



## UvA-DARE (Digital Academic Repository)

### Grothendieck inequalities, nonlocal games and optimization

Briët, J.

**Publication date**  
2011

[Link to publication](#)

#### **Citation for published version (APA):**

Briët, J. (2011). *Grothendieck inequalities, nonlocal games and optimization*. [Thesis, fully internal, Universiteit van Amsterdam]. Institute for Logic, Language and Computation.

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

## Chapter 4

---

# The positive semidefinite Grothendieck problem with rank constraint

The content of this chapter is based on joint work with Fernando Mário de Oliveira Filho and Frank Vallentin [BOFV10a].

### 4.1 Introduction

In this chapter we study computational aspects of an optimization problem called *the positive semidefinite Grothendieck problem with rank- $r$  constraint*. This problem is defined as follows.

---

**Problem 4.1 (The positive semidefinite Grothendieck problem with rank- $r$  constraint).** Takes as input a positive integer  $n$  and a real  $n$ -by- $n$  positive semidefinite matrix  $A$ .

---

$$\begin{aligned} & \text{maximize} && \sum_{i,j=1}^n A_{ij} X_{ij} \\ & \text{subject to} && X \in \mathcal{S}_n^+ \\ & && X_{ii} = 1, \forall i = 1, \dots, n \\ & && \text{rank}(X) = r \end{aligned}$$

---

This optimization problem looks almost like a semidefinite program (see Section 1.7). However, the constraint on the rank makes that it is not always efficiently solvable. In particular, the case  $r = 1$  contains the maximum cut problem (MAX CUT) as an instance. When the matrix  $A$  appearing in the problem is the Laplacian matrix of a graph then the optimum gives the size of

a maximum cut in the graph (see Section 1.7.2). As MAX CUT is one of Karp's celebrated 21 NP-complete problems [Kar72] it follows that Problem 4.1 is NP-hard for  $r = 1$ . If we drop the rank constraint then the problem does become a semidefinite program, which can be solved efficiently. We will refer to this semidefinite program as the case  $r = \infty$ .

The problem can be visualized in a geometric way. A matrix  $X$  of rank  $r$  with ones on the diagonal is positive semidefinite if and only if there exist  $r$ -dimensional unit vectors  $x_1, \dots, x_n$  such that for each coordinate of  $X$ , we have  $X_{ij} = x_i \cdot x_j$ . The problem thus asks to position  $n$  points on a real  $r$ -dimensional unit sphere in such a way that a certain weighted sum of their inner products is maximized. The special case  $r = 1$  has a more combinatorial nature, since the one-dimensional unit sphere consists only of  $-1$  and  $1$ . The following proposition now follows easily and will simplify some of the notation later on.

**4.1.1. PROPOSITION.** *For all positive integers  $n, r$  and any matrix  $A \in \mathcal{S}_n^+$ , the optimum of Problem 4.1 equals  $\text{SDP}_r(A)$ , defined as in Definition 2.1, by*

$$\text{SDP}_r(A) = \max \left\{ \sum_{i,j=1}^n A_{ij} x_i \cdot y_j : x_1, \dots, x_n, y_1, \dots, y_n \in S^{r-1} \right\}.$$

PROOF: As argued above, the optimum of the problem involves has *one* sequence of unit vectors  $x_1, \dots, x_n$ . But  $\text{SDP}_r(A)$  has a maximization over *two* sequences of unit vectors. Hence,  $\text{SDP}_r(A)$  is at least the optimum of the problem. Suppose that the vectors  $x_1, \dots, x_n$  and  $y_1, \dots, y_n \in S^{r-1}$  are optimal for  $\text{SDP}_r(A)$ . Since  $A$  is positive semidefinite, there are vectors  $a_1, \dots, a_n \in S^{n-1}$  such that  $A_{ij} = a_i \cdot a_j$ . The argument of  $\text{SDP}_r(A)$  can thus be written as

$$\begin{aligned} \sum_{i,j=1}^n (a_i \cdot a_j)(x_i \cdot y_j) &= \sum_{i,j=1}^n (a_i \otimes x_i) \cdot (a_j \otimes y_j) \\ &= \left( \sum_{i=1}^n a_i \otimes x_i \right) \cdot \left( \sum_{j=1}^n a_j \otimes y_j \right). \end{aligned}$$

The last inner product is maximal if and only if the two vectors  $\sum_{i=1}^n a_i \otimes x_i$  and  $\sum_{j=1}^n a_j \otimes y_j$  are equal. Hence, we must have  $y_1 = x_1, \dots, y_n = x_n$ .  $\square$

It follows from Proposition 4.1.1 and Definition 2.3.2 that  $K_G^{\geq}(\infty \mapsto r)$  is an upper bound on the ratio of the optimum of the natural semidefinite relaxation of Problem 4.1 (the case  $r = \infty$ ), and its true optimum. Moreover,  $K_G^L(\infty \mapsto r)$

(see Definition 2.3.3) is an upper bound on this ratio if the matrix  $A$  is the Laplacian of a graph.

The case  $r = 1$  was dealt with extensively in previous works. It was studied by Rietz [Rie74] in the context of Grothendieck's inequality and by Nesterov [Nes97, Nes98] in the context of semidefinite relaxations for nonconvex quadratic optimization problems. Both proved that  $K_G^{\geq}(\infty \mapsto 1) \leq \pi/2$ , meaning that the optimum is always within a factor  $2/\pi$  of the optimum of the natural semidefinite relaxation, and Nesterov [Nes98] gave a randomized polynomial-time  $2/\pi$ -approximation algorithm for the case  $r = 1$  based on this fact (see also Section 4.2.2). Grothendieck [Gro53] proved that  $K_G^{\geq}(\infty \mapsto 1) \geq \pi/2$ , which shows that Rietz and Nesterov's result are in fact optimal (see also [AN06, Section 5.3]). Under the assumption that the UGC is true (see Section 1.7.4), Khot and Naor [KN09] proved that there is no polynomial-time approximation algorithm that has approximation ratio  $2/\pi + \varepsilon$  for any  $\varepsilon > 0$  that is independent of the matrix size  $n$ . For the special case of Laplacian matrices we saw in Section 1.7.2 that Goemans and Williamson's .878-approximation result is the best possible for polynomial-time algorithms, provided that the UGC is true. Recall that Goemans and Williamson's result together with those of Karloff [Kar96] and Feige and Schechtman's [FS02] imply that  $K_G^L(\infty \mapsto 1) = (.878\dots)^{-1}$ . Avidor and Zwick [AZ05] proved that  $K_G^L(q \mapsto 1) < (.878\dots)^{-1}$  when  $q = 2, 3$ , which means that better approximation results are possible when the semidefinite relaxation has an optimal solution of rank 2 or 3.

Much less seems to be known about the more geometric cases of Problem 4.1, where  $r \geq 2$ . In this chapter we extend most of the known complexity results for the case  $r = 1$  to larger values of  $r$ .

### 4.1.1 An optimal approximation algorithm?

In this section we present the main results of this chapter. As mentioned above, the natural semidefinite relaxation of Problem 4.1 is simply the same optimization problem without the rank constraint (the case  $r = \infty$ ). Based on this semidefinite relaxation we construct a simple polynomial-time approximation algorithm for Problem 4.1, Algorithm 4.1 below. For the case  $r = 1$  this algorithm is Goemans and Williamson's celebrated *randomized hyperplane rounding* algorithm. For this case the algorithm can be derandomized using the techniques of Mahajan and Ramesh [MR95].

---

**Algorithm 4.1** Takes as input positive integers  $n, r$  and  $n$ -by- $n$  positive semidefinite matrix  $A$ , and returns a feasible solution  $y_1, \dots, y_n \in S^{r-1}$  for Problem 4.1.

---

- (1) Solve the semidefinite relaxation of Problem 4.1 for the matrix  $A$ , obtaining vectors  $x_1, \dots, x_n \in S^{n-1}$ .
  - (2) Sample matrix  $Z \in \mathbb{R}^{r \times n}$  according to  $N(0, 1)^{r \times n}$ , that is, the entries  $Z_{ij}$  are i.i.d. random variables with mean 0 and variance 1.
  - (3) Define  $y_1, \dots, y_n \in S^{r-1}$  by  $y_i = Zx_i / \|Zx_i\|_2$  for  $i = 1, \dots, n$ .
- 

**The approximation ratio.** The analysis of Algorithm 4.1 gives the following approximation result for Problem 4.1.

**4.1.2. THEOREM.** *For every positive integer  $r$  we have*

$$1 \leq K_G^{\succ}(\infty \mapsto r) \leq \frac{1}{\gamma(r)} = 1 + \Theta\left(\frac{1}{r}\right),$$

where

$$\gamma(r) = \frac{2}{r} \left( \frac{\Gamma\left(\frac{r+1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)} \right)^2,$$

and there is a randomized polynomial-time  $\gamma(r)$ -approximation algorithm for Problem 4.1 that is based on its natural semidefinite relaxation.

We prove this theorem in Section 4.2

**A refined, dimension-dependent analysis.** If we take into account the size of the matrix  $A$  appearing in Problem 4.1 then the upper bounds given in Theorem 4.1.2 can be tightened for the combinatorial case  $r = 1$ . This gives a slight improvement on the bounds of Nesterov [Nes97] and Rietz [Rie74]. Note that if a positive semidefinite matrix  $A$  has size  $n$ -by- $n$ , then  $\text{SDP}_\infty(A) = \text{SDP}_n(A)$ . Upper bounds on  $K_G^{\succ}(n \mapsto 1)$  therefore imply upper bounds on the ratio  $\text{SDP}_\infty(A) / \text{SDP}_1(A)$  whenever  $A$  is positive semidefinite and of size  $n$ -by- $n$ .

**4.1.3. THEOREM.** *For every positive integer  $n$  we have*

$$1 \leq K_G^{\succ}(n \mapsto 1) \leq \frac{\pi\gamma(n)}{2} = \frac{\pi}{2} - \Theta\left(\frac{1}{n}\right),$$

and there is a polynomial-time  $2/(\pi\gamma(n))$ -approximation algorithm for the case  $r = 1$  of Problem 4.1.

We prove this theorem in Section 4.3. Together with Theorem 3.3.1, Theorems 4.1.2 and 4.1.3 imply that we now know the exact values of  $K_G^{\geq}(\infty \mapsto r)$  and  $K_G^{\geq}(n \mapsto 1)$ . To see this, notice that Grothendieck's operator, which we used to prove the lower bounds of Theorem 3.3.1, can be seen as an infinite-dimensional matrix given by  $A(x, y) = x \cdot y$  where  $x, y$  are  $n$ -dimensional unit vectors. Clearly this matrix is positive semidefinite. The problem of approximating this matrix by a finite matrix while preserving positive-semidefiniteness can be dealt with using an  $\varepsilon$ -net argument of Alon and Naor [AN06, Section 5.2]. The first ten values of  $K_G(\infty \mapsto r)$  are summarized in Table 4.1.

**Table 4.1:** The table shows the exact values of  $K_G^{\geq}(\infty \mapsto r)$  for  $r = 1, \dots, 10$ . For  $r = 1$ , the lower bound is due to Grothendieck [Gro53] and the upper bound due to Nesterov [Nes98] and Rietz [Rie74].

$r$	$K_G^{\geq}(\infty \mapsto r)$
1	1.570796...
2	1.273239...
3	1.178097...
4	1.131768...
5	1.104466...
6	1.086497...
7	1.073786...
8	1.064324...
9	1.057008...
10	1.051184...

**Unique-Games hardness of approximation.** By using arguments from the proof of Theorem 4.1.3 and by Khot and Naor's [KN09] UGC hardness result for approximating the case  $r = 1$ , we obtain the following hardness result for approximating Problem 4.1.

**4.1.4. THEOREM.** *Under the assumption of the Unique Games Conjecture, there is no polynomial-time approximation algorithm for Problem 4.1 that has ratio  $\gamma(r) + \varepsilon$  for any  $\varepsilon > 0$  that is independent of  $n$  (where  $n$  is the size of the matrix in Problem 4.1).*

We prove this theorem in Section 4.4. With this, the current complexity status of the  $r = 1$  case of Problem 4.1 is similar to the one of the minimum

vertex cover problem: given a graph, find a subset of the vertices of minimal size, such that every edge has at least one endpoint in the subset. On the one hand, Karakostas [Kar05] showed that this problem can be approximated to within a factor  $2 - \Theta(1/\sqrt{\log |V|})$  in polynomial time. On the other hand, Khot and Regev [KR08] showed that under the assumption of the UGC, the size of a minimum vertex cover cannot be approximated in polynomial time to within a factor  $2 - \varepsilon$  for any  $\varepsilon > 0$  that is independent of  $|V|$ .

### 4.1.2 Interpretations

We give two interpretations of Problem 4.1, one in classical statistical physics and one in nonlocal games. The objective function of Problem 4.1 can be interpreted as the energy of a system of interacting particles. Stanley [Sta68] introduced a model of  $n$  interacting particles in a spin glass with ferromagnetic and antiferromagnetic interactions, where the particles are represented by  $r$ -dimensional unit vectors  $x_1, \dots, x_n$ . The case  $r = 1$  corresponds to the Ising model, the case  $r = 2$  to the XY (or planar) model, the case  $r = 3$  to the Heisenberg model, and the case  $r = \infty$  to the Berlin-Kac spherical model. The potential function  $(A_{ij})_{i,j=1}^n$  is 0 if particles  $i$  and  $j$  do not interact, it is positive if there is ferromagnetic interaction between particles  $i$  and  $j$ , and it is negative if there is antiferromagnetic interaction. In the absence of an external field, the energy of the system is given by the *Hamiltonian*

$$- \sum_{i,j=1}^n A_{ij} x_i \cdot x_j.$$

The *ground state* of this model is a configuration of spins  $x_1, \dots, x_n \in S^{r-1}$  which minimizes the total energy. If  $A$  is positive semidefinite, finding the ground state is the same as solving Problem 4.1. Of course, considering only positive semidefinite potential functions may be rather restrictive and in Chapter 5 we deal with the most general setting of Stanley's model (which requires a fair bit more work). However, if the potential function is indeed positive semidefinite then the approximation results for the ground state energy given in Theorem 4.1.2 are stronger than those presented in Chapter 5.

Proposition 4.1.1 creates a bridge between the optimum of Problem 4.1 and the bias of certain two-player XOR games based on dimensional-restricted entangled strategies. Let  $\mathcal{G} = (\pi, \Sigma)$  be a two-player XOR game given by a probability distribution  $\pi$  on  $\{1, \dots, n\} \times \{1, \dots, n\}$  and  $n$ -by- $n$  sign matrix  $\Sigma$ . Define an  $n$ -by- $n$  matrix  $A$  by  $A_{ij} = \pi(i, j)\Sigma_{ij}$ . It follows from Tsirelson's Theorem

(see Section 1.5) that  $\text{SDP}_r(A)$  is a *lower* bound on the entangled bias of  $\mathcal{G}$  when the players are restricted to sharing a state with local dimension  $2^{\lceil r/2 \rceil}$ , and an *upper* bound on the entangled bias when the players are restricted to local dimensions  $\sqrt{r/2}$ . If the matrix  $A$  is positive semidefinite, then the approximation results of Theorem 4.1.2 now allow us to estimate these biases. Of course, considering only games for which the game matrix is positive semidefinite is rather restrictive and the results of Chapter 5 will allow us to drop this assumption (see Section 5.1.1). The results of this chapter give better approximation results for these bounds with this restriction.

### 4.1.3 More related work

A few variations of Problem 4.1 that were previously considered in the context of optimization are as follows.

**Quadratic programming.** If we allow the matrix  $A$  that appears in Problem 4.1 to also have negative eigenvalues, then the case  $r = 1$  corresponds to the well-studied problem of quadratic programming [BBC04, CW04, ABH<sup>+</sup>05, AN06, AMMN06, KO06, RS09, KN10]. We will consider this problem and its generalization for larger values of  $r$  in detail in Chapter 5.

**The  $\ell_p$ -Grothendieck problem.** Allowing the matrix  $A$  to have negative eigenvalues and optimizing over matrices of the form  $X = xx^T$  for  $x \in \mathbb{R}^n$  such that  $\|x\|_p \leq 1$ , gives the  $\ell_p$ -Grothendieck problem. For  $p \geq 2$ , Kindler, Naor and Schechtman [KNS10] gave a polynomial-time  $(p/e + 30 \log p)$ -approximation algorithm and showed that under the assumption of the UGC, it is NP-hard to approximate the optimum to within factor  $p/e + 1/4$ .

**The Kernel-Clustering problem.** In the *kernel clustering problem*, introduced by Song et al. [SSGB07], in addition to an  $n$ -by- $n$  positive semidefinite matrix  $A$ , we are given a smaller  $k$ -by- $k$  positive semidefinite matrix  $B$ . The goal is to find a partition  $S_1, \dots, S_k$  of the set  $\{1, \dots, n\}$  so as to maximize

$$\sum_{i,j=1}^k \left( \sum_{(i',j') \in S_i \times S_j} A_{i'j'} \right) B_{ij}.$$

The case where  $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$  corresponds to Problem 4.1 with  $r = 1$ . Khot and Naor [KN10] gave polynomial-time (in  $n$ ) approximation algorithms for ev-

ery choice of the matrix  $B$  and showed their approximation ratios are optimal under the assumption of the UGC.

**Outline of the rest of this chapter.** In Section 4.2 we give a detailed analysis of the approximation algorithm given in Section 4.1.1, leading to a proof of Theorem 4.1.2. In Section 4.3 we prove Theorem 4.1.3. In Section 4.4 we prove the UGC hardness results for Problem 4.1 given in Theorem 4.1.4. In Section 4.5 we specialize some of our results to the case of Laplacian matrices and we briefly summarize this chapter in Section 4.6.

## 4.2 The approximation ratio

In this section we prove Theorem 4.1.2. We achieve this by analyzing Algorithm 4.1, which converts solution vectors  $x_1, \dots, x_n \in S^{n-1}$  of the semidefinite relaxation of Problem 4.1 into a feasible solution in the form of vector-valued random variables  $y_1, \dots, y_n \in S^{r-1}$ . Our techniques are inspired by the approach used by Nesterov [Nes97] for the case  $r = 1$ .

### 4.2.1 The expectation function

By linearity of expectation, the expected quality of the solution of Algorithm 4.1 is given by

$$\mathbb{E} \left[ \sum_{i,j=1}^n A_{ij} y_i \cdot y_j \right] = \sum_{i,j=1}^n A_{ij} \mathbb{E}[y_i \cdot y_j]. \quad (4.1)$$

Let us have a closer look at the expectation  $\mathbb{E}[y_i \cdot y_j]$  for some arbitrary pair  $i, j$ . By the definition of the random vectors  $y_i, y_j$  the expectation equals

$$\mathbb{E} \left[ \frac{Zu}{\|Zu\|_2} \cdot \frac{Zv}{\|Zv\|_2} \right], \quad (4.2)$$

where the expectation is over random Gaussian matrix  $Z \sim N(0, 1)^{r \times n}$  and  $u, v$  are some  $n$ -dimensional unit vectors. The distribution of  $Z$  is invariant under orthogonal transformations, that is, for any orthogonal matrix  $U \in O(\mathbb{R}^n)$ , the random matrix  $ZU$  has the same distribution. To see this, note that each row of  $Z$  is an independent random vector whose direction with respect to the origin is uniformly distributed. An orthogonal transformation simply rotates these vectors about the origin, thus leaving their distributions unchanged. It follows that we can pick  $U$  such that  $Uu = (1, 0, \dots, 0)^T$  and  $Uv = (t, \sqrt{1-t^2}, 0, \dots, 0)^T$

for  $t = u \cdot v$ , and leave the expectation (4.2) unchanged, showing that it depends on the inner product  $u \cdot v$  only. This justifies defining the function  $E_r : [-1, 1] \rightarrow [-1, 1]$  by

$$E_r(u \cdot v) = \mathbb{E} \left[ \frac{Zu}{\|Zu\|_2} \cdot \frac{Zv}{\|Zv\|_2} \right].$$

Then, since we had  $y_i = Zx_i / \|Zx_i\|_2$  where  $x_1, \dots, x_n \in S^{n-1}$  are optimal for  $\text{SDP}_\infty(A)$ , we can write the right-hand side of Eq. (4.1) as

$$\sum_{i,j=1}^n A_{ij} E_r(x_i \cdot x_j). \quad (4.3)$$

The following lemma shows that the function  $E_r$  enjoys a special property that will allow us to derive lower bounds for  $\text{SDP}_r(A)$  in terms of  $\text{SDP}_\infty(A)$ .

**4.2.1. LEMMA.** *There exists a real number  $c > 0$  such that for every positive integer  $k$  and any real  $n$ -dimensional unit vectors  $u_1, \dots, u_k$ , the matrix*

$$(E_r(u_i \cdot u_j) - cu_i \cdot u_j)_{i,j=1}^k$$

*is positive semidefinite.*

Recall that for positive semidefinite matrices  $A, B$ , we have  $\sum_{i,j} A_{ij} B_{ij} = \langle A, B \rangle \geq 0$ . Hence, by Eq. (4.3) and Lemma 4.2.1, we have

$$\begin{aligned} \text{SDP}_r(A) &\geq \sum_{i,j=1}^n A_{ij} E_r(x_i \cdot x_j) \\ &= c \sum_{i,j=1}^n A_{ij} x_i \cdot x_j + \sum_{i,j=1}^n A_{ij} (E_r(x_i \cdot x_j) - cx_i \cdot x_j) \\ &\geq c \text{SDP}_\infty(A), \end{aligned}$$

where the factor  $c$  comes from Lemma 4.2.1. From this it follows that the second term on the second line is at least 0. The second inequality follows since the vectors  $x_1, \dots, x_n$  are optimal for  $\text{SDP}_\infty(A)$ . Lemma 4.2.1 thus enables us to prove that  $\text{SDP}_\infty(A) / \text{SDP}_r(A) \leq 1/c$ . As  $\text{SDP}_\infty(A) \geq \text{SDP}_r(A)$ , we also get that the approximation ratio of Algorithm 4.1 is at least  $c$ . In the next two sections we prove Lemma 4.2.1 and compute the number  $c$ .

## 4.2.2 Positive functions for spheres

Lemma 4.2.1 states that the function  $t \mapsto E_r(t) - ct$  has a special property, namely that it is of *positive type for unit spheres*.

**4.2.2. DEFINITION.** A continuous function  $f : [-1, 1] \rightarrow [-1, 1]$  is of *positive type for  $S^\infty$* , if for all positive integers  $n, k$  and any real  $n$ -dimensional unit vectors  $u_1, \dots, u_k$ , the matrix

$$(f(u_i \cdot u_j))_{i,j=1}^k$$

is positive semidefinite.

Functions of positive type were extensively studied by Schoenberg [Sch42], who gave a very useful characterization of them in terms of their Taylor series.

**4.2.3. THEOREM (SCHOENBERG).** A continuous function  $f : [-1, 1] \rightarrow \mathbb{R}$  is of positive type for  $S^\infty$  if and only if it is of the form

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

where  $c_0, c_1, \dots \geq 0$  and the series  $\sum_{k=0}^{\infty} c_k$  converges.

**The rank-1 case.** The analysis for the case  $r = 1$  relies on Grothendieck's Identity (Lemma 1.7.1), which we restate below for convenience. This identity gives the exact form of the function  $E_1$  and allows us to obtain our lower bound on the number  $c$  from Lemma 4.2.1. In turn we get a lower bound on the approximation ratio of Algorithm 4.1.

**4.2.4. LEMMA (GROTHENDIECK'S IDENTITY).** Let  $u, v$  be real unit vectors and let  $z$  be a random vector with independently distributed entries that have mean 0 and variance 1. Then, we have

$$\mathbb{E}[\text{sign}(z \cdot u) \text{sign}(z \cdot v)] = \frac{2}{\pi} \arcsin(u \cdot v).$$

Grothendieck's Identity and the Taylor expansion of the arcsine function thus give

$$E_1(t) = \frac{2}{\pi} \arcsin t = \frac{2}{\pi} \sum_{k=0}^{\infty} \frac{(2k)!}{2^{2k} (k!)^2 (2k+1)} t^{2k+1}.$$

Notice that all the coefficients in this expansion are nonnegative (also the series converges on  $[-1, 1]$ ). Hence, by Schoenberg's Theorem (Theorem 4.2.3), this

expansion shows that  $E_1$  is indeed a function of positive type for  $S^\infty$ . As the linear term in this expansion is  $2t/\pi$ , it follows that the function  $(E_1(t) - 2t/\pi)$  is of positive type for  $S^\infty$ . For the case  $r = 1$ , the number  $c$  from Lemma 4.2.1 giving the upper bound  $\text{SDP}_\infty(A)/\text{SDP}_1(A) \leq 1/c$  can thus be taken to be  $2/\pi$ . We have just derived Nesterov's  $2/\pi$  upper bound on the approximation ratio of Algorithm 4.1 for the case  $r = 1$ .

**Extension to higher ranks.** For the cases  $r \geq 2$  it takes quite a bit of work to obtain an explicit form of the function  $E_r$ . We obtain this in Chapter 5 (see Lemma 5.2.1). For the moment we take the following approach. We first argue that for every positive integer  $r$  the function  $E_r$  is of positive type for  $S^\infty$ . To see this, note that for every positive integer  $k$  and any choice of unit vectors  $u_1, \dots, u_k$ , we have that the matrix

$$(E_r(u_i \cdot u_j))_{i,j=1}^k$$

is a convex combination of positive semidefinite matrices, since

$$\begin{aligned} (E_r(u_i \cdot u_j))_{i,j=1}^k &= \left( \mathbb{E} \left[ \frac{Zu_i}{\|Zu_i\|_2} \cdot \frac{Zu_j}{\|Zu_j\|_2} \right] \right)_{i,j=1}^k \\ &= \mathbb{E} \left[ \left( \frac{Zu_i}{\|Zu_i\|_2} \cdot \frac{Zu_j}{\|Zu_j\|_2} \right)_{i,j=1}^k \right]. \end{aligned}$$

Clearly each of the matrices inside the square brackets is positive semidefinite. Convex combinations of positive semidefinite matrices are again positive semidefinite, showing that the function  $E_r$  is indeed of positive type for  $S^\infty$ .

Now, by Schoenberg's Theorem (Theorem 4.2.3) there exist  $c_0, c_1, \dots \geq 0$  such that  $E_r(t) = \sum_{k=0}^{\infty} c_k t^k$ . A second application of Schoenberg's Theorem then gives that the function  $E_r(t) - c_1 t$  is of positive type for  $S^\infty$  as well. It follows that Lemma 4.2.1 holds for  $c$  the coefficient  $c_1$  multiplying the linear term in the Taylor series of  $E_r$ . For our purposes, it therefore suffices just to compute this term instead of the whole Taylor expansion of  $E_r$ .

### 4.2.3 The Wishart distribution

What is left to do, is to compute the coefficient multiplying the linear term in the Taylor series expansion of the function  $E_r$  for  $r \geq 2$ . To this end, we simplify the expression for  $E_r$  and evaluate its first derivative at  $t = 0$ .

Towards simplifying the expression for  $E_r$ , let for some angle  $\theta \in [0, 2\pi]$ ,  $u = (\cos \theta, \sin \theta, 0, \dots, 0)^T$  and  $v = (\cos \theta, -\sin \theta, 0, \dots, 0)^T$  be  $n$ -dimensional unit vectors. Notice that any pair of unit vectors can be simultaneously put into this form by an orthogonal transformation. Assuming that the vectors have this form will bring the number of dimensions involved in the expression for  $E_r$  down to two, because all terms that appear in the expression

$$\frac{Zu}{\|Zu\|_2} \cdot \frac{Zv}{\|Zv\|_2} = \frac{u^T Z^T Z v}{\sqrt{(u^T Z Z^T u)(v^T Z Z^T v)'}}$$

involve only the upper-left 2-by-2 sub-matrix of the matrix  $Z^T Z$ . This sub-matrix is distributed according to a (standard) Wishart distribution from multivariate statistics. The *Wishart distribution*  $W_2(r)$  is the distribution of a 2-by-2 positive semidefinite matrix of the form  $H^T H$  where  $H$  is an  $r$ -by-2 random matrix with independent  $N(0, 1)$  entries (see for example [Mui82]). This distribution may be seen as a matrix variant of the chi-square distribution. The probability density function of  $W_2(r)$  is given by

$$\frac{1}{2^r \Gamma_2(r/2)} e^{\text{Tr}(W)/2} (\det W)^{(r-3)/2},$$

where  $\Gamma_q$  is the *multivariate gamma function*, defined as

$$\Gamma_q(x) = \pi^{q(q-1)/4} \prod_{i=1}^q \Gamma\left(x - \frac{i-1}{2}\right).$$

Hence, for  $x = (\cos \theta, \sin \theta)^T$ ,  $y = (\cos \theta, -\sin \theta)^T$ ,  $t = \cos 2\theta$  and  $W \sim W_2(r)$ , we now have a more explicit form for the function  $E_r$ , given by

$$\begin{aligned} E_r(t) &= \mathbb{E}_{W \sim W_2(r)} \left[ \frac{x^T W y}{\sqrt{(x^T W x)(y^T W y)}} \right] \\ &= \frac{1}{2^r \Gamma_2(r/2)} \int_{\mathcal{S}_2^+} \frac{x^T W y}{\sqrt{(x^T W x)(y^T W y)}} e^{\text{Tr}(W)/2} (\det W)^{(r-3)/2} dW \end{aligned} \quad (4.4)$$

The integral above can be simplified by using the parametrization of the cone of 2-by-2 positive semidefinite matrices given by

$$\mathcal{S}_2^+ = \left\{ \begin{pmatrix} \frac{a}{2} + \cos \phi & \alpha \sin \phi \\ \alpha \sin \phi & \frac{a}{2} - \cos \phi \end{pmatrix} : a \in \mathbb{R}_+, \phi \in [0, 2\pi], \alpha \in [0, a/2] \right\}.$$

This parametrization can easily be obtained from the characteristic polynomial  $t^2 - \text{Tr}(W)t + \det(W)$  of a generic element  $W \in \mathcal{S}_2^+$ . We then have

$$\text{Tr}(W) = a, \quad \det W = \frac{a^2}{4} - \alpha^2, \quad dW = \alpha d\phi d\alpha da$$

and

$$\begin{aligned} x^T W y &= \frac{at}{2} + \alpha \cos \phi \\ x^T W x &= \frac{a}{2} + \alpha(t \cos \phi + 2 \sin \theta \cos \theta \sin \phi) \\ y^T W y &= \frac{a}{2} + \alpha(t \cos \phi - 2 \sin \theta \cos \theta \sin \phi). \end{aligned}$$

Plugging this back into the form for  $E_r(t)$  obtained in Eq. (4.4) gives the large, but manageable, triple integral

$$E_r(t) = \frac{1}{2^r \Gamma_2(r/2)} \int_0^\infty \int_0^{a/2} \int_0^{2\pi} \frac{\frac{at}{2} + \alpha \cos \phi}{\sqrt{\left(\frac{a}{2} + \alpha t \cos \phi\right)^2 - \alpha^2(1-t^2)(\sin \phi)^2}} \cdot e^{-a/2} \left(\frac{a^2}{4} - \alpha^2\right)^{(r-3)/2} \alpha d\phi dx da.$$

Making the substitution  $\alpha = (a/2)s$  and integrating over  $a$  already reduces the integral to

$$\frac{\Gamma(r)}{2^{r-1} \Gamma_2(r/2)} \int_0^1 \int_0^{2\pi} \frac{(t + s \cos \phi) s (1-t^2)^{(r-3)/2}}{\sqrt{(1 + st \cos \phi)^2 - s^2(1-t^2)(\sin \phi)^2}} d\phi ds. \quad (4.5)$$

Another simplification follows from Legendre's duplication formula [AAR99, Theorem 1.5.1],  $\Gamma(2m)\Gamma(1/2) = 2^{2m-1}\Gamma(m)\Gamma(m+1/2)$ , which gives

$$\frac{\Gamma(r)}{2^{r-1} \Gamma_2(r/2)} = \frac{r-1}{2\pi}.$$

Recall that our objective was to compute the coefficient multiplying the linear term in the Taylor expansion of  $E_r$ . Evaluating the derivative of Eq. (4.5) with respect to  $t$  at  $t = 0$ , gives that this coefficient is given by the integral

$$c_1 = \frac{r-1}{2\pi} \int_0^1 \int_0^{2\pi} \frac{s(1-s^2)^{(r-1)/2}}{(1-s^2(\sin \phi)^2)^{3/2}} d\phi ds.$$

Using Euler's integral representation of the hypergeometric function [AAR99, Theorem 2.2.1] and by a substitution of variables, we get

$$\begin{aligned} c_1 &= \frac{r-1}{2\pi} \int_0^{2\pi} \frac{\Gamma(1)\Gamma((r+1)/2)}{2\Gamma((r+3)/2)} {}_2F_1 \left( \begin{matrix} 3/2, 1 \\ (r+3)/2 \end{matrix}; \sin^2 \phi \right) d\phi \\ &= \frac{r-1}{4\pi} \frac{\Gamma((r+1)/2)}{\Gamma((r+3)/2)} 4 \int_0^1 {}_2F_1 \left( \begin{matrix} 3/2, 1 \\ (r+3)/2 \end{matrix}; t^2 \right) (1-t^2)^{-1/2} dt \\ &= \frac{r-1}{\pi} \frac{\Gamma((r+1)/2)}{\Gamma((r+3)/2)} \frac{1}{2} \int_0^1 {}_2F_1 \left( \begin{matrix} 3/2, 1 \\ (r+3)/2 \end{matrix}; t \right) (1-t)^{-1/2} t^{-1/2} dt. \end{aligned}$$

This simplifies further by Euler's generalized integral [AAR99, (2.2.2)], and Gauss's summation formula [AAR99, Theorem 2.2.2]

$$\begin{aligned}
c_1 &= \frac{r-1}{2\pi} \frac{\Gamma((r+1)/2)}{\Gamma((r+3)/2)} \frac{\Gamma(1/2)\Gamma(1/2)}{\Gamma(1)} {}_3F_2 \left( \begin{matrix} 3/2, 1, 1/2 \\ (r+3)/2, 1 \end{matrix}; 1 \right) \\
&= \frac{r-1}{2} \frac{\Gamma((r+1)/2)}{\Gamma((r+3)/2)} {}_2F_1 \left( \begin{matrix} 3/2, 1/2 \\ (r+3)/2 \end{matrix}; 1 \right) \\
&= \frac{r-1}{2} \frac{\Gamma((r+1)/2)}{\Gamma((r+3)/2)} \frac{\Gamma((r+3)/2)\Gamma((r-1)/2)}{\Gamma(r/2)\Gamma((r+2)/2)} \\
&= \frac{2}{r} \left( \frac{\Gamma((r+1)/2)}{\Gamma(r/2)} \right)^2.
\end{aligned}$$

This proves both Lemma 4.2.1 and Theorem 4.1.2, as it shows that  $c_1 = \gamma(r)$ .

### 4.3 A refined, dimension-dependent analysis

In this section we show that one can slightly improve Nesterov and Rietz's approximation ratio for Algorithm 4.1 for the case  $r = 1$  when we take into account the size of the matrix. This result is key to the hardness results for approximating  $\text{SDP}_r(A)$  presented in the next section.

We will use another theorem of Schoenberg [Sch42], which gives a characterization of positive functions on spheres of specific dimension.<sup>1</sup> The Taylor series that appears in Theorem 4.2.3 will be replaced by a series expansion in terms of Gegenbauer polynomials. These polynomials form a complete orthogonal basis for  $L^2([-1, 1])$ , the space of square-integrable functions on  $[-1, 1]$ , endowed with the inner product

$$(f, g)_n = \int_{-1}^1 f(t)g(t)(1-t^2)^{(n-3)/2} dt. \quad (4.6)$$

The *Gegenbauer polynomials*  $P_0^n, P_1^n, P_2^n, \dots$  are the polynomials obtained by performing a Gram-Schmidt orthogonalization procedure to the sequence of linearly independent functions  $1, t, t^2, \dots$  (see for example [Sze75, Chapter IV]).

**4.3.1. THEOREM (SCHOENBERG).** *A continuous function  $f : [-1, 1] \rightarrow \mathbb{R}$  is of positive type for  $S^{n-1}$  if and only if it is of the form*

$$f(t) = \sum_{k=0}^{\infty} c_k P_k^n(t),$$

<sup>1</sup>A nice proof of this theorem can be found in [OF09].

for  $c_0, c_1, \dots \geq 0$  such that the series  $\sum_{k=0}^{\infty} c_k$  converges.

PROOF OF THEOREM 4.1.3: Let  $x_1, \dots, x_n \in S^{n-1}$  be optimal for  $\text{SDP}_{\infty}(A)$ . By Grothendieck's Identity, Algorithm 4.1 gives  $\{-1, 1\}$ -valued random variables  $\chi_1, \dots, \chi_n$  that satisfy

$$\mathbb{E}[\chi_i \chi_j] = \frac{2}{\pi} \arcsin(x_i \cdot x_j).$$

Since the arcsin function is positive for  $S^{\infty}$ , in particular it is positive for  $S^{n-1}$ . Therefore, by Theorem 4.3.1, arcsin can be expanded in terms of the Gegenbauer polynomials as

$$\arcsin(t) = \sum_{k=0}^{\infty} c_k P_k^n(t),$$

where  $c_0, \dots, c_k \geq 0$  and  $\sum_{k=0}^{\infty} c_k$  converges. Then, since  $P_1^n(t) = t$ , the function  $\arcsin(t) - c_1 t$  is positive for  $S^{n-1}$  as well. Arguing as before, we get

$$\text{SDP}_1(A) \geq \frac{2}{\pi} \sum_{i,j=1}^n A_{ij} \arcsin(x_i \cdot x_j) \geq \frac{2c_1}{\pi} \text{SDP}_{\infty}(A).$$

What is left is to compute the constant  $c_1$ . Since the Gegenbauer polynomials are orthonormal with respect to the inner product (4.6) and  $P_1^n(t) = t$ , we have  $c_1 = c(n) = (\arcsin, P_1^n)_n / (P_1^n, P_1^n)_n$ . The numerator of  $c(n)$  equals

$$\begin{aligned} (\arcsin t, P_1^n)_n &= \int_{-1}^1 \arcsin(t) t (1-t^2)^{(n-3)/2} dt \\ &= \int_{-\pi/2}^{\pi/2} \theta \sin \theta (\cos \theta)^{n-2} d\theta \\ &= \frac{\Gamma(1/2) \Gamma(\frac{n}{2})}{(n-1) \Gamma(\frac{n+1}{2})}. \end{aligned}$$

The denominator of  $c(n)$  equals

$$\begin{aligned} (P_1^n, P_1^n)_n &= \int_{-1}^1 t^2 (1-t^2)^{(n-3)/2} dt \\ &= \frac{\Gamma(3/2) \Gamma(\frac{n-1}{2})}{\Gamma(\frac{n+2}{2})}, \end{aligned}$$

where we used the Beta integral of Eq. (3.14). Now, by using the functional equation  $x\Gamma(x) = \Gamma(x+1)$ , the desired equality  $c(n) = 1/\gamma(n)$  follows.  $\square$

## 4.4 Unique-Games hardness of approximation

In this section, we prove the hardness of approximation result for Problem 4.1 given in Theorem 4.1.4. The idea behind the proof is that a good approximation algorithm for the case  $r > 1$  can be converted into a good approximation algorithm for the case  $r = 1$ . By Khot and Naor's [KN09] UGC hardness results for the case  $r = 1$ , the algorithm for  $r > 1$  cannot be too good.

PROOF OF THEOREM 4.1.4: Suppose that  $\rho$  is the smallest approximation ratio a polynomial-time algorithm can achieve for Problem 4.1. Given positive integer  $n$  and  $n$ -by- $n$  positive semidefinite matrix  $A$ , let  $x_1, \dots, x_n \in S^{r-1}$  be an approximate solution coming from such a polynomial-time algorithm. Then,

$$\sum_{i,j=1}^n A_{ij} x_i \cdot x_j \geq \rho \text{SDP}_r(A).$$

Applying the hyperplane rounding technique to  $x_1, \dots, x_n \in S^{r-1}$  gives  $\{-1, 1\}$ -valued random variables  $\chi_1, \dots, \chi_n$  such that

$$\begin{aligned} \mathbb{E} \left[ \sum_{i,j=1}^n A_{ij} \chi_i \chi_j \right] &= \frac{2}{\pi} \sum_{i,j=1}^n A_{ij} \arcsin x_i \cdot x_j \\ &\geq \frac{2\rho}{\pi\gamma(r)} \text{SDP}_r(A), \end{aligned}$$

where we used the fact that the function  $\arcsin(t) - t/\gamma(r)$  is of positive type for  $S^{r-1}$ , as was established in the previous section in the proof of Theorem 4.1.3. Since  $\text{SDP}_r(A) \geq \text{SDP}_1(A)$ , this is a polynomial-time approximation algorithm for the  $r = 1$  case of Problem 4.1 with approximation ratio  $\pi\gamma(r)/(2\rho)$ . The hardness result of [KN09] for approximating this case with ratio  $\pi/2 - \varepsilon$  for  $\varepsilon > 0$  independent of  $r$  now gives that the UGC implies  $\rho \leq \gamma(r)$ .  $\square$

## 4.5 The case of graphs

In this section we show that one can improve the approximation ratio of Algorithm 4.1 if the positive semidefinite matrix  $A = (A_{ij}) \in \mathbb{R}^{n \times n}$  has the following special structure:

$$A_{ij} \leq 0, \quad \text{if } i \neq j, \quad (4.7)$$

$$\sum_{i=1}^n A_{ij} = 0, \quad \text{for every } j = 1, \dots, n. \quad (4.8)$$

This happens for instance when  $A$  is the Laplacian matrix of a graph.

**4.5.1. PROPOSITION.** *For positive integers  $n, r$  with  $r \leq n$  and real  $n$ -by- $n$  positive semidefinite matrix  $A$  that satisfies Eq.'s (4.7) and (4.8), we have*

$$\frac{\text{SDP}_\infty(A)}{\text{SDP}_r(A)} \leq \frac{1}{\rho(r)},$$

where  $\rho(r)$  is given by

$$\rho(r) = \min \left\{ \frac{1 - E_r(t)}{1 - t} : t \in [-1, 1] \right\}.$$

In particular, the above proposition implies  $K_G^L(\infty \mapsto r) \leq 1/\rho(r)$ . The proof follows a standard argument of Goemans and Williamson [GW95] (see also Section 1.7.2).

PROOF: Applying Algorithm 4.1 gives  $S^{r-1}$ -valued random variables  $y_1, \dots, y_n$  such that

$$\text{SDP}_r(A) \geq \mathbb{E} \left[ \sum_{i,j=1}^n A_{ij} y_i \cdot y_j \right] = \sum_{i,j=1}^n A_{ij} E_r(x_i \cdot x_j),$$

where  $x_1, \dots, x_n \in S^{n-1}$  are optimal vectors for  $\text{SDP}_\infty(A)$ . Note that we have  $E_r(1) = 1$ , which follows easily from the definition of this function. Using this, and the fact that  $A$  satisfies Eq.'s (4.7) and (4.8), we have

$$\begin{aligned} \sum_{i,j=1}^n A_{ij} E_r(x_i \cdot x_j) &= \sum_{i,j=1}^n (-A_{ij})(1 - E_r(x_i \cdot x_j)) \\ &= \sum_{i \neq j} (-A_{ij}) \frac{1 - E_r(x_i \cdot x_j)}{1 - x_i \cdot x_j} (1 - x_i \cdot x_j) \\ &\geq \rho(r) \sum_{i \neq j} (-A_{ij})(1 - x_i \cdot x_j) \\ &= \rho(r) \sum_{i,j=1}^n (-A_{ij})(1 - x_i \cdot x_j) \\ &= \rho(r) \text{SDP}_\infty(A), \end{aligned}$$

where we used Eq. (4.8) on the first line,  $E_r(1) = 1$  on the second line, Eq. (4.7) and the definition of  $\rho(r)$  on the third line,  $x_i \cdot x_i = 1$  on the fourth line and Eq. (4.8) on the last line.  $\square$

The first ten numerical values of the above upper bounds are given in Table 4.2. The numerical values suggest that as  $r \rightarrow \infty$ , the value of  $t$  for which the minimum appearing in the function  $\rho(r)$  is attained approaches 0.5.

**Table 4.2:** The table shows numerical estimates for the approximation ratio of Algorithm 4.1 for the case of Laplacian matrices of graphs for  $r = 1, \dots, 10$ . The case  $r = 1$  corresponds to the MAX CUT approximation algorithm of Goemans and Williamson [GW95].

$r$	$\rho(r)$	minimum attained at
1	0.87856...	-0.68915
2	0.93494...	-0.61712
3	0.95633...	-0.58426
4	0.96733...	-0.56556
5	0.97397...	-0.55353
6	0.97839...	-0.54518
7	0.98154...	-0.53905
8	0.98389...	-0.53437
9	0.98572...	-0.53068
10	0.98717...	-0.52770

## 4.6 Summary

We studied computational aspects of the positive semidefinite Grothendieck problem with rank- $r$  constraint (Problem 4.1). We showed that:

1. There is an efficient randomized approximation algorithm, Algorithm 4.1 for this problem that achieves approximation ratio  $\gamma(r) = 1 - \Theta(1/r)$ .
2. This approximation ratio can be improved to  $2/(\pi\gamma(n))$  when the matrix has size  $n$ -by- $n$  for the case  $r = 1$ .
3. Assuming the Unique Games Conjecture, there is no polynomial-time approximation algorithm with approximation ratio  $\gamma(r) + \varepsilon$  for any  $\varepsilon > 0$  independent of the matrix size.

The results of this chapter show that there is a relatively small ratio between  $\text{SDP}_r(A)$  and  $\text{SDP}_\infty(A)$ . Fortunately, this leaves just enough room for two interesting consequences: the existence of XOR games that can serve to test Hilbert space dimension of entangled states (see Chapter 3) and the existence of efficient and accurate approximation algorithms (the results of this chapter).