THEORY OF COMPUTING, Volume 10 (4), 2014, pp. 77–105
www.theoryofcomputing.org

Grothendieck Inequalities for Semidefinite Programs with Rank Constraint

Jop Briët* Fernando Mário de Oliveira Filho† Frank Vallentin‡

Received May 4, 2012; Revised January 21, 2014; Published May 31, 2014

Abstract: Grothendieck inequalities are fundamental inequalities which are frequently used in many areas of mathematics and computer science. They can be interpreted as upper bounds for the integrality gap between two optimization problems: a difficult semidefinite program with rank-1 constraint and its easy semidefinite relaxation where the rank constraint is dropped. For instance, the integrality gap of the Goemans-Williamson approximation algorithm for MAX CUT can be seen as a Grothendieck inequality. In this paper we consider Grothendieck inequalities for ranks greater than 1 and we give two applications: approximating ground states in the n-vector model in statistical mechanics and XOR games in quantum information theory.

ACM Classification: G.1.6, G.2.2

AMS Classification: 68W25, 90C22

Key words and phrases: randomized rounding, Grothendieck inequality, n-vector model, XOR games

1 Introduction

Let $G = (V, E)$ be a graph with finite vertex set V and edge set $E \subseteq \binom{V}{2}$. Let $A\colon V \times V \to \mathbb{R}$ be a symmetric matrix whose rows and columns are indexed by the vertex set of G, and r be a positive integer. The

*Supported by Vici grant 639.023.302 from the Netherlands Organization for Scientific Research (NWO), by the European Commission under the Integrated Project Qubit Applications (QAP) funded by the IST directorate as Contract Number 015848, by the Dutch BSIK/BRICKS project and by the European grant QCS.

†Supported by NWO Rubicon grant 680-50-1014.

‡Supported by Vidi grant 639.032.917 from the Netherlands Organization for Scientific Research (NWO).

© 2014 Jop Briët, Fernando Mário de Oliveira Filho, and Frank Vallentin
Licensed under a Creative Commons Attribution License (CC-BY)
DOI: 10.4086/toc.2014.v010a004

graphical Grothendieck problem with rank-r constraint is the following optimization problem:

$$\mathrm{SDP}_r(G,A) = \max\left\{ \sum_{\{u,v\}\in E} A(u,v) f(u)\cdot f(v) \;:\; f\colon V \to S^{r-1} \right\},$$

where $S^{r-1} = \{x \in \mathbb{R}^r : x \cdot x = 1\}$ is the $(r-1)$-dimensional unit sphere. The *rank-r Grothendieck constant of the graph G* is the smallest constant $K(r,G)$ so that for all symmetric matrices $A\colon V\times V \to \mathbb{R}$ the following inequality holds:

$$\mathrm{SDP}_\infty(G,A) \le K(r,G)\,\mathrm{SDP}_r(G,A)\,. \tag{1.1}$$

Here S^∞ denotes the unit sphere of the Hilbert space $l^2(\mathbb{R})$ of square summable sequences, which contains $\mathbb{R}^n$ as the subspace of the first n components. It is easy to see that $K(r,G) \ge 1$. In this paper, we prove new upper bounds for $K(r,G)$.

1.1 Some history

Inequality (1.1) is called a *Grothendieck inequality* because it first appeared in the work [23] of Grothendieck on the metric theory of tensor products. More precisely, Grothendieck considered the case $r = 1$ for bipartite graphs, although in quite a different language. Grothendieck proved that in this case $K(1,G)$ is upper bounded by a constant that is independent of the size of G.

Later, Lindenstrauss and Pełczyński [34] reformulated Grothendieck's inequality for bipartite graphs in a way that is very close to the formulation we gave above. The graphical Grothendieck problem with rank-1 constraint was introduced by Alon, Makarychev, Makarychev, and Naor [3]. Haagerup [24] considered the complex case of Grothendieck's inequality; his upper bound is also valid for the real case $r = 2$. The higher rank case for bipartite graphs was introduced by Briët, Buhrman, and Toner [11].

1.2 Computational perspective

There has been a recent surge of interest in Grothendieck inequalities by the computer science community. The problem $\mathrm{SDP}_r(G,A)$ is a semidefinite maximization problem with rank-r constraint:

$$\begin{aligned}\mathrm{SDP}_r(G,A) = \max\Bigg\{ \sum_{\{u,v\}\in E} A(u,v)X(u,v) \;:\; & X \in \mathbb{R}^{V\times V}_{\succeq 0},\\ & X(u,u) = 1 \text{ for all } u \in V,\\ & \operatorname{rank} X \le r \Bigg\},\end{aligned}$$

where $\mathbb{R}^{V\times V}_{\succeq 0}$ is the set of matrices $X\colon V\times V \to \mathbb{R}$ that are positive semidefinite.

On the one hand, $\mathrm{SDP}_r(G,A)$ is generally a difficult computational problem. For instance, if $r = 1$ and G is the complete bipartite graph $K_{n,n}$ on $2n$ nodes, and if A is the Laplacian matrix of a graph G' on n nodes, then computing $\mathrm{SDP}_1(K_{n,n},A)$ is equivalent to computing the weight of a maximum cut of G'. The maximum cut problem (MAX CUT) is one of Karp's 21 NP-complete problems. On the other hand, if we relax the rank-r constraint, then we deal with $\mathrm{SDP}_\infty(G,A)$, which is an easy computational problem: Obviously, one has $\mathrm{SDP}_\infty(G,A) = \mathrm{SDP}_{|V|}(G,A)$ and computing $\mathrm{SDP}_{|V|}(G,A)$ amounts to solving a semidefinite programming problem (see, e. g., Vandenberghe, Boyd [47]). Therefore one may approximate it to any fixed precision in polynomial time by using the ellipsoid method or interior point algorithms.

In many cases the optimal constant $K(r,G)$ is not known and so one is interested in finding upper bounds for $K(r,G)$. Usually, proving an upper bound amounts to giving a randomized polynomial-time approximation algorithm for $\mathrm{SDP}_r(G,A)$. In the case of the MAX CUT problem, Goemans and Williamson [22] pioneered an approach based on randomized rounding: One rounds an optimal solution of $\mathrm{SDP}_\infty(G,A)$ to a feasible solution of $\mathrm{SDP}_r(G,A)$. The expected value of the rounded solution is then related to the one of the original solution, and this gives an upper bound for $K(r,G)$. Using this basic idea, Goemans and Williamson [22] showed that for any matrix $A\colon V \times V \to \mathbb{R}$ that is a Laplacian matrix of a weighted graph with nonnegative edge weights one has

$$\mathrm{SDP}_\infty(K_{n,n},A) \le (0.878\ldots)^{-1}\,\mathrm{SDP}_1(K_{n,n},A)\,.$$

1.3 Applications and references

Grothendieck's inequality is a fundamental inequality in the theory of Banach spaces. Many books on the geometry of Banach spaces contain a substantial treatment of the result. We refer for instance to the books by Pisier [41], Jameson [26], and Garling [21].

During the last years, especially after Alon and Naor [4] pointed out the connection between the inequality and approximation algorithms using semidefinite programs, Grothendieck's inequality has also become a unifying and fundamental tool outside of functional analysis.

It has applications in optimization (Nesterov [40], Nemirovski, Roos, Terlaky [39], Megretski [37]), extremal combinatorics (Alon, Naor [4]), system theory (Ben-Tal, Nemirovski [9]), machine learning (Charikar, Wirth [14], Khot, Naor [28, 29]), communication complexity (Linial, Shraibman [35]), quantum information theory (Tsirel'son [46], Regev, Toner [44]), and computational complexity (Khot, O'Donnell [30], Arora, Berger, Kindler, Safra, Hazan [6], Khot and Naor [27], Raghavendra, Steurer [42]).

The references above mainly deal with the combinatorial rank $r = 1$ case, when $S^0 = \{-1,+1\}$. For applications in quantum information (Briët, Buhrman, Toner [11]) and in statistical mechanics (mentioned in Alon, Makarychev, Makarychev, Naor [3], Kindler, Naor, Schechtman [31]) the more geometrical case when $r > 1$ is of interest—this case is the subject of this paper.

In statistical mechanics, the problem of computing $\mathrm{SDP}_n(G,A)$ is known as finding ground-states of the n-vector model. Introduced by Stanley [45], the *n-vector model*[1] describes the interaction of particles in a spin glass with ferromagnetic and anti-ferromagnetic interactions.

Let $G = (V,E)$ be the interaction graph where the vertices are particles and where edges indicate which particles interact. The potential function $A\colon V \times V \to \mathbb{R}$ is 0 if u and v are not adjacent, it is positive if there is ferromagnetic interaction between u and v, and it is negative if there is anti-ferromagnetic interaction. The particles possess a vector-valued spin $f\colon V \to S^{n-1}$. In the absence of an external field,

[1] The case $n = 1$ is known as the Ising model, the case $n = 2$ as the XY model, the case $n = 3$ as the Heisenberg model, and the case $n = \infty$ as the Berlin-Kac spherical model.

the total energy of the system is given by the *Hamiltonian*

$$H(f) = -\sum_{\{u,v\}\in E} A(u,v) f(u)\cdot f(v)\,.$$

The ground state of this model is a configuration of spins $f\colon V \to S^{n-1}$ which minimizes the total energy, so finding the ground state is the same as solving $\mathrm{SDP}_n(G,A)$. Typically, the interaction graph has small chromatic number, e. g., the most common case is when G is a finite subgraph of the integer lattice $\mathbb{Z}^n$ where the vertices are the lattice points and where two vertices are connected if their Euclidean distance is one. These graphs are bipartite since they can be partitioned into even and odd vertices, corresponding to the parity of the sum of the coordinates.

We also briefly describe the connection to quantum information theory. In an influential paper, Einstein, Podolsky, and Rosen [18] pointed out an anomaly of quantum mechanics that allows spatially separated parties to establish peculiar correlations by each performing measurements on a private quantum system: *entanglement*. Later, Bell [8] proved that local measurements on a pair of spatially separated, entangled quantum systems, can give rise to joint probability distributions of measurement outcomes that violate certain inequalities (now called Bell inequalities), satisfied by any classical distribution. Experimental results of Aspect, Grangier, and Roger [7] give strong evidence that nature indeed allows distant physical systems to be correlated in such non-classical ways.

XOR games, first formalized by Cleve, Høyer, Toner, and Watrous [15], constitute the simplest model in which entanglement can be studied quantitatively. In an XOR game, two players, Alice and Bob, receive questions u and v (resp.) that are picked by a referee according to some probability distribution $\pi(u,v)$ known to everybody in advance. Without sharing their questions, the players have to answer the referee with bits a and b (resp.), and win the game if and only if the exclusive-OR of their answers $a\oplus b$ equals the value of a Boolean function $g(u,v)$; the function g is also known in advance to all three parties.

In a quantum-mechanical setting, the players determine their answers by performing measurements on their shares of a pair of entangled quantum systems. A *state* of a pair of d-dimensional quantum systems is a trace-1 positive semidefinite operator $\rho \in \mathbb{C}^{d^2\times d^2}_{\succeq 0}$. The systems are *entangled* if ρ cannot be written as a convex combination of tensor products of d-by-d positive semidefinite matrices. For each question u, Alice has a two-outcome measurement defined by a pair of d-by-d positive semidefinite matrices $\{A^0_u, A^1_u\}$ that satisfies $A^0_u + A^1_u = I$, where I is the identity matrix. Bob has a similar pair $\{B^0_v, B^1_v\}$ for each question v. When the players perform their measurements, the probability that they obtain bits a and b is given by $\mathrm{Tr}(A^a_u \otimes B^b_v \rho)$.

The case $d = 1$ corresponds to a classical setting. In this case, the maximum winning probability equals $\big(1+\mathrm{SDP}_1(G,A)\big)/2$, where G is the complete bipartite graph with Alice and Bob's questions on opposite sides of the partition, and $A(u,v) = (-1)^{g(u,v)}\pi(u,v)/2$ for pairs $\{u,v\}\in E$ and $A(u,v) = 0$ everywhere else.

Tsirel'son [46] related the maximum winning probability $\omega^*_d(\pi,g)$ of the game (π,g), when the players are restricted to measurements on d-dimensional quantum systems, to the quantity $\mathrm{SDP}_r(G,A)$. In particular, he proved that

$$\frac{1+\mathrm{SDP}_{\lfloor \log d\rfloor}(G,A)}{2} \le \omega^*_d(\pi,g) \le \frac{1+\mathrm{SDP}_{2d}(G,A)}{2}\,.$$

The quantity $\mathrm{SDP}_r(G,A)$ thus gives bounds on the maximum winning probability of XOR games when players are limited in the amount of entanglement they are allowed to use. The rank-r Grothendieck constant $K(r,G)$ of the bipartite graph G described above gives a quantitative bound on the advantage that unbounded entanglement gives over finite entanglement in XOR games.

1.4 Our results and methods

The purpose of this paper is to prove explicit upper bounds for $K(r,G)$. We are especially interested in the case of small r and graphs with small chromatic number, although our methods are not restricted to this. Our main theorem, Theorem 1.2, which will be stated shortly, can be used to compute the bounds for $K(r,G)$ shown in Table 1 below.

r	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	$1.093456\ldots$	$2.124024\ldots$

Table 1: Bounds on Grothendieck's constant.

Theorem 1.2 actually gives, for every r, a randomized polynomial-time approximation algorithm for the optimization problem $\mathrm{SDP}_r(G,A)$. So in particular it provides a randomized polynomial-time approximation algorithm for computing the ground states of the 3-vector model, also known as the Heisenberg model, in the lattice $\mathbb{Z}^3$ with approximation ratio $0.78\ldots = (1.28\ldots)^{-1}$. This result can be regarded as one of the main contributions of this paper.

To prove the main theorem we use the framework of Krivine and Haagerup which we explain below. Our main technical contributions are a matrix version of Grothendieck's identity (Lemma 2.1) and a method to construct new unit vectors which can also deal with nonbipartite graphs (Lemma 4.1).

The strategy of Haagerup and Krivine is based on the following embedding lemma:

Lemma 1.1. *Let $G = (V,E)$ be a graph and choose $Z = (Z_{ij}) \in \mathbb{R}^{r\times|V|}$ at random so that each entry is distributed independently according to the normal distribution with mean 0 and variance 1, that is, $Z_{ij} \sim N(0,1)$.*

Given $f\colon V \to S^{|V|-1}$, there is a function $g\colon V \to S^{|V|-1}$ such that whenever u and v are adjacent in

G, then

$$\mathbb{E}\left[\frac{Zg(u)}{\|Zg(u)\|}\cdot\frac{Zg(v)}{\|Zg(v)\|}\right]=\beta(r,G)f(u)\cdot f(v)$$

for some constant $\beta(r,G)$ depending only on r and G.

In the statement above we are vague regarding the constant $\beta(r,G)$. We will soon define more precisely the constant $\beta(r,G)$, and in Section 4 we will provide a precise statement of this lemma (cf. Lemma 4.1 there). Right now this precise statement is not relevant to our discussion.

Now, the strategy of Haagerup and Krivine amounts to analyzing the following four-step procedure that yields a randomized polynomial-time approximation algorithm for $\mathrm{SDP}_r(G,A)$:

Algorithm A. Takes as input a finite graph $G=(V,E)$ with at least one edge and a symmetric matrix $A\colon V\times V\to\mathbb{R}$, and returns a feasible solution $h\colon V\to S^{r-1}$ of $\mathrm{SDP}_r(G,A)$.

1. Solve $\mathrm{SDP}_\infty(G,A)$, obtaining an optimal solution $f\colon V\to S^{|V|-1}$.
2. Use f to construct $g\colon V\to S^{|V|-1}$ according to Lemma 1.1.
3. Choose $Z=(Z_{ij})\in\mathbb{R}^{r\times|V|}$ at random so that every matrix entry Z_{ij} is distributed independently according to the standard normal distribution with mean 0 and variance 1, that is, $Z_{ij}\sim N(0,1)$.
4. Define $h\colon V\to S^{r-1}$ by setting $h(u)=Zg(u)/\|Zg(u)\|$.

To analyze this procedure, we compute the expected value of the feasible solution h. Using Lemma 1.1 we obtain

$$\begin{aligned}\mathrm{SDP}_r(G,A)&\geq\mathbb{E}\left[\sum_{\{u,v\}\in E}A(u,v)h(u)\cdot h(v)\right]\\&=\sum_{\{u,v\}\in E}A(u,v)\mathbb{E}[h(u)\cdot h(v)]\\&=\beta(r,G)\sum_{\{u,v\}\in E}A(u,v)f(u)\cdot f(v)\\&=\beta(r,G)\,\mathrm{SDP}_\infty(G,A)\,,\end{aligned}\tag{1.2}$$

and so we have $K(r,G)\leq\beta(r,G)^{-1}$.

If we were to skip step (2) and apply step (4) to f directly, then the expectation $\mathbb{E}[h(u)\cdot h(v)]$ would be a non-linear function of $f(u)\cdot f(v)$, which would make it difficult to assess the quality of the feasible solution h. The purpose of step (2) is to linearize this expectation, which allows us to estimate the quality of h in terms of a linear function of $\mathrm{SDP}_r(G,A)$.

The constant $\beta(r,G)$ in Lemma 1.1 is defined in terms of the Taylor expansion of the inverse of the function $E_r\colon[-1,1]\to[-1,1]$ given by

$$E_r(x\cdot y)=\mathbb{E}\left[\frac{Zx}{\|Zx\|}\cdot\frac{Zy}{\|Zy\|}\right],$$

where $x, y \in S^\infty$ and $Z = (Z_{ij}) \in \mathbb{R}^{r\times\infty}$ is chosen so that its entries are independently distributed according to the normal distribution with mean 0 and variance 1. The function E_r is well-defined since the expectation above is invariant under orthogonal transformations.

The function E_r^{-1} has the Taylor expansion

$$E_r^{-1}(t) = \sum_{k=0}^{\infty} b_{2k+1} t^{2k+1},$$

with a positive radius of convergence around zero, as will be shown in Section 3. Our main theorem can be thus stated:

Theorem 1.2. *The Grothendieck constant $K(r,G)$ is at most $\beta(r,G)^{-1}$ where the number $\beta(r,G)$ is defined by the equation*

$$\sum_{k=0}^{\infty} |b_{2k+1}| \beta(r,G)^{2k+1} = \frac{1}{\vartheta(\overline{G})-1}, \tag{1.3}$$

where b_{2k+1} are the coefficients of the Taylor expansion of E_r^{-1} and where $\vartheta(\overline{G})$ is the theta number of the complement of the graph G. (In particular, there exists a number satisfying (1.3).)

(For a definition of the Lovász theta number of a graph, see (4.1) in Section 4.)

The Taylor expansion of E_r is computed in Section 2 and the Taylor expansion of E_r^{-1} is treated in Section 3. A precise version of Lemma 1.1 is stated and proved in Section 4, following Krivine [33]. As we showed above, this lemma, together with Algorithm A, then implies Theorem 1.2. In Section 6 we discuss how we computed Table 1. We computed the entries numerically and we strongly believe that all digits are correct even though we do not have a formal mathematical proof.

We finish this section with some remarks. When $r = 1$ and G is bipartite, Theorem 1.2 specializes to Krivine's [33] bound for the original Grothendieck constant $K_G = \lim_{n\to\infty} K(1, K_{n,n})$. For more than thirty years this was the best known upper bound, and it was conjectured by many to be optimal. However, shortly after our work appeared in preprint form, Braverman, Makarychev, Makarychev and Naor [10] showed that Krivine's bound can be slightly improved. In this light we now believe that the upper bound in Theorem 1.2 is not tight.

The best known *lower* bound on K_G is $1.676956\ldots$, due to Davie [16] and Reeds [43] (see also Khot and O'Donnell [30]).

When $r = 2$ and G is bipartite, Theorem 1.2 specializes to Haagerup's [24] upper bound for the complex Grothendieck constant $K_G^{\mathbb{C}}$; this is currently the best known upper bound for this constant.

Using different techniques, in [13] we proved for the asymptotic regime where r is large that $K(r, K_{n,n}) = 1 + \Theta(1/r)$ holds. A recent argument of Naor and Regev [38] (which was used to show that specific variations of Algorithm A exist whose approximation quality become arbitrary close to the Grothendieck constant) implies that Theorem 1.2 can also be used to prove an upper bound of $1 + O(1/r)$.

For graphs with large chromatic number Alon, Makarychev, Makarychev, and Naor [3] give the best known bounds for $K(1,G)$. They prove a logarithmic dependence on the chromatic number of the graph (actually on the theta number of the complement of G, cf. Section 4) whereas our methods only give a linear dependence. Although our main focus is on small chromatic numbers, for completeness we extend the results of [3] for large chromatic numbers to $r \geq 2$ in Section 5.

2 A matrix version of Grothendieck's identity

In the analysis of many approximation algorithms that use semidefinite programming the following identity plays a central role: Let u, v be unit (column) vectors in $\mathbb{R}^n$ and let $Z \in \mathbb{R}^{1\times n}$ be a random (row) vector whose entries are distributed independently according to the standard normal distribution with mean 0 and variance 1. Then,

$$\mathbb{E}[\operatorname{sign}(Zu)\operatorname{sign}(Zv)] = \mathbb{E}\left[\frac{Zu}{\|Zu\|}\cdot\frac{Zv}{\|Zv\|}\right] = \frac{2}{\pi}\arcsin(u\cdot v)\,.$$

For instance, the celebrated algorithm of Goemans and Williamson [22] for approximating the MAX CUT problem is based on this. The identity is called *Grothendieck's identity* since it appeared for the first time in Grothendieck's work on the metric theory of tensor products [23, Proposition 4, p. 63] (see also Diestel, Fourie, and Swart [17]).

In this section we extend Grothendieck's identity from vectors to matrices by replacing the arcsine function by a *hypergeometric function*, defined as follows. For any nonnegative integers p,q, real numbers $a_1,a_2,\dots,a_p$ and strictly positive real numbers $b_1,b_2,\dots,b_q$, there is a hypergeometric function

$${}_pF_q\left(\begin{matrix}a_1,a_2,\dots,a_p\\ b_1,b_2,\dots,b_q\end{matrix};x\right) = \sum_{k=0}^{\infty}\frac{(a_1)_k(a_2)_k\cdots(a_p)_k}{(b_1)_k(b_2)_k\cdots(b_q)_k}\frac{x^k}{k!}\,,$$

where

$$(c)_k = c(c+1)(c+2)\cdots(c+k-1)$$

denotes the rising factorial function. Conforming with the notation of Andrews, Askey and Roy [5], if $p=0$ we substitute the absent parameters a_i by a horizontal line:

$${}_0F_q\left(\begin{matrix}\text{---}\\ b_1,b_2,\dots,b_q\end{matrix};x\right).$$

Lemma 2.1. *Let u, v be unit vectors in $\mathbb{R}^n$ and let $Z \in \mathbb{R}^{r\times n}$ be a random matrix whose entries are distributed independently according to the standard normal distribution with mean 0 and variance 1. Then,*

$$\begin{aligned}\mathbb{E}\left[\frac{Zu}{\|Zu\|}\cdot\frac{Zv}{\|Zv\|}\right] &= \frac{2}{r}\left(\frac{\Gamma((r+1)/2)}{\Gamma(r/2)}\right)^2(u\cdot v)\,{}_2F_1\left(\begin{matrix}1/2,1/2\\ r/2+1\end{matrix};(u\cdot v)^2\right)\\ &= \frac{2}{r}\left(\frac{\Gamma((r+1)/2)}{\Gamma(r/2)}\right)^2\sum_{k=0}^{\infty}\frac{(1/2)_k(1/2)_k}{(r/2+1)_k}\frac{(u\cdot v)^{2k+1}}{k!}\,.\end{aligned} \tag{2.1}$$

Before proving the lemma we review special cases known in the literature. If $r=1$, then we get the original Grothendieck's identity:

$$\begin{aligned}\mathbb{E}[\operatorname{sign}(Zu)\operatorname{sign}(Zv)] &= \frac{2}{\pi}\arcsin(u\cdot v)\\ &= \frac{2}{\pi}\left(u\cdot v+\left(\frac{1}{2}\right)\frac{(u\cdot v)^3}{3}+\left(\frac{1\cdot 3}{2\cdot 4}\right)\frac{(u\cdot v)^5}{5}+\cdots\right).\end{aligned}$$

The case $r = 2$ is due to Haagerup [24]:

$$\begin{aligned}\mathbb{E}\left[\frac{Zu}{\|Zu\|}\cdot\frac{Zv}{\|Zv\|}\right] &= \frac{1}{u\cdot v}\left(E(u\cdot v)-(1-(u\cdot v)^2)K(u\cdot v)\right)\\ &= \frac{\pi}{4}\left(u\cdot v+\left(\frac{1}{2}\right)^2\frac{(u\cdot v)^3}{2}+\left(\frac{1\cdot 3}{2\cdot 4}\right)^2\frac{(u\cdot v)^5}{3}+\cdots\right),\end{aligned}$$

where K and E are the complete elliptic integrals of the first and second kind. Note that on page 201 of Haagerup [24] $\pi/2$ has to be $\pi/4$. Briët, Oliveira, and Vallentin [12] computed, for every r, the zeroth coefficient of the series in (2.1), which is the Taylor series of the expectation.

The following elegant proof of Grothendieck's identity has become a classic: We have

$$\operatorname{sign}(Zu)\operatorname{sign}(Zv) = 1$$

if and only if the vectors u and v lie on the same side of the hyperplane orthogonal to the vector $Z \in \mathbb{R}^{1\times n}$. Now we project this n-dimensional situation to the plane spanned by u and v. Then the projected random hyperplane becomes a random line. This random line is distributed according to the uniform probability measure on the unit circle because Z is normally distributed. Now one obtains the final result by measuring intervals on the unit circle: The probability that u and v lie on the same side of the line is $1 - \arccos(u\cdot v)/\pi$.

We do not have such a picture proof for our matrix version. Our proof is based on the rotational invariance of the normal distribution and integration with respect to spherical coordinates together with some identities for hypergeometric functions. A similar calculation was done by König and Tomczak-Jaegermann [32]. It would be interesting to find a more geometrical proof of the lemma.

For computing the first coefficient of the Taylor series in [12] we took a slightly different route: We integrated using the Wishart distribution of 2×2-matrices.

Proof of Lemma 2.1. Let $Z_i \in \mathbb{R}^n$ be the i-th row of the matrix Z, with $i = 1,\ldots r$. We define vectors

$$x = \begin{pmatrix} Z_1\cdot u\\ Z_2\cdot u\\ \vdots\\ Z_r\cdot u\end{pmatrix} \qquad \text{and} \qquad y = \begin{pmatrix} Z_1\cdot v\\ Z_2\cdot v\\ \vdots\\ Z_r\cdot v\end{pmatrix}$$

so that we have $x\cdot y = (Zu)\cdot(Zv)$. Since the probability distribution of the vectors Z_i is invariant under orthogonal transformations we may assume that $u = (1,0,\ldots,0)$ and $v = (t,\sqrt{1-t^2},0,\ldots,0)$ and so the pair $(x,y) \in \mathbb{R}^r\times\mathbb{R}^r$ is distributed according to the probability density function (see, e. g., Feller [20, p. 69])

$$(2\pi\sqrt{1-t^2})^{-r}\exp\left(-\frac{x\cdot x-2tx\cdot y+y\cdot y}{2(1-t^2)}\right).$$

Hence,

$$\mathbb{E}\left[\frac{x}{\|x\|}\cdot\frac{y}{\|y\|}\right] = (2\pi\sqrt{1-t^2})^{-r}\int_{\mathbb{R}^r}\int_{\mathbb{R}^r}\frac{x}{\|x\|}\cdot\frac{y}{\|y\|}\exp\left(-\frac{x\cdot x-2tx\cdot y+y\cdot y}{2(1-t^2)}\right)dx\,dy.$$

By using spherical coordinates $x = \alpha\xi$, $y = \beta\eta$, where $\alpha, \beta \in [0,\infty)$ and $\xi, \eta \in S^{r-1}$, we rewrite the above integral as

$$\int_0^\infty \int_0^\infty (\alpha\beta)^{r-1} \exp\left(-\frac{\alpha^2+\beta^2}{2(1-t^2)}\right) \int_{S^{r-1}} \int_{S^{r-1}} \xi \cdot \eta \exp\left(\frac{\alpha\beta t \xi \cdot \eta}{1-t^2}\right) d\omega(\xi)\, d\omega(\eta)\, d\alpha\, d\beta\,,$$

where ω is the surface-area measure, such that $\omega(S^{n-1}) = 2\pi^{n/2}/\Gamma(n/2)$.

If $r = 1$, we get for the inner double integral

$$\begin{aligned}\int_{S^0} \int_{S^0} \xi \cdot \eta \exp\left(\frac{\alpha\beta t \xi \cdot \eta}{1-t^2}\right) d\omega(\xi)\, d\omega(\eta) &= 4 \sinh\left(\frac{\alpha\beta t}{1-t^2}\right) \\ &= 4\frac{\alpha\beta t}{1-t^2}\, {}_0F_1\left(\begin{matrix} — \\ 3/2 \end{matrix}; \left(\frac{\alpha\beta t}{2(1-t^2)}\right)^2\right).\end{aligned}$$

Now we consider the case when $r \geq 2$. Since the inner double integral over the spheres only depends on the inner product $p = \xi \cdot \eta$ it can be rewritten as

$$\omega(S^{r-2})\omega(S^{r-1}) \int_{-1}^1 p \exp\left(\frac{\alpha\beta t p}{1-t^2}\right) (1-p^2)^{(r-3)/2}\, dp\,,$$

where

$$\omega(S^{r-2})\omega(S^{r-1}) = \frac{4\pi^{r-1/2}}{\Gamma(r/2)\Gamma((r-1)/2)}.$$

Integration by parts yields

$$\begin{aligned}&\int_{-1}^1 p(1-p^2)^{(r-3)/2} \exp\left(\frac{\alpha\beta t p}{1-t^2}\right) dp \\ &\quad = \frac{\alpha\beta t}{(r-1)(1-t^2)} \int_{-1}^1 (1-p^2)^{(r-1)/2} \exp\left(\frac{\alpha\beta t p}{1-t^2}\right) dp.\end{aligned}$$

The last integral can be rewritten using the modified Bessel function of the first kind (cf. Andrews, Askey, Roy [5, p. 235, Exercise 9])

$$\begin{aligned}&\int_{-1}^1 (1-p^2)^{(r-1)/2} \exp\left(\frac{\alpha\beta t p}{1-t^2}\right) dp \\ &\quad = \Gamma((r+1)/2)\sqrt{\pi} \left(\frac{2(1-t^2)}{\alpha\beta t}\right)^{r/2} I_{r/2}\left(\frac{\alpha\beta t}{1-t^2}\right).\end{aligned}$$

One can write $I_{r/2}$ as a hypergeometric function (cf. Andrews, Askey, and Roy [5, (4.12.2)])

$$I_{r/2}(x) = (x/2)^{r/2} \sum_{k=0}^\infty \frac{(x/2)^{2k}}{k!\Gamma(r/2+k+1)} = \frac{(x/2)^{r/2}}{\Gamma((r+2)/2)}\, {}_0F_1\left(\begin{matrix} — \\ (r+2)/2 \end{matrix}; \left(\frac{x}{2}\right)^2\right).$$

Putting things together, we get

$$\begin{aligned}\omega(S^{r-2})\omega(S^{r-1})\int_{-1}^{1} p\exp\left(\frac{\alpha\beta t p}{1-t^2}\right)(1-p^2)^{(r-3)/2}\,dp \\ = \frac{4\pi^r}{\Gamma(r/2)^2 r}\frac{\alpha\beta t}{1-t^2}{}_0F_1\left(\begin{matrix}\text{---}\\(r+2)/2\end{matrix};\left(\frac{\alpha\beta t}{2(1-t^2)}\right)^2\right).\end{aligned}$$

Notice that the last formula also holds for $r=1$. So we can continue without case distinction.

Now we evaluate the outer double integral

$$\int_0^\infty\int_0^\infty(\alpha\beta)^r\exp\left(-\frac{\alpha^2+\beta^2}{2(1-t^2)}\right){}_0F_1\left(\begin{matrix}\text{---}\\(r+2)/2\end{matrix};\left(\frac{\alpha\beta t}{2(1-t^2)}\right)^2\right)d\alpha\,d\beta\,.$$

Here the inner integral equals

$$\int_0^\infty\alpha^r\exp\left(-\frac{\alpha^2}{2(1-t^2)}\right){}_0F_1\left(\begin{matrix}\text{---}\\(r+2)/2\end{matrix};\left(\frac{\alpha\beta t}{2(1-t^2)}\right)^2\right)d\alpha\,,$$

and doing the substitution $\gamma=\alpha^2/(2(1-t^2))$ gives

$$2^{(r-1)/2}(1-t^2)^{(r+1)/2}\int_0^\infty\gamma^{(r-1)/2}\exp(-\gamma)\,{}_0F_1\left(\begin{matrix}\text{---}\\(r+2)/2\end{matrix};\frac{\gamma(\beta t)^2}{2(1-t^2)}\right)d\gamma,$$

which is by the Bateman Manuscript Project [19, p. 337 (11)] equal to

$$2^{(r-1)/2}(1-t^2)^{(r+1)/2}\Gamma((r+1)/2)\,{}_1F_1\left(\begin{matrix}(r+1)/2\\(r+2)/2\end{matrix};\frac{(\beta t)^2}{2(1-t^2)}\right).$$

Now we treat the remaining outer integral in a similar way, using [19, p. 219 (17)], and get that

$$\begin{aligned}\int_0^\infty\beta^r\exp\left(-\frac{\beta^2}{2(1-t^2)}\right){}_1F_1\left(\begin{matrix}(r+1)/2\\(r+2)/2\end{matrix};\frac{(\beta t)^2}{2(1-t^2)}\right)d\beta \\ = 2^{(r-1)/2}(1-t^2)^{(r+1)/2}\Gamma((r+1)/2)\,{}_2F_1\left(\begin{matrix}(r+1)/2,(r+1)/2\\(r+2)/2\end{matrix};t^2\right).\end{aligned}$$

By applying Euler's transformation (cf. Andrews, Askey, and Roy [5, (2.2.7)])

$${}_2F_1\left(\begin{matrix}(r+1)/2,(r+1)/2\\(r+2)/2\end{matrix};t^2\right)=(1-t^2)^{-r/2}\,{}_2F_1\left(\begin{matrix}1/2,1/2\\(r+2)/2\end{matrix};t^2\right)$$

and after collecting the remaining factors we arrive at the result. □

3 Convergence radius

To construct the new vectors in the second step of Algorithm A we will make use of the Taylor series expansion of the inverse of E_r. Locally around zero we can expand the function E_r^{-1} as

$$E_r^{-1}(t) = \sum_{k=0}^{\infty} b_{2k+1} t^{2k+1}, \tag{3.1}$$

but in the proof of Lemma 4.1 it will be essential that this expansion be valid on $[-\beta(r,G), \beta(r,G)]$. In the case $r = 1$ we have $E_1^{-1}(t) = \sin((\pi/2)t)$, whose Taylor expansion has infinite convergence radius. In this section we show that for all $r \geq 2$ the convergence radius of the Taylor series of E_r^{-1} is also large enough for our purposes. The case $r = 2$ was previously dealt with by Haagerup [24], who proved that the convergence radius is at least 1. Our proof, which applies uniformly for all cases $r \geq 2$ (but gives a smaller radius for $r = 2$), is based on elementary techniques from complex analysis.

Let $\mathbb{D} = \{ z \in \mathbb{C} : |z| < 1 \}$ denote the open unit disc and for a real number $c > 0$ define $c\mathbb{D} = \{ z \in \mathbb{C} : |z| < c \}$. Since the function E_r can be represented by a Taylor series in $[-1,1]$, it has an analytic extension $\mathcal{E}_r$ in $\mathbb{D}$ given by

$$\mathcal{E}_r(z) = C_r z\, {}_2F_1\left(\begin{matrix} 1/2, 1/2 \\ r/2+1 \end{matrix} ; z^2 \right), \qquad C_r = \frac{2}{r} \left(\frac{\Gamma\big((r+1)/2\big)}{\Gamma(r/2)} \right)^2. \tag{3.2}$$

Theorem 3.1. *Let r be a positive integer. Then, the Taylor series (3.1) has convergence radius at least $|\mathcal{E}_r(i)|$.*

Theorem 3.1 follows from Lemma 3.2 and Lemma 3.5 below, by observing that since $\mathcal{E}_r$ equals E_r on $[-1,1]$, $\mathcal{E}_r^{-1}$ equals E_r^{-1} on $[E_r(-1), E_r(1)]$.

Lemma 3.2. *Let r be a positive integer and let c_r be the number*

$$c_r = \min\{ |\mathcal{E}_r(\mathrm{e}^{it})| : t \in [0, 2\pi] \}. \tag{3.3}$$

Then the Taylor series at the origin of the function $\mathcal{E}_r^{-1}$ has convergence radius at least c_r.

For the proof of Lemma 3.2 we collect the following two basic facts (Proposition 3.3 and Proposition 3.4) about the function $\mathcal{E}_r$, which are consequences of Rouché's theorem. The proof strategy can also be found in the classical lectures [25] by Hurwitz.

Proposition 3.3. *The function $\mathcal{E}_r$ has exactly one root in $\mathbb{D}$ and this is a simple root located at the origin.*

Proof. Since $\mathcal{E}_r(0) = 0$ and $\mathcal{E}_r'(0) = C_r \neq 0$, the function $\mathcal{E}_r$ has a simple root at the origin. Recall that the Taylor coefficients $a_1, a_2, a_3, \ldots$ of $\mathcal{E}_r$ are nonnegative, that $a_1 = C_r$ and that $\mathcal{E}_r(1) = E_r(1) = \sum_{k=1}^{\infty} a_k = 1$. For any $z \in \partial\mathbb{D}$, the triangle inequality therefore gives

$$|\mathcal{E}_r(z) - C_r z| = \left| \sum_{k=2}^{\infty} a_k z^k \right| \leq \sum_{k=2}^{\infty} a_k = 1 - C_r. \tag{3.4}$$

In [11] it is shown that C_r increases with r. Now, since $C_1 = 2/\pi > 1/2$, we have $1 - C_r < C_r$ and (3.4) thus implies $|\mathcal{E}_r(z) - C_r z| < C_r = |C_r z|$. By Rouché's theorem, the function $\mathcal{E}_r$ therefore has the same number (counting multiplicities) of zeros in $\mathbb{D}$ as the function $C_r z$ does: one. $\square$

Proposition 3.4. *For any point $z \in c_r\mathbb{D}$ there is exactly one point $w \in \mathbb{D}$ such that $\mathcal{E}_r(w) = z$.*

Proof. If $z = 0$ the claim follows by Proposition 3.3. Fix a point z in the punctured disc $c_r\mathbb{D} \setminus \{0\}$ and define the function g by $g(w) = \mathcal{E}_r(w) - z$. For any $w \in \partial\mathbb{D}$ on the boundary of the unit disc,

$$|\mathcal{E}_r(w) - g(w)| = |z| < c_r \leq |\mathcal{E}_r(w)|\,.$$

Hence, by Rouché's Theorem the functions $\mathcal{E}_r$ and g have an equal number of roots in $\mathbb{D}$. It now follows from Proposition 3.3 that the function g has exactly one root in the punctured unit disc $\mathbb{D} \setminus \{0\}$ and that this is a simple root, which proves the claim. □

Proof of Lemma 3.2. To get the Taylor series of $\mathcal{E}_r^{-1}$ at the origin we express this function as a contour integral whose integrand we develop into a geometric series.

Let f be any function that is analytic in an open set of $\mathbb{C}$ that contains the closed unit disc $\overline{\mathbb{D}}$. For $z \in c_r\mathbb{D}$ consider the contour integral

$$I(z) = \frac{1}{2\pi i} \int_{\partial\mathbb{D}} f(w)\, \frac{\mathcal{E}_r'(w)}{\mathcal{E}_r(w) - z}\, dw\,,$$

where the integral is over the counter-clockwise path around the unit circle. By Proposition 3.4 the function $g(w) = \mathcal{E}_r(w) - z$ has exactly one root in $\mathbb{D}$ and this is a root of order one. Hence, by the residue theorem, $I(z)$ is the value of f at the root of g. This root is $\mathcal{E}_r^{-1}(z)$, so we have $I(z) = f\big(\mathcal{E}_r^{-1}(z)\big)$. By taking the function $f(w) = w$ we thus get

$$\mathcal{E}_r^{-1}(z) = \frac{1}{2\pi i} \int_{\partial\mathbb{D}} w\, \frac{\mathcal{E}_r'(w)}{\mathcal{E}_r(w) - z}\, dw\,. \tag{3.5}$$

We expand the fraction appearing in the integrand of (3.5) as

$$\frac{\mathcal{E}_r'(w)}{\mathcal{E}_r(w) - z} = \frac{\mathcal{E}_r'(w)}{\mathcal{E}_r(w)} \left(1 + \frac{z}{\mathcal{E}_r(w)} + \frac{z^2}{\mathcal{E}_r(w)^2} + \cdots\right). \tag{3.6}$$

For any $w \in \partial\mathbb{D}$ the above geometric series converges uniformly inside any disc $c_r'\mathbb{D}$, where $c_r' < c_r$, since then $|z| < c_r$ and by the definition of c_r (given in (3.3)), we have $|\mathcal{E}_r(w)| \geq c_r$. Substituting the fraction in the integrand of (3.5) by the right-hand side of (3.6) gives the Taylor series at the origin

$$\sum_{j=0}^{\infty} \left(\frac{1}{2\pi i} \int_{\partial\mathbb{D}} w\, \frac{\mathcal{E}_r'(w)}{\mathcal{E}_r(w)^{j+1}}\, dw\right) z^j\,, \tag{3.7}$$

which converges to $\mathcal{E}_r^{-1}(z)$ in $c_r\mathbb{D}$. □

Lemma 3.5. *Let $r \geq 2$ and c_r be as in (3.3). Then $c_r = |\mathcal{E}_r(i)|$.*

Proof. Inspection of the definition of $\mathcal{E}_r$ (given in (3.2)) shows that it suffices to consider the function

$$\mathcal{F}_r(t) = {}_2F_1\left(\begin{matrix}1/2, 1/2\\ r/2+1\end{matrix}; e^{it}\right)$$

and show that $|\mathcal{F}_r(t)|$ is minimized at $t = \pi$. To this end we write $|\mathcal{F}_r(t)|^2 = R_r(t)^2 + I_r(t)^2$, where R_r and I_r are the real and the imaginary part of this function:

$$R_r(t) = \sum_{k=0}^{\infty} \frac{(1/2)_k^2}{(r/2+1)_k} \frac{\cos(kt)}{k!},$$

$$I_r(t) = \sum_{k=0}^{\infty} \frac{(1/2)_k^2}{(r/2+1)_k} \frac{\sin(kt)}{k!}.$$

We have $I_r(\pi) = 0$ so if $R_r(t)^2$ attains a minimum at $t = \pi$ we are done. Notice that for $t \in [0, \pi]$ we have $R_r(\pi + t) = R_r(\pi - t)$. We claim that the derivative $\big(R_r(t)^2\big)' = 2R_r(t)R_r'(t)$ is strictly negative for $t \in (0, \pi)$. Since $R_r(t)^2$ is nonnegative and symmetric around π it then follows that its minimum on $[0, \pi]$ is indeed attained at π.

Claim 3.6. *The function R_r is strictly positive on* $(0, \pi)$.

Proof. Vietoris's theorem (see [5, Theorem 7.3.5]) states that for any positive integer n and real numbers $d_1, d_2, \dots, d_n$ that satisfy $d_1 \geq d_2 \geq \cdots \geq d_n > 0$ and $2kd_{2k} \leq (2k-1)d_{2k-1}$, $k \geq 1$, we have

$$\sum_{k=0}^{n} d_k \cos(kt) > 0 \qquad \text{for } 0 < t < \pi.$$

It is easy to check that the series $R_r(t)$ satisfies the conditions of Vietoris's theorem with

$$d_k = \frac{(1/2)_k^2}{(r/2+1)_k k!}$$

when $r \geq 2$, so the claim follows. □

Claim 3.7. *The derivative R_r' is strictly negative on* $(0, \pi)$.

Proof. The function R_r' is given by

$$R_r'(t) = -\sum_{k=0}^{\infty} \frac{(1/2)_k^2}{(r/2+1)_k \, k!} k \sin(kt). \tag{3.8}$$

We show that the ratios appearing on the right-hand side of (3.8) are the moments of a finite nonnegative (Borel) measure μ on $[0, 1]$,

$$\frac{(1/2)_k^2}{(r/2+1)_k \, k!} = \int_0^1 s^k \, d\mu(s). \tag{3.9}$$

To this end, we write

$$\frac{(1/2)_k^2}{(r/2+1)_k\,k!} = \left(\frac{\Gamma\big(\frac{r}{2}+1\big)}{\Gamma(\frac{1}{2})^2}\right)\left(\frac{\Gamma\big(k+\frac{1}{2}\big)}{\Gamma\big(k+\frac{r}{2}+1\big)}\right)\left(\frac{\Gamma\big(k+\frac{1}{2}\big)}{\Gamma(k+1)}\right). \tag{3.10}$$

For real numbers $a,b>0$ such that $a-b<0$, define the sequence $(d_k)_{k=0}^{\infty}$ by $d_k=\Gamma(k+a)/\Gamma(k+b)$. Let $\Delta d_k = d_{k+1}-d_k$ be the linear *difference operator* and for positive integer ℓ recursively define $\Delta^{\ell} d_k = \Delta(\Delta^{\ell-1} d_k)$. By the formula $\Gamma(x+1)=x\Gamma(x)$, we have

$$\Delta d_k = \left(\frac{k+a}{k+b}-1\right)d_k = \left(\frac{a-b}{k+b}\right)d_k\,,$$

and induction on ℓ gives

$$\Delta^{\ell} d_k = \frac{(a-b)(a-b-1)\cdots(a-b-\ell+1)}{(k+b)_{\ell}}\,d_k\,.$$

Since $a-b$ is negative this shows that $(-1)^{\ell}\Delta^{\ell} d_k \geq 0$ for every ℓ, which is to say that the sequence $(d_k)_k$ is completely monotonic. Hausdorff's theorem [20, pp. 223] says that a sequence is completely monotonic if and only if it is the moment sequence of some finite nonnegative Borel measure on $[0,1]$. In other words, there exist independent $[0,1]$-valued random variables X and Y and normalization constants $C_X, C_Y > 0$ such that for every integer $k \geq 0$, the right-hand side of (3.10) can be written as

$$\frac{\Gamma\big(\frac{r}{2}+1\big)}{\Gamma(\frac{1}{2})^2}\,C_X\mathbb{E}[X^k]\,C_Y\mathbb{E}[Y^k]\,.$$

By defining the $[0,1]$-valued random variable $Z=XY$, the above can be written as

$$C_X C_Y\big(\Gamma(r/2+1)/\Gamma(1/2)^2\big)\mathbb{E}[Z^k]\,,$$

which gives (3.9).

With this, (3.8) becomes

$$\begin{aligned}
R_r'(t) &= -\sum_{k=0}^{\infty}\left(\int_0^1 s^k\,d\mu(s)\right)k\sin(kt) \\
&= -\int_0^1\left(\sum_{k=0}^{\infty} s^k k\sin(kt)\right)d\mu(s) \\
&= -\int_0^1\left(\frac{s(1-s^2)\sin(t)}{\big(1-2s\cos(t)+s^2\big)^2}\right)d\mu(s) < 0\,, \quad 0<t<\pi\,,
\end{aligned}$$

where in last line we used the identity

$$\sum_{k=0}^{\infty} s^k k\sin(kt) = \frac{s(1-s^2)\sin(t)}{\big(1-2s\cos(t)+s^2\big)^2}\,, \quad 0 \leq s < 1\,,$$

which follows by differentiating the imaginary part of the Poisson kernel. This completes the proof. □

By combining the two claims, we get that $\big(R_r(t)^2\big)' = 2R_r(t)R_r'(t)$ is strictly negative on $(0,\pi)$, which gives the result. $\square$

Theorem 3.1 follows by combining Lemmas 3.2 and 3.5.

4 Constructing new vectors

In this section we use the Taylor expansion of the inverse of the function E_r to give a precise statement and proof of Lemma 1.1; this is done in Lemma 4.1. For this we follow Krivine [33], who proved the statement of the lemma in the case of bipartite graphs. We comment on how his ideas are related to our construction, which can also deal with nonbipartite graphs, after we prove the lemma.

For the nonbipartite case we need to use the theta number, which is a graph parameter introduced by Lovász [36]. Let $G = (V,E)$ be a graph. The *theta number* of the complement of G, denoted by $\vartheta(\overline{G})$, is the optimal value of the following semidefinite program:

$$\begin{aligned}\vartheta(\overline{G}) = \min\Big\{ \lambda :\ & Z \in \mathbb{R}^{V\times V}_{\succeq 0},\\ & Z(u,u) = \lambda - 1 \ \text{ for } u \in V,\\ & Z(u,v) = -1 \ \text{ for } \{u,v\} \in E \Big\}.\end{aligned} \tag{4.1}$$

It is known that the theta number of the complement of G provides a lower bound for the chromatic number of G. This can be easily seen as follows. Any proper k-coloring of G defines a mapping of V to the vertices of a $(k-1)$-dimensional regular simplex whose vertices lie on a sphere of radius $\sqrt{k-1}$: Vertices in the graph having the same color are sent to the same vertex in the regular simplex and vertices of different colors are sent to different vertices in the regular simplex. The Gram matrix of these vectors gives a feasible solution of (4.1).

Lemma 4.1. *Let $G = (V,E)$ be a graph with at least one edge. Given a function $f\colon V \to S^{|V|-1}$, there is a function $g\colon V \to S^{|V|-1}$ such that if u and v are adjacent,*

$$E_r\big(g(u)\cdot g(v)\big) = \beta(r,G) f(u)\cdot f(v)\,.$$

The constant $\beta(r,G)$ is defined as the solution of the equation

$$\sum_{k=0}^{\infty} |b_{2k+1}|\beta(r,G)^{2k+1} = \frac{1}{\vartheta(\overline{G})-1}\,, \tag{4.2}$$

where the coefficients b_{2k+1} are those of the Taylor series (3.1).

Recall from Theorem 3.1 and Lemma 3.5 that the series (3.1) has convergence radius at least $c_r = |\mathcal{E}_r(i)|$. The proof of Lemma 4.1 relies on the following proposition.

Proposition 4.2. *Let r be a positive integer and G be a graph with at least one edge. Then, for any $t \in [-1,1]$, the series*

$$\sum_{k=0}^{\infty} b_{2k+1}\big(t\,\beta(r,G)\big)^{2k+1}$$

converges to $E_r^{-1}\big(t\,\beta(r,G)\big)$.

Proof. As described in the beginning of Section 3, the case $r = 1$ follows from the fact that in that case the convergence radius of the series (3.1) is infinity. For $r \geq 2$, we consider the series

$$f(t) = \sum_{k=0}^{\infty} |b_{2k+1}| t^{2k+1} . \tag{4.3}$$

Let $\beta = \beta(r, G)$ be as in (4.2). To prove the claim it suffices to show that β indeed exists and that it lies in an interval where the series (3.1) converges to E_r^{-1}. Since G has at least one edge, we have $\vartheta(\overline{G}) \geq 2$. Hence, by (4.2) the number β should satisfy $f(\beta) \leq 1$. Note that f is well defined for any $t \in (-c_r, c_r)$, which follows from Theorem 3.1 and Lemma 3.5 showing that the series (4.3) converges in $c_r\mathbb{D}$.

We distinguish two cases based on the behavior of f at the point c_r. The first case is: $f(c_r) = +\infty$. In this case notice that $f(0) = 0$ and that f is continuous and increasing on the interval $[0, c_r)$. Since $f(c_r) > 1$ it follows that there exists a $t \in (0, c_r)$ such that $f(t) = 1$. Now, since $0 \leq f(\beta) \leq 1$, we see that β exists and lies in the radius of convergence of the series (3.1).

The second case is that $f(c_r)$ is finite. Recall that the Taylor series at the origin of the complex function $\mathcal{E}_r^{-1}$ is given by

$$\sum_{k=0}^{\infty} b_{2k+1} z^{2k+1} . \tag{4.4}$$

Then, since for any $z \in \overline{\mathbb{D}}$ the triangle inequality gives

$$\Big| \sum_{k=0}^{\infty} b_{2k+1} (c_r z)^{2k+1} \Big| \leq f(c_r) ,$$

the series (4.4) converges absolutely in the *closed* disc $c_r\overline{\mathbb{D}}$ and thus defines a continuous function $g : c_r\overline{\mathbb{D}} \to \mathbb{C}$. By Lemma 3.2 g equals $\mathcal{E}_r^{-1}$ in the open disc $c_r\mathbb{D}$, but by continuity of both $\mathcal{E}_r^{-1}$ and g, this equality must hold even in its closure $c_r\overline{\mathbb{D}}$. In particular, this implies that the series (3.1) converges to E_r^{-1} on $[-c_r, c_r]$.

Next, we argue that $\beta \leq c_r$. Since $\mathcal{E}_r$ is an odd function, and using Lemma 3.5,

$$\mathcal{E}_r(i) = \sum_{k=0}^{\infty} a_{2k+1} i^{2k+1} = i \sum_{k=0}^{\infty} a_{2k+1} (-1)^k = \pm i c_r . \tag{4.5}$$

Suppose that $\mathcal{E}_r(i) = i c_r$ holds (the other case $\mathcal{E}_r(i) = -i c_r$ follows by the same argument). Then, the above discussion implies that applying $\mathcal{E}_r^{-1}$ to both sides of (4.5) gives

$$i = \big(\mathcal{E}_r\big)^{-1}(i c_r) = \sum_{k=0}^{\infty} b_{2k+1} (i c_r)^{2k+1} = i \sum_{k=0}^{\infty} b_{2k+1} (-1)^k c_r^{2k+1} . \tag{4.6}$$

Taking absolute values of the left- and right-hand sides of (4.6) gives

$$1 = \Big| i \sum_{k=0}^{\infty} b_{2k+1} (-1)^k c_r^{2k+1} \Big| \leq \sum_{k=0}^{\infty} |b_{2k+1}| c_r^{2k+1} = f(c_r) .$$

Hence, $f(\beta) \leq 1 \leq f(c_r)$ and from the fact that f is increasing and zero at the origin we conclude that β exists and that $\beta \in [0, c_r]$. $\square$

Now we prove Lemma 4.1.

Proof of Lemma 4.1. We construct the vectors $g(u) \in S^{|V|-1}$ by constructing vectors $R(u)$ in an infinite-dimensional Hilbert space whose inner product matrix coincides with the one of the $g(u)$. We do this in three steps.

In the first step, set $H = \mathbb{R}^{|V|}$ and consider the Hilbert space

$$\mathcal{H} = \bigoplus_{k=0}^{\infty} H^{\otimes(2k+1)}.$$

For a unit vector $x \in H$, consider the vectors $S(x)$, $T(x) \in \mathcal{H}$ given componentwise by

$$S(x)_k = \sqrt{|b_{2k+1}|\beta(r,G)^{2k+1}}\, x^{\otimes(2k+1)}$$

and

$$T(x)_k = \operatorname{sign}(b_{2k+1})\sqrt{|b_{2k+1}|\beta(r,G)^{2k+1}}\, x^{\otimes(2k+1)}.$$

From Proposition 4.2 it follows that for any vectors $x, y \in S^{|V|-1}$ we have

$$S(x) \cdot T(y) = E_r^{-1}(\beta(r,G) x \cdot y),$$

and moreover

$$S(x) \cdot S(x) = T(x) \cdot T(x) = \sum_{k=0}^{\infty} |b_{2k+1}|\beta(r,G)^{2k+1} = \frac{1}{\vartheta(\overline{G}) - 1}.$$

In the second step, let $\lambda = \vartheta(\overline{G})$, and Z be an optimal solution of (4.1). We have $\lambda \geq 2$ since G has at least one edge. Let J denote the $|V| \times |V|$ all-ones matrix and set

$$A = \frac{(\lambda - 1)(J + Z)}{2\lambda} \qquad \text{and} \qquad B = \frac{(\lambda - 1)J - Z}{2\lambda},$$

and consider the matrix

$$U = \begin{pmatrix} A & B \\ B & A \end{pmatrix}.$$

By applying a Hadamard transformation

$$\frac{1}{\sqrt{2}}\begin{pmatrix} I & I \\ I & -I \end{pmatrix} U \frac{1}{\sqrt{2}}\begin{pmatrix} I & I \\ I & -I \end{pmatrix} = \begin{pmatrix} A + B & 0 \\ 0 & A - B \end{pmatrix}$$

one sees that U is positive semidefinite, since both $A + B$ and $A - B$ are positive semidefinite. Define $s \colon V \to \mathbb{R}^{2|V|}$ and $t \colon V \to \mathbb{R}^{2|V|}$ so that

$$s(u) \cdot s(v) = t(u) \cdot t(v) = A(u,v) \qquad \text{and} \qquad s(u) \cdot t(v) = B(u,v).$$

Matrix U is the Gram matrix of vectors $\big(s(u)\big)_{u \in V}$ and $\big(t(v)\big)_{v \in V}$. It follows that these maps have the following properties:

1. $s(u) \cdot t(u) = 0$ for all $u \in V$,
2. $s(u) \cdot s(u) = t(u) \cdot t(u) = (\vartheta(\overline{G}) - 1)/2$ for all $u \in V$,
3. $s(u) \cdot s(v) = t(u) \cdot t(v) = 0$ whenever $\{u, v\} \in E$,
4. $s(u) \cdot t(v) = s(v) \cdot t(u) = 1/2$ whenever $\{u, v\} \in E$.

In the third step we combine the previous two. We define the vectors

$$R(u) = s(u) \otimes S(f(u)) + t(u) \otimes T(f(u)).$$

For adjacent vertices $u, v \in V$ we have

$$R(u) \cdot R(v) = E_r^{-1}(\beta(r, G) f(u) \cdot f(v)),$$

and moreover the $R(u)$ are unit vectors. Hence, one can use the Cholesky decomposition of the matrix $(R(u) \cdot R(v))_{u,v \in V} \in \mathbb{R}^{V \times V}_{\succeq 0}$ to define the desired function $g \colon V \to S^{|V|-1}$. □

We conclude this section with a few remarks on Lemma 4.1 and its proof:

1. To approximate the Gram matrix $(R(u) \cdot R(v))$ it is enough to compute the series expansion of E_r^{-1} and the matrix U to the desired precision. The latter is found by solving a semidefinite program.
2. Krivine proved the statement of the lemma in the case $r = 1$ and for bipartite graphs G; then, $\vartheta(\overline{G}) = 2$ holds. Here, one only needs the first step of the proof. Also, $\beta(1, G)$ can be computed analytically. We have $E_1^{-1}(t) = \sin(\pi/2t)$ and

$$\sum_{k=0}^{\infty} \left| (-1)^{2k+1} \frac{(\pi/2)^{2k+1}}{(2k+1)!} \right| t^{2k+1} = \sinh(\pi/2t).$$

   Hence, $\beta(1, G) = 2 \operatorname{arcsinh}(1)/\pi = 2 \ln(1 + \sqrt{2})/\pi$.
3. In the second step one can also work with any feasible solution of the semidefinite program (4.1). For instance one can replace $\vartheta(\overline{G})$ in the lemma by the chromatic number $\chi(G)$ albeit getting a potentially weaker bound.
4. Alon, Makarychev, Makarychev, and Naor [3] also gave an upper bound for $K(1, G)$ using the theta number of the complement of G. They prove that

$$K(1, G) \leq O(\log \vartheta(\overline{G})),$$

   which is much better than our result in the case of large $\vartheta(\overline{G})$. However, our bound is favourable when $\vartheta(\overline{G})$ is small. In Section 5 we generalize the methods of Alon, Makarychev, Makarychev, and Naor [3] to obtain better upper bounds on $K(r, G)$ for $r \geq 2$ and large $\vartheta(\overline{G})$.

5 Better bounds for large chromatic numbers

For graphs with large chromatic number, or more precisely with large $\vartheta(\overline{G})$, our bounds on $K(r,G)$ proved above can be improved using the techniques of Alon, Makarychev, Makarychev, and Naor [3]. In this section, we show how their bounds on $K(1,G)$ generalize to higher values of r.

Theorem 5.1. *For a graph $G=(V,E)$ and integer $1 \le r \le \log \vartheta(\overline{G})$, we have*

$$K(r,G) \le O\left(\frac{\log \vartheta(\overline{G})}{r}\right).$$

Proof. It suffices to show that for any matrix $A : V \times V \to \mathbb{R}$, we have

$$\mathrm{SDP}_r(G,A) \ge \Omega\left(\frac{r}{\log \vartheta(\overline{G})}\right) \mathrm{SDP}_\infty(G,A).$$

Fix a matrix $A : V \times V \to \mathbb{R}$. Let $f : V \to S^{|V|-1}$ be optimal for $\mathrm{SDP}_\infty(G,A)$, so that

$$\sum_{\{u,v\}\in E} A(u,v) f(u)\cdot f(v) = \mathrm{SDP}_\infty(G,A).$$

Let $\lambda = \vartheta(\overline{G})$, and $\widetilde{Z} : V \times V \to \mathbb{R}$ be an optimal solution of (4.1). Since the matrix $\widetilde{Z}$ is positive semidefinite we get from its Gram decomposition column vectors $z(u) \in \mathbb{R}^{|V|}$ for $u \in V$. From the properties of $\widetilde{Z}$ it follows that $z(u)\cdot z(u) = \lambda - 1$ and $z(u)\cdot z(v) = -1$ if $\{u,v\} \in E$. Denote by $\mathbf{0} \in \mathbb{R}^{|V|}$ the all-zero vector. We now define vectors $s(u), t(u) \in \mathbb{R}^{2|V|+1}$ as

$$s(u) = \frac{1}{\sqrt{\lambda}}\begin{bmatrix} z(u) \\ \mathbf{0} \\ 1 \end{bmatrix} \quad \text{and} \quad t(u) = \frac{1}{\sqrt{\lambda}}\begin{bmatrix} \mathbf{0} \\ z(u) \\ 1 \end{bmatrix}.$$

It is easy to verify that these vectors have the following dot products:

1. $s(u)\cdot s(u) = t(u)\cdot t(u) = 1$ for all $u \in V$.
2. $s(u)\cdot t(u) = 1/\lambda$ for all $u,v \in V$.
3. $s(u)\cdot s(v) = t(u)\cdot t(v) = 0$ for all $\{u,v\} \in E$.

Let $\mathcal{H}$ be the Hilbert space of vector-valued functions $h : \mathbb{R}^{r\times|V|} \to \mathbb{R}^r$ with inner product

$$(g,h) = \mathbb{E}[g(Z)\cdot h(Z)],$$

where the expectation is taken over random $r \times |V|$ matrices Z whose entries are i. i. d. $N(0,1/r)$ random variables.

Let $R \ge 2$ be some real number to be set later. Define for every $u \in V$ the function $g_u \in \mathcal{H}$ by

$$g_u(Z) = \begin{cases} \frac{Zf(u)}{R} & \text{if } \|Zf(u)\| \le R, \\ \frac{Zf(u)}{\|Zf(u)\|} & \text{otherwise.} \end{cases}$$

Notice that for every matrix $Z \in \mathbb{R}^{r\times|V|}$, the vector $g_u(Z) \in \mathbb{R}^r$ has Euclidean norm at most 1. It follows by linearity of expectation that

$$\mathrm{SDP}_r(G,A) \geq \mathbb{E}\Bigg[\sum_{\{u,v\}\in E} A(u,v)\, g_u(Z)\cdot g_v(Z)\Bigg] = \sum_{\{u,v\}\in E} A(u,v)(g_u,g_v)\,.$$

We proceed by lower bounding the right-hand side of the above inequality.

Based on the definition of g_u we define two functions $h_u^0, h_u^1 \in \mathcal{H}$ by

$$h_u^0(Z) = \frac{Zf(u)}{R} + g_u(Z) \quad \text{and} \quad h_u^1(Z) = \frac{Zf(u)}{R} - g_u(Z)\,.$$

For every $u \in V$, define the function $H_u \in \mathbb{R}^{2|V|}\otimes\mathcal{H}$ by

$$H_u = \frac{1}{4}s(u)\otimes h_u^0 + 2\lambda\, t(u)\otimes h_u^1\,.$$

We expand the inner products (g_u,g_v) in terms of $f(u)\cdot f(v)$ and $\langle H_u,H_v\rangle$.

Claim 5.2. *For every $\{u,v\}\in E$ we have*

$$(g_u,g_v) \quad = \quad \frac{1}{R^2}f(u)\cdot f(v) - \langle H_u,H_v\rangle\,.$$

Proof. Simply expanding the inner product $\langle H_u,H_v\rangle$ gives

$$\begin{aligned}\langle H_u,H_v\rangle = {} & \frac{s(u)\cdot s(v)}{16}(h_u^0,h_v^0) + 4\lambda^2\big(t(u)\cdot t(v)\big)\,(h_u^1,h_v^1) \\ & + \frac{\lambda}{2}\Big[\big(s(u)\cdot t(v)\big)\,(h_u^0,h_v^1) + \big(t(u)\cdot s(v)\big)\,(h_u^1,h_v^0)\Big].\end{aligned}$$

It follows from property 3 of s and t that the above terms involving $s(u)\cdot s(v)$ and $t(u)\cdot t(v)$ vanish. By property 2, the remaining terms reduce to

$$\begin{aligned}\frac{1}{2}\Big((h_u^0,h_v^1)+(h_u^1,h_u^0)\Big) = {} & \frac{1}{2}\mathbb{E}\left[\left(\frac{Zf(u)}{R}+g_u(Z)\right)\cdot\left(\frac{Zf(v)}{R}-g_v(Z)\right)\right] \\ & + \frac{1}{2}\mathbb{E}\left[\left(\frac{Zf(u)}{R}-g_u(Z)\right)\cdot\left(\frac{Zf(v)}{R}+g_v(Z)\right)\right].\end{aligned}$$

Expanding the first expectation gives

$$\frac{1}{R^2}\mathbb{E}[f(u)^{\mathsf{T}}Z^{\mathsf{T}}Zf(v)] - (g_u,g_v) - \mathbb{E}\left[\frac{Zf(u)}{R}\cdot g_v(Z)\right] + \mathbb{E}\left[g_u(Z)\cdot\frac{Zf(v)}{R}\right]$$

and expanding the second gives

$$\frac{1}{R^2}\mathbb{E}[f(u)^{\mathsf{T}}Z^{\mathsf{T}}Zf(v)] - (g_u,g_v) + \mathbb{E}\left[\frac{Zf(u)}{R}\cdot g_v(Z)\right] - \mathbb{E}\left[g_u(Z)\cdot\frac{Zf(v)}{R}\right].$$

Adding these two gives that the last two terms cancel. Since $\mathbb{E}[Z^{\mathsf{T}}Z] = I$, what remains equals

$$\frac{1}{R^2} f(u)\cdot f(v) - (g_u, g_v)\,,$$

which proves the claim. □

From the above claim it follows that

$$\begin{aligned}\sum_{\{u,v\}\in E} A(u,v)(g_u,g_v) &= \frac{1}{R^2}\,\mathrm{SDP}_\infty(G,A) - \sum_{\{u,v\}\in E} A(u,v)\langle H_u, H_v\rangle \\ &\geq \left(\frac{1}{R^2} - \max_{u\in V}\|H_u\|^2\right)\mathrm{SDP}_\infty(G,A)\,,\end{aligned}$$

where $\|H_u\|^2 = \langle H_u, H_u\rangle$.

By the triangle inequality, we have for every $u \in V$,

$$\|H_u\|^2 \leq \left(\frac{1}{4}\|h_u^0\| + 2\lambda\,\|h_u^1\|\right)^2 \leq \frac{1}{R^2}\left(\frac{1}{2} + 2\lambda R\,\mathbb{E}\left[\left\|\frac{Zf(u)}{R} - g_u(Z)\right\|\right]\right)^2.$$

By the definition of g_u, the vectors $Zf(u)$ and g_u are parallel. Moreover, they are equal if $\|Zf(u)\| \leq R$. Since $f(u)$ is a unit vector, the r entries of the random vector $Zf(u)$ are i. i. d. $N(0,1/r)$ random variables. Hence,

$$\begin{aligned}\mathbb{E}\left[\left\|\frac{Zf(u)}{R} - g_u(Z)\right\|\right] &= \int_{\mathbb{R}^r} \mathbf{1}[\|x\| \geq R]\Big(\frac{\|x\|}{R} - 1\Big)\Big(\frac{r}{2\pi}\Big)^{r/2} \mathrm{e}^{-r\|x\|^2/2}\,dx \\ &= \int_R^\infty \int_{S^{r-1}} \rho^{r-1}\Big(\frac{\rho}{R} - 1\Big)\Big(\frac{r}{2\pi}\Big)^{r/2}\mathrm{e}^{-r\rho^2/2}\,d\rho\, d\tilde{\omega}_r(\xi) \\ &\leq \frac{r^{r/2}}{R\Gamma\big(\frac{r}{2}\big)}\int_R^\infty \rho^r \mathrm{e}^{-r\rho^2/2}\,d\rho\,,\end{aligned}$$

where $\tilde{\omega}_r$ is the unique rotationally invariant measure on S^{r-1}, normalized so that $\tilde{\omega}_r(S^{r-1}) = r^{r/2}/\Gamma(r/2)$. Using a substitution of variables, we get

$$\int_R^\infty \rho^r \mathrm{e}^{-r\rho^2/2} d\rho = \frac{1}{2}\Big(\frac{2}{r}\Big)^{(r+1)/2}\Gamma\Big(\frac{r+1}{2}, \frac{rR^2}{2}\Big)\,,$$

where $\Gamma(a,x)$ is the lower incomplete Gamma function [5, Eq. (4.4.5)].

Collecting the terms from above then gives the bound

$$\mathrm{SDP}_r(G,A) \geq \frac{1}{R^2}\left(1 - \left(\frac{1}{2} + \lambda\,\frac{2^{(r+1)/2}}{\sqrt{r}\,\Gamma\big(\frac{r}{2}\big)}\Gamma\Big(\frac{r+1}{2}, \frac{rR^2}{2}\Big)\right)^2\right)\mathrm{SDP}_\infty(G,A)\,. \tag{5.1}$$

The bound in the theorem follows by setting R as small as possible such that the above factor between brackets is some positive constant.

By Stirling's formula, there is a constant $C_1 > 0$ such that $\Gamma(x) \geq C_1 \mathrm{e}^{-x} x^{x-1/2}$ holds (see for example Abramowitz and Stegun [2, Eq. (6.1.37)]). Hence, for some constants $c, C > 0$, we have

$$\frac{2^{(r+1)/2}}{\sqrt{r}\,\Gamma\left(\frac{r}{2}\right)} \leq C\left(\frac{c}{r}\right)^{r/2}. \tag{5.2}$$

The power series of the incomplete gamma function [2, Eq. (6.5.32)] gives that if $a \leq x$, for some constant $C_2 > 0$, the inequality $\Gamma(a,x) \leq C_2 x^a \mathrm{e}^{-x}$ holds. As $R \geq 2$, for some constants $d, D > 0$, we have

$$\Gamma\left(\frac{r+1}{2}, \frac{rR^2}{2}\right) \leq D\sqrt{r}\left(\frac{r}{d^{R^2}}\right)^{r/2}. \tag{5.3}$$

Putting together estimates (5.2) and (5.3) gives

$$\lambda \frac{2^{(r+1)/2}}{\sqrt{r}\,\Gamma\left(\frac{r}{2}\right)} \Gamma\left(\frac{r+1}{2}, \frac{rR^2}{2}\right) \quad \leq \quad CD\sqrt{r}\lambda\left(\frac{c}{d^{R^2}}\right)^{r/2}.$$

Since $r \leq \log\lambda$ there is some constant C' such that for $R^2 = C'\big(\log\lambda\big)/r$, the above value is less than $1/4$. It follows that for this value of R, inequality (5.1) is nontrivial and we get the result. □

6 Numerical computation of Table 1

Table 1 shows numerical estimates of $1/\beta(r,G)$ which can be obtained using the Taylor polynomials of f given by

$$f_M(\beta) = \sum_{k=0}^{M} |b_{2k+1}|\beta^{2k+1}$$

for positive integers M. The numerical estimates of $\beta(r,G)$ can be obtained by solving for $f_M(\beta) = 1/\big(\vartheta(\overline{G})-1\big)$ using a computer program such as the PARI/GP [1] code given below. We strongly believe that all the digits of the table are correct but our computations are just numerical and do not yield a formal proof. Since we also believe that our bound is not sharp, $K(r,G) < \beta(r,G)^{-1}$, we did not make the effort to transform the numerical computations into rigorous proofs.

Acknowledgements

The first author thanks Stamatis Koumandos for helpful suggestions regarding the proof of Claim 3.7. The third author thanks Assaf Naor for helpful comments, and the Institute for Pure & Applied Mathematics at UCLA for its hospitality and support during the program "Modern Trends in Optimization and Its Application," September 17–December 17, 2010. We thank the anonymous referees for many helpful comments and suggestions.

Source code 1 PARI/GP [1] code to generate Table 1.

```
\p 100
M = 100

r = 2
rhs = 1
Er = 0 + O(t^(2*n+2))
for(k=0, M, Er = Er + (gamma(1/2+k)^2 *
gamma(r/2+1))/(gamma(1/2)^2 * gamma(r/2+1+k) * gamma(k+1)) *
t^(2*k+1))
Er = 2/r * (gamma((r+1)/2)/gamma(r/2))^2 * Er
Erinv = serreverse(Er)

fpol = 0
for(k=0, M, fpol = fpol + abs(polcoeff(Erinv,2*k+1))*t^(2*k+1))

beta_rG = polroots(fpol-rhs)[1]
```

References

[1] PARI/GP computer algebra system, version 2.5.3, Bordeaux, 2012. PARI/GP. 99, 100

[2] MILTON ABRAMOWITZ AND IRENE ANN STEGUN: *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. U.S. Department of Commerce, National Bureau of Standards, 1964. Available at Internet Archive. 99

[3] NOGA ALON, KONSTANTIN MAKARYCHEV, YURY MAKARYCHEV, AND ASSAF NAOR: Quadratic forms on graphs. *Invent. Math.*, 163(3):499–522, 2006. Preliminary version in STOC'05. [doi:10.1007/s00222-005-0465-9] 78, 79, 83, 95, 96

[4] NOGA ALON AND ASSAF NAOR: Approximating the cut-norm via Grothendieck's inequality. *SIAM J. Comput.*, 35(4):787–803, 2006. Preliminary version in STOC'04. [doi:10.1137/S0097539704441629] 79

[5] GEORGE E. ANDREWS, RICHARD ASKEY, AND RANJAN ROY: *Special Functions*. Volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge Univ. Press, 1999. [doi:10.1017/CBO9781107325937] 84, 86, 87, 90, 98

[6] SANJEEV ARORA, ELI BERGER, ELAD HAZAN, GUY KINDLER, AND MULI SAFRA: On non-approximability for quadratic programs. In *Proc. 46th FOCS*, pp. 206–215. IEEE Comp. Soc. Press, 2005. See also at ECCC. [doi:10.1109/SFCS.2005.57] 79

[7] ALAIN ASPECT, PHILIPPE GRANGIER, AND GÉRARD ROGER: Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.*, 47(7):460–463, 1981. [doi:10.1103/PhysRevLett.47.460] 80

[8] JOHN S. BELL: On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(3):195–200, 1964. Available at Wissenschaftstheorie und Wissenschaftsgeschichte. 80

[9] AHARON BEN-TAL AND ARKADI NEMIROVSKI: On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. *SIAM J. on Optim.*, 12(3):811–833, 2002. [doi:10.1137/S1052623400374756] 79

[10] MARK BRAVERMAN, KONSTANTIN MAKARYCHEV, YURY MAKARYCHEV, AND ASSAF NAOR: The Grothendieck constant is strictly smaller than Krivine's bound. In *Proc. 52nd FOCS*, pp. 453–462. IEEE Comp. Soc. Press, 2011. See also at arXiv. [doi:10.1109/FOCS.2011.77] 83

[11] JOP BRIËT, HARRY BUHRMAN, AND BEN TONER: A generalized Grothendieck inequality and nonlocal correlations that require high entanglement. *Comm. Math. Phys.*, 305(3):827–843, 2011. [doi:10.1007/s00220-011-1280-3] 78, 79, 88

[12] JOP BRIËT, FERNANDO MÁRIO DE OLIVEIRA FILHO, AND FRANK VALLENTIN: The positive semidefinite Grothendieck problem with rank constraint. In *Proc. 37th Internat. Colloq. on Automata, Languages and Programming (ICALP'10)*, pp. 31–42. Springer, 2010. [doi:10.1007/978-3-642-14165-2_4] 85

[13] JOP BRIËT, FERNANDO MÁRIO DE OLIVEIRA FILHO, AND FRANK VALLENTIN: The Grothendieck problem with rank constraint. In *Proceedings of the 19th Symposium on Mathematical Theory of Networks and Systems (MTNS'10)*, pp. 111–113, 2010. MTNS. 83

[14] MOSES CHARIKAR AND ANTHONY WIRTH: Maximizing quadratic programs: Extending Grothendieck's inequality. In *Proc. 45th FOCS*, pp. 54–60. IEEE Comp. Soc. Press, 2004. [doi:10.1109/FOCS.2004.39] 79

[15] RICHARD CLEVE, PETER HØYER, BENJAMIN TONER, AND JOHN WATROUS: Consequences and limits of nonlocal strategies. In *Proc. 19th IEEE Conf. on Computational Complexity (CCC'04)*, pp. 236–249. IEEE Comp. Soc. Press, 2004. See also at arXiv. [doi:10.1109/CCC.2004.1313847] 80

[16] ALEXANDER DAVIE: Lower bound for K_G. Unpublished, 1984. 83

[17] JOE DIESTEL, JAN H. FOURIE, AND JOHAN SWART: *The metric theory of tensor products: Grothendieck's résumé revisited.* Amer. Math. Soc., 2008. AMS. 84

[18] ALBERT EINSTEIN, BORIS PODOLSKY, AND NATHAN ROSEN: Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47(10):777–780, 1935. [doi:10.1103/PhysRev.47.777] 80

[19] ARTHUR ERDÉLYI, WILHELM MAGNUS, FRITZ OBERHETTINGER, AND FRANCESCO G. TRICOMI: *Tables of integral transforms, Vol. 1.* McGraw-Hill, New York, 1954. 87

[20] WILLIAM FELLER: *An introduction to probability theory and applications*. Volume II. John Wiley and Sons, New York, 1966. 85, 91

[21] DAVID J. H. GARLING: *Inequalities: a journey into linear analysis*. Cambridge Univ. Press, 2007. [doi:10.1017/CBO9780511755217] 79

[22] MICHEL X. GOEMANS AND DAVID P. WILLIAMSON: Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995. Preliminary version in STOC'94. [doi:10.1145/227683.227684] 79, 84

[23] ALEXANDER GROTHENDIECK: Résumé de la théorie métrique des produits tensoriels topologiques (French). *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953. Available at Instituto de Matemática e Estatística da Universidade de São Paulo. 78, 84

[24] UFFE HAAGERUP: A new upper bound for the complex Grothendieck constant. *Israel Journal of Mathematics*, 60(2):199–224, 1987. [doi:10.1007/BF02790792] 78, 83, 85, 88

[25] ADOLF HURWITZ: *Vorlesungen über allgemeine Funktionentheorie und elliptische Funktionen*. Spring, 1925. Available at Göttinger Digitalisierungszentrum. 88

[26] GRAHAM J. O. JAMESON: *Summing and nuclear norms in Banach space theory*. Cambridge Univ. Press, 1987. [doi:10.1017/CBO9780511569166] 79

[27] SUBHASH KHOT AND ASSAF NAOR: Linear equations modulo 2 and the L_1 diameter of convex bodies. *SIAM J. Comput.*, 38(4):1448–1463, 2008. Preliminary version in FOCS'07. [doi:10.1137/070691140] 79

[28] SUBHASH KHOT AND ASSAF NAOR: Approximate kernel clustering. *Mathematika*, 55(1-2):129–165, 2009. Preliminary version in FOCS'08. [doi:10.1112/S002557930000098X] 79

[29] SUBHASH KHOT AND ASSAF NAOR: Sharp kernel clustering algorithms and their associated Grothendieck inequalities. *Random Structures & Algorithms*, 42(3):269–300, 2013. Preliminary version in SODA'10. See also at arXiv. [doi:10.1002/rsa.20398] 79

[30] SUBHASH KHOT AND RYAN O'DONNELL: SDP gaps and UGC-hardness for Max-Cut-Gain. *Theory of Computing*, 5(4):83–117, 2009. Preliminary version in FOCS'06. [doi:10.4086/toc.2009.v005a004] 79, 83

[31] GUY KINDLER, ASSAF NAOR, AND GIDEON SCHECHTMAN: The UGC hardness threshold of the L_p Grothendieck problem. *Math. Oper. Res.*, 35(2):267–283, 2010. Preliminary version in SODA'08. [doi:10.1287/moor.1090.0425] 79

[32] HERMANN KÖNIG: On an extremal problem originating in questions of unconditional convergence. In W. HAUSSMANN, K. JETTER, AND M. REIMER, editors, *Recent Progress in Multivariate Approximation: 4th Internat. Conf., Witten-Bommerholz (Germany), Sep. 2000*, volume 137 of *Internat. Ser. Numer. Math.*, pp. 185–192. Birkhäuser, 2001. [doi:10.1007/978-3-0348-8272-9_14] 85

[33] JEAN-LOUIS KRIVINE: Constantes de Grothendieck et fonctions de type positif sur les sphères. *Advances in Math.*, 31(1):16–30, 1979. [doi:10.1016/0001-8708(79)90017-3] 83, 92

[34] JORAM LINDENSTRAUSS AND ALEKSANDER PEŁCZYŃSKI: Absolutely summing operators in L_p-spaces and their applications. *Studia Math.*, 29(3):275–326, 1968. Available at EuDML. 78

[35] NATI LINIAL AND ADI SHRAIBMAN: Lower bounds in communication complexity based on factorization norms. *Random Structures & Algorithms*, 34(3):368–394, 2009. Preliminary version in STOC'07. [doi:10.1002/rsa.20232] 79

[36] LÁSZLÓ LOVÁSZ: On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979. [doi:10.1109/TIT.1979.1055985] 92

[37] ALEXANDRE MEGRETSKI: Relaxations of quadratic programs in operator theory and system analysis. In *Systems, Approximation, Singular Integral Operators, and Related Topics: International Workshop on Operator Theory and Applications (IWOTA'00)*, volume 129, pp. 365–392. Birkhäuser Basel, 2001. [doi:10.1007/978-3-0348-8362-7_15] 79

[38] ASSAF NAOR AND ODED REGEV: Krivine schemes are optimal. Technical report, 2012. To appear in Proc. AMS. [arXiv:1205.6415] 83

[39] ARKADI NEMIROVSKI, CORNELIS ROOS, AND TAMÁS TERLAKY: On maximization of quadratic form over intersection of ellipsoids with common center. *Mathematical Programming*, 86(3):463–473, 1999. [doi:10.1007/s101070050100] 79

[40] YURII NESTEROV: Semidefinite relaxations and nonconvex quadratic optimization. *Optimization Methods and Software*, 9(1-3):141–160, 1998. [doi:10.1080/10556789808805690] 79

[41] GILLES PISIER: *Factorization of linear operators and geometry of Banach spaces*. Number 60 in CBMS regional conference series. Amer. Math. Soc., Providence, Rhode Island, 1986. 79

[42] PRASAD RAGHAVENDRA AND DAVID STEURER: Towards computing the Grothendieck constant. In *Proc. 20th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'09)*, pp. 525–534. ACM Press, 2009. [ACM:1496770.1496828] 79

[43] JAMES REEDS: A new lower bound on the real Grothendieck constant. Available at author's home page, 1991. 83

[44] ODED REGEV AND BEN TONER: Simulating quantum correlations with finite communication. *SIAM J. Comput.*, 39(4):1562–1580, 2009. Preliminary version in FOCS'07. [doi:10.1137/080723909] 79

[45] HARRY EUGENE STANLEY: Spherical model as the limit of infinite spin dimensionality. *Physical Review*, 176(2):718–722, 1968. [doi:10.1103/PhysRev.176.718] 79

[46] BORIS TSIRELSON: Quantum analogues of the Bell inequalities. The case of two spatially separated domains. *J. Soviet Math.*, 36:557–570, 1987. Translated from Problems of the Theory of Probability Distributions IX, Math. Inst. Steklov (LOMI), 142 (1985), pp. 174–194. (In Russian.) Available at author's homepage. 79, 80

[47] LIEVEN VANDENBERGHE AND STEPHEN BOYD: Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996. [doi:10.1137/1038003] 79

AUTHORS

Jop Briët
Postdoc
Courant Institute
New York University
jop.briet@cims.nyu.edu
http://cims.nyu.edu/~jb4808/

Fernando Mário de Oliveira Filho
Assistant professor
Instituto de Matemática e Estatística
Universidade de São Paulo
fmario@gmail.com
http://www.ime.usp.br/~fmario/

Frank Vallentin
Professor
Mathematisches Institut
Universität zu Köln
frank.vallentin@uni-koeln.de
http://www.mi.uni-koeln.de/opt/frank-vallentin/

ABOUT THE AUTHORS

JOP BRIËT is a postdoc at the Courant Institute of New York University. He received his Ph. D. from C.W.I. and the University of Amsterdam in 2011, under the wonderful supervision Harry Buhrman. He is very grateful to Peter Høyer, who coached him into theoretical computer science and encouraged him to pursue a Ph. D. degree after obtaining B. Sc. and M. Sc. degrees in physics at the University of Calgary. He likes problems that lie in the intersection of theoretical computer science and pure mathematics. He recently started taking yoga classes where he tries handstands and does many downward dogs.

FERNANDO MÁRIO DE OLIVEIRA FILHO is assistant professor at the Department of Computer Science of the Institute of Mathematics and Statistics of the University of São Paulo. In 2009, he received his Ph. D. in mathematics and computer science from the University of Amsterdam after working for 4 years at the Centrum Wiskunde & Informatica under Alexander Schrijver and Frank Vallentin. Past appointments include a postdoc at Tilburg University and a postdoc at the Freie Universität Berlin.

FRANK VALLENTIN is a professor of applied mathematics and computer science at Universität zu Köln. In 2003 he received his Ph. D. in mathematics from Technische Universität München under the supervision of Jürgen Richter-Gebert. Past appointments include assistant and associate professor at Technische Universiteit Delft, postdoc at Centrum Wiskunde & Informatica in Amsterdam and postdoc at the Hebrew University of Jerusalem.