



## 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 3

---

# A generalized Grothendieck constant and nonlocal games that require high entanglement

The content of this chapter is based on joint work with Harry Buhrman and Ben Toner [BBT11].

### 3.1 Introduction

The Clauser-Horne-Shimony-Holt (CHSH) game, briefly introduced in Section 1.4.1, is a simple nonlocal game that classical players can win with probability no greater than 0.75, but for which entangled players can produce correlated answers such that their probability of winning is roughly 0.85. In principle, nonlocal games can thus be used to witness a key feature of quantum systems: entangled states. If a joint distribution on pairs of answers that result from local measurements on a shared quantum state, could be used to win the CHSH game with probability strictly greater than 0.75, then entanglement must have been present. Motivated by the fact that in quantum information theory, *dimensionality* of quantum systems is viewed as a fundamental resource (see for example [BKCD02, WCD08]), Brunner et al. [BPA<sup>+</sup>08] asked if a more refined deduction is also possible:

*Given a set of correlations originating from measurements on a quantum state of unknown Hilbert-space dimension, can we determine the minimal dimension necessary to produce such correlations?*

A more concrete motivation for this problem comes from quantum key distribution [BB84, Eke91], where usually security can only be proved if the assumption is made that the local dimension of the shared entangled state is known to both honest parties or that their state can be used to violate a Bell inequality [BHK05, AGM06].

In this chapter, we address the above question via a connection between local Hilbert space dimensions required to play certain two-player nonlocal games optimally, and a generalization of the Grothendieck inequality. To illustrate this connection, consider the following alternative to the “CHSH test” for the cruder problem of detecting any entanglement whatsoever. This test is based on the Grothendieck constant. The fact that the Grothendieck constant is strictly greater than 1 is established by proving that for some  $n$ -by- $n$  matrix  $A$  (for some  $n$ ) and some real unit vectors  $x_1, \dots, x_n, y_1, \dots, y_n$ , the inequality

$$\sum_{i,j=1}^n A_{ij} x_i \cdot y_j \geq K \max \left\{ \sum_{i,j=1}^n A_{ij} \chi_i \psi_j : \chi_1, \dots, \chi_n, \psi_1, \dots, \psi_n \in \{-1, 1\} \right\}, \quad (3.1)$$

holds for some real number  $K > 1$ . The first to prove this was Grothendieck himself [Gro53]. He gave an example of a matrix  $A$  for which the above inequality holds with  $K = \pi/2 = 1.5707\dots$

Suppose that we normalize Grothendieck’s example such that it can be decomposed as  $A_{ij} = \pi(i, j) \cdot \Sigma_{ij}$  for some probability distribution  $\pi$  on pairs  $\{1, \dots, n\} \times \{1, \dots, n\}$  and  $n$ -by- $n$  sign matrix  $\Sigma$  (this can be done by simply dividing each of the elements of  $A$  by  $\sum_{i,j=1}^n |A_{ij}|$ ). Then, the pair  $(\pi, \Sigma)$  defines a two-player nonlocal game as follows. A referee samples a pair  $(i, j)$  according to  $\pi$  and asks Alice question “ $i$ ”, and Bob question “ $j$ ”. Alice and Bob answer with signs  $\chi_i$  and  $\psi_j$ , respectively, and win the game if  $\chi_i \psi_j = \Sigma_{ij}$ . A simple calculation (see Section 1.4) shows that the maximum on the right-hand side of Inequality (3.1) equals the *classical bias*, defined as the maximum difference between the probability of winning and the probability of losing with classical strategies. On the other hand, Tsirelson’s Theorem (see Section 1.5) shows that the *entangled bias* is at least the value on the left-hand side of (3.1). Hence, the entangled bias of this “Grothendieck game”  $(\pi, \Sigma)$  is at least  $\pi/2$  times greater than the classical bias, and therefore, this game can be used to witness the fact that entanglement is present among Alice and Bob.

Like the CHSH game, the game described above is an XOR game, where  $\{-1, 1\}$  is used for the binary basis. Brunner et al. conjectured that the more refined problem of testing Hilbert space dimensions could also be dealt with

by considering simple nonlocal games.

**3.1.1. CONJECTURE (BRUNNER ET AL.).** *For every positive integer  $d$ , there exists a two-player XOR game  $\mathcal{G}$ , probability  $p$  and constant  $\varepsilon > 0$ , such that  $\mathcal{G}$  can be won with probability  $p$  with an entangled strategy, but any entangled strategy with local Hilbert space dimensions less than  $d$  achieves winning probability at most  $p - \varepsilon$ .*

The main result of this chapter is a proof of this conjecture. In [BPA<sup>+</sup>08] it is observed that the truth of Conjecture 3.1.1 would follow if  $K_G(q)$  is strictly increasing with  $q$ , which is plausible, but is currently unknown to be true. We avoid this issue by using the new generalization of the Grothendieck constant  $K_G(q \mapsto r)$  and proving that it is strictly greater than 1. This enables us to obtain the result with an application of Tsirelson's Theorem. For convenience, let us recall that  $K_G(q \mapsto r)$  is defined by

$$K_G(q \mapsto r) = \sup \left\{ \frac{\text{SDP}_q(A)}{\text{SDP}_r(A)} : n \in \mathbb{N}, A \in \mathbb{R}^{n \times n} \right\}$$

and that  $\text{SDP}_r(A)$  is defined by

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

## 3.2 Grothendieck's Inequality with operators

In the next section, we prove lower bounds on the constant  $K_G(q \mapsto r)$ . This is done by showing that for some matrix  $A$  and some constant  $K > 1$ , we have

$$\text{SDP}_q(A) \geq K \text{SDP}_r(A),$$

implying that  $K_G(q \mapsto r) \geq K$ . However, the matrix we consider is of a special kind that is not obviously covered in the definition of  $K_G(q \mapsto r)$ , because it has uncountably many rows and columns. Slightly more precisely, the matrix we consider has rows and columns that are indexed by real *unit vectors*. The purpose of this section is to show that this is not a problem. In fact, all lower bounds on the original Grothendieck constant were obtained by using similar kinds of infinite matrices. Moreover, the form of Grothendieck's Inequality that results from this is much closer to the form in which it was originally formulated in [Gro53]. The matrix we use in the next section is the one with which Grothendieck himself proved the first lower bound of  $\pi/2$  on  $K_G$ . The

current section will also prepare us for some more general results regarding extremal examples for our constant  $K_G(q \mapsto r)$ , given in Section 3.5.

We now describe precisely what we mean by these infinite matrices. An  $n$ -by- $n$  matrix  $A$  defines a linear operator from  $\mathbb{R}^n$  to  $\mathbb{R}^n$ , as it maps a vector  $x \in \mathbb{R}^n$  to a vector  $Ax \in \mathbb{R}^n$  via matrix-vector multiplication. Conversely, a linear operator  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  defines a real  $n$ -by- $n$  matrix given by  $(e_i \cdot Ae_j)_{i,j=1}^n$ , where  $e_1, \dots, e_n$  are the canonical unit vectors. Hence, in the definition of our constant  $K_G(q \mapsto r)$ , we could have used linear operators instead of matrices.

In the subsequent sections we will work with linear operators, instead of infinite-dimensional analogues of matrices. More specifically, we will work with linear operators that map *functions* to *functions*. These functions are of the form  $f : S^{n-1} \rightarrow \mathbb{R}$  and are continuous. Informally speaking, a function  $f$  is continuous if  $f(x)$  is close to  $f(y)$  whenever  $x$  is close to  $y$ . We formalize this by endowing  $\mathbb{R}^n$  with the metric given by the Euclidean distance  $\|x - y\|_2$ , so that it makes sense to speak of continuous functions (see for example Appendix A or [Rud86]). We denote the space of real-valued continuous functions on the real  $n$ -dimensional unit sphere by  $C(S^{n-1})$ .

The linear operators that will take the place of finite matrices in our lower bounds on  $K_G(q \mapsto r)$  are of the form  $A : C(S^{n-1}) \rightarrow C(S^{n-1})$ . Morally, we can think of such a linear operator as a matrix whose rows and columns are indexed by  $n$ -dimensional unit vectors. However, in order to be able to do so formally, we would need to deal with “problematic” cases that give rise to generalized functions (or distributions) such as the Dirac delta function (see for example [RS72, p. 148]). We choose to stick with linear operators instead of the analogues of matrices that would be needed in order to avoid that discussion.

Now that we have specified the kind of linear operators that we will use, we continue by extending the definition of  $\text{SDP}_r(A)$  for the case where  $A$  is a linear operator of the form  $A : C(S^{n-1}) \rightarrow C(S^{n-1})$ . The goal of this is to establish that  $K_G(q \mapsto r) \geq \text{SDP}_q(A) / \text{SDP}_r(A)$  for any such operator  $A$ . Let us recall that the definition of  $\text{SDP}_r(A)$  when  $A$  is an  $n$ -by- $n$  matrix is

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

The argument of this maximum can be rewritten as

$$\sum_{i,j=1}^n A_{ij} x_i \cdot y_j = \sum_{i=1}^n x_i \cdot \left( \sum_{j=1}^n A_{ij} y_j \right). \quad (3.2)$$

The first thing we do towards extending the definition of  $\text{SDP}_r$  is to give the analogue of the last term  $\sum_{j=1}^n A_{ij}y_j$  for the case of linear operators on functions defined on the  $n$ -dimensional unit sphere.

For linear operator  $A : C(S^{n-1}) \rightarrow C(S^{n-1})$  and continuous *vector-valued* function  $f : S^{n-1} \rightarrow \mathbb{R}^r$ , the expression  $Af$  should be interpreted as follows. We can view the function  $f$  as a collection of  $r$  real-valued functions  $f_1, \dots, f_r \in C(S^{n-1})$ , such that  $f(x) = (f_1(x), \dots, f_r(x))^T$ . By  $Af$  we mean that  $A$  acts on each of these  $r$  functions simultaneously, giving another continuous vector-valued function  $(Af) : S^{n-1} \rightarrow \mathbb{R}^r$  defined by  $(Af_1, \dots, Af_r)^T$ . Now, the analogue of the term  $\sum_{j=1}^n A_{ij}y_j$  above, where each  $y_j$  is an  $r$ -dimensional unit vector, is given by  $(Ag)$ , where  $g$  is a function of the form  $g : S^{n-1} \rightarrow S^{r-1}$ . The sum over  $j$  on the right-hand side of Eq. (3.2) will therefore “disappear” when we consider linear operators on functions.

The remaining sum over  $i$  appearing on the right-hand side of Eq. (3.2) will be replaced by an integral over the  $n$ -dimensional unit sphere. For this, we use the following standard tools from measure theory (see for example [Mat99, Rud86]). We let  $O(\mathbb{R}^n) = \{U \in \mathbb{R}^{n \times n} : U^T U = I\}$  denote the orthogonal group on  $\mathbb{R}^n$ . A measure  $\nu$  on  $S^{n-1}$  (which we endow with the Borel  $\sigma$ -algebra) is *rotationally invariant* if for any measurable subset  $R \subseteq S^{n-1}$  and orthogonal matrix  $U \in O(\mathbb{R}^n)$ , we have  $\nu(\{Ua : a \in R\}) = \nu(R)$ . A measure  $\nu$  on a measurable space  $\Omega$  is a *probability measure* if it is normalized so that  $\nu(\Omega) = 1$ . Let  $\omega_n$  be the (unique) rotationally invariant probability measure on  $S^{n-1}$  (see for example [MS86] for a proof of the uniqueness property).

With this, we can now extend the definition of  $\text{SDP}_r$ . For linear operator  $A : C(S^{n-1}) \rightarrow C(S^{n-1})$  and integer  $r \geq 2$ , define

$$\text{SDP}_r(A) = \sup \left\{ \int_{S^{n-1}} f(x) \cdot (Ag)(x) d\omega_n(x) : f, g : S^{n-1} \rightarrow S^{r-1} \right\}, \quad (3.3)$$

where the supremum is taken over all functions  $f, g$  that are continuous and measurable. We define  $\text{SDP}_1(A)$  as the supremum over continuous measurable functions  $f, g$  taking values in  $[-1, 1]$ . The reason for this is that the only continuous  $\{-1, 1\}$ -valued functions are constant functions. We define  $\text{SDP}_\infty(A)$  analogous to the finite setting.

The fact that it is possible to prove lower bounds on  $K_G(q \mapsto r)$  by considering linear operators on  $C(S^{n-1})$  follows directly from the following lemma.

**3.2.1. LEMMA.** *For all positive integers  $n, q, r$  with  $q > r$ , any linear operator  $A : C(S^{n-1}) \rightarrow C(S^{n-1})$  and any  $\eta > 0$ , there exists positive integer  $N = N(\eta)$  and*

real  $N$ -by- $N$  matrix  $B$  such that,

$$\frac{\text{SDP}_q(B)}{\text{SDP}_r(B)} \geq \frac{\text{SDP}_q(A)}{\text{SDP}_r(A)} - \eta. \quad (3.4)$$

We defer the proof of this lemma, which uses a standard  $\varepsilon$ -net argument, to the end of this chapter (Section 3.8), so that we can move on to prove our lower bounds on  $K_G(q \mapsto r)$ . We only mention that the converse of Lemma 3.2.1 also holds. In Section 3.5 we show that in order to prove lower bounds on  $K_G(q \mapsto r)$ , it is sufficient to restrict to linear operators on the sphere of a special kind: rotationally invariant operators.

### 3.3 Lower bounds on the generalized Grothendieck constant

In this section, we prove lower bounds on the constant  $K_G(q \mapsto r)$ .

**3.3.1. THEOREM.** *For all positive integers  $q, r$  such that  $q > r$ , we have*

$$K_G(q \mapsto r) \geq \frac{\gamma(q)}{\gamma(r)} \geq 1 + \frac{1}{2r} - \frac{1}{2q} - O\left(\frac{1}{r^2}\right),$$

where the function  $\gamma : \mathbb{R} \rightarrow \mathbb{R}$  is defined by

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

where  $\Gamma : \mathbb{R} \rightarrow \mathbb{R}$  is the Gamma function, defined by

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt.$$

The theorem follows by considering the operator  $A : C(S^{n-1}) \rightarrow C(S^{n-1})$  defined by

$$(Af)(x) = \int_{S^{n-1}} x \cdot y f(y) d\omega_n(y). \quad (3.5)$$

With this operator, Grothendieck proved the  $\pi/2$  lower bound on  $K_G$ , which we obtain by letting  $q \rightarrow \infty$  and  $r = 1$ . For this operator we can compute the value  $\text{SDP}_r(A)$  exactly, giving the bounds  $K_G(q \mapsto r) \geq \text{SDP}_q(A) / \text{SDP}_r(A)$  of Theorem 3.3.1. The value of  $\text{SDP}_r(A)$  is given in the following lemma.

**3.3.2. LEMMA.** *Let  $A : C(S^{n-1}) \rightarrow C(S^{n-1})$  be the linear operator defined in Eq. (3.5). Then, for every integer  $1 \leq r \leq n$ , we have*

$$\text{SDP}_r(A) = \frac{1}{r} \left( \frac{\Gamma(\frac{r+1}{2})\Gamma(\frac{n}{2})}{\Gamma(\frac{r}{2})\Gamma(\frac{n+1}{2})} \right)^2 = \frac{1}{n} \frac{\gamma(r)}{\gamma(n)}. \quad (3.6)$$

PROOF: We prove the lemma in two parts. First, we show that the problem of computing  $\text{SDP}_r(A)$  can be reduced to computing a particular integral over the  $n$ -dimensional unit sphere. This is the content of Claim 4. Second, we compute that integral. The result of this computation is given in Claim 5.

**4. CLAIM.** *For every  $1 \leq r \leq n$ , we have*

$$\text{SDP}_r(A) = \frac{1}{r} \left( \int_{S^{n-1}} \left( \sum_{i=1}^r x_i^2 \right)^{1/2} d\omega_n(x) \right)^2. \quad (3.7)$$

PROOF: The value  $\text{SDP}_r(A)$  for the operator  $A$  of Eq. (3.5) is given by

$$\sup \left\{ \int_{S^{n-1}} \int_{S^{n-1}} (x \cdot y) (f(x) \cdot g(y)) d\omega_n(x) d\omega_n(y) \right\}, \quad (3.8)$$

where the supremum is over functions  $f, g : S^{n-1} \rightarrow S^{r-1}$  that are measurable and continuous.

We start by rewriting the double integral in Eq. (3.8) as the trace inner-product between two  $n$ -by- $r$  matrices. Invariance of the trace function under cyclic permutations of its arguments gives the simple identity

$$(x \cdot y) (f(x) \cdot g(y)) = \text{Tr}(f(x)x^T y g(y)^T),$$

where we used  $x \cdot y = x^T y$  and  $f(x) \cdot g(y) = g(y)^T f(x)$ . By linearity of the trace function, this identity allows us to rewrite the argument of Eq. (3.8) as the trace inner-product of two  $n$ -by- $r$  matrices:

$$\left\langle \int_{S^{n-1}} x f(x)^T d\omega_n(x), \int_{S^{n-1}} y g(y)^T d\omega_n(y) \right\rangle. \quad (3.9)$$

The Cauchy-Schwarz inequality shows that this value is at most the product of the Hilbert-Schmidt norms of the two matrices. Since equality in Cauchy-Schwarz holds if and only if the matrices are scalar multiples of each other, we may assume that the functions  $f, g$  satisfy  $f = g$ . It follows that

$$\text{SDP}_r(A) = \sup \left\{ \left\| \int_{S^{n-1}} x f(x)^T d\omega_n(x) \right\|_{\text{HS}}^2 \right\}, \quad (3.10)$$

where the supremum is over measurable and continuous  $f : S^{n-1} \rightarrow S^{r-1}$ .

For arbitrary  $\varepsilon > 0$ , let  $f : S^{n-1} \rightarrow S^{r-1}$  be a measurable continuous function such that

$$\left\| \int_{S^{n-1}} x f(x)^T d\omega_n(x) \right\|_{\text{HS}}^2 \leq \text{SDP}_r(A) \leq \left\| \int_{S^{n-1}} x f(x)^T d\omega_n(x) \right\|_{\text{HS}}^2 + \varepsilon. \quad (3.11)$$

Let  $\int_{S^{n-1}} x f(x)^T d\omega_n(x) = \chi F$  where  $\chi > 0$  and  $F$  is an  $n$ -by- $r$  matrix satisfying  $\|F\|_{\text{HS}} = 1$ . By the singular value decomposition, we have  $F = U^T D V$  where  $U \in O(\mathbb{R}^n)$ ,  $V \in O(\mathbb{R}^r)$  and  $D$  is a real  $n$ -by- $r$  diagonal matrix with diagonal entries  $\lambda_1 \geq \dots \geq \lambda_r \geq 0$  satisfying  $\|F\|_{\text{HS}}^2 = \lambda_1^2 + \dots + \lambda_r^2 = 1$ .

By linearity of the trace inner product, we have

$$\begin{aligned} \chi &= \left\langle \int_{S^{n-1}} x f(x)^T d\omega_n(x), F \right\rangle \\ &= \int_{S^{n-1}} \left\langle x f(x)^T, F \right\rangle d\omega_n(x) \\ &= \int_{S^{n-1}} f(x) \cdot (F^T x) d\omega_n(x). \end{aligned}$$

The Cauchy-Schwarz inequality and the fact that the  $f(x)$  has unit norm shows that the above expression is maximized if  $f$  is of the form  $f(x) = (F^T x) / \|F^T x\|_2$ , which is a normalized projection onto an  $r$ -dimensional subspace. Without loss of generality, we may assume that  $f$  is of this form. This gives

$$\chi = \int_{S^{n-1}} \|F^T x\|_2 d\omega_n(x).$$

Since both the Euclidean norm and the measure  $\omega_n$  are invariant under orthogonal transformations, the singular value decomposition of  $F$  gives

$$\begin{aligned} \int_{S^{n-1}} \|F^T x\|_2 d\omega_n(x) &= \int_{S^{n-1}} \|V^T D^T U x\|_2 d\omega_n(x) \\ &= \int_{S^{n-1}} \|D^T x\|_2 d\omega_n(x) \\ &= \chi(\lambda_1, \dots, \lambda_r), \end{aligned} \quad (3.12)$$

where

$$\chi(\lambda_1, \dots, \lambda_r) = \int_{S^{n-1}} \left( \sum_{i=1}^r \lambda_i^2 x_i^2 \right)^{1/2} d\omega_n(x).$$

It remains to show that the weights  $\lambda_1, \dots, \lambda_r$  can be taken to be equal. By invariance of  $\omega_n$  under permutations of the coordinates (which are orthogonal

transformations), we have  $\chi(\lambda_1, \lambda_2, \dots, \lambda_r) = \chi(\lambda_2, \lambda_1, \dots, \lambda_r)$ , and indeed, such an identity holds for any other permutation of the indices  $1, \dots, r$ . We now use a symmetrization argument to show that, without loss of generality, we may assume  $\lambda_1 = \dots = \lambda_r$ .

Let  $\sigma : \{1, \dots, r\} \rightarrow \{1, \dots, r\}$  be a random permutation, uniformly distributed over all  $r!$  possible choices. Let  $\bar{\lambda} = \sqrt{\mathbb{E}_\sigma[\lambda_{\sigma(1)}^2]} = 1/r$ . Then, by Jensen's inequality and concavity of the square-root function, we have

$$\begin{aligned} \chi(\lambda_1, \dots, \lambda_r) &= \mathbb{E}_\sigma[\chi(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(r)})] \\ &= \mathbb{E}_\sigma \left[ \int_{S^{n-1}} \left( \sum_{i=1}^r \lambda_{\sigma(i)}^2 x_i^2 \right)^{1/2} d\omega_n(x) \right] \\ &= \int_{S^{n-1}} \mathbb{E}_\sigma \left[ \left( \sum_{i=1}^r \lambda_{\sigma(i)}^2 x_i^2 \right)^{1/2} \right] d\omega_n(x) \\ &\leq \int_{S^{n-1}} \left( \sum_{i=1}^r \bar{\lambda}^2 x_i^2 \right)^{1/2} d\omega_n(x) \\ &= \chi(\bar{\lambda}, \dots, \bar{\lambda}), \end{aligned}$$

giving  $\chi(1/r, \dots, 1/r)^2 \leq \text{SDP}_r(A) \leq \chi(1/r, \dots, 1/r)^2 + \varepsilon$  for any  $\varepsilon > 0$   $\blacklozenge$

What is left to do in order to prove Lemma 4 is to compute the integral given in Claim 4.

**5. CLAIM.** *For every integer  $1 \leq r \leq n$ , we have*

$$\int_{S^{n-1}} \left( \sum_{i=1}^r x_i^2 \right)^{1/2} d\omega_n(x) = \frac{\Gamma(\frac{n}{2})\Gamma(\frac{r+1}{2})}{\Gamma(\frac{n+1}{2})\Gamma(\frac{r}{2})}. \quad (3.13)$$

**PROOF:** For  $\phi, \theta_1, \dots, \theta_{n-2}$  the angles of the hyperspherical coordinate system for  $\mathbb{R}^n$ , we have that the volume element  $d\omega_n$  can be decomposed as

$$\frac{\pi^{r/2}\Gamma(\frac{n}{2})}{\pi^{n/2}\Gamma(\frac{r}{2})} \sin^{n-2} \theta_{n-2} \sin^{n-3} \theta_{n-3} \cdots \sin \theta_{r-1} d\theta_{n-2} d\theta_{n-3} \cdots d\theta_{r-1} d\omega_r$$

(see for example [AAR99, p. 456]; note that we have labeled the angles in reverse order and normalized  $\omega_n$ ). After applying a substitution of variables, this allows us to write the left-hand side of Eq. (3.13) as

$$\frac{\pi^{r/2}\Gamma(\frac{n}{2})}{\pi^{n/2}\Gamma(\frac{r}{2})} \left( \prod_{i=1}^{n-r-1} \int_{-1}^1 (1-t_i^2)^{(n-2-i)/2} dt_i \right) \left( \int_{S^{r-1}} \left( \sum_{i=1}^r x_i^2 \right)^{1/2} d\omega_r(x) \right)$$

The integral over  $S^{r-1}$  equals  $\omega_r(S^{r-1}) = 1$ , as its integrand is simply the Euclidean norm of the vector  $x$ . The remaining product of integrals can be dealt with using the following version of the Beta integral (see for example [AAR99, Eq. (1.1.12)])

$$B(\alpha, \beta) = \int_0^1 s^{2\alpha-1} (1-s^2)^{\beta-1} ds = \frac{\Gamma(\alpha)\Gamma(\beta)}{2\Gamma(\alpha+\beta)}. \quad (3.14)$$

Setting  $\alpha$  and  $\beta$  to the appropriate values,  $\alpha = 1/2$  and  $\beta = (n-i)/2$ , gives

$$\begin{aligned} \prod_{i=1}^{n-r-1} \int_{-1}^1 (1-t_i^2)^{(n-2-i)/2} dt_i &= 2^{n-r-1} \prod_{i=1}^{n-r-1} \int_0^1 (1-t_i^2)^{(n-2-i)/2} dt_i \\ &= \pi^{(n-r-1)/2} \prod_{i=1}^{n-r} \frac{\Gamma(\frac{n-i}{2})}{\Gamma(\frac{n-i+1}{2})} \\ &= \frac{\pi^{n/2} \Gamma(\frac{r+1}{2})}{\pi^{r/2} \Gamma(\frac{n+1}{2})} \end{aligned}$$

Multiplying this by the left-over factor from above then gives result.  $\blacklozenge$

Combining the two claims gives

$$\text{SDP}_r(A) = \frac{1}{r} \left( \frac{\Gamma(\frac{r+1}{2})\Gamma(\frac{n}{2})}{\Gamma(\frac{r}{2})\Gamma(\frac{n+1}{2})} \right)^2,$$

which proves the lemma.  $\square$

With Lemma 3.3.2 in hand, the proof of Theorem 3.3.1 is straightforward.

PROOF OF THEOREM 3.3.1: By Lemma 3.3.2, we have

$$K_G(q \mapsto r) \geq \frac{\text{SDP}_q(A)}{\text{SDP}_r(A)} = \frac{\gamma(q)}{\gamma(r)}.$$

The asymptotic lower bound follows from the duplication formula for the Gamma function  $\Gamma(z)\Gamma(z+1/2) = 2^{1-2z}\sqrt{\pi}\Gamma(2z)$ , which gives [KVR90, GKP94]

$$\frac{\Gamma(z+1/2)}{\Gamma(z)} = \sqrt{z} \left( 1 - \frac{1}{8z} + \frac{1}{128z^2} + \dots \right).$$

This proves the theorem.  $\square$

Next, we show that the lower bounds established in Theorem 3.3.1 are strictly greater than 1 for all  $q > r$ . This fact follows from the following lemma.

**3.3.3. LEMMA.** *The function  $\gamma(r)$  is strictly increasing on integers  $r = 1, 2, \dots$*

PROOF: For  $r \leq 9$ , just evaluate  $\gamma(r)$ . For  $r > 9$ , we use the following bound on  $\log \Gamma(z)$  (where  $\log$  is the natural logarithm), first proved by Robbins [Rob55] for integer values of  $z$ , but which Matsunawa observed [Mat76, Remark 4.1] is also valid for real values of  $z \geq 2$ :

$$\sqrt{2\pi}z^{z+1/2}e^{-z+1/(12z+1)} < \Gamma(z+1) < \sqrt{2\pi}z^{z+1/2}e^{-z+1/(12z)}. \quad (3.15)$$

Using this bound, we obtain

$$\begin{aligned} \log \frac{\gamma(r+1)}{\gamma(r)} &= 2 \log \left(\frac{r}{2}\right) - \log \left(1 + \frac{1}{r}\right) + 4 \log \Gamma\left(\frac{r}{2}\right) - 4 \log \Gamma\left(\frac{r+1}{2}\right) \\ &\geq 2 \log \left(1 + \frac{1}{(r/2) - 2}\right) - \log \left(1 + \frac{1}{r}\right) - 2r \log \left(1 + \frac{1}{r-2}\right) + \\ &\quad \frac{4}{6r-11} + \frac{6r-8}{3r-3}. \end{aligned}$$

Now use

$$\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \frac{1}{4n^4} \leq \log \left(1 + \frac{1}{n}\right) \leq \frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3},$$

(which is valid for all  $n \geq 1$ ), and we obtain

$$\begin{aligned} \log \frac{\gamma(t+10)}{\gamma(t+9)} &\geq \left(14t^7 + 679t^6 + 13923t^5 + 155346t^4 + 1005620t^3 + \right. \\ &\quad \left. + 3684139t^2 + 6679947t + 3828140\right) / \left(3(t+7)^4(t+8)(t+9)^3(6t+43)\right), \end{aligned}$$

which is positive for  $t \geq 0$ , i.e., for  $r \geq 9$ . Thus  $\gamma(r)$  is strictly increasing.  $\square$

## 3.4 Nonlocal games that require high entanglement

In this section, we prove Conjecture 3.1.1.

**3.4.1. THEOREM.** *For every positive integer  $d$ , there exists a two-player XOR game  $\mathcal{G}$ , probability  $p$  and constant  $\varepsilon > 0$ , such that  $\mathcal{G}$  can be won with probability  $p$  with an entangled strategy if the local Hilbert space dimensions are at least  $2^{d^2+1}$ , but any entangled strategy with local Hilbert space dimensions less than  $d$  achieves winning probability at most  $p - \varepsilon$ .*

We prove Theorem 3.4.1 using Tsirelson's Theorem, which gives a correspondence relation between the entangled bias of an XOR game  $(\pi, \Sigma)$  where the players are restricted to sharing a state with local dimension at most  $d$ , and the value  $\text{SDP}_r(\pi \circ \Sigma)$  for some  $r = r(d)$  (where  $\circ$  denotes the entrywise product for matrices). For convenience, we restate Tsirelson's Theorem here.

**3.4.2. THEOREM (TSIRELSON).** (*Hard direction*) For all positive integers  $n, r$  and any real  $r$ -dimensional unit vectors  $x_1, \dots, x_n, y_1, \dots, y_n$ , there exists a positive integer  $d$  that depends on  $r$  only, a state  $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$  and  $\{-1, 1\}$ -observables  $F_1, \dots, F_n, G_1, \dots, G_n \in \mathcal{O}(\mathbb{C}^d)$ , such that for every  $i, j \in \{1, \dots, n\}$ , we have

$$\langle \psi | F_i \otimes G_j | \psi \rangle = x_i \cdot y_j.$$

Moreover,  $d \leq 2^{\lceil r/2 \rceil}$ .

(*Easy direction*) Conversely, for all positive integers  $n, d$ , state  $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$  and  $\{-1, 1\}$ -observables  $F_1, \dots, F_n, G_1, \dots, G_n \in \mathcal{O}(\mathbb{C}^d)$ , there exist a positive integer  $r$  that depends on  $d$  only and real  $r$ -dimensional unit vectors  $x_1, \dots, x_n, y_1, \dots, y_n$  such that for every  $i, j \in \{1, \dots, n\}$ , we have

$$x_i \cdot y_j = \langle \psi | F_i \otimes G_j | \psi \rangle.$$

Moreover,  $r \leq 2d^2$ .

**PROOF OF THEOREM 3.4.1:** From the previous section, we know that for every positive integer  $r$ , we have

$$K_G(r+1 \mapsto r) > 1.$$

Hence, there exists some positive integer  $n$  and real  $n$ -by- $n$  matrix  $A$  such that

$$\frac{\text{SDP}_{r+1}(A)}{\text{SDP}_r(A)} > 1. \quad (3.16)$$

Note that the existence of such a matrix follows directly from Lemma 3.2.1 and the fact that this bound holds for Grothendieck's operator, as was shown above. By suitably normalizing matrix  $A$ , we can decompose it entrywise as

$$A_{ij} = \pi(i, j) \Sigma_{ij},$$

where  $\pi : \{1, \dots, n\} \times \{1, \dots, n\} \rightarrow [0, 1]$  is a probability distribution and  $\Sigma$  is an  $n$ -by- $n$  sign matrix. Note that the pair  $(\pi, \Sigma)$  defines a two-player XOR game and that such normalization does not change the ratio (3.16).

Let us denote by  $\beta_m^*(\pi, \Sigma)$  the entangled bias attainable with a state of local dimension at most  $m$ , and by  $\beta_\infty^*(\pi, \Sigma)$  the entangled bias when there is no restriction on the dimension.

On the one hand, the easy direction of Tsirelson's Theorem shows that the bias attainable for game  $(\pi, \Sigma)$  by players who share an entangled state with local dimension  $d = \lfloor \sqrt{r/2} \rfloor$ , is at most  $\text{SDP}_r(A)$ . To see this, note that what the lemma tells us is that for every optimal  $d$ -dimensional strategy for the game, there exist real  $r$ -dimensional unit vectors  $x_1, \dots, x_n, y_1, \dots, y_n$  such that

$$\text{SDP}_r(A) \geq \sum_{i,j=1}^n A_{ij} x_i \cdot y_j = \beta_d^*(\pi, \Sigma).$$

On the other hand, the hard direction of Tsirelson's Theorem tells us that for  $D = 2^{\lceil (r+1)/2 \rceil}$ , there exist state  $|\psi\rangle \in \mathbb{C}^D \otimes \mathbb{C}^D$  and observables  $F_1, \dots, F_n, G_1, \dots, G_n \in \mathcal{O}(\mathbb{C}^D)$ , such that

$$\beta_\infty^*(\pi, \Sigma) \geq \mathbb{E}_{(i,j) \sim \pi} [\Sigma_{ij} \langle \psi | F_i \otimes G_j | \psi \rangle] = \text{SDP}_{r+1}(A).$$

Hence, we have

$$\frac{\beta_\infty^*(\pi, \Sigma)}{\beta_d^*(\pi, \Sigma)} \geq \frac{\text{SDP}_{r+1}(A)}{\text{SDP}_r(A)} > 1.$$

We conclude that entangled players can win the game  $(\pi, \Sigma)$  with probability  $p = (1 + \text{SDP}_{r+1}(A))/2$ , but *not* with a state that has local dimension strictly less than  $d$ . This completes the proof.  $\square$

We conclude this section with a couple of comments regarding Theorem 3.4.1 and its proof.

- In Theorem 3.4.1 there is an exponential separation between the local Hilbert space dimensions that can be separated by looking at the maximal bias of two-player XOR games. A result of Slofstra [Slo10] shows that this separation cannot be decreased by much.
- After a preliminary version of this result was submitted to the twelfth workshop on Quantum Information Processing (QIP 2009) on 20 October, 2008, we learned of a paper by Pál and Vértesi [PV08], who obtain similar results independently. Without explicitly defining  $K_G(q \mapsto r)$ , they prove that this quantity is strictly increasing with  $m$  when  $n \rightarrow \infty$  using essentially the same methods that we do, and use this result to confirm Conjecture 3.1.1, giving an XOR game that has an infinite number of questions; they obtain dimension witnesses with finite number of questions using different methods.

### 3.5 Invariant operators and Grothendieck's constant

The operator with which we proved lower bounds on  $K_G(q \mapsto r)$ , let's call it Grothendieck's operator, has a special property, namely that it is *rotationally invariant*. Intuitively, this means that if we were to think of the operator as a matrix whose rows and columns are indexed by  $n$ -dimensional unit vectors, then the  $(x, y)$ -entry of the matrix depends only on the inner product  $x \cdot y$ . To define more formally what it means for an operator to be rotationally invariant, let us for continuous function  $f$  on the  $n$ -dimensional unit sphere and  $n$ -by- $n$  orthogonal matrix  $U$  denote by  $f^U$  the function  $f^U(x) = f(U^T x)$ . Then, a linear operator  $A : C(S^{n-1}) \rightarrow C(S^{n-1})$  is rotationally invariant if for any continuous function  $f$  and orthogonal matrix  $U$ , we have  $(Af^U)(Ux) = (Af)(x)$ .

The main message of this section is that there exists a rotationally invariant operator  $A$  for which the ratio  $\text{SDP}_q(A) / \text{SDP}_r(A)$  equals  $K_G(q \mapsto r)$ . In order to establish tight lower bounds for  $K_G(q \mapsto r)$ , it therefore suffices to restrict our attention to rotationally invariant operators. Since all rotationally invariant operators share the same set eigenfunctions, differing only in their eigenvalue spectrum, the search space can be reduced quite dramatically. A similar fact about operators on functions on Gaussian spaces was used by Raghavendra and Steurer [RS09] to show that the exact value of  $K_G$  can be approximated to within an error  $\varepsilon$  in time  $O(\exp(\exp(1/\varepsilon^3)))$  by a linear program.

**3.5.1. LEMMA.** *For all positive integers  $n, q, r$  with  $q > r$  and any real  $n$ -by- $n$  matrix  $A$ , there exists a rotationally invariant linear operator  $B : C(S^{q-1}) \rightarrow C(S^{q-1})$  such that*

$$\frac{\text{SDP}_q(B)}{\text{SDP}_r(B)} \geq \frac{\text{SDP}_q(A)}{\text{SDP}_r(A)}.$$

The proof of this fact closely follows that of the similar statement about  $K_G$  and operators on Gaussian spaces, due to Raghavendra and Steurer [RS09].

The proof relies on the use of a linear operator that would give rise to the kind of generalized function alluded to in Section 3.2. In order to be able to introduce the operator swiftly, we fix the following notation. For  $q$ -dimensional unit vector  $x$ , let  $x^\perp$  denote the set of all  $q$ -dimensional unit vectors that are orthogonal to  $x$  and let  $\omega_{x^\perp}$  be the rotationally invariant probability measure on  $x^\perp$ . Let  $\mu_q$  be the rotationally invariant probability measure on  $O(\mathbb{R}^q)$ .

PROOF: Let  $u_1, \dots, u_n, v_1, \dots, v_n \in S^{q-1}$  be the optimal vectors for  $\text{SDP}_q(A)$ . We construct the invariant operator  $B$  using linear combinations of the auxil-

ary operator  $T_\rho : C(S^{q-1}) \rightarrow C(S^{q-1})$ , defined for all  $\rho \in [-1, 1]$  by

$$(T_\rho \chi)(x) = \int_{x^\perp} \chi(\rho x + \sqrt{1 - \rho^2} y) d\omega_{x^\perp}(y). \quad (3.17)$$

To get some intuition for this operator, observe that the value  $(T_\rho \chi)(x)$  is the average of  $\chi$  over the perimeter of a spherical cap of radius  $\sqrt{1 - \rho^2}$  with pole  $x$ . Putting  $\rho = 1$  gives the identity and  $\rho = 0$  gives the Radon transform (see for example [Hel99, KR11]). Moreover, this operator is rotationally invariant.

Now, we define  $B$  by

$$B = \sum_{i,j=1}^n A_{ij} T_{u_i \cdot v_j}.$$

Clearly, this operator is also rotation invariant.

In order to bound the value  $\text{SDP}_q(B)$  from below, consider the action of  $T_\rho$  on the linear function  $\chi$  given by  $\chi(x) = x_1$ . We have

$$\begin{aligned} (T_\rho \chi)(x) &= \int_{x^\perp} (\rho x_1 + \sqrt{1 - \rho^2} y_1) d\omega_{x^\perp}(y) \\ &= \rho x_1 \\ &= \rho \chi(x). \end{aligned}$$

Hence,  $\chi$  is an eigenfunction of  $T_\rho$  with eigenvalue  $\rho$ . It is not hard to see that in fact *any* linear function is an eigenfunction of  $T_\rho$  with eigenvalue  $\rho$ . From this, it follows that for  $f, g : S^{q-1} \rightarrow S^{q-1}$  given by  $f(x) = g(x) = x$ , we have

$$\begin{aligned} \text{SDP}_q(B) &\geq \sum_{i,j=1}^n A_{ij} \int_{S^{q-1}} f(x) \cdot (T_{u_i \cdot v_j} g)(y) d\omega_q(x) \\ &= \sum_{i,j=1}^n A_{ij} u_i \cdot v_j \int_{S^{q-1}} f(x) \cdot g(x) d\omega_q(x) \\ &= \text{SDP}_q(A), \end{aligned} \quad (3.18)$$

where we used that  $f$  and  $g$  have linear functions at each of their coordinates.

In order to bound the value  $\text{SDP}_r(B)$  from above, we use the following claim. This claim will enable us to convert optimal functions  $f', g' : S^{q-1} \rightarrow S^{r-1}$  for  $\text{SDP}_r(B)$  into a sequence of  $r$ -dimensional unit vectors for  $\text{SDP}_r(A)$ .

**6. CLAIM.** For any  $u, v \in S^{q-1}$  and  $\chi, \psi \in C(S^{q-1})$ , we have

$$\int_{S^{q-1}} \chi(x) (T_{u \cdot v} \psi)(x) d\omega_q(x) = \int_{O(\mathbb{R}^q)} \chi(U \cdot u) \psi(U \cdot v) d\mu_q(U). \quad (3.19)$$

PROOF: Set  $\rho = u \cdot v$ . As the measure  $\mu_q$  is rotationally invariant, it suffices to consider  $u = (1, 0, \dots, 0)^T$  and  $v = (\rho, \sqrt{1 - \rho^2}, 0, \dots, 0)^T$ . Let us denote an orthogonal matrix  $U$  as  $U = [x, y, z_1, \dots, z_{q-2}]$ , where  $x, y, z_1, \dots, z_{q-2} \in S^{q-1}$  are its columns. Then, for random  $U$  distributed according to  $\mu_q$ , we have that the vector  $U \cdot u = x$  is uniformly distributed over the  $q$ -dimensional unit sphere, and  $U \cdot v = \rho x + \sqrt{1 - \rho^2} y$  has the vector  $y$  uniformly distributed over the  $(q - 1)$ -dimensional unit sphere  $x^\perp$ . This shows that the right-hand side of Eq. (3.19) equals

$$\int_{S^{q-1}} \chi(x) \left( \int_{x^\perp} \psi(\rho x + \sqrt{1 - \rho^2} y) d\omega_{x^\perp}(y) \right) d\omega_q(x),$$

which in turn equals the left-hand side by the definition of  $T_\rho$ .  $\blacklozenge$

Let  $f', g' : S^{q-1} \rightarrow S^{r-1}$  be optimal functions for  $\text{SDP}_r(B)$ . Then, the claim above allows us to upper bound  $\text{SDP}_r(B)$  by

$$\begin{aligned} \sum_{i,j=1}^n A_{ij} \int_{S^{q-1}} f'(x) \cdot (T_{u_i \cdot v_j} g')(x) d\omega_q(x) &= \\ \sum_{i,j=1}^n A_{ij} \int_{O(\mathbb{R}^q)} f'(Uu_i) \cdot g'(Uv_j) d\mu_q(U) &= \\ \int_{O(\mathbb{R}^q)} \left( \sum_{i,j=1}^m A_{ij} f'(Uu_i) \cdot g'(Uv_j) \right) d\mu_q(U) &\leq \text{SDP}_r(A), \end{aligned}$$

where the last inequality follows because the last integral is a convex combination over the values attained by sequences of real  $r$ -dimensional unit vectors given by  $u'_i = f'(Uu_i)$  and  $v'_j = g'(Uv_j)$ .

The result follows by putting this together with the lower bound on  $\text{SDP}_q(B)$  given in Eq. (3.18).  $\square$

From the point of view of XOR games, the proof of the lemma shows that entangled players who may use an unbounded amount of entanglement, can use Tsirelson's Theorem in order to construct observables from the question vectors  $x$  and  $y$ , and win a game  $\mathcal{G}_B = (\pi, \Sigma)$  such that  $\pi \circ \Sigma = B$  with bias at least as large as their bias for a game  $\mathcal{G}_A = (\pi', \Sigma')$  with  $\pi' \circ \Sigma' = A$ . On the other hand, Claim 6 shows that entangled players who are restricted in the amount of entanglement they are allowed to use, can transform any strategy for game  $\mathcal{G}_B$  into a strategy for game  $\mathcal{G}_A$  by using shared randomness in the form of a uniformly distributed orthogonal matrix, which implies that their bias for game  $\mathcal{G}_B$  is at most that of game  $\mathcal{G}_A$ .

### 3.6 Open problems

Davie [Dav84] and Reeds [Ree91] (independently) showed that Grothendieck's lower bound of  $\pi/2$  on  $K_G$  can be improved to  $1.6769\dots$ . Both authors achieve this using a modification of Grothendieck's operator, which we call the Davie-Reeds operator. For  $\rho \in [0, 1]$ , the *Davie-Reeds operator*  $A_\rho : C(S^{n-1}) \rightarrow C(S^{n-1})$  is defined by

$$(A_\rho f)(x) = n \int_{S^{n-1}} x \cdot y f(y) d\omega_n(y) - \rho f(x).$$

The number  $\rho$  is a parameter that can be optimized over in order to obtain the best bounds. It is not hard to see that this operator is rotationally invariant.

There is an important difference between Grothendieck's operator and the Davie-Reeds operator. The former belongs to the class of "positive semidefinite operators" (think matrices), for which it is possible to prove upper bounds on the ratios  $\text{SDP}_\infty(A) / \text{SDP}_r(A)$  that match the lower bounds of Theorem 3.3.1 (see Chapter 4). Grothendieck's operator is thus an extreme example for this special class of operators. The Davie-Reeds operator shows that it is possible to achieve strictly larger ratios between  $\text{SDP}_\infty$  and  $\text{SDP}_1$  with *non*-positive-semidefinite operators. A natural question is: Can the Davie-Reeds operator be used to improve the lower bounds on  $K_G(q \mapsto r)$  proved in this chapter for values of  $q$  and  $r$  other than  $\infty$  and  $1$ , respectively?

### 3.7 Summary

In this chapter, we introduced a new generalization of the Grothendieck constant, which we denoted by  $K_G(q \mapsto r)$ . We proved that for any choice of positive integers  $q > r$ , it is strictly greater than 1, and used this fact to show that for any positive integer  $d$ , there exists a two-player XOR game for which the entangled bias cannot be attained if the local Hilbert space dimensions are less than  $d$ , thereby confirming a conjecture of [BPA<sup>+</sup>08].

### 3.8 Proof of Lemma 3.2.1

In this section, we prove Lemma 3.2.1, which we restate here for convenience.

**3.8.1. LEMMA.** *For all positive integers  $n, q, r$  with  $q > r$ , any linear operator  $A : C(S^{n-1}) \rightarrow C(S^{n-1})$  and any  $\eta > 0$ , there exists positive integer  $N = N(\eta)$  and real  $N$ -by- $N$  matrix  $B$  such that,*

$$\frac{\text{SDP}_q(B)}{\text{SDP}_r(B)} \geq \frac{\text{SDP}_q(A)}{\text{SDP}_r(A)} - \eta. \quad (3.20)$$

For the proof of the lemma we use an  $\varepsilon$ -net for  $S^{n-1}$ , which is a finite set of  $n$ -dimensional unit vectors  $\mathcal{Z}_\varepsilon = \{z_1, \dots, z_N\}$  that satisfies that for any  $x \in S^{n-1}$ , there exists  $z \in \mathcal{Z}_\varepsilon$  such that  $\|z - x\|_2 \leq \varepsilon$ . The following lemma gives a bound on the size of such a set. We omit a proof of this fact, which follows from a standard volume argument (see for example [Pis99, Lemma 4.10]).

**3.8.2. LEMMA.** *For every positive integer  $n$  and any  $\varepsilon > 0$  there exists an  $\varepsilon$ -net  $\mathcal{Z}_\varepsilon = \{z_1, \dots, z_N\} \subseteq S^{n-1}$  of size*

$$N \leq \left(\frac{3}{\varepsilon}\right)^n.$$

**PROOF OF LEMMA 3.8.1:** Define for continuous function  $f : S^{n-1} \rightarrow \mathbb{R}^q$  the norm  $\|f\|_\infty = \max\{\|f(x)\|_2 : x \in S^{n-1}\}$ . Without loss of generality, we may assume that  $A$  is normalized such that for all continuous  $f : S^{n-1} \rightarrow \mathbb{R}^q$ , we have  $\|Af\|_\infty / \|f\|_\infty \leq 1$ .

We define the finite operator  $B : \mathbb{R}^N \rightarrow \mathbb{R}^N$  to be a discretized version of  $A$  as follows. Let  $\mathcal{Z}_\varepsilon = \{z_1, \dots, z_N\}$  be an  $\varepsilon$ -net for the  $n$ -dimensional unit sphere, for some  $\varepsilon$  to be chosen later. Let for each  $i \in \{1, \dots, N\}$  the region  $R_i \subseteq S^{n-1}$  be the set of vectors for which point  $z_i$  of  $\mathcal{Z}_\varepsilon$  is closest in Euclidean distance (with ties distributed arbitrarily) and let  $I_{R_i} : S^{n-1} \rightarrow \{0, 1\}$  be the indicator function for region  $R_i$ . The idea is to take  $B$  of the form

$$B_{ij} = \int_{R_i} (AI_{R_j})(x) d\omega_n(x).$$

However, there is the technical problem that the indicator functions are discontinuous while  $A$  is defined to act only on continuous functions. For this, we use the fact that indicator functions on metric spaces can be approximated by continuous functions arbitrarily well (see e.g., [Rud86, p. 39]). We will denote by  $\tilde{I}_{R_j}$  an arbitrary continuous approximation of  $I_{R_j}$  that suffices for our needs and instead define

$$B_{ij} = \int_{R_i} (A\tilde{I}_{R_j})(x) d\omega_n(x).$$

We start by showing that  $\text{SDP}_q(B)$  is not much *smaller* than  $\text{SDP}_q(A)$ . (Since these quantities appear in the numerator of Eq. (3.20), there is no problem if  $\text{SDP}_q(B)$  is larger than  $\text{SDP}_q(A)$ .) To this end, let  $f, g : S^{n-1} \rightarrow S^{q-1}$  be optimal for  $\text{SDP}_q(A)$ . Trivially,

$$\text{SDP}_q(B) \geq \sum_{i,j=1}^N B_{ij} f(z_i) \cdot g(z_j).$$

Define the continuous function  $h : S^{q-1} \rightarrow S^{q-1}$  by  $h = \sum_{j=1}^N g(z_j) \tilde{I}_{R_j}$ . Then, by expanding the definition of  $B_{ij}$  in the above right-hand side, we get

$$\begin{aligned} \sum_{i,j=1}^N B_{ij} f(z_i) \cdot g(z_j) &= \sum_{i,j=1}^N \left( \int_{R_i} (A \tilde{I}_{R_j})(x) d\omega_n(x) \right) f(z_i) \cdot g(z_j) \\ &= \sum_{i=1}^N f(z_i) \cdot \int_{R_i} (Ah)(x) d\omega_n(x). \end{aligned}$$

With this, the difference  $\text{SDP}_q(A) - \text{SDP}_q(B)$  is bounded from above by

$$\sum_{i=1}^N \int_{R_i} (f(x) \cdot (Ag)(x) - f(z_i) \cdot (Ah)(x)) d\omega_n(x).$$

By our assumed normalization of operator  $A$  and the Cauchy-Schwarz inequality, we can write and bound the above integrand as

$$\begin{aligned} f(x) \cdot (Ag)(x) - f(z_i) \cdot (Ah)(x) &= \\ (f(x) - f(z_i)) \cdot (Ag)(x) + f(z_i) \cdot ((Ag)(x) - (Ah)(x)) &\leq \\ \|f(x) - f(z_i)\|_2 + \|(Ag)(x) - (Ah)(x)\|_2. \end{aligned}$$

Since the function  $f$  is continuous, we can make  $\|f(x) - f(z_i)\|_2$  arbitrarily small for every  $i \in \{1, \dots, N\}$  and  $x \in R_i$  by varying  $\varepsilon$ . Moreover, again using the normalization of  $A$ , we have that

$$\|(Ag)(x) - (Ah)(x)\|_2 = \|(A(g-h))(x)\|_2 \leq \|g-h\|_\infty,$$

which can also be made arbitrarily small by virtue of the fact that  $g$  is continuous and by suitably setting  $\varepsilon$ . Hence, for any  $\delta_1 > 0$  we can define  $B$  as above such that  $\text{SDP}_q(A) - \text{SDP}_q(B) \leq \delta_1$ .

Next, we show that  $\text{SDP}_r(B)$  cannot be much *larger* than  $\text{SDP}_r(A)$ . (Since these quantities appear in the denominator of Eq. (3.20), there is no problem if

$\text{SDP}_q(B)$  is smaller than  $\text{SDP}_q(A)$ .) To this end, let  $x_1, \dots, x_N, y_1, \dots, y_N \in S^{r-1}$  be optimal for  $\text{SDP}_r(B)$ . Then, the candidate functions  $f = \sum_{i=1}^N x_i \tilde{I}_{R_i}$  and  $g = \sum_{j=1}^N y_j \tilde{I}_{R_j}$  for  $\text{SDP}_r(A)$  give

$$\begin{aligned} \text{SDP}_r(B) &= \sum_{i,j=1}^N B_{ij} x_i \cdot y_j \\ &= \sum_{i,j=1}^N \left( \int_{R_i} (A \tilde{I}_{R_j})(x) d\omega_n(x) \right) x_i \cdot y_j \\ &= \int_{S^{n-1}} x_i I_{R_i}(x) \cdot (Ag) d\omega_n(x) \\ &\leq \text{SDP}_r(A) + \delta_2 \end{aligned}$$

for arbitrary  $\delta_2 > 0$  depending on the choice of  $\tilde{I}_{R_i}$ , since the function  $f$  can be made to approximate the (discontinuous) function  $x_i I_{R_i}(x)$  arbitrarily well.

In conclusion, we have that for any  $\delta_1, \delta_2 > 0$ , there exist positive integer  $N$  and finite operator  $B : \mathbb{R}^N \rightarrow \mathbb{R}^N$  such that,

$$\begin{aligned} \text{SDP}_q(B) &\geq \text{SDP}_q(A) - \delta_1 \\ \text{SDP}_r(B) &\leq \text{SDP}_r(A) + \delta_2, \end{aligned}$$

from which the claim follows by taking the ratios of the two inequalities.  $\square$