Grothendieck inequalities for SDPs with rank constraint

Jop Briët (CWI)

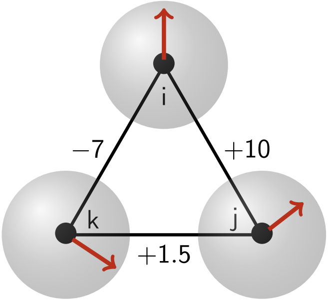
Joint work with **Fernando de Oliveira Filho (FU Berlin)**
and **Frank Vallentin (TU Delft)**

Phenomena in high dimensions in geometric analysis,
random matrices and computational geometry

Roscoff, June 26, 2012

The Grothendieck problem

- Given a weighted graph and dimension d, find an optimal configuration of unit vectors in $\mathbb{R}^d$ associated to the vertices



$$10\, x_i \cdot x_j - 7\, x_i \cdot x_k + 1.5\, x_j \cdot x_k$$

- An optimal configuration maximizes the weighted sum of dot products between the vectors on the edges

$$\sum_{ij \in E} W_{ij}\, x_i \cdot x_j$$

Example 1: Ground-state energy of a spin glass

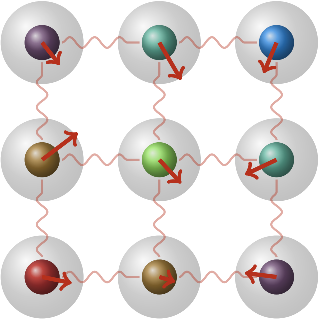
- Model of interacting particles introduced by Stanley (1968)
- Particles form the vertices of a weighted interaction graph

Particles are unit vectors $x_i \in \mathbb{R}^d$

d=1: Ising model $x_i \in \{-1, 1\}$

d=2: planar model

d=3: Heisenberg model



Edge weights W_{ij} give their interaction type and strength

The ground-state energy is given by

$$\min \sum_{ij \in E} W_{ij} \, x_i \cdot x_j$$

Example 2: XOR games

- Abstract model of the simplest physics experiments used to test predictions of quantum mechanics
- $W \in \{-1,1\}^{I \times J}$, probability distribution π over $I \times J$

$i \in I$ $\qquad\qquad$ $j \in J$

random questions $(i,j) \sim \pi$

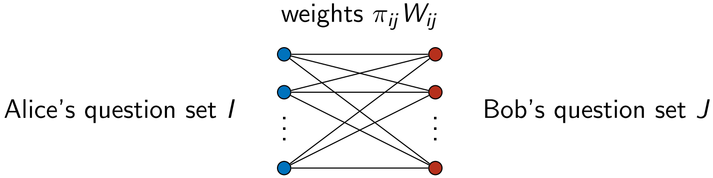
non-communicating players

answers

$a_i = \pm 1$ $\qquad\qquad$ $b_j = \pm 1$

- The players win if $W_{ij} a_i b_j = +1$ and lose if $W_{ij} a_i b_j = -1$
- The goal: Maximize the bias $\mathbb{E}[W_{ij} a_i b_j]$

XOR games are bipartite instances



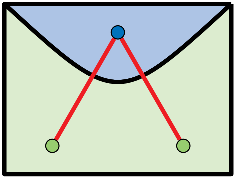
- **(Tsirelson).** For any quantum strategy there exist unit vectors x_i, y_j such that for all i, j

$$\mathbb{E}[a_i b_j] = x_i \cdot y_j$$

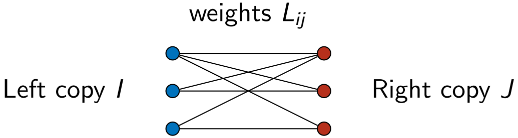
- The bias is of the form $\sum_{ij} \pi_{ij} W_{ij} x_i \cdot y_j$
- Dimension of the vectors is related to the quantum systems used in the strategy

Example 3: The maximum cut problem

- *Given a graph, partition its vertices such that the number of edges crossing the cut is maximized*



- Combinatorial ($d = 1$) instance of the Grothendieck problem



- Max $\sum_{i \in I, j \in J} L_{ij} x_i y_j$ over $x_i, y_j \in \{-1, 1\}$ gives a max cut

Computational perspective

- If P≠NP, then MaxCut cannot be solved in polynomial time
- Hence, worst-case instances of the Grothendieck problem are likely to be intractable

Then, can we efficiently find d-dimensional unit vectors that achieve a value very close *to the optimum?*

- **(Håstad)** In general, no! If P≠NP, then MaxCut cannot even be *approximated* to within a factor $16/17 - \varepsilon$ in poly-time
- But if the required dimension is greater than the number of vectors, we *can*!
- Since N vectors span a space of dimension at most N, there is no effective constraint on the dimension

SDP formulation of the Grothendieck problem

- A matrix $X \in \mathbb{R}^{N \times N}$ is positive semidefinite iff there exist vectors $x_1, \ldots, x_N$ such that $X_{ij} = x_i \cdot x_j$

- The Grothendieck problem as a semidefinite program (SDP) with rank constraint:

Maximize	$\sum_{ij \in E} W_{ij} X_{ij}$
Subject to	X is positive semidefinite
	$X_{ii} = 1$ for all vertices i
	~~rank $X \leq d$~~

- Without the rank constraint, this problem can be solved (approximated arbitrarily well) in poly-time

SDP-based approximation algorithms

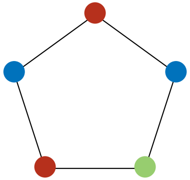
- To approximate the optimum value of the Grothendieck problem, drop the rank constraint and solve the SDP

Def. The maximum ratio of this optimum and the SDP value, $K(d, G)$, is the rank-d Grothendieck constant of the graph G

$$K(d, G) = \sup_{W} \frac{\mathrm{SDP}_{\infty}(W, G)}{\mathrm{SDP}_{d}(W, G)}$$

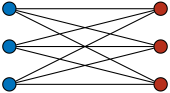
- How large $K(d, G)$ is highly depends on the *chromatic number* $\chi(G)$ of the graph G

Def. $\chi(G) =$ the least number of colors with which the vertices can be colored, such that adjacent vertices get different colors



$\chi(C_5) = 3$

Previous results for bipartite graphs



- **(Grothendieck, Alon-Naor).** $\exists$ universal constant K_G such that $K(1, G) \leq K_G$ for all bipartite graphs G
- **(Braverman-Makarychev-Makarychev-Naor).** $K_G < \frac{\pi}{2\ln(1+\sqrt{2})} = 1.782\ldots$
- **(Goemans-Williamson).** For MaxCut this upper bound can be improved to $1/0.878\ldots = 1.138\ldots$
- **(Davie-Reeds).** $K_G \geq 1.676\ldots$
- **(Raghavendra-Steurer).** Assuming the Unique Games Conjecture, K_G is the poly-time approximation threshold

Previous results for general graphs

- **(Alon-Makarychev-Makarychev-Naor).** $K(1, G) \leq O\big(\log \vartheta(\overline{G})\big)$

$$\vartheta(\overline{G}) = \text{Min} \quad \lambda$$
$$\text{Subject to} \quad Z_{ii} = \lambda - 1 \text{ and positive semidefinite}$$
$$Z_{ij} = -1 \text{ for all edges } ij$$

- The Lovász theta number satisfies $\vartheta(\overline{G}) \leq \chi(G)$

- **(AMMN, Arora-Berger-Hazan-Kindler-Safra, Khot-O'Donnell).** $K(1, K_N) = \Theta(\log N)$

- **(Alon-Berger).** $K\big(1, G(N, p)\big) = \Theta(\log Np)$ almost surely.

Our main results

- **(B-de Oliveira Filho-Vallentin).** New upper bounds on $K(1, G)$ for graphs with small chromatic number

$$K(1, G) \leq \frac{\pi}{2 \operatorname{arcsinh}\left((\vartheta(\overline{G}) - 1)^{-1}\right)}$$

- This bound grows linearly with $\vartheta(\overline{G})$, but is favorable over the $O\big(\log \vartheta(\overline{G})\big)$ when $\vartheta(\overline{G})$ is small

Our main results

- For higher dimensions, we give numeric upper bounds

d	*bipartite G*	*tripartite G*
1	$1.782213\ldots$	$3.264251\ldots$
2	$1.404909\ldots$	$2.621596\ldots$
3	$1.280812\ldots$	$2.412700\ldots$
4	$1.216786\ldots$	$2.309224\ldots$
5	$1.177179\ldots$	$2.247399\ldots$
6	$1.150060\ldots$	$2.206258\ldots$
7	$1.130249\ldots$	$2.176891\ldots$
8	$1.115110\ldots$	$2.154868\ldots$
9	$1.103150\ldots$	$2.137736\ldots$
10	$1.093456\ldots$	$2.124024\ldots$

- **(Naor-Regev).** The first column goes as $1 + O(1/d)$
- **(B-Buhrman-Toner).** $1 + \Omega(1/d)$ lower bound for bipartite graphs
- So our first column is asymptotically optimal

High chromatic numbers

- **(B-de Oliveira Filho-Vallentin).** If the graph has large chromatic number the method of [AMMN] gives the better bound

$$K(d, G) \leq O\left(\frac{\log \vartheta(\overline{G})}{d}\right)$$

Proof outline

The Grothendieck problem

Maximize	$\sum_{ij \in E} W_{ij} x_i \cdot x_j$
Subject to	$x_1, \ldots, x_N \in \mathbb{R}^d$
	$\|x_i\|_2 = 1$ for all vertices $i \in [N]$

- If we let $\mathrm{SDP}_d(W, G)$ be the optimum, then

$$K(d, G) = \sup_W \frac{\mathrm{SDP}_N(W, G)}{\mathrm{SDP}_d(W, G)}$$

- To upper bound $K(d, G)$, we convert N-dimensional SDP vectors x_i into random d-dimensional unit vectors y_i such that

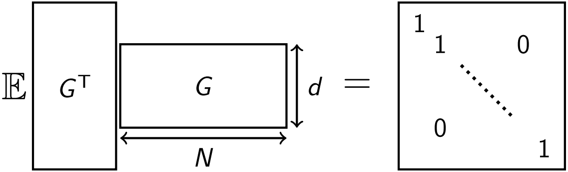
$$\mathbb{E}[y_i \cdot y_j] = c\, x_i \cdot x_j$$

  for all adjacent vertices i and j, for some $c = c(G)$

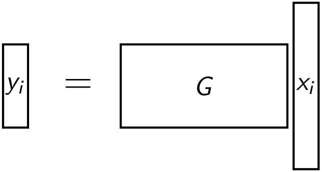
- ...implying $K(d, G) \leq 1/c$ by the averaging principle

Gaussian dimension reduction

- For a random $d \times N$ matrix G with i.i.d. $\mathcal{N}(0, 1/d)$ entries,



- If we use G to map the N-dimensional SDP vectors $x_1, \ldots, x_N$ to d-dimensional vectors

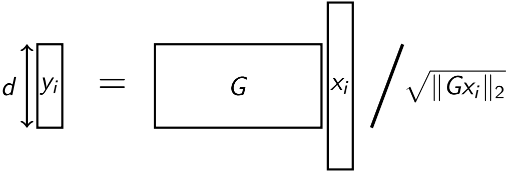


- then $\mathbb{E}[y_i \cdot y_j] = x_i^\mathsf{T}\, \mathbb{E}[G^\mathsf{T} G]\, x_j = x_i \cdot x_j$ for all vertices i, j

- Great! But these new vectors may not have unit norm. . .

Can't we just normalize the new vectors?

- Suppose we normalize the new vectors



- Gives unit vectors of the desired dimension, but their expected dot products depends non-linearly on the original dot products

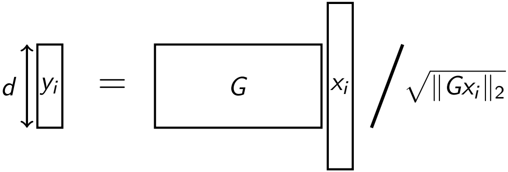
- For example, if $d = 1$ we have **Grothendieck's identity**:

$$\mathbb{E}[y_i \cdot y_j] = \frac{2}{\pi} \arcsin(x_i \cdot x_j)$$

- If the SDP value is high, the new vectors can give a low (expected) value due to non-linear scaling of the dot products

A generalized Grothendieck identity

Lemma (B-de Oliveira Filho-Vallentin). For



we have $\mathbb{E}[y_i \cdot y_j] = E_d(x_i \cdot x_j)$, where

$$E_d(t) = \sum_{k=0}^{\infty} \underbrace{\frac{2}{d}\left(\frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\right)^2 \prod_{\ell=1}^{k} \frac{(2\ell-1)^2}{2\ell(d-2\ell)}}_{\text{Taylor coefficients}} \, t^{2k+1}$$

- Grothendieck's identity: $E_1(t) = (2/\pi) \arcsin t$

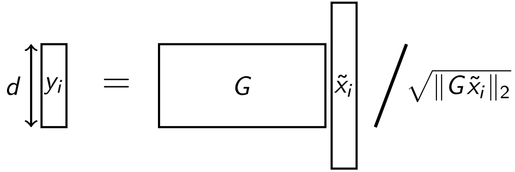
Krivine's method: Preprocess before embedding

- Main idea: Preprocess the SDP vectors $x_i \mapsto \tilde{x}_i$ so that

$$E_d(\tilde{x}_i \cdot \tilde{x}_j) = c\, x_i \cdot x_j$$

for all adjacent vertices i and j, and some $c = c(G)$

- Then, rounding the preprocessed vectors as before



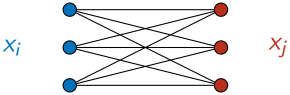
gives d-dimensional unit vectors with expected dot product

$$\mathbb{E}[y_i \cdot y_j] = E_d(\tilde{x}_i \cdot \tilde{x}_j) = c\, x_i \cdot x_j$$

(by our lemma)

Preprocessing for bipartite graphs

- To preprocess the SDP vectors for a bipartite graph, we group them according to a 2-coloring



- It suffices to find maps BLUE, RED that send unit vectors to unit vectors and satisfy

$$E_d\big(\mathrm{BLUE}(x_i)\cdot \mathrm{RED}(x_j)\big) = c\, x_i \cdot x_j$$

  for some $c > 0$

The maps BLUE and RED

- Take the Taylor coefficients $a_0, a_1, a_2, \ldots$ of the *inverse* E_d^{-1} of E_d (can be computed from those of E_d)
- By the tensor-power trick: $x^{\otimes k} \cdot y^{\otimes k} = (x \cdot y)^k$, the vectors

$$\begin{bmatrix} \sqrt{|a_0|} \\ \sqrt{|a_1|} c x_i \\ \sqrt{|a_2|} (c x_i)^{\otimes 2} \\ \vdots \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \operatorname{sgn}(a_0)\sqrt{|a_0|} \\ \operatorname{sgn}(a_1)\sqrt{|a_1|} c x_j \\ \operatorname{sgn}(a_2)\sqrt{|a_2|} (c x_j)^{\otimes 2} \\ \vdots \end{bmatrix}$$

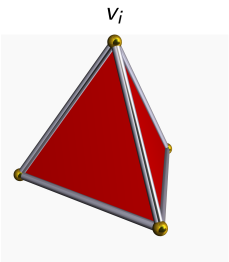
  have dot product $a_0 + a_1(c\, x_i \cdot x_j) + a_2(c\, x_i \cdot x_j)^2 + \cdots$
- This equals $E_d^{-1}(c\, x_i \cdot x_j)$ since the Taylor series converges
- $E_d\big(\mathrm{BLUE}(x_i) \cdot \mathrm{RED}(x_j)\big) = E_d\left(E_d^{-1}(c\, x_i \cdot x_j)\right) = c\, x_i \cdot x_j$
- Set c such that $\mathrm{BLUE}(x_i)$ and $\mathrm{RED}(x_j)$ are unit vectors

Dealing with larger chromatic numbers

- For graphs with chromatic number larger than 2, we preprocess using linear combinations of the BLUE and RED maps

$$x_i \mapsto \text{BLUE}(x_i) \otimes s_i + \text{RED}(x_i) \otimes t_i$$

- The vectors s_i are t_i are "orthogonal tails"
- Can be obtained from a feasible solution for $\vartheta(\overline{G})$
- For example if G is k-colorable, we can use vertices of a $(k-1)$-dimensional simplex of radius $\sqrt{k-1}$

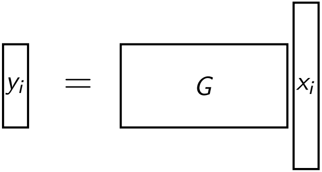


$$t_i = (v_i, 0, 1)$$

$$s_i = (v_i, 1, 0)$$

A better method for very large chromatic numbers

- The main problem was to turn the d-dimensional vectors



  into unit vectors

- However, $\|y_i\|_2$ is large with only tiny probability
- The [AMMN] method:
  - Scale y_i to y_i/R if y_i lies in the ball of radius R
  - Set $y_i = 0$ otherwise
- Optimizing over R shows that there exist vectors in the unit ball giving value $1/R^2 \approx d/\log \vartheta(G)$ times the SDP value

Open problems and future/current work

- The Krivine and AMMN methods are favorable for small/large chromatic number resp.
- *Is there some hybrid scheme of these two?*

- Extend the known $\Omega\big(\log \omega(G)\big)$ lower bound on $K(1, G)$ [AMMN] to the higher-dimensional cases

Thank you!

arXiv:1011.1754 [math.OC]