Hardness of Grothendieck Problems

Jop Briët, Oded Regev and Rishi Saket
(CWI) (NYU) (IBM)

Marne-la-Vallée, Oct. 27 2015

The Grothendieck Inequality

Given a matrix $M \in \mathbb{R}^{d \times d}$ (bilinear form on ℓ_∞) let

$$\mathrm{OPT}(M) = \max_{\alpha_i, \beta_j \in \{\pm 1\}} \sum M_{ij} \alpha_i \beta_j$$

$$\mathrm{SDP}(M) = \sup_{\|a_i\|_{\ell_2} = \|b_j\|_{\ell_2} = 1} \sum M_{ij} \langle a_i, b_j \rangle$$

Thm. (Grothendieck '53)

There exists an absolute constant K_G such that for any M,

$$\mathrm{OPT}(M) \leq \mathrm{SDP}(M) \leq K_G \, \mathrm{OPT}(M)$$

$K_G \leq 1.78\ldots$ (Krivine '79) $-\ \varepsilon$ (Braverman et al. '12)

The Grothendieck Inequality

Given a matrix $M \in \mathbb{R}^{d \times d}$ (bilinear form on ℓ_∞) let

$$\mathrm{OPT}(M) = \max_{\alpha_i, \beta_j \in \{\pm 1\}} \sum M_{ij} \alpha_i \beta_j$$

$$\mathrm{SDP}(M) = \sup_{\|a_i\|_{\ell_2} = \|b_j\|_{\ell_2} = 1} \sum M_{ij} \langle a_i, b_j \rangle$$

Thm. (Grothendieck '53)

There exists an absolute constant K_G such that for any M,

$$\mathrm{OPT}(M) \le \mathrm{SDP}(M) \le K_G \, \mathrm{OPT}(M)$$

$K_G \le 1.78\ldots$ (Krivine '79) $-\ \varepsilon$ (Braverman et al. '12)

The Grothendieck Inequality

Given a matrix $M \in \mathbb{R}^{d \times d}$ (bilinear form on ℓ_∞) let

$$\mathrm{OPT}(M) = \max_{\alpha_i, \beta_j \in \{\pm 1\}} \sum M_{ij} \alpha_i \beta_j$$

$$\mathrm{SDP}(M) = \sup_{\|a_i\|_{\ell_2} = \|b_j\|_{\ell_2} = 1} \sum M_{ij} \langle a_i, b_j \rangle$$

Thm. (Grothendieck '53)

There exists an absolute constant K_G such that for any M,

$$\mathrm{OPT}(M) \le \mathrm{SDP}(M) \le K_G \, \mathrm{OPT}(M)$$

$K_G \le 1.78\ldots$ (Krivine '79) $-\ \varepsilon$ (Braverman et al. '12)

The Grothendieck Inequality

Given a matrix $M \in \mathbb{R}^{d \times d}$ (bilinear form on ℓ_∞) let

$$\mathrm{OPT}(M) = \max_{\alpha_i, \beta_j \in \{\pm 1\}} \sum M_{ij} \alpha_i \beta_j$$

$$\mathrm{SDP}(M) = \sup_{\|a_i\|_{\ell_2} = \|b_j\|_{\ell_2} = 1} \sum M_{ij} \langle a_i, b_j \rangle$$

Thm. (Grothendieck '53)

There exists an absolute constant K_G such that for any M,

$$\mathrm{OPT}(M) \leq \mathrm{SDP}(M) \leq K_G \, \mathrm{OPT}(M)$$

$K_G \leq 1.78\ldots$ (Krivine '79) $- \varepsilon$ (Braverman et al. '12)

The Grothendieck Inequality

Given a matrix $M \in \mathbb{R}^{d \times d}$ (bilinear form on ℓ_∞) let

$$\mathrm{OPT}(M) = \max_{\alpha_i, \beta_j \in \{\pm 1\}} \sum M_{ij} \alpha_i \beta_j$$

$$\mathrm{SDP}(M) = \sup_{\|a_i\|_{\ell_2} = \|b_j\|_{\ell_2} = 1} \sum M_{ij} \langle a_i, b_j \rangle$$

Thm. (Grothendieck '53)

There exists an absolute constant K_G such that for any M,

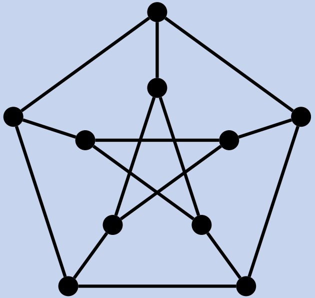
$$\mathrm{OPT}(M) \leq \mathrm{SDP}(M) \leq K_G \, \mathrm{OPT}(M)$$

$K_G \leq 1.78\ldots$ (Krivine '79) $- \varepsilon$ (Braverman et al. '12)

Important computer science example: *MAX CUT*

Maximum Cut Problem

Given a graph $G = (V, E)$, compute the maximum number of edges crossing a bi-partition of V:



$L = \text{Deg}_G - \text{Adj}_G$ (Laplacian)

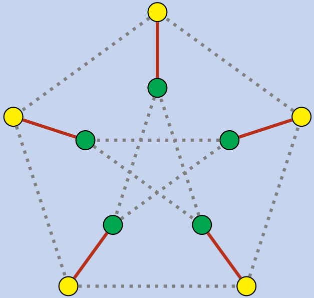
$\text{MAX CUT(G)} = \frac{1}{4}\,\text{OPT}(L)$

- Efficiently algo. for $\text{OPT}(M) \Rightarrow$ P=NP $\Rightarrow$ **\$1,000,000**!

Important computer science example: *MAX CUT*

Maximum Cut Problem

Given a graph $G = (V, E)$, compute the maximum number of edges crossing a bi-partition of V:



$L = \mathrm{Deg}_G - \mathrm{Adj}_G$ (Laplacian)

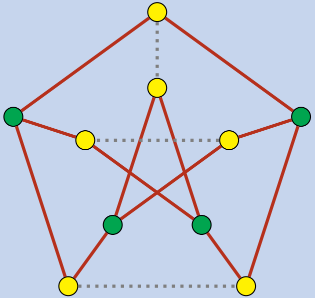
$\text{MAX CUT(G)} = \frac{1}{4}\,\mathrm{OPT}(L)$

- Efficiently algo. for $\mathrm{OPT}(M)$ $\Rightarrow$ P=NP $\Rightarrow$ **\$1,000,000**!

Important computer science example: *MAX CUT*

Maximum Cut Problem

Given a graph $G = (V, E)$, compute the maximum number of edges crossing a bi-partition of V:



$L = \text{Deg}_G - \text{Adj}_G$ (Laplacian)

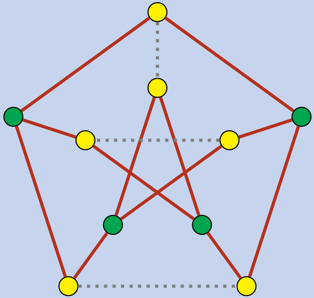
$\text{MAX CUT(G)} = \frac{1}{4}\,\text{OPT}(L)$

- Efficiently algo. for OPT(M) $\Rightarrow$ P=NP $\Rightarrow$ **$1,000,000**!

Important computer science example: *MAX CUT*

Maximum Cut Problem

Given a graph $G = (V, E)$, compute the maximum number of edges crossing a bi-partition of V:



$L = \mathrm{Deg}_G - \mathrm{Adj}_G$ (Laplacian)

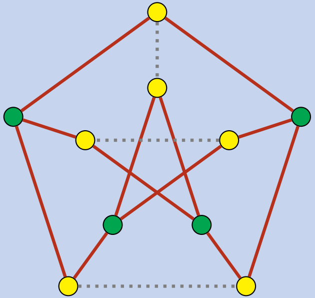
$\text{MAX CUT(G)} = \frac{1}{4}\,\mathrm{OPT}(L)$

- Efficiently algo. for $\mathrm{OPT}(M)$ $\Rightarrow$ P=NP $\Rightarrow$ **$1,000,000**!

Important computer science example: *MAX CUT*

Maximum Cut Problem

Given a graph $G = (V, E)$, compute the maximum number of edges crossing a bi-partition of V:



$L = \mathrm{Deg}_G - \mathrm{Adj}_G$ (Laplacian)

$\text{MAX CUT(G)} = \frac{1}{4}\,\mathrm{OPT}(L)$

- Efficiently algo. for $\mathrm{OPT}(M) \Rightarrow$ P=NP $\Rightarrow$ **\$1,000,000**!

Relevance to CS: Alon–Naor '06

- Computing OPT(M) is interesting!
  - ⋆ MAX CUT.
  - ⋆ Computing the cut-norm of a matrix; graph partitioning.
  - ⋆ Statistical physics; ground state energy in spin glasses.

- SDP(M) can be computed efficiently using... "SDP"!

- Groth. Ineq. ⇒ OPT(M) can be approximated efficiently to within a factor K_G.

- Algorithmic proof of Groth. Ineq. gives an efficient algorithm that, given M, outputs $\alpha_i, \beta_j \in \{\pm 1\}$ s.t.

$$\frac{1}{K_G} \mathrm{OPT}(M) \leq \sum M_{ij} \alpha_i \beta_j \leq \mathrm{OPT}(M)$$

Relevance to CS: Alon–Naor '06

- Computing OPT(M) is interesting!
  - ⋆ MAX CUT.
  - ⋆ Computing the cut-norm of a matrix; graph partitioning.
  - ⋆ Statistical physics; ground state energy in spin glasses.

- SDP(M) can be computed efficiently using... "SDP"!

- Groth. Ineq. ⇒ OPT(M) can be approximated efficiently to within a factor K_G.

- Algorithmic proof of Groth. Ineq. gives an efficient algorithm that, given M, outputs $\alpha_i, \beta_j \in \{\pm 1\}$ s.t.

$$\frac{1}{K_G}\,\mathrm{OPT}(M) \leq \sum M_{ij}\alpha_i\beta_j \leq \mathrm{OPT}(M)$$

Relevance to CS: Alon–Naor '06

- Computing OPT(M) is interesting!
  - ⋆ MAX CUT.
  - ⋆ Computing the cut-norm of a matrix; graph partitioning.
  - ⋆ Statistical physics; ground state energy in spin glasses.

- SDP(M) can be computed efficiently using... "SDP"!

- Groth. Ineq. ⇒ OPT(M) can be approximated efficiently to within a factor K_G.

- Algorithmic proof of Groth. Ineq. gives an efficient algorithm that, given M, outputs $\alpha_i, \beta_j \in \{\pm 1\}$ s.t.

$$\frac{1}{K_G}\,\mathrm{OPT}(M) \le \sum M_{ij}\alpha_i\beta_j \le \mathrm{OPT}(M)$$

Relevance to CS: Alon–Naor '06

- Computing OPT(M) is interesting!
  - ⋆ MAX CUT.
  - ⋆ Computing the cut-norm of a matrix; graph partitioning.
  - ⋆ Statistical physics; ground state energy in spin glasses.

- SDP(M) can be computed efficiently using… "SDP"!

- Groth. Ineq. $\Rightarrow$ OPT(M) can be approximated efficiently to within a factor K_G.

- Algorithmic proof of Groth. Ineq. gives an efficient algorithm that, given M, outputs $\alpha_i, \beta_j \in \{\pm 1\}$ s.t.

$$\frac{1}{K_G}\,\mathrm{OPT}(M) \le \sum M_{ij}\alpha_i\beta_j \le \mathrm{OPT}(M)$$

Hardness of the Grothendieck Problem

- Better approximation algorithms for OPT(M)?
- Probably not if Khot's *Unique Games Conjecture* holds.

Theorem (Khot–O'Donnell '09), (Raghavendra–Steurer '09).

If the UGC is true and P$\neq$NP, then for any $\varepsilon > 0$, there is **no** efficient algo. that approximates OPT(M) to within a factor $K_G - \varepsilon$.

- (Whatever the value of K_G may be!)

Hardness of the Grothendieck Problem

- Better approximation algorithms for OPT(M)?
- Probably not if Khot's *Unique Games Conjecture* holds.

Theorem (Khot–O'Donnell '09), (Raghavendra–Steurer '09).

If the UGC is true and P$\neq$NP, then for any $\varepsilon > 0$, there is **no** efficient algo. that approximates OPT(M) to within a factor $K_G - \varepsilon$.

- (Whatever the value of K_G may be!)

Hardness of the Grothendieck Problem

- Better approximation algorithms for $\mathrm{OPT}(M)$?
- Probably not if Khot's *Unique Games Conjecture* holds.

Theorem (Khot–O'Donnell '09), (Raghavendra–Steurer '09).

If the UGC is true and P$\neq$NP, then for any $\varepsilon > 0$, there is **no** efficient algo. that approximates $\mathrm{OPT}(M)$ to within a factor $K_G - \varepsilon$.

- (Whatever the value of K_G may be!)

The NC Grothendieck Inequality

Given a tensor $T \in \mathbb{R}^{d \times d \times d \times d}$ (bilinear form on matrices) let

$$\mathrm{OPT}(T) = \max_{\substack{A,B \in \mathbb{C}^{d \times d} \\ \text{unitary}}} \left| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \right|$$

$$\mathrm{SDP}(T) = \max_{\vec{A},\vec{B}} \left| \sum T_{ijkl} \langle \vec{A}_{ij}, \vec{B}_{kl} \rangle \right|$$

Conjecture (Groth. '53) Theorem (Pisier '78) (Haagerup '85)

Haagerup–Itoh '95: 2 is optimal! (NC Groth. constant *is* known.)

The NC Grothendieck Inequality

Given a tensor $T \in \mathbb{R}^{d\times d\times d\times d}$ (bilinear form on matrices) let

$$\mathrm{OPT}(T) = \max_{\substack{A,B\in\mathbb{C}^{d\times d}\\ \text{unitary}}} \left| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \right|$$

$$\mathrm{SDP}(T) = \max_{\vec{A},\vec{B}} \left| \sum T_{ijkl} \langle \vec{A}_{ij}, \vec{B}_{kl} \rangle \right|$$

Conjecture (Groth. '53) Theorem (Pisier '78) (Haagerup '85)

Haagerup–Itoh '95: 2 is optimal! (NC Groth. constant *is* known.)

The NC Grothendieck Inequality

Given a tensor $T \in \mathbb{R}^{d\times d\times d\times d}$ (bilinear form on matrices) let

$$\mathrm{OPT}(T) = \max_{\substack{A,B\in\mathbb{C}^{d\times d}\\ \text{unitary}}} \left| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \right|$$

$$\mathrm{SDP}(T) = \max_{\vec{A},\vec{B}} \left| \sum T_{ijkl} \langle \vec{A}_{ij}, \vec{B}_{kl} \rangle \right|$$

Conjecture (Groth. '53) Theorem (Pisier '78) (Haagerup '85)

There exists an absolute constant K such that for any T,

$$\mathrm{OPT}(T) \le \mathrm{SDP}(T) \le K\,\mathrm{OPT}(T)$$

Haagerup–Itoh '95: 2 is optimal! (NC Groth. constant *is* known.)

The NC Grothendieck Inequality

Given a tensor $T \in \mathbb{R}^{d\times d\times d\times d}$ (bilinear form on matrices) let

$$\mathrm{OPT}(T) = \max_{\substack{A,B\in\mathbb{C}^{d\times d}\\ \text{unitary}}} \left| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \right|$$

$$\mathrm{SDP}(T) = \max_{\vec{A},\vec{B}} \left| \sum T_{ijkl} \langle \vec{A}_{ij}, \vec{B}_{kl} \rangle \right|$$

Conjecture (Groth. '53) Theorem (Pisier '78) (Haagerup '85)

There exists an absolute constant K such that for any T,

$$\mathrm{OPT}(T) \le \mathrm{SDP}(T) \le K\,\mathrm{OPT}(T)$$

Haagerup–Itoh '95: 2 is optimal! (NC Groth. constant *is* known.)

The NC Grothendieck Inequality

Given a tensor $T \in \mathbb{R}^{d\times d\times d\times d}$ (bilinear form on matrices) let

$$\mathrm{OPT}(T) = \max_{\substack{A,B\in\mathbb{C}^{d\times d}\\ \text{unitary}}} \left| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \right|$$

$$\mathrm{SDP}(T) = \max_{\vec{A},\vec{B}} \left| \sum T_{ijkl} \langle \vec{A}_{ij}, \vec{B}_{kl} \rangle \right|$$

Conjecture (Groth. '53) Theorem (Pisier '78) (Haagerup '85)

For any T,

$$\mathrm{OPT}(T) \le \mathrm{SDP}(T) \le 2\,\mathrm{OPT}(T)$$

Haagerup–Itoh '95: 2 is optimal! (NC Groth. constant *is* known.)

The NC Grothendieck Inequality

Given a tensor $T \in \mathbb{R}^{d \times d \times d \times d}$ (bilinear form on matrices) let

$$\mathrm{OPT}(T) = \max_{\substack{A,B \in \mathbb{C}^{d\times d} \\ \text{unitary}}} \left| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \right|$$

$$\mathrm{SDP}(T) = \max_{\vec{A},\vec{B}} \left| \sum T_{ijkl} \langle \vec{A}_{ij}, \vec{B}_{kl} \rangle \right|$$

Conjecture (Groth. '53) Theorem (Pisier '78) (Haagerup '85)

For any T,

$$\mathrm{OPT}(T) \leq \mathrm{SDP}(T) \leq 2\,\mathrm{OPT}(T)$$

Haagerup–Itoh '95: 2 is optimal! (NC Groth. constant *is* known.)

Relevance to CS: Naor–Regev–Vidick '14

- Computing OPT(T) is interesting!
  - ⋆ Procrustes problem
  - ⋆ Principle component analysis
  - ⋆ Quantum XOR games
  - ⋆ Bandeira–Kennedy–Singer '13: Global Registration

- SDP(T) can be computed efficiently using… "SDP"!

- NCG ⇒ can efficiently approx. OPT(T) to within a factor **2**.

- Algorithmic version of Haagerup's proof gives an efficient algorithm that, given T, outputs unitary matrices A, B s.t.

$$\frac{1}{2}\mathrm{OPT}(T) \leq \left| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \right| \leq \mathrm{OPT}(T)$$

Relevance to CS: Naor–Regev–Vidick '14

- Computing OPT(T) is interesting!
  - ⋆ Procrustes problem
  - ⋆ Principle component analysis
  - ⋆ Quantum XOR games
  - ⋆ Bandeira–Kennedy–Singer '13: Global Registration
- SDP(T) can be computed efficiently using... "SDP"!
- NCG ⇒ can efficiently approx. OPT(T) to within a factor **2**.
- Algorithmic version of Haagerup's proof gives an efficient algorithm that, given T, outputs unitary matrices A, B s.t.

$$\frac{1}{2}\mathrm{OPT}(T) \leq \left| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \right| \leq \mathrm{OPT}(T)$$

Relevance to CS: Naor–Regev–Vidick '14

- Computing OPT(T) is interesting!
  - ⋆ Procrustes problem
  - ⋆ Principle component analysis
  - ⋆ Quantum XOR games
  - ⋆ Bandeira–Kennedy–Singer '13: Global Registration

- SDP(T) can be computed efficiently using... "SDP"!

- NCG ⇒ can efficiently approx. OPT(T) to within a factor **2**.

- Algorithmic version of Haagerup's proof gives an efficient algorithm that, given T, outputs unitary matrices A, B s.t.

$$\frac{1}{2}\mathrm{OPT}(T) \leq \left|\sum T_{ijkl} A_{ij} \overline{B_{kl}}\right| \leq \mathrm{OPT}(T)$$

Relevance to CS: Naor–Regev–Vidick '14

- Computing $\mathrm{OPT}(T)$ is interesting!
  - ⋆ Procrustes problem
  - ⋆ Principle component analysis
  - ⋆ Quantum XOR games
  - ⋆ Bandeira–Kennedy–Singer '13: Global Registration
- $\mathrm{SDP}(T)$ can be computed efficiently using… "SDP"!
- NCG ⇒ can efficiently approx. $\mathrm{OPT}(T)$ to within a factor **2**.
- Algorithmic version of Haagerup's proof gives an efficient algorithm that, given T, outputs unitary matrices A, B s.t.

$$\frac{1}{2}\mathrm{OPT}(T) \leq \Big| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \Big| \leq \mathrm{OPT}(T)$$

Main result: tight hardness of the NCG

- No hardness for the NCG was known besides $1.4\ldots$- UGC hardness of approx. due to the (complex) Groth. Problem.

Theorem (B–Regev–Saket).

If P≠NP, then for any $\varepsilon > 0$, there is **no** efficient algo. that approximates $\mathrm{OPT}(T)$ to within a factor $\mathbf{2}-\varepsilon$.

Main result: tight hardness of the NCG

- No hardness for the NCG was known besides $1.4\ldots$- UGC hardness of approx. due to the (complex) Groth. Problem.

Theorem (B–Regev–Saket).

If P$\neq$NP, then for any $\varepsilon > 0$, there is **no** efficient algo. that approximates OPT(T) to within a factor $\mathbf{2}-\varepsilon$.

Hardness of the *Little* NCG

- We prove our main result by proving something stronger: optimal hardness for a special case.

Theorem 2 (B–Regev–Saket).

If P≠NP, then for any $\varepsilon > 0$, there is no efficient $(\sqrt{2} - \varepsilon)$-approx. algo. for the *Little NCG*.

- Implies the main result since the translation involves a square.

Hardness of the *Little* NCG

- We prove our main result by proving something stronger: optimal hardness for a special case.

Theorem 2 (B–Regev–Saket).

If P$\neq$NP, then for any $\varepsilon > 0$, there is no efficient $(\sqrt{2} - \varepsilon)$-approx. algo. for the *Little NCG*.

- Implies the main result since the translation involves a square.

Hardness of the *Little* NCG

- We prove our main result by proving something stronger: optimal hardness for a special case.

Theorem 2 (B–Regev–Saket).

If P≠NP, then for any $\varepsilon > 0$, there is no efficient $(\sqrt{2} - \varepsilon)$-approx. algo. for the *Little NCG*.

- Implies the main result since the translation involves a square.

Warm-up: Hardness for Little Groth

Theorem 3 (B–Regev–Saket).

If P$\neq$NP, for any $\varepsilon > 0$, there is no efficient $(\sqrt{\pi/2} - \varepsilon)$-approx. algo. for the *Little Groth. Problem*.

- Tight! Previously known under the UGC (Khot–Naor '09).

Warm-up: Hardness for Little Groth

Theorem 3 (B–Regev–Saket).

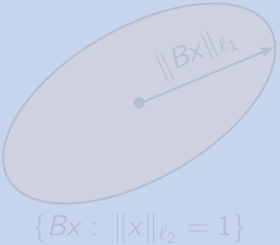
If P$\neq$NP, for any $\varepsilon > 0$, there is no efficient $(\sqrt{\pi/2} - \varepsilon)$-approx. algo. for the *Little Groth. Problem*.

- Tight! Previously known under the UGC (Khot–Naor '09).

The Little Groth. Problem.

Given a matrix B, compute:

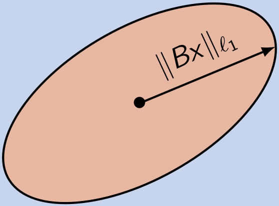
$$\|B\|_{\ell_2 \to \ell_1} := \max_{\|x\|_{\ell_2}=1} \|Bx\|_{\ell_1}$$



The Little Groth. Problem.

Given a matrix B, compute:

$$\|B\|_{\ell_2 \to \ell_1} := \max_{\|x\|_{\ell_2}=1} \|Bx\|_{\ell_1}$$



$\{Bx : \|x\|_{\ell_2} = 1\}$

Why is this a special case?

The Little Groth. Problem.

$$\begin{aligned}\|B\|_{\ell_2\to\ell_1} &:= \max_{\|x\|_{\ell_2}=1} \|Bx\|_{\ell_1} \\ &= \max_{\|x\|_{\ell_2}=\|y\|_{\ell_\infty}=1} y^\mathsf{T} Bx \\ &= \max_{\|y\|_{\ell_\infty}=1} \|y^\mathsf{T} B\|_{\ell_2} \qquad \|B\|_{\ell_2\to\ell_1} = \|B^*\|_{\ell_\infty\to\ell_2} \\ &= \max_{\|y\|_{\ell_\infty}=1} \sqrt{y^\mathsf{T} BB^\mathsf{T} y} \\ &= \max_{\|y\|_{\ell_\infty}=\|z\|_{\ell_\infty}=1} \sqrt{y^\mathsf{T} BB^\mathsf{T} z} \quad \text{Cauchy–Schwarz}\end{aligned}$$

Hence, for $M = BB^\mathsf{T}$, we have $\|B\|^2_{\ell_2\to\ell_1} = \mathrm{OPT}(M)$.

Why is this a special case?

The Little Groth. Problem.

$$\begin{aligned}
\|B\|_{\ell_2\to\ell_1} &:= \max_{\|x\|_{\ell_2}=1} \|Bx\|_{\ell_1} \\
&= \max_{\|x\|_{\ell_2}=\|y\|_{\ell_\infty}=1} y^\mathsf{T} Bx \\
&= \max_{\|y\|_{\ell_\infty}=1} \|y^\mathsf{T} B\|_{\ell_2} \qquad \|B\|_{\ell_2\to\ell_1} = \|B^*\|_{\ell_\infty\to\ell_2} \\
&= \max_{\|y\|_{\ell_\infty}=1} \sqrt{y^\mathsf{T} BB^\mathsf{T} y} \\
&= \max_{\|y\|_{\ell_\infty}=\|z\|_{\ell_\infty}=1} \sqrt{y^\mathsf{T} BB^\mathsf{T} z} \quad \text{Cauchy–Schwarz}
\end{aligned}$$

Hence, for $M = BB^\mathsf{T}$, we have $\|B\|^2_{\ell_2\to\ell_1} = \mathrm{OPT}(M)$.

Why is this a special case?

The Little Groth. Problem.

$$\begin{aligned}\|B\|_{\ell_2\to\ell_1} &:= \max_{\|x\|_{\ell_2}=1} \|Bx\|_{\ell_1} \\ &= \max_{\|x\|_{\ell_2}=\|y\|_{\ell_\infty}=1} y^{\mathsf{T}}Bx \\ &= \max_{\|y\|_{\ell_\infty}=1} \|y^{\mathsf{T}}B\|_{\ell_2} \qquad \|B\|_{\ell_2\to\ell_1} = \|B^*\|_{\ell_\infty\to\ell_2} \\ &= \max_{\|y\|_{\ell_\infty}=1} \sqrt{y^{\mathsf{T}}BB^{\mathsf{T}}y} \\ &= \max_{\|y\|_{\ell_\infty}=\|z\|_{\ell_\infty}=1} \sqrt{y^{\mathsf{T}}BB^{\mathsf{T}}z} \quad \text{Cauchy–Schwarz}\end{aligned}$$

Hence, for $M = BB^{\mathsf{T}}$, we have $\|B\|^2_{\ell_2\to\ell_1} = \mathrm{OPT}(M)$.

Why is this a special case?

The Little Groth. Problem.

$$\|B\|_{\ell_2\to\ell_1} := \max_{\|x\|_{\ell_2}=1} \|Bx\|_{\ell_1}$$

$$= \max_{\|x\|_{\ell_2}=\|y\|_{\ell_\infty}=1} y^{\mathsf{T}}Bx$$

$$= \max_{\|y\|_{\ell_\infty}=1} \|y^{\mathsf{T}}B\|_{\ell_2} \qquad \|B\|_{\ell_2\to\ell_1} = \|B^*\|_{\ell_\infty\to\ell_2}$$

$$= \max_{\|y\|_{\ell_\infty}=1} \sqrt{y^{\mathsf{T}}BB^{\mathsf{T}}y}$$

$$= \max_{\|y\|_{\ell_\infty}=\|z\|_{\ell_\infty}=1} \sqrt{y^{\mathsf{T}}BB^{\mathsf{T}}z} \quad \text{Cauchy–Schwarz}$$

Hence, for $M = BB^{\mathsf{T}}$, we have $\|B\|^2_{\ell_2\to\ell_1} = \mathrm{OPT}(M)$.

Why is this a special case?

The Little Groth. Problem.

$$\|B\|_{\ell_2\to\ell_1} := \max_{\|x\|_{\ell_2}=1} \|Bx\|_{\ell_1}$$

$$= \max_{\|x\|_{\ell_2}=\|y\|_{\ell_\infty}=1} y^\top Bx$$

$$= \max_{\|y\|_{\ell_\infty}=1} \|y^\top B\|_{\ell_2} \qquad \|B\|_{\ell_2\to\ell_1} = \|B^*\|_{\ell_\infty\to\ell_2}$$

$$= \max_{\|y\|_{\ell_\infty}=1} \sqrt{y^\top BB^\top y}$$

$$= \max_{\|y\|_{\ell_\infty}=\|z\|_{\ell_\infty}=1} \sqrt{y^\top BB^\top z} \quad \text{Cauchy–Schwarz}$$

Hence, for $M = BB^\top$, we have $\|B\|^2_{\ell_2\to\ell_1} = \mathrm{OPT}(M)$.

Why is this a special case?

The Little Groth. Problem.

$$\begin{aligned}
\|B\|_{\ell_2\to\ell_1} &:= \max_{\|x\|_{\ell_2}=1} \|Bx\|_{\ell_1} \\
&= \max_{\|x\|_{\ell_2}=\|y\|_{\ell_\infty}=1} y^{\mathsf{T}}Bx \\
&= \max_{\|y\|_{\ell_\infty}=1} \|y^{\mathsf{T}}B\|_{\ell_2} \qquad \|B\|_{\ell_2\to\ell_1} = \|B^*\|_{\ell_\infty\to\ell_2} \\
&= \max_{\|y\|_{\ell_\infty}=1} \sqrt{y^{\mathsf{T}}BB^{\mathsf{T}}y} \\
&= \max_{\|y\|_{\ell_\infty}=\|z\|_{\ell_\infty}=1} \sqrt{y^{\mathsf{T}}BB^{\mathsf{T}}z} \qquad \text{Cauchy–Schwarz}
\end{aligned}$$

Hence, for $M = BB^{\mathsf{T}}$, we have $\|B\|^2_{\ell_2\to\ell_1} = \mathrm{OPT}(M)$.

Why is this a special case?

The Little Groth. Problem.

$$\begin{aligned}
\|B\|_{\ell_2\to\ell_1} &:= \max_{\|x\|_{\ell_2}=1} \|Bx\|_{\ell_1} \\
&= \max_{\|x\|_{\ell_2}=\|y\|_{\ell_\infty}=1} y^\mathsf{T} Bx \\
&= \max_{\|y\|_{\ell_\infty}=1} \|y^\mathsf{T} B\|_{\ell_2} \qquad \|B\|_{\ell_2\to\ell_1} = \|B^*\|_{\ell_\infty\to\ell_2} \\
&= \max_{\|y\|_{\ell_\infty}=1} \sqrt{y^\mathsf{T} BB^\mathsf{T} y} \\
&= \max_{\|y\|_{\ell_\infty}=\|z\|_{\ell_\infty}=1} \sqrt{y^\mathsf{T} BB^\mathsf{T} z} \qquad \text{Cauchy–Schwarz}
\end{aligned}$$

Hence, for $M = BB^\mathsf{T}$, we have $\|B\|^2_{\ell_2\to\ell_1} = \mathrm{OPT}(M)$.

Embedding $\mathbb{R}^n$ into L_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little Groth. to the following lemma.

Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{R}^n \to L_1$ s.t.

1.) $\forall i \in [n]$, $\quad \|f(e_i)\|_{L_1} = 1$

2.) $\forall \|a\|_{\ell_2} = 1$, $\quad \|a\|_{\ell_\infty} < \frac{\varepsilon}{2} \Rightarrow \|f(a)\|_{L_1} < \sqrt{\frac{2}{\pi}} + \varepsilon.$

Subspace of L_1 whose unit ball looks like:

Embedding $\mathbb{R}^n$ into L_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little Groth. to the following lemma.

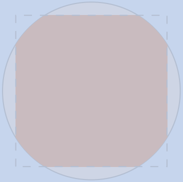
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{R}^n \to L_1$ s.t.

1.) $\forall i \in [n]$, $\quad \|f(e_i)\|_{L_1} = 1$

2.) $\forall \|a\|_{\ell_2} = 1$, $\quad \|a\|_{\ell_\infty} < \frac{\varepsilon}{2} \Rightarrow \|f(a)\|_{L_1} < \sqrt{\frac{2}{\pi}} + \varepsilon.$

Subspace of L_1 whose unit ball looks like:



Embedding $\mathbb{R}^n$ into L_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little Groth. to the following lemma.

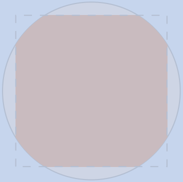
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{R}^n \to L_1$ s.t.

1.) $\forall i \in [n], \quad \|f(e_i)\|_{L_1} = 1$

2.) $\forall \|a\|_{\ell_2} = 1, \quad \|a\|_{\ell_\infty} < \frac{\varepsilon}{2} \Rightarrow \|f(a)\|_{L_1} < \sqrt{\frac{2}{\pi}} + \varepsilon.$

Subspace of L_1 whose unit ball looks like:



Embedding $\mathbb{R}^n$ into L_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little Groth. to the following lemma.

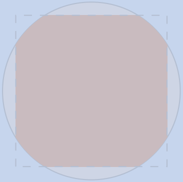
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{R}^n \to L_1$ s.t.

1.) $\forall i \in [n], \qquad \|f(e_i)\|_{L_1} = 1$

2.) $\forall \|a\|_{\ell_2} = 1, \qquad \|a\|_{\ell_\infty} < \frac{\varepsilon}{2} \Rightarrow \|f(a)\|_{L_1} < \sqrt{\frac{2}{\pi}} + \varepsilon.$

Subspace of L_1 whose unit ball looks like:



Embedding $\mathbb{R}^n$ into L_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little Groth. to the following lemma.

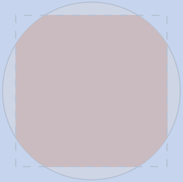
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{R}^n \to L_1$ s.t.

1.) $\forall i \in [n], \qquad \|f(e_i)\|_{L_1} = 1$

2.) $\forall \|a\|_{\ell_2} = 1, \qquad \|a\|_{\ell_\infty} < \frac{\varepsilon}{2} \Rightarrow \|f(a)\|_{L_1} < \sqrt{\frac{2}{\pi}} + \varepsilon.$

Subspace of L_1 whose unit ball looks like:



Embedding $\mathbb{R}^n$ into L_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little Groth. to the following lemma.

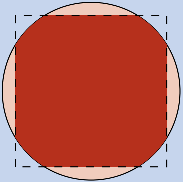
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{R}^n \to L_1$ s.t.

1.) $\forall i \in [n], \qquad \|f(e_i)\|_{L_1} = 1$

2.) $\forall \|a\|_{\ell_2} = 1, \qquad \|a\|_{\ell_\infty} < \frac{\varepsilon}{2} \;\Rightarrow\; \|f(a)\|_{L_1} < \sqrt{\frac{2}{\pi}} + \varepsilon.$

Subspace of L_1 whose unit ball looks like:



Embed using the Central Limit Theorem

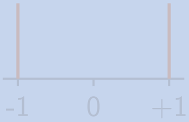
Proof of Lemma:

Define $f : \mathbb{R}^n \to L_1(\{-1,1\}^n)$ by

$$f(a)(Z_1, \dots, Z_n) = a_1 Z_1 + \cdots + a_n Z_n.$$

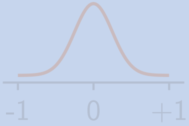
Then, $\quad \|f(a)\|_{L_1} = \mathbb{E}\big[|a_1 Z_1 + \cdots + a_n Z_n|\big]$

1.) $\|f(e_i)\|_{L_1} = 1$



2.) By the Central Limit Theorem (Berry–Esséen),

$$\|a\|_{\ell_\infty} < \tfrac{\varepsilon}{2} \;\;\Rightarrow\;\; \|f(a)\|_{L_1} < \sqrt{\tfrac{2}{\pi}} + \varepsilon$$



Embed using the Central Limit Theorem

Proof of Lemma:

Define $f : \mathbb{R}^n \to L_1(\{-1,1\}^n)$ by

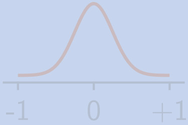
$$f(a)(Z_1, \ldots, Z_n) = a_1 Z_1 + \cdots + a_n Z_n.$$

Then, $$\|f(a)\|_{L_1} = \mathbb{E}\big[|a_1 Z_1 + \cdots + a_n Z_n|\big]$$

1.) $\|f(e_i)\|_{L_1} = 1$

2.) By the Central Limit Theorem (Berry–Esséen),

$$\|a\|_{\ell_\infty} < \tfrac{\varepsilon}{2} \;\Rightarrow\; \|f(a)\|_{L_1} < \sqrt{\tfrac{2}{\pi}} + \varepsilon$$



Embed using the Central Limit Theorem

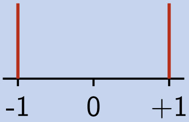
Proof of Lemma:

Define $f : \mathbb{R}^n \to L_1(\{-1,1\}^n)$ by

$$f(a)(Z_1, \ldots, Z_n) = a_1 Z_1 + \cdots + a_n Z_n.$$

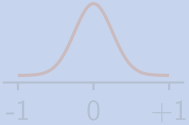
Then, $$\|f(a)\|_{L_1} = \mathbb{E}\big[|a_1 Z_1 + \cdots + a_n Z_n|\big]$$

1.) $\|f(e_i)\|_{L_1} = 1$



2.) By the Central Limit Theorem (Berry–Esséen),

$$\|a\|_{\ell_\infty} < \tfrac{\varepsilon}{2} \;\; \Rightarrow \;\; \|f(a)\|_{L_1} < \sqrt{\tfrac{2}{\pi}} + \varepsilon$$



Embed using the Central Limit Theorem

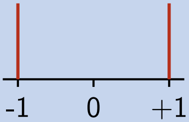
Proof of Lemma:

Define $f : \mathbb{R}^n \to L_1(\{-1,1\}^n)$ by

$$f(a)(Z_1,\ldots,Z_n) = a_1 Z_1 + \cdots + a_n Z_n.$$

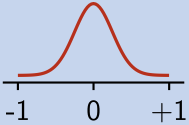
Then, $$\|f(a)\|_{L_1} = \mathbb{E}\big[|a_1 Z_1 + \cdots + a_n Z_n|\big]$$

1.) $\|f(e_i)\|_{L_1} = 1$



2.) By the Central Limit Theorem (Berry–Esséen),

$$\|a\|_{\ell_\infty} < \tfrac{\varepsilon}{2} \;\Rightarrow\; \|f(a)\|_{L_1} < \sqrt{\tfrac{2}{\pi}} + \varepsilon$$



The Little NCG

- The *Schatten-1 norm* $\|A\|_{S_1}$ of a matrix A is the *average of its singular values*.

The Little NCG.

Given a linear map $\mathcal{B} : \mathbb{C}^n \to S_1$, compute:

$$\|\mathcal{B}\|_{\ell_2 \to S_1} := \max_{\|x\|_{\ell_2}=1} \|\mathcal{B}x\|_{S_1}$$

The Little NCG

- The *Schatten-1 norm* $\|A\|_{S_1}$ of a matrix A is the *average of its singular values*.

The Little NCG.

Given a linear map $\mathcal{B} : \mathbb{C}^n \to S_1$, compute:

$$\|\mathcal{B}\|_{\ell_2 \to S_1} := \max_{\|x\|_{\ell_2}=1} \|\mathcal{B}x\|_{S_1}$$

Embedding $\mathbb{C}^n$ into S_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little NCG to the following lemma.

Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{C}^n \to S_1$ s.t.

$$1.)\ \forall i \in [n], \qquad \|f(e_i)\|_{S_1} = 1$$

$$2.)\ \forall \|a\|_{\ell_2} = 1, \qquad \|a\|_{\ell_\infty} < \tfrac{\varepsilon}{2} \Rightarrow \|f(a)\|_{S_1} < \tfrac{1}{\sqrt{2}} + \varepsilon.$$

Subspace of S_1 whose unit ball looks like:

Embedding $\mathbb{C}^n$ into S_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little NCG to the following lemma.

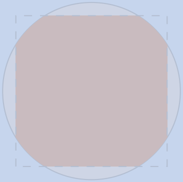
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{C}^n \to S_1$ s.t.

1.) $\forall i \in [n], \qquad \|f(e_i)\|_{S_1} = 1$

2.) $\forall \|a\|_{\ell_2} = 1, \qquad \|a\|_{\ell_\infty} < \frac{\varepsilon}{2} \Rightarrow \|f(a)\|_{S_1} < \frac{1}{\sqrt{2}} + \varepsilon.$

Subspace of S_1 whose unit ball looks like:



Embedding $\mathbb{C}^n$ into S_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little NCG to the following lemma.

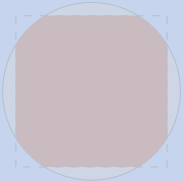
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{C}^n \to S_1$ s.t.

1.) $\forall i \in [n]$, $\quad \|f(e_i)\|_{S_1} = 1$

2.) $\forall \|a\|_{\ell_2} = 1$, $\quad \|a\|_{\ell_\infty} < \frac{\varepsilon}{2} \Rightarrow \|f(a)\|_{S_1} < \frac{1}{\sqrt{2}} + \varepsilon$.

Subspace of S_1 whose unit ball looks like:



Embedding $\mathbb{C}^n$ into S_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little NCG to the following lemma.

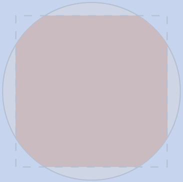
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{C}^n \to S_1$ s.t.

$$1.)\ \forall i \in [n], \qquad \|f(e_i)\|_{S_1} = 1$$

$$2.)\ \forall \|a\|_{\ell_2} = 1, \qquad \|a\|_{\ell_\infty} < \tfrac{\varepsilon}{2} \ \Rightarrow\ \|f(a)\|_{S_1} < \tfrac{1}{\sqrt{2}} + \varepsilon.$$

Subspace of S_1 whose unit ball looks like:



Embedding $\mathbb{C}^n$ into S_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little NCG to the following lemma.

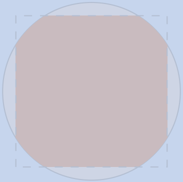
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{C}^n \to S_1$ s.t.

$$1.)\ \forall i \in [n], \qquad \|f(e_i)\|_{S_1} = 1$$

$$2.)\ \forall \|a\|_{\ell_2} = 1, \qquad \|a\|_{\ell_\infty} < \frac{\varepsilon}{2} \Rightarrow \|f(a)\|_{S_1} < \frac{1}{\sqrt{2}} + \varepsilon.$$

Subspace of S_1 whose unit ball looks like:



Embedding $\mathbb{C}^n$ into S_1

- Machinery of Guruswami–Raghavendra–Saket–Wu '12 reduces hardness of approx. Little NCG to the following lemma.

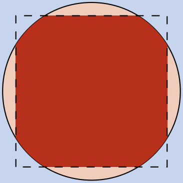
Lemma.

For every $n \in \mathbb{N}$ there exists a linear $f : \mathbb{C}^n \to S_1$ s.t.

$$1.)\ \forall i \in [n], \qquad \|f(e_i)\|_{S_1} = 1$$

$$2.)\ \forall \|a\|_{\ell_2} = 1, \qquad \|a\|_{\ell_\infty} < \tfrac{\varepsilon}{2} \Rightarrow \|f(a)\|_{S_1} < \tfrac{1}{\sqrt{2}} + \varepsilon.$$

Subspace of S_1 whose unit ball looks like:



Attempt 1: Random unitaries

- The matrix analogue of Bernoullis are random unitaries.
- So a natural guess for the function f would be

$$f(a)(U_1, \ldots, U_n) = a_1 U_1 + \cdots + a_n U_n.$$

- Then, $\|f(e_i)\|_{S_1} = 1$. ✓
- But if $a \in \mathbb{C}^n$ has no large coordinates then $f(a)$ is roughly a random Gaussian matrix G.
- By Wigner's semi-circle law, $\mathbb{E}[\|G\|_{S_1}] \approx 0.848 > 1/\sqrt{2}$. ✗

Attempt 1: Random unitaries

- The matrix analogue of Bernoullis are random unitaries.

- So a natural guess for the function f would be

$$f(a)(U_1, \ldots, U_n) = a_1 U_1 + \cdots + a_n U_n.$$

- Then, $\|f(e_i)\|_{S_1} = 1$. ✓

- But if $a \in \mathbb{C}^n$ has no large coordinates then $f(a)$ is roughly a random Gaussian matrix G.

- By Wigner's semi-circle law, $\mathbb{E}[\|G\|_{S_1}] \approx 0.848 > 1/\sqrt{2}$. ✗

Attempt 1: Random unitaries

- The matrix analogue of Bernoullis are random unitaries.
- So a natural guess for the function f would be

$$f(a)(U_1, \ldots, U_n) = a_1 U_1 + \cdots + a_n U_n.$$

- Then, $\|f(e_i)\|_{S_1} = 1$. ✓
- But if $a \in \mathbb{C}^n$ has no large coordinates then $f(a)$ is roughly a random Gaussian matrix G.
- By Wigner's semi-circle law, $\mathbb{E}[\|G\|_{S_1}] \approx 0.848 > 1/\sqrt{2}$. ✗

Attempt 1: Random unitaries

- The matrix analogue of Bernoullis are random unitaries.
- So a natural guess for the function f would be

$$f(a)(U_1, \ldots, U_n) = a_1 U_1 + \cdots + a_n U_n.$$

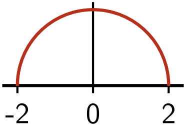
- Then, $\|f(e_i)\|_{S_1} = 1$. ✓
- But if $a \in \mathbb{C}^n$ has no large coordinates then $f(a)$ is roughly a random Gaussian matrix G.
- By Wigner's semi-circle law, $\mathbb{E}[\|G\|_{S_1}] \approx 0.848 > 1/\sqrt{2}$. ✗

Attempt 1: Random unitaries

- The matrix analogue of Bernoullis are random unitaries.
- So a natural guess for the function f would be

$$f(a)(U_1, \ldots, U_n) = a_1 U_1 + \cdots + a_n U_n.$$

- Then, $\|f(e_i)\|_{S_1} = 1$. ✓
- But if $a \in \mathbb{C}^n$ has no large coordinates then $f(a)$ is roughly a random Gaussian matrix G.
- By Wigner's semi-circle law, $\mathbb{E}[\|G\|_{S_1}] \approx 0.848 > 1/\sqrt{2}$. ✗



The Haagerup–Itoh example

- Integrality-gap examples often provide useful hints for hardness constructions.
- Haagerup–Itoh use an *isometric embedding* $f : \mathbb{C}^n \to S_1$ based on the generators of the C.A.R. algebra.
- Another dead end... But *anti-commutation* is the key.
- The generators of the Clifford algebra $C_1, \ldots, C_n \in \mathbb{C}^{N \times N}$ make everything works perfectly!
- $f(a) = a_1 C_1 + \cdots + a_n C_n$ is an *isometry from* $\mathbb{R}^n$ *to* S_1.
- But for complex vectors something *magic* happens!

The Haagerup–Itoh example

- Integrality-gap examples often provide useful hints for hardness constructions.
- Haagerup–Itoh use an *isometric embedding* $f : \mathbb{C}^n \to S_1$ based on the generators of the C.A.R. algebra.
- Another dead end... But *anti-commutation* is the key.
- The generators of the Clifford algebra $C_1, \ldots, C_n \in \mathbb{C}^{N \times N}$ make everything works perfectly!
- $f(a) = a_1 C_1 + \cdots + a_n C_n$ is an *isometry from* $\mathbb{R}^n$ *to* S_1.
- But for complex vectors something *magic* happens!

The Haagerup–Itoh example

- Integrality-gap examples often provide useful hints for hardness constructions.
- Haagerup–Itoh use an *isometric embedding* $f : \mathbb{C}^n \to S_1$ based on the generators of the C.A.R. algebra.
- Another dead end... But *anti-commutation* is the key.
- The generators of the Clifford algebra $C_1, \ldots, C_n \in \mathbb{C}^{N \times N}$ make everything works perfectly!
- $f(a) = a_1 C_1 + \cdots + a_n C_n$ is an *isometry from* $\mathbb{R}^n$ *to* S_1.
- But for complex vectors something *magic* happens!

The Haagerup–Itoh example

- Integrality-gap examples often provide useful hints for hardness constructions.
- Haagerup–Itoh use an *isometric embedding* $f : \mathbb{C}^n \to S_1$ based on the generators of the C.A.R. algebra.
- Another dead end. . . But *anti-commutation* is the key.
- The generators of the Clifford algebra $C_1, \ldots, C_n \in \mathbb{C}^{N \times N}$ make everything works perfectly!
- $f(a) = a_1 C_1 + \cdots + a_n C_n$ is an *isometry from* $\mathbb{R}^n$ *to* S_1.
- But for complex vectors something *magic* happens!

The Haagerup–Itoh example

- Integrality-gap examples often provide useful hints for hardness constructions.
- Haagerup–Itoh use an *isometric embedding* $f : \mathbb{C}^n \to S_1$ based on the generators of the C.A.R. algebra.
- Another dead end... But *anti-commutation* is the key.
- The generators of the Clifford algebra $C_1, \ldots, C_n \in \mathbb{C}^{N \times N}$ make everything works perfectly!
- $f(a) = a_1 C_1 + \cdots + a_n C_n$ is an *isometry from* $\mathbb{R}^n$ *to* S_1*.*
- But for complex vectors something *magic* happens!

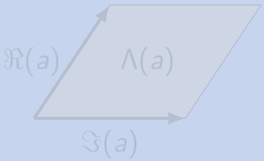
The Haagerup–Itoh example

- Integrality-gap examples often provide useful hints for hardness constructions.
- Haagerup–Itoh use an *isometric embedding* $f : \mathbb{C}^n \to S_1$ based on the generators of the C.A.R. algebra.
- Another dead end... But *anti-commutation* is the key.
- The generators of the Clifford algebra $C_1, \ldots, C_n \in \mathbb{C}^{N \times N}$ make everything works perfectly!
- $f(a) = a_1 C_1 + \cdots + a_n C_n$ is an *isometry from* $\mathbb{R}^n$ *to* S_1.
- But for complex vectors something *magic* happens!

The Magic Parallelogram Lemma

For $a \in \mathbb{C}^n$, the map $f(a) = a_1 C_1 + \cdots + a_n C_n$ satisfies

$$\|f(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



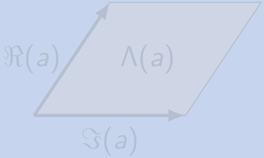
$\Lambda(e_i) = 0 \quad \Rightarrow \quad \|f(e_i)\|_{S_1} = 1$ ✓

$\forall a \in \mathbb{R}^n \ \Lambda(a) = 0 \quad \Rightarrow \quad \|f(a)\|_{S_1} = \|a\|_{\ell_2}$ ✗

The Magic Parallelogram Lemma

For $a \in \mathbb{C}^n$, the map $f(a) = a_1 C_1 + \cdots + a_n C_n$ satisfies

$$\|f(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



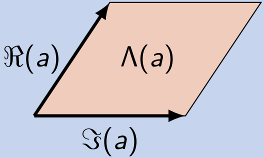
$\Lambda(e_i) = 0 \quad \Rightarrow \quad \|f(e_i)\|_{S_1} = 1$ ✓

$\forall a \in \mathbb{R}^n \ \Lambda(a) = 0 \quad \Rightarrow \quad \|f(a)\|_{S_1} = \|a\|_{\ell_2}$ ✗

The Magic Parallelogram Lemma

For $a \in \mathbb{C}^n$, the map $f(a) = a_1 C_1 + \cdots + a_n C_n$ satisfies

$$\|f(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



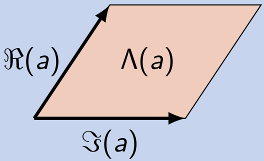
$\Lambda(e_i) = 0 \quad \Rightarrow \quad \|f(e_i)\|_{S_1} = 1$ ✓

$\forall a \in \mathbb{R}^n \;\; \Lambda(a) = 0 \quad \Rightarrow \quad \|f(a)\|_{S_1} = \|a\|_{\ell_2}$ ✗

The Magic Parallelogram Lemma

For $a \in \mathbb{C}^n$, the map $f(a) = a_1 C_1 + \cdots + a_n C_n$ satisfies

$$\|f(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



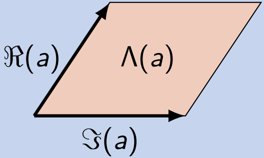
$\Lambda(e_i) = 0 \quad \Rightarrow \quad \|f(e_i)\|_{S_1} = 1$ ✓

$\forall a \in \mathbb{R}^n \ \Lambda(a) = 0 \quad \Rightarrow \quad \|f(a)\|_{S_1} = \|a\|_{\ell_2}$ ✗

The Magic Parallelogram Lemma

For $a \in \mathbb{C}^n$, the map $f(a) = a_1 C_1 + \cdots + a_n C_n$ satisfies

$$\|f(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



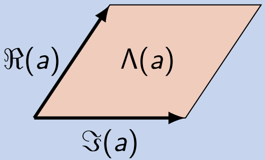
$\Lambda(e_i) = 0 \quad \Rightarrow \quad \|f(e_i)\|_{S_1} = 1$ ✓

$\forall a \in \mathbb{R}^n \ \Lambda(a) = 0 \quad \Rightarrow \quad \|f(a)\|_{S_1} = \|a\|_{\ell_2}$ ✗

The Magic Parallelogram Lemma

For $a \in \mathbb{C}^n$, the map $f(a) = a_1 C_1 + \cdots + a_n C_n$ satisfies

$$\|f(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



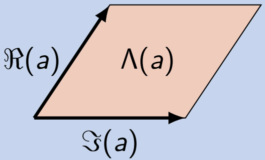
$\Lambda(e_i) = 0 \quad \Rightarrow \quad \|f(e_i)\|_{S_1} = 1$ ✓

$\forall a \in \mathbb{R}^n \; \Lambda(a) = 0 \quad \Rightarrow \quad \|f(a)\|_{S_1} = \|a\|_{\ell_2}$ ✗

The Magic Parallelogram Lemma

For $a \in \mathbb{C}^n$, the map $f(a) = a_1 C_1 + \cdots + a_n C_n$ satisfies

$$\|f(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$

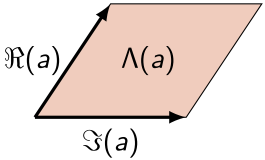


$\Lambda(e_i) = 0 \quad \Rightarrow \quad \|f(e_i)\|_{S_1} = 1$ ✓

$\forall a \in \mathbb{R}^n \ \Lambda(a) = 0 \quad \Rightarrow \quad \|f(a)\|_{S_1} = \|a\|_{\ell_2}$ ✗

Final step: random phases

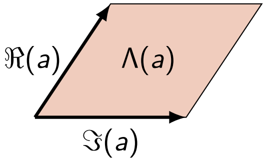
$$\|f(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



- $\omega_1, \ldots, \omega_n \sim \{1, i, -1, -i\}$.
- Then $\Lambda(\omega_i e_i) = 0 \Rightarrow \|f(\omega_i e_i)\|_{S_1} = 1$.
- $a = (\frac{\pm 1}{\sqrt{n}}, \ldots, \frac{\pm 1}{\sqrt{n}}) \Rightarrow \mathbb{E}\Lambda(\omega \circ a) = \frac{1}{2} \Rightarrow \mathbb{E}\|f(\omega \circ a)\|_{S_1} = \frac{1}{\sqrt{2}}$.
- Final function: $f(a) = \bigoplus_\omega \left(a_1\omega_1 C_1 + \cdots + a_n\omega_n C_n\right)$.

Final step: random phases

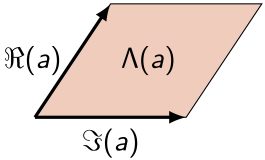
$$\|f(a)\|_{S_1} = \tfrac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \tfrac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



- $\omega_1, \ldots, \omega_n \sim \{1, i, -1, -i\}$.
- Then $\Lambda(\omega_i e_i) = 0 \Rightarrow \|f(\omega_i e_i)\|_{S_1} = 1$.
- $a = (\frac{\pm 1}{\sqrt{n}}, \ldots, \frac{\pm 1}{\sqrt{n}}) \Rightarrow \mathbb{E}\Lambda(\omega \circ a) = \frac{1}{2} \Rightarrow \mathbb{E}\|f(\omega \circ a)\|_{S_1} = \frac{1}{\sqrt{2}}$.
- Final function: $f(a) = \bigoplus_\omega \left(a_1 \omega_1 C_1 + \cdots + a_n \omega_n C_n\right)$.

Final step: random phases

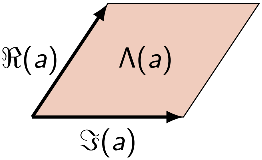
$$\|f(a)\|_{S_1} = \tfrac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \tfrac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



- $\omega_1, \ldots, \omega_n \sim \{1, i, -1, -i\}$.
- Then $\Lambda(\omega_i e_i) = 0 \Rightarrow \|f(\omega_i e_i)\|_{S_1} = 1$.
- $a = (\frac{\pm 1}{\sqrt{n}}, \ldots, \frac{\pm 1}{\sqrt{n}}) \Rightarrow \mathbb{E}\Lambda(\omega \circ a) = \frac{1}{2} \Rightarrow \mathbb{E}\|f(\omega \circ a)\|_{S_1} = \frac{1}{\sqrt{2}}$.
- Final function: $f(a) = \bigoplus_\omega \left(a_1\omega_1 C_1 + \cdots + a_n\omega_n C_n\right)$.

Final step: random phases

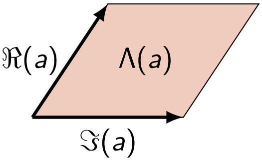
$$\|f(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



- $\omega_1, \ldots, \omega_n \sim \{1, i, -1, -i\}$.
- Then $\Lambda(\omega_i e_i) = 0 \Rightarrow \|f(\omega_i e_i)\|_{S_1} = 1$.
- $a = (\frac{\pm 1}{\sqrt{n}}, \ldots, \frac{\pm 1}{\sqrt{n}}) \Rightarrow \mathbb{E}\Lambda(\omega \circ a) = \frac{1}{2} \Rightarrow \mathbb{E}\|f(\omega \circ a)\|_{S_1} = \frac{1}{\sqrt{2}}$.
- Final function: $f(a) = \bigoplus_\omega \left(a_1\omega_1 C_1 + \cdots + a_n\omega_n C_n\right)$.

Final step: random phases

$$\|f(a)\|_{S_1} = \tfrac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \tfrac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}$$



- $\omega_1, \ldots, \omega_n \sim \{1, i, -1, -i\}$.
- Then $\Lambda(\omega_i e_i) = 0 \Rightarrow \|f(\omega_i e_i)\|_{S_1} = 1$.
- $a = (\frac{\pm 1}{\sqrt{n}}, \ldots, \frac{\pm 1}{\sqrt{n}}) \Rightarrow \mathbb{E}\Lambda(\omega \circ a) = \frac{1}{2} \Rightarrow \mathbb{E}\|f(\omega \circ a)\|_{S_1} = \frac{1}{\sqrt{2}}$.
- Final function: $f(a) = \bigoplus_\omega \left(a_1\omega_1 C_1 + \cdots + a_n\omega_n C_n\right)$.

Summary / Open questions

Main Theorem (optimal hardness for the NCG)

For any $\varepsilon > 0$, given a tensor $T \in \mathbb{R}^{d\times d\times d\times d}$, it is NP-hard to approximate the value

$$\mathrm{OPT}(T) = \max_{\substack{A,B\in\mathbb{C}^{d\times d}\\ \text{unitary}}} \Big| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \Big|$$

to within a factor $2-\varepsilon$. Moreover, this is optimal.

- NP-hardness of K_G-approx. the Groth. problem?
- Close gaps for real and Hermitian variants of the NCG.
- Explore "operator space" variants of the Groth. Ineq.

Summary / Open questions

Main Theorem (optimal hardness for the NCG)

For any $\varepsilon > 0$, given a tensor $T \in \mathbb{R}^{d\times d\times d\times d}$, it is NP-hard to approximate the value

$$\mathrm{OPT}(T) = \max_{\substack{A,B\in\mathbb{C}^{d\times d}\\ \text{unitary}}} \left| \sum T_{ijkl} A_{ij} \overline{B_{kl}} \right|$$

to within a factor $2-\varepsilon$. Moreover, this is optimal.

- NP-hardness of K_G-approx. the Groth. problem?
- Close gaps for real and Hermitian variants of the NCG.
- Explore "operator space" variants of the Groth. Ineq.