sigma| \geq (1 - \epsilon)s \) such that

\[
P_\sigma(\mathcal{E}) \supseteq c\sqrt{\epsilon}B_\sigma,
\]

where \( B_\sigma \) is the Euclidean unit ball in \( \mathbb{R}^\sigma \) and \( c > 0 \) is an absolute constant.
The first proof by Bourgain and Szarek gave a weaker dependence on $\epsilon$. The work of Szarek and Talagrand improved the dependence on $\epsilon$ to $\epsilon^{-2}$.

The best known dependence on $\epsilon$ is $c(\epsilon) = \frac{c}{\epsilon}$. The tools that are used are factorization arguments related to Grothendieck’s inequality and the following stronger version of the isomorphic Sauer-Shelah lemma.

Let $u_1, \ldots, u_s \in B_2^n$ and define $\mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \sum_{j=1}^s \delta_j u_j \leq 1 \right\}$. For every $\epsilon \in (0, 1)$ we can find $\sigma \subseteq \{1, \ldots, s\}$ with $|\sigma| \geq (1 - \epsilon)s$ such that

$$P_\sigma(\mathcal{E}) \supseteq c\sqrt{\epsilon}B_\sigma,$$

where $B_\sigma$ is the Euclidean unit ball in $\mathbb{R}^\sigma$ and $c > 0$ is an absolute constant.

The $\sqrt{\epsilon}$-dependence on $\epsilon$ in the previous result is best possible.
The first proof by Bourgain and Szarek gave a weaker dependence on $\epsilon$. The work of Szarek and Talagrand improved the dependence on $\epsilon$ to $\epsilon^{-2}$.

The best known dependence on $\epsilon$ is $c(\epsilon) = \frac{c}{\epsilon}$. The tools that are used are factorization arguments related to Grothendieck’s inequality and the following stronger version of the isomorphic Sauer-Shelah lemma.

Let $u_1, \ldots, u_s \in B_2^n$ and define $\mathcal{E} = \left\{ (\delta_j)_{j \leq s} \in \mathbb{R}^s : \left| \sum_{j=1}^s \delta_j u_j \right| \leq 1 \right\}$. For every $\epsilon \in (0, 1)$ we can find $\sigma \subseteq \{1, \ldots, s\}$ with $|\sigma| \geq (1 - \epsilon)s$ such that

$$P_{\sigma}(\mathcal{E}) \supseteq c\sqrt{\epsilon}B_{\sigma},$$

where $B_{\sigma}$ is the Euclidean unit ball in $\mathbb{R}^\sigma$ and $c > 0$ is an absolute constant.

The $\sqrt{\epsilon}$-dependence on $\epsilon$ in the previous result is best possible.

Having the proportional Dvoretzky-Rogers factorization theorem, by an application of the Cauchy-Schwarz inequality we receive the main proposition that we used to prove the estimate $R_{\infty}^n \leq cn^{5/6}$ for the Banach-Mazur distance to the cube.
Asymptotic centers of the Banach-Mazur compactum

- As an application of the proportional Dvoretzky-Rogers factorization theorem, Bourgain and Szarek gave a final answer to the problem of the uniqueness up to constant of the center of the Banach-Mazur compactum.
As an application of the proportional Dvoretzky-Rogers factorization theorem, Bourgain and Szarek gave a final answer to the problem of the uniqueness up to constant of the center of the Banach-Mazur compactum.

**Question**

Does there exist a function $f(\lambda)$, $\lambda \geq 1$, such that for every $X \in B_n$ with $R(X) \leq \lambda \sqrt{n}$ we must have $d(X, \ell_2^n) \leq f(\lambda)$?
Asymptotic centers of the Banach-Mazur compactum

- As an application of the proportional Dvoretzky-Rogers factorization theorem, Bourgain and Szarek gave a final answer to the problem of the uniqueness up to constant of the center of the Banach-Mazur compactum.

**Question**

Does there exist a function \( f(\lambda), \lambda \geq 1 \), such that for every \( X \in B_n \) with \( R(X) \leq \lambda \sqrt{n} \) we must have \( d(X, \ell^n_2) \leq f(\lambda) \)?

In other words, the question is if all the “asymptotic centers” of the Banach-Mazur compactum are close to Euclidean space.
Asymptotic centers of the Banach-Mazur compactum

- As an application of the proportional Dvoretzky-Rogers factorization theorem, Bourgain and Szarek gave a final answer to the problem of the uniqueness up to constant of the center of the Banach-Mazur compactum.

**Question**

Does there exist a function \( f(\lambda), \lambda \geq 1 \), such that for every \( X \in B_n \) with \( R(X) \leq \lambda \sqrt{n} \) we must have \( d(X, \ell^2_n) \leq f(\lambda) \)?

In other words, the question is if all the “asymptotic centers” of the Banach-Mazur compactum are close to Euclidean space.

- The answer is negative:

**Bourgain-Szarek**

Let \( X_0 = \ell^s_2 \oplus \ell^{n-s}_1 \) where \( s = \lfloor n/2 \rfloor \). Then \( R(X_0) \leq c\sqrt{n} \) for some absolute constant but \( d(X_0, \ell^2_n) \geq \sqrt{n/2} \).
Asymptotic centers of the Banach-Mazur compactum

- As an application of the proportional Dvoretzky-Rogers factorization theorem, Bourgain and Szarek gave a final answer to the problem of the uniqueness up to constant of the center of the Banach-Mazur compactum.

**Question**

Does there exist a function $f(\lambda)$, $\lambda \geq 1$, such that for every $X \in B_n$ with $R(X) \leq \lambda \sqrt{n}$ we must have $d(X, \ell_2^n) \leq f(\lambda)$?

In other words, the question is if all the “asymptotic centers” of the Banach-Mazur compactum are close to Euclidean space.

- The answer is negative:

**Bourgain-Szarek**

Let $X_0 = \ell_2^s \oplus \ell_1^{n-s}$ where $s = \lfloor n/2 \rfloor$. Then $R(X_0) \leq c \sqrt{n}$ for some absolute constant but $d(X_0, \ell_2^n) \geq \sqrt{n}/2$.

- The main tool in the proof is the proportional Dvoretzky-Rogers theorem.
An alternative approach

A second proof of the bound $R_n^\infty \leq cn^{5/6}$
An alternative approach

A second proof of the bound $\mathcal{R}_\infty^n \leq cn^{5/6}$

- Next we discuss an alternative proof of the proportional Dvoretzky-Rogers factorization theorem, which is due to P. Youssef.
An alternative approach

A second proof of the bound $R_n^\infty \leq cn^{5/6}$

- Next we discuss an alternative proof of the proportional Dvoretzky-Rogers factorization theorem, which is due to P. Youssef.
- We have seen that this also implies the upper bound $R_n^\infty \leq cn^{5/6}$. 
An alternative approach

A second proof of the bound $R_\infty^n \leq cn^{5/6}$

- Next we discuss an alternative proof of the proportional Dvoretzky-Rogers factorization theorem, which is due to P. Youssef.
- We have seen that this also implies the upper bound $R_\infty^n \leq cn^{5/6}$.
- Youssef exploited the method introduced in previous work of Spielman and Srivastava.
Spectral sparsification

- We start with the work of Batson, Spielman and Srivastava on the question to approximate a graph $G = (V, E, w)$ by a sparse graph $G'$. 

Recall that the Laplacian matrix $L_G$ of a graph $G = (V, E, w)$ is defined by 

$$ \langle L_G x, x \rangle = \sum_{(u, v) \in E} w_{u, v} (x_u - x_v)^2. $$

Here, $V$ is the set of vertices of $G$, $E$ is the set of edges of $G$, and $w_{u, v}$ is the weight of the edge $(u, v) \in E$.

Formally, one says that $G'$ is a $\gamma$-approximation of $G$ (for some $\gamma > 1$) if 

$$ \langle L_G x, x \rangle \leq \langle L_{G'} x, x \rangle \leq \gamma \langle L_G x, x \rangle $$

for all $x \in \mathbb{R}^V$.

Batson, Spielman and Srivastava developed a method which shows that for every $d > 1$, every undirected weighted graph $G = (V, E, w)$ with $n$ vertices and $m$ edges contains a weighted subgraph $G' = (V', F', \tilde{w})$ with $\lceil d(n - 1) \rceil$ edges that satisfies 

$$ \langle L_G x, x \rangle \leq \langle L_{G'} x, x \rangle \leq \gamma_d \langle L_G x, x \rangle $$

for all $x \in \mathbb{R}^n$, where 

$$ \gamma_d = \left( \sqrt{d} + \sqrt{d - 1} \right)^2. $$

The proof also provided a deterministic algorithm for computing the graph $G'$ in time $O(dn^3m)$. 

Apostolos Giannopoulos (University of Athens)
Spectral sparsification

- We start with the work of Batson, Spielman and Srivastava on the question to approximate a graph \( G = (V, E, w) \) by a sparse graph \( G' \).
- Recall that the Laplacian matrix \( L_G \) of a graph \( G = (V, E, w) \) is defined by
  \[
  \langle L_G x, x \rangle = \sum_{(u,v) \in E} w_{u,v} (x_u - x_v)^2.
  \]

Here, \( V \) is the set of vertices of \( G \), \( E \) is the set of edges of \( G \), and \( w_{u,v} \) is the weight of the edge \((u, v) \in E\).
Spectral sparsification

- We start with the work of Batson, Spielman and Srivastava on the question to approximate a graph \( G = (V, E, w) \) by a sparse graph \( G' \).
- Recall that the Laplacian matrix \( L_G \) of a graph \( G = (V, E, w) \) is defined by
  \[
  \langle L_G x, x \rangle = \sum_{(u,v) \in E} w_{u,v} (x_u - x_v)^2.
  \]
  Here, \( V \) is the set of vertices of \( G \), \( E \) is the set of edges of \( G \), and \( w_{u,v} \) is the weight of the edge \((u, v) \in E\).
- Formally, one says that \( G' \) is a \( \gamma \)-approximation of \( G \) (for some \( \gamma > 1 \)) if
  \[
  \langle L_G x, x \rangle \leq \langle L_G' x, x \rangle \leq \gamma \langle L_G x, x \rangle
  \]
  for all \( x \in \mathbb{R}^V \).
Spectral sparsification

- We start with the work of Batson, Spielman and Srivastava on the question to approximate a graph $G = (V, E, w)$ by a sparse graph $G'$.
- Recall that the Laplacian matrix $L_G$ of a graph $G = (V, E, w)$ is defined by
  \[
  \langle L_G x, x \rangle = \sum_{(u,v) \in E} w_{u,v} (x_u - x_v)^2.
  \]
  Here, $V$ is the set of vertices of $G$, $E$ is the set of edges of $G$, and $w_{u,v}$ is the weight of the edge $(u, v) \in E$.
- Formally, one says that $G'$ is a $\gamma$-approximation of $G$ (for some $\gamma > 1$) if
  \[
  \langle L_G x, x \rangle \leq \langle L'_{G} x, x \rangle \leq \gamma \langle L_G x, x \rangle
  \]
  for all $x \in \mathbb{R}^V$.
- Batson, Spielman and Srivastava developed a method which shows that for every $d > 1$, every undirected weighted graph $G = (V, E, w)$ with $n$ vertices and $m$ edges contains a weighted subgraph $G' = (V, F', \tilde{w})$ with $\lceil d(n-1) \rceil$ edges that satisfies
  \[
  \langle L_G x, x \rangle \leq \langle L'_{G} x, x \rangle \leq \gamma_d \langle L_G x, x \rangle
  \]
  for all $x \in \mathbb{R}^n$, where $\gamma_d := \left( \frac{\sqrt{d+1}}{\sqrt{d-1}} \right)^2$. 

Spectral sparsification

- We start with the work of Batson, Spielman and Srivastava on the question to approximate a graph \( G = (V, E, w) \) by a sparse graph \( G' \).
- Recall that the Laplacian matrix \( L_G \) of a graph \( G = (V, E, w) \) is defined by
  \[
  \langle L_G x, x \rangle = \sum_{(u,v) \in E} w_{u,v} (x_u - x_v)^2 .
  \]
  Here, \( V \) is the set of vertices of \( G \), \( E \) is the set of edges of \( G \), and \( w_{u,v} \) is the weight of the edge \((u, v) \in E\).
- Formally, one says that \( G' \) is a \( \gamma \)-approximation of \( G \) (for some \( \gamma > 1 \)) if
  \[
  \langle L_G x, x \rangle \leq \langle L_{G'} x, x \rangle \leq \gamma \langle L_G x, x \rangle
  \]
  for all \( x \in \mathbb{R}^V \).
- Batson, Spielman and Srivastava developed a method which shows that for every \( d > 1 \), every undirected weighted graph \( G = (V, E, w) \) with \( n \) vertices and \( m \) edges contains a weighted subgraph \( G' = (V, F', \tilde{w}) \) with \( \lceil d(n - 1) \rceil \) edges that satisfies
  \[
  \langle L_G x, x \rangle \leq \langle L_{G'} x, x \rangle \leq \gamma_d \langle L_G x, x \rangle
  \]
  for all \( x \in \mathbb{R}^n \), where \( \gamma_d := \left( \frac{\sqrt{d+1}}{\sqrt{d-1}} \right)^2 \).
- The proof also provided a deterministic algorithm for computing the graph \( G' \) in time \( O(dn^3 m) \).
For notational convenience, from now on $v$ denotes a column vector in $\mathbb{R}^n$ (an $n \times 1$ matrix) and $v^T$ denotes a row vector (a $1 \times n$ matrix). We write $I$ for the identity matrix of the appropriate dimension. If $A, B$ are two $n \times n$ matrices then the notation $A \preceq B$ means that the matrix $B - A$ is positive semidefinite, while $A \prec B$ means that $B - A$ is positive definite.
Spectral sparsification

- For notational convenience, from now on \( v \) denotes a column vector in \( \mathbb{R}^n \) (an \( n \times 1 \) matrix) and \( v^T \) denotes a row vector (a \( 1 \times n \) matrix). We write \( I \) for the identity matrix of the appropriate dimension. If \( A, B \) are two \( n \times n \) matrices then the notation \( A \preceq B \) means that the matrix \( B - A \) is positive semidefinite, while \( A \prec B \) means that \( B - A \) is positive definite.

- The main technical result of Batson, Spielman and Srivastava is the following purely linear algebraic theorem.

**Batson-Spielman-Srivastava, \( \sim 2009 \)**

Let \( d > 1 \), \( \gamma_d := \left( \frac{\sqrt{d+1}}{\sqrt{d-1}} \right)^2 \) and \( v_1, \ldots, v_m \in \mathbb{R}^n \) such that

\[
I = \sum_{j=1}^{m} v_j v_j^T.
\]

There exist non-negative reals \( \{s_j\}_{1 \leq j \leq m}, \) with \( |\{j : s_j \neq 0\}| \leq dn \), such that

\[
I \preceq \sum_{j=1}^{m} s_j v_j v_j^T \preceq \gamma_d I.
\]
It was soon understood that the theorem of Batson, Spielman and Srivastava is closely related to John decompositions and should have important applications to convex geometry.
Geometric applications

- It was soon understood that the theorem of Batson, Spielman and Srivastava is closely related to John decompositions and should have important applications to convex geometry.
- A sample of applications (chronologically the first):

**Srivastava, ∼ 2010**

Let $K$ be a symmetric convex body in $\mathbb{R}^n$. For any $0 < \epsilon < 1$ there exists a symmetric convex body $D$ in $\mathbb{R}^n$ such that $D \subseteq K \subseteq (1 + \epsilon)D$ and $D$ has at most $cn/\epsilon^2$ contact points with its John ellipsoid, where $c > 0$ is an absolute constant.
Geometric applications

- It was soon understood that the theorem of Batson, Spielman and Srivastava is closely related to John decompositions and should have important applications to convex geometry.
- A sample of applications (chronologically the first):

**Srivastava, ∼ 2010**

Let $K$ be a symmetric convex body in $\mathbb{R}^n$. For any $0 < \epsilon < 1$ there exists a symmetric convex body $D$ in $\mathbb{R}^n$ such that $D \subseteq K \subseteq (1 + \epsilon)D$ and $D$ has at most $cn/\epsilon^2$ contact points with its John ellipsoid, where $c > 0$ is an absolute constant.

- Using completely different methods, Rudelson had proved that one can do the same with a convex body $D$ whose number of contact points with its John ellipsoid is less than $Cn \log n/\epsilon^2$. 
It was soon understood that the theorem of Batson, Spielman and Srivastava is closely related to John decompositions and should have important applications to convex geometry.

A sample of applications (chronologically the first):

**Srivastava, ~ 2010**

Let \( K \) be a symmetric convex body in \( \mathbb{R}^n \). For any \( 0 < \epsilon < 1 \) there exists a symmetric convex body \( D \) in \( \mathbb{R}^n \) such that \( D \subseteq K \subseteq (1 + \epsilon)D \) and \( D \) has at most \( \frac{cn}{\epsilon^2} \) contact points with its John ellipsoid, where \( c > 0 \) is an absolute constant.

- Using completely different methods, Rudelson had proved that one can do the same with a convex body \( D \) whose number of contact points with its John ellipsoid is less than \( Cn \log n/\epsilon^2 \).
- Srivastava also obtained a non-symmetric analogue of this theorem. Later, it took an optimal form:

**Friedland-Youssef, ~ 2016**

Let \( K \) be a convex body in \( \mathbb{R}^n \). For any \( 0 < \epsilon < 1 \) there exists a convex body \( D \) in \( \mathbb{R}^n \) such that \( d(K, D) \leq 1 + \epsilon \) and \( D \) has at most \( \frac{cn}{\epsilon^2} \) contact points with its John ellipsoid, where \( c > 0 \) is an absolute constant.
Let $d > 1$. If $K$ is a symmetric convex body whose minimal volume ellipsoid is the Euclidean unit ball, then there is a subset $X \subset K \cap S^{n-1}$ of cardinality $\text{card}(X) \leq dn$ such that

$$K \subseteq B_2^n \subseteq \gamma_d \sqrt{n} \text{conv}(X).$$
Let $d > 1$. If $K$ is a symmetric convex body whose minimal volume ellipsoid is the Euclidean unit ball, then there is a subset $X \subset K \cap S^{n-1}$ of cardinality $\text{card}(X) \leq dn$ such that

$$K \subseteq B_n^2 \subseteq \gamma_d \sqrt{n} \text{conv}(X).$$

Barvinok applied this fact to prove that there exist $C, \epsilon_0 > 0$ such that for any $0 < \epsilon < \epsilon_0$ and any symmetric convex body $C$ in $\mathbb{R}^n$, $n \geq 1$, there exists a symmetric polytope $P$ in $\mathbb{R}^d$ with at most $\left(\frac{C}{\sqrt{\epsilon}} \log \frac{1}{\epsilon}\right)^n$ vertices such that $P \subseteq C \subseteq (1 + \epsilon)P$.

One should compare this estimate with the standard bound $(3/\epsilon)^n$ which follows by a simple volumetric argument.
Let $d > 1$. If $K$ is a symmetric convex body whose minimal volume ellipsoid is the Euclidean unit ball, then there is a subset $X \subset K \cap S^{n-1}$ of cardinality $\text{card}(X) \leq dn$ such that

$$K \subseteq B_2^n \subseteq \gamma_d \sqrt{n} \text{conv}(X).$$

- Barvinok applied this fact to prove that there exist $C, \varepsilon_0 > 0$ such that for any $0 < \varepsilon < \varepsilon_0$ and any symmetric convex body $C$ in $\mathbb{R}^n$, $n \geq 1$, there exists a symmetric polytope $P$ in $\mathbb{R}^d$ with at most $\left(\frac{C}{\sqrt{\varepsilon}} \log \frac{1}{\varepsilon}\right)^n$ vertices such that $P \subseteq C \subseteq (1 + \varepsilon)P$. One should compare this estimate with the standard bound $(3/\varepsilon)^n$ which follows by a simple volumetric argument.

- Gluskin and Litvak applied the same fact to obtain the optimal form of an estimate of Bezdek and Litvak for the vertex index of a convex body, defined by

$$\text{vein}(K) = \inf \left\{ \sum_{j=1}^{N} \|y_j\|_K : K \subseteq \text{conv}\{y_1, \ldots, y_N\} \right\}.$$
Let $d > 1$. If $K$ is a symmetric convex body whose minimal volume ellipsoid is the Euclidean unit ball, then there is a subset $X \subseteq K \cap S^{n-1}$ of cardinality $\text{card}(X) \leq dn$ such that

$$K \subseteq B_2^n \subseteq \gamma_d \sqrt{n} \text{conv}(X).$$

- Barvinok applied this fact to prove that there exist $C, \epsilon_0 > 0$ such that for any $0 < \epsilon < \epsilon_0$ and any symmetric convex body $C$ in $\mathbb{R}^n$, $n \geq 1$, there exists a symmetric polytope $P$ in $\mathbb{R}^d$ with at most $\left( \frac{C}{\sqrt{\epsilon}} \log \frac{1}{\epsilon} \right)^n$ vertices such that $P \subseteq C \subseteq (1 + \epsilon)P$. One should compare this estimate with the standard bound $(3/\epsilon)^n$ which follows by a simple volumetric argument.

- Gluskin and Litvak applied the same fact to obtain the optimal form of an estimate of Bezdek and Litvak for the vertex index of a convex body, defined by

$$\text{vein}(K) = \inf \left\{ \sum_{j=1}^{N} \|y_j\|_K : K \subseteq \text{conv}\{y_1, \ldots, y_N\} \right\}.$$

- They proved that if $K$ is a centrally symmetric convex body in $\mathbb{R}^n$ then $\text{vein}(K) \leq 24n^{3/2}$. The example of the Euclidean ball shows that the bound $O(n^{3/2})$ is optimal.
The restricted invertibility principle of Bourgain and Tzafriri states that if $A$ is an $n \times n$ matrix whose columns $Ae_j$ have Euclidean norm equal to 1 then there exists $\sigma \subset [n]$ of cardinality $|\sigma| \geq cn/\|A\|^2_2$ such that the restriction $A_\sigma$ of $A$ to $\text{span}\{e_j : j \in \sigma\}$ is well-invertible.
The restricted invertibility principle of Bourgain and Tzafriri states that if $A$ is an $n \times n$ matrix whose columns $Ae_j$ have Euclidean norm equal to 1 then there exists $\sigma \subset [n]$ of cardinality $|\sigma| \geq cn/\|A\|_2^2$ such that the restriction $A_\sigma$ of $A$ to $\text{span}\{e_j : j \in \sigma\}$ is well-invertible.

Bourgain-Tzafriri, 1987

There exist absolute constants $\delta, \kappa > 0$ such that if $A : \ell^n_2 \rightarrow \ell^n_2$ is a linear operator with $|Ae_j| = 1$ for all $j = 1, \ldots, n$ then one may find a subset $\sigma \subseteq [n]$ of cardinality $|\sigma| \geq \delta n/\|A\|_2^2$ such that

$$\left| \sum_{j \in \sigma} t_j Ae_j \right|^2 \geq \kappa \sum_{j \in \sigma} |t_j|^2$$

(2)

for any choice of scalars $\{t_j\}_{j \in \sigma}$. 
The restricted invertibility principle of Bourgain and Tzafriri states that if $A$ is an $n \times n$ matrix whose columns $Ae_j$ have Euclidean norm equal to 1 then there exists a subset $\sigma \subset [n]$ of cardinality $|\sigma| \geq cn/\|A\|_2^2$ such that the restriction $A_\sigma$ of $A$ to $\text{span}\{e_j : j \in \sigma\}$ is well-invertible.

Bourgain-Tzafriri, 1987

There exist absolute constants $\delta, \kappa > 0$ such that if $A : \ell_2^n \rightarrow \ell_2^n$ is a linear operator with $|Ae_j| = 1$ for all $j = 1, \ldots, n$ then one may find a subset $\sigma \subset [n]$ of cardinality $|\sigma| \geq \delta n/\|A\|_2^2$ such that

$$\left| \sum_{j \in \sigma} t_j A e_j \right|^2 \geq \kappa \sum_{j \in \sigma} |t_j|^2$$

for any choice of scalars $\{t_j\}_{j \in \sigma}$.

If $A_\sigma$ is the restriction of $A$ to $\text{span}\{e_j : j \in \sigma\}$ then (2) is equivalent to the fact that $s_{\text{min}}(A_\sigma) \geq \kappa$, where $s_{\text{min}}(A)$ denotes the smallest singular number of an operator $A$. 
Vershynin generalized the restricted invertibility theorem as follows.

\textbf{Vershynin, \sim 2000}

Let \( I = \sum_{j=1}^{m} v_j v_j^T \) is an arbitrary decomposition of the identity and \( A : \ell_2^n \to \ell_2^n \) be a linear operator. Then, for any \( \epsilon \in (0, 1) \) one can find \( \sigma \subset [m] \) of cardinality \( |\sigma| \geq (1 - \epsilon) \|A\|_{\text{HS}}^2 / \|A\|_2^2 \) such that for any choice of scalars \( (t_j)_{j \in \sigma} \),

\[
\left| \sum_{j \in \sigma} t_j \frac{A v_j}{|A v_j|} \right| \geq c(\epsilon) \left( \sum_{j \in \sigma} t_j^2 \right)^{1/2}, \tag{3}
\]

where \( c(\epsilon) > 0 \) is a constant depending only on \( \epsilon \).
Restricted invertibility principle

Vershynin generalized the restricted invertibility theorem as follows.

**Vershynin, ~ 2000**

Let \( I = \sum_{j=1}^{m} v_j v_j^T \) is an arbitrary decomposition of the identity and \( A : \ell_2^n \rightarrow \ell_2^n \) be a linear operator. Then, for any \( \epsilon \in (0, 1) \) one can find \( \sigma \subset [m] \) of cardinality
\[
|\sigma| \geq (1 - \epsilon) \frac{\|A\|_{HS}^2}{\|A\|_2^2}
\]
such that for any choice of scalars \( (t_j)_{j \in \sigma} \),
\[
\left| \sum_{j \in \sigma} t_j \frac{Av_j}{|Av_j|} \right| \geq c(\epsilon) \left( \sum_{j \in \sigma} t_j^2 \right)^{1/2},
\]
where \( c(\epsilon) > 0 \) is a constant depending only on \( \epsilon \).

- Note that if \( |Ae_j| = 1 \) for all \( j \) then, applying Vershynin’s theorem for the standard decomposition \( I = \sum_{j=1}^{n} e_j e_j^T \) we recover the theorem of Bourgain and Tzafriri.
Vershynin generalized the restricted invertibility theorem as follows.

**Vershynin, ~ 2000**

Let $I = \sum_{j=1}^{m} v_j v_j^T$ is an arbitrary decomposition of the identity and $A : \ell_2^n \to \ell_2^n$ be a linear operator. Then, for any $\epsilon \in (0, 1)$ one can find $\sigma \subset [m]$ of cardinality $|\sigma| \geq (1 - \epsilon) \|A\|_{\text{HS}}^2 / \|A\|^2_2$ such that for any choice of scalars $(t_j)_{j \in \sigma}$,

$$\left| \sum_{j \in \sigma} t_j \frac{Av_j}{|Av_j|} \right| \geq c(\epsilon) \left( \sum_{j \in \sigma} t_j^2 \right)^{1/2},$$

(3)

where $c(\epsilon) > 0$ is a constant depending only on $\epsilon$.

- Note that if $|Ae_j| = 1$ for all $j$ then, applying Vershynin’s theorem for the standard decomposition $I = \sum_{j=1}^{n} e_j e_j^T$ we recover the theorem of Bourgain and Tzafriri.

- Moreover, we may now find $\sigma \subseteq [n]$ of cardinality greater than $(1 - \epsilon)n / \|A\|^2_2$ for any $\epsilon \in (0, 1)$ so that (2) will hold true, of course with a constant $\delta = c(\epsilon)$ depending on $\epsilon$. 
Vershynin generalized the restricted invertibility theorem as follows.

**Vershynin, ∼ 2000**

Let \( I = \sum_{j=1}^{m} v_j v_j^T \) is an arbitrary decomposition of the identity and \( A : \ell_2^n \to \ell_2^n \) be a linear operator. Then, for any \( \epsilon \in (0, 1) \) one can find \( \sigma \subset [m] \) of cardinality \( |\sigma| \geq (1 - \epsilon) \|A\|_{\text{HS}}^2 / \|A\|_2^2 \) such that for any choice of scalars \( (t_j)_{j \in \sigma} \),

\[
\left| \sum_{j \in \sigma} t_j \frac{A v_j}{|A v_j|} \right| \geq c(\epsilon) \left( \sum_{j \in \sigma} t_j^2 \right)^{1/2},
\]

where \( c(\epsilon) > 0 \) is a constant depending only on \( \epsilon \).

- Note that if \( |A e_j| = 1 \) for all \( j \) then, applying Vershynin’s theorem for the standard decomposition \( I = \sum_{j=1}^{n} e_j e_j^T \) we recover the theorem of Bourgain and Tzafriri.

- Moreover, we may now find \( \sigma \subseteq [n] \) of cardinality greater than \( (1 - \epsilon)n / \|A\|_2^2 \) for any \( \epsilon \in (0, 1) \) so that (2) will hold true, of course with a constant \( \delta = c(\epsilon) \) depending on \( \epsilon \).

- Vershynin’s argument is based on an iteration of the Bourgain-Tzafriri theorem and a result of Kashin-Tzafriri, and this affects the final dependence of \( c(\epsilon) \) on \( \epsilon \).
Spielman and Srivastava gave a generalization of the Bourgain-Tzafriri theorem, in the spirit of Vershynin’s theorem, with optimal dependence on $\epsilon$, exploiting the method of their previous work with Batson.
Spielman and Srivastava gave a generalization of the Bourgain-Tzafriri theorem, in the spirit of Vershynin’s theorem, with optimal dependence on $\epsilon$, exploiting the method of their previous work with Batson.

\textbf{Spielman-Srivastava, $\sim$ 2010}

Let $\epsilon \in (0, 1)$ and $v_1, \ldots, v_m \in \mathbb{R}^n$ such that $I = \sum_{j=1}^{m} v_j v_j^T$. Let $A : \ell_n^2 \to \ell_n^2$ be a linear operator. We can find $\sigma \subseteq [m]$ of cardinality $|\sigma| \geq \lfloor (1 - \epsilon)^2 \|A\|_{\text{HS}}^2 / \|A\|_2^2 \rfloor$ such that the set $\{Av_j : j \in \sigma\}$ is linearly independent and

$$\lambda_{\min} \left( \sum_{j \in \sigma} (Av_j)(Av_j)^T \right) \geq \epsilon^2 \frac{\|A\|_{\text{HS}}^2}{m},$$

where the smallest eigenvalue $\lambda_{\min}$ is computed on the subspace $\text{span}\{Av_j : j \in \sigma\}$. 
Spielman and Srivastava gave a generalization of the Bourgain-Tzafriri theorem, in the spirit of Vershynin’s theorem, with optimal dependence on $\epsilon$, exploiting the method of their previous work with Batson.

### Spielman-Srivastava, $\sim$ 2010

Let $\epsilon \in (0, 1)$ and $v_1, \ldots, v_m \in \mathbb{R}^n$ such that $I = \sum_{j=1}^{m} v_j v_j^T$. Let $A : \ell_2^n \to \ell_2^n$ be a linear operator. We can find $\sigma \subseteq [m]$ of cardinality $|\sigma| \geq \left\lfloor (1 - \epsilon)^2 \frac{\|A\|_2^2}{\|A\|_2^2} \right\rfloor$ such that the set $\{Av_j : j \in \sigma\}$ is linearly independent and

$$\lambda_{\min}\left(\sum_{j \in \sigma} (Av_j)(Av_j)^T\right) \geq \epsilon^2 \frac{\|A\|_2^2}{m},$$

where the smallest eigenvalue $\lambda_{\min}$ is computed on the subspace $\text{span}\{Av_j : j \in \sigma\}$.

The statement above is equivalent to the fact that, for any choice of scalars $(t_j)_{j \in \sigma}$,

$$\left| \sum_{j \in \sigma} t_j Av_j \right| \geq \epsilon \frac{\|A\|_{\text{HS}}}{\sqrt{m}} \left( \sum_{j \in \sigma} t_j^2 \right)^{1/2}.$$
Restricted invertibility principle

- Spielman and Srivastava gave a generalization of the Bourgain-Tzafriri theorem, in the spirit of Vershynin’s theorem, with optimal dependence on $\epsilon$, exploiting the method of their previous work with Batson.

**Spielman-Srivastava, \sim 2010**

Let $\epsilon \in (0, 1)$ and $v_1, \ldots, v_m \in \mathbb{R}^n$ such that $I = \sum_{j=1}^{m} v_j v_j^T$. Let $A : \ell_2^n \to \ell_2^n$ be a linear operator. We can find $\sigma \subseteq [m]$ of cardinality $|\sigma| \geq \lceil (1 - \epsilon)^2 \|A\|_{\text{HS}}^2 / \|A\|_2^2 \rceil$ such that the set $\{Av_j : j \in \sigma\}$ is linearly independent and

$$\lambda_{\text{min}} \left( \sum_{j \in \sigma} (Av_j)(Av_j)^T \right) \geq \epsilon^2 \frac{\|A\|_{\text{HS}}^2}{m},$$

where the smallest eigenvalue $\lambda_{\text{min}}$ is computed on the subspace $\text{span}\{Av_j : j \in \sigma\}$.

- The statement above is equivalent to the fact that, for any choice of scalars $(t_j)_{j \in \sigma}$,

$$\left| \sum_{j \in \sigma} t_j Av_j \right| \geq \epsilon \frac{\|A\|_{\text{HS}}}{\sqrt{m}} \left( \sum_{j \in \sigma} t_j^2 \right)^{1/2}. $$

- The Bourgain-Tzafriri theorem follows from this one, with constants $\delta(\epsilon) = (1 - \epsilon)^2$ and $\kappa(\epsilon) = \epsilon^2$. 
Comparing the previous results we see that both generalize the Bourgain-Tzafriri theorem but in a different way.
Comparing the previous results we see that both generalize the Bourgain-Tzafriri theorem but in a different way.

- Vershynin: the vectors that are chosen are normalized but the dependence on $\epsilon$ is weak.
Comparing the previous results we see that both generalize the Bourgain-Tzafriri theorem but in a different way.

- Vershynin: the vectors that are chosen are normalized but the dependence on $\epsilon$ is weak.
- Spielman-Srivastava: optimal dependence on $\epsilon$ but the vectors are not normalized.
Comparing the previous results we see that both generalize the Bourgain-Tzafriri theorem but in a different way.

- Vershynin: the vectors that are chosen are normalized but the dependence on $\epsilon$ is weak.
- Spielman-Srivastava: optimal dependence on $\epsilon$ but the vectors are not normalized.

Youssef obtained a restricted invertibility theorem for any rectangular matrix and any normalization, with a good dependence on $\epsilon$ at the same time.

### Youssef, 2012

Let $A$ be an $n \times m$ matrix and $D = \text{diag}(\alpha_1, \ldots, \alpha_m)$ be a diagonal $m \times m$ matrix such that $\text{Ker}(D) \subset \text{Ker}(A)$. Then, for any $\epsilon \in (0, 1)$ there exists $\sigma \subset \{1, \ldots, m\}$ with $|\sigma| \geq (1 - \epsilon)^2 \|A\|_{\text{HS}}^2 / \|A\|_2^2$ such that

$$s_{\min}(A_{\sigma} D_{\sigma}^{-1}) > \epsilon \|A\|_{\text{HS}} / \|D\|_{\text{HS}},$$

where $s_{\min}$ denotes the smallest singular value.
Proportional Dvoretzky-Rogers factorization

Comparing the previous results we see that both generalize the Bourgain-Tzafriri theorem but in a different way.

- Vershynin: the vectors that are chosen are normalized but the dependence on $\epsilon$ is weak.
- Spielman-Srivastava: optimal dependence on $\epsilon$ but the vectors are not normalized.

Youssef obtained a restricted invertibility theorem for any rectangular matrix and any normalization, with a good dependence on $\epsilon$ at the same time.

Youssef, 2012

Let $A$ be an $n \times m$ matrix and $D = \text{diag}(\alpha_1, \ldots, \alpha_m)$ be a diagonal $m \times m$ matrix such that $\text{Ker}(D) \subseteq \text{Ker}(A)$. Then, for any $\epsilon \in (0, 1)$ there exists $\sigma \subset \{1, \ldots, m\}$ with $|\sigma| \geq (1 - \epsilon)^2 \|A\|_{\text{HS}}^2 / \|A\|_2^2$ such that

$$s_{\min}(A_\sigma D_\sigma^{-1}) > \epsilon \|A\|_{\text{HS}} / \|D\|_{\text{HS}},$$

where $s_{\min}$ denotes the smallest singular value.

Equivalently, for any choice of reals $(t_j)_{j \in \sigma}$ one has

$$\left| \sum_{j \in \sigma} t_j \frac{A e_j}{\alpha_j} \right| \geq \epsilon \frac{\|A\|_{\text{HS}}}{\|D\|_{\text{HS}}} \left( \sum_{j \in \sigma} t_j^2 \right)^{1/2}.$$
Theorem

Assume that $B^n_2$ is the minimal volume ellipsoid of $K$, For every $\epsilon \in (0, 1)$ there exist $k \geq (1 - \epsilon)^2 n$ and $y_1, \ldots, y_k \in B^n_2$ such that, for any choice of scalars $(t_j)_{j \leq k}$,

$$\epsilon \left( \sum_{j=1}^{k} t_j^2 \right)^{1/2} \leq \left\| \sum_{j=1}^{k} t_j y_j \right\| \leq \sum_{j=1}^{k} |t_j|.$$ 

- We start from John’s decomposition $I = \sum_{j \leq m} c_j x_j x_j^T$ where $x_j \in \partial(K) \cap S^{n-1}$. 

Proof of the proportional Dvoretzky-Rogers factorization theorem
Theorem

Assume that $B_2^n$ is the minimal volume ellipsoid of $K$, For every $\epsilon \in (0, 1)$ there exist $k \geq (1 - \epsilon)^2 n$ and $y_1, \ldots, y_k \in B_2^n$ such that, for any choice of scalars $(t_j)_{j \leq k}$,

$$\epsilon \left( \sum_{j=1}^{k} t_j^2 \right)^{1/2} \leq \left\| \sum_{j=1}^{k} t_j y_j \right\| \leq \sum_{j=1}^{k} |t_j|.$$ 

- We start from John’s decomposition $I = \sum_{j \leq m} c_j x_j x_j^T$ where $x_j \in \partial(K) \cap \mathbb{S}^{n-1}$.
- We consider the $n \times m$ matrix $A = (\sqrt{c_1} x_1, \ldots, \sqrt{c_m} x_m)$ with columns $\sqrt{c_j} x_j$ and the diagonal matrix $D = \text{diag}(\sqrt{c_1}, \ldots, \sqrt{c_m})$. Then, $AA^T = I$ and $\|A\|_{\text{HS}} = \|D\|_{\text{HS}} = \sqrt{n}$. 

Apostolos Giannopoulos (University of Athens)
Proof of the proportional Dvoretzky-Rogers factorization theorem

**Theorem**

Assume that $B_2^n$ is the minimal volume ellipsoid of $K$, for every $\epsilon \in (0, 1)$ there exist $k \geq (1 - \epsilon)^2 n$ and $y_1, \ldots, y_k \in B_2^n$ such that, for any choice of scalars $(t_j)_{j \leq k}$,

\[ \epsilon \left( \sum_{j=1}^{k} t_j^2 \right)^{1/2} \leq \left\| \sum_{j=1}^{k} t_j y_j \right\| \leq \sum_{j=1}^{k} |t_j|. \]

- We start from John’s decomposition $I = \sum_{j \leq m} c_j x_j x_j^T$ where $x_j \in \partial(K) \cap S^{n-1}$.
- We consider the $n \times m$ matrix $A = (\sqrt{c_1} x_1, \ldots, \sqrt{c_m} x_m)$ with columns $\sqrt{c_j} x_j$ and the diagonal matrix $D = \text{diag}(\sqrt{c_1}, \ldots, \sqrt{c_m})$. Then, $AA^T = I$ and $\|A\|_{HS} = \|D\|_{HS} = \sqrt{n}$.
- Given $\epsilon \in (0, 1)$ we apply Youssef’s theorem to $A$ and $D$ to find $\sigma \subset \{1, \ldots, m\}$ with $|\sigma| = k \geq (1 - \epsilon)^2 n$ such that, for any choice of scalars $t = (t_j)_{j \in \sigma}$,

\[ |A_{\sigma} D_{\sigma}^{-1} t| = \left| \sum_{j \in \sigma} t_j x_j \right| \geq \epsilon \left( \sum_{j \in \sigma} t_j^2 \right)^{1/2}. \]
Proof of the proportional Dvoretzky-Rogers factorization theorem

Theorem

Assume that $B_2^n$ is the minimal volume ellipsoid of $K$, for every $\epsilon \in (0, 1)$ there exist $k \geq (1 - \epsilon)^2 n$ and $y_1, \ldots, y_k \in B_2^n$ such that, for any choice of scalars $(t_j)_{j \leq k}$,

$$\epsilon \left( \sum_{j=1}^{k} t_j^2 \right)^{1/2} \leq \left\| \sum_{j=1}^{k} t_j y_j \right\| \leq \sum_{j=1}^{k} |t_j|.$$ 

- We start from John’s decomposition $l = \sum_{j \leq m} c_j x_j x_j^T$ where $x_j \in \partial(K) \cap S^{n-1}$.
- We consider the $n \times m$ matrix $A = (\sqrt{c_1 x_1}, \ldots, \sqrt{c_m x_m})$ with columns $\sqrt{c_j} x_j$ and the diagonal matrix $D = \text{diag}(\sqrt{c_1}, \ldots, \sqrt{c_m})$. Then, $AA^T = I$ and $\|A\|_{HS} = \|D\|_{HS} = \sqrt{n}$.
- Given $\epsilon \in (0, 1)$ we apply Youssef’s theorem to $A$ and $D$ to find $\sigma \subset \{1, \ldots, m\}$ with $|\sigma| = k \geq (1 - \epsilon)^2 n$ such that, for any choice of scalars $t = (t_j)_{j \in \sigma}$,

$$|A_{\sigma} D_{\sigma}^{-1} t| = \left| \sum_{j \in \sigma} t_j x_j \right| \geq \epsilon \left( \sum_{j \in \sigma} t_j^2 \right)^{1/2}.$$ 

- Since $K \subseteq B_2^n$ and $\|x_j\| = 1$, we also have

$$\left| \sum_{j \in \sigma} t_j x_j \right| \leq \left\| \sum_{j \in \sigma} t_j x_j \right\| \leq \sum_{j \in \sigma} |t_j| \|x_j\| \leq \sum_{j \in \sigma} |t_j|. $$

Apostolos Giannopoulos (University of Athens) Banach-Mazur distance to the cube Castro Urdiales, September 2018 32 / 43
Idea of the proof

Youssef

Let $A$ be an $n \times m$ matrix and $D = \text{diag}(\alpha_1, \ldots, \alpha_m)$ be a diagonal $m \times m$ matrix such that $\text{Ker}(D) \subset \text{Ker}(A)$. Then, for any $\epsilon \in (0, 1)$ there exists $\sigma \subset \{1, \ldots, m\}$ with $|\sigma| \geq (1 - \epsilon)^2 \|A\|_{\text{HS}}^2 / \|A\|_2^2$ such that

$$s_{\min}(A_{\sigma} D_{\sigma}^{-1}) > \frac{\epsilon \|A\|_{\text{HS}}}{\|D\|_{\text{HS}}}.$$

where $s_{\min}$ denotes the smallest singular value.

- It suffices to find $\sigma \subset \{1, \ldots, m\}$ with $|\sigma| \geq (1 - \epsilon)^2 \|A\|_{\text{HS}}^2 / \|A\|_2^2$ such that

$$(A_{\sigma} D_{\sigma}^{-1}) \cdot (A_{\sigma} D_{\sigma}^{-1})^T = \sum_{j \in \sigma} (A D_{\sigma}^{-1} e_j) \cdot (A D_{\sigma}^{-1} e_j)^T = \sum_{j \in \sigma} \left( \frac{Ae_j}{\alpha_j} \right) \cdot \left( \frac{Ae_j}{\alpha_j} \right)^T$$

has rank equal to $k_0 = |\sigma|$ and its smallest positive eigenvalue is greater than $\epsilon^2 \|A\|_{\text{HS}}^2 / \|D\|_{\text{HS}}^2$. 

Apostolos Giannopoulos (University of Athens) Banach-Mazur distance to the cube Castro Urdiales, September 2018
Let $A$ be an $n \times m$ matrix and $D = \text{diag}(\alpha_1, \ldots, \alpha_m)$ be a diagonal $m \times m$ matrix such that $\text{Ker}(D) \subset \text{Ker}(A)$. Then, for any $\epsilon \in (0, 1)$ there exists $\sigma \subset \{1, \ldots, m\}$ with $|\sigma| \geq (1 - \epsilon)^2 \|A\|_{\text{HS}}^2 / \|A\|_2^2$ such that

$$s_{\min}(A_{\sigma}D^{-1}_{\sigma}) > \frac{\epsilon\|A\|_{\text{HS}}}{\|D\|_{\text{HS}}} ,$$

where $s_{\min}$ denotes the smallest singular value.

- It suffices to find $\sigma \subset \{1, \ldots, m\}$ with $|\sigma| \geq (1 - \epsilon)^2 \|A\|_{\text{HS}}^2 / \|A\|_2^2$ such that

$$(A_{\sigma}D^{-1}_{\sigma}) \cdot (A_{\sigma}D^{-1}_{\sigma})^T = \sum_{j \in \sigma} (AD^{-1}_{\sigma}e_j) \cdot (AD^{-1}_{\sigma}e_j)^T = \sum_{j \in \sigma} \left(\frac{Ae_j}{\alpha_j}\right) \cdot \left(\frac{Ae_j}{\alpha_j}\right)^T$$

has rank equal to $k_0 = |\sigma|$ and its smallest positive eigenvalue is greater than $\epsilon^2 \|A\|_{\text{HS}}^2 / \|D\|_{\text{HS}}^2$.

- The matrix $M_{k_0} = \sum_{j \in \sigma} \left(\frac{Ae_j}{\alpha_j}\right) \cdot \left(\frac{Ae_j}{\alpha_j}\right)^T$ is defined by an inductive scheme. We start with $M_0 = 0$ and at each step we add a rank one matrix $\left(\frac{Ae_j}{\alpha_j}\right) \cdot \left(\frac{Ae_j}{\alpha_j}\right)^T$ for a suitable $j$, which will give a new positive eigenvalue.
Let $A$ be an invertible $n \times n$ matrix. For any $v \in \mathbb{R}^n$ we have

$$(A + vv^T)^{-1} = A^{-1} - \frac{A^{-1}vv^T A^{-1}}{1 + v^T A^{-1}v}.$$
Facts from linear algebra

**Sherman-Morrison formula**

Let $A$ be an invertible $n \times n$ matrix. For any $v \in \mathbb{R}^n$ we have

$$(A + vv^T)^{-1} = A^{-1} - \frac{A^{-1}v v^T A^{-1}}{1 + v^T A^{-1}v}.$$

**Matrix determinant formula**

Let $A$ be an invertible $n \times n$ matrix. For any $v \in \mathbb{R}^n$ we have

$$\det(A + vv^T) = \det(A)(1 + v^T A^{-1}v).$$
Sherman-Morrison formula
Let $A$ be an invertible $n \times n$ matrix. For any $v \in \mathbb{R}^n$ we have

$$(A + vv^T)^{-1} = A^{-1} - \frac{A^{-1}vv^TA^{-1}}{1 + v^TA^{-1}v}.$$ 

Matrix determinant formula
Let $A$ be an invertible $n \times n$ matrix. For any $v \in \mathbb{R}^n$ we have

$$\det(A + vv^T) = \det(A)(1 + v^T A^{-1}v).$$

Cauchy’s interlacing theorem
Let $\chi(A)(x) = \det(xI - A)$ denote the characteristic polynomial of $A$. If $A$ is a symmetric $n \times n$ matrix and $v \in \mathbb{R}^n$ then $\chi(A)$ interlaces $\chi(A + vv^T)$: if $\lambda_i, \lambda'_i$ are their eigenvalues in decreasing order then

$$\lambda'_1 \geq \lambda_1 \geq \lambda'_2 \geq \lambda_2 \geq \cdots \geq \lambda'_n \geq \lambda_n.$$
Facts from linear algebra

Condition for eigenvalues

Let $M \succeq 0$ be a positive semidefinite $n \times n$ matrix with $k$ positive eigenvalues, all of them greater than $b' > 0$. If $w \neq 0$ and $1 + w^T (M - b'I)^{-1} w < 0$ then $M + ww^T$ has exactly $k + 1$ positive eigenvalues, all of them greater than $b'$.

- Let $\lambda_1 \geq \cdots \geq \lambda_k$ be the non-zero eigenvalues of the matrix $M$ and $\lambda'_1 \geq \cdots \geq \lambda'_{k+1}$ be the largest (in decreasing order) eigenvalues of $M + ww^T$. 

Apostolos Giannopoulos (University of Athens)
Facts from linear algebra

Condition for eigenvalues

Let $M \succeq 0$ be a positive semidefinite $n \times n$ matrix with $k$ positive eigenvalues, all of them greater than $b' > 0$. If $w \neq 0$ and $1 + w^T(M - b'I)^{-1}w < 0$ then $M + ww^T$ has exactly $k + 1$ positive eigenvalues, all of them greater than $b'$.

- Let $\lambda_1 \geq \cdots \geq \lambda_k$ be the non-zero eigenvalues of the matrix $M$ and $\lambda'_1 \geq \cdots \geq \lambda'_{k+1}$ be the largest (in decreasing order) eigenvalues of $M + ww^T$.
- Consider the quantity

$$\text{tr}((M - b'I)^{-1}) = \sum_{i=1}^{k} \frac{1}{\lambda_i - b'} + \sum_{i=k+1}^{n} \frac{1}{0 - b'}.$$
Facts from linear algebra

Condition for eigenvalues

Let $M \succeq 0$ be a positive semidefinite $n \times n$ matrix with $k$ positive eigenvalues, all of them greater than $b' > 0$. If $w \neq 0$ and $1 + w^T(M - b'I)^{-1}w < 0$ then $M + ww^T$ has exactly $k + 1$ positive eigenvalues, all of them greater than $b'$.

- Let $\lambda_1 \geq \cdots \geq \lambda_k$ be the non-zero eigenvalues of the matrix $M$ and $\lambda'_1 \geq \cdots \geq \lambda'_{k+1}$ be the largest (in decreasing order) eigenvalues of $M + ww^T$.
- Consider the quantity

$$\text{tr}((M - b'I)^{-1}) = \sum_{i=1}^{k} \frac{1}{\lambda_i - b'} + \sum_{i=k+1}^{n} \frac{1}{0 - b'}.$$  

- From the Sherman-Morrisson formula we have

$$\text{tr}((M + ww^T - b'I)^{-1}) - \text{tr}((M - b'I)^{-1}) = -\frac{w^T(M - b'I)^{-2}w}{1 + w^T(M - b'I)^{-1}w} > 0$$

because the assumption implies that the denominator on the right hand side is negative, and the numerator is positive since $M - b'I$ is non-singular, therefore $(M - b'I)^{-2}$ is positive definite.
Condition for eigenvalues

Let $M \succeq 0$ be a positive semidefinite $n \times n$ matrix with $k$ positive eigenvalues, all of them greater than $b' > 0$. If $w \neq 0$ and $1 + w^T (M - b' I)^{-1} w < 0$ then $M + ww^T$ has exactly $k + 1$ positive eigenvalues, all of them greater than $b'$.

- Computing directly the same difference we get

\[
0 < \text{tr}((M + ww^T - b' I)^{-1}) - \text{tr}((M - b' I)^{-1}) = \frac{1}{\lambda'_{k+1} - b'} - \frac{1}{0 - b'} + \sum_{i=1}^{k} \frac{1}{\lambda'_i - b'} - \sum_{i=1}^{k} \frac{1}{\lambda_i - b'} \leq \frac{1}{\lambda'_{k+1} - b'} + \frac{1}{b'},
\]

because, by Cauchy’s interlacing theorem,

\[
\lambda'_1 \geq \lambda_1 \geq \lambda'_2 \geq \cdots \geq \lambda_k \geq \lambda'_{k+1} \geq 0
\]

and hence

\[
\frac{1}{\lambda'_i - b'} - \frac{1}{\lambda_i - b'} \leq 0
\]

for every $i \leq k$. 

---

Apostolos Giannopoulos (University of Athens) Banach-Mazur distance to the cube Castro Urdiales, September 2018 36 / 43
Facts from linear algebra

Condition for eigenvalues

Let $M \succeq 0$ be a positive semidefinite $n \times n$ matrix with $k$ positive eigenvalues, all of them greater than $b' > 0$. If $w \neq 0$ and $1 + w^T (M - b' I)^{-1} w < 0$ then $M + w w^T$ has exactly $k + 1$ positive eigenvalues, all of them greater than $b'$.

- Computing directly the same difference we get

$$0 < \text{tr}((M + w w^T - b' I)^{-1}) - \text{tr}((M - b' I)^{-1}) = \frac{1}{\lambda'_{k+1} - b'} - \frac{1}{0 - b'} + \sum_{i=1}^{k} \frac{1}{\lambda'_i - b'} - \sum_{i=1}^{k} \frac{1}{\lambda_i - b'} \leq \frac{1}{\lambda'_{k+1} - b'} + \frac{1}{b'},$$

because, by Cauchy’s interlacing theorem,

$$\lambda'_1 \geq \lambda_1 \geq \lambda'_2 \geq \cdots \geq \lambda_k \geq \lambda'_{k+1} \geq 0$$

and hence

$$\frac{1}{\lambda'_i - b'} - \frac{1}{\lambda_i - b'} \leq 0$$

for every $i \leq k$.

- Since $\lambda'_{k+1} \geq 0$, we conclude that $\lambda'_{k+1} > b'$.
Proof

- For any symmetric matrix $M$ and any $b > 0$, we define the potential with barrier $b$ by

$$
\Phi_b(M) = \text{tr}\left( A^T (M - bl)^{-1} A \right).
$$
Proof

- For any symmetric matrix $M$ and any $b > 0$, we define the potential with barrier $b$ by

$$
\Phi_b(M) = \text{tr}\left( A^T (M - bI)^{-1} A \right).
$$

- We fix $\delta > 0$ to be chosen, and write $M_k$ for the matrix that has been constructed at the $k$-th step. We assume that $M_k$ has $k$ nonzero eigenvalues, all of them greater than $b_k > 0$. We set $\Phi_k(M_k) := \Phi_{b_k}(M_k)$. 

Apostolos Giannopoulos (University of Athens)
Proof

For any symmetric matrix $M$ and any $b > 0$, we define the potential with barrier $b$ by

$$\Phi_b(M) = \text{tr} \left( A^T (M - bI)^{-1} A \right).$$

We fix $\delta > 0$ to be chosen, and write $M_k$ for the matrix that has been constructed at the $k$-th step. We assume that $M_k$ has $k$ nonzero eigenvalues, all of them greater than $b_k > 0$. We set $\Phi_k(M_k) := \Phi_{b_k}(M_k)$.

Our aim is to add a rank one matrix $v \cdot v^T$ to $M_k$ so that $M_{k+1} = M_k + vv^T$ has $k + 1$ nonzero eigenvalues, all of them greater than $b_{k+1} = b_k - \delta$ and $\Phi_{k+1}(M_{k+1}) \leq \Phi_k(M_k)$. 
Proof

- For any symmetric matrix $M$ and any $b > 0$, we define the potential with barrier $b$ by
  \[ \Phi_b(M) = \text{tr} \left( A^T (M - bI)^{-1} A \right). \]

- We fix $\delta > 0$ to be chosen, and write $M_k$ for the matrix that has been constructed at the $k$-th step. We assume that $M_k$ has $k$ nonzero eigenvalues, all of them greater than $b_k > 0$. We set $\Phi_k(M_k) := \Phi_{b_k}(M_k)$.

- Our aim is to add a rank one matrix $v \cdot v^T$ to $M_k$ so that $M_{k+1} = M_k + vv^T$ has $k + 1$ nonzero eigenvalues, all of them greater than $b_{k+1} = b_k - \delta$ and $\Phi_{k+1}(M_{k+1}) \leq \Phi_k(M_k)$.

- We compute
  \[ \Phi_{k+1}(M_{k+1}) = \Phi_{k+1}(M_k) - \frac{v^T(M_k - b_{k+1}I)^{-1}AA^T(M_k - b_{k+1}I)^{-1}v}{1 + v^T(M_k - b_{k+1}I)^{-1}v}. \]
Proof

- For any symmetric matrix $M$ and any $b > 0$, we define the potential with barrier $b$ by

$$
\Phi_b(M) = \text{tr} \left( A^T (M - bI)^{-1} A \right).
$$

- We fix $\delta > 0$ to be chosen, and write $M_k$ for the matrix that has been constructed at the $k$-th step. We assume that $M_k$ has $k$ nonzero eigenvalues, all of them greater than $b_k > 0$. We set $\Phi_k(M_k) := \Phi_{b_k}(M_k)$.

- Our aim is to add a rank one matrix $v \cdot v^T$ to $M_k$ so that $M_{k+1} = M_k + vv^T$ has $k + 1$ nonzero eigenvalues, all of them greater than $b_{k+1} = b_k - \delta$ and $\Phi_{k+1}(M_{k+1}) \leq \Phi_k(M_k)$.

- We compute

$$
\Phi_{k+1}(M_{k+1}) = \Phi_{k+1}(M_k) - \frac{v^T(M_k - b_{k+1}I)^{-1}AA^T(M_k - b_{k+1}I)^{-1}v}{1 + v^T(M_k - b_{k+1}I)^{-1}v}.
$$

- So, in order to have $\Phi_{k+1}(M_{k+1}) \leq \Phi_k(M_k)$, we need to choose a vector $v$ such that

$$
- \frac{v^T(M_k - b_{k+1}I)^{-1}AA^T(M_k - b_{k+1}I)^{-1}v}{1 + v^T(M_k - b_{k+1}I)^{-1}v} \leq \Phi_k(M_k) - \Phi_{k+1}(M_k).
$$
Proof

- We saw that a sufficient condition so that $M_k + vv^T$ will have exactly $k + 1$ positive eigenvalues, all of them greater than $b_{k+1}$, is

\[
1 + v^T (M_k - b_{k+1} I)^{-1} v < 0.
\]
Proof

- We saw that a sufficient condition so that $M_k + vv^T$ will have exactly $k + 1$ positive eigenvalues, all of them greater than $b_{k+1}$, is

$$1 + v^T(M_k - b_{k+1}I)^{-1}v < 0.$$ 

- Choosing a vector $v$ that verifies both this inequality and

$$- \frac{v^T(M_k - b_{k+1}I)^{-1}AA^T(M_k - b_{k+1}I)^{-1}v}{1 + v^T(M_k - b_{k+1}I)^{-1}v} \leq \Phi_k(M_k) - \Phi_{k+1}(M_k).$$

is equivalent to choosing $v$ so that

$$v^T(M_k - b_{k+1}I)^{-1}AA^T(M_k - b_{k+1}I)^{-1}v$$

$$\leq \left(\Phi_k(M_k) - \Phi_{k+1}(M_k)\right)\left(-1 - v^T(M_k - b_{k+1}I)^{-1}v\right).$$
Proof

- We saw that a sufficient condition so that $M_k + vv^T$ will have exactly $k + 1$ positive eigenvalues, all of them greater than $b_{k+1}$, is

$$1 + v^T(M_k - b_{k+1}I)^{-1}v < 0.$$  

- Choosing a vector $v$ that verifies both this inequality and

$$-\frac{v^T(M_k - b_{k+1}I)^{-1}AA^T(M_k - b_{k+1}I)^{-1}v}{1 + v^T(M_k - b_{k+1}I)^{-1}v} \leq \Phi_k(M_k) - \Phi_{k+1}(M_k).$$

is equivalent to choosing $v$ so that

$$v^T(M_k - b_{k+1}I)^{-1}AA^T(M_k - b_{k+1}I)^{-1}v \leq \left(\Phi_k(M_k) - \Phi_{k+1}(M_k)\right)\left(-1 - v^T(M_k - b_{k+1}I)^{-1}v\right).$$

- Since $AA^T \preceq \|A\|_2^2 I$ and $(M_k - b_{k+1}I)^{-1}$ is symmetric, it is sufficient to choose $v$ so that

$$v^T(M_k - b_{k+1}I)^{-2}v \leq \frac{1}{\|A\|_2^2} \left(\Phi_k(M_k) - \Phi_{k+1}(M_k)\right)\left(-1 - v^T(M_k - b_{k+1}I)^{-1}v\right).$$
Proof

- We set $\tau_D := \{j \leq m \mid \alpha_j \neq 0\}$ where $(\alpha_j)_{j \leq m}$ are the diagonal entries of $D$. Since we have assumed that $\ker(D) \subseteq \ker(A)$, we have

$$\|A\|_{\text{HS}}^2 = \sum_{j \leq m} |Ae_j|^2 = \sum_{j \in \tau_D} |Ae_j|^2 \leq |\tau_D| \cdot \|A\|_2^2,$$

and thus $|\tau_D| \geq \|A\|_{\text{HS}}^2/\|A\|_2^2$. 

Apostolos Giannopoulos (University of Athens)  
Banach-Mazur distance to the cube  
Castro Urdiales, September 2018
Proof

- We set $\tau_D := \{ j \leq m \mid \alpha_j \neq 0 \}$ where $(\alpha_j)_{j \leq m}$ are the diagonal entries of $D$. Since we have assumed that $\text{Ker}(D) \subseteq \text{Ker}(A)$, we have

$$\|A\|_{\text{HS}}^2 = \sum_{j \leq m} |A e_j|^2 = \sum_{j \in \tau_D} |A e_j|^2 \leq |\tau_D| \cdot \|A\|_2^2,$$

and thus $|\tau_D| \geq \|A\|_{\text{HS}}^2 / \|A\|_2^2$.

- At each step, we will select a vector $v$ satisfying the condition among $(\frac{A e_j}{\alpha_j})_{j \in \tau_D}$. What we need is to find $j \in \tau_D$ such that

$$\left( A e_j \right)^T (M_k - b_{k+1} I)^{-2} A e_j \leq \frac{\Phi_k(M_k) - \Phi_{k+1}(M_k)}{\|A\|_2^2} \left( - \alpha_j^2 - \left( A e_j \right)^T (M_k - b_{k+1} I)^{-1} A e_j \right).$$
Proof

- We set $\tau_D := \{j \leq m \mid \alpha_j \neq 0\}$ where $(\alpha_j)_{j \leq m}$ are the diagonal entries of $D$. Since we have assumed that $\text{Ker}(D) \subseteq \text{Ker}(A)$, we have

$$\|A\|_{\text{HS}}^2 = \sum_{j \leq m} |Ae_j|^2 = \sum_{j \in \tau_D} |Ae_j|^2 \leq |\tau_D| \cdot \|A\|_2^2,$$

and thus $|\tau_D| \geq \|A\|_{\text{HS}}^2 / \|A\|_2^2$.

- At each step, we will select a vector $v$ satisfying the condition among $(\frac{Ae_j}{\alpha_j})_{j \in \tau_D}$. What we need is to find $j \in \tau_D$ such that

$$(Ae_j)^T (M_k - b_{k+1} I)^{-2} Ae_j$$

$$\leq \frac{\Phi_k(M_k) - \Phi_{k+1}(M_k)}{\|A\|_2^2} \left( -\alpha_j^2 - (Ae_j)^T (M_k - b_{k+1} I)^{-1} Ae_j \right).$$

- The existence of such a $j \in \tau_D$ is guaranteed by the fact that the condition holds true if we take the sum over all $(\frac{Ae_j}{\alpha_j})_{j \in \tau_D}$.
Proof

The hypothesis $\text{Ker}(D) \subset \text{Ker}(A)$ implies that

1. \[ \sum_{j \in \tau_D} (Ae_j)^T(M_k - b_{k+1}I)^{-2}Ae_j = \text{tr}\left(A^T(M_k - b_{k+1}I)^{-2}A\right), \]

2. \[ \sum_{j \in \tau_D} (Ae_j)^T(M_k - b_{k+1}I)^{-1}Ae_j = \text{tr}\left(A^T(M_k - b_{k+1}I)^{-1}A\right) = \Phi_{k+1}(M_k). \]
Proof

The hypothesis \( \ker(D) \subset \ker(A) \) implies that

\[
\sum_{j \in \tau_D} (Ae_j)^T (M_k - b_{k+1}I)^{-2} Ae_j = \text{tr} \left( A^T (M_k - b_{k+1}I)^{-2} A \right),
\]

\[
\sum_{j \in \tau_D} (Ae_j)^T (M_k - b_{k+1}I)^{-1} Ae_j = \text{tr} \left( A^T (M_k - b_{k+1}I)^{-1} A \right) = \Phi_{k+1}(M_k).
\]

Therefore it is enough to prove that, at each step,

\[
\text{tr} \left( A^T (M_k - b_{k+1}I)^{-2} A \right) \leq \frac{\Phi_k(M_k) - \Phi_{k+1}(M_k)}{\|A\|_2^2} \left( - \|D\|_{\text{HS}}^2 - \Phi_{k+1}(M_k) \right).
\]
Proof

The next lemma provides the conditions that are required at each step in order to prove

$$\text{tr}(A^T(M_k - b_{k+1}I)^{-2}A) \leq \frac{\Phi_k(M_k) - \Phi_{k+1}(M_k)}{\|A\|^2} \left( -\|D\|_{HS}^2 - \Phi_{k+1}(M_k) \right).$$
Proof

The next lemma provides the conditions that are required at each step in order to prove

\[
\text{tr}(A^T (M_k - b_{k+1} I)^{-2} A) \leq \frac{\Phi_k(M_k) - \Phi_{k+1}(M_k)}{\|A\|^2} \left( - \|D\|_{\text{HS}}^2 - \Phi_{k+1}(M_k) \right).
\]

Lemma

Suppose that \( M_k \) has \( k \) nonzero eigenvalues all greater than \( b_k \), and write \( Z_k \) for the orthogonal projection onto the kernel of \( M_k \). If

\[
\Phi_k(M_k) \leq -\|D\|_{\text{HS}}^2 - \frac{\|A\|^2}{\delta}
\]

and

\[
0 < \delta < b_k \leq \delta \frac{\|Z_k A\|_{\text{HS}}^2}{\|A\|^2},
\]

then there exists \( i \in \tau_D \) such that \( M_{k+1} := M_k + \left( \frac{Ae_i}{\alpha_i} \right) \cdot \left( \frac{Ae_i}{\alpha_i} \right)^T \) has \( k + 1 \) nonzero eigenvalues all greater than \( b_{k+1} := b_k - \delta \) and \( \Phi_{k+1}(M_{k+1}) \leq \Phi_k(M_k) \).
We are now able to complete the proof of the theorem. We must verify that the two conditions

\[ \Phi_k(M_k) \leq -\|D\|_{HS}^2 - \frac{\|A\|_2^2}{\delta} \]

and

\[ 0 < \delta < b_k \leq \delta \frac{\|Z_kA\|_{HS}^2}{\|A\|_2^2}, \]

of the Lemma hold at each step.
We are now able to complete the proof of the theorem. We must verify that the two conditions

$$\Phi_k(M_k) \leq -\|D\|_{HS}^2 - \frac{\|A\|_2^2}{\delta}$$

and

$$0 < \delta < b_k \leq \delta \frac{\|Z_kA\|_{HS}^2}{\|A\|_2^2},$$

of the Lemma hold at each step.

At the beginning we have $M_0 = 0$ and $Z_k = I$, so we must choose a barrier $b_0$ such that:

$$-\frac{\|A\|_{HS}^2}{b_0} \leq -\|D\|_{HS}^2 - \frac{\|A\|_2^2}{\delta}$$

and

$$b_0 \leq \delta \frac{\|A\|_{HS}^2}{\|A\|_2^2}.$$
Proof

- We are now able to complete the proof of the theorem. We must verify that the two conditions

\[ \Phi_k(M_k) \leq -\|D\|_{HS}^2 - \frac{\|A\|_2^2}{\delta} \]

and

\[ 0 < \delta < b_k \leq \delta \frac{\|Z_k A\|_{HS}^2}{\|A\|_2^2}, \]

of the Lemma hold at each step.

- At the beginning we have \( M_0 = 0 \) and \( Z_k = I \), so we must choose a barrier \( b_0 \) such that:

\[ -\frac{\|A\|_{HS}^2}{b_0} \leq -\|D\|_{HS}^2 - \frac{\|A\|_2^2}{\delta} \]

and

\[ b_0 \leq \delta \frac{\|A\|_{HS}^2}{\|A\|_2^2} . \]

- We choose

\[ b_0 := \epsilon \frac{\|A\|_{HS}^2}{\|D\|_{HS}^2} \quad \text{and} \quad \delta := \frac{\epsilon}{1 - \epsilon} \frac{\|A\|_2^2}{\|D\|_{HS}^2}. \]
Proof

We choose

\[ b_0 := \epsilon \frac{\|A\|_{HS}^2}{\|D\|_{HS}^2} \quad \text{and} \quad \delta := \frac{\epsilon}{1 - \epsilon} \frac{\|A\|_{2}^2}{\|D\|_{HS}^2}. \]
Proof

- We choose

\[ b_0 := \epsilon \frac{\|A\|_{\text{HS}}^2}{\|D\|_{\text{HS}}^2} \quad \text{and} \quad \delta := \frac{\epsilon}{1 - \epsilon} \frac{\|A\|_{\text{HS}}^2}{\|D\|_{\text{HS}}^2}. \]

- At the \((k + 1)\)-th step

\[ \Phi_{k+1}(M_{k+1}) \leq -\|D\|_{\text{HS}}^2 - \frac{\|A\|_{\text{HS}}^2}{\delta} \]

holds because \(\Phi_{k+1}(M_{k+1}) \leq \Phi_k(M_k)\).
Proof

- We choose 
  \[ b_0 := \epsilon \|A\|_{\text{HS}}^2 / \|D\|_{\text{HS}}^2 \quad \text{and} \quad \delta := \frac{\epsilon}{1 - \epsilon} \|A\|_{\text{HS}}^2 / \|D\|_{\text{HS}}^2. \]

- At the \((k + 1)\)-th step
  \[ \Phi_{k+1}(M_{k+1}) \leq -\|D\|_{\text{HS}}^2 - \frac{\|A\|_{\text{HS}}^2}{\delta} \]
  holds because \( \Phi_{k+1}(M_{k+1}) \leq \Phi_k(M_k) \).

- Since \( \|Z_kA\|_{\text{HS}}^2 \) decreases at each step by at most \( \|A\|_{\text{HS}}^2 \), the right-hand side of
  \[ 0 < \delta < b_k \leq \delta \frac{\|Z_kA\|_{\text{HS}}^2}{\|A\|_{\text{HS}}^2} \]
  decreases by at most \( \delta \), and therefore \( b_{k+1} \leq \delta \frac{\|Z_{k+1}A\|_{\text{HS}}^2}{\|A\|_{\text{HS}}^2} \) also holds.
Proof

- We choose
  \[ b_0 := \epsilon \|A\|_{\text{HS}}^2 / \|D\|_{\text{HS}}^2 \text{ and } \delta := \frac{\epsilon}{1 - \epsilon} \|A\|_{\text{HS}}^2 / \|D\|_{\text{HS}}^2. \]

- At the \((k + 1)\)-th step
  \[ \Phi_{k+1}(M_{k+1}) \leq -\|D\|_{\text{HS}}^2 - \frac{\|A\|_{\text{HS}}^2}{\delta} \]
  holds because \(\Phi_{k+1}(M_{k+1}) \leq \Phi_k(M_k)\).

- Since \(\|Z_k A\|_{\text{HS}}^2\) decreases at each step by at most \(\|A\|_{\text{HS}}^2\), the right-hand side of
  \[ 0 < \delta < b_k < \delta \frac{\|Z_k A\|_{\text{HS}}^2}{\|A\|_{\text{HS}}^2}, \]
  decreases by at most \(\delta\), and therefore \(b_{k+1} \leq \delta \frac{\|Z_{k+1} A\|_{\text{HS}}^2}{\|A\|_{\text{HS}}^2}\) also holds.

- Finally note that, after \(k_0 = (1 - \epsilon)^2 \|A\|_{\text{HS}}^2 / \|A\|_{\text{HS}}^2\) steps, the barrier will be
  \[ b_{k_0} = b_0 - k_0 \delta = \epsilon^2 \|A\|_{\text{HS}}^2 / \|D\|_{\text{HS}}^2. \]

This completes the proof.