

# GROTHENDIECK INEQUALITIES FOR SEMIDEFINITE PROGRAMS WITH RANK CONSTRAINT

JOP BRIËT, FERNANDO MÁRIO DE OLIVEIRA FILHO, AND FRANK VALLENTIN

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 constrained 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 one application in statistical mechanics: Approximating ground states in the  $n$ -vector model.

## 1. INTRODUCTION

Let  $G = (V, E)$  be a graph with finite vertex set  $V$  and edge set  $E \subseteq \binom{V}{2}$ . Let  $A: V \times V \rightarrow \mathbb{R}$  be a symmetric matrix whose rows and columns are indexed by the vertex set of  $G$ . Let  $r$  be a positive integer. The *graphical Grothendieck problem with rank- $r$  constraint* is the following optimization problem:

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

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 matrices  $A: V \times V \rightarrow \mathbb{R}$  the following inequality holds:

$$(1) \quad \text{SDP}_\infty(G, A) \leq K(r, G) \text{SDP}_r(G, A).$$

Here  $S^\infty$  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  $K(r, G) \geq 1$ . In this paper, we prove new upper bounds for  $K(r, G)$ .

**1.1. Some history.** Inequality (1) is called a *Grothendieck inequality* because it first appeared in the work [17] of Grothendieck on the metric theory of tensor products. More precisely, Grothendieck considered the case  $r = 1$  for 2-chromatic (bipartite) graphs, although in quite a different language. (A *k-chromatic graph* is

---

*Date:* November 4, 2010.

*1991 Mathematics Subject Classification.* 68W25, 90C22.

*Key words and phrases.* Grothendieck inequality,  $n$ -vector model, randomized rounding.

The first author is 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, and by the Dutch BSIK/BRICKS project. The third author was supported by Vidi grant 639.032.917 from the Netherlands Organization for Scientific Research (NWO).

a graph whose chromatic number is  $k$ , i.e., one can color its vertices with  $k$  colors so that adjacent vertices get different colors, but  $k - 1$  colors do not suffice for this.) 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 [28] 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 [1]. Haagerup [18] considered the complex case of Grothendieck's inequality which amounts to the case  $r = 2$ . The higher rank case for bipartite graphs was introduced by Briët, Buhrman, and Toner [7].

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

$$\text{SDP}_r(G, A) = \max \left\{ \sum_{\{u,v\} \in E} A(u,v)X(u,v) : X \in \mathbb{R}_{\geq 0}^{V \times V}, \right. \\ \left. X(u,u) = 1 \text{ for all } u \in V, \right. \\ \left. \text{rank } X \leq r \right\},$$

where  $\mathbb{R}_{\geq 0}^{V \times V}$  is the set of matrices  $A: V \times V \rightarrow \mathbb{R}$  that are positive semidefinite.

On the one hand,  $\text{SDP}_r(G, A)$  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  $\text{SDP}_1(K_{n,n}, A)$  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  $\text{SDP}_\infty(G, A)$ , which is an easy computational problem: Obviously, one has  $\text{SDP}_\infty(G, A) = \text{SDP}_{|V|}(G, A)$  and computing  $\text{SDP}_{|V|}(G, A)$  amounts to solving a semidefinite programming problem (see e.g. Vandenberghe, Boyd [42]). Therefore one may approximate it to any fixed precision in polynomial time by using the ellipsoid method or interior point algorithms.

When  $r = 1$ , inequality (1) measures the integrality gap between the two problems  $\text{SDP}_r$  and  $\text{SDP}_\infty$ . For  $r > 1$ , we refer to this gap as the *dimensionality gap*. In many cases the optimal constant  $K(r, G)$  is not known and so one is interested in finding upper bounds of for  $K(r, G)$ . Usually, proving an upper bound amounts to giving a randomized polynomial-time approximation algorithm for  $\text{SDP}_r(G, A)$ . In the case of the MAX CUT problem, Goemans and Williamson [16] pioneered an approach based on randomized rounding: One rounds an optimal solution of  $\text{SDP}_\infty(G, A)$  to a feasible solution of  $\text{SDP}_r(G, A)$ . 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 [16] showed that for all symmetric matrices  $A: V \times V \rightarrow \mathbb{R}$  which have the properties  $A(u, v) \leq 0$  for  $u$  distinct from  $v$  and  $\sum_{u \in V} A(u, v) = 0$  for all  $v \in V$ , we have the inequality

$$\text{SDP}_1(K_{n,n}, A) \leq (0.878 \dots)^{-1} \text{SDP}_\infty(K_{n,n}, A).$$

**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 [35], Jameson [19], and Garling [15].

During the last years, especially after Alon and Naor [2] 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 [34], Nemirovski, Roos, Terlaky [33], Megretski [30]), extremal combinatorics (Alon, Naor [2]), system theory (Bental, Nemirovski [6]), machine learning (Charikar, Wirth [10], Khot, Naor [20, 21]), communication complexity (Linial, Shraibman [27]), quantum information theory (Tsirel'son [41], Regev, Toner [38]), and computational complexity (Khot, O'Donnell [23], Arora, Berger, Kindler, Safra, Hazan [4], Khot and Naor [22], Raghavendra, Steurer [36]).

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 [7]) and in statistical mechanics (mentioned in Alon, Makarychev, Makarychev, Naor [1], Kindler, Naor, Schechtman [24]) the more geometrical case when  $r > 1$  is of interest. This case is the subject of this paper.

Before we present our results we consider the application to statistical mechanics: The *n-vector model*, introduced by Stanley [40], describes the interaction of particles in a spin glass with ferromagnetic and antiferromagnetic interactions. The case  $n = 1$  corresponds to the Ising model, the case  $n = 2$  to the XY model, the case  $n = 3$  to the Heisenberg model, and the case  $n = \infty$  to the Berlin-Kac spherical model.

Let  $G = (V, E)$  be the interaction graph where the vertices are particles and where edges indicate which particles interact. The potential function  $A: V \times V \rightarrow \mathbb{R}$  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 antiferromagnetic interaction. The particles possess a vector-valued spin  $f: V \rightarrow S^{n-1}$ . In the absence of an external field, the total energy of the system is given by the *Hamiltonian*

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

The ground state of this model is a configuration of spins  $f: V \rightarrow S^{n-1}$  which minimizes the total energy. Finding the ground state is the same as solving  $\text{SDP}_n(G, A)$ . 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 briefly mention the relation to quantum information theory. In quantum information theory, *XOR games*, first considered by Clauser, Horne, Shimony, and Holt [11], constitute the simplest model in which the predictions of quantum mechanics can be contrasted with those of classical mechanics, both theoretically and experimentally. In an XOR game, two players, Alice and Bob, receive questions  $u$  and  $v$  respectively, picked by a referee according to some probability distribution

$\pi(u, v)$  which is known to everybody in advance. Without communicating with each other, they answer the referee with bits  $a, b \in \{0, 1\}$ . Alice and Bob win the game if the exclusive-OR (XOR) of their answers  $a \oplus b$  equals the value of a Boolean function  $g(u, v)$  also known in advance, otherwise they lose.

In a quantum-mechanical setting, the players may use *entangled strategies*: The players “share” a positive semidefinite operator  $\rho \in \mathbb{C}_{\geq 0}^{d^2 \times d^2}$  which represents the state of a pair of  $d$ -dimensional entangled quantum systems, and for each question  $u \in V$ , Alice has a pair of  $d$ -by- $d$  positive semidefinite matrices  $\{A_u^0, A_u^1\}$  that satisfy  $A_u^0 + A_u^1 = I$ , where  $I$  is the  $d$ -by- $d$  identity matrix, and for each question  $v \in V$ , Bob has a similar pair of matrices  $\{B_v^0, B_v^1\}$ . To determine her answer to question  $u$ , Alice performs a *measurement* corresponding to the matrices associated to her question. The probability that she measures  $a \in \{0, 1\}$  equals  $\text{trace}(\rho A_u^a \otimes I)$ . Similarly, the probability that Bob measures  $b$  equals  $\text{trace}(\rho I \otimes B_v^b)$ .

Tsirel’son [41] proved that for  $G = (V_A, V_B; E)$  the complete bipartite graph with Alice and Bob’s questions on opposite sides of the partition, and  $A(u, v) = \frac{1}{2}\pi(u, v) \cdot (-1)^{g(u, v)}$  for all  $\{u, v\} \in V_A \times V_B$  and  $A(u, v) = 0$  otherwise, the difference between the probability of winning and the probability of losing when the players use optimal strategies for fixed dimension  $d$ , is bounded from above by  $\text{SDP}_r(G, A)$  with  $r = 2d$  and bounded from below by  $\text{SDP}_r(G, A)$  with  $r = \lfloor \log d \rfloor$ .

**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. The proof of the following theorem gives a randomized polynomial-time approximation algorithm for approximating ground states in the Heisenberg model in the lattice  $\mathbb{Z}^3$  with approximation ratio  $0.78\dots = (1.28\dots)^{-1}$ . This result can be regarded as the principal contribution of this paper.

**Theorem 1.1.** *For  $r = 1, \dots, 10$  and in the case of a bipartite or a tripartite graph  $G$  the rank- $r$  Grothendieck constant is at most:*

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

The bound for the original Grothendieck constant  $K(1, G)$  is due to Krivine [26]. Currently, it is the best known upper bound, the best known lower bound is  $1.676956\dots$  due to Davie [12] and Reeds [37] (see also Khot and O’Donnell [23]). The bound for  $K(2, G)$  is due to Haagerup [18].

When the graph  $G$  has large chromatic number, then the result of Alon, Makarychev, Makarychev, and Naor [1] gives 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 method only gives a linear dependence.

For the proof of Theorem 1.1 we use the framework of Krivine and Haagerup which we explain below. Our main technical contributions are a matrix version of Grothendieck's identity given in Lemma 2.1. Furthermore in Lemma 4.1 we provide a method to construct new unit vectors linearizing an expectation which can also deal with nonbipartite graphs.

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

**Lemma 1.2.** *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  $f: V \rightarrow S^{|V|-1}$ , there is a function  $g: V \rightarrow S^{|V|-1}$  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 give the precise statement of the lemma in Section 4 (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  $\text{SDP}_r(G, A)$ :

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

- (1) Solve  $\text{SDP}_\infty(G, A)$ , obtaining an optimal solution  $f: V \rightarrow S^{|V|-1}$ .
- (2) Use  $f$  to construct  $g: V \rightarrow S^{|V|-1}$  according to Lemma 1.2.
- (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  $h: V \rightarrow S^{r-1}$  by setting  $h(u) = Zg(u)/\|Zg(u)\|$ .

To analyze this procedure, we compute the expected value of the feasible solution  $h$ . Using Lemma 1.2 we obtain

$$\begin{aligned} \text{SDP}_r(G, A) &\geq \mathbb{E} \left[ \sum_{\{u,v\} \in E} A(u, v) h(u) \cdot h(v) \right] \\ &= \sum_{\{u,v\} \in E} A(u, v) \mathbb{E}[h(u) \cdot h(v)] \\ (2) \quad &= \beta(r, G) \sum_{\{u,v\} \in E} A(u, v) f(u) \cdot f(v) \\ &= \beta(r, G) \text{SDP}_\infty(G, A), \end{aligned}$$

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  $\text{SDP}_r(G, A)$ .

The constant  $\beta(r, G)$  in Lemma 1.2 is defined in terms of the Taylor expansion of the inverse of the function  $E_r: [-1, 1] \rightarrow [-1, 1]$  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 Taylor expansion of  $E_r$  is computed in Section 2. The Taylor expansion of  $E_r^{-1}$  is treated in Section 3, where we basically follow Haagerup [18]. A precise version of Lemma 1.2 is stated and proved in Section 4, following Krivine [26].

Finally, in Section 5 we show that one can refine this analysis and can (strictly) improve the upper bound if one takes the dimension of the matrix  $A: V \times V \rightarrow \mathbb{R}$  into account. In particular, we compare the problems  $\text{SDP}_q$  and  $\text{SDP}_r$  for  $q \geq r$ . Earlier, Avidor and Zwick [5] considered the problem of bounding the ratio  $\text{SDP}_r(G, A)/\text{SDP}_1(G, A)$  for  $r = 2, 3$  and  $A$  the Laplacian matrix of a graph.

## 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,

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

For instance, the celebrated algorithm of Goemans and Williamson [16] 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 [17, Proposition 4, p. 63] (see also Diestel, Fourie, and Swart [13]).

In this section we extend Grothendieck’s identity from vectors to matrices by replacing the arcsine function by a hypergeometric function.

**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,*

$$\mathbb{E} \left[ \frac{Zu}{\|Zu\|} \cdot \frac{Zv}{\|Zv\|} \right] = \frac{2}{r} \left( \frac{\Gamma((r+1)/2)}{\Gamma(r/2)} \right)^2 (u \cdot v) {}_2F_1 \left( \frac{1}{2}, 1/2; r/2 + 1; (u \cdot v)^2 \right).$$

Here,

$$\begin{aligned} & (u \cdot v) {}_2F_1 \left( \frac{1}{2}, \frac{1}{2}; \frac{r}{2} + 1; (u \cdot v)^2 \right) \\ &= \sum_{k=0}^{\infty} \frac{(1 \cdot 3 \cdots (2k-1))^2}{(2 \cdot 4 \cdots 2k)((r+2) \cdot (r+4) \cdots (r+2k))} (u \cdot v)^{2k+1}. \end{aligned}$$

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

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

The case  $r = 2$  is due to Haagerup [18]:

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

where  $K$  and  $E$  are the complete elliptic integrals of the first and second kind. Note that on page 201 of Haagerup [18]  $\pi/2$  has to be  $\pi/4$ . Briët, Oliveira, and Vallentin [8] computed the first coefficient  $2/r(\Gamma((r+1)/2)/\Gamma(r/2))^2$  of the Taylor series of the expectation for every  $r$ .

The following elegant proof of Grothendieck's identity has become a classic: We have  $\text{sign}(Zu) \text{sign}(Zv) = 1$  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 regions on the unit circle: The probability that  $u$  and  $v$  lie on the same side of the line is  $1 - \arccos(u \cdot v)/\pi$ .

However, 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 [25]. It would be interesting to find a more geometrical proof of the lemma.

For computing the first coefficient of the Taylor series in [8] we took a slightly different route: We integrated using the Wishart distribution of  $2 \times 2$ -matrices.

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

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

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  $u = (1, 0, \dots, 0)$

and  $v = (t, \sqrt{1-t^2}, 0, \dots, 0)$  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. Muirhead [32, p. 10, eq. (7)])

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

Hence,

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

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

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

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

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

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

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

The last integral can be rewritten using the modified Bessel function of the first kind (cf. Andrews, Askey, Roy [3, p. 235, Exercise 9])

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



One can write  $I_{r/2}$  as a hypergeometric function (cf. Andrews, Askey, and Roy [3, (4.12.2)])

$$I_{r/2}(x) = (x/2)^{r/2} \sum_{k=0}^{\infty} \frac{(x/2)^{2k}}{k! \Gamma(r/2 + k + 1)} = \frac{(x/2)^{r/2}}{\Gamma((r+2)/2)} {}_0F_1 \left( \frac{\text{---}}{(r+2)/2}; \left(\frac{x}{2}\right)^2 \right).$$

Putting things together, we get

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

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

Now we evaluate the outer double integral

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

Here the inner integral equals

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

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

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

which is by the Bateman Manuscript Project [14, p. 337 (11)] equal to

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

Now we treat the remaining outer integral in a similar way, using [14, p. 219 (17)], and get that

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

By applying Euler's transformation (cf. Andrews, Askey, and Roy [3, (2.2.7)])

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

and after collecting the remaining factors we arrive at the result.  $\square$

## 3. CONVERGENCE RADIUS

To construct the new vectors in the third step of the algorithm that are used to linearize the expectation 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

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

but in the proof of Lemma 4.1 it will be essential that this expansion be valid for all  $t \in [-1, 1]$ .

In the case  $r = 1$  we have  $E_1^{-1}(t) = \sin(\pi/2t)$  and here the convergence radius is even infinity. The case  $r = 2$  was treated by Haagerup and it requires quite some technical work which we sketch very briefly now. He shows that  $|b_k| \leq C/k^2$  for some constant  $C$ , independent of  $k$ , using tools from complex analysis. Using Cauchy's integral formula and after doing some simplifications [18, p. 208] one can express  $b_k$  as

$$b_k = \frac{2}{\pi k} \int_1^\alpha \Im(E_2(z)^{-k}) dz + \frac{2}{\pi k} \Im \left( \int_{C'_\alpha} E_2(z)^{-k} dz \right),$$

where  $C'_\alpha$  is the quarter circle  $\{\alpha e^{i\theta} : \theta \in [0, \pi/2]\}$ .

For an appropriate choice of  $\alpha$  the first integral is in absolute value bounded above by  $C/k$  and the second integral is in absolute value exponentially small in  $k$ . We refer to the original paper for the details. One key point in the arguments is the following integral representation of  $E_2$  giving an analytic continuation of  $E_2$  on the complex plane slit along the half line  $(1, \infty)$ :

$$E_2(z) = \int_0^{\pi/2} \sin \theta \arcsin(z \sin \theta) d\theta.$$

Here, the term  $\arcsin(z \sin \theta)$  gives the main contribution in the estimates.

Now we derive a similar representation of  $E_r$  and using it in Haagerup's analysis with obvious changes shows that also for  $r > 2$  we have  $b_k \leq C/k^2$  for some constant  $C$ , independent of  $k$ .

**Lemma 3.1.** *For  $r \geq 2$  we have*

$$E_r(z) = \frac{2(r-1)\Gamma((r+1)/2)}{\Gamma(1/2)\Gamma(r/2)} \int_0^{\pi/2} \cos^{r-2} \theta \sin \theta \arcsin(z \sin \theta) d\theta.$$

*Proof.* Using Euler's integral representation of the hypergeometric function (cf. Andrews, Askey, and Roy [3, Theorem 2.2.1]) we can rewrite  $E_r$  as

$$E_r(z) = \frac{\Gamma((r+1)/2)}{\Gamma(1/2)\Gamma(r/2)} \int_0^1 \frac{(1-t)^{(r-1)/2} z}{\sqrt{t(1-z^2t)}} dt,$$

which is valid in the complex plane slit along the half line  $(1, \infty)$ . Using the substitution  $t = \sin^2 \theta$  we get

$$E_r(z) = 2 \frac{\Gamma((r+1)/2)}{\Gamma(1/2)\Gamma(r/2)} \int_0^{\pi/2} \frac{\cos^r \theta z}{\sqrt{1-z^2 \sin^2 \theta}} d\theta.$$

Now integration by parts and the identity

$$\frac{d}{d\theta} \arcsin(z \sin \theta) = \frac{z \cos \theta}{\sqrt{1 - z^2 \sin^2 \theta}}$$

gives the result.  $\square$

#### 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 1.2; this is done in Lemma 4.1. For this we follow Krivine [26], 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 [29]. 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:

$$(3) \quad \vartheta(\overline{G}) = \min \left\{ \lambda : Z \in \mathbb{R}_{\geq 0}^{V \times V}, \right. \\ \left. \begin{aligned} Z(u, u) &= \lambda - 1 \text{ for } u \in V, \\ Z(u, v) &= -1 \text{ for } \{u, v\} \in E \end{aligned} \right\}.$$

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 in 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 (3).

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

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

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

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

where

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

With this lemma, we can give a proof of Theorem 1.1.

*Proof of Theorem 1.1.* We combine Lemma 4.1 with the analysis of Algorithm A from Section 1.4. To compute the table in the theorem, we use the formula

$$b_k = \frac{1}{k! a_1^k} \left[ \frac{d^{k-1}}{dt^{k-1}} \left( 1 + \frac{a_2}{a_1} t + \dots + \frac{a_k}{a_1} t^{k-1} \right)^{-k} \right]_{t=0},$$

where  $a_i$  are the Taylor coefficients of  $E_r$  (cf. Morse and Feshbach [31, (4.5.13)]).  $\square$

Now we give a proof of the lemma.

*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

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

Then for 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 (3). We have  $\lambda \geq 2$  since  $G$  has at least one edge. Set

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

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  $s: V \rightarrow \mathbb{R}^{2|V|}$  and  $t: V \rightarrow \mathbb{R}^{2|V|}$  so that

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

The matrix  $U$  is the Gram matrix of the vectors  $(s(u))_{u \in V}$  and  $(t(v))_{v \in V}$ . 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  $(R(u) \cdot R(v)) \in \mathbb{R}_{\geq 0}^{V \times V}$  to define the desired function  $g: V \rightarrow S^{|V|-1}$ .  $\square$

We conclude this section with a few remarks on the lemma and its proof:

- (1) Krivine proved the statement of the lemma in the case  $r = 1$  and for bipartite graphs  $G$ . Then,  $\vartheta(\overline{G}) = 2$  holds. In this case one has various simplifications: 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,  $\beta(1, G) = 2 \operatorname{arcsinh}(1)/\pi = 2 \ln(1 + \sqrt{2})/\pi$ .

- (2) In the second step one can also work with any feasible solution of the semidefinite program (3). For instance one can replace  $\vartheta(\overline{G})$  in the lemma by the chromatic number  $\chi(G)$  albeit getting a potentially weaker bound.
- (3) Alon, Makarychev, Makarychev, and Naor [1] also provide an upper bound for  $K(1, G)$  using the theta number of the complement of  $G$ . They show 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.

- (4) Finally, notice that in the first step it was essential that the Taylor expansion of  $E_r^{-1}$  has convergence radius of at least one.

## 5. A REFINED, DIMENSION-DEPENDENT ANALYSIS

So far we only compared the two problems  $\operatorname{SDP}_{\infty}$  and  $\operatorname{SDP}_r$ . One can perform a refined, dimension-dependent analysis by comparing  $\operatorname{SDP}_q$  and  $\operatorname{SDP}_r$  when  $q \geq r$ . This is of interest because for instance  $\operatorname{SDP}_{\infty}(G, A) = \operatorname{SDP}_{|V|}(G, A)$ .

Let  $K(q \mapsto r, G)$ , where  $q \geq r$ , be the least number such that

$$\operatorname{SDP}_q(G, A) \leq K(q \mapsto r, G) \operatorname{SDP}_r(G, A)$$

for all matrices  $A: V \times V \rightarrow \mathbb{R}$ . In this section we give an upper bound for  $K(q \mapsto r, G)$  that depends on  $q$  and  $r$ . For fixed  $r$ , this upper bound will become smaller as  $q$  comes closer to  $r$ . Krivine [26] gave such a refined, dimension-dependent analysis in the bipartite case. Our contribution is that we show that the upper bound for  $K(q \mapsto r, G)$  is *strictly* smaller than the upper bound for  $K(q+1 \mapsto r, G)$ . If  $K(\infty \mapsto 1, G) = \pi/(2 \ln(1 + \sqrt{2}))$  in the case of bipartite  $G$ , then our result would imply a conjecture of Brunner, Pironio, Acin, Gisin, Méthot, and Scarani [9]: Two-outcome measurements are sufficient to test the dimension of any bipartite quantum system.

Our upper bound comes from the following lemma:

**Lemma 5.1.** *Let  $G = (V, E)$  be a graph with at least one edge. Given  $f: V \rightarrow S^{q-1}$ , there is a function  $g: V \rightarrow S^{|V|-1}$  such that whenever  $u$  and  $v$  are adjacent, then*

$$E_r(g(u) \cdot g(v)) = \beta(q \mapsto r, G) f(u) \cdot f(v),$$

where  $0 < \beta(q \mapsto r, G) \leq 1$  is such that  $\beta(q \mapsto r, G) > \beta(q+1 \mapsto r, G)$  and  $\beta(q \mapsto r, G) > \beta(r, G)$  for all  $q \geq 2$ .

So we have the theorem:

**Theorem 5.2.** *Let  $G = (V, E)$  be a graph with at least one edge and let  $q \geq r \geq 1$  be integers. Then  $K(q \mapsto r, G) \leq \beta(q \mapsto r, G)^{-1}$ .*

*Proof.* Combine Lemma 5.1 with Algorithm A from Section 1.4.  $\square$

The proof of the lemma uses some basic facts from harmonic analysis, which we now summarize. For measurable functions  $f, g: [-1, 1] \rightarrow \mathbb{R}$  we consider the inner product

$$(4) \quad \langle f, g \rangle_n = \int_{-1}^1 f(t)g(t)(1-t^2)^{(n-3)/2} dt.$$

We say that a continuous function  $f: [-1, 1] \rightarrow \mathbb{R}$  is of *positive type for  $S^{n-1}$*  if for any choice  $x_1, \dots, x_N$  of points in  $S^{n-1}$  we have that the matrix  $(f(x_i \cdot x_j))_{i,j=1}^N$  is positive semidefinite. If two continuous functions  $f, g: [-1, 1] \rightarrow \mathbb{R}$  are of positive type for  $S^{n-1}$ , then  $\langle f, g \rangle_n \geq 0$ .

Schoenberg [39] characterized the continuous functions of positive type in terms of Gegenbauer polynomials. We denote by  $P_k^n$  the Gegenbauer polynomial of degree  $k$  and parameter  $(n-2)/2$  which is normalized so that  $P_k^n(1) = 1$ . Notice that this normalization is not the one commonly found in the literature.

The Gegenbauer polynomials  $P_0^n, P_1^n, P_2^n, \dots$  are pairwise orthogonal with respect to the inner product (4), and they form a complete orthogonal system for the space  $L^2([-1, 1])$ , equipped with the inner product (4).

Schoenberg's characterization of the functions of positive type is as follows: A function  $f: [-1, 1] \rightarrow \mathbb{R}$  is continuous and of positive type for  $S^{n-1}$  if and only if

$$(5) \quad f(t) = \sum_{k=0}^{\infty} a_k P_k^n(t)$$

for some nonnegative numbers  $a_0, a_1, a_2, \dots$  such that  $\sum_{k=0}^{\infty} a_k$  converges, in which case the series in (5) converges absolutely and uniformly in  $[-1, 1]$ .

A continuous function  $f: [-1, 1] \rightarrow \mathbb{R}$  can also be of positive type for spheres of every dimension. Schoenberg [39] also characterized these functions. They are the ones that can be decomposed as

$$f(t) = \sum_{k=0}^{\infty} a_k t^k$$

for some nonnegative numbers  $a_0, a_1, a_2, \dots$  such that  $\sum_{k=0}^{\infty} a_k$  converges.

A polynomial in  $\mathbb{R}[x_1, \dots, x_n]$  is *harmonic* if it is homogeneous and vanishes under the Laplace operator  $\Delta = \partial^2/\partial x_1^2 + \dots + \partial^2/\partial x_n^2$ . Harmonic polynomials are related to Gegenbauer polynomials by the *addition formula* (see e.g. Andrews, Askey, and Roy [3, Theorem 9.6.3]): Let  $H_k$  be the space of degree  $k$  harmonic polynomials on  $n$  variables. Any orthonormal basis of  $H_k$  can be scaled so as to give a basis  $e_{k,1}, \dots, e_{k,h_k}$  of  $H_k$  for which the following holds: For every  $u, v \in S^{n-1}$  we have that

$$P_k^n(u \cdot v) = \sum_{i=1}^{h_k} e_{k,i}(u)e_{k,i}(v).$$

With this we have all that we need to prove the lemma. We only consider the bipartite case in the proof in order to simplify the notation and to make the argument more transparent. One can handle the nonbipartite case exactly in the same way as in the proof of Lemma 4.1.

*Proof of Lemma 5.1.* Fix  $0 < \beta \leq 1$  and consider the expansion

$$E_r^{-1}(\beta t) = \sum_{k=0}^{\infty} g_k^q(\beta) P_k^q(t),$$

which converges in the  $L^2$  sense. We claim that  $\sum_{k=0}^{\infty} |g_k^q(\beta)|$  converges for every  $0 < \beta \leq 1$ , and hence the above expansion converges absolutely and uniformly for  $t \in [-1, 1]$ .

To see the claim, consider the expansion  $E_r^{-1}(t) = \sum_{k=0}^{\infty} b_k t^k$  and recall that  $\sum_{k=0}^{\infty} |b_k|$  converges. We may define the function  $\bar{E}_r^{-1}(t) = \sum_{k=0}^{\infty} |b_k| t^k$ , which is then of positive type for every sphere. So by Schoenberg's theorem we can write

$$\bar{E}_r^{-1}(t) = \sum_{k=0}^{\infty} \bar{g}_k^q P_k^q(t)$$

for nonnegative numbers  $\bar{g}_k^q$  such that  $\sum_{k=0}^{\infty} \bar{g}_k^q$  converges. Now notice that

$$g_k^q(\beta) = \|P_k^q\|_q^{-2} \langle E_r^{-1}(\beta t), P_k^q \rangle_q = \|P_k^q\|_q^{-2} \sum_{l=0}^{\infty} b_l \beta^l \langle t^l, P_k^q \rangle_q,$$

where  $\|P_k^q\|_q = \langle P_k^q, P_k^q \rangle_q^{1/2}$ . Above, since  $t^l$  is a function of positive type for every sphere, we have that  $\langle t^l, P_k^q \rangle_q \geq 0$ . But we also have that

$$\bar{g}_k^q = \|P_k^q\|_q^{-2} \langle \bar{E}_r^{-1}, P_k^q \rangle_q = \|P_k^q\|_q^{-2} \sum_{l=0}^{\infty} |b_l| \langle t^l, P_k^q \rangle_q,$$

and we see that  $|g_k^q(\beta)| \leq \bar{g}_k^q$  for all  $k \geq 0$  and  $0 < \beta \leq 1$ . This finishes the proof of the claim.

From the claim, it also follows that the function

$$(6) \quad \sum_{k=0}^{\infty} |g_k^q(\beta)|$$

is a continuous function of  $\beta$ .

Now, let  $\beta(q \mapsto r, G)$  be the maximum number in  $(0, 1]$  that is such that

$$\sum_{k=0}^{\infty} |g_k^q(\beta(q \mapsto r, G))| = 1.$$

Such a number exists because (6) is continuous as a function of  $\beta$ , being equal to 0 when  $\beta = 0$  and at least  $E_r^{-1}(1) = 1$  when  $\beta = 1$ .

Consider the Hilbert space

$$\mathcal{H} = \bigoplus_{k=0}^{\infty} \mathbb{R}^{h_k},$$

equipped with the Euclidean inner product and where  $h_k$  is the dimension of  $H_k$ , the space of harmonic polynomials of degree  $k$  on  $q$  variables. For a vector  $x \in S^{q-1}$ ,

consider the vectors  $S(x)$  and  $T(x) \in \mathcal{H}$  given componentwise by

$$\begin{aligned} S(x)_k &= \sqrt{|g_k^q(\beta(q \mapsto r, G))|} (e_{k,1}(x), \dots, e_{k,h_k}(x)) \quad \text{and} \\ T(x)_k &= \text{sign}(g_k^q(\beta(q \mapsto r, G))) \sqrt{|g_k^q(\beta(q \mapsto r, G))|} (e_{k,1}(x), \dots, e_{k,h_k}(x)). \end{aligned}$$

By the addition formula we have that

$$S(f(u)) \cdot T(f(v)) = E_r^{-1}(\beta(q \mapsto r, G)) f(u) \cdot f(v).$$

Moreover, we also have that

$$\|S(f(u))\|^2 = \|T(f(v))\|^2 = \sum_{k=0}^{\infty} |g_k^q(\beta(q \mapsto r, G))| = 1,$$

and so from the Gram matrix of the vectors  $S(f(u))$  and  $T(f(v))$  we may obtain the function  $g: V \rightarrow S^{|V|-1}$  as we wanted.

Now we show that  $\beta(q \mapsto r, G) > \beta(q+1 \mapsto r, G)$  for all  $q \geq 2$ . To this end, consider the function

$$F(\beta, t) = \sum_{k=0}^{\infty} |g_k^{q+1}(\beta)| P_k^{q+1}(t).$$

Since  $\sum_{k=0}^{\infty} |g_k^{q+1}(\beta)|$  converges, from Schoenberg's theorem we see that  $F$  is a continuous function of positive type for the sphere  $S^q$ . Notice moreover that, by definition,  $\beta(q+1 \mapsto r, G)$  is the maximum number in  $(0, 1]$  such that  $F(\beta(q+1 \mapsto r, G), 1) = 1$ .

Since  $F$  is of positive type for  $S^q$ , it is also of positive type for  $S^{q-1}$ , and then we may write

$$F(\beta, t) = \sum_{k=0}^{\infty} a_k(\beta) P_k^q(t),$$

and we have that  $\sum_{k=0}^{\infty} a_k(\beta(q+1 \mapsto r, G)) = 1$ . We also have the expression

$$(7) \quad a_k(\beta) = \|P_k^q\|_q^{-2} \langle F(\beta, t), P_k^q \rangle_q = \|P_k^q\|_q^{-2} \sum_{l=0}^{\infty} |g_l^{q+1}(\beta)| \langle P_l^{q+1}, P_k^q \rangle_q.$$

Notice that, since both  $P_l^{q+1}$  and  $P_k^q$  are of positive type for  $S^{q-1}$ , we have that  $\langle P_l^{q+1}, P_k^q \rangle_q \geq 0$  for all  $l$  and  $k$ .

Now, from the expansion

$$E_r^{-1}(\beta t) = \sum_{k=0}^{\infty} g_k^{q+1}(\beta) P_k^{q+1}(t)$$

we see that

$$(8) \quad g_k^q(\beta) = \|P_k^q\|_q^{-2} \langle E_r^{-1}(\beta t), P_k^q \rangle_q = \|P_k^q\|_q^{-2} \sum_{l=0}^{\infty} g_l^{q+1}(\beta) \langle P_l^{q+1}, P_k^q \rangle_q.$$

It is not hard to show that  $E_r^{-1}$  is not of positive type, and so some of the  $g_l^{q+1}(\beta)$  must be negative. This, together with (7) and (8), implies that  $|g_k^q(\beta)| < a_k(\beta)$  for all  $0 < \beta \leq 1$ . So we must have that

$$\sum_{k=0}^{\infty} |g_k^q(\beta(q+1 \mapsto r, G))| < \sum_{k=0}^{\infty} a_k(\beta(q+1 \mapsto r, G)) = 1,$$



and we see that  $\beta(q \mapsto r, G) > \beta(q + 1 \mapsto r, G)$ . In a similar way, one may show that  $\beta(q \mapsto r, G) > \beta(r, G)$ .  $\square$

#### ACKNOWLEDGEMENTS

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.

#### REFERENCES

- [1] N. Alon, K. Makarychev, Y. Makarychev, A. Naor, *Quadratic forms on graphs*, Invent. Math. **163** (2006), 499–522.
- [2] N. Alon, A. Naor, *Approximating the cut-norm via Grothendieck’s inequality*, SIAM J. Comp. **35** (2006), 787–803.
- [3] G.E. Andrews, R. Askey, R. Roy, *Special functions*, Cambridge University Press, 1999.
- [4] S. Arora, E. Berger, G. Kindler, E. Hazan, S. Safra, *On non-approximability for quadratic programs*, pp. 205–215 in 46th Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press, 2005.
- [5] A. Avidor, U. Zwick, *Rounding two and three dimensional solutions of the SDP relaxation of MAX CUT*, pp. 14–25 in 8th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2005 and 9th International Workshop on Randomization and Computation, RANDOM, LNCS 3624, 2005.
- [6] A. Ben-Tal, A. Nemirovski, *On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty*, SIAM J. Optim. **12** (2002), 811–833
- [7] J. Briët, H. Buhrman, B. Toner, *A generalized Grothendieck inequality and entanglement in XOR games*, arXiv:0901.2009 [quant-ph], January 2009.
- [8] J. Briët, F.M. de Oliveira Filho, F. Vallentin, *The positive semidefinite Grothendieck problem with rank constraint*, pp. 31–42 in: *Proceedings of the 37th International Colloquium on Automata, Languages and Programming, ICALP 2010 (S. Abramsky et al. eds.), Part I*, LNCS 6198, 2010.
- [9] N. Brunner, S. Pironio, A. Acin, N. Gisin, A.A. Méthot, V. Scarani *Testing the Dimension of Hilbert Spaces*, Phys. Rev. Lett. **100**, (2008), 210503, 4 pages
- [10] M. Charikar, A. Wirth, *Maximizing quadratic programs: extending Grothendieck’s inequality*, pp. 54–60 in 45th Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press, 2004.
- [11] J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett. **23** (1969), 880–884.
- [12] A. M. Davie, *Lower bound for  $K_G$* , unpublished note (1984).
- [13] J. Diestel, J.H. Fourie, J. Swart, *The metric theory of tensor products — Grothendieck’s Résumé revisited*, American Mathematical Society, 2008.
- [14] A. Erdélyi, W. Magnus, F. Oberhettinger, F.G. Tricomi, *Tables of Integral Transforms, Vol. I*, McGraw-Hill, 1954.
- [15] D.J.H. Garling, *Inequalities: a journey into linear analysis*, Cambridge University Press, 2007.
- [16] M.X. Goemans, D.P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. ACM **42** (1995), 1115–1145.
- [17] A. Grothendieck, *Résumé de la théorie métrique des produits tensoriels topologiques*, Bol. Soc. Mat. São Paulo **8** (1953), 1–79.
- [18] U. Haagerup, *A new upper bound for the complex Grothendieck constant*, Israel J. Math. **60** (1987), 199–224.
- [19] G.J.O. Jameson, *Summing and Nuclear Norms in Banach Space Theory*, Cambridge University Press, 1987.
- [20] S. Khot, A. Naor, *Approximate kernel clustering*, pp. 561–570 in *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, IEEE Computer Society, 2008.

- [21] S. Khot, A. Naor, *Sharp kernel clustering algorithms and their associated Grothendieck inequalities*, arXiv:0906.4816 [cs.DS], June 2009.
- [22] S. Khot, A. Naor, *Linear Equations Modulo 2 and the L1 Diameter of Convex Bodies*, pp. 318–328 in *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, IEEE Computer Society, 2007.
- [23] S. Khot, R. O’Donnell, *SDP gaps and UGC-hardness for max-cut-gain*, *Theory Comput.* **5** (2009), 83–117.
- [24] G. Kindler, A. Naor, G. Schechtman, *The UGC hardness threshold of the  $l_p$  Grothendieck problem*, pp. 64–73 in *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ACM, 2008.
- [25] H. König, *On an extremal problem originating in questions of unconditional convergences*, *Intern. Ser. Numer. Math.* **137** (2001), 185–192.
- [26] J.-L. Krivine, *Constantes de Grothendieck et fonctions de type positif sur les sphères*, *Adv. Math.* **31** (1979), 16–30.
- [27] N. Linial, A. Shraibman, *Lower bounds in communication complexity based on factorization norms*, pp. 699–708 in *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, 2007.
- [28] J. Lindenstrauss, A. Pełczyński, *Absolutely summing operators in  $L_p$ -spaces and their applications*, *Studia Math.* **29** (1968), 275–326.
- [29] L. Lovász, *On the Shannon capacity of a graph*, *IEEE Trans. Inf. Th.* **25** (1979), 1–7.
- [30] A. Megretski, *Relaxations of quadratic programs in operator theory and system analysis*, *Oper. Theory Adv. Appl.* **129** (2001), 365–392.
- [31] P. M. Morse, H. Feshbach, *Methods of Theoretical Physics, Part I*, McGraw-Hill, 1953.
- [32] R.J. Muirhead, *Aspects of multivariate statistical theory*, Jon Wiley & Sons, 1982.
- [33] A. Nemirovski, C. Roos, T. Terlaky, *On maximization of quadratic form over intersection of ellipsoids with common center*, *Math. Program., Ser. A* **86** (1999), 463–473.
- [34] Y.E. Nesterov, *Semidefinite relaxation and nonconvex quadratic optimization*, *Optimization Methods and Software* **9** (1998), 141–160.
- [35] G. Pisier, *Factorization of linear operators and geometry of Banach spaces*, CBMS Regional Conference Series in Mathematics, 60, AMS, 1986.
- [36] P. Raghavendra, D. Steurer, *Towards computing the Grothendieck constant*, pp. 525–534 in *ACM-SIAM Symposium on Discrete Algorithms*, 2009.
- [37] J. A. Reeds, *A new lower bound on the real Grothendieck constant*, unpublished note, available at <http://www.dtc.umn.edu/reedsj/bound2.dvi> (1991).
- [38] O. Regev, B. Toner, *Simulating quantum correlations with finite communication*, *SIAM J. Comput.* **39** (2009/10), 1562–1580.
- [39] I.J. Schoenberg, *Positive definite functions on spheres*, *Duke Math. J.* **9** (1942), 96–108.
- [40] H.E. Stanley, *Spherical model as the limit of infinite spin dimensionality*, *Phys. Rev.* **176** (1968), 718–721.
- [41] B. Tsirel’son, *Quantum analogues of the Bell inequalities. The case of two spatially separated domains*, *J. Soviet Math.* **36**, 557–570.
- [42] L. Vandenberghe, S. Boyd, *Semidefinite Programming*, *SIAM Review* **38** (1996), 49–95.

J. BRIËT, CENTRUM WISKUNDE & INFORMATICA (CWI), SCIENCE PARK 123, 1098 SJ AMSTERDAM, THE NETHERLANDS

*E-mail address:* `j.briet@cwi.nl`

F.M. DE OLIVEIRA FILHO, DEPARTMENT OF ECONOMETRICS AND OR, TILBURG UNIVERSITY, P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS

*E-mail address:* `F.M.deOliveiraFilho@uvt.nl`

F. VALLENTIN, DELFT INSTITUTE OF APPLIED MATHEMATICS, TECHNICAL UNIVERSITY OF DELFT, P.O. BOX 5031, 2600 GA DELFT, THE NETHERLANDS

*E-mail address:* `f.vallentin@tudelft.nl`