Grothendieck inequalities, nonlocal games and optimization

Briët, J.

Publication date
2011
Document Version
Final published version

Grothendieck Inequalities, Nonlocal Games and Optimization

Jop Briët

Jop Briët Grothendieck Inequalities, Nonlocal Games and Optimization

Grothendieck Inequalities, Nonlocal Games and Optimization

Jop Briët

Grothendieck Inequalities, Nonlocal Games and Optimization

ILLC Dissertation Series DS-2011-07

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 904
1098 XH Amsterdam
phone: +31-20-525 6051
fax: +31-20-525 5206
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

Grothendieck Inequalities, Nonlocal Games and Optimization

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. D.C. van den Boom
ten overstaan van een door het college voor
promoties ingestelde commissie, in het openbaar
te verdedigen in de Agnietenkapel
op donderdag 27 oktober 2011, te 12.00 uur

door

Jop Briët

geboren te Leiden.

Promotiecommissie:

Promotor:
prof.dr. H. M. Buhrman

Overige leden:
prof. dr. F. A. Bais
prof. dr. M. Laurent
dr. D. Pérez-García
dr. O. Regev
prof. dr. C. J. M. Schoutens
dr. F. Vallentin
prof. dr. R. M. de Wolf

Faculteit der Natuurwetenschappen, Wiskunde en Informatica
Universiteit van Amsterdam
Science Park 904
1098 XH Amsterdam

The investigations were performed at the Centrum Wiskunde & Informatica (CWI) and were 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 EU QCS grant.

Printed and bound by Ipskamp Drukkers.

ISBN: 978-90-5776-228-4

to my family

Contents

Acknowledgments

First of all I would like to thank Harry Buhrman, my PhD advisor. Being a student in his group has been one of the most exciting experiences of my life. He got me involved in fascinating research projects, introduced me to many other interesting researchers and gave me the opportunity to explore many beautiful topics in mathematics and computer science. During our discussions he not only taught me valuable tricks of the trade, but he also instilled in me greater and greater levels of tolerance for silly jokes of which I am the subject. Before coming to CWI I was a research assistant for Peter Høyer at the University of Calgary. If I had not met Peter I would have missed out on some great personal inspiration to do scientific research. I am truly grateful for his kind guidance and generosity. For willing to be on my PhD committee and giving many very helpful comments on an earlier version of this thesis I thank Harry Buhrman, Sander Bais, Monique Laurent, David Pérez-García, Oded Regev, Kareljan Schoutens, Frank Vallentin and Ronald de Wolf.

My other colleagues at CWI are also largely responsible for having made my last four years so enjoyable. Ronald de Wolf often gave me much-needed advice on research, scientific writing and giving presentations. For the last few years he has passed on his newspaper to me so I could save some money and help my poor state of knowledge on world affairs. I also very much enjoyed my interactions, both scientific and informal, with Sourav Chakraborty, David García-Soriano, Peter Harremoës, Monique Laurent, Bruno Lof, Arie Matsliah, Fernando de Olivera-Filho, Giannicola Scarpa, Christian Schaffner, Florian Speelman, Ben Toner, Falk Unger, Frank Vallentin, Antonis Varvitsiotis and Stephanie Wehner. It was an honor to have collaborated on papers with Harry Buhrman, Sourav Chakraborty, David García-Soriano, Peter Har-

remoës, Troy Lee, Arie Matsliah, Fernando de Oliveira-Filho, Ben Toner, Frank Vallentin, Thomas Vidick and Ronald de Wolf. In particular I would like to thank Harry, Ben, Fernando, Frank, Thomas and Troy, who worked with me on the papers that form most of this thesis. I thank Peter Høyer, David Pérez-García and Oded Regev for hosting me during interesting and fun research visits at the University of Calgary, Universidad Complutense, LRI and ENS. I would also like to thank the CWI support staff, in particular the librarians for their help with all the obscure books I wanted to read and Maarten Dijkema for his IT support and hitchhiking to Fontaine Bleau to come bouldering.

Last but not least, I thank my family and friends for their help and for giving me so many things to enjoy outside of my studies.

Amsterdam Jop Briët
August, 2011.

Chapter 1

Nonlocal Games and Optimization

1.1 Introduction

Nonlocal games. To gain a better understanding of the physical world, physicists use mathematical frameworks to model it. Within these frameworks it is often possible to describe everything from interactions of a subatomic particles to the orbits of planets flying around the sun. Such frameworks can be used to predict what can or cannot happen in certain real-world situations and the quality of a framework can be measured by how well its their predictions match what is actually observed. Two of the most important frameworks are Classical Mechanics and Quantum Mechanics, the latter being a refinement of the former. Einstein's General Relativity is another celebrated framework, but it does not play a role in this thesis. A large part of this thesis is devoted to studying within the frameworks of both Classical and Quantum Mechanics an abstraction of a physical experiment called a *nonlocal game*, introduced first by Cleve, Høyer, Toner and Watrous [CHTW04].[1] The main reason for considering these games is that they provide an excellent way to study the most important feature unique to Quantum Mechanics: *entanglement*. A nonlocal game involves two or more players who are not allowed to communicate with each other, but do interact with an extra party usually referred to as the referee. At the start of the game the referee asks each of the players a question, upon which they each reply to him with some answer. Then, the referee decides if the players win or lose based only on the questions he asked and the answers

[1] The organization of the bibliography in the back of this thesis follows alphabetical order of the abbreviations used for references in the text.

he received. The players know in advance what set of answers would cause them to win, which of course is their objective. The catch is that they only know the question that was aimed directly at them and not any of the other players' questions, so they may not have enough information to know what to answer in order to win. The players thus don't play against each other, but rather have to try to coordinate their strategies to win. The way we study nonlocal games in the frameworks of Classical and Quantum Mechanics is by analyzing the winning probabilities for optimal strategies. Probabilities come into play here because we assume that the referee randomly picks the questions and because the players' strategies may involve some random processes. It turns out that the best course of action for players who live in a world described by Classical Mechanics is the simplest kind imaginable: decide before the game begins what to answer to each question and stick with that strategy throughout the game. In a Quantum Mechanical world, more sophisticated strategies sometimes give better results. Each player can base their answer on the outcome of an experiment done on some private physical system. Such an experiment may be, for example, measuring the orientation of the intrinsic magnetic field (the spin) of an electron. Such strategies typically give rise to some randomness in the players' answers, meaning that what a player answers to a particular question is not determined in advance. But this is not what separates quantum strategies from classical strategies. The key feature of quantum strategies is that they can cause the players to produce answers that are *correlated* in ways that are impossible in a classical world, as was shown for the first time by Bell [Bel64] in a slightly different language. Physical systems that allow players to obtain such correlations are said to be *entangled*. The fact that Quantum Mechanics predicts such a phenomenon was used by Einstein, Podolski and Rosen [EPR35] to argue that this framework must be incomplete, because according to them entanglement could not be part of a reasonable description of Nature. Surprisingly, experiments done by Aspect et al. [AGR81, ADR82, AGR82] gave convincing evidence that the world we live does in fact allow for this!

Optimization. And now for something completely different. An important type of problem in computer science is that of *optimization* under constraints. One example of such a problem is finding an optimal strategy for a nonlocal game, subject to the constraint that the strategy obeys the rules of Classical Mechanics (classical strategies). For a typical nonlocal game, finding an optimal classical strategy may involve searching over a huge number of possibilities.

For example, if in a two-player nonlocal game the referee can choose from n different questions and the players can choose from two possible answers per questions, then there are 2^{2n} possible (deterministic) classical strategies. Another example of such a problem that we will encounter in this thesis originates from (classical) statistical physics. Here, the problem is to optimize spatial configurations of interacting particles so as to minimize the energy of the total system. A spin can point in two possible directions, so in an array of n particles there are 2^n possible configurations. Again a huge number of possibilities to search over. Problems like the above two likely can't be solved exactly by any computer in a reasonable amount of time, where time is measured by the number of elementary steps a computer makes and where by "reasonable time" we mean a polynomial number of steps in the size of the problem.[2] The next-best thing to exactly solving an optimization problem in polynomial time is to *approximate* it. In this case we are willing to settle for any solution (e.g., a strategy for a nonlocal game or a configuration of spins) that is near-optimal, but can be found in a reasonable amount of time. A computer algorithm that finds such a solution in polynomial time is referred to as a polynomial-time approximation algorithm. The second major theme in this thesis deals with analyzing new approximation algorithms for a general type of optimization problem that will allow us, for example, to approximate the kind of energy-minimization problem mentioned above.

Grothendieck Inequalities. Nonlocal games and optimization may at first sight seem to be quite unrelated. However, it turns out that the problems discussed above can be treated in a very similar fashion, using mathematical tools we call *Grothendieck Inequalities*. This name derives from the fact that these tools have their origin in a celebrated paper of Grothendieck [Gro53]. Grothendieck Inequalities are the fibers pulling the other topics in this thesis together.

1.2 Quantum information theory

In this section, we give some basic mathematical background information on the aspects of quantum information theory relevant to this thesis. More infor-

[2] In more technical terms, if P$\neq$NP then there exists no polynomial-time algorithm for these problems. This follows from a translation of specific instances these problems to one of Karp's [Kar72] NP-complete problems. In fact, Håstad [Hås99] showed that the situation regarding these problems is even gloomier. The details of his result will be discussed later.

mation can be found in Appendix A, the book of Nielsen and Chuang [NC00] and the excellent lectures notes of Watrous [Wat08].

1.2.1 States and quantum systems

A *state* is a complex positive semidefinite matrix ρ that satisfies $\mathrm{Tr}(\rho) = 1$. Any n-by-n state ρ can be decomposed as

$$\rho = \sum_{i=1}^{n} \lambda_i |\psi_i\rangle\langle\psi_i|,$$

where $\lambda_1, \ldots, \lambda_n \geq 0$ are its eigenvalues and $|\psi_1\rangle, \ldots, |\psi_n\rangle$ are corresponding eigenvectors, which follows from the Spectral Theorem and positive semidefiniteness. A state ρ is *pure* if it has rank 1, that is, if $\rho = |\psi\rangle\langle\psi|$ for some complex unit vector $|\psi\rangle$. The trace of a positive semidefinite matrix equals the sum of its eigenvalues. Hence, a state is a convex combination of pure states. A state with rank greater than 1 is sometimes referred to as a *mixed state*. It is common to refer to a complex unit vector $|\psi\rangle$ as a state. What is implicitly referred to in this case is the pure state $\rho = |\psi\rangle\langle\psi|$. We will follow this custom when we are working in the context of quantum information theory.

Although a state can be treated as a purely mathematical object, it should be thought of as describing the configuration of some *quantum system*, which is an abstract physical object, or collection of objects, on which one can perform experiments. Associated with a quantum system X is a positive integer n and a copy of the vector space $\mathbb{C}^n$. The possible configurations of X are given by the states in $\mathbb{C}^{n\times n}$. The reason why we associate $\mathbb{C}^n$ with a quantum system instead of $\mathbb{C}^{n\times n}$ is that we will be working mostly with pure states. The integer n is referred to as the *dimension*, or *Hilbert space dimension* of X. A quantum system X is said to be *in* state ρ.

1.2.2 Measurements and observables

Let n be a positive integer and $\mathcal{A}$ be a finite set. A *measurement* on an n-dimensional quantum system with *outcomes* in $\mathcal{A}$ is defined by a set of positive semidefinite matrices $\{F^a\}_{a\in\mathcal{A}} \subseteq \mathbb{C}^{n\times n}$ that satisfy

$$\sum_{a\in\mathcal{A}} F^a = I.$$

If the matrices F^a also satisfy $F^a F^b = \delta_{ab} F^a$ for every $a, b \in \mathcal{A}$, then they define a *projective measurement*.

A measurement represents an experiment that one can perform on a quantum system. A measurement $\{F^a\}_{a\in\mathcal{A}} \subseteq \mathbb{C}^{n\times n}$ performed on an n-dimensional quantum system in state $\rho \in \mathbb{C}^{n\times n}$ yields a random variable χ over the set $\mathcal{A}$ whose probability distribution is given by

$$\Pr[\chi = a] = \mathrm{Tr}(\rho F^a).$$

The random variable χ is referred to as the *measurement outcome*.

If the set $\mathcal{A}$ consists of real numbers, then the expected value of the random variable resulting from a projective measurement $\{F^a\}_{a\in\mathcal{A}}$ is given by

$$\mathbb{E}[\chi] = \sum_{a\in\mathcal{A}} a\mathrm{Tr}(\rho F^a) = \mathrm{Tr}\Big(\rho\big(\sum_{a\in\mathcal{A}} aF^a\big)\Big). \tag{1.1}$$

The matrix $\sum_{a\in\mathcal{A}} aF^a$ appearing on the right-hand side of Eq. (1.1) is then called the *observable* associated to the projective measurement $\{F^a\}_{a\in\mathcal{A}}$.

We will mostly work with observables associated to projective measurements with only two outcomes. A $\{-1,1\}$*-valued observable* is an observable corresponding to a projective measurement with outcomes in the set $\{-1,1\}$. We denote the set of $\{-1,1\}$-valued observables in $\mathbb{C}^{n\times n}$ by $\mathcal{O}(\mathbb{C}^n)$.

We note the following useful fact about $\{-1,1\}$-valued observables, which we use again later on. It follows from the definition that such an observable can be written as the difference $F^+ - F^-$ of two orthogonal projectors. Squaring such an observable thus gives

$$(F^+ - F^-)^2 = F^+ + F^- = I.$$

A $\{-1,1\}$-valued observable is therefore both Hermitian and unitary. Since any matrix that is Hermitian and unitary has its eigenvalues in $\{-1,1\}$, the converse is also true.

1.2.3 Entangled states and local measurements

A quantum system X may consist of subsystems $\mathsf{X}_1,\ldots,\mathsf{X}_N$. In this case, we associate with each subsystem X_i a copy of the vector space $\mathbb{C}^{n_i}$, and we associate with X the vector space $\mathbb{C}^{n_1}\otimes\cdots\otimes\mathbb{C}^{n_N}$. So, if X is in state ρ then ρ is a matrix of size $n_1\cdots n_N$.

The subsystems may be distributed among N parties, who may be located at different places anywhere in the universe. If the overall quantum system X is in state ρ, then we say that the parties *share* the state ρ.

If the first party performs measurement $\{F^{a_1}\}_{a_1 \in \mathcal{A}_1} \subseteq \mathbb{C}^{n_1 \times n_1}$ on her subsystem X_1, while the second party performs measurement $\{F^{a_2}\}_{a_2 \in \mathcal{A}_2} \subseteq \mathbb{C}^{n_2 \times n_2}$ on his subsystem X_2, etc., then the joint probability distribution of the N measurement outcomes $\chi_1, \chi_2, \ldots, \chi_N$ is, by definition, given by

$$\Pr[\chi_1 = a_1, \chi_2 = a_2, \ldots, \chi_N = a_N] = \mathrm{Tr}(\rho F_1^{a_1} \otimes F_2^{a_2} \otimes \cdots \otimes F_N^{a_N}).$$

A pure state $|\psi\rangle \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_N}$ is a *product state* if it is of the form

$$|\psi\rangle = |\psi_1\rangle|\psi_2\rangle \cdots |\psi_N\rangle.$$

(Tensor product symbols are usually omitted when using Dirac notation.) If $|\psi\rangle$ is not a product state then it is said to be *entangled*. If a mixed state is a convex combination of pure product states then it is *separable*. The most famous entangled state is the so-called EPR pair

$$|\mathrm{EPR}\rangle = \frac{|0\rangle|0\rangle + |1\rangle|1\rangle}{\sqrt{2}} \in \mathbb{C}^2 \otimes \mathbb{C}^2,$$

named after Einstein, Podolski and Rosen. This is a pure state of a pair of two-dimensional quantum systems (usually referred to as *qubits*).

The most important difference between pure product states and pure entangled states is that the former type always gives rise to product distributions on local measurement outcomes, while this may not be the case for the latter type of states. In other words, product states give uncorrelated measurement outcomes, but entangled states can give correlated measurement outcomes.

Suppose that two parties, call them Alice and Bob, share a bi-partite product state $|\psi\rangle = |\psi_A\rangle|\psi_B\rangle$ and perform measurements $\{F^a\}_{a \in \mathcal{A}}$ and $\{G^b\}_{b \in \mathcal{B}}$ on their respective quantum systems. Then, the probability that Alice's measurement outcome χ_A is a and Bob's measurement outcome χ_B is b, equals

$$\begin{aligned} \mathrm{Tr}\big(|\psi\rangle\langle\psi| F^a \otimes G^b\big) &= \langle\psi| F^a \otimes G^b |\psi\rangle \\ &= \langle\psi_A| \langle\psi_B| F^a \otimes G^b |\psi_A\rangle |\psi_B\rangle \\ &= \langle\psi_A| F^a |\psi_A\rangle \langle\psi_B| G^b |\psi_B\rangle. \end{aligned} \tag{1.2}$$

Since $\langle\psi_A| F^a |\psi_A\rangle$ is the probability of Alice obtaining a and $\langle\psi_B| G^b |\psi_B\rangle$ is the probability of Bob obtaining b, it follows that the distribution defined by Eq. (1.2) is a product distribution and in particular, that the measurement outcomes are uncorrelated.

Below, we give some examples in which parties produce correlated measurement outcomes using entangled states.

1.3 Nonlocal games

A two-player nonlocal game is defined by four finite sets $\mathcal{A}, \mathcal{B}, \mathcal{S}$ and $\mathcal{T}$, a joint probability distribution $\pi : \mathcal{S} \times \mathcal{T} \to [0,1]$ and a map $V : \mathcal{A} \times \mathcal{B} \times \mathcal{S} \times \mathcal{T} \to \{0,1\}$. The map V is usually referred to as the *predicate*. As the underlying sets are implicit in the probability distribution π and the predicate V, a nonlocal game can be uniquely defined by π and V.

A nonlocal game $\mathcal{G} = (\pi, V)$ involves three parties: A person called the *referee* and two players, usually called Alice and Bob. The probability distribution and predicate are known to the three parties in advance. Before the game begins, Alice and Bob may come together to decide on a strategy to play the game. But after the game has begun, they are not allowed to communicate with each other anymore.

At the start of the game, the referee picks a pair $(s,t) \in \mathcal{S} \times \mathcal{T}$ according to the probability distribution π, and sends s to Alice and t to Bob. Based on their strategies, the two players then answer the referee with $a \in \mathcal{A}$ and $b \in \mathcal{B}$, respectively. The players win the game if $V(a,b,s,t) = 1$, and lose otherwise. The players' objective is of course to maximize their chance of winning.

1.3.1 Classical strategies

A *deterministic classical strategy* refers to a strategy where the players simply use deterministic maps $a : \mathcal{S} \to \mathcal{A}$ and $b : \mathcal{T} \to \mathcal{B}$ to decide what to answer the referee after receiving their questions. In this case, their probability of winning a nonlocal game $\mathcal{G} = (\pi, V)$ is given by

$$\mathbb{E}_{(s,t)\sim\pi}\big[V\big(a(s), b(t), s, t\big)\big].$$

A slightly more sophisticated classical strategy involves shared and private randomness. Here, the players flip coins (some of which both can see and others that are private) to determine their answers. However, since such a course of action results in a probability distribution over deterministic classical strategies, it cannot increase the maximal chance of winning (see for example [CHTW04]).

1.3.2 Entangled strategies

We will contrast classical strategies with entangled strategies, in which Alice and Bob may share an entangled state on which they perform local measurements to determine their answers.

An entangled strategy consists of a positive integer n, a pair of n-dimensional quantum systems X_A and X_B in some entangled state ρ and measurements $\{F_s^a\}_{a\in\mathcal{A}}$ and $\{G_t^b\}_{b\in\mathcal{B}} \subseteq \mathbb{C}^{n\times n}$. The system X_A belongs to Alice and the system X_B to Bob. The players thus share the entangled state ρ.

Upon receiving question s, Alice performs measurement $\{F_s^a\}_{a\in\mathcal{A}}$ on X_A, and upon receiving question t, Bob performs measurement $\{G_t^b\}_{b\in\mathcal{B}}$ on X_B. The answers that Alice and Bob send back to the referee are their measurement outcomes. Since the probability that Alice answers a and Bob answers b is given by $\mathrm{Tr}(\rho F_s^a \otimes G_t^b)$, their probability of winning the game equals

$$\mathbb{E}_{(s,t)\sim\pi}\Big[\sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}} \mathrm{Tr}(\rho F_s^a \otimes G_t^b)\, V(a,b,s,t)\Big].$$

It follows easily from linearity of the trace function and the fact that states are convex combinations of pure states, that pure entangled states suffice in order to maximize the winning probability with an entangled strategy. Additionally, in order to possibly have any advantage over classical classical strategies, the state ρ should be entangled, as separable states give rise to random uncorrelated answers, that is, randomized classical strategies.

1.4 Two-player XOR games

An XOR game is a nonlocal game in which the answer sets $\mathcal{A}$ and $\mathcal{B}$ are $\{0,1\}$ and the predicate V depends only on the exclusive-OR (XOR) of the answers given by the players and the value of a boolean function $f : \mathcal{S}\times\mathcal{T} \to \{0,1\}$. More precisely, the predicate is given by $V(a,b,s,t) = [a\oplus b = f(s,t)]$ where the square brackets denote the $0/1$ truth value of the statement.

The truth table of the XOR function is as follows:

$\oplus$	0	1
0	0	1
1	1	0

An XOR game is thus defined by a pair $\mathcal{G} = (\pi, f)$ consisting of a probability distribution π and boolean function f.

The bias and the violation ratio. In an XOR game, the players (quantum or classical) can always win with probability $1/2$ by answering every question simply by flipping an unbiased coin. For the case of XOR games it therefore

makes more sense to look at the amount by which the maximum winning probability is bounded away from $1/2$.

We define the *classical bias* of an XOR game $\mathcal{G}$ to be the difference between the probability of winning and the probability of losing for optimal classical strategy. We denote the classical bias by $\beta(\mathcal{G})$. We define the *entangled bias* similarly, and denote it by $\beta^*(\mathcal{G})$. The (classical or entangled) bias then equals *twice* the amount by which the maximal classical winning probability is greater than $1/2$. The reason to consider this definition is given in the next paragraph.

As a measure of the advantage entangled strategies give over classical strategies we define the *violation ratio* of $\mathcal{G}$ to be the fraction $\beta^*(\mathcal{G})/\beta(\mathcal{G})$.

Signs and observables. XOR games are more easily analyzed using the $\{-1,1\}$-basis instead of the $\{0,1\}$-basis for boolean-valued objects. Let (π, f) be some XOR game. For any classical strategy $a : \mathcal{S} \to \{0,1\}$ and $b : \mathcal{T} \to \{0,1\}$, the bias is given by the probability under π that $a(s) \oplus b(t) = f(s,t)$ minus the probability under π that $a(s) \oplus b(t) \neq f(s,t)$. Concisely, the bias equals

$$\begin{aligned}\mathbb{E}_{(s,t)\sim\pi}\Big[(-1)^{[a(s)\oplus b(t)=f(s,t)]}\Big] &= \mathbb{E}_{(s,t)\sim\pi}\Big[(-1)^{a(s)\oplus b(t)+f(s,t)}\Big] \\ &= \mathbb{E}_{(s,t)\sim\pi}\Big[(-1)^{a(s)}(-1)^{b(t)}(-1)^{f(s,t)}\Big].\end{aligned}$$

Hence, if we define sign matrix $\Sigma_{st} = (-1)^{f(s,t)}$ and functions $\chi(s) = (-1)^{a(s)}$ and $\psi(t) = (-1)^{b(t)}$, the bias becomes

$$\mathbb{E}_{(s,t)\sim\pi}\Big[\chi(s)\psi(t)\,\Sigma_{st}\Big].$$

Let us now consider an entangled strategy consisting of a shared (pure) entangled state $|\psi\rangle$ and projective measurements $\{F_s^0, F_s^1\}$ and $\{G_t^0, G_t^1\}$. The probability that Alice answers bit a upon receiving question s and Bob answers bit b upon receiving question t equals $\langle\psi|F_s^a \otimes F_t^b|\psi\rangle$. Hence, the expected value of the sign $(-1)^{a\oplus b}$ equals

$$\begin{aligned}&\Pr[a=b] - \Pr[a\neq b] = \\ &\quad \langle\psi|F_s^0\otimes G_t^0|\psi\rangle + \langle\psi|F_s^1\otimes G_t^1|\psi\rangle - \langle\psi|F_s^0\otimes G_t^1|\psi\rangle - \langle\psi|F_s^1\otimes G_t^0|\psi\rangle = \\ &\qquad\qquad \langle\psi|(F_s^0 - F_s^1)\otimes(G_t^0 - G_t^1)|\psi\rangle.\end{aligned}$$

Defining the $\{-1,1\}$-valued observables $F_s = F_s^0 - F_s^1$ and $G_t = G_t^0 - G_t^1$, we get that the bias based on this strategy equals

$$\mathbb{E}_{(s,t)\sim\pi}\Big[\langle\psi|F_s\otimes G_t|\psi\rangle\,\Sigma_{st}\Big].$$

We will often replace the boolean function f by the matrix Σ, and say that the pair (π, Σ) defines an XOR game. By the above calculations, the classical bias of such a game is given by

$$\max\left\{\mathbb{E}_{(s,t)\sim\pi}\left[\Sigma_{st}\,\chi(s)\psi(t)\right] : \chi : \mathcal{S} \to \{-1,1\},\, \psi : \mathcal{T} \to \{-1,1\}\right\}$$

and the entangled bias is given by

$$\sup_{n\in\mathbb{N}}\left\{\mathbb{E}_{(s,t)\sim\pi}\left[\Sigma_{st}\langle\psi|F_s \otimes G_t|\psi\rangle\right] : |\psi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^n,\, F_s, G_t \in \mathcal{O}(\mathbb{C}^n)\right\}.$$

The supremum is used in the entangled bias because the possibility exists that the maximal winning probability increases indefinitely with the dimension of the quantum systems.

This reformulation will prove to be a great convenience later on. The reason why we only considered projective measurements is that general measurements do not give an advantage over projective measurements, as shown by Cleve, Høyer, Toner and Watrous [CHTW04, Proposition 2].

1.4.1 The CHSH game

The CHSH game, named after Clauser, Horne, Shimony and Holt [CHSH69], is a two-player XOR with two possible questions per player, 0 and 1. The probability distribution π on $\{0,1\} \times \{0,1\}$ is the uniform distribution, so every pair of questions is asked with probability $1/4$. The predicate V evaluates to 1 if and only if $a \oplus b = s \wedge t$, where $\wedge$ denotes the AND function (which is 1 if and only if $s = t = 1$). Classical players can win this game with probability no greater than $3/4$, which can be seen by observing that the system of equations

$$\begin{aligned} a_0 \oplus b_0 &= 0 \\ a_0 \oplus b_1 &= 0 \\ a_1 \oplus b_0 &= 0 \\ a_1 \oplus b_1 &= 1 \end{aligned}$$

is overdetermined and only three equations can be satisfied simultaneously.

By sharing an EPR pair, Alice and Bob can win the CHSH game with probability $\cos(\pi/8)^2 \approx 0.85$. An entangled strategy based on $\{-1,1\}$-valued observables that achieves this is as follows. Define the matrices $X = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ and $Y = \left[\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right]$. These matrices satisfy $X^2 = Y^2 = I$, so they are observables, and they anti-commute, meaning that $XY + YX = 0$. Define Alice's observables

for questions 0 and 1 by $F_0 = X$ and $F_1 = Y$, respectively. Define Bob's observables for questions 0 and 1 to be $G_0 = (X - Y)/\sqrt{2}$ and $G_1 = (X + Y)\sqrt{2}$, respectively. The matrices X and Y should be thought of as being given in the basis $|0\rangle, |1\rangle$ in which the EPR pair

$$|\mathrm{EPR}\rangle = \frac{|0\rangle|0\rangle + |1\rangle|1\rangle}{\sqrt{2}}$$

is given. The following relations then follow easily:

$$\begin{aligned} \langle \mathrm{EPR}|X \otimes X|\mathrm{EPR}\rangle &= 1 & \langle \mathrm{EPR}|Y \otimes Y|\mathrm{EPR}\rangle &= -1 \\ \langle \mathrm{EPR}|X \otimes Y|\mathrm{EPR}\rangle &= 0 & \langle \mathrm{EPR}|Y \otimes X|\mathrm{EPR}\rangle &= 0. \end{aligned}$$

From these equations we get $\langle \mathrm{EPR}|F_s \otimes G_t|\mathrm{EPR}\rangle = (-1)^{s \wedge t}/\sqrt{2}$ for every $s, t \in \{0, 1\}$ and it follows that the bias based on the above entangled strategy equals

$$\frac{1}{4} \sum_{s,t=0}^{1} (-1)^{s \wedge t} \langle \mathrm{EPR}|F_s \otimes G_t|\mathrm{EPR}\rangle = \frac{1}{\sqrt{2}},$$

making the winning probability $1/2 + 1/(2\sqrt{2}) = \cos(\pi/8)^2$.

1.5 Tsirelson's Theorem

Tsirelson's Theorem [Tsi87] gives an extremely useful characterization of entangled strategies in two-player XOR games. It forms the basis of many results in this thesis. Roughly speaking, the theorem gives a correspondence relation between entangled strategies consisting of a shared entangled state and $\{-1, 1\}$-valued observables on the one hand, and pairs of sequences of real unit vectors on the other. The correspondence relation is given by the following theorem, which is commonly referred to as Tsirelson's Theorem. We will refer to the two parts of the correspondence as the "hard direction" and the "easy direction".

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

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

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

PROOF OF THEOREM 1.5.1: We start by proving the hard direction. Let

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(These matrices are called the *Pauli matrices*.) Note that each of them squares to the identity matrix I. This implies that they have eigenvalues in $\{-1,1\}$. Additionally, note that the last three of them, X, Y and Z, pair-wise *anti-commute*, meaning that $XY + YX = XZ + ZX = YZ + ZY = 0$.

Define for each $\ell = 1, \ldots, \lceil r/2 \rceil$, the d-by-d *Clifford matrices*,

$$\begin{aligned} S_{2\ell+1} &= Z^{\otimes(\ell-1)} \otimes X \otimes I^{\otimes(\lceil r/2 \rceil - \ell)}, \\ S_{2\ell} &= Z^{\otimes(\ell-1)} \otimes Y \otimes I^{\otimes(\lceil r/2 \rceil - \ell)}. \end{aligned}$$

From the properties satisfied by the Pauli matrices, the Clifford matrices satisfy that they square to the identity matrix (of size d-by-d) and pair-wise anti-commute. So, for every $k, \ell \in \{1, \ldots, \lceil r/2 \rceil\}$, we have $S_k S_\ell + S_\ell S_k = 2\delta_{k\ell} I$. Additionally, for every $k \neq \ell$, we have $\mathrm{Tr}(S_k S_\ell) = 0$.

Define $F_1, \ldots, F_n, G_1, \ldots, G_n \in \mathbb{C}^{d \times d}$ by

$$\begin{aligned} F_i &= \sum_{k=1}^{r} (x_i)_k S_k, \\ G_j &= \sum_{k=1}^{r} (y_j)_k S_k^T. \end{aligned}$$

1. CLAIM. *The matrices $F_1, \ldots, F_n, G_1, \ldots, G_n$ are $\{-1,1\}$-observables.*

PROOF: (Hard direction) It suffices to show that $F_i^2 = G_j^2 = I$ for each $i, j \in \{1, \ldots, n\}$, as this implies that the matrices have eigenvalues in $\{-1,1\}$. To this end, consider the expansion of F_i^2,

$$\sum_{k,\ell=1}^{r} (x_i)_k (x_i)_\ell S_k S_\ell = x \cdot x I + \sum_{k>\ell} (x_i)_k (x_i)_\ell (S_k S_\ell + S_\ell S_k).$$

From the anti-commutation relations satisfied by the Clifford matrices, it follows that the second sum on the right-hand side equals zero. What remains is the identity, as x is a unit vector.

Of course, the same argument works for G_j. This proves the claim. ♦

2. CLAIM. *For every $i, j \in \{1, \ldots, n\}$, we have* $\mathrm{Tr}\big(F_i G_j^T\big)/d = x_i \cdot y_j$.

PROOF: Fix $i, j \in \{1, \ldots, n\}$ Similarly as in the proof of the previous claim, consider the expansion of the product $F_i G_j^T$,

$$\sum_{k,\ell=1}^{r} (x_i)_k (y_j)_\ell S_k S_\ell. \tag{1.3}$$

Since $\mathrm{Tr}(S_k S_\ell) = d\delta_{k\ell}$, the only terms in (1.3) that contribute nontrivially to $\mathrm{Tr}(F_i G_j^T)$, are those for which $k = \ell$. The sum of those terms is exactly $dx \cdot y$. ♦

We now consider the expansion of $\mathrm{Tr}\big(F_i G_j^T\big)/d$. Let $\{|1\rangle, \ldots, |d\rangle\} \subseteq \mathbb{C}^d$ be an orthonormal basis for $\mathbb{C}^d$. Let

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{s=1}^{d} |s\rangle \otimes |s\rangle,$$

be the maximally entangled state.

We have

$$\begin{aligned} \langle\psi|F_i \otimes G_j|\psi\rangle &= \frac{1}{d} \sum_{s,t=1}^{d} \langle s| \otimes \langle s|F_i \otimes G_j|t\rangle \otimes |t\rangle \\ &= \frac{1}{d} \sum_{s,t=1}^{d} \langle s|F_i|t\rangle \langle s|G_j|t\rangle \\ &= \frac{1}{d} \mathrm{Tr}\big(F_i G_j^T\big). \end{aligned}$$

Combining this with the two claims then proves the hard direction.

(Easy direction) Note that since $|\psi\rangle$ has norm 1 and the observables F_i and G_j are unitary operators, $F_i \otimes I|\psi\rangle$ and $I \otimes G_j|\psi\rangle$ are unit vectors in $\mathbb{C}^{d^2}$. Additionally, note that since F_i and G_j are Hermitian, we have that the inner product

$$\big(\langle\psi|F_i \otimes I\big) \cdot \big(I \otimes G_j|\psi\rangle\big) = \langle\psi|F_i \otimes G_j|\psi\rangle,$$

is a real number. For $v \in \mathbb{C}^{d^2}$ we let $\Re(v)$ denote its real part and $\Im(v)$ its complex part, so that

$$\begin{aligned} F_i \otimes I|\psi\rangle &= \Re\big(F_i \otimes I|\psi\rangle\big) + i\Im\big(F_i \otimes I|\psi\rangle\big) \\ I \otimes G_j|\psi\rangle &= \Re\big(I \otimes G_j|\psi\rangle\big) + i\Im\big(I \otimes G_j|\psi\rangle\big). \end{aligned}$$

Define vectors $2d^2$-dimensional unit vectors x_i, y_j by

$$\begin{aligned} x_i &= \Re\big(F_i \otimes I|\psi\rangle\big) \oplus \Im\big(F_i \otimes I|\psi\rangle\big) \\ y_j &= \Re\big(G_j \otimes I|\psi\rangle\big) \oplus \Im\big(-G_j \otimes I|\psi\rangle\big) \end{aligned}$$

Then, since $\langle\psi|F_i \otimes G_j|\psi\rangle$ is a real number, we have

$$\begin{aligned} x_i \cdot y_j &= \Re\big(\langle\psi|F_i \otimes G_j|\psi\rangle\big) - \Im\big(\langle\psi|F_i \otimes G_j|\psi\rangle\big) \\ &= \langle\psi|F_i \otimes G_j|\psi\rangle, \end{aligned}$$

as desired. □

1.6 Multiplayer XOR games

By a multiplayer XOR game we generally mean an XOR game involving more than two players. For convenience, we will only consider N-player XOR games in which the question sets are all the same finite set $\mathcal{S}$. Let π be a probability distribution on $\mathcal{S}^N$ and $f : \mathcal{S}^N \to \{0,1\}$ be a boolean function. In an N-player XOR game $\mathcal{G} = (\pi, f)$, the referee picks an N-tuple of questions $(s_1, \ldots, s_N)$ according to π and sends s_1 to the first player, s_2 to the second, and so on. The players answer with $a_1, \ldots, a_N \in \{0,1\}^N$, respectively and win the game if

$$a_1 \oplus \cdots \oplus a_N = f(s_1, \ldots, s_N).$$

The classical and entangled biases are given in terms of the map $\Sigma : \mathcal{S}^N \to \{-1,1\}$ defined by $\Sigma[s_1, \ldots, s_N] = (-1)^{f(s_1,\ldots,s_N)}$. The map Σ will often be referred to as a *sign tensor* and if $N = 2$ it will be called a *sign matrix*. The classical bias of the game $\mathcal{G} = (\pi, \Sigma)$ is then given by

$$\beta(\mathcal{G}) = \max\left\{\mathbb{E}_{(s_1,\ldots,s_N)\sim\pi}\Big[\Sigma[s_1, \ldots, s_N]\chi_1(s_1)\cdots\chi_N(s_N)\Big]\right\},$$

where the maximum is taken over maps $\chi_1, \ldots, \chi_N : \mathcal{S} \to \{-1,1\}$.

Then entangled bias is given by

$$\beta^*(\mathcal{G}) = \sup\left\{\mathbb{E}_{(s_1,\ldots,s_N)\sim\pi}\Big[\Sigma[s_1, \ldots, s_N]\langle\psi|F_1(s_1) \otimes \cdots \otimes F_N(s_N)|\psi\rangle\Big]\right\},$$

where the supremum is over positive integers n, states $|\psi\rangle \in \mathbb{C}^n \otimes \cdots \otimes \mathbb{C}^n$ and observable-valued maps $F_1, \ldots, F_N : \mathcal{S} \to \mathcal{O}(\mathbb{C}^n)$.

1.6.1 Mermin's Game

Mermin [Mer90] gave a sequence of XOR games, one for every number N of players, in which the violation ratio grows exponentially with N. Entangled players can play these games perfectly by sharing an N-qubit GHZ state

$$|\mathrm{GHZ}\rangle = \frac{|0\rangle \cdots |0\rangle + |1\rangle \cdots |1\rangle}{\sqrt{2}},$$

named after its inventors Greenberger, Horne and Zeilinger [GHZ89]. Mermin's game is described as follows. The referee picks an N-bit string $x = x_1 x_2 \ldots x_N$ uniformly at random from all strings with even Hamming weight $|x|$ (i.e., the number of 1s appearing in x is even). He sends x_1 to the first player, x_2 to the second, etc. In order to win the game, the players must answer bits $a_1, \ldots, a_N$ (resp.) such that $a_1 \oplus \cdots \oplus a_N = |x|/2 \bmod 2$.

1.6.1. PROPOSITION. *The classical bias of Mermin's game is at most $2^{-(N-1)/2}$ if N is odd and at most $2^{-(N-2)/2}$ if N is even.*

PROOF: Without loss of generality, we may assume that the players use a deterministic strategy in order to play the game. Let $a_k(0)$ and $a_k(1)$ denote the answers of the k^{th} player to questions 0 and 1, respectively.

A simple calculation shows that the players' bias is given by the formula

$$\frac{1}{2^{N-1}} \sum_{x \in \{0,1\}^N :\, |x| \text{even}} (-1)^{|x|/2} (-1)^{a_1(x_1)+\cdots+a_k(x_k)} = \frac{1}{2^{N-1}} \Re \left[\prod_{k=1}^{N} \left((-1)^{a_k(0)} + i(-1)^{a_k(1)} \right) \right],$$

where $\Re$ denotes the real part of a complex number. Note that each complex number $(-1)^{a_k(0)} + i(-1)^{a_k(1)}$ has modulus $\sqrt{2}$ and argument a multiple of $\pi/4$. If N is odd, then the product of these complex numbers makes a 45 degree angle with the real axis in the complex plane, making their real part equal to $\pm 2^{(N-1)/2}$. If N is even, then their product is either parallel to the imaginary axis or parallel to the real axis. Hence, the real part of their product is at most $2^{N/2}$. Dividing by the above factor 2^{N-1} gives the result. □

1.6.2. PROPOSITION. *The entangled bias of Mermin's game is 1.*

PROOF: Let $X = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ and $Y = \left[\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right]$. As these matrices satisfy $X^2 = Y^2 = I$, they are $\{-1, 1\}$-valued observables. Moreover, they satisfy

$$X|0\rangle = |1\rangle \qquad X|1\rangle = |0\rangle$$
$$Y|0\rangle = i|1\rangle \qquad Y|1\rangle = -i|0\rangle.$$

We claim that N entangled players can play the game perfectly by measuring their respective qubits of the N-qubit GHZ state

$$|\text{GHZ}\rangle = \frac{|0\rangle \cdots |0\rangle + |1\rangle \cdots |1\rangle}{\sqrt{2}},$$

using observable $F(0) = X$ on question 0 and $F(1) = Y$ on question 1.

To see this, notice that we have

$$\begin{aligned} X \otimes X \otimes X \otimes \cdots \otimes X|\psi\rangle &= -|\text{GHZ}\rangle \\ Y \otimes Y \otimes X \otimes \cdots \otimes X|\psi\rangle &= |\text{GHZ}\rangle. \end{aligned}$$

In general, if the number of Y's that appear in the tensor products above is an odd multiple of 2, then $|\psi\rangle$ is an eigenvector with eigenvalue -1, and if the number of Y's is a multiple of 4, then $|\psi\rangle$ has eigenvalue +1. Hence, for $x \in \{0,1\}^N$ with $|x|$ even, we have

$$\bigotimes_{k=1}^{N} F(x_k)|\text{GHZ}\rangle = (-1)^{|x|/2}|\text{GHZ}\rangle.$$

The result now follows from the fact that the players' bias based on this strategy equals

$$\frac{1}{2^{N-1}} \sum_{x \in \{0,1\}^N : |x| \text{even}} (-1)^{|x|/2} \langle \text{GHZ}| \bigotimes_{k=1}^{N} F(x_k)|\text{GHZ}\rangle = 1,$$

which completes the proof. □

1.6.2 Stabilizer states

The GHZ state, defined in the previous subsection, is a special case of a general class of states known as stabilizer states. An N-qubit *stabilizer state* $|\psi\rangle$ is the unique common eigenvector of the elements of an abelian subgroup $S \subseteq \{I, X, Y, Z\}^{\otimes N}$ of order 2^N, such that $M|\psi\rangle = |\psi\rangle$ for every $M \in S$. Here, I, X, Y, Z are the 2-by-2 Pauli matrices (see Section 1.5) and the group operation of S is regular matrix multiplication. By a *tripartite stabilizer state*, we mean a stabilizer state whose qubits are distributed among three parties. These states are discussed in the context of multiplayer XOR games in Chapter 6.

1.7 Semidefinite programs and relaxations

A generic *semidefinite program* (SDP) has the following form. Given positive integers k, n, real n-by-n matrices $A, B_1, \ldots, B_k$ and real numbers $c_1, \ldots, c_k$

$$\begin{aligned} \text{maximize} \quad & \langle A, X \rangle \\ \text{subject to} \quad & X \in \mathcal{S}_n^+ \\ & \langle B_i, X \rangle = c_i, \end{aligned}$$

for $i = 1, \ldots, k$. Here $\langle C, X \rangle = \mathrm{Tr}(C^T X)$ denotes the trace inner product of the matrices C and X and $\mathcal{S}_n^+$ denotes the set of real n-by-n positive semidefinite matrices.

We will use the following standard terminology and facts of semidefinite programs (see for example the books of Grötschel, Lovász and Schrijver [GLS93] and Boyd and Vandenberghe [BV04], or the survey of Laurent and Rendl [LR05]).

The quantity $\langle A, X \rangle$ above is referred to as the *objective value* of the SDP. The conditions $X \in \mathcal{S}_n^+$ and $\langle B_i, X \rangle = c_i$ imposed on the matrix X are the *constraints*. If a matrix X satisfies all the constraints of an SDP, then it is said to be a *feasible solution*, or simply *feasible* for short. If a matrix X is a feasible and it maximizes the objective value, then it is said to be an *optimal solution* for the SDP, or *optimal* for short. The value $\langle A, X \rangle$ for optimal solution X is the *optimum* of the SDP.

The most important fact about SDPs is that their optimum can be approximated to within arbitrary fixed precision in polynomial time, as testing whether a rational matrix is positive semidefinite can be done efficiently using for example Gaussian elimination.

1.7.1 Approximation algorithms

One of the most important uses of semidefinite programs is in approximation algorithms for combinatorial optimization problems that are unknown to be solvable exactly in polynomial time. The philosophy behind such algorithms is that it is often good enough to have a solution that is close to optimal. The advantage gained by relaxing exact optimality is that near-optimal solutions can sometimes be found much faster.

We distinguish semidefinite programs from approximation algorithms by requiring from the latter that they return a feasible solution for the optimization problem they approximate. A semidefinite program which serves as a

relaxation for an optimization problem can sometimes be turned into an approximation algorithm by adding a procedure which turns an optimal solution to the SDP (some positive semidefinite matrix) into a feasible one for the optimization problem.

If the optimum of an SDP is c times the optimum of some optimization problem OPT, then we say that the SDP has *approximation ratio c* for OPT. If the output of an approximation algorithm gives a value of δ times the optimum of an optimization problem, then we say that the approximation algorithm gives a *δ-approximation*. Here, c is typically greater than 1 and δ lies in $[0, 1]$.

Below we give two examples of applications of semidefinite programs for well-known combinatorial optimization problems: the maximum cut problem and the problem of computing the chromatic number of a graph.

1.7.2 MAX CUT

The *maximum cut problem* (MAX CUT) refers to the following combinatorial optimization problem. Given an undirected graph $G = (V, E)$ with finite vertex set V and edge set $E \subset V \times V$ (with no self-loops), find a bi-partitioning of V such that the number of edges crossing the partition is maximal. Such a bi-partitioning is also referred to as a *cut*, and the number of edges crossing it as the *size* of the cut.

The MAX CUT problem is one of Karp's 21 NP-complete problems [Kar72] (see also [GJ76]). It is therefore unlikely that a polynomial-time algorithm exists that solves it exactly in the worst case. To make matters worse, Håstad [Hås99] proved that even finding a cut of size $16/17 - \varepsilon$ times the size of a maximum cut, for any constant $\varepsilon > 0$, cannot be done in polynomial time unless P=NP.

Good upper bounds on the size of a maximum cut of a graph can be found using a semidefinite program and a matrix called the Laplacian. Given a graph $G = (V, E)$, its *Laplacian* $A : V \times V \to \mathbb{R}$ is defined by

$$A(u,v) = \begin{cases} \deg(u) & \text{if } v = u \\ -1 & \text{if } \{u,v\} \in E \\ 0 & \text{otherwise,} \end{cases}$$

where $\deg(u) = |\{v \in V : \{u,v\} \in E\}|$ denotes the *degree* of vertex u.

The semidefinite program is then given by:

$$\begin{array}{ll} \text{maximize} & \frac{1}{4}\langle A, X\rangle \\ \text{subject to} & X \in \mathcal{S}_V^+ \\ & X(u,u) = 1 \text{ for every } u \in V, \end{array} \tag{1.4}$$

where $\mathcal{S}_V^+$ denotes the set of real positive semidefinite matrices whose rows and columns are indexed by the vertices of G. The fact that the optimum of this SDP upper bounds the size of a maximum cut can be shown as follows. Suppose that $S \subseteq V$ defines a cut $(S, V\backslash S)$ of maximal size. Define the function $\chi : V \to \{-1, 1\}$ by setting $\chi(u) = +1$ if $u \in S$ and $\chi(u) = -1$ otherwise. Then, the matrix $X(u,v) = \chi(u)\chi(v)$ is feasible for SDP (1.4) since it is positive semidefinite and has ones on the diagonal. For its objective value we compute

$$\begin{aligned} \langle A, X\rangle &= \sum_{u,v\in V} A(u,v)\chi(u)\chi(v) \\ &= \sum_{u\in V} \deg(u) - 2\sum_{\{u,v\}\in E} \chi(u)\chi(v) \\ &= 2\sum_{\{u,v\}\in E} \big(1 - \chi(u)\chi(v)\big). \end{aligned} \tag{1.5}$$

Each of the terms $1 - \chi(u)\chi(v)$ in the last sum equals 2 if the edge $\{u, v\}$ crosses the cut and 0 otherwise. Hence, the objective value of X is exactly the size of the maximum cut. Note that the optimum of SDP (1.4) may be higher.

In a celebrated paper, Goemans and Williamson [GW94] turned SDP (1.4) into a .878-approximation algorithm for MAX CUT, Algorithm 1.1 shown below. The description of the algorithm uses that for any $X \in \mathcal{S}_V^+$ satisfying $X(u,u) = 1$ there is a function $f : V \to S^{|V|-1}$ such that $X(u,v) = f(u) \cdot f(v)$ for every $u, v \in V$, where

$$S^{n-1} = \{x \in \mathbb{R}^n : x \cdot x = 1\}$$

denotes the real n-dimensional unit sphere (see for example Appendix A).

To analyze Algorithm 1.1 we define a function $\chi : V \to \{-1, 1\}$ by setting $\chi(u) = +1$ if u belongs to the set S returned by the algorithm and setting $\chi(u) = -1$ otherwise. Based on the vector z sampled in the algorithm we have

$$\chi(u) = \text{sign}\big(z \cdot f(u)\big).$$

Let $A : V \times V \to \mathbb{R}$ be the Laplacian matrix of the graph G given to the algorithm. By running the sequence of equations in Eq. (1.5) backwards we

Algorithm 1.1 (Goemans and Williamson) Takes as input a graph $G = (V, E)$ and returns a cut $(S, V \backslash S)$ for some $S \subseteq V$ in G.

(1) Solve SDP (1.4), obtaining a function $f : V \to S^{|V|-1}$.

(2) Sample a vector $z \in \mathbb{R}^{|V|}$ such that the entries of z are independently distributed Gaussian random variables with mean 0 and variance 1.

(3) Put $u \in S$ if and only if $z \cdot f(u) \geq 0$.

get that on expectation over the vector z, the size of the cut returned by the algorithm is given by

$$\begin{aligned} \mathbb{E}_z\left[\frac{1}{2}\sum_{\{u,v\}\in E}(1-\chi(u)\chi(v))\right] &= \mathbb{E}_z\left[\frac{1}{4}\sum_{u,v\in V}A(u,v)\chi(u)\chi(v)\right] \\ &= \frac{1}{4}\sum_{u,v\in V}A(u,v)\mathbb{E}_z\big[\chi(u)\chi(v)\big], \quad (1.6) \end{aligned}$$

where we used linearity of expectation for the second identity.

The next step of the analysis uses a useful identity often referred to as Grothendieck's Identity, as it appeared first in [Gro53, Proposition 4, p. 63].

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

$$\mathbb{E}_z[\operatorname{sign}(z \cdot x)\operatorname{sign}(z \cdot y)] = \frac{2}{\pi}\arcsin(x \cdot y).$$

PROOF: We have $\operatorname{sign}(z \cdot x)\operatorname{sign}(z \cdot y) = +1$ if and only if the vectors x and y lie on the same side of the hyperplane orthogonal to the vector z. Now we project this n-dimensional situation to the plane spanned by x and y. 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. We obtain the result by measuring regions on the unit circle and using the identity $\arcsin(t) = \pi/2 - \arccos(t)$: The probability that x and y lie on the same side of the line is $1 - \arccos(x \cdot y)/\pi$. □

Using Grothendieck's Identity and $\chi(u) = \operatorname{sign}\big(z \cdot f(u)\big)$, we get that the sum appearing on the right-hand side of Eq. (1.6) equals

$$\sum_{u,v\in V}A(u,v)\mathbb{E}_z\big[\chi(u)\chi(v)\big] = \sum_{u,v\in V}A(u,v)\frac{2}{\pi}\arcsin\big(f(u)\cdot f(v)\big).$$

The fact that the matrix A satisfies $\sum_{u \in V} A(u,v) = 0$ for every $v \in V$ then gives

$$\sum_{u,v \in V} A(u,v) \frac{2}{\pi} \arcsin\big(f(u) \cdot f(v)\big) = \sum_{u,v \in V} \big(-A(u,v)\big)\Big(1 - \frac{2}{\pi} \arcsin\big(f(u) \cdot f(v)\big)\Big). \quad (1.7)$$

Define

$$\alpha_{\mathrm{GW}} = \min\left\{\frac{\arccos(t)}{1-t} : t \in [-1,1]\right\} = .878\ldots.$$

Using the trigonometric identity $1 - 2\arcsin(t)/\pi = \arccos(t)$ and $A(u,v) \leq 0$ for all $u \neq v$, we can now write and bound the right-hand side of Eq. (1.7) as

$$\sum_{\{u,v\} \in E} \big(-A(u,v)\big) \left(\frac{\arccos\big(f(u) \cdot f(v)\big)}{1 - f(u) \cdot f(v)}\right) \big(1 - f(u) \cdot f(v)\big) \geq \alpha_{\mathrm{GW}} \sum_{\{u,v\} \in E} \big(-A(u,v)\big)\big(1 - f(u) \cdot f(v)\big).$$

Now using $1 - f(u) \cdot f(u) = 0$ and $A(u,v) = 0$ for all $\{u,v\} \notin E$ allows us to sum over all pairs of vertices, making the above sum equal to

$$\sum_{u,v \in V} \big(-A(u,v)\big)\big(1 - f(u) \cdot f(v)\big) = \sum_{u,v \in V} A(u,v) f(u) \cdot f(v),$$

where in the identity we again used that $\sum_{v \in V} A(u,v) = 0$ for all $u \in V$. The last sum above is simply 4 times the optimum of SDP (1.4), which is in turn at least as large as the size of a maximum cut. Collecting the factor $1/4$ left behind in Eq. (1.6) gives that the expected size of a cut returned by Algorithm 1.1 is at least $.878\ldots$ times the the size of a maximum cut.

Optimality of Goemans and Williamson's approximation algorithm. By exhibiting an explicit family of graphs, Karloff [Kar96], and later Feige and Schechtman [FS02], proved that Goemans and Williamson's analysis of their algorithm is in fact optimal, showing that strange-appearing number $.878\ldots$ is an upper bound on the approximation ratio of the algorithm for those graphs. Khot, Kindler, Mossel and O'Donnell [KKMO04] showed that based on the assumption of a complexity-theoretic conjecture known as the Unique Games Conjecture (cf. Section 1.7.4), $.878\ldots$ is in fact the best-possible approximation ratio achievable by *any* polynomial-time approximation algorithm.

1.7.3 The chromatic number and the Lovász theta number

The *chromatic number* of a graph is defined as the smallest number of colors needed to color its vertices such that no two adjacent vertices receive the same color. A coloring of the vertices that assigns different colors to adjacent pairs and uses k colors is said to be a *proper k-coloring* of the graph. Computing the chromatic number is a well-known NP-hard problem.

The theta number refers to the optimum of a celebrated semidefinite program introduced by Lovász [Lov79]. One of its many applications is that it gives a lower bound on the chromatic number of a graph. For this, we consider the *complement* of a graph $G = (V, E)$, denoted $\overline{G}$, which is the graph with vertex set V in which a pair of distinct vertices are an edge if and only if they are not an edge in G. The *theta number* of the complement of a graph $G = (V, E)$, denoted by $\vartheta(\overline{G})$, is the optimum of the following semidefinite program:

$$\begin{aligned} \text{minimize} \quad & \lambda \\ \text{subject to} \quad & Z \in \mathcal{S}_V^+ \\ & Z(u,u) = \lambda - 1 \text{ for every } u \in V \\ & Z(u,v) = -1 \text{ for every } \{u,v\} \in E, \end{aligned}$$

where $\mathcal{S}_V^+$ denotes the set of real positive semidefinite matrices whose rows and columns are indexed by the vertices of G.

The fact that the value $\vartheta(\overline{G})$ provides a lower bound for the chromatic number of G can be seen as follows. Suppose that G has a proper k-coloring. We associate with each vertex $v \in V$ a vector $f(v) \in \mathbb{R}^{\binom{k}{2}}$ whose coordinates are indexed by all unordered pairs $\{i,j\} \in \binom{\{1,\ldots,k\}}{2}$.[3] If the coloring assigns color i to v then we define $f(v)$ by

$$f(v)_{\{i,j\}} = \begin{cases} 1 & \text{if } j > i \\ -1 & \text{if } j < i \end{cases}$$

and setting all other entries to zero. The matrix $Z(u,v) = f(u) \cdot f(v)$ is feasible for the above SDP and has objective value k. It follows that $\vartheta(\overline{G}) \leq \chi(G)$.

Notice that there are only k different vectors in the set $\big(f(v)\big)_{v \in V}$. So, although the vectors $f(v)$ have dimension $\binom{k}{2}$, they only span a k-dimensional space. Geometrically, the vectors $f(v)$ define 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.

[3] Throughout we denote by $\binom{S}{t}$ the family of all t-element subsets of a finite set S.

1.7.4 A little on the Unique Games Conjecture

In 2002 Khot [Kho02] introduced the Unique Games Conjecture (UGC) in order to make progress on the problem of obtaining hardness of approximation results for NP-complete problems. Before that, Håstad [Hås99] made significant advances in this area. However, for many problems exact approximation results remained unknown. Since its introduction, it has been shown that the UGC would imply many inapproximability results unknown to be obtainable otherwise [KN08, KN09]. Often such results are highly accurate, matching the approximation ratios of known algorithms. Examples of problems where exact UGC hardness results are known are MAX CUT [KKMO04], minimum vertex cover [KR08], kernel clustering [KN10], max-kCSP [ST09]. Perhaps the most striking result is due to Raghavendra [Rag08], who showed that truth of the UGC implies that there is a single generic SDP-based polynomial-time approximation algorithm for all constraint satisfaction problems that achieves the optimal approximation ratio.

One of several equivalent formulations of the UGC [Kho10] is as follows. For positive integer n, an instance of a *linear unique game over* $\mathbb{Z}_n$ is a two-player nonlocal game given by a positive integer N and numbers $c_{ij} \in \mathbb{Z}_n$ for $i, j \in \{1, \ldots, N\}$. At the start of the game a referee uniformly samples a pair i, j from the set $\{1, \ldots, N\}$ and sends question "i" to Alice and question "j" to Bob. The players answer $a_i, b_j \in \mathbb{Z}_n$, respectively, and win if $a_i - b_j = c_{ij} \pmod{n}$.

1.7.2. CONJECTURE (UNIQUE GAMES CONJECTURE). *For any $0 < \varepsilon < 1$, there exists positive integer $n = n(\varepsilon)$ such that given a linear unique game over $\mathbb{Z}_n$ with maximum classical winning probability $1 - \varepsilon$, there is no polynomial-time algorithm that finds a classical strategy whose winning probability is greater than ε.*

Recently, Arora, Barak and Steurer [ABS10] gave a sub-exponential-time algorithm with performance guarantee better than is allowed in the conjecture for any polynomial-time algorithm. Though this does not disprove the conjecture, it does show that it is on somewhat shaky ground.

In the context of nonlocal games it is natural to ask what happens to the UGC when we allow for entangled strategies. Kempe, Regev and Toner [KRT08] examined exactly this situation and showed that showed that in this case, conjecture is false.

Chapter 2

Grothendieck inequalities

2.1 Introduction

Grothendieck's Inequality is a unifying theme for the chapters in this thesis. Many of the mathematical tools we use to deal with the problems addressed here are variations or extensions of this celebrated inequality. The inequality arose for the first time in Grothendieck's 1953 paper *Résumé de la théorie métrique des produits tensoriels topologiques* [Gro53], nowadays often referred to simply as the *Résumé*. The influence this paper has had until now is difficult to overstate. In particular its main result, Grothendieck's Inequality, has had important applications in huge number of different areas in pure mathematics, theoretical computer science and theoretical physics. A few important examples of such applications are as follows. Tsirelson [Tsi87] showed that the inequality can be interpreted as comparing the classical and quantum biases of a two-player XOR game, which becomes clear after one puts together Tsirelson's Theorem (see Section 1.5) and the form of Grothendieck's Inequality given below. We discuss this application in further detail in Chapters 3 and 6. Alon and Naor [AN06] realized that the inequality gives an upper bound on the ratio of the optima of certain integer optimization problems and their semidefinite relaxations. They showed that this implies the existence of constant-factor approximation algorithms for the problem of computing the cut-norm of a matrix. Their paper kindled a large amount of research on to connections between Grothendieck's Inequality and approximation algorithms based on semidefinite programming. These results are discussed in greater detail in Chapters 4 and 5. Linial and Shraibman [LS09] and Lee, Shraibman and

Schechtman [LSS09] showed that the inequality has connections to communication complexity and Regev and Toner [RT09] adapted techniques used in a proof of the inequality to simulate quantum correlations with classical communication. Pérez-García [PG06] applied the inequality in the context of Banach algebras, a result we discuss in more detail in Chapters 7. Not surprisingly, many equivalent forms of the inequality have been discovered since its first appearance. Arguably its most elementary shape was found by Lindenstrauss and Pełczyński [LP68], which is the shape in which we present the inequality below. In this chapter we present most of the variations of Grothendieck's Inequality that appear in subsequent chapters, though for convenience definitions will often be repeated when they are needed. Many more references regarding applications of Grothendieck's Inequality can be found throughout this thesis. We also refer to the recent extensive surveys of Pisier [Pis11] and Khot and Naor [KN11] for more information on this inequality, variations of it and applications to combinatorial optimization.

2.2 Grothendieck's Inequality

To suppress the space needed to state Grothendieck's Inequality and some of the modifications of it that feature in this thesis we introduce the following notation.

2.2.1. DEFINITION. For positive integers n, r and real n-by-n matrix A, define

$$\mathrm{SDP}_r(A) = \max\left\{ \sum_{i,j=1}^{n} A_{ij} x_i \cdot y_j : x_1, \ldots, x_n, y_1, \ldots, y_n \in S^{r-1} \right\}. \tag{2.1}$$

Define $\mathrm{SDP}_\infty(A)$ analogously, with a maximum over the unit sphere of $\ell_2(\mathbb{R})$.

With regard to the above definition, let us note that since any collection of vectors $x_1, \ldots, x_n, y_1, \ldots, y_n$ span a space of dimension at most $2n$, we have $\mathrm{SDP}_\infty(A) = \mathrm{SDP}_{2n}(A)$ for every n-by-n matrix A. We also note that the set S^0 consists just of the numbers 1 and -1. The reason for the abbreviation SDP is a connection to semidefinite programs which will become more explicit in the subsequent chapters. Grothendieck's Inequality can now be stated as follows.

2.2.2. THEOREM (GROTHENDIECK [GRO53]). *There exists a real number $K > 0$ such that for every positive integer n and any real n-by-n matrix A, we have*

$$\mathrm{SDP}_\infty(A) \leq K \, \mathrm{SDP}_1(A). \tag{2.2}$$

Inequality (2.2) is nowadays known as *Grothendieck's Inequality*. Associated to Grothendieck's Inequality is the smallest number K for which it holds.

2.2.3. DEFINITION. The *Grothendieck constant* K_G is defined by

$$K_G = \sup \left\{ \frac{\mathrm{SDP}_\infty(A)}{\mathrm{SDP}_1(A)} : n \in \mathbb{N},\, A \in \mathbb{R}^{n \times n} \right\}.$$

Despite many efforts the exact value of K_G is currently not known. However, it is known to be bounded by

$$1.676\ldots \lesssim K_G < \frac{\pi}{2\ln(1+\sqrt{2})} = 1.782\ldots$$

The lower bound is due to Davie [Dav84] and Reeds [Ree91], who independently found the same result. The upper bound of $\pi/\big(2\ln(1+\sqrt{2})\big)$ due to Krivine [Kri79] was the best known for over thirty years and by many believed to be tight. However, an exciting development took place while this thesis was being written. Using an extension of Krivine's techniques, Braverman, Makarychev, Makarychev and Naor [BMMN11] proved that his upper bound can be improved, disproving Krivine's conjecture that his bound gave the exact value of K_G. Although they do not give a numerical bound, they prove that $K_G \leq \pi/\big(2\ln(1+\sqrt{2})\big) - \varepsilon$ holds for some constant $\varepsilon > 0$. A proof of Krivine's upper bound is given in Chapter 5 as part of a more general result regarding a generalization of K_G based on graphs, which is described below.

Despite the fact that the exact value of K_G is unknown, Raghavendra and Steurer [RS09] were able to show that K_G is the UGC hardness threshold for computing the value $\mathrm{SDP}_1(A)$ for any real matrix A. Moreover, they show that the exact value of K_G can be approximated to within an error ε in time $O\big(\exp(\exp(1/\varepsilon^3))\big)$ by a linear program.

2.3 Generalizations of Grothendieck's Inequality

In this section we define various generalizations of Grothendieck's Inequality that will appear in the subsequent chapters.

2.3.1 The rank-r Grothendieck constant

The first generalization we consider relates $\mathrm{SDP}_r(A)$ for values of r that may differ from ∞ and 1. This generalization appeared for the first time in the

paper [BBT11], whose content will be presented in Chapter 3. We first define the following generalization of the Grothendieck constant.

2.3.1. DEFINITION. For every pair $q, r \in \mathbb{N} \cup \{\infty\}$ such that $q \geq r$, define $K_G(q \mapsto r)$ by

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

We refer to $K_G(\infty \mapsto r)$ as the *rank-r Grothendieck constant*. The reason for the word *rank* is that a matrix $X \in \mathbb{R}^{n \times n}$ has rank r if and only if there exist r-dimensional vectors $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ such that $X_{ij} = x_i \cdot y_j$. It follows that $\mathrm{SDP}_r(A)$ is the maximum of $\langle A, X \rangle$ over rank-r matrices X. Based on Definition 2.3.1 we get the following generalization of Grothendieck's Inequality: For every positive integer n and any real n-by-n matrix A, we have

$$\mathrm{SDP}_q(A) \leq K_G(q \mapsto r)\, \mathrm{SDP}_r(A). \tag{2.3}$$

The constant $K_G(q \mapsto 1)$ is known as *the Grothendieck constant of order q* and is usually denoted $K_G(q)$. It was studied before by Krivine [Kri77], who proved that $K_G(2) = \sqrt{2}$ and $K_G(4) \leq \pi/2$, and numerically computed upper bounds for other values of q, including $K_G(3) < 1.57$.

Variations of $K_G(q \mapsto r)$ that will appear in Chapter 4 are based on positive semidefinite matrices and Laplacian matrices.

2.3.2. DEFINITION. For every pair $q, r \in \mathbb{N} \cup \{\infty\}$ such that $q \geq r$, define $K_G^{\succeq}(q \mapsto r)$ by

$$K_G^{\succeq}(q \mapsto r) = \sup \left\{ \frac{\mathrm{SDP}_q(A)}{\mathrm{SDP}_r(A)} : n \in \mathbb{N},\, A \in \mathcal{S}_n^+ \right\}.$$

2.3.3. DEFINITION. For every pair $q, r \in \mathbb{N} \cup \{\infty\}$ such that $q \geq r$, define $K_G^L(q \mapsto r)$ by

$$K_G^L(q \mapsto r) = \sup \left\{ \frac{\mathrm{SDP}_q(A)}{\mathrm{SDP}_r(A)} : n \in \mathbb{N},\, A \in \mathcal{S}_n^+ \text{ and Laplacian} \right\}.$$

We have the following easy relations between the above constants:

$$K_G^L(q \mapsto r) \leq K_G^{\succeq}(q \mapsto r) \leq K_G(q \mapsto r) \leq K_G.$$

The calculations done in Section 1.7.2 to analyze the Goemans-Williamson approximation algorithm for MAX-CUT show that the constant $K_G^L(\infty \mapsto 1)$ is bounded from above by $(.878\ldots)^{-1} = 1.138\ldots$. Upper bounds on the constants $K_G^L(q \mapsto 1)$ for $q = 2, 3$ were computed by Avidor and Zwick [AZ05], who showed that for these values of q, we have $K_G^L(q \mapsto 1) < K_G^L(\infty \mapsto 1)$.

2.3.2 The Grothendieck constant of a graph

We also consider a generalization of Grothendieck's Inequality based on graphs. For this, we introduce a variation of the quantity $\mathrm{SDP}_r(A)$ based on graphs and matrices whose rows and columns are indexed by the vertices of those graphs.

2.3.4. DEFINITION. For a graph $G = (V, E)$, positive integer r and matrix $A : V \times V \to \mathbb{R}$, define

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

Define $\mathrm{SDP}_\infty(G, A)$ analogously with a maximum over functions $f : V \to S^\infty$ where S^∞ is the unit sphere of $\ell_2(\mathbb{R})$.

Since $\ell_2(\mathbb{R})$ contains $\mathbb{R}^n$ as the subspace of its first n components, we have that $|V|$-dimensional unit vectors suffice to achieve the maximum above (note that a collection of $|V|$ vectors span a space of dimension at most $|V|$). That is, $\mathrm{SDP}_\infty(G, A) = \mathrm{SDP}_{|V|}(G, A)$. An important difference between $\mathrm{SDP}_r(G, A)$ and $\mathrm{SDP}_r(A)$ defined above is that the latter has a maximum over *two* sequences of unit vectors, while the the former has only *one* such sequence.

2.3.5. DEFINITION. For a graph $G = (V, E)$ and pair $q, r \in \mathbb{N} \cup \{\infty\}$ such that $q \geq r$, define $K(q \mapsto r, G)$ by

$$K(q \mapsto r, G) = \sup \left\{ \frac{\mathrm{SDP}_q(G, A)}{\mathrm{SDP}_r(G, A)} : A : V \times V \to \mathbb{R} \right\}.$$

The *rank-r Grothendieck constant of the graph G*, denoted $K(r, G)$, is defined by $K(r, G) = K(\infty \mapsto r, G)$. This number plays a major role in Chapter 5, where we establish new upper bounds for $r > 1$. The constant $K(G) = K(1, G)$ was considered by by Alon, Makarychev, Makarychev and Naor [AMMN06], who called it simply *the Grothendieck constant of the graph G*. They proved that

$$\Omega\big(\log \omega(G)\big) \leq K(G) \leq O\big(\log \vartheta(\overline{G})\big),$$

where $\omega(G)$ is the size of the largest clique in G. This shows in particular that $K(G)$ depends strongly on the graph, and is not a universal constant like the Grothendieck constant K_G. Laurent and Varvitsiotis [LV11] showed that for specific graph classes, it is possible to compute $K(G)$ exactly. In particular, if G is an n-cycle, then $K(G) = n \cos(\pi/n)/(n-2)$ and if G has no K_5 minor and is not a forest, then $K(G)$ equals the maximum value of $K(C)$ for C a cycle graph appearing as an induced subgraph in G. We refer to Chapter 5 for more results on these numbers.

2.3.3 The complex Grothendieck constant

Perhaps the most natural generalization of Grothendieck's Inequality is obtained by allowing all quantities involved to take complex values. Let us denote by $S_{\mathbb{C}}^{r-1}$ the r-dimensional complex unit sphere.

2.3.6. DEFINITION. For positive integers n, r and complex n-by-n matrix A, define

$$\mathrm{SDP}_r^{\mathbb{C}}(A) = \max\left\{ \Big| \sum_{i,j=1}^{n} A_{ij}\langle x_i, y_j\rangle \Big| : x_1, \ldots, x_n, y_1, \ldots, y_n \in S_{\mathbb{C}}^{r-1} \right\}, \qquad (2.4)$$

where $\langle \cdot, \cdot \rangle$ denotes the regular inner product on $\mathbb{C}^d$. Define $\mathrm{SDP}_\infty^{\mathbb{C}}$ analogously but with a supremum over the unit sphere of the Hilbert space $\ell_2(\mathbb{C})$ of complex square-summable sequences.

2.3.7. DEFINITION. For every pair $q, r \in \mathbb{N} \cup \{\infty\}$ such that $q \geq r$, define

$$K_G^{\mathbb{C}}(q \mapsto r) = \sup\left\{ \frac{\mathrm{SDP}_q^{\mathbb{C}}(A)}{\mathrm{SDP}_r^{\mathbb{C}}(A)} : n \in \mathbb{N},\ A \in \mathbb{C}^{n\times n} \right\}. \qquad (2.5)$$

The *complex Grothendieck constant* $K_G^{\mathbb{C}}$ is defined by $K_G^{\mathbb{C}} = K_G^{\mathbb{C}}(\infty \mapsto 1)$. The corresponding variant of Grothendieck's Inequality is then as follows: For any positive integer n and any matrix $A \in \mathbb{C}^{n\times n}$, we have

$$\mathrm{SDP}_\infty^{\mathbb{C}}(A) \leq K_G^{\mathbb{C}} \mathrm{SDP}_1(A).$$

The currently best lower and upper bounds on $K_G^{\mathbb{C}}$, proved by Davie [Dav84] and Haagerup [Haa87], respectively, are given by $1.33807 \leq K_G^{\mathbb{C}} \lesssim 1.40491$. König [Kön91] obtained the numerical bounds $1.152 \leq K_G^{\mathbb{C}}(2 \mapsto 1) \leq 1.216$. Another related result of Davie [Dav85] shows that for every positive integer n, any complex n-by-n matrix A and r the integer part of $\sqrt{2n-1}$, we have $\mathrm{SDP}_\infty^{\mathbb{C}}(A) = \mathrm{SDP}_r^{\mathbb{C}}(A)$; he noted that similar results hold for the real setting.

2.3.4 Tonge's Inequality

Blei [Ble79] and Tonge [Ton78] considered certain multilinear generalizations of Grothendieck's inequality, where the matrix A is replaced by a higher-order tensor, and the inner product function replaced by a multilinear functional on more than two unit vectors. We use these generalizations in Chapters 6 and 7.

By a (real) *N-tensor* we mean a map $A : [n]^N \to \mathbb{R}$, which can be seen as an array of reals whose coordinates are indexed by N-tuples of integers $(i_1, \ldots, i_N)$ over the set $[n]$. The case $N = 2$ thus gives ordinary real matrices.

We introduce two quantities, reminiscent of the quantities $\mathrm{SDP}_1(A)$ and $\mathrm{SDP}_\infty(A)$ appearing on opposing sides of Grothendieck's Inequality.

2.3.8. DEFINITION. For positive integers n, N and N-tensor $A : [n]^N \to \mathbb{R}$, define

$$\mathrm{OPT}(A) = \max \left\{ \sum_{I \in [n]^N} A[I]\chi_1(i_1) \cdots \chi(i_N) : \right.$$
$$\left. \chi_1, \ldots, \chi_N : [n] \to \{-1, 1\} \right\}. \quad (2.6)$$

It may be helpful to note that for $N = 2$, we have $\mathrm{OPT}(A) = \mathrm{SDP}_1(A)$.

We introduce a multilinear functional, which replaces the regular inner product appearing in Grothendieck's Inequality. The *generalized inner product* of vectors $x_1, \ldots, x_n \in \mathbb{C}^d$ is defined by

$$\langle x_1, \ldots, x_N \rangle = \sum_i (x_1)_i \cdots (x_N)_i,$$

where $(x_1)_i$ denotes the ith coordinate of the vector x_1 in the canonical basis. Note that for the case $N = 2$, $\langle \cdot, \cdot \rangle$ is linear in both arguments, as opposed to conjugate linear in the first and linear in the second. This conflicting notation with the standard inner product will not be an issue later on and will only occur in Chapters 6 and 7 where the cases $N \geq 3$ are of main interest. Let $B_{\mathbb{C}^d}$ denote the d-dimensional complex unit ball.

2.3.9. DEFINITION. For positive integers n, N and N-tensor $A : [n]^N \to \mathbb{R}$, we define

$$\mathrm{GIP}(A) = \sup \left\{ \left| \sum_{I \in [n]^N} A[i_1, \ldots, i_N] \langle f_1(i_1), \ldots, f_N(i_N) \rangle \right| : \right.$$
$$\left. d \in \mathbb{N},\ f_1, \ldots, f_N : [n] \to B_{\mathbb{C}^d} \right\}. \quad (2.7)$$

Note carefully that in the definition of $\mathrm{GIP}(A)$, the supremum is taken over *complex* vectors, while the tensor A is assumed to be *real*. For $N = 2$, the identity $\mathrm{GIP}(A) = \mathrm{SDP}^{\mathbb{C}}_{\infty}(A)$ holds, as there is a 1-1 correspondence between complex unit vectors and their conjugates.

The multilinear generalization of Grothendieck's Inequality given below is a slight variation of a result due to Tonge [Ton78].[1]

2.3.10. THEOREM (TONGE). *Let $n, N \geq 2$ be positive integers. Then, for any N-tensor $A : [n]^N \to \mathbb{R}$, we have*

$$\mathrm{GIP}(A) \leq 2^{(3N-5)/2} K_G^{\mathbb{C}} \, \mathrm{OPT}(A). \tag{2.8}$$

In the original inequality proved by Tonge the tensor A may be complex and the maximization on the right-hand side is over variables in the complex unit disc. The version stated above is tailored specifically to our needs.

The proof of Theorem 2.3.10 that we give here is longer than Tonge's original proof, but more elementary. Both proofs use induction on N. The base case, $N = 2$, is derived from the complex version of Grothendieck's Inequality, which we restate here in its strongest form.

2.3.11. LEMMA (HAAGERUP). *For all positive integers n, d, any* complex *n-by-n matrix A and* complex *vectors $x_1, \ldots, x_n$ and $y_1, \ldots, y_n \in B_{\mathbb{C}^d}$, the inequality*

$$\Big| \sum_{i,j=1}^{n} A_{ij} \langle x_i, y_j \rangle \Big| \leq K_G^{\mathbb{C}} \max \left\{ \Big| \sum_{i,j=1}^{n} A_{ij} \sigma_1(i) \sigma_2(j) \Big| : \sigma_1, \sigma_2 : [n] \to B_{\mathbb{C}} \right\}, \tag{2.9}$$

where $K_G^{\mathbb{C}} \lesssim 1.40491$ is independent of n and d.

Note that the maximization on the right-hand side of Eq. (2.9) is over sequences $\sigma_1(1), \ldots, \sigma_1(n)$ and $\sigma_2(1), \ldots, \sigma_2(n)$ of scalars in the *complex* unit disc.

The inductive step relies on a slight modification of an Inequality of Littlewood (Lemma 2.3.12 below) [Lit30] (see also [Pie72, page 43] and [Sza76]).

2.3.12. LEMMA (LITTLEWOOD). *For all positive integers n, d and any complex n-by-d matrix B, we have*

$$\sum_{i=1}^{n} \Big(\sum_{j=1}^{d} |B_{ij}|^2 \Big)^{1/2} \leq 2^{3/2} \max \left\{ \Big| \sum_{i=1}^{n} \sum_{j=1}^{d} B_{ij} \chi(i) \xi(j) \Big| : \chi(i), \xi(j) \in \{-1, 1\} \right\}. \tag{2.10}$$

[1] A weaker version of Tonge's result was proved earlier by Blei [Ble79] (though it was published shortly after Tonge's paper was).

We first prove Theorem 2.3.10. Afterwards, we prove Lemma 2.3.12

PROOF OF THEOREM 2.3.10: By induction on N. For the base case, $N = 2$, we use Lemma 2.3.11 and relate the right-hand side of Inequality (2.9) to $\mathrm{OPT}(A)$ using the following version of [MST99, Proposition 15]:

3. CLAIM. *For any real n-by-n matrix A and sequences of scalars $\sigma_1(1), \ldots, \sigma_1(n)$ and $\sigma_2(1), \ldots, \sigma_2(n)$ in the* complex *unit disc $B_{\mathbb{C}}$, we have*

$$\Big| \sum_{i,j=1}^{n} A_{ij}\sigma_1(i)\sigma_2(j) \Big| \leq \max \left\{ \sum_{i,j=1}^{n} A_{ij} \Re\Big(\sigma_1'(i)\sigma_2'(j)\Big) : \sigma_1', \sigma_2' : [n] \to B_{\mathbb{C}} \right\}. \tag{2.11}$$

PROOF: Using polar coordinates, the complex number $\sum_{i,j=1}^{n} A_{ij}\sigma_1(i)\sigma_2(j)$ can be written as $re^{i\phi}$ for some non-negative real number r and angle ϕ. This gives

$$\begin{aligned} \Big| \sum_{i,j=1}^{n} A_{ij}\sigma_1(i)\sigma_2(j) \Big| &= \Big| e^{-i\phi} \sum_{i,j=1}^{n} A_{ij}\sigma_1(i)\sigma_2(j) \Big| \\ &= \Re\left(e^{-i\phi} \sum_{i,j=1}^{n} A_{ij}\sigma_1(i)\sigma_2(j) \right) \\ &= \sum_{i,j=1}^{n} A_{ij} \Re\Big(e^{-i\phi}\sigma_1(i)\sigma_2(j) \Big), \end{aligned}$$

where the second identity follows because the number between brackets is real and nonnegative (it is r), and the third identity follows because A is real. The result follows by defining $\sigma_1' = e^{-i\phi/2}\sigma_1$ and $\sigma_2' = e^{-i\phi/2}\sigma_2$. □

We can write the real part $\Re(\sigma_1\sigma_2)$ of the product of two complex numbers σ_1, σ_2 as the inner product between real vectors $a = \big(\Re(\sigma_1), \Im(\sigma_1)\big)^T$ and $b = \big(\Re(\sigma_2), -\Im(\sigma_2)\big)^T$. Using this, Lemma 2.3.11 and Claim 3, we get that for every sequence of unit vectors $x_1, \ldots, x_n$ and $y_1, \ldots, y_n \in B_{\mathbb{C}^d}$,

$$\begin{aligned} \Big| \sum_{i,j=1}^{n} A_{ij}\langle x_i, y_j\rangle \Big| &\leq K_G^{\mathbb{C}} \max_{\sigma_1,\sigma_2:[n]\to B_(\mathbb{C})} \Big| \sum_{i,j=1}^{n} A_{ij}\sigma_1(i)\sigma_2(j) \Big| \\ &\leq K_G^{\mathbb{C}} \max_{\sigma_1,\sigma_2:[n]\to B_(\mathbb{C})} \sum_{i,j=1}^{n} A_{ij} \Re\Big(\sigma_1(i)\sigma_2(j)\Big) \\ &\leq K_G^{\mathbb{C}} \max_{a,b:[n]\to B_{\mathbb{R}^2}} \sum_{i,j=1}^{n} A_{ij}\, a(i)\cdot b(j) \end{aligned}$$

The base case now follows from Krivine's [Kri79] bound $K_G(2) \leq \sqrt{2}$ on the Grothendieck constant of order 2 (see Section 2.3.1). This implies that

$$\sum_{i,j=1}^{n} A_{ij} a(i) \cdot b(j) \leq \sqrt{2} \max \left\{ \sum_{i,j=1}^{n} A_{ij} \chi_1(i) \chi_2(j) : \chi_1, \chi_2 : [n] \to \{-1, 1\} \right\}$$

holds for any sequence of vectors $a(1), \ldots, a(2)$ and $b(1), \ldots, b(2) \in B_{\mathbb{R}^2}$, and hence proves the base case.

We continue with the induction step. Suppose that Inequality (2.8) holds for some $N \geq 2$. Let $A : [n]^{N+1} \to \mathbb{R}$ be a real $(N+1)$-tensor. Define the complex n-by-d matrix B by

$$B_{ij} = \sum_{i_1, \ldots, i_N = 1}^{n} A[i_1, \ldots, i_N, i] f_1(i_1)_j \cdots f_N(i_N)_j,$$

where $f_1(i_1)_j$ stands for the j^{th} coordinate of the d-dimensional complex vector $f_1(i_1)$. Then, we can write the left-hand side of Inequality (2.8) as

$$\sum_{i_1, \ldots, i_N, i = 1}^{n} A[i_1, \ldots, i_N, i] \langle f_1(i_1), \ldots, f_{N+1}(i) \rangle = \sum_{i=1}^{n} \sum_{j=1}^{d} B_{ij} f_{N+1}(i)_j.$$

By the triangle inequality, the Cauchy-Schwarz inequality and Inequality (2.10) from Lemma 2.3.12, we can bound the absolute value of this quantity by

$$\begin{aligned} \Big| \sum_{i=1}^{n} \sum_{j=1}^{d} B_{ij} f_{N+1}(i)_j \Big| &\leq \sum_{i=1}^{n} \Big| \sum_{j=1}^{d} B_{ij} f_{N+1}(i)_j \Big| \\ &\leq \sum_{i=1}^{n} \Big(\sum_{j=1}^{d} |B_{ij}|^2 \Big)^{1/2} \\ &\leq 2^{3/2} \max \left\{ \Big| \sum_{i=1}^{n} \sum_{j=1}^{d} B_{ij} \chi(i) \xi(j) \Big| \right\}, \end{aligned} \tag{2.12}$$

where the maximum is taken over $\chi : [n] \to \{-1, 1\}$ and $\xi : [d] \to \{-1, 1\}$. Let χ and ξ be the functions with which this maximum is achieved.

Expanding the definition of the matrix B gives

$$\begin{aligned} &2^{3/2} \Big| \sum_{i=1}^{n} \sum_{j=1}^{d} B_{ij} \chi(i) \xi(j) \Big| \\ &\quad = 2^{3/2} \Big| \sum_{i=1}^{n} \sum_{j=1}^{d} \Big(\sum_{i_1, \ldots, i_N = 1}^{n} A[i_1, \ldots, i_N, i] f_1(i_1)_j \cdots f_N(i_N)_j \Big) \chi(i) \xi(j) \Big| \end{aligned} \tag{2.13}$$

Define the real N-tensor $A' : [n]^N \to \mathbb{R}$ defined by

$$A'[i_1, \ldots, i_N] = \sum_{i=1}^{n} A[i_1, \ldots, i_N, i]\chi(i).$$

Then, we can write the right-hand side of Eq. (2.13) as

$$2^{3/2}\Big|\sum_{j=1}^{d}\sum_{I\in[n]^N} A'[I]f_1(i_1)_j \cdots f_N(i_N)_j\xi(j)\Big| = 2^{3/2}\Big|\sum_{I\in[n]^N} A'[I]\langle f_1(i_1)_j, \cdots, f_N(i_N)\circ\xi\rangle\Big|,$$

where by $f_N(i_N)\circ\xi$ we mean the entry-wise product of the d-dimensional complex vectors $f_N(i_N)$ and $\big(\xi(1),\ldots,\xi(d)\big)$. By the induction hypothesis, the last quantity is bounded from above by $2^{3/2}2^{(3N-5)/2}K_G^{\mathbb{C}}\,\mathrm{OPT}(A')$. Since $\mathrm{OPT}(A)$ involves a re-maximization over χ that appears in the definition of A', we have $\mathrm{OPT}(A') \le \mathrm{OPT}(A)$. This completes the proof. □

We now prove Inequality (2.10) of Lemma 2.3.12. We derive it from Khintchine's Inequality (see for example [MS86]), which states that there exists a constant κ such that for any any finite sequence of complex scalars $\sigma(1),\ldots,\sigma(n)$ the inequality

$$\left(\sum_{i=1}^{n}|\sigma(i)|^2\right)^{1/2} \le \kappa\int_{t=0}^{1}\Big|\sum_{i=1}^{n}\sigma(i)r_i(t)\Big|dt, \tag{2.14}$$

where $r_i(t) = \mathrm{sign}\big(\sin(2^i\pi t)\big)$ denotes the i^{th} Rademacher function. The best value of κ is due to Szarek [Sza76] (see also [LO94]), who proved that $\kappa \le \sqrt{2}$.

PROOF OF LEMMA 2.3.12: By Inequality (2.14), we have

$$\begin{aligned}\sum_{i=1}^{n}\Big(\sum_{j=1}^{d}|B_{ij}|^2\Big)^{1/2} &\le \sqrt{2}\int_{t=0}^{1}\sum_{i=1}^{n}\Big|\sum_{j=1}^{d}B_{ij}r_j(t)\Big|dt \\ &\le \sqrt{2}\max\left\{\sum_{i=1}^{n}\Big|\sum_{j=1}^{d}B_{ij}\xi(j)\Big|\right\}, \end{aligned}\tag{2.15}$$

where the above maximum is over maps $\xi : [d] \to \{-1, 1\}$. Let ξ achieve this maximum. Define $\chi : [n] \to B_{\mathbb{C}}$ by

$$\chi(i) = \frac{\Big(\sum_{j=1}^{d}B_{ij}\xi(j)\Big)^*}{\big|\sum_{j=1}^{d}B_{ij}\xi(j)\big|}$$

Then, we have that the maximum on the right-hand side of Eq. (2.15) equals

$$\sum_{i=1}^{n} \chi(i) \sum_{j=1}^{d} B_{ij} \xi(j) = \sum_{i=1}^{n} \sum_{j=1}^{d} B_{ij} \chi(i) \xi(j).$$

Now, we make χ real-valued, at the cost of a factor at most 2. By the triangle inequality, we have

$$\sum_{i=1}^{n} \chi(i) \sum_{j=1}^{d} B_{ij} \xi(j) \leq \Big| \sum_{i=1}^{n} \Re\big(\chi(i)\big) \sum_{j=1}^{d} B_{ij} \xi(j) \Big| + \Big| \sum_{i=1}^{n} \Im\big(\chi(i)\big) \sum_{j=1}^{d} B_{ij} \xi(j) \Big| \quad (2.16)$$

Set χ' to be either the real or imaginary part of χ, whichever gives the largest value on the right-hand side of Eq. (2.16). Then, $\chi' : [n] \to [-1, 1]$ and we have

$$\begin{aligned} \sum_{i=1}^{n} \Big(\sum_{j=1}^{d} |B_{ij}|^2 \Big)^{1/2} &\leq 2^{3/2} \Big| \sum_{i=1}^{n} \sum_{j=1}^{d} B_{ij} \chi'(i) \xi(j) \Big| \\ &\leq 2^{3/2} \max \left\{ \Big| \sum_{i=1}^{n} \sum_{j=1}^{d} B_{ij} \chi''(i) \xi(j) \Big| \right\}, \end{aligned}$$

where the maximum is taken over $\chi'' : [n] \to \{-1, 1\}$ and $\xi : [d] \to \{-1, 1\}$. Here, the second inequality follows because χ' can be written as a convex combination of functions $\chi'' : [n] \to \{-1, 1\}$ and by the triangle inequality. This completes the proof. □

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$^+$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 $x_1, \ldots, x_n, y_1, \ldots, y_n$, the inequality

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

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 $K = \pi/2 = 1.5707\ldots$.

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 π on pairs $\{1, \ldots, n\} \times \{1, \ldots, n\}$ and n-by-n sign matrix Σ (this can be done by simply dividing each of the elements of A by $\sum_{i,j=1}^{n} |A_{ij}|$). Then, the pair (π, Σ) defines a two-player nonlocal game as follows. A referee samples a pair (i,j) according to π and asks Alice question "i", and Bob question "j". Alice and Bob answer with signs χ_i and ψ_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" (π, Σ) 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+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

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

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

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

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

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

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 $A : \mathbb{R}^n \to \mathbb{R}^n$ defines a real n-by-n matrix given by $(e_i \cdot Ae_j)_{i,j=1}^n$, where $e_1, \ldots, e_n$ 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 $f : S^{n-1} \to \mathbb{R}$ 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 $A : C(S^{n-1}) \to C(S^{n-1})$. 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 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 $\mathrm{SDP}_r(A)$ for the case where A is a linear operator of the form $A : C(S^{n-1}) \to C(S^{n-1})$. The goal of this is to establish that $K_G(q \mapsto r) \geq \mathrm{SDP}_q(A)/\,\mathrm{SDP}_r(A)$ for any such operator A. Let us recall that the definition of $\mathrm{SDP}_r(A)$ when A is an n-by-n *matrix* is

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

The argument of this maximum can be rewritten as

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

The first thing we do towards extending the definition of 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 $A : C(S^{n-1}) \to C(S^{n-1})$ and continuous *vector-valued* function $f : S^{n-1} \to \mathbb{R}^r$, the expression Af should be interpreted as follows. We can view the function f as a collection of r real-valued functions $f_1, \ldots, f_r \in C(S^{n-1})$, such that $f(x) = \big(f_1(x), \ldots, f_r(x)\big)^T$. By Af we mean that A acts on each of these r functions simultaneously, giving another continuous vector-valued function $(Af) : S^{n-1} \to \mathbb{R}^r$ defined by $(Af_1, \ldots, Af_r)^T$. 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 $g : S^{n-1} \to S^{r-1}$. 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 $O(\mathbb{R}^n) = \{U \in \mathbb{R}^{n\times n} : U^TU = I\}$ denote the orthogonal group on $\mathbb{R}^n$. A measure ν on S^{n-1} (which we endow with the Borel σ-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 ν on a measurable space Ω is a *probability measure* if it is normalized so that $\nu(\Omega) = 1$. Let ω_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 SDP_r. For linear operator $A : C(S^{n-1}) \to C(S^{n-1})$ and integer $r \geq 2$, define

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

where the supremum is taken over all functions f, g that are continuous and measurable. We define $\mathrm{SDP}_1(A)$ 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 $\mathrm{SDP}_\infty(A)$ 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 $A : C(S^{n-1}) \to C(S^{n-1})$ and any $\eta > 0$, there exists positive integer $N = N(\eta)$ and*

real N-by-N matrix B such that,

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

We defer the proof of this lemma, which uses a standard ε-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 $\gamma : \mathbb{R} \to \mathbb{R}$ 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 $\Gamma : \mathbb{R} \to \mathbb{R}$ is the Gamma function, defined by

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

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

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

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

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

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

PROOF: We prove the lemma in two parts. First, we show that the problem of computing $\mathrm{SDP}_r(A)$ 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*

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

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

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

where the supremum is over functions $f, g : S^{n-1} \to S^{r-1}$ 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

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

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:

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

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

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

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

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

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

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 $\|F\|_{\mathrm{HS}} = 1$. 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 $\lambda_1 \geq \cdots \geq \lambda_r \geq 0$ satisfying $\|F\|_{\mathrm{HS}}^2 = \lambda_1^2 + \cdots + \lambda_r^2 = 1$.

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 ω_n are invariant under orthogonal transformations, the singular value decomposition of F gives

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

where

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

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

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

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

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

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

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*

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

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

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

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

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

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)]

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

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

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

Multiplying this by the left-over factor from above then gives result. ♦

Combining the two claims gives

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

which proves the lemma. □

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

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

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]

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

This proves the theorem. □

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 $r = 1, 2, \ldots$.*

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 $z \geq 2$:

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

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

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

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

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 (π, Σ) where the players are restricted to sharing a state with local dimension at most d, and the value $\mathrm{SDP}_r(\pi \circ \Sigma)$ 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 $x_1, \ldots, x_n$, $y_1, \ldots, y_n$, 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 $F_1, \ldots, F_n$, $G_1, \ldots, G_n \in \mathcal{O}(\mathbb{C}^d)$, such that for every $i, j \in \{1, \ldots, n\}$, 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 $F_1, \ldots, F_n$, $G_1, \ldots, G_n \in \mathcal{O}(\mathbb{C}^d)$, there exist a positive integer r that depends on d only and real r-dimensional unit vectors $x_1, \ldots, x_n$, $y_1, \ldots, y_n$ such that for every $i, j \in \{1, \ldots, n\}$, 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

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

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 $\pi : \{1, \ldots, n\} \times \{1, \ldots, n\} \to [0, 1]$ is a probability distribution and Σ is an n-by-n sign matrix. Note that the pair (π, Σ) 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 (π, Σ) by players who share an entangled state with local dimension $d = \lfloor \sqrt{r/2} \rfloor$, is at most $\mathrm{SDP}_r(A)$. 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 $x_1, \ldots, x_n, y_1, \ldots, y_n$ such that

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

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

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

Hence, we have

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

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

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 $n \to \infty$ 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 $A : C(S^{n-1}) \to C(S^{n-1})$ 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 $\mathrm{SDP}_q(A)/\mathrm{SDP}_r(A)$ 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 ε in time $O\big(\exp(\exp(1/\varepsilon^3))\big)$ 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 $B : C(S^{q-1}) \to C(S^{q-1})$ such that*

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

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 μ_q be the rotationally invariant probability measure on $O(\mathbb{R}^q)$.

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

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

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

To get some intuition for this operator, observe that the value $(T_\rho\chi)(x)$ is the average of χ 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 $\mathrm{SDP}_q(B)$ from below, consider the action of T_ρ on the linear function χ given by $\chi(x) = x_1$. We have

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

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

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

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

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

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

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

PROOF: Set $\rho = u \cdot v$. As the measure μ_q is rotationally invariant, it suffices to consider $u = (1, 0, \ldots, 0)^T$ and $v = (\rho, \sqrt{1-\rho^2}, 0, \ldots, 0)^T$. Let us denote an orthogonal matrix U as $U = [x, y, z_1, \ldots, z_{q-2}]$, where $x, y, z_1 \ldots, z_{q-2} \in S^{q-1}$ are its columns. Then, for random U distributed according to μ_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

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

which in turn equals the left-hand side by the definition of T_ρ. ♦

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

$$\sum_{i,j=1}^{n} A_{ij} \int_{S^{q-1}} f'(x) \cdot (T_{u_i \cdot v_j} g')(x) d\omega_q(x) =$$

$$\sum_{i,j=1}^{n} A_{ij} \int_{O(\mathbb{R}^q)} f'(Uu_i) \cdot g'(Uv_j) d\mu_q(U) =$$

$$\int_{O(\mathbb{R}^q)} \left(\sum_{i,j=1}^{m} A_{ij} f'(Uu_i) \cdot g'(Uv_j) \right) d\mu_q(U) \leq \mathrm{SDP}_r(A),$$

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 $\mathrm{SDP}_q(B)$ given in Eq. (3.18). □

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 $1.6769\ldots$. 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* $A_\rho : C(S^{n-1}) \to C(S^{n-1})$ 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 ρ 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 $\mathrm{SDP}_\infty(A)/\mathrm{SDP}_r(A)$ 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 SDP_∞ and 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 ∞ 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$^+$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 $A : C(S^{n-1}) \to C(S^{n-1})$ and any $\eta > 0$, there exists positive integer $N = N(\eta)$ and real N-by-N matrix B such that,*

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

For the proof of the lemma we use an *ε-net* for S^{n-1}, which is a finite set of n-dimensional unit vectors $\mathcal{Z}_\varepsilon = \{z_1, \ldots, z_N\}$ 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 ε-net $\mathcal{Z}_\varepsilon = \{z_1, \ldots, z_N\} \subseteq S^{n-1}$ of size*

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

PROOF OF LEMMA 3.8.1: Define for continuous function $f : S^{n-1} \to \mathbb{R}^q$ 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 $f : S^{n-1} \to \mathbb{R}^q$, we have $\|Af\|_\infty / \|f\|_\infty \leq 1$.

We define the finite operator $B : \mathbb{R}^N \to \mathbb{R}^N$ to be a discretized version of A as follows. Let $\mathcal{Z}_\varepsilon = \{z_1, \ldots, z_N\}$ be an ε-net for the n-dimensional unit sphere, for some ε to be chosen later. Let for each $i \in \{1, \ldots, N\}$ the region $R_i \subseteq S^{n-1}$ be the set of vectors for which point z_i of Z_ε is closest in Euclidean distance (with ties distributed arbitrarily) and let $I_{R_i} : S^{n-1} \to \{0, 1\}$ 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 $\mathrm{SDP}_q(B)$ is not much *smaller* than $\mathrm{SDP}_q(A)$. (Since these quantities appear in the numerator of Eq. (3.20), there is no problem if $\mathrm{SDP}_q(B)$ is larger than $\mathrm{SDP}_q(A)$.) To this end, let $f, g : S^{n-1} \to S^{q-1}$ be optimal for $\mathrm{SDP}_q(A)$. Trivially,

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

Define the continuous function $h : S^{q-1} \to S^{q-1}$ 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 $\mathrm{SDP}_q(A) - \mathrm{SDP}_q(B)$ is bounded from above by

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

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

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

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

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

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

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

$\mathrm{SDP}_q(B)$ is smaller than $\mathrm{SDP}_q(A)$.) To this end, let $x_1, \ldots, x_N, y_1, \ldots, y_N \in S^{r-1}$ be optimal for $\mathrm{SDP}_r(B)$. 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 $\mathrm{SDP}_r(A)$ give

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

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 $B : \mathbb{R}^N \to \mathbb{R}^N$ such that,

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

from which the claim follows by taking the ratios of the two inequalities. □

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.

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

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 $x_1, \ldots, x_n$ 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* $\mathrm{SDP}_r(A)$*, defined as in Definition 2.1, by*

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

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

It follows from Proposition 4.1.1 and Definition 2.3.2 that $K_G^{\succeq}(\infty \mapsto r)$ 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 $K_G^{\succeq}(\infty \mapsto 1) \leq \pi/2$, 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 $K_G^{\succeq}(\infty \mapsto 1) \geq \pi/2$, 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 $K_G^L(\infty \mapsto 1) = (.878\ldots)^{-1}$. Avidor and Zwick [AZ05] proved that $K_G^L(q \mapsto 1) < (.878\ldots)^{-1}$ 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 $y_1, \ldots, y_n \in S^{r-1}$ for Problem 4.1.

(1) Solve the semidefinite relaxation of Problem 4.1 for the matrix A, obtaining vectors $x_1, \ldots, x_n \in S^{n-1}$.

(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 $y_1, \ldots, y_n \in S^{r-1}$ by $y_i = Zx_i / \|Zx_i\|_2$ for $i = 1, \ldots, n$.

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*

$$1 \leq K_G^{\succeq}(\infty \mapsto r) \leq \frac{1}{\gamma(r)} = 1 + \Theta\Big(\frac{1}{r}\Big),$$

where

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

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 $\mathrm{SDP}_\infty(A) = \mathrm{SDP}_n(A)$. Upper bounds on $K_G^{\succeq}(n \mapsto 1)$ therefore imply upper bounds on the ratio $\mathrm{SDP}_\infty(A) / \mathrm{SDP}_1(A)$ whenever A is positive semidefinite and of size n-by-n.

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

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

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 $K_G^{\succeq}(\infty \mapsto r)$ and $K_G^{\succeq}(n \mapsto 1)$. 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 ε-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 $K_G^{\succeq}(\infty \mapsto r)$ for $r = 1, \ldots, 10$. For $r = 1$, the lower bound is due to Grothendieck [Gro53] and the upper bound due to Nesterov [Nes98] and Rietz [Rie74].

r	$K_G^{\succeq}(\infty \mapsto r)$
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 $x_1, \ldots, x_n$. 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 $x_1, \ldots, x_n \in S^{r-1}$ 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 π on $\{1, \ldots, n\} \times \{1, \ldots, n\}$ and n-by-n sign matrix Σ. 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 $\mathrm{SDP}_r(\mathrm{A})$ 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$^+$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 ℓ_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 *ℓ_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 $S_1, \ldots, S_k$ of the set $\{1, \ldots, n\}$ 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 $B = \left[\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right]$ 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 $x_1, \ldots, x_n \in S^{n-1}$ of the semidefinite relaxation of Problem 4.1 into a feasible solution in the form of vector-valued random variables $y_1, \ldots, y_n \in S^{r-1}$. 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

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

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

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

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 $Uu = (1, 0, \ldots, 0)^T$ and $Uv = (t, \sqrt{1-t^2}, 0 \ldots, 0)^T$

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 $E_r : [-1, 1] \to [-1, 1]$ 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 $x_1, \ldots, x_n \in S^{n-1}$ are optimal for $\mathrm{SDP}_\infty(A)$, we can write the right-hand side of Eq. (4.1) as

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

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

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 $u_1, \ldots, u_k$, the matrix*

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

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

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

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 $x_1, \ldots, x_n$ are optimal for $\mathrm{SDP}_\infty(A)$. Lemma 4.2.1 thus enables us to prove that $\mathrm{SDP}_\infty(A)/\,\mathrm{SDP}_r(A) \leq 1/c$. As $\mathrm{SDP}_\infty(A) \geq \mathrm{SDP}_r(A)$, 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 $f : [-1,1] \to [-1,1]$ is *of positive type for S^∞*, if for all positive integers n, k and any real n-dimensional unit vectors $u_1, \ldots, u_k$, the matrix

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

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 $f : [-1,1] \to \mathbb{R}$ is of positive type for S^∞ if and only if it is of the form*

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

where $c_0, c_1, \cdots \geq 0$ 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*

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

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^∞. As the linear term in this expansion is $2t/\pi$, it follows that the function $\big(E_1(t) - 2t/\pi\big)$ is of positive type for S^∞. For the case $r = 1$, the number c from Lemma 4.2.1 giving the upper bound $\mathrm{SDP}_\infty(A)/\,\mathrm{SDP}_1(A) \leq 1/c$ 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^∞. To see this, note that for every positive integer k and any choice of unit vectors $u_1, \ldots, u_k$, we have that the matrix

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

is a convex combination of positive semidefinite matrices, since

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

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^∞.

Now, by Schoenberg's Theorem (Theorem 4.2.3) there exist $c_0, c_1, \cdots \geq 0$ 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^∞ 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 $\theta \in [0, 2\pi]$, $u = (\cos\theta, \sin\theta, 0, \ldots, 0)^T$ and $v = (\cos\theta, -\sin\theta, 0, \ldots, 0)^T$ 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

$$\frac{Zu}{\|Zu\|_2} \cdot \frac{Zv}{\|Zv\|_2} = \frac{u^T Z^T Z v}{\sqrt{(u^T Z Z^T u)(v^T Z Z v)}},$$

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

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

where Γ_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

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

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 $t^2 - \mathrm{Tr}(W)t + \det(W)$ of a generic element $W \in \mathcal{S}_2^+$. We then have

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

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

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

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

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

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 $\mathrm{SDP}_r(A)$ 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.[1] 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

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

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

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

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

[1] A nice proof of this theorem can be found in [OF09].

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

PROOF OF THEOREM 4.1.3: Let $x_1, \ldots, x_n \in S^{n-1}$ be optimal for $\mathrm{SDP}_\infty(A)$. By Grothendieck's Identity, Algorithm 4.1 gives $\{-1, 1\}$-valued random variables $\chi_1, \ldots, \chi_n$ 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^∞, 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 $c_0, \ldots, c_k \geq 0$ and $\sum_{k=0}^{\infty} c_0$ 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

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

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 $c_1 = c(n) = \big(\arcsin, P_1^n\big)_n / (P_1^n, P_1^n)_n$. The numerator of $c(n)$ equals

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

The denominator of $c(n)$ equals

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

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. □

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 ρ 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 $x_1, \ldots, x_n \in S^{r-1}$ be an approximate solution coming from such a polynomial-time algorithm. Then,

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

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

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

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 $\mathrm{SDP}_r(A) \geq \mathrm{SDP}_1(A)$, 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 $\rho \leq \gamma(r)$. □

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:

$$A_{ij} \leq 0, \qquad \text{if } i \neq j, \tag{4.7}$$

$$\sum_{i=1}^{n} A_{ij} = 0, \qquad \text{for every } j = 1, \ldots, n. \tag{4.8}$$

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*

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

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 $y_1, \ldots, y_n$ such that

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

where $x_1, \ldots, x_n \in S^{n-1}$ are optimal vectors for $\mathrm{SDP}_\infty(A)$. 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

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

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. □

The first ten numerical values of the above upper bounds are given in Table 4.2. The numerical values suggest that as $r \to \infty$, 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 $r = 1, \ldots, 10$. 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 ration $\gamma(r) = 1 - \Theta(1/r)$.

2. This approximation ratio can be improved to $2/\big(\pi\gamma(n)\big)$ 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 $\mathrm{SDP}_r(A)$ and $\mathrm{SDP}_\infty(A)$. 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).

Chapter 5

The graphical Grothendieck problem with rank constraint

The content of this chapter is based on joint work with Fernando Mário de Oliveira Filho and Frank Vallentin [BOFV10b]

5.1 Introduction

In this chapter, we study computational aspects of another optimization problem, the *graphical Grothendieck problem with rank-r constraint*. This problem is based on a graph $G = (V, E)$ with finite vertex set V, edge set $E \in \binom{V}{2}$ and a symmetric matrix A whose rows and columns are indexed by V. Let us recall that $\mathcal{S}_V^+$ denotes the cone of positive semidefinite matrices whose rows and columns are indexed by V. The problem is defined as follows.

Problem 5.1 (The graphical Grothendieck problem with rank-r constraint). Takes as input a graph $G = (V, E)$, positive integer r and symmetric matrix $A : V \times V \to \mathbb{R}$.

$$\begin{aligned} \text{maximize} \quad & \textstyle\sum_{\{u,v\} \in E} A(u,v) X(u,v) \\ \text{subject to} \quad & X \in \mathcal{S}_V^+ \\ & X(u,u) = 1 \; \forall u \in V \\ & \operatorname{rank}(X) = r \end{aligned}$$

Like the variant considered in Chapter 4 this problem is almost a semidefinite program. But due to the rank constraint it may not be efficiently solvable

or approximable to within arbitrary precision. The case $r = 1$ has MAX CUT (see Section 1.7.2) as a special case, and is therefore NP-hard. To obtain the MAX CUT problem we set the graph G to be the complete bipartite graph $K_{n,n}$ on $2n$ vertices. Take a Laplacian matrix B of some graph on n vertices and set $A = \left[\begin{smallmatrix} 0 & B \\ B^T & 0 \end{smallmatrix}\right]$. An optimal solution for Problem 5.1 then gives a cut of maximal size for the n-vertex graph. The problem reduces to the *positive semidefinite Grothendieck problem* treated in Chapter 4 when we replace the above matrix B by an arbitrary positive semidefinite matrix. If we remove the rank constraint (we will denote this by $r = \infty$) then the problem does become a semidefinite program, which can be solved efficiently regardless of the graph G.

We can interpret the problem geometrically using the 1-1 correspondence between rank-r positive semidefinite matrices $X : V \times V \to \mathbb{R}$ and matrices of the form $\big(f(u) \cdot f(v)\big)_{u,v \in V}$ where each $f(u)$ is an r-dimensional unit vector. The problem thus asks to position $|V|$ vectors on a real r-dimensional unit sphere in such a way that a weighted sum of their inner products is maximized. It follows from this that the optimum of Problem 5.1 is given by $\mathrm{SDP}_r(G, A)$ (see Definition 2.3.4) and that the largest possible ratio $\mathrm{SDP}_q(G, A) / \mathrm{SDP}_r(G, A)$ for matrices $A : V \times V \to \mathbb{R}$ is given by $K(q \mapsto r, G)$ (see Definition 2.3.5). In particular, the rank-r Grothendieck constant of the graph G, defined by $K(r, G) = K(\infty \mapsto r, G)$, gives the largest possible ratio of the optimum of the natural semidefinite relaxation of Problem 5.1, and its actual optimum.

The case $r = 1$ of Problem 5.1 was studied extensively by the computer science community. The case of bipartite graphs was studied by Alon and Naor [AN06] in the context of computing the cut norm of matrices and finding Szemerédi partitions of graphs. Based on the fact that $K(1, K_{n,n}) \leq K_G$, they gave a polynomial-time $(1/K_G)$-approximation algorithm for computing the cut-norm, thereby kindling a large mass of research related to connections between optimization, semidefinite programming and Grothendieck-like inequalities. Let K_n denote the complete graph on n vertices. For $G = K_n$ Problem 5.1 is known as the *quadratic programming problem with* $\{-1, 1\}$*-constraint*. Independently, Nemirovski, Roos and Terlaky [NRT99], Megretski [Meg01] and Charikar and Wirth [CW04] proved that $K(1, K_n) \leq O(\log n)$. Khot and O'Donnell [KO08] proved that $K(1, K_n) \geq \Omega(\log n)$, showing that in contrast to K_G, its graphical versions are not in general constants (see also [AMMN06, ABH$^+$05]). Hardness-of-approximation results for the quadratic programming problem were obtained by Arora et al. [ABH$^+$05]. We refer to Section 2.3.2 for more details on these numbers.

Much less seems to be known about the computational aspects of the more geometric cases of the graphical Grothendieck problem with rank-r constraint, where $r \geq 2$. The main results of this chapter are new upper bounds on the numbers $K(r, G)$ for arbitrary ranks r and graphs G with small chromatic number. These upper bounds are obtained by analyzing an efficient $K(r, G)^{-1}$-approximation algorithm for Problem 5.1, given in Section 5.1.2. At the end of this chapter we derive new upper bounds on $K(r, G)$ for graphs with large chromatic number, by analyzing a straight-forward modification of an efficient approximation algorithm due to Alon, Makarychev, Makarychev, and Naor [AMMN06] (see Section 5.6). Before giving details of the main results we discuss two applications.

5.1.1 Applications

We give two interpretations of Problem 5.1, one in ground state energies and one in XOR games. Similar to the problem considered in Chapter 4, the objective function of Problem 5.1 can be interpreted as a kind of energy. Stanley's *n-vector model* [Sta68] describes the interaction of particles in a spin glass with ferromagnetic and antiferromagnetic interactions. Let $G = (V, E)$ be the interaction graph where the vertices represent particles and where edges indicate which particles interact. The potential function $A\colon V \times V \to \mathbb{R}$ is 0 if u and v are not adjacent, positive if there is ferromagnetic interaction between u and v, and negative if there is antiferromagnetic interaction. The particles possess a vector-valued spin $f\colon V \to S^{n-1}$. The case $n = 1$ corresponds to the Ising model, the case $n = 2$ to the XY (or classical planar) model, the case $n = 3$ to the Heisenberg model, and the case $n = \infty$ to the Berlin-Kac spherical model. In the absence of an external field, the total energy of the system is given by the *Hamiltonian*

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

The *ground state* of this model is a configuration of spins $f\colon V \to S^{n-1}$ which minimizes the total energy. Finding the ground state is the same as solving $\mathrm{SDP}_n(G, A)$. The much-studied Ising model (the case $n = 1$) is a simplification of the spin glass model in which the vectors are two- or three-dimensional (i.e., the XY model and the Heisenberg model) [Sta68, BGJR88, KNS10]. Typically, the interaction graph has small chromatic number. The most common case is when this graph 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 refer to Talagrand's book [Tal03] and the paper Bansal, Bravyi and Terhal [BBT08] for more extensive introductions and mathematical/computational treatments of spin glasses.

The case of bipartite graphs in Problem 5.1 is also of interest to us because it is related to the setting of two-player nonlocal games. Let L and R be disjoint finite sets and let $\mathcal{G} = (\pi, \Sigma)$ be a two-player XOR game given by probability distribution π on $L \times R$ and sign matrix $\Sigma : L \times R \to \{-1, 1\}$. The set L contains Alice's questions and the set R Bob's. Let $G = (V, E)$ be the complete bipartite graph on vertex set $V = L \cup R$ where all edges are between the sets L and R. Define the matrix $A : L \cap R \to \mathbb{R}$ by setting $A(u, v) = \pi(u, v)\Sigma(u, v)$ if $\{u, v\} \in E$ and $A(u, v) = 0$ otherwise. The optimum of Problem 5.1 is of the form

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

for some functions $f : V \to S^{r-1}$. Since our graph G is bipartite, we can split the collection of vectors $f(u)$ into two groups corresponding to whether $u \in L$ or $u \in R$. By renaming the vectors $f(u)$ for every $u \in R$ to, say, $g(u)$ we get that the sum above equals

$$\sum_{u \in L} \sum_{v \in R} A(u, v) f(u) \cdot g(v) = \mathbb{E}_{(u,v) \sim \pi} \big[\Sigma(u, v) f(u) \cdot g(v) \big].$$

By Tsirelson's Theorem (see Section 1.5) we thus have that the optimum above is a *lower bound* on the entangled bias of $\mathcal{G}$ when the players have quantum systems of local dimension $2^{\lceil r/2 \rceil}$, and an *upper bound* on the bias when the local dimensions are $\sqrt{r/2}$.

5.1.2 An efficient approximation algorithm for graphs with small chromatic number

In this chapter we prove explicit upper bounds for $K(r, G)$. For the most part, we will focus on the case of small r and graphs with small chromatic number, although our methods for such cases 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 $0.78\ldots = (1.28\ldots)^{-1}$. This result can be regarded as the principal contribution of this chapter.

5.1.1. THEOREM. *For $r = 1, \ldots, 10$ 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...

Our bound for the original Grothendieck constant K_G, which corresponds to the case where $r = 1$ and G is a complete bipartite graph $K_{n,n}$ of any size n, is due to Krivine [Kri79]. Our bound for $K(2, K_{n,n})$ coincides with Haagerup's [Haa87] upper bound on $K_G^{\mathbb{C}}$. Though these numbers may be different, it should not be a surprise that the bounds are equal, since we use some of Haagerup's techniques. When the graph G has large chromatic number, then the result of [AMMN06] gives the best known bounds for $K(1, G)$ (see Section 2.3.2). They prove a logarithmic dependence on the chromatic number of the graph whereas the first row in the table has a linear dependence on the chromatic number. We extend the results of [AMMN06] for large chromatic numbers for $r \geq 2$ in Section 5.6.

For the proof of Theorem 5.1.1 we use the framework developed by Krivine and Haagerup for the case of bipartite graphs, explained below. The main new technical tool used in the proof is a matrix version of Grothendieck's Identity given in Lemma 5.2.1. To develop some intuition for the proof we begin by considering the natural strategy for proving upper bounds on $K(r, G)$. Based on the Goemans and Williamson approximation algorithm for MAX CUT and Algorithm 4.1 for Problem 4.1, the natural strategy is to embed a collection of $|V|$-dimensional vectors $\big(f(u)\big)_{u \in V}$ for which the value $\mathrm{SDP}_\infty(G, A)$ is achieved[1] into S^{r-1} using a random projection based on an r-by-$|V|$ matrix Z with i.i.d.

[1] Recall that $|V|$-dimensional vectors always suffice since $|V|$ vectors span a space of dimension at most $|V|$.

Gaussian entries. This amounts to defining the random r-dimensional unit vectors $g(u) = Zf(u)/\|Zf(u)\|_2$. By linearity of expectation, the expected objective value of this solution for the rank-r case of Problem 5.1 is given by

$$\mathbb{E}\left[\sum_{\{u,v\}\in E} A(u,v)g(u)\cdot g(v)\right] = \sum_{\{u,v\}\in E} A(u,v)\mathbb{E}\big[g(u)\cdot g(v)\big].$$

The problem now is lower bound this quantity in terms of the optimum of the problem $\mathrm{SDP}_r(G,A)$. If $r=1$, Grothendieck's Identity gives $\mathbb{E}\big[g(u)\cdot g(v)\big] = (2/\pi)\arcsin\big(f(u)\cdot f(v)\big)$. For larger values of r this expectation is also some nonlinear function of $f(u)\cdot f(v)$. The strategy of Krivine and Haagerup is to "linearize" these functions by using the following new embedding lemma.

5.1.2. LEMMA. *Let $G=(V,E)$ be a graph and choose $Z=(Z_{ij})\in\mathbb{R}^{r\times|V|}$ at random so that the entries are i.i.d. $N(0,1)$ random variables. Given $f\colon V\to S^{|V|-1}$, there is a function $g\colon V\to S^{|V|-1}$ such that whenever u and v are adjacent in G, then*

$$\mathbb{E}\left[\frac{Zg(u)}{\|Zg(u)\|_2}\cdot\frac{Zg(v)}{\|Zg(v)\|_2}\right] = \beta(r,G)f(u)\cdot f(v)$$

for some constant $\beta(r,G)$ depending only on r and G. Moreover, the function g can be found in polynomial time in $|V|$.

In the statement above we are vague regarding the constant $\beta(r,G)$. We will give the precise statement of the lemma in Section 5.4 (Lemma 5.4.1 there), right now this precise statement is not relevant to our discussion. Now, the strategy of Krivine and Haagerup amounts to analyzing a following four-step procedure that yields a randomized polynomial-time approximation algorithm for Problem 5.1, Algorithm 5.1 shown below.

To analyze this algorithm, we compute the expected value of the feasible solution h. By linearity of expectation, we get

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

as before. But now, using Lemma 5.1.2, we get that the above sum equals

$$\beta(r,G)\sum_{\{u,v\}\in E} A(u,v)f(u)\cdot f(v) \;=\; \beta(r,G)\,\mathrm{SDP}_\infty(G,A), \tag{5.1}$$

Algorithm 5.1 Takes as input graph $G = (V, E)$, positive integer r and symmetric matrix $A : V \times V \to \mathbb{R}$, and returns a feasible solution $h : V \to S^{r-1}$ for $\mathrm{SDP}_r(G, A)$.

(1) Solve the semidefinite relaxation of Problem 5.1, obtaining $f : V \to S^{n-1}$.

(2) Use f to construct $g : V \to S^{|V|-1}$ according to Lemma 5.1.2.

(3) Sample matrix $Z \in \mathbb{R}^{r \times |V|}$ such that the entries Z_{ij} are independently distributed Gaussian random variables with mean 0 and variance 1.

(4) Define $h : V \to S^{r-1}$ by $h(u) = Zg(u)/\|Zg(u)\|_2$ for every $u \in V$.

and hence $K(r, G) \leq \beta(r, G)^{-1}$. Since $\mathrm{SDP}_\infty(G, A) \geq \mathrm{SDP}_r(G, A)$ it also follows that Algorithm 5.1 is a $\beta(r, G)$-approximation algorithm for Problem 5.1.

The constant $\beta(r, G)$ in Lemma 5.1.2 is defined in terms of the Taylor expansion of the inverse of the function $E_r \colon [-1, 1] \to [-1, 1]$ given by

$$E_r(x \cdot y) = \mathbb{E}\left[\frac{Zx}{\|Zx\|_2} \cdot \frac{Zy}{\|Zy\|_2}\right],$$

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. In Section 4.2.1 of Chapter 4 we argued that the function E_r is indeed well-defined, which follows because the expectation above is invariant under orthogonal transformations.

Outline of the rest of this chapter. The Taylor expansion of E_r is computed in Section 5.2. The Taylor expansion of E_r^{-1} is treated in Section 5.3, where we basically follow Haagerup [Haa87]. A precise version of Lemma 5.1.2 is stated and proved in Section 5.4, following Krivine [Kri79]. In Section 5.5 we show that one can refine this analysis and can (strictly) improve the upper bounds on $K(r, G)$ if one takes the size of the vertex set into account. In particular, there we prove upper bounds on $K(q \mapsto r, G)$. In Section 5.6 we show how to generalize the technique of [AMMN06] to deal with graphs with large chromatic numbers and higher values of r and we briefly summarize this chapter in Section 5.7.

5.2 A matrix version of Grothendieck's Identity

In this section we prove a generalization of Grothendieck's Identity. This gives the Taylor coefficients of the function E_r, which we need to prove Lemma 5.1.2.

5.2.1. LEMMA. *For positive integers r, n, let u, v be real n-dimensional unit vectors and let Z be a random real r-by-n matrix with independent $N(0,1)$ entries. Then,*

$$\mathbb{E}\left[\frac{Zu}{\|Zu\|_2}\cdot\frac{Zv}{\|Zv\|_2}\right] = \gamma(r)\,(u\cdot v)\,{}_2F_1\left(\begin{matrix}1/2,\,1/2\\ r/2+1\end{matrix}\;;\,(u\cdot v)^2\right),$$

where

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

and

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

is a hypergeometric function (see for example [AAR99]).

Before proving this lemma, we note a couple of special cases. For the case $r = 1$, we obtain Grothendieck's Identity (Lemma 4.2.4):

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

The case $r = 2$ gives a function used by Haagerup [Haa87] to upper bound $K_G^{\mathbb{C}}$:

$$\begin{aligned}\mathbb{E}\left[\frac{Zu}{\|Zu\|_2}\cdot\frac{Zv}{\|Zv\|_2}\right] &= \frac{1}{u\cdot v}\Big(E(u\cdot v) - (1-(u\cdot v)^2)K(u\cdot v)\Big)\\ &= \frac{\pi}{4}\left(u\cdot v + \left(\frac{1}{2}\right)^2\frac{(u\cdot v)^3}{2} + \left(\frac{1\cdot 3}{2\cdot 4}\right)^2\frac{(u\cdot v)^5}{3} + \cdots\right),\end{aligned}$$

where K and E are the complete elliptic integrals of the first and second kind (see for example [AAR99]). Note that on page 201 of Haagerup [Haa87] $\pi/2$

should be $\pi/4$. In the previous chapter we computed the first coefficient of the Taylor series of the expectation for every r, which turned out to be $\gamma(r)$.

Unfortunately, for $r \geq 2$ we don't have a nice geometric proof as we do for the case $r = 1$. The proof we give here 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 [Kön01]. It would be interesting to find a more geometrical proof of the lemma.[2]

PROOF OF LEMMA 5.2.1: Let $Z_i \in \mathbb{R}^n$ be the i-th row of the matrix Z, with $i = 1, \dots r$. We define vectors

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

so that we have $x \cdot y = (Zu) \cdot (Zv)$. Since the probability distribution of the vectors Z_i is invariant under orthogonal transformations we may assume that $u = (1, 0, \dots, 0)^T$ and $v = (t, \sqrt{1-t^2}, 0, \dots, 0)^T$ and so the pair $(x, y) \in \mathbb{R}^r \times \mathbb{R}^r$ is distributed according to the probability density function (see for example [Mui82, Theorem 1.2.9])

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

Hence,

$$\mathbb{E}\left[\frac{x}{\|x\|_2} \cdot \frac{y}{\|y\|_2}\right] = (2\pi\sqrt{1-t^2})^{-r} \int_{\mathbb{R}^r}\int_{\mathbb{R}^r} \frac{x}{\|x\|_2} \cdot \frac{y}{\|y\|_2} \exp\left(-\frac{x \cdot x - 2tx \cdot y + y \cdot y}{2(1-t^2)}\right) dxdy.$$

By using spherical coordinates $x = \alpha\xi$, $y = \beta\eta$, where $\alpha, \beta \in [0, \infty)$ and $\xi, \eta \in S^{r-1}$, and the rotationally invariant (surface area) measure $\tilde{\omega}_r$ on the r-dimensional unit sphere, normalized such that $\tilde{\omega}_r(S^{r-1}) = \pi^{r/2}/\Gamma\big(r/2\big)$, we rewrite the above integral as

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

[2] Oded Regev gave a more intuitive proof based on well-known probabilistic estimates, but we won't give the details of his proof here.

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

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

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

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

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 (see for example [AAR99, 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 [AAR99, Eq. (4.12.2)]

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

Putting things together, we get

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

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

Now we evaluate the outer double integral

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

Here the inner integral equals

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

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

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

which, by the Bateman Manuscript Project [EMOT54, p. 337 Eq. (11)], equals

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

Now we treat the remaining outer integral in a similar way, using [EMOT54, p. 219 Eq. (17)], and get that

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

By applying Euler's transformation (see for example [AAR99, Eq. (2.2.7)])

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

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

5.3 Convergence radius

To construct the new vectors in the step (2) of Algorithm 5.1 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

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

for some coefficients b_{2k+1}, but in the proof of Lemma 5.1.2 it will be essential that this expansion be valid for all $t \in [-1, 1]$.

In the case $r = 1$ we have $E_1^{-1}(t) = \sin(\pi t/2)$ 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 [Haa87, p. 208] one can express b_k for any choice of $\alpha > 1$ 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'_α is the quarter circle $\{ \alpha e^{i\theta} : \theta \in [0, \pi/2] \}$.

For an appropriate choice of α 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.

5.3.1. LEMMA. *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 (see for example [AAR99, Theorem 2.2.1]) we can rewrite E_r as

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

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. □

5.4 Constructing new vectors

In this section, we prove Lemma 5.1.2, of which Lemma 5.4.1 below is the detailed version. Roughly speaking, we define the function $g : V \to S^{|V|-1}$ such that the inner product $g(u) \cdot g(v)$, for adjacent vertices u and v, inverts the function E_r and leaves a linear function of $f(u) \cdot f(v)$. For this, we use the Taylor expansion of the inverse of E_r and build on a construction of Krivine [Kri79], who proved the lemma for the case of bipartite graphs.

For the nonbipartite case we use the theta number, which is a graph parameter introduced by Lovász [Lov79]. Let $G = (V, E)$ be a graph. The *theta number* of the complement of G, denoted by $\vartheta(\overline{G})$, introduced in Section 1.7.3. We restate it here for convenience. It is the optimal value of the following semidefinite program:

$$\begin{aligned} \vartheta(\overline{G}) = \min \Big\{ \lambda :\ & Z \in \mathcal{S}_V^+, \\ & Z(u,u) = \lambda - 1 \ \text{ for } u \in V, \\ & Z(u,v) = -1 \ \text{ for } \{u,v\} \in E \Big\}. \end{aligned} \tag{5.2}$$

5.4.1. LEMMA. *Let $G = (V, E)$ be a graph with at least one edge. Given $f : V \to S^{|V|-1}$, there exists $g : V \to S^{|V|-1}$ such that for all $\{u, v\} \in E$, we have*

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

where the constant $\beta(r, G)$ is defined by the unique positive solution of the equation

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

where the coefficients b_{2k+1} come from the expansion

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

With this lemma we can now prove Theorem 5.1.1.

PROOF OF THEOREM 5.1.1: We combine Lemma 5.4.1 with the analysis of Algorithm 5.1. To compute the table in the theorem, we use the formula

$$b_k = \frac{1}{k! a_1^k} \left[\frac{d^{k-1}}{dt^{k-1}} \left(1 + \frac{a_2}{a_1} t + \cdots + \frac{a_k}{a_1} t^{k-1} \right)^{-k} \right]_{t=0}, \tag{5.3}$$

where a_i are the Taylor coefficients of E_r (see for example Morse and Feshbach [MF53, (4.5.13)]). □

PROOF OF LEMMA 5.4.1: We construct the vectors $g(u) \in S^{|V|-1}$ by constructing vectors $R(u)$ in an infinite-dimensional Hilbert space $\mathcal{H}$ whose inner product matrix coincides with the one of the $g(u)$. We construct the vectors $R(u)$ from two pairs of vector-valued functions, *inner functions* $S, T : \mathbb{R}^{|V|} \to \mathcal{H}$, and *outer functions* $s, t : V \to \mathbb{R}^{2|V|}$. The inner functions serve to invert the function E_r and the outer functions serve to control the pairwise inner products for adjacent vertices u and v. We proceed in three steps.

In the first step, we construct the inner functions. Set $H = \mathbb{R}^{|V|}$ and

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

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

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

and

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

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 by the definition of $\beta(r, G)$ given in the lemma,

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

Notice that here it is essential that the Taylor expansion of E_r^{-1} has a convergence radius of at least one.

In the second step, we define the outer functions. Let $\lambda = \vartheta(\overline{G})$, and Z be an optimal solution for (5.2). We have $\lambda \geq 2$ since G has at least one edge. Set

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

where J is the all-ones matrix, and consider the matrix

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

By applying a Hadamard transformation

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

we see that U is positive semidefinite, since both $A+B$ and $A-B$ are positive semidefinite. We define the functions $s\colon V \to \mathbb{R}^{2|V|}$ and $t\colon V \to \mathbb{R}^{2|V|}$ so that U is the Gram matrix of the vectors $\big(s(u)\big)_{u\in V}$ and $\big(t(v)\big)_{v\in V}$ with inner products

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

It follows that the functions s and t have the following properties:

1. $s(u)\cdot t(u) = 0$ for every $u \in V$,
2. $s(u)\cdot s(u) = t(u)\cdot t(u) = (\vartheta(\overline{G})-1)/2$ for every $u \in V$,
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 two pairs of functions S, T and s, t to define

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

Then, 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. Finally, we use the Gram decomposition of $(R(u)\cdot R(v)) \in \mathcal{S}_V^+$ to define the function $g : V \to S^{|V|-1}$. □

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

1. The last sentence of the above proof of Lemma 5.4.1 states that there is a positive semidefinite matrix $Y \in \mathcal{S}_V^+$ which satisfies $Y(u,u) = 1$ and

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

for every edge $\{u,v\}$ of G. As this matrix only has to satisfy linear constraints, it can be found in polynomial time in $|V|$ using a semidefinite program. Hence, the function $g : V \to S^{|V|-1}$ of the lemma, defined by the Gram decomposition of Y, can be found in polynomial time.

2. 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, $\beta(1,G) = 2\,\mathrm{arcsinh}(1)/\pi = 2\ln(1+\sqrt{2})/\pi$.

3. In the second step one can also work with any feasible solution of the semidefinite program (5.2). For instance one can replace $\vartheta(\overline{G})$ in the lemma by the chromatic number $\chi(G)$ albeit getting a potentially weaker bound.

4. Alon, Makarychev, Makarychev, and Naor [AMMN06] 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. In particular, we obtain

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

5.5 A refined, dimension-dependent analysis

So far we only bounded $K(\infty \mapsto r, G)$. One can perform a refined, dimension-dependent analysis by bounding $K(q \mapsto r, G)$ when $q \geq r$. This is of interest

because for instance $\mathrm{SDP}_\infty(G, A) = \mathrm{SDP}_{|V|}(G, A)$. In this section we prove an upper bound for $K(q \mapsto r, G)$ that depends on q and r. For fixed r, this upper bound will approach 1 as q approaches r. Krivine [Kri79] gave such a refined analysis for bipartite graphs. We show that our upper bound on $K(q \mapsto r, G)$ is *strictly* smaller than our upper bound for $K(q+1 \mapsto r, G)$.

5.5.1. LEMMA. *Let $G = (V, E)$ be a graph with at least one edge. Given $f\colon V \to S^{q-1}$, there is a function $g\colon V \to S^{|V|-1}$ 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 $q \geq 2$.

Together with the analysis of Algorithm 5.1, this lemma implies the following bounds on $K(q \mapsto r, G)$.

5.5.2. THEOREM. *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 5.5.1 with Algorithm 5.1. □

The proof of the lemma uses a few more basic facts from harmonic analysis, which we now summarize. Let $\overline{P}_k^n$ denote the renormalized version of the Gegenbauer polynomial P_k^n (introduced in Section 4.3) such that $\overline{P}_k^n(1) = 1$. Let us recall the the Gegenbauer polynomials form a completely orthonormal basis for $L^2([-1, 1])$ for the inner product

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

A polynomial in $\mathbb{R}[x_1, \ldots, x_n]$ is *harmonic* if it is homogeneous and vanishes under the Laplace operator $\Delta = \partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_n^2$. When restricted to S^{n-1}, harmonic polynomials are usually referred to as *spherical harmonics*. We endow the space of measurable functions on S^{n-1} with the inner product

$$(f, g) = \int_{S^{n-1}} f(x) g(x) d\omega_n(x).$$

Spherical harmonics are related to Gegenbauer polynomials by the *addition formula* (see for example [AAR99, Theorem 9.6.3]): Let H_k be the space of degree-k

spherical harmonics on n variables. Any orthonormal basis of H_k can be scaled to give a basis $e_{k,1}, \ldots, e_{k,d_k}$ of H_k such that for every $x, y \in S^{n-1}$, we have

$$\overline{P}_k^n(x \cdot y) = \sum_{i=1}^{d_k} e_{k,i}(x) e_{k,i}(y).$$

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 5.4.1.

PROOF OF LEMMA 5.5.1: As before, we construct the function $g : V \to S^{|V|-1}$ from functions S and T that satisfy $S(x) \cdot T(y) = E_r^{-1}(\beta x \cdot y)$ for some real number β. Consider the expansion

$$E_r^{-1}(\beta t) = \sum_{k=0}^{\infty} c_k^q(\beta) \overline{P}_k^q(t).$$

7. CLAIM. *The function $h_q : [0,1] \to \mathbb{R}$ given by*

$$h_q(\beta) = \sum_{k=1}^{\infty} |c_k^q(\beta)|$$

is continuous on its domain.

Before proving the claim, we show how it is used to prove Lemma 5.5.1. Let $\beta(q \mapsto r, G)$ be the largest number $\beta \in [0,1]$ such that $h_q(\beta) = 1$. The fact that such a β exists follows by the Intermediate Value Theorem, from the fact that $h_q(0) = 0$, $h_q(1) \geq E_r^{-1}(1) = 1$ and continuity of h_q.

Consider the Hilbert space

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

where d_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

$$\begin{aligned} S(x)_k &= \sqrt{|c_k^q(\beta(q \mapsto r, G))|}\,(e_{k,1}(x), \ldots, e_{k,d_k}(x)) \\ T(x)_k &= \operatorname{sign}\left(c_k^q(\beta(q \mapsto r, G))\right) \sqrt{|c_k^q(\beta(q \mapsto r, G))|}\,(e_{k,1}(x), \ldots, e_{k,d_k}(x)). \end{aligned}$$

By the addition formula, we have

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

Moreover, from our normalization of the Gegenbauer polynomials $\overline{P}_k^q$ and the addition formula, it follows that we have

$$\|S\big(f(u)\big)\|_2^2 = \|T\big(f(v)\big)\|_2^2 = h_q\big(\beta(q \mapsto r, G)\big) = 1.$$

The desired function $g : V \to S^{|V|-1}$ can be obtained from the $2|V|$-by-$2|V|$ Gram matrix of the vectors $S\big(f(u)\big)$ and $T\big(f(v)\big)$.

Next, we show that for every $q \geq 2$, we have

$$\beta(q \mapsto r, G) > \beta(q+1 \mapsto r, G).$$

We prove this by showing that $h_q\big(\beta(q+1 \mapsto r, G)\big) < 1$, which is sufficient since, by definition, $\beta(q \mapsto r, G)$ is the largest $\beta \in [0,1]$ such that $h_q(\beta) = 1$. Recall that

$$h_q(\beta) \;=\; \sum_{k=0}^{\infty} |c_k^q(\beta)|,$$

where the functions c_k^q came from the expansion

$$E_r^{-1}(\beta t) = \sum_{k=0}^{\infty} c_k^q(\beta) \overline{P}_k^q(t).$$

Using the expansion of $E_r^{-1}(\beta t)$ in terms of the polynomials $\overline{P}_1^{q+1}, \overline{P}_2^{q+1}, \ldots$, we can thus write

$$c_k^q(\beta) = \frac{1}{\|\overline{P}_k^q\|_q^2} \big(E_r^{-1}(\beta t), \overline{P}_k^q\big)_q = \frac{1}{\|\overline{P}_k^q\|_q^2} \sum_{\ell=0}^{\infty} c_\ell^{q+1}(\beta) (\overline{P}_\ell^{q+1}, \overline{P}_k^q)_q.$$

The function E_r^{-1} is not of positive type because the coefficient b_3 of its Taylor expansion is always negative (this can easily be checked using Eq. (5.3)). It follows that some of the $c_k^{q+1}(\beta)$ are negative. Hence,

$$h_q(\beta) = \sum_{k=0}^{\infty} |c_k^q(\beta)| < \frac{1}{\|\overline{P}_k^q\|_q^2} \sum_{k,\ell=0}^{\infty} |c_\ell^{q+1}(\beta)| (\overline{P}_\ell^{q+1}, \overline{P}_k^q)_q = \sum_{\ell=0}^{\infty} |c_\ell^{q+1}(\beta)| \left(\frac{1}{\|\overline{P}_k^q\|_q^2} \sum_{k=1}^{\infty} (\overline{P}_\ell^{q+1}, \overline{P}_k^q)_q \right).$$

It follows from the fact that the polynomials $\overline{P}_1^q, \overline{P}_2^q, \ldots$ form a complete orthogonal basis for $L^2([-1,1])$ with respect to the inner product $(\cdot,\cdot)_q$ and our choice of normalization $\overline{P}_k^{q+1}(1) = \overline{P}_k^q(1) = 1$ that the expression between brackets equals 1. This establishes that for any $\beta \in (0,1)$, we have $h_q(\beta) < h_{q+1}(\beta)$.

What is left to do, is to prove the claim.

PROOF OF CLAIM 7: We begin by showing that for any $\beta \in [0,1]$, the series

$$\sum_{k=0}^{\infty} |c_k^q(\beta)| \tag{5.4}$$

converges. For this, we use the comparison test. Consider the Taylor expansion of the function E_r^{-1}, given by

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

and recall that the series $\sum_{k=0}^{\infty} |b_k|$ converges. Hence, by Schoenberg's Theorem (Theorem 4.2.3), the function

$$\overline{E}_r^{-1}(t) = \sum_{k=0}^{\infty} |b_k| t^k$$

is of positive type for S^∞. In particular, this function is of positive type for S^{q-1}, and can therefore, by Theorem 4.3.1, be expanded in terms of the Gegenbauer polynomials as $\sum_{k=0}^{\infty} \bar{c}_k^q \overline{P}_k^q(t)$, for some $\bar{c}_0^q, \bar{c}_1^q, \cdots \geq 0$ such that $\sum_{k=0}^{\infty} \bar{c}_k^q$ converges.

By orthogonality of the Gegenbauer polynomials with respect to the inner product $(\cdot,\cdot)_q$, we have

$$\bar{c}_k^q = \frac{1}{\|\overline{P}_k^q\|_q^2}(\overline{E}_r^{-1}, \overline{P}_k^q)_q = \frac{1}{\|\overline{P}_k^q\|_q^2} \sum_{\ell=0}^{\infty} |b_\ell| (t^\ell, \overline{P}_k^q)_q \tag{5.5}$$

$$c_k^q(\beta) = \frac{1}{\|\overline{P}_k^q\|_q^2}(E_r^{-1}(\beta t), \overline{P}_k^q)_q = \frac{1}{\|\overline{P}_k^q\|_q^2} \sum_{\ell=0}^{\infty} b_\ell \beta^\ell (t^\ell, \overline{P}_k^q)_q. \tag{5.6}$$

Since for every ℓ, the function $t \mapsto t^\ell$ is of positive type for S^{q-1} (since it is of positive type for every dimension), we have $(t^\ell, \overline{P}_k^q)_q \geq 0$. Comparing Eq.'s (5.5) and (5.6), we see that for every k and any $\beta \in [0,1]$, we have $|c_k^q(\beta)| \leq \bar{c}_k$. The fact that the series (5.4) converges now follows from the fact that $\sum_{k=0}^{\infty} \bar{c}_k$ converges.

The above discussion also establishes that for every every k, the function c_k^q is continuous in β on the interval $[0,1]$, from which it follows that $\beta \mapsto |c_k^q(\beta)|$ is continuous there as well. The fact that the function h_q is continuous now follows because, by convergence of (5.4), it can be approximated arbitrarily well by a finite sum of continuous functions. ♦

This completes the proof. □

5.6 Bounds for graphs with large chromatic number

For graphs with large chromatic number, our bounds on $K(r,G)$ proved above can be improved using the techniques of [AMMN06], which rely on so-called Gaussian Hilbert spaces (see also [JL01, AN06, KNS10]). In this section, we show how their bounds on $K(1,G)$ can be generalized to higher values of r.

5.6.1. THEOREM. *Given graph $G = (V,E)$ and positive integer $1 \leq r \leq \log \vartheta(\overline{G})$, we have*

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

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

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

Fix a matrix $A : V \times V \to \mathbb{R}$. Let $f : V \to S^{|V|-1}$ be optimal for the semidefinite relaxation of Problem 5.1 given the matrix A, so that

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

Let $\widetilde{Z} : V \times V \to \mathbb{R}$ be an optimal solution for the Lovász-theta SDP. Let J be the $2|V|$-by-$2|V|$ all-ones matrix and I the 2-by-2 identity matrix. Since the matrix $(I \otimes \widetilde{Z} + J)/\lambda$ is positive semidefinite, we obtain from its Gram decomposition functions $s, t : V \to \mathbb{R}^{2|V|}$ that satisfy

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

Let $\mathcal{H}$ be the Hilbert space of *vector-valued* functions $h : \mathbb{R}^{r\times|V|} \to \mathbb{R}^r$ such that for a random r-by-$|V|$ matrix Z whose entries are i.i.d. $N(0, 1/r)$ distributed random variables, the inner product on $\mathcal{H}$ is defined by

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

We emphasize that elements of $\mathcal{H}$ map matrices to r-dimensional vectors.

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

$$g_u(Z) = \begin{cases} \frac{Zf(u)}{R} & \text{if } \|Zf(u)\|_2 \leq R \\ \frac{Zf(u)}{\|Zf(u)\|_2} & \text{otherwise,} \end{cases}$$

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

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

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

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

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

Next, we define a function in the space $\mathbb{R}^{2|V|} \otimes \mathcal{H}$ by combining the vectors $s(u), t(v) \in \mathbb{R}^{2|V|}$ and $h_u^0, h_u^1 \in \mathcal{H}$. We endow this space with the natural inner product: For $x \otimes g, y \otimes h \in \mathbb{R}^{2|V|} \otimes \mathcal{H}$, define $\langle x \otimes g, y \otimes h\rangle = (x \cdot y)\,(g, h)$, and extend this inner product linearly so that it is defined for all of $\mathbb{R}^{2|V|} \otimes \mathcal{H}$. For every $u \in V$, define the function $H_u \in \mathbb{R}^{2|V|} \otimes \mathcal{H}$ by

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

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

8. CLAIM. *For every $\{u, v\} \in E$ we have*

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

PROOF: Simply expanding the inner product $\langle H_u, H_v \rangle$ gives

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

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

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

Expanding the first expectation gives

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

and expanding the second gives

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

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

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

which proves the claim. ♦

From the above claim it follows that

$$\begin{aligned}\sum_{\{u,v\}\in E} A(u,v)(g_u, g_v) &= \frac{1}{R^2}\,\mathrm{SDP}_\infty(G, A) - \sum_{\{u,v\}\in E} A(u,v)\langle H_u, H_v\rangle \\ &\geq \left(\frac{1}{R^2} - \max_{u\in V}\|H_u\|_2^2\right)\mathrm{SDP}_\infty(G, A),\end{aligned}$$

where $\|H_u\|_2^2 = \langle H_u, H_u\rangle$.

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

$$\|H_u\|_2^2 \leq \frac{1}{R^2}\left(\frac{1}{2} + 2\vartheta(\overline{G})R\,\mathbb{E}_Z\left[\left\|\frac{Zf(u)}{R} - g_u(Z)\right\|_2^2\right]\right)^2.$$

By the definition of g_u, the vectors $Zf(u)$ and g_u are parallel. Moreover, they are equal if $\|Zf(u)\|_2 \leq R$. Since $f(u)$ is a unit vector, the r entries of the random vector $Zf(u)$ are i.i.d. $N(0, 1/r)$ random variables. Hence,

$$\begin{aligned}\mathbb{E}_Z\left[\left\|\frac{Zf(u)}{R} - g_u(Z)\right\|_2^2\right] &= \int_{\mathbb{R}^r} \mathbf{1}[\|x\| \geq R]\Big(\frac{\|x\|}{R} - 1\Big)\Big(\frac{r}{2\pi}\Big)^{r/2} e^{-r\|x\|^2/2} dx \\ &= \int_R^\infty \int_{S^{r-1}} \rho^{r-1}\Big(\frac{\rho}{R} - 1\Big)\Big(\frac{r}{2\pi}\Big)^{r/2} e^{-r\rho^2/2} d\rho d\tilde{\omega}_r(\xi) \\ &\leq \frac{r^{r/2}}{R\Gamma\big(\frac{r}{2}\big)} \int_R^\infty \rho^r e^{-r\rho^2/2} d\rho,\end{aligned}$$

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

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

where $\Gamma(a, x)$ is the lower incomplete Gamma function [AAR99, Eq. (4.4.5)].

Collecting the terms from above then gives the bound

$$\mathrm{SDP}_r(G, A) \geq \frac{1}{R^2}\left(1 - \left(\frac{1}{2} + \vartheta(\overline{G})\frac{2^{(r+1)/2}}{\sqrt{r}\Gamma\big(\frac{r}{2}\big)}\Gamma\Big(\frac{r+1}{2}, \frac{rR^2}{2}\Big)\right)^2\right) \mathrm{SDP}_\infty(G, A). \tag{5.7}$$

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

By Stirling's formula, we have that for some constant $C_1 > 0$, the inequality $\Gamma(x) \geq C_1 e^{-x} x^{x-1/2}$ holds (see for example [AS64, Eq. (6.1.37)]). Hence, for some constants $c, C > 0$, we have

$$\frac{2^{(r+1)/2}}{\sqrt{r}\Gamma\big(\frac{r}{2}\big)} \leq C\left(\frac{c}{r}\right)^{r/2} \tag{5.8}$$

The power series of the incomplete gamma function (see for example [AS64, Eq. (6.5.32)]) gives that if $a \leq x$, for some constant $C_2 > 0$, the inequality $\Gamma(a, x) \leq C_2 x^a e^{-x}$ holds. As $R \geq 2$, for some constants $d, D > 0$, we have

$$\Gamma\left(\frac{r+1}{2}, \frac{rR^2}{2}\right) \leq D\sqrt{r}\left(\frac{r}{d^{R^2}}\right)^{r/2}. \tag{5.9}$$

Putting together estimates (5.8) and (5.9) gives

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

Since $r \leq \log \vartheta(\overline{G})$ there is some constant C' such that for $R^2 = C'\big(\log \vartheta(\overline{G})\big)/r$, the above value is less than $1/4$. It follows that for this value of R, Inequality (5.7) is nontrivial and we get the result. □

5.7 Summary

In this chapter, we proved the first upper bounds depending on r on the rank-r graphical Grothendieck constants $K(r, G)$ for $r > 1$, giving a $1/K(r, G)$-approximation algorithm for the graphical Grothendieck problem with rank-r constraint based on its natural semidefinite relaxation. In particular, we obtained the best known approximation results for approximating the ground state energy for the Heisenberg model when the interaction graph has small chromatic number.

Chapter 6

Entanglement in multiplayer XOR games

The content of this chapter is based on joint work with Harry Buhrman, Troy Lee and Thomas Vidick [BBLV09].

6.1 Introduction

Due to Tsirelson's Theorem, the role of entanglement in two-player XOR games is reasonably well understood. As we have seen in the previous chapters, it implies that the violation ratio is always bounded by the Grothendieck constant. Moreover, it implies that there is a semidefinite program of size polynomial in the number of questions whose optimum value is exactly the entangled bias of a two-player XOR game. This contrasts with the classical setting, where even approximating the bias to within a small constant is NP-hard [Hås01].

Unfortunately, our understanding of entangled games does not extend far beyond the setting of two-player XOR games. Two-player games with larger answer sizes seem to be much harder to get a handle on (see however [BRSW10, JP11, Reg11] for some recent results on the violation ratios achievable in this setting). Even less is known about games involving three players or more. This is in part a reflection of the fact that multipartite entanglement is much less well understood, and seemingly much more diverse than its bipartite counterpart. A simple example of a three-player XOR game that exhibits properties of tripartite entanglement which bipartite entanglement cannot possess is Mermin's game (see Section 1.6.1). In this game classical players can attain bias at most $1/2$, but by sharing the three-qubit GHZ state entangled players can play the game perfectly by performing two-outcome measurements on their local

qubits. That is, the entangled bias equals 1. This kind of separation between entangled and classical biases is impossible in a two-player scenario. Cleve, Høyer, Toner and Watrous [CHTW04, Theorem 8] showed that if the entangled bias of a two-player XOR game equals 1, then the classical bias must be 1 as well. Another important example of a property unique to multipartite entanglement is that of *monogamy* [Ton09], which shows that there is a trade-off in the amount of entanglement between two quantum systems and the amount of entanglement between either one of them and a third system. Monogamy plays a role in many multiplayer nonlocal games [TV06].

Entanglement is not a resource that is easily created or manipulated, and when studying violation ratios of nonlocal games one may ask which types of entanglement are the most useful; in fact this is a question that has preoccupied physicists for the past four decades (see e.g. [MS07] for a survey). In the bipartite setting the most natural measure of entanglement of a quantum state $|\psi\rangle_{AB}$ is its von Neumann entropy $S(\psi) = -\mathrm{Tr}\rho_A \log \rho_A$, where $\rho_A = \mathrm{Tr}_B|\psi\rangle\langle\psi|$ is the reduced density matrix on Alice's subsystem. However, from the point of view of nonlocality, states with higher entropy are not always the most useful: while in any dimension d the maximally entangled state $|\psi_d\rangle = d^{-1/2}\sum_{i=1}^{d}|i\rangle|i\rangle$ has the largest entropy, for some games it is known not to be the best resource [AGG05, JP11, VW10, Reg11]. In fact, there is a different "maximally nonlocal" state, the *embezzlement state* [DH03, Oli10] $|\phi_d\rangle$ which is the state proportional to $\sum_{i=1}^{d} i^{-1/2}|i\rangle|i\rangle$: for any nonlocal game (not necessarily XOR) for which there is an optimal finite-dimensional strategy achieving the violation ratio, and any $\varepsilon > 0$, there is also a d and a strategy using $|\phi_d\rangle$ which achieves the violation ratio up to precision ε. Interestingly, this distinction is not apparent in two-player XOR games, for which it is known that the maximally entangled state is also optimal.

In the case of XOR games with more than two players, little is known about the power of specific states as a resource to produce nonlocal correlations. The most striking recent results in this area are due to Pérez-García et al. [PGWP$^+$08], who show that for every positive integer d, there exists a three-player XOR game with violation ratio $\Omega(\sqrt{d})$. Here d refers to the smallest local dimension of the entangled state for any of the three players. An interesting feature of their work is that it makes use of techniques from operator space theory, a field that had not been connected to problems related to nonlocality before. Unfortunately their proof techniques don't give much insight into what *kind* of states can be used to achieve such "unbounded violations".

6.2 Bounded violations for a large class of states

In this chapter we consider multiplayer XOR games in which the players are restricted to sharing specific patterns of entanglement. For this, we introduce two main types of N-partite entanglement. The first is a generalization of GHZ states that we call *Schmidt states*, named so because they admit a sort of tripartite Schmidt decomposition[1]: states of the form $|\psi\rangle = \sum_i \alpha_i |i\rangle^{\otimes N}$, for any sequence of positive (normalized) coefficients α_i.[2] Note that these also contain a natural generalization of the "universal" embezzlement states to more than two parties, and as such one might expect that they are the most highly nonlocal in the context of multiplayer games. The second type of states are formed by what we will refer to as *clique-wise entanglement*. Here, we consider the general setting where the N players are organized in k coalitions of r players each (a given player can take part in any number of coalitions). The members of each of the coalitions are allowed to share a GHZ state of arbitrary dimension, i.e., a state of the form $d^{-1/2} \sum_{i=1}^{d} |i\rangle \otimes \cdots \otimes |i\rangle$, among themselves. Note that this includes possible collections of EPR pairs shared among two-party coalitions, as these states are simply higher-dimensional two-party GHZ states. Clique-wise entanglement includes states that have been covered extensively in the literature on entanglement, such as GHZ states [GHZ89, Mer90, Zuk93, CB97, BCD01, RW08], which have even been realized experimentally [BPD$^+$99, PBD$^+$00], and tripartite stabilizer states (see Section 1.6.2), which are of fundamental importance to the theory of quantum error correction [Got97, Nes05, BFG06] and also appear in the context of nonlocal games [GTHB05, TGB06].

We denote by $\beta_{\mathrm{S}}^*(\mathcal{G})$ (resp. $\beta_{\mathrm{C}}^*(\mathcal{G})$) the maximal bias achievable in game $\mathcal{G}$ by players who are restricted to sharing a Schmidt state of arbitrary dimension (resp. arbitrary clique-wise entanglement). We note the following obvious relationships between the biases:

$$\beta(\mathcal{G}) \leq \beta_{\mathrm{S}}^*(\mathcal{G}) \leq \beta^*(\mathcal{G}) \qquad \text{and} \qquad \beta(\mathcal{G}) \leq \beta_{\mathrm{C}}^*(\mathcal{G}) \leq \beta^*(\mathcal{G}).$$

The main results of this chapter are constant upper bounds on the violation ratios of these quantities.

Concerning Schmidt states we prove the following.

[1] For bipartite states, the Schmidt decomposition is simply the singular value decomposition when the state is represented by a matrix.

[2] The assumption that the α_is are real and positive is not a restriction, as complex arguments can be introduced to them via a local unitary transformation done by one of the N players.

6.2.1. THEOREM. *Let $\mathcal{G}$ be an N-player XOR game. Then the maximum bias achievable by players sharing a Schmidt state $|\psi\rangle = \sum_{i=1}^{d} \alpha_i |i\rangle^{\otimes N}$, for an arbitrary dimension d, is at most a constant factor greater than the classical bias. More precisely,*

$$\beta_S^*(\mathcal{G}) \leq 2^{(3N-5)/2} K_G^{\mathbb{C}} \beta(\mathcal{G}),$$

where $K_G^{\mathbb{C}} \lesssim 1.40491$ is the complex Grothendieck constant (see Section 2.3.3).

The exponential dependence on the number of players is necessary in this theorem as Zukowski [Zuk93] gave an explicit sequence of N-player XOR games where players sharing an N-partite GHZ state can a achieve a bias that is $2^{-1}(\pi/2)^N$ times larger than the classical bias. The same conclusion follows from Mermin's game, but Zukowski's games give slightly larger separations between the two biases.

Theorem 6.2.1 generalizes–with slightly improved constants–a result of Pérez-García et al. who show a constant violation ratio for the case of GHZ states of arbitrary local dimension. The proof of the theorem also uses fairly elementary techniques compared to those used in [PGWP$^+$08].

Our second result deals with the case where the players share clique-wise entanglement. Even in this complex setting, we can show that the violation ratio is bounded by a constant depending only on the number of coalitions, and the number of players taking part in each of them, but independent of the dimension of the various states shared among the parties.

6.2.2. THEOREM. *Let $\mathcal{G}$ be an N-player XOR game. Then the maximum bias achievable by players sharing clique-wise entanglement, in which the players are organized in k coalitions of r players each, is greater than the classical bias by at most a constant factor depending only on k and r. More precisely,*

$$\beta_C^*(\mathcal{G}) \leq 2^{k(3r-5)/2} (K_G^{\mathbb{C}})^k \beta(\mathcal{G}).$$

Stabilizer states were considered in the context of XOR games in [GTHB05, TGB06], where it is shown that they allow for violations that grow exponentially with the number of players sharing them. In view of these results, one might hope to obtain explicit examples of three-player XOR games that exhibit the unbounded violation ratios proved possible by Pérez-García et al. by cleverly grouping some large numbers of players sharing a stabilizer state into three sets that, when treated as three players, can still obtain large violations. Based on a result by Bravyi et al. [BFG06] we obtain the following corollary of Theorem 6.2.2, showing that, unfortunately, such a construction impossible.

6.2.3. COROLLARY. *Let $\mathcal{G}$ be a 3-player XOR game in which the players are restricted to using a stabilizer state. Then the maximum bias achievable is bounded by a universal constant, independent of the specific state used, or its dimension. More formally, if $|\psi\rangle$ is an arbitrary stabilizer state, then the following inequality holds:*

$$\beta^*_{|\psi\rangle}(\mathcal{G}) \leq 8\,(K_G^{\mathbb{C}})^4\,\beta(\mathcal{G}).$$

The above two theorems and corollary provide a perhaps surprising counterpoint to another of [PGWP+08]'s results, mentioned above, which shows that some states can achieve much larger gaps. Together, these results indicate a large variation in nonlocality for multipartite states, which is already apparent through their use in XOR games. This contrasts with the bipartite scenario, where all states give at most constant violation ratios, and both the maximally entangled state and the embezzlement states are optimal resources.

In the following section we outline implications of the above results for Banach algebras, for hardness of approximation, and for parallel repetitions.

6.2.1 Implications

Implications for Banach algebras. Theorem 6.2.1 answers an open question of Pérez-García et al. They were particularly interested in this question because they were able to relate the violation ratio with Schmidt states to an old open problem of Varopoulos in Banach algebras [Var75]. Via the reductions given in [Dav73, PGWP+08] and in conjunction with the partial answers of Le-Merdy [LM98] and Pérez-García [PG06], Theorem 6.2.1 settles Varopoulos's question completely. We discuss this result in detail in the next chapter, where we explain our contribution separate from the context of nonlocal games, and sketch the connection to Schmidt states made in [PGWP+08].

Implications for hardness of approximation. On the one hand, Tsirelson's characterization of two-player entangled XOR games gives a means to efficiently compute the bias $\beta^*(\mathcal{G})$ to high accuracy via semidefinite programming. On the other hand, approximating the *classical* bias of two-player XOR games within a sufficiently small constant is NP-hard [Hås01]. Hence the natural relaxation that corresponds to allowing the players to share entanglement marks the transition from a hard optimization problem to a tractable one.

As our results show, for multiplayer XOR games the violation ratio can be tightly bounded when the players share specific forms of entanglement,

and it is interesting to ask whether the quantum bias can again be efficiently approximated. It turns out, however, that the situation in this case is quite different. In fact, our results imply the following:

6.2.4. THEOREM. *Unless P=NP, for any integer $N \geq 3$ there is no polynomial-time algorithm that approximates the maximum bias of an entangled N-player game in which the players are restricted to sharing either a Schmidt state or clique-wise entanglement to within a factor c for any constant $c > 1$.*

Our results only hold for the specific types of entanglement that we consider, and it could very well be the case that $\beta^*(\mathcal{G})$ can be computed exactly or approximated closely in polynomial-time for general entanglement. The proof of Theorem 6.2.4 follows from a hardness-of-approximation result for Max-E3-Lin2 due to Håstad and Venkatesh [HV04], and we give it in Section 6.7.

Implications for parallel repetition. *Parallel repetition* of a general two-player nonlocal game $\mathcal{G}$ refers to the following situation: The referee samples independently some number ℓ of question pairs $(i^1, j^1), \ldots, (i^\ell, j^\ell)$ from the probability distribution π associated to $\mathcal{G}$, and sends the ℓ questions $i^\ell, \ldots, i^\ell$ to Alice and $j^1, \ldots, j^\ell$ to Bob. The players are then expected to each return ℓ answers, one corresponding to each of their questions. They win this parallel repetition version of $\mathcal{G}$ if their answers win each of the ℓ instances of $\mathcal{G}$.

If the maximal winning probability of one round of $\mathcal{G}$ is at most $\omega(\mathcal{G}) < 1$, then it is intuitively clear that an ℓ-fold parallel repetition of it is even more difficult. Determining just how much more difficult parallel repetitions make a game, turns out to be a very non-trivial matter. In general, it is not true that the winning probability simply scales as one would expect, i.e. as $\omega(\mathcal{G})^\ell$, which would be the case if the repetitions of the game are performed sequentially [Fei91, CHTW04, Raz08, BHH$^+$08, KR10]. For two-party nonlocal games the celebrated parallel repetition theorem [Raz98, Hol07] states that the winning probability does decrease exponentially in the number of parallel repetitions of the game. Only recently, Kempe and Vidick [KV11] proved that the entangled winning probability of general nonlocal games decreases at all under parallel repetitions. This was shown to hold before for XOR games by Cleve et al. [CSUU08] and unique games by Kempe, Regev and Toner [KRT08].

Closely related to parallel repetition theorems are XOR lemmas. Let us recall that an XOR game can be defined by a probability distribution π on $[n] \times [n]$ and a sign matrix $\Sigma \in \{-1, 1\}^{n \times n}$. The ℓ-fold *XOR repetition* of an XOR

game $\mathcal{G} = (\pi, \Sigma)$ is again an XOR game, and is defined as $\mathcal{G}^{\otimes\ell} = (\pi^{\otimes\ell}, \Sigma^{\otimes\ell})$. Thus, in this game ℓ question pairs $(i^k, j^k)_{k=1\ldots\ell}$ are picked independently with respect to π, and all i^k are sent to Alice, j^k to Bob. In order to win the game, they should answer signs a and b respectively such that $ab = \Sigma_{i^1 j^1} \cdots \Sigma_{i^k j^k}$.

Cleve et al. [CSUU08] show that for any two-player XOR game $\mathcal{G}$, the game $\mathcal{G}^{\otimes\ell}$ has entangled bias exactly $\beta^*(\mathcal{G})^\ell$. Since the classical and quantum biases are within a constant factor of each other, this also implies that if $\beta^*(\mathcal{G}) < 1$, then the classical bias $\beta(\mathcal{G}^{\otimes\ell})$ must go down exponentially with ℓ (although it does not behave as nicely as the quantum bias with respect to taking XORs). Cleve et al. further use this XOR lemma for the entangled bias to show that the *winning probability* with entanglement of two-player XOR games behaves perfectly under parallel repetition, i.e., as $\omega(\mathcal{G})^\ell$ (such behavior is usually referred to as strong parallel repetition). In fact, quite generally XOR lemmas imply parallel repetition theorems [Ung09].

Surprisingly, our results (as well as the previous results by Pérez-García et al. [PGWP$^+$08]) imply that there is no such XOR lemma for classical XOR games in the N-player setting for $N > 2$. This can be seen as follows. Suppose that $\beta^*_S(\mathcal{G}) = 1$ and $\beta(\mathcal{G}) < 1$ for some game $\mathcal{G}$. Then clearly $\beta^*_S(\mathcal{G}^{\otimes\ell}) = 1$, and so by Theorem 6.2.1 it must be the case that the classical bias satisfies

$$\beta(\mathcal{G}) \geq \frac{1}{2^{(3N-5)/2}\, K_G^{\mathbb{C}}},$$

which is independent of ℓ. Mermin's game is an example of such a game.

6.3 Proof overview and techniques

The main technical contribution of this chapter is the expansion of the connection between violation ratios of two-player XOR games and Grothendieck's Inequality established by Tsirelson's Theorem. We relate violation ratios for N-player XOR games with the patterns of entanglement discussed above and certain multilinear extensions of Grothendieck's inequality. Let us briefly recall that in the case of two players sharing an entangled state $|\psi\rangle$, the easy direction of Tsirelson's Theorem follows from the simple observation that the expected value of the product of the players' answers determined by $\{-1, 1\}$-valued observables F and G, given by $\langle\psi|F \otimes G|\psi\rangle$, can be written as the inner-product of two complex vectors $(\langle\psi|F \otimes I) \cdot (I \otimes G|\psi\rangle)$. Hence, the optimization over entangled strategies is readily upper-bounded by an optimization

over complex unit vectors, which can in turn be related to the classical bias via Grothendieck's Inequality.

In the multiplayer case, this kind of connection between the entangled and classical biases is not so obvious. The bipartite structure needed to express the expectation as an inner product between vectors is lost already when we consider the case of three players, where this expectation has the form $\langle\psi|F\otimes G\otimes H|\psi\rangle$. The results of this chapter stem from the observation that the function $(F,G,H)\mapsto\langle\psi|F\otimes G\otimes H|\psi\rangle$ is still a multilinear functional, whose exact dependence on the coefficients of F,G and H will depend on the state $|\psi\rangle$. Hence we isolate certain classes of states $|\psi\rangle$ (Schmidt states and cliquewise entanglement, already described above) and study the functionals that arise from them.

Our proofs proceed in two steps. In the first step, we show that, given a class of states $|\psi\rangle$, a maximization over observables F,G,H can be upper-bounded by the maximization of a certain generalized inner product over unit vectors. This step greatly depends on the class of states $|\psi\rangle$ under consideration. In the second step, we bound this last optimization as a function of the classical bias, which is the maximization over products of $\{-1,1\}$-valued functions. This step involves a constant-factor loss, as indeed in general the classical bias is smaller than the quantum bias that we started with.

We illustrate those two steps in more detail below by giving an overview of the proof of Theorem 6.2.1 for the case of three-player games in which entangled players share GHZ states of arbitrary local dimension. We use the following definitions introduced in Section 2.3.4. Let us recall that for N-tensor $A:[n]^N\to\mathbb{R}$, we defined

$$\mathrm{OPT}(A)=\max\left\{\sum_{I\in[n]^N}A[I]\chi_1(i_1)\cdots\chi(i_N):\chi_1,\ldots,\chi_N:[n]\to\{-1,1\}\right\}$$

and

$$\mathrm{GIP}(A)=\sup\left\{\left|\sum_{i_1,\ldots,i_N=1}^{N}A[i_1,\ldots,i_N]\,\langle f_1(i_1),\ldots,f_N(i_N)\rangle\right|:\right.$$
$$\left.d\in\mathbb{N},\,f_1,\ldots,f_N:[n]\to B_{\mathbb{C}^d}\right\},$$

where $\langle x_1,\ldots,x_N\rangle=\sum_i(x_1)_i\cdots(x_N)_i$ is the generalized inner product. Note that if $A=\pi\circ\Sigma$ for some N-player XOR game $\mathcal{G}=(\pi,\Sigma)$, then $\mathrm{OPT}(A)$ is precisely the classical bias $\beta(\mathcal{G})$ of $\mathcal{G}$.

6.3.1 First step: relating the entangled bias to the GIP bias

We show that, when $|\psi\rangle$ has a certain structure, one can relate the expected value $\langle\psi|F \otimes G \otimes H|\psi\rangle$ to a certain natural *trilinear functional* over unit vectors. For instance, for the simplest case of GHZ states one obtains the generalized inner-product $\langle x, y, z\rangle$ defined above. Other types of states may lead to more complicated functionals, and hence this step crucially depends on the type of entanglement that the players are allowed to use. Note that, in contrast, the Schmidt decomposition implies that for the case of two-player games the only bilinear functional which arises is essentially a weighted inner product. The many inequivalent classes of multilinear functionals that one can obtain for the case of three or more players are a reflection of the much richer structure of multipartite entanglement.

For the case of GHZ states $|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} |\ell\rangle|\ell\rangle|\ell\rangle$ we show the following:

6.3.1. LEMMA. *Let $\mathcal{G} = (\pi, \Sigma)$ be a 3-player XOR game. Assume that the players are restricted to sharing a GHZ state, and denote the resulting maximum bias by $\beta^*_Z(\mathcal{G})$. Then the following bound holds:*

$$\beta^*_Z(\mathcal{G}) \leq \mathrm{GIP}(\pi \circ \Sigma).$$

PROOF: Fix an optimal strategy of the players based on the shared entangled state $|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} |\ell\rangle|\ell\rangle|\ell\rangle$, and let F_i, G_j, H_k be each player's $\{-1, 1\}$-valued observables in that strategy. Let $A = \pi \circ \Sigma$. The players' bias is given by

$$\begin{aligned}
\beta^*_Z(\mathcal{G}) &= \sum_{(i,j,k)\in[n]^3} A[i,j,k]\langle\psi|F_i \otimes G_j \otimes H_k|\psi\rangle \\
&= \frac{1}{d} \sum_{(i,j,k)\in[n]^3} A[i,j,k] \sum_{\ell,m=1}^{d} \langle\ell|F_i|m\rangle\langle\ell|G_j|m\rangle\langle\ell|H_k|m\rangle \\
&= \frac{1}{d} \sum_{m=1}^{d} \left(\sum_{(i,j,k)\in[n]^3} A[i,j,k] \sum_{\ell=1}^{d} \langle\ell|F_i|m\rangle\langle\ell|G_j|m\rangle\langle\ell|H_k|m\rangle \right) \\
&\leq \frac{1}{d} \sum_{m=1}^{d} \mathrm{GIP}(A) = \mathrm{GIP}(A).
\end{aligned}$$

The inequality holds as the inner sum on the third line is a generalized inner product of the m^{th} columns of the $\{-1, 1\}$-valued observables F_i, G_j, H_k, which are unit vectors since these matrices are unitary. □

6.3.2 Second step: relating the GIP bias to the classical bias

The second step in our proofs consists of upper-bounding the multilinear expression resulting from the first step by a similar optimization over real numbers of absolute value less than 1. This step involves a constant-factor loss, and is based on Tonge's Inequality, Theorem 2.3.10. For our second result, Theorem 6.2.2, we use an inequality proved by Carne [Car80] in the context of Banach lattices, combined with Grothendieck's inequality. Specialized to the case of real rank-3 tensors, Tonge's Inequality reads:

6.3.2. THEOREM. *For every positive integer $n \geq 2$ and any 3-tensor $A : [n]^3 \to \mathbb{R}$, we have*

$$\mathrm{GIP}(A) \leq 4K_G^{\mathbb{C}} \, \mathrm{OPT}(A). \tag{6.1}$$

Combining Lemma 6.3.1 with Theorem 6.3.2 gives $\beta_Z^*(\mathcal{G}) \leq 4\, K_G^{\mathbb{C}}\, \beta(\mathcal{G})$, since putting $A = \pi \circ \Sigma$ makes the maximum on the right-hand side of Eq. (6.1) exactly the classical bias $\beta(\mathcal{G})$. This proves Theorem 6.2.1 for the case of three entangled players sharing GHZ states.

Tonge's Inequality also plays a role in the proof of [PGWP$^+$08] showing a constant violation ratio in the case of GHZ states. It is used there as an intermediate step to show a relationship between different tensor norms that are in turn used to prove the bound. Their technique, however, does not seem to be easily adapted to the case of Schmidt states or clique-wise entanglement.

Outline of the rest of this Chapter In Section 6.4 we introduce a few more notational conventions and definition. In Section 6.5 we prove Theorem 6.2.1, extending the above techniques to the case of Schmidt states. In Section 6.6 we prove Theorem 6.2.2 and Corollary 6.2.3, extending the above techniques to the case of clique-wise entanglement and stabilizer states. In Section 6.8 prove Carne's Theorem. In Section 6.9 we pose an open question and we give a brief summary of this chapter in Section 6.10

6.4 Notation and definitions

The following definition will be useful in studying the different biases achievable by players who are restricted to sharing a specific type of entanglement.

6.4.1. DEFINITION. Let $\mathcal{G} = (\pi, \Sigma)$ be an N-player XOR game and $|\psi\rangle \in \mathcal{H}^{\otimes N}$ be a fixed entangled state shared by N players. Then the *bias restricted to* $|\psi\rangle$, denoted $\beta^*_{|\psi\rangle}(\mathcal{G})$, is defined as

$$\beta^*_{|\psi\rangle}(\mathcal{G}) \;=\; \max_{F_1,\ldots,F_N} \mathbb{E}_{I\sim\pi}\Big[\Sigma[I]\,\langle\psi|F_1(i_1)\otimes\cdots\otimes F_N(i_N)|\psi\rangle\Big]$$

where the maximum is taken over $F_1, \ldots, F_N : [n] \to \mathcal{O}(\mathcal{H})$.

The following setups are the ones that we will encounter most frequently, and for each we introduce a special notation for the bias. For the case of GHZ states $|\psi\rangle = d^{-1/2}\sum_{i=1}^{d}|i\rangle_1\cdots|i\rangle_N$ (of arbitrary dimension d) we will denote the maximum bias by $\beta^*_{\mathrm{Z}}(\mathcal{G})$, while for Schmidt states $|\psi\rangle = \sum_{i=1}^{d}\alpha_i|i\rangle_1\cdots|i\rangle_N$ (with arbitrary dimension d and real positive coefficients α_i satisfying $\sum_{i=1}^{d}\alpha_i^2 = 1$) we will use the notation $\beta^*_{\mathrm{S}}(\mathcal{G})$. Finally, clique-wise entanglement is any type of entanglement that can be obtained by grouping the N players into k coalitions of r players each (a given player can take part in any number of coalitions), and allowing the members of each of the coalitions to share a GHZ state of arbitrary dimension (recall that collections of EPR pairs shared among a two-party coalition are simply higher dimensional two-party GHZ states). In that case, we denote the maximal bias by $\beta^*_{\mathrm{C}}(\mathcal{G})$. This may depend on the parameters k and r, which are kept implicit so as not to overload the notation, but will always be clear from context.

6.5 Bounded violations for Schmidt states

In this section we prove Theorem 6.2.1. As this chapter is rather heavy on notation, we present the proof of Theorem 6.2.1 in three steps, in order to let the reader get accustomed to the various quantities involved. First, in Section 6.5.1 we analyze the maximum bias $\beta^*_{\mathrm{Z}}(\mathcal{G})$ achievable by strategies that are limited to sharing a GHZ state for games with an arbitrary number of players. In Section 6.5.2, we extend our proof to cover the case where the players are allowed to share a Schmidt state.

6.5.1 Strategies with GHZ states.

We prove the following lemma, which is a (straightforward) generalization of Lemma 6.3.1 proved in Section 6.3:

6.5.1. LEMMA. *Let $\mathcal{G} = (\pi, \Sigma)$ be an N-player game. Assume that the players are restricted to sharing a GHZ state (i.e., a state of the form $|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} |\ell\rangle^{\otimes N}$ where d is arbitrary). Then the maximum bias the players can achieve is upper-bounded by*

$$\beta_{\mathrm{Z}}^{*}(\mathcal{G}) \leq \mathrm{GIP}(\pi \circ \Sigma).$$

PROOF: For $i_1, \ldots, i_N \in [n]$, let $F_1(i_1), \ldots, F_N(i_N)$ be the $\{-1, 1\}$-valued observables used by the N players to play $\mathcal{G}$ while sharing state $|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} |\ell\rangle^{\otimes N}$. Let $A = \pi \circ \Sigma$. The players' bias is then given by

$$\begin{aligned}
\beta_{\mathrm{Z}}^{*}(\mathcal{G}) &= \sum_{I \in [n]^N} A[I] \langle \psi | F_1(i_1) \otimes \cdots \otimes F_N(i_N) | \psi \rangle \\
&= \frac{1}{d} \sum_{I \in [n]^N} A[I] \sum_{\ell, m=1}^{d} \langle \ell | F_1(i_1) | m \rangle \cdots \langle \ell | F_N(i_N) | m \rangle \\
&\leq \frac{1}{d} \sum_{m=1}^{d} \left(\sum_{I \in [n]^N} A[I] \sum_{\ell=1}^{d} \langle \ell | F_1(i_1) | m \rangle \cdots \langle \ell | F_N(i_N) | m \rangle \right) \\
&\leq \frac{1}{d} \sum_{m=1}^{d} \mathrm{GIP}(A) = \mathrm{GIP}(A).
\end{aligned}$$

The inequality holds as the inner sum on the third line is a generalized inner product of the m^{th} columns of the observables $F_1(i_1), \ldots, F_N(i_N)$, which are unit vectors since these matrices are unitary. □

The inequality $\beta_{\mathrm{Z}}^{*}(\mathcal{G}) \leq 2^{(3N-5)/2} K_{\mathrm{G}}^{\mathbb{C}} \beta(\mathcal{G})$ now follows from Lemma 6.5.1 and Theorem 2.3.10, since for $A = \pi \circ \Sigma$, we have that $\mathrm{OPT}(A)$ is precisely the classical bias of the game $\mathcal{G} = (\pi, \Sigma)$. This proves Theorem 6.2.1 for the special case of GHZ states.

6.5.2 Extension to Schmidt states.

We extend the result of Section 6.5.1 to the case of Schmidt states, thus proving Theorem 6.2.1 in full generality. For this, analogous to Lemma 6.5.1, it is sufficient to show that for a Schmidt state $|\psi\rangle = \sum_{\ell=1}^{d} \alpha_\ell |\ell\rangle^{\otimes N}$, we have

$$\beta_{|\psi\rangle}^{*}(\mathcal{G}) \leq \mathrm{GIP}(\pi \circ \Sigma)$$

The theorem then follows by setting $A = \pi \circ \Sigma$ and applying Theorem 2.3.10.

PROOF OF THEOREM 6.2.1: For N-player XOR game $\mathcal{G} = (\pi, \Sigma)$, define the tensor $A = \pi \circ \Sigma$. Suppose that the N players share the Schmidt state $|\psi\rangle = \sum_{\ell=1}^{d} \alpha_\ell |\ell\rangle^{\otimes N}$. For $i_1, \ldots, i_N \in [n]$, let $F_1(i_1), \ldots, F_N(i_N)$ be a choice of $\{-1, 1\}$-valued observables used by the players to achieve bias $\beta^*_{|\psi\rangle}(\mathcal{G})$.

We use the following claim, which shows that $|\psi\rangle$ can be expressed as a weighted sum of GHZ-type states.

9. CLAIM. *There exist nonnegative reals $\nu_1, \ldots, \nu_d$ such that $|\psi\rangle = \sum_{m=1}^{d} \nu_m |\phi_m\rangle$, where $|\phi_m\rangle = \sum_{\ell=1}^{m} |\ell\rangle^{\otimes N}$ for $m = 1, \ldots, d$ is a "partial" (un-normalized) GHZ state. Moreover, the ν_ℓ satisfy the following equation:*

$$\sum_{m,k=1}^{d} \nu_m \nu_k \cdot \min\{m, k\} = 1 \tag{6.2}$$

PROOF: Renaming the basis vectors as necessary, we can assume that $\alpha_1 \geq \cdots \geq \alpha_d$. Let $\nu_d = \alpha_d$ and $\nu_m = \alpha_m - \alpha_{m+1}$ for $\ell = 1, \ldots, d-1$. Then we have

$$|\psi\rangle = \sum_{m=1}^{d} \nu_m |\phi_m\rangle.$$

Moreover, Eq. (6.2) is immediate from the fact that $|\langle\psi|\psi\rangle| = 1$ and $\langle\phi_m|\phi_k\rangle = \min\{m, k\}$ (recall that $|\phi_{m'}\rangle$ itself was not normalized). ♦

This reformulation of $|\psi\rangle$ reduces the task of showing an upper bound on $\beta^*_{|\psi\rangle}(\mathcal{G})$ to a form similar to what we had before. Namely,

$$\beta^*_{|\psi\rangle}(\mathcal{G}) = \sum_{m,k} \nu_m \nu_k \sum_{I \in [n]^N} A[I] \langle\phi_m| F_1(i_1) \otimes \cdots \otimes F_N(i_N) |\phi_k\rangle.$$

For fixed m, k, each term of the sum involves unnormalized "partial" GHZ states, which can be handled in the same fashion as Lemma 6.5.1.

10. CLAIM. *For tensor A and states $|\phi_m\rangle$ as defined above, we have*

$$\sum_{I \in [n]^N} A[I] \langle\phi_m| F_1(i_1) \otimes \cdots \otimes F_N(i_N) |\phi_k\rangle \leq \min\{m, k\} \, \mathrm{GIP}(A).$$

PROOF: Writing out $|\phi_m\rangle$ and $|\phi_k\rangle$, we have

$$\sum_{I \in [n]^N} A[I] \langle\phi_m| F_1(i_1) \otimes \cdots \otimes F_N(i_N) |\phi_k\rangle =$$

$$\sum_{I \in [n]^N} A[I] \sum_{s=1}^{m} \sum_{t=1}^{k} \langle s|F_1(i_1)|t\rangle \cdots \langle s|F_N(i_N)|t\rangle.$$

We will order the double sum over s, t depending on whether m or k is smaller—we want the outer sum to be over the smaller one. Suppose that $m \leq k$. The other case is completely analogous. Then

$$\sum_{I\in[n]^N} A[I] \sum_{s=1}^{m}\sum_{t=1}^{k} \langle s|F_1(i_1)|t\rangle \cdots \langle s|F_N(i_N)|t\rangle = \sum_{s=1}^{m} \left(\sum_{I\in[n]^N} A[I] \sum_{t=1}^{k} \langle s|F_1(i_1)|t\rangle \cdots \langle s|F_N(i_N)|t\rangle \right). \quad (6.3)$$

For each fixed s, the inner sum is now a generalized inner product of the first k entries of the s^{th} rows of the matrices $F_1(i_1), \ldots, F_N(i_N)$. Since the full rows of these matrices have norm at most 1, we obtain complex vectors of norm at most 1 by taking only their first k coordinates. Hence, we have

$$\sum_{I\in[n]^N} A[I]\langle \phi_m|F_1(i_1)\otimes\cdots\otimes F_N(i_N)|\phi_k\rangle \leq \min\{m,k\}\,\mathrm{GIP}(A).$$

This proves the claim. ♦

We can now finish the proof of the theorem. Combining the above two claims gives

$$\begin{aligned}\beta^*_{|\psi\rangle}(\mathcal{G}) &= \sum_{m,k} \nu_m\nu_k \sum_{I\in[n]^N} A[I]\langle \phi_m|F_1(i_1)\otimes\cdots\otimes F_N(i_N)|\phi_k\rangle \\ &\leq \sum_{m,k} \nu_m\nu_k \min\{m,k\}\,\mathrm{GIP}(A) \\ &= \mathrm{GIP}(A).\end{aligned}$$

The first line follows from Claim 10 and the last from Claim 9. □

6.6 Bounded violations for clique-wise entanglement

The proof of Theorem 6.2.2 is based on a result by Carne [Car80] that essentially shows how Grothendieck-type inequalities can be composed in order to prove new inequalities of the same type. This will let us prove bounds on the entangled bias when the players are allowed to share any combination of EPR pairs and GHZ states. We explain Carne's theorem in Section 6.6.1, we explain how it is applied to prove Theorem 6.2.2 in Section 6.6.2 and we end this section with a proof of Corollary 6.2.3.

6.6.1 Carne's Theorem

Carne's Theorem is most easily explained with the use of hypergraphs. A *hypergraph* $H = (V, E)$ consists of a finite set V of vertices and a family E of subsets (called hyper-edges) of V. In a normal graph, the edge set E consists of pairs of vertices, but in a hypergraph, the hyper-edges are allowed to have any size ranging from 1 to $|V|$. For a vertex $u \in V$, we denote by $E(u)$ the set of hyper-edges $e \in E$ that contain u as an element.

Towards understanding Carne's Theorem, let $H = (V, E)$ be a hypergraph. We associate with each hyper-edge $e \in E$ and vertex $u \in e$ a complex Hilbert space $\mathcal{H}(u, e)$. Furthermore, we associate with every edge $e \in E$ a linear functional $\phi_e : \bigotimes_{u \in e} \mathcal{H}(u, e) \to \mathbb{C}$. Later, every vertex $u \in V$ will correspond to a player in a $|V|$-player XOR game and for every hyper edge e containing u, the space $\mathcal{H}(u, e)$ will be u's local Hilbert space for some state $|\phi_e\rangle \in \bigotimes_{v \in e} \mathcal{H}(v, e)$ that u shares with the other members of e. The linear functionals ϕ_e will correspond to generalized inner products that arise when the $|\phi_e\rangle$ are GHZ states.

Suppose that every ϕ_e satisfies a Grothendieck-type inequality, by which we mean that for every $|e|$-tensor $A : [n]^{|e|} \to \mathbb{R}$ and functions $f_u : [n] \to B_{\mathcal{H}(u,e)}$, for each $u \in e$, the inequality

$$\Big| \sum_{I \in [n]^{|e|}} A[I] \phi_e\Big(\bigotimes_{u \in e} f_u(i_u) \Big) \Big| \leq C_e \ \mathrm{OPT}(A), \tag{6.4}$$

holds for some constant C_e independent of A and the f_u. The functionals ϕ_e that we will encounter below are those for which $\phi_e\big(\bigotimes_{u \in e} x_u \big)$ is the generalized inner product between the vectors $x_u \in \mathcal{H}(u, e)$.

Define for every $u \in V$ the Hilbert space $\mathcal{H}_u = \bigotimes_{e \in E(u)} \mathcal{H}(u, e)$. Carne's Theorem then states that a certain natural combination of the linear functionals ϕ_e in a general multilinear functional Φ defined over the entire Hilbert space $\mathcal{H} = \bigotimes_{u \in V} \mathcal{H}_u$ also satisfies a Grothendieck-type inequality with a constant equal to the product of the C_e. This combination of the ϕ_e is precisely the type we obtain by allowing the players in each hyper-edge to share a GHZ state. Since a vertex u can be part of many different edges, there can be many functionals ϕ_e that act on the same space $\mathcal{H}_u$. This is what makes Carne's Theorem non-trivial. We need one last thing, which is the linear *re-arranging map*

$$\sigma : \bigotimes_{u \in V} \Big(\bigotimes_{e \in E(u)} \mathcal{H}(u, e) \Big) \to \bigotimes_{e \in E} \Big(\bigotimes_{u \in e} \mathcal{H}(u, e) \Big),$$

which simply permutes the elements of a vector $x \in \bigotimes_{u \in V} \mathcal{H}_u$.

6.6.1. THEOREM (CARNE). *Define the linear functional* $\Phi : \bigotimes_{u\in V} \mathcal{H}_u \to \mathbb{C}$ *as* $\Phi = \Big(\bigotimes_{e\in E} \phi_e\Big) \circ \sigma$, *where* $\circ$ *denotes the composition of the two maps. Then, for any* $|V|$*-tensor* $A : [n]^{|V|} \to \mathbb{R}$ *and set of functions* $f_u : [n] \to B_{\mathcal{H}_u}$, *for* $u \in V$, *we have*

$$\sum_{I\in[n]^V} A[I]\, \Phi\Big(\bigotimes_{u\in V} f_u(i_u)\Big) \leq \Big(\prod_{e\in E} C_e\Big)\, \mathrm{OPT}(A), \tag{6.5}$$

where the C_e *are as in Eq.* (6.4).

If for each $e \in E$, the functional ϕ_e gives the generalized inner product between vectors $x_u \in \mathcal{H}(u,e) = \mathbb{C}^d$, then by Theorem 2.3.10 (Tonge's Inequality), we get that Inequality 6.4 holds with $C_e = 2^{(3|e|-5)/2} K_G^{\mathbb{C}}$.

6.6.2 Bounding the violations achievable by strategies with clique-wise entanglement

Consider an N-player XOR game $\mathcal{G} = (\pi, \Sigma)$. Let the players be organized in k coalitions of r players each[3], where each player can take part in any number of coalitions. Each coalition is allowed to share a GHZ state between its members.

To model this setup, we associate a hypergraph $H = (V, E)$ to the coalition structure, with $V = [N]$ and there is a hyperedge for every coalition. For every hyper edge e we introduce a Hilbert space $\mathcal{H}(e) = \bigotimes_{u\in e} \mathcal{H}(u,e)$, where $\mathcal{H}(u,e)$ is a local space of player u corresponding to edge e. The state of the players in this space is a GHZ state $|\phi_e\rangle = d^{-1/2} \sum_{j=1}^{d} |j\rangle^{\otimes|e|}$. The global entangled state shared by the players at the start of the game is then

$$|\tilde{\Phi}\rangle = \bigotimes_{e\in E} |\phi_e\rangle \in \bigotimes_{e\in E} \Big(\bigotimes_{u\in e} \mathcal{H}(u,e)\Big) \tag{6.6}$$

Finally, each player u has an observable $F_u(i_u)$ corresponding to question $i_u \in [N]$. These act on player u's local space $\mathcal{H}_u = \bigotimes_{e\in E(u)} \mathcal{H}(u,e)$.

Theorem 6.2.2 states that the maximum bias achievable by a strategy of the form that we have just described is at most a constant times the classical bias of the game. In order to prove it, we first relate the bias with any $\{-1,1\}$-valued observables $\{F_u(i_u)\}_{u\in V}$ to an expression similar to the one appearing on the left-hand side of Inequality (6.9) in Carne's Theorem, where ϕ_e will be the linear functional associated with the GHZ state. More precisely, for every

[3]The organization of these coalitions is independent of the game itself; rather it is used to define the structure of the entanglement that is shared between the players.

$e \in E$, the Hilbert spaces $\mathcal{H}(u,e)$ for $u \in e$ will be $\mathbb{C}^d$ (for some d) and ϕ_e will be such that for any set of vectors $\{x_u\}_{u\in e} \subseteq \mathbb{C}^d$ the value $\phi_e\big(\bigotimes_{u\in e} x_u\big)$ equals the generalized inner product of the x_us. Applying Theorem 6.6.1 will conclude the argument.

PROOF OF THEOREM 6.2.2: Fix observables $\{F_u\}_{u\in V}$ and an entangled state $|\Phi\rangle = \sigma^{-1}(|\tilde{\Phi}\rangle)$, where $|\tilde{\Phi}\rangle$ is described in Eq. (6.6) and σ is the rearrangement map that appears in Carne's Theorem. This map appears because we need to re-arrange the terms of $|\tilde{\Phi}\rangle$ to correspond to the decomposition of space $\bigotimes_{u\in V} \mathcal{H}_u$. (We omit the arguments i_u in the F_u for now to suppress notation and because they do not play a role at this moment.)

We begin by expanding the expectation $\langle\Phi| \bigotimes_{u\in V} F_u|\Phi\rangle$, with the goal of relating it to the map Φ of Theorem 6.6.1. Let $[d]^E$ denote the set of $|E|$-tuples of the form $(j_e)_{e\in E}$ where each j_e is an integer in $[d]$. Recall that the members of an edge $e \in E$ share a state of the form $|\phi_e\rangle = d^{-1/2}\sum_{j_e=1}^{d} |j_e\rangle^{\otimes|e|}$. We have

$$\begin{aligned} |\Phi\rangle &= \sigma^{-1}\left(\frac{1}{\sqrt{d^{|E|}}} \bigotimes_{e\in E}\Big(\sum_{j_e=1}^{d}\bigotimes_{u\in e}|j_e\rangle\Big)\right) \\ &= \frac{1}{\sqrt{d^{|E|}}}\sum_{J\in[d]^E}\bigotimes_{u\in V}|J_{|E(u)}\rangle \end{aligned}$$

where $J_{|E(u)}$ denotes the tuple $(j_e)_{e\in E(u)}$ and $|J_{|E(u)}\rangle = \bigotimes_{e\in E(u)}|j_e\rangle$ is a state in the Hilbert space $\mathcal{H}_u$ of player u.

Since observables are Hermitian, the expected value $\langle\Phi| \bigotimes_{u\in V} F_u|\Phi\rangle$ equals

$$\begin{aligned} \langle\Phi|\bigotimes_{u\in V}F_u|\Phi\rangle &= \frac{1}{2\cdot d^{|E|}}\sum_{J',J\in[d]^E}\left(\prod_{u\in V}\langle J'_{|E(u)}|F_u|J_{|E(u)}\rangle + \prod_{u\in V}\langle J_{|E(u)}|F_u|J'_{|E(u)}\rangle\right) \\ &= \frac{1}{2\cdot d^{|E|}}\sum_{J',J\in[d]^E}\left(\prod_{u\in V}\langle J'_{|E(u)}|F_u|J_{|E(u)}\rangle + \prod_{v\in V}\langle J'_{|E(v)}|F_v^*|J_{|E(v)}\rangle\right) \\ &= \frac{1}{d^{|E|}}\sum_{J'\in[d]^E}\left(\sum_{J\in[d]^E}\Re\Big(\prod_{u\in V}\langle J'_{|E(u)}|F_u|J_{|E(u)}\rangle\Big)\right). \end{aligned} \tag{6.7}$$

Note that, since the expression on the left-hand side is real, the one on the right-hand side is too, and we can safely ignore the $\Re$ symbol on the right. Since the F_u are unitary matrices, their columns are unit vectors. This implies that there exist unit vectors $x_u \in \bigotimes_{e\in E(u)}\mathcal{H}(u,e)$ (depending on J') such that the expression between the brackets in equation (6.7) is of the form

$$\sum_{J\in[d]^E}\prod_{u\in V}(x_u)_{J_{|E(u)}}$$

where $(x_u)_{J_{|E(u)}}$ denotes $J_{|E(u)}$-coordinate of the vector x_u, when written in the basis defined by the vectors $|J_{|E(u)}\rangle$.

11. CLAIM. *For $\Phi = \Big(\bigotimes_{e\in E}\phi_e\Big)\circ\sigma$ where each ϕ_e corresponds to the generalized inner product function on $\bigotimes_{u\in e}\mathcal{H}(u,e)$, we have*

$$\sum_{J\in[d]^E}\prod_{u\in V}(x_u)_{J_{|E(u)}} = \Phi\Big(\bigotimes_{u\in V}x_u\Big).$$

PROOF: Since Φ is linear, it suffices to prove the claim for vectors of the form $x_u = \bigotimes_{e\in E(u)}x_{u,e}$, where each $x_{u,e}\in\mathcal{H}(u,e)$. In this case, we have

$$\begin{aligned}
\Big(\bigotimes_{e\in E}\psi_e\Big)\circ\sigma\Big(\bigotimes_{u\in V}\big(\bigotimes_{e\in E(u)}x_{u,e}\big)\Big) &= \bigotimes_{e\in E}\Big(\psi_e\big(\bigotimes_{u\in e}x_{u,e}\big)\Big)\\
&= \prod_{e\in E}\Big(\sum_{j_e=1}^{d}\Big(\prod_{u\in e}(x_{u,e})_{j_e}\Big)\Big)\\
&= \sum_{J\in[d]^E}\prod_{e\in E}\Big(\prod_{u\in e}(x_{u,e})_{j_e}\Big)\\
&= \sum_{J\in[d]^E}\prod_{u\in V}\Big(\prod_{e\in E(u)}(x_{u,e})_{j_e}\Big),
\end{aligned}$$

where the last product is $\prod_{e\in E(u)}(x_{u,e})_{j_e} = (x_u)_{J_{|E(u)}}$. □

Let $F_u(i_u)$ be the observable used by player u on question i_u, so that the bias achieved by this strategy in the game $\mathcal{G} = (\pi,\Sigma)$ is

$$\sum_{I\in[n]^V}A[I]\,\langle\Phi|\bigotimes_{u\in V}F_u(i_u)|\Phi\rangle$$

where $A = \pi\circ\Sigma$. We can bound this expression by

$$\begin{aligned}
&\Big|\sum_{I\in[n]^V}A[I]\left(\frac{1}{d^{|E|}}\sum_{J'\in[d]^E}\sum_{J\in[d]^E}\prod_{u\in V}\big[F_u(i_u)\big]_{J_{|E(u)},J'_{|E(u)}}\right)\Big|\\
&\qquad\le\frac{1}{d^{|E|}}\sum_{J'\in[d]^E}\Big|\sum_{I\in[n]^V}A[I]\cdot\sum_{J\in[d]^E}\prod_{u\in V}\big[F_u(i_u)\big]_{J_{|E(u)},J'_{|E(u)}}\Big|\\
&\qquad\le\max_{J'\in[d]^E}\Big|\sum_{I\in[n]^V}A[I]\cdot\sum_{J\in[d]^E}\prod_{u\in V}\big[F_u(i_u)\big]_{J_{|E(u)},J'_{|E(u)}}\Big|\\
&\qquad\le\max_{f_u:[n]\to B_{\mathcal{H}_u}:\,u\in V}\Big|\sum_{I\in[n]^V}A[I]\,\Phi\Big(\bigotimes_{u\in V}f_u(i_u)\Big)\Big|, \qquad (6.8)
\end{aligned}$$

where the first equality is (6.7), and the last inequality follows from Claim 11. The result then follows directly from Theorem 6.6.1 combined with the bound in Theorem 2.3.10, giving the last part of the theorem. □

We end this section with a proof of Corollary 6.2.3.

PROOF OF COROLLARY 6.2.3: Theorem 5 in [BFG06] states that, if $|\Psi\rangle$ is any stabilizer state shared in an arbitrary way among three parties, then there exist unitary matrices U_1, U_2 and U_3 on the Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$, respectively, such that $U_1 \otimes U_2 \otimes U_3|\Psi\rangle$ is a state of the form $|\Phi\rangle$ considered above. In other words, $|\Psi\rangle$ is local-unitarily equivalent to a number of EPR pairs shared between each of the three pairs of players, together with a GHZ state shared in common. By defining local observables $U_1 F_1(i_1) U^*$, etc, it is not difficult to see that for any three-player XOR game $\mathcal{G}$, the bias $\beta^*_{|\Psi\rangle}(\mathcal{G})$ is at most the bias attainable with clique-wise entanglement shared among the three players.

It now suffices to consider the hypergraph H with vertex set $V = \{1,2,3\}$, and edge set $E = \{\{1,2\},\{2,3\},\{1,3\},\{1,2,3\}\}$. In the notation of Theorem 6.2.2, this hypergraph has $k = 4$ and $r \leq 3$, which gives the bound $2^8(K_G^{\mathbb{C}})^4$. However, a careful examination of the proof of Theorem 6.2.2 easily reveals that the inequality holds with the smaller constant $8(K_G^{\mathbb{C}})^4$. □

6.7 Hardness of approximation of the entangled bias

Khot and Naor [KN08] observed that the hardness-of-approximation results for Max-E3-Lin2 of Håstad and Venkatesh [HV04] can be extended to:

6.7.1. THEOREM (HÅSTAD-VENKATESH-KHOT-NAOR). *Unless P=NP, there is no polynomial-time algorithm that approximates the classical bias of a three-party XOR game to within a multiplicative factor c for any constant $c > 1$.*

The inapproximability results in [HV04] only hold for *symmetric* strategies, in which the players all share the same strategy. However, Khot and Naor show that the inapproximability result holds even when restricted to games $\mathcal{G} = (\pi, \Sigma)$ that are invariant under permutations of the three players (i.e. for $A = \pi \circ \Sigma$ we have $A[i,j,k] = A[i,k,j] = A[j,i,k] = A[j,k,i] = A[k,i,j] = A[k,j,i]$) and are such that the same question is never asked to two players simultaneously (i.e. $A[i,j,j] = A[j,i,j] = A[j,j,i] = 0$). In this case Lemma 2.1 in [KN08] shows that the optimum with respect to symmetric strategies is

within a factor 10 of the general optimum. Combining this result with Theorems 6.2.1 and 6.2.2 immediately proves Theorem 6.2.4. Indeed, Theorem 6.2.1 (resp. Theorem 6.2.2) shows that, as long as the players are restricted to using an arbitrary Schmidt state (resp. clique-wise entanglement), the quantum bias is at most a constant times the classical bias. Hence any constant-factor approximation to the quantum bias would give a constant approximation to the classical bias, which is ruled out by the hardness result from [HV04].

6.8 Proof of Carne's Theorem

In this section we prove Theorem 6.6.1, which we restate here for convenience.

6.8.1. THEOREM (CARNE). *Define the linear functional* $\Phi : \bigotimes_{u\in V} \mathcal{H}_u \to \mathbb{C}$ *as* $\Phi = \Big(\bigotimes_{e\in E} \phi_e\Big) \circ \sigma$, *where* $\circ$ *denotes the composition of the two maps. Then, for any* $|V|$*-tensor* $A : [n]^{|V|} \to \mathbb{R}$ *and set of functions* $f_u : [n] \to B_{\mathcal{H}_u}$, *for* $u \in V$, *we have*

$$\sum_{I\in[n]^V} A[I]\,\Phi\Big(\bigotimes_{u\in V} f_u(i_u)\Big) \leq \Big(\prod_{e\in E} C_e\Big)\,\mathrm{OPT}(A), \tag{6.9}$$

where the C_e *are as in Eq.* (6.4).

PROOF OF THEOREM 6.8.1: The proof is by induction on the number of edges $|E|$. If the edge set is empty, then there is nothing to prove. Let e_0 be any edge in the hypergraph H, and consider the graph $H_0 = (V, E\backslash\{e_0\})$. To re-write the expression, first assume that each vector $f_u(i_u) \in \mathcal{H}_u = \bigotimes_{e\in E(u)} \mathcal{H}(u,e)$ has the following tensor structure:

$$f_u(i_u) = f_u^0(i_u) \otimes f_u^1(i_u),$$

where $f_u^0(i_x) \in \otimes_{e\in E\backslash\{e_0\}} \mathcal{H}(u,e)$ and $f_u^1(i_u) \in \mathcal{H}(u,e_0)$.

Define $\Phi_{H_0} = \Big(\bigotimes_{e\in E\backslash\{e_0\}} \phi_e\Big) \circ \sigma_{H_0}$, where σ_{H_0} is the re-arranging map for H_0. With this notation we have

$$\begin{aligned}\Phi\Big(\bigotimes_{u\in V} f_u(i_u)\Big) &= \Phi\Big(\bigotimes_{u\in V} f_u^0(i_u) \otimes f_u^1(i_u)\Big)\\ &= \Phi_{H_0}\Big(\bigotimes_{u\in V} f_u^0(i_u)\Big) \cdot \phi_{e_0}\Big(\bigotimes_{u\in e_0} f_u^1(i_u)\Big)\end{aligned}$$

Define the tensor $B[I] = A[I] \cdot \phi_{e_0}\left(\bigotimes_{u \in e_0} f_u^1(i_u)\right)$. Applying the induction hypothesis to $B[I]$ and the graph H_0 (note that the $\phi_{e_0}(\cdots)$ term is simply a number, dependent on I) gives

$$\sum_{I \in [n]^V} B[I] \cdot \Phi_{H_0}\Big(\bigotimes_{u \in V} f_u^0(i_u)\Big) \leq \Big(\prod_{e \in E \setminus \{e_0\}} C_e\Big) \operatorname{OPT}(B) \tag{6.10}$$

By definition,

$$\begin{aligned}\operatorname{OPT}(B) &= \max\left\{\sum_I B[I] \prod_{u \in V} \chi_u(i_u) : \chi_u : [n] \to \{-1,1\}\right\} \\ &= \max\left\{\sum_I A[I] \left(\prod_{u \in V} \chi_u(i_u)\right) \phi_{e_0}\left(\bigotimes_{u \in e_0} f_u^1(i_u)\right) : \chi_u : [n] \to \{-1,1\}\right\}.\end{aligned}$$

Fix χ_u that achieve this maximum, and define the tensor $C[I] = A[I] \prod_{u \in V} \chi_u(i_i)$. By hypothesis, the function ϕ_e enjoys a Grothendieck-type inequality, hence the expression above can be bounded by

$$\operatorname{OPT}(B) = \sum_I C[I] \cdot \phi_{e_0}\left(\bigotimes_{x \in e_0} f_x^1(i_x)\right) \leq C_{e_0} \operatorname{OPT}(C) \tag{6.11}$$

To conclude, we can relate $\operatorname{OPT}(C)$ to $\operatorname{OPT}(A)$ in the following way:

$$\begin{aligned}\operatorname{OPT}(C) &= \max\left\{\sum_I C[I] \prod_{u \in V} \chi'_u(i_u) : \chi'_u : [n] \to \{-1,1\}\right\} \\ &= \max\left\{\sum_I A[I] \prod_{u \in V} \chi_u(i_u)\chi'_u(i_u) : \chi'_u : [n] \to \{-1,1\}\right\} \\ &= \max\left\{\sum_I A[I] \prod_{u \in V} \chi''_u(i_u) : \chi''_u : [n] \to \{-1,1\}\right\} \\ &= \operatorname{OPT}(A).\end{aligned}$$

Combining Eqs. (6.10) and (6.11) gives the result in the case where all $f_u(i_u)$ have the tensor structure we described earlier. If not, since Φ is linear, writing their Schmidt decomposition will result in a weighted sum of expressions involving only unit vectors of this form. The weighted sum can be bounded by its maximum component, for which we can apply the reasoning above. □

6.9 Open questions

We proved that Schmidt states admit only constant-factor violation ratios. We also proved that clique-wise entanglement admits violation ratios bounded from above by a factor depending only on the number of coalitions. Clique-wise entanglement consists of combinations of GHZ states, which are special cases of Schmidt states. Unfortunately, we were not able to unify these results in the sense that we defined clique-wise entanglement as combinations of Schmidt states. A natural open question thus is: Do the bounds on the violation ratios for clique-wise entanglement still hold when we allow the players in each coalition to share general Schmidt states, instead of restricting them to sharing GHZ states?

6.10 Summary

In this chapter, we considered the problem of *upper bounding* the largest possible violation ratio for XOR games that involve possibly many players and where entangled players are restricted to using one of two types of entanglement: Schmidt states, or clique-wise entanglement. We proved that when the players use these types of entanglement, their advantage over classical players is at most a *constant* factor, depending only on the number of players. The case of Schmidt states settled an open problem of [PGWP$^+$08] and by a reduction given in that paper, a much older problem of [Var75] (see Chapter 7). The case of clique-wise entanglement shows that, perhaps surprisingly, entanglement consisting of arbitrary combinations of EPR pairs and GHZ states shared among the players is *insufficient* to reproduce the violation ratios proved possible in [PGWP$^+$08]. A theorem of [BFG06] implies that the same holds for stabilizer states.

Chapter 7

A problem of Varopoulos: Schatten spaces with the Schur product are Q-algebras

The content of this chapter is based on joint work with Harry Buhrman, Troy Lee and Thomas Vidick [BBLV11].

7.1 Introduction

And now for something completely different. In this chapter, we discuss an old problem posed by Varopoulos [Var75] in the context of Banach algebras. Our contribution to this problem is the solution to a part of it that, when put in conjunction with a series of previous results due to Pietsch and Triebel [PT68], Varopoulos [Var72], Davie [Dav73], Le-Merdy [LM98] and Pérez-García [PG06], leads to its complete resolution.

We begin by giving an informal explanation of what Varopoulos's question is about. Put briefly, the question asks for the existence of isomorphisms between Banach algebras. Roughly speaking, a Banach algebra is a vector space in which one can add the elements and multiply them by scalars as usual, but in addition one can multiply the elements themselves. Two Banach algebras are isomorphic if there is a linear bijection (i.e., a one-to-one correspondence defined by a linear function) between the underlying vector spaces that preserves the multiplication operations. Varopoulos's question concerns two types Banach algebras:

1. algebras in which the vector spaces are formed by sets of matrices and the additional multiplication operation is the entry-wise multiplication

(known as the Schur or Hadamard product),

2. algebras formed by a vector space of complex-valued functions which can be multiplied in the obvious way (the product fg defined by $(fg)(x) = f(x)g(x)$).

Roughly, the problem is to determine whether the first kind (1) are isomorphic to Banach algebras (called Q-algebras) formed by *cosets* of the second kind (2) (details follow below).

Although this problem may appear completely unrelated to the rest of this thesis, Pérez-García et al. [PGWP$^+$08] showed that part of it is equivalent to the problem of determining whether Schmidt states (defined in Chapter 6) allow for arbitrarily large violation ratios in multiplayer XOR games: a problem that we solved in the negative in the previous chapter. The negative answer to the Schmidt state problem (Theorem 6.2.1) in conjunction with the results mentioned above, implies a positive answer for the Banach algebra problem.

The main purpose of this chapter is to explain Varopoulos's question in more detail and to explain which part of it was solved by Theorem 6.2.1. We give a proof of our contribution separate from the context of XOR games and explain the relation to the Schmidt state problem found in [PGWP$^+$08] afterwards. Last, we briefly sketch why the whole problem is solved when our result is put in conjunction with the results mentioned above.

Before continuing, I want to confess that I am a layman in the subject matter of this chapter. Clearly, the problem about to be discussed was solved in the most part due to more significant partial results and reformulations of others. Nevertheless, the hope is that this presentation may be useful in some way.

In the remainder of this section we gather the mathematical tools needed to explain Varopoulos's problem precisely. More details of the following information can be found in the excellent books by Diestel, Jarchow and Tonge [DJT95], Reed and Simon [RS72], Rudin [Rud86] and Simon [Sim05].

7.1.1 Banach algebras

A *complex algebra* $\mathcal{X} = (\mathcal{V}, *)$ is a vector space $\mathcal{V}$ over $\mathbb{C}$ in which a multiplication $*$ is defined that is distributive and associative. For $A, B \in \mathcal{V}$, we have $A * B \in \mathcal{V}$. If $\mathcal{V}$ has a norm $\| \ \|$ defined on it that satisfies for all $A, B \in \mathcal{V}$

$$\|A * B\| \leq \|A\|\|B\|,$$

then $\mathcal{X}$ is called a *normed complex algebra*. If $\mathcal{V}$ is complete with respect to this norm, then $\mathcal{X}$ is called a *Banach algebra*.

A Banach algebra is *commutative* if the multiplication operation is commutative, that is, $A * B = B * A$ for all $A, B \in \mathcal{V}$.

An important example is the space $\mathcal{C}(K)$ of continuous functions $f : K \to \mathbb{C}$ on a metric space K, which is a Banach algebra when endowed with the *uniform norm* (also called the supremum norm), defined as $\|f\| = \sup\{|f(x)| : x \in K\}$, and the pointwise multiplication, given by $(f * g)(x) = f(x)g(x)$. Often this Banach algebra is also denoted by $\mathcal{C}(K)$, as we will do here.

7.1.2 Q-algebras

Two complex algebras $\mathcal{X} = (\mathcal{V}, *)$ and $\mathcal{Y} = (\mathcal{W}, \cdot)$ are *isomorphic* if there exists a linear bijective map $\varphi : \mathcal{V} \to \mathcal{W}$ that preserves the multiplication in the sense that $\varphi(A * B) = \varphi(A) \cdot \varphi(B)$ for all $A \in \mathcal{V}$ and $B \in \mathcal{W}$. Such a map φ is called an isomorphism. If $\mathcal{X}$ and $\mathcal{Y}$ are normed algebras, then they are said to be *isometrically isomorphic* if there exists an isomorphism $\varphi : \mathcal{V} \to \mathcal{W}$ that is norm-preserving, that is, $\|A\|_{\mathcal{V}} = \|\varphi(A)\|_{\mathcal{W}}$.

A Banach algebra is a *uniform algebra* if it is isometrically isomorphic to a closed subspace of the Banach algebra $\mathcal{C}(K)$ for some space K which is allowed to be a slightly more general space than a metric space (namely a compact Hausdorff space; see for example [Rud86, p. 36]).

Q-algebras are closely related to uniform algebras. Roughly speaking, a Q-algebra is a Banach algebra formed by cosets in a uniform algebra. To define Q-algebras precisely, we need two more definitions, that of an ideal, and that of a quotient algebra.

A subset $\mathcal{I}$ of a commutative complex algebra $\mathcal{X}$ is an *ideal* in $\mathcal{X}$ if $\mathcal{I}$ is a subspace of $\mathcal{X}$ (in the sense of a vector space), and if for every $A \in \mathcal{I}$ and $B \in \mathcal{X}$, we have $A * B \in \mathcal{I}$.

Given an ideal $\mathcal{I}$, we can associate with each $A \in \mathcal{X}$ the *coset* $\varphi(A) = A + \mathcal{I} = \{A + B : B \in \mathcal{I}\}$. A vector space is obtained out of such cosets by defining

$$\begin{aligned} \varphi(A) + \varphi(B) &= \varphi(A + B) \\ \alpha\varphi(A) &= \varphi(\alpha A) \end{aligned}$$

for every complex scalar α. Moreover, if $\mathcal{I}$ is closed and properly contained

in $\mathcal{X}$ (i.e., $\mathcal{I} \neq \mathcal{X}$), then a commutative Banach algebra is obtained by defining

$$\begin{aligned} \varphi(A)\varphi(B) &= \varphi(A * B) \\ \|\varphi(A)\| &= \inf\{\|B\| : B \in \varphi(A)\}. \end{aligned}$$

The Banach algebra obtained this way is called a *quotient algebra* of $\mathcal{X}$, and is denoted by $\mathcal{X}/\mathcal{I}$.

7.1.1. DEFINITION. Let $\mathcal{X}$ be a commutative Banach algebra. Then $\mathcal{X}$ is a *Q-algebra* if there exists a uniform algebra $\mathcal{Y}$ and a closed ideal $\mathcal{I} \subset \mathcal{Y}$ such that $\mathcal{X}$ is isomorphic to the quotient algebra $\mathcal{Y}/\mathcal{I}$.

The most interesting feature of Q-algebras, discovered by Cole (see [Wer69]), is that they are isometrically isomorphic to a closed (commutative) subalgebra of $\mathcal{B}(\mathcal{H})$, the algebra of bounded operators on a Hilbert space (where the multiplication is the regular matrix product). In other words, Q-algebras are commutative operator algebras. In general, the converse is false [Var74], but Tonge [Ton78] showed that it is true for every algebra generated by a set of commuting Hilbert-Schmidt operators (endowed with the matrix product). We refer to the Notes and Remarks section of [DJT95, Chapter 18] for more information on the significance and historical developments of Q-algebras.

7.1.3 Schatten spaces and the Schur product

Varopoulos's question involves Banach algebras formed by Schatten spaces and the Schur product, which we introduce next.

The Spectral Theorem asserts that the Banach space of compact operators on ℓ_2, which we denote by S_∞, consists of the operators A that admit a representation of the form

$$A = \sum_{i=1}^{\infty} \lambda_i \langle \cdot\,, e_i \rangle f_i, \tag{7.1}$$

where $(e_i)_i$ and $(f_i)_i$ are orthonormal bases for ℓ_2 and the sequence $(\lambda_i)_i \subset \mathbb{R}$ satisfies $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ and $\lim_{i\to\infty} \lambda_i = 0$ (see for example [RS72, Theorem VI.17]). The space S_∞ is endowed with the operator norm $\|A\| = \sup\{|\langle x, Ay\rangle| : \|x\|, \|y\| \leq 1\}$.

For $1 \leq p < \infty$, the *Schatten p-norm* of a compact operator A is given by the ℓ_p-norm $(|\lambda_1|^p + |\lambda_2|^p + \cdots)^{1/p}$ of the sequence $(\lambda_i)_i$ appearing in Eq. (7.1). It is a well-known, but nontrivial fact, proved by Schatten and von Neumann [Sch46, SvN46, SvN48], that these functions are indeed norms. The

p-Schatten space $S_p \subseteq S_\infty$ is the normed vector space formed by the set of compact operators that have finite Schatten p-norm, where the norm is the Schatten p-norm. In [SvN48] it was first proved that these spaces are Banach spaces. Much-studied examples of these spaces are the trace class S_1 and the Hilbert-Schmidt operators S_2.

The *Schur product*, for which we henceforth fix the symbol $\circ$, is a commutative multiplication for S_∞ defined as the entry-wise product when the elements of S_∞ are represented by matrices using the canonical basis for ℓ_2. That is, for $A, B \in S_\infty$ such that $A = (A_{ij})_{i,j}$ and $B = (B_{ij})_{i,j}$, we have

$$A \circ B = (A_{ij}B_{ij})_{i,j}.$$

7.2 Varopoulos's question and our part of the answer

Davie [Dav73] and Varopoulos [Var72] proved that the Banach algebra $(\ell_p, \circ)$ is a Q-algebra for all $1 \leq p \leq \infty$. Here the multiplication $\circ$ is the pointwise multiplication. Note that this notation is consistent with the symbol used for the Schur product (defined in the previous section) when we represent an element $x \in \ell_p$ as a linear combination of the canonical basis vectors. This result implies immediately that the algebra of Hilbert-Schmidt operators $(S_2, \circ)$ (with the Schur product) is a Q-algebra. To see this, notice that a matrix can be seen as a vector by simply appending all of its columns underneath each other. The Hilbert-Schmidt norm (also known as the Frobenius norm) of the matrix then coincides with the ℓ_2 norm of that vector, and the Schur product of two matrices corresponds to the pointwise product of their corresponding vectors. In other words, $(S_2, \circ)$ is isometrically isomorphic to $(\ell_2, \circ)$.

Varopoulos's question [Var75] is the natural one following these facts:

Is it true that $(S_p, \circ)$ *is a Q-algebra for* all $1 \leq p \leq \infty$*?*

Progress was made by Le-Merdy [LM98] and Pérez-García [PG06], who proved that the property holds *true* for all $2 \leq p \leq 4$ and $1 \leq p \leq 2$, respectively. Mantero and Tonge [MT80] proved that $(S_\infty, \circ)$ *fails* to be a so-called 1-summing algebra, which requires only slightly stronger conditions than for being a Q-algebra. Nevertheless, our contribution gives a positive result for the high end of the spectrum.

7.2.1. THEOREM. *The Banach algebra $(S_\infty, \circ)$ is a Q-algebra.*

The proof of Theorem 7.2.1 relies on an important result of Davie [Dav73, Theorem 3.3], which gives a simple characterization of Q-algebras . We use a slight reformulation of it, as given in [DJT95, Lemma 18.5 and Proposition 18.6]. For tensor $T : [n]^N \to \mathbb{C}$, define

$$\mathrm{OPT}^{\mathbb{C}}(T) = \sup\left\{ \Big| \sum_{I\in[n]^N} T[I]\xi_1(i_1)\cdots\xi_N(i_N)\Big| : \xi_1,\ldots,\xi_N : [n] \to B_{\mathbb{C}} \right\}.$$

7.2.2. THEOREM (DAVIE). *Let $\mathcal{X} = (\mathcal{V}, *)$ be a commutative Banach algebra. Then $\mathcal{X}$ is a Q-algebra if and only if there exists a universal constant $K > 0$, such that for any choice of positive integers n, N, complex tensor $T : [n]^N \to \mathbb{C}$, and $\mathcal{V}$-valued sequences $A_1, \ldots, A_N : [n] \to B_{\mathcal{V}}$, the inequality*

$$\Big\| \sum_{I\in[n]^N} T[I]A_1(i_1) * \cdots * A_N(i_N)\Big\|_{\mathcal{V}} \le K^N \,\mathrm{OPT}^{\mathbb{C}}(T), \tag{7.2}$$

holds.

We prove that the Banach algebra $(S_\infty, *)$ satisfies Davie's criterion using the multilinear extension of the complex version of Grothendieck's inequality, due to Blei [Ble79] and Tonge [Ton78], which we encountered in a slightly different form Section 2.3.4.

7.2.3. THEOREM (TONGE). *Let $n, N \ge 2$ and d be positive integers. Then, for any tensor $T : [n]^N \to \mathbb{C}$ and functions $f_1, \ldots, f_N : [n] \to B_{\mathbb{C}^d}$, we have*

$$\Big| \sum_{I\in[n]^N} T[I]\big\langle f_1(i_1), \ldots, f_N(i_N)\big\rangle\Big| \le 2^{(N-2)/2} K_G^{\mathbb{C}} \,\mathrm{OPT}^{\mathbb{C}}(T). \tag{7.3}$$

A proof of this theorem can be obtained with some minor modifications of the proof of the variant presented in Section 2.3.4. This theorem was also used by Pérez-García [PG06] to prove that $(S_1, \circ)$ is a Q-algebra.

PROOF OF THEOREM 7.2.1: The case $N = 1$ is trivial and holds for $K = 1$, as for any sequence $A(1), \ldots, A(n) \in B_{S_\infty}$, we have

$$\Big\| \sum_{i=1}^n T[i]A(i)\Big\| = \sup\left\{ \Big| \sum_{i=1}^n T[i]\langle u, A(i)v\rangle\Big| : u, v \in B_{\ell_2} \right\} = \mathrm{OPT}^{\mathbb{C}}(T).$$

From now on, we fix integers $n, N \geq 2$, tensor $T : [n]^N \to \mathbb{C}$ and operator-valued maps $A_1, \ldots, A_N : [n] \to B_{S_\infty}$. Define

$$M = \sum_{I \in [n]^N} T[I] A_1(i_1) \circ \cdots \circ A_N(i_N).$$

By Theorem 7.2.2 (Davie's criterion) it suffices to show that the inequality

$$\|M\| \leq K^N \|T\|_\infty, \tag{7.4}$$

holds for some constant K independent of n, N, T and $A_1, \ldots, A_N$.

We begin by making four small preliminary steps to show that without loss of generality we may assume that T is real valued and the A_i are finite-dimensional Hermitian matrices. Afterwards we will be able to apply Theorem 7.2.3 in order to prove Eq. (7.4). In the first step we show that without loss of generality, we may assume that the tensor T is real-valued. To this end, define the real-valued tensors T_R and T_C by $T_R[I] = \Re(T[I])$ and $T_C[I] = \Im(T[I])$ for every $I \in [n]^N$. Define

$$\begin{aligned} M_R &= \sum_{I \in [n]^N} T_R[I] A_1(i_1) \circ \cdots \circ A_N(i_N) \\ M_C &= \sum_{I \in [n]^N} T_C[I] A_1(i_1) \circ \cdots \circ A_N(i_N) \end{aligned}$$

Since $M = M_R + iM_C$, we have $\|M\| \leq 2 \max\{\|M_R\|, \|M_C\|\}$. Proving Eq. (7.4) for real-valued tensors thus suffices.

In the second step we show that it suffices to consider the case where the operators $A_1(i_1), \ldots, A_N(i_N) \in B_{S_\infty}$ are finite-dimensional matrices (in the canonical basis for ℓ_2). Recall that norm of M is given by

$$\|M\| = \sup\{|\langle u, Mv\rangle| : u, v \in B_{\ell_2}\}.$$

For any $u \in \ell_2$ with $\|u\| \leq 1$ and any $\varepsilon > 0$ there exists a $D \in \mathbb{N}$ such that the vector $u' = \sum_{\ell=1}^{D} u_\ell e_\ell$ has norm at least $1 - \varepsilon$. Hence, for any $u, v \in B_{\ell_2}$ and $\varepsilon > 0$ there exist $D \in \mathbb{N}$ and $u', v' \in B_{\ell_2}$ supported only on $e_1, \ldots, e_D$ such that

$$|\langle u, Mv\rangle| \leq |\langle u', Mv'\rangle| + \big(2\varepsilon(1-\varepsilon) + \varepsilon^2\big)|\langle u, Mv\rangle|.$$

It follows that for some $D \in \mathbb{N}$ and vectors $u', v' \in B_{\ell_2}$ supported only on $e_1, \ldots, e_D$, we have

$$\|M\| \leq 2|\langle u', Mv'\rangle|. \tag{7.5}$$

Define for every $k = 1, \ldots, N$ and $i_k = 1, \ldots, n$ the D-by-D complex matrix $A'_k(i_k) = (\langle e_\ell, A_k(i_k)e_m\rangle)_{\ell,m=1}^D$. Note that $\|A'_k(i_k)\| \leq \|A_k(i_k)\| \leq 1$. Expanding the definition of M then gives

$$\begin{aligned}\langle u', Mv'\rangle = \Big\langle u', \sum_{I\in[n]^N} T[I]A_1(i_1)\circ\cdots\circ A_N(i_N)v'\Big\rangle = \\ \sum_{I\in[n]^N} T[I]\langle u', A_1(i_1)\circ\cdots\circ A_N(i_N)v'\rangle = \\ \sum_{I\in[n]^N} T[I]\langle u', A'_1(i_1)\circ\cdots\circ A'_N(i_N)v'\rangle. \quad (7.6)\end{aligned}$$

Define the complex number $\Theta = \langle u', Mv'\rangle$. Eq. (7.5) shows that to prove the theorem, it suffices to show that the inequality

$$|\Theta| \leq K^N \|T\|_\infty, \quad (7.7)$$

holds for some constant K, and Eq. (7.6) shows that we can write Θ using the matrix-valued maps $A'_1, \ldots, A'_N$.

In the third step we absorb the complex part of the number Θ into the matrix-valued map A'_1. Let us write Θ in polar coordinates as $|\Theta|e^{i\phi}$ for some $\phi \in [0, 2\pi]$. Define $A''_1(i_1) = e^{-i\phi}A'_1(i_1)$. Then by Eq. (7.6), we have

$$\sum_{I\in[n]^N} T[I]\langle u', A''_1(i_1)\circ A'_2(i_2)\circ\cdots\circ A'_N(i_N)v'\rangle = |\Theta|. \quad (7.8)$$

In the fourth step we symmetrize the situation by making the matrices Hermitian. To this end, define the map $\rho : \mathbb{C}^{D\times D} \to \mathbb{C}^{2D\times 2D}$ by

$$\rho(A) = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}.$$

Define matrix-valued maps $B_1, \ldots, B_N : [n] \to \mathbb{C}^{2D\times 2D}$ by

$$\begin{aligned} B_1(i_1) &= \rho\big(A''_1(i_1)\big) \\ B_2(i_2) &= \rho\big(A'_2(i_2)\big) \\ &\vdots \\ B_N(i_N) &= \rho\big(A'_N(i_N)\big). \end{aligned}$$

Note that $\|B_k(i_k)\| \leq 1$ for all $k = 1, \ldots, N$ and $i_k = 1, \ldots, n$, since the map ρ leaves the norm unchanged. Define the matrices

$$\begin{aligned} M' &= \sum_{I\in[n]^N} T[I]A''_1(i_1)\circ A'_2(i_2)\circ\cdots\circ A'_N(i_N) \\ M'' &= \sum_{I\in[n]^N} T[I]B_1(i_1)\circ B_2(i_2)\circ\cdots\circ B_N(i_N). \end{aligned}$$

Since the tensor T is real-valued we have $M'' = \rho(M')$.

Define the vector $w = (v' \oplus u')/\sqrt{2}$ and note that $\|w\| \leq 1$. We have

$$\begin{aligned}
\langle w, M''w\rangle &= \frac{1}{2}[(u')^*, (v')^*] \begin{bmatrix} 0 & M' \\ (M')^* & 0 \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix} \\
&= \Re\big(\langle u', M'v'\rangle\big) \\
&= \Re\left(\sum_{I\in[n]^N} T[I]\langle u', A_1''(i_1)\circ\cdots\circ A_N'(i_N)v'\rangle\right) \\
&= |\Theta|, \qquad (7.9)
\end{aligned}$$

where the last identity follows from Eq. (7.8), which shows that the term between brackets on the third line is the real number $|\Theta|$.

Next, we absorb the complex parts of the vector w into the matrix-valued map B_1. Using polar coordinates we can write

$$w = \sum_{\ell=1}^{2D} w_\ell e^{i\psi_\ell} e_\ell$$

for some moduli $w_\ell \in \mathbb{R}_+$ and arguments $\psi_\ell \in [0, 2\pi]$. Let $U \in \mathbb{C}^{D\times D}$ be the diagonal unitary matrix given by $U = \operatorname{diag}(e^{i\psi_1}, \ldots, e^{i\psi_D})$. Define the non-negative real vector $w' = U^*w = \sum_{\ell=1}^{2D} w_\ell e_\ell$ and define the matrix-valued map B_1' by $B_1'(i_1) = U^*B_1(i_1)U$. Note that $\|B_1'(i_1)\| \leq \|B_1(i_1)\| \leq 1$.

Then, by Eq. (7.9) and by expanding the definition of M'' we have

$$\sum_{I\in[n]^N} T[I]\langle w', B_1'(i_1)\circ B_2(i_2)\circ\cdots B_N(i_N)w'\rangle = \langle w, M''w\rangle = |\Theta|. \qquad (7.10)$$

We can now make a connection to Theorem 7.2.3 using the following two claims.

12. CLAIM. *There exist real numbers $\mu_1, \ldots, \mu_{2D} \geq 0$ such that*

$$0 \leq \sum_{\ell,m=1}^{2D} \mu_\ell \mu_m \min\{\ell, m\} \leq 1 \qquad (7.11)$$

and for $1_\ell = e_1 + \cdots + e_\ell$,

$$|\Theta| = \sum_{\ell,m=1}^{2D} \mu_\ell \mu_m \theta_{\ell,m}, \qquad (7.12)$$

where

$$\theta_{\ell,m} = \sum_{I\in[n]^N} T[I]\langle 1_\ell, B_1'(i_1)\circ B_2(i_2)\circ\cdots\circ B_N(i_N)1_m\rangle.$$

PROOF: By relabeling the basis vectors $e_1, \ldots, e_{2D}$ appropriately, we may assume that the coefficients of the above vector w' satisfy $w_1 \geq w_2 \geq \cdots \geq w_{2D}$. Setting $\mu_\ell = (w_\ell - w_{\ell-1})$ for $\ell = 1, \ldots, 2D-1$ and $\mu_{2D} = w_{2D}$ gives

$$w' = \sum_{\ell=1}^{2D} \mu_\ell 1_\ell,$$

since $\langle w', e_k \rangle = \mu_k + \mu_{k+1} + \cdots + \mu_{2D} = w_k$. Eq. (7.11) follows from the fact that $0 \leq \langle w', w' \rangle \leq 1$ and $\langle 1_\ell, 1_m \rangle = \min\{\ell, m\}$, and Eq. (7.12) follows by expanding w' in Eq. (7.10). □

13. CLAIM. *For every $1 \leq \ell, m \leq 2D$, we have*

$$|\theta_{\ell,m}| \leq C_N \min\{\ell, m\} \|T\|_\infty, \tag{7.13}$$

where $C_N = 2^{(N-2)/2} K_G$.

PROOF: Expanding the vectors 1_ℓ in the canonical basis gives

$$\Big\langle 1_\ell, B_1'(i_1) \circ B_2(i_2) \circ \cdots \circ B_N(i_N) 1_m \Big\rangle = \sum_{s=1}^{\ell} \sum_{t=1}^{m} \Big\langle e_s, B_1'(i_1) \circ B_2(i_2) \circ \cdots \circ B_N(i_N) e_t \Big\rangle. \tag{7.14}$$

Note that each term in the double sum on the right-hand side of Eq. (7.14) is simply the product of (s,t)-entries of the matrices $B_1'(i_1), B_2(i_2), \ldots, B_N(i_N)$.

Suppose that $\ell \leq m$. Since the matrices $B_1'(i_1), B_2(i_2), \ldots, B_N(i_N)$ have norm at most 1, their rows belong to $B_{\ell_2^m}$ (where ℓ_2^m is the set of length-m 2-summable sequences). Hence, the inner sum on the right-hand side of Eq. (7.14),

$$\sum_{t=1}^{m} \Big\langle e_s, B_1'(i_1) \circ B_2(i_2) \cdots \circ B_N(i_N) e_t \Big\rangle = \sum_{t=1}^{m} \langle e_s, B_1'(i_1) e_t \rangle \langle e_s, B_2(i_2) e_t) \rangle \cdots \langle e_s, B_N(i_N) e_t \rangle,$$

is the generalized inner product of a set of N vectors in $B_{\ell_2^m}$. The result for the case $\ell \leq m$ now follows from the triangle inequality and Theorem 7.2.3, as

$$|\theta_{\ell,m}| = \left| \sum_{I \in [n]^N} T[I] \Big\langle 1_\ell, B_1'(i_1) \circ B_2(i_2) \circ \cdots \circ B_N(i_N) 1_m \Big\rangle \right| \leq \sum_{s=1}^{\ell} \left| \sum_{I \in [n]^N} T[I] \sum_{t=1}^{m} \langle e_s, B_1'(i_1) e_t \rangle \langle e_s, B_2(i_2) e_t) \rangle \cdots \langle e_s, B_N(i_N) e_t \rangle \right| \leq \ell 2^{(N-2)/2} K_G \|T\|_\infty.$$

The case $\ell \geq m$ is proved in the same manner. ♦

Putting Claim 12 and Claim 13 together gives

$$\begin{aligned} |\Theta| &= \sum_{\ell,m=1}^{2D} \mu_\ell \mu_m \theta_{\ell,m} \\ &\leq \sum_{\ell,m=1}^{2D} \mu_\ell \mu_m |\theta_{\ell,m}| \\ &\leq C_N \|T\|_\infty \sum_{\ell,m=1}^{2D} \mu_\ell \mu_m \min\{\ell, m\} \\ &\leq C_N \|T\|_\infty. \end{aligned}$$

We conclude that Eq. (7.7) (Davie's criterion) holds for $K \leq 4$. □

7.2.1 The connection to the Schmidt states

For completeness, we now sketch the connection made in [PGWP+08] to the problem of determining whether Schmidt states, which are states of the form $|\psi\rangle = \sum_{\ell=1}^{d} \alpha_\ell |\ell\rangle^{\otimes N}$ for arbitrary real nonnegative coefficients α_i, allow for arbitrary large violation ratios in N-player XOR games.

The starting point is the last line of Eq. (7.6) in the proof above, which is of the form:

$$\sum_{I \in [n]^N} T[I] \Big\langle u, A_1(i_1) \circ \cdots \circ A_N(i_N) v \Big\rangle,$$

for some d-dimensional vectors u, v and matrices $A_\ell(i_\ell)$. Theorem 7.2.1 was proved by showing that the absolute value of this quantity is bounded from above by $K^N \, \mathrm{OPT}^{\mathbb{C}}(T)$ for some universal constant K.

Renaming the basis vectors $e_1, \ldots, e_d$ for $\mathbb{C}^d$ as $|1\rangle, \ldots, |d\rangle$ gives $u = \sum_{\ell=1}^{d} \alpha_\ell |\ell\rangle$ and $v = \sum_{\ell=1}^{d} \beta_\ell |\ell\rangle$. The crucial observation is now that for Schmidt states

$$|\psi\rangle = \sum_{\ell=1}^{d} \alpha_\ell |\ell\rangle^{\otimes N} \qquad \text{and} \qquad |\phi\rangle = \sum_{\ell=1}^{d} \beta_\ell |\ell\rangle^{\otimes N},$$

we have

$$\sum_{I\in[n]^N} T[I]\Big\langle u, A_1(i_1)\circ\cdots\circ A_N(i_N)v\Big\rangle =$$

$$\sum_{I\in[n]^N} T[I]\sum_{\ell,m=1}^{d}\alpha_\ell^*\beta_m\langle\ell|A_1(i_1)\circ\cdots\circ A_N(i_N)|m\rangle =$$

$$\sum_{I\in[n]^N} T[I]\langle\psi|A_1(i_1)\otimes\cdots\otimes A_N(i_N)|\phi\rangle.$$

The form of the last quantity above and the fact that $|\psi\rangle$ and $|\phi\rangle$ are Schmidt states already give a strong indication that it cannot be far from the entangled bias of some N-player XOR game where the players are restricted to sharing a Schmidt state. As shown in [PGWP$^+$08], this is indeed the case. The fact that this bias is at most K^N times larger than the classical bias (for some universal constant K), as stated in Theorem 6.2.1, then implies the required bound, as the classical bias of N-player XOR game (π,Σ) equals $\mathrm{OPT}(\pi\circ\Sigma)\leq\mathrm{OPT}^{\mathbb{C}}(\pi\circ\Sigma)$.

7.3 The intermediate cases

It turns out that once the cases $p=1$ and $p=\infty$ of Varopoulos's question are answered in the positive, the same results for intermediate ones $1<p<\infty$ are obtained essentially for free. The reason for this comes from a pair of very useful results of Pietsch and Triebel [PT68] and Varopoulos [Var72], which give that the Banach algebras $(S_p,\circ)$ can be characterized as algebras "between" $(S_1,\circ)$ and $(S_\infty,\circ)$. What is meant by "between" is that there is a way to obtain the spaces S_p for $1<p<\infty$ by taking certain combinations of S_1 and S_∞. This method is known as the *complex interpolation method*; we refer to Berg and Löfström [BL76] for a detailed account.

We give rough a description of what the complex interpolation method entails in the current setting. Consider the space $\mathcal{F}$ of functions $f:\mathbb{C}\to S_\infty$ that are analytic in the open strip $\{0<\Re(\xi)<1:\ \xi\in\mathbb{C}\}$ and continuous on the closed strip $\{0\leq\Re(\xi)\leq 1:\ \xi\in\mathbb{C}\}$ (additionally, the functions in $\mathcal{F}$ have to approach 0 sufficiently rapidly when their argument moves away from the real line; see [BL76] for details). We endow $\mathcal{F}$ with the norm

$$\|f\|_{\mathcal{F}}=\max\Big\{\sup_{t\in\mathbb{R}}\|f(it)\|_{S_\infty},\sup_{t\in\mathbb{R}}\|f(1+it)\|_{S_1}\Big\}.$$

For $0<\theta<1$, the *interpolation space* $(S_\infty,S_1)_{[\theta]}$ is defined as the subset of elements $A\in S_\infty$ such that $A=f(\theta)$ for some $f\in\mathcal{F}$. The norm on this space

is defined by

$$\|A\|_{[\theta]} = \inf\left\{\|f(\theta)\|_{\mathcal{F}} : f(\theta) = A,\, f \in \mathcal{F}\right\}.$$

Surprisingly, the p-Schatten spaces for the intermediate values $1 < p < \infty$ can be *characterized* in this way.

7.3.1. LEMMA (PIETSCH AND TRIEBEL). *For any $1 < p < \infty$, we have*

$$(S_\infty, S_1)_{[1/p]} = S_p.$$

Varopoulos [Var72] proved that the property of being a Q-algebra is inherited under the complex interpolation method. Specialized to the current setting, his result says the following.

7.3.2. LEMMA (VAROPOULOS). *If $(S_1, \circ)$ and $(S_\infty, \circ)$ are Q-algebras, then for any value $0 < p < 1$, we have that $\big((S_\infty, S_1)_{[1/p]}, \circ\big)$ is a Q-algebra.*

Combining the above two lemmas with the result of Pérez-García [PG06] showing that $(S_1, \circ)$ is a Q-algebra and Theorem 7.2.1 thus gives the following corollary, showing that Varopoulos's question is now completely answered.

7.3.3. COROLLARY. *For any $1 \leq p \leq \infty$, the Banach algebra $(S_p, \circ)$ is a Q-algebra.*

Appendix A

Some useful linear algebra and analysis

In this section, we provide some basic facts and definitions from linear algebra and analysis which are used in this thesis.

A.1 Vector spaces

Euclidean vector spaces Let n be a positive integer. The vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ consist of column vectors of the form

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

where $x_1, \ldots, x_n$ are real or complex scalars, respectively. Addition and multiplication by scalars are defined by

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + x_n \end{pmatrix}, \qquad \alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

Transpose The *transpose* of a vector x in a Euclidean vector space, denoted x^T, is defined to be the row-vector $(x_1, \ldots, x_n)$.

Conjugate transpose The *conjugate transpose* of a vector x in a complex Euclidean vector space, denoted x^*, is defined to be the row-vector $(\bar{x}_1, \ldots, \bar{x}_n)$.

Normed vector spaces A *norm* on a vector space $\mathcal{V}$ is a function $\| \ \| : \mathcal{V} \to \mathbb{R}$ which satisfies for every $x, y \in \mathcal{V}$ and scalar α,

1. $\|\alpha x\| = |\alpha|\|x\|$
2. $\|x\| = 0$ if and only $x = 0$
3. $\|x + y\| \leq \|x\| + \|y\|$

The last property is referred to as the *triangle inequality*. A vector space endowed with a norm is a *normed vector space*.

The 2-norm on a Euclidean vector space is defined by

$$\|x\|_2 = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}.$$

Inner product spaces An *inner product* on a complex vector space $\mathcal{V}$ is a map of the form $\langle \ , \ \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ which satisfies for $x, y, z \in \mathcal{V}$ and scalar α,

1. $\langle x, y \rangle = \overline{\langle y, x \rangle}$
2. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
3. $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$
4. $\langle x, x \rangle \geq 0$
5. $\langle x, x \rangle = 0$ if and only $x = 0$

A vector space endowed with an inner product is an *inner product space*.

The *Euclidean inner product* on $\mathbb{R}^n$ is by $x \cdot y = x_1 y_1 + \cdots + x_n y_n$. Using the transpose, this can also be denoted as $x^T y$.

The Euclidean inner product on $\mathbb{C}^n$ is defined by $\langle x, y \rangle = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n$. Using the conjugate transpose, this can also be written as $x^* y$.

Metric spaces For a vector space $\mathcal{V}$ a *metric* is a function $d : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ which satisfies for any $x, y, z \in \mathcal{V}$,

1. $d(x, y) \geq 0$
2. $d(x, y) = 0$ if and only if $x = y$
3. $d(x, z) \leq d(x, y) + d(y, z)$

The last property is also referred to as the *triangle inequality*. A vector space endowed with a metric is a *metric space*.

Hilbert spaces Let $\mathcal{H}$ be an inner product space. We can make $\mathcal{H}$ into a normed vector space by endowing it with the norm $\|x\| = \sqrt{\langle x, x\rangle}$. We can make $\mathcal{H}$ a metric space by endowing it with the metric $d(x, y) = \|x - y\|$. A sequence $(x_i)_{i=1}^{\infty} \subseteq \mathcal{H}$ is a *Cauchy sequence* if for any $\varepsilon > 0$ there is an integer N such that $d(x_i, x_j) \leq \varepsilon$ for all $i, j > N$. Then, we have that $\mathcal{H}$ is a *Hilbert space* if every Cauchy sequence converges to an element of $\mathcal{H}$ (i.e., if $\mathcal{H}$ is *complete*).

The Euclidean spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ are Hilbert spaces when endowed with the Euclidean inner product. The Hilbert space $L^2([-1, 1])$ consists of the functions $f : [-1, 1] \to \mathbb{R}$ with finite norm, where the inner product is defined by

$$(f, g) = \int_{-1}^{1} f(t)g(t)dt.$$

Cauchy-Schwarz inequality For Hilbert space $\mathcal{H}$, the *Cauchy-Schwarz inequality* states that for any $x, y \in \mathcal{H}$, we have $|\langle x, y\rangle| \leq \|x\|\|y\|$.

Continuous functions on metric spaces Let $\mathcal{X}, \mathcal{Y}$ be metric spaces. A function $f : \mathcal{X} \to \mathcal{Y}$ is *continuous* if for any $\varepsilon > 0$ there is a $\delta > 0$, such that for any $x, y \in \mathcal{X}$ satisfying $d_{\mathcal{X}}(x, y) < \delta$, we have $d_{\mathcal{Y}}\big(f(x), f(y)\big) < \varepsilon$.

A.2 Matrices

Transpose The *transpose* of a complex matrix $A \in \mathbb{C}^{n\times m}$ is the complex matrix $A^T \in \mathbb{C}^{m\times n}$ defined by $(A^T)_{ij} = A_{ji}$.

Conjugate transpose The *conjugate transpose* of a complex matrix $A \in \mathbb{C}^{m\times n}$, denoted A^*, is the complex n-by-m matrix defined by $(A^*)_{ij} = A^*_{ji}$.

Trace The *trace function* $\mathrm{Tr} : \mathbb{C}^{n\times n} \to \mathbb{C}$ is defined by $\mathrm{Tr}(A) = A_{11} + \cdots + A_{nn}$.

Trace inner product The *trace inner product* (also known as the Hilbert-Schmidt inner product) is an inner product on the vector space of matrices $\mathbb{C}^{n\times n}$ defined by $\langle A, B\rangle = \mathrm{Tr}(A^*B)$. Endowed with this inner product, $\mathbb{C}^{n\times n}$ forms an n^2-dimensional Hilbert space.

Rank The *rank* of a matrix is defined to be its largest number of linearly independent columns.

Outer product The *outer product* of two vectors $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^m$ is the matrix $xy^* \in \mathbb{C}^{n \times m}$ given by $(xy^*)_{ij} = x_i \bar{y}_j$.

Hermitian matrix A complex matrix $A \in \mathbb{C}^{n \times n}$ is *Hermitian* if $A^* = A$.

Unitary matrices A complex matrix $U \in \mathbb{C}^{n \times n}$ is *unitary* if it satisfies

$$U^*U = I.$$

Unitary matrices have the property that they preserve inner products between vectors. In fact, this property is equivalent to being unitary. For any pair of vectors $x, y \in \mathbb{C}^n$, we have $\langle Ux, Uy \rangle = \langle x, y \rangle$. It follows that unitary matrices are also norm-preserving: $\|Ux\| = \|x\|$.

Positive semidefinite matrices A complex Hermitian matrix $A \in \mathbb{C}^{n \times n}$ is *positive semidefinite* if one of the following holds.

1. The matrix A has only real nonnegative eigenvalues.
2. There exist a complex n-dimensional vectors $z_1, \ldots, z_n$ such that for every $i, j \in \{1, \ldots, n\}$, we have $A_{ij} = z_i \cdot z_j$.
3. For any vector $z \in \mathbb{C}^n$, we have $z^* A z \geq 0$.
4. There exists a complex matrix B such that $A = B^* B$.

In fact, Items 1-4 are equivalent (see for example [Bha07]). The factorization given in item 2 is called the *Gram decomposition* of A.

The set of positive semidefinite matrices forms a *convex cone*, meaning that for any n-by-n positive semidefinite matrices A, B and nonnegative scalars $\alpha, \beta \in \mathbb{R}_+$, we have that the matrix $\alpha A + \beta B$ is also positive semidefinite. Sometimes the notation $A \succeq 0$ will be used to denote that A is positive semidefinite.

A positive semidefinite matrix A satisfying $A^2 = A$ is an *orthogonal projector*. An orthogonal projector corresponds to a subspace of $\mathbb{C}^n$ defined by the space spanned by its nonzero eigenvectors.

In the case of real matrices, we have the following analogous characterization of positive semidefinite matrices. A real symmetric matrix $A \in \mathbb{R}^{n \times n}$ is *positive semidefinite* if one of the following holds.

1. The matrix A has only real nonnegative eigenvalues.
2. There exist a real n-dimensional vectors $z_1, \ldots, z_n$ such that for every $i, j \in \{1, \ldots, n\}$, we have $A_{ij} = z_i \cdot z_j$.

3. For any vector $z \in \mathbb{R}^n$, we have $z^T A z \geq 0$.
4. There exists a real matrix B such that $A = B^T B$.

We denote the cone of real n-by-n positive semidefinite matrices by $\mathcal{S}_n^+$.

The rank of a positive semidefinite matrix equals the smallest positive integer d such that there exists a Gram decomposition of it in $\mathbb{R}^d$.

Laplacian matrices Let $G = (V, E)$ be a graph with finite vertex set V and edge set $E \subseteq \binom{V}{2}$. Then, the *Laplacian matrix* of G is the matrix $A : V \times V \to \mathbb{R}$ (this matrix has rows and columns indexed by the vertices of V) defined by

$$A(u,v) = \begin{cases} \deg(u) & \text{if } v = u \\ -1 & \text{if } \{u,v\} \in E \\ 0 & \text{otherwise,} \end{cases}$$

where $\deg(u) = |\{v \in V : \{u,v\} \in E\}|$ denotes the *degree* of vertex u.

The Laplacian matrix of a graph is always a positive semidefinite matrix. To see this, let $G = (V, E)$ be some graph and let us define for each edge $\{u, v\}$ in the graph the vector $x_{uv} = e_u - e_v$, where the e_u are the $|V|$-dimensional canonical unit vectors and the choice of which of the two unit vectors in x_{uv} is subtracted from the other is arbitrary. Then, we have that the matrix

$$A = \sum_{\{u,v\} \in E} x_{uv} x_{uv}^T$$

satisfies

$$\begin{aligned} A(u,v) &= e_u^T \sum_{\{u',v'\} \in E} (e_{u'} - e_{v'})(e_{u'} - e_{v'})^T e_v \\ &= \begin{cases} \deg(u) & \text{if } v = u \\ -1 & \text{if } \{u,v\} \in E \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Hence, A is the Laplacian matrix of G. This matrix is positive semidefinite because it is a positive linear combination of the rank-1 positive semidefinite matrices $x_{uv} x_{uv}^T$.

A.3 Tensor products

If $\mathcal{X} = \mathbb{C}^{n_1 \times m_1}$ and $\mathcal{Y} = \mathbb{C}^{n_2 \times m_2}$ then the tensor product of the vector spaces $\mathcal{X}$ and $\mathcal{Y}$ is defined as $\mathcal{X} \otimes \mathcal{Y} = \mathbb{C}^{n_1 n_2 \times m_1 m_2}$.

To define the tensor product of complex matrices it is convenient to index the rows and columns of a matrix by sets $\mathcal{R}$ and $\mathcal{C}$, respectively, and view the matrix as a map from $\mathcal{R} \times \mathcal{C}$ to $\mathbb{C}$. An n-by-m matrix A is thus viewed as a map $A : \{1, \ldots, n\} \times \{1, \ldots, m\} \to \mathbb{C}$ and its (i, j)-entry is written as $A(i, j)$

Let $\mathcal{R}_1, \mathcal{C}_1$ and $\mathcal{R}_2, \mathcal{C}_2$ be sets and let $A : \mathcal{R}_1 \times \mathcal{C}_1 \to \mathbb{C}$ and $B : \mathcal{R}_2 \times \mathcal{C}_2 \to \mathbb{C}$ be complex matrices. Then, their *tensor product* is the matrix

$$A \otimes B : (\mathcal{R}_1 \times \mathcal{R}_2) \times (\mathcal{C}_1 \times \mathcal{C}_2) \to \mathbb{C}$$

is defined by

$$(A \otimes B)\big((r_1, r_2), (c_1, c_2)\big) = A(r_1, c_1)B(r_2, c_2).$$

It follows easily that the tensor product satisfies for any matrices A, B, C, D:

$$\begin{aligned} (A \otimes B) \otimes C &= A \otimes (B \otimes C) \\ A \otimes (B + C) &= A \otimes B + A \otimes C \\ (A \otimes B)(C \otimes D) &= (AC) \otimes (BD), \end{aligned}$$

where for the last identity we assumed that A and C have equal size and that B and D have equal size.

We also have for $x_1, y_1 \in \mathbb{C}^n$ and $x_2, y_n \in \mathbb{C}^m$, the easy identity

$$\langle x_1 \otimes x_2, y_1 \otimes y_2 \rangle = \langle x_1, y_1 \rangle \langle x_2, y_2 \rangle.$$

A.4 Dirac notation

Dirac notation refers to a notational convention used for the Hilbert space $\mathbb{C}^n$ in the context of quantum information theory. Vectors are usually denoted by a Greek symbol or a non-negative integer wedged between a "|" and a "⟩". We thus write for example $|\psi\rangle \in \mathbb{C}^n$ or $|1\rangle \in \mathbb{C}^n$. The non-negative integers are reserved for the canonical basis vectors, that is

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \qquad \ldots, \qquad |n-1\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

The conjugate transpose of a vector $|\psi\rangle \in \mathbb{C}^n$ is denoted by $\langle\psi|$. Usually the tensor product symbol is omitted when we take the tensor product of two vectors $|\psi\rangle$ and $|\phi\rangle$. So $|\psi\rangle \otimes |\phi\rangle$ is abbreviated to $|\psi\rangle|\phi\rangle$.

Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$ be matrices and let $|\psi\rangle \in \mathbb{C}^n$ and $|\phi\rangle \in \mathbb{C}^m$ be vectors. It follows easily from the properties of the tensor product that

$$\langle\psi|\langle\phi|A \otimes B|\psi\rangle|\phi\rangle = \langle\psi|A|\psi\rangle\langle\phi|B|\phi\rangle.$$

Bibliography

[AAR99] G. Andrews, R. Askey, and R. Roy. Special functions. volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1999.

[ABH+05] S. Arora, E. Berger, E. Hazan, G. Kindler, and M. Safra. On non-approximability for quadratic programs. In *Proceedings of the 46th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2005)*, pages 206–215. IEEE Computer Society, 2005. doi: 10.1109/SFCS.2005.57.

[ABS10] S. Arora, B. Barak, and D. Steurer. Subexponential algorithms for unique games and related problems. In *Proceedings of the 51st Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2010)*, pages 563–572. IEEE Computer Society, 2010.

[ADR82] A. Aspect, J. Dalibard, and G. Roger. Experimental test of Bell's inequalities using time- varying analyzers. *Phys. Rev. Lett.*, 49(25):1804–1807, 1982. doi:10.1103/PhysRevLett.49.1804.

[AGG05] A. Acín, R. Gill, and N. Gisin. Optimal Bell tests do not require maximally entangled states. *Phys. Rev. Lett.*, 95(21):1079–7114, 2005.

[AGM06] A. Acín, A., N. Gisin, and L. Masanes. From Bell's theorem to secure quantum key distribution. *Phys. Rev. Lett.*, 97(12):120405, 2006. ISSN 1079-7114.

[AGR81] A. Aspect, P. Grangier, and G. Roger. Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.*, 47(7):460–463, 1981. doi:10.1103/PhysRevLett.47.460.

[AGR82] A. Aspect, P. Grangier, and G. Roger. Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities. *Phys. Rev. Lett.*, 49(2):91–94, 1982. doi:10.1103/PhysRevLett.49.91.

[AMMN06] N. Alon, K. Makarychev, Y. Makarychev, and A. Naor. Quadratic forms on graphs. *Invent. Math.*, 163(3), 2006. doi:10.1007/s00222-005-0465-9.

[AN06] N. Alon and A. Naor. Approximating the Cut-Norm via Grothendieck's inequality. *SIAM Journal on Computing*, 35(4):787–803, 2006. doi:10.1137/S0097539704441629.

[AS64] M. Abramowitz and I. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover publications, 1964. ISBN 0486612724.

[AZ05] A. Avidor and U. Zwick. Rounding two and three dimensional solutions of the sdp relaxation of max cut. In C. Chekuri, K. Jansen, J. D. P. Rolim, and L. Trevisan, editors, *Approximation, Randomization and Combinatorial Optimization*, volume 3624 of *Lecture Notes in Computer Science*, pages 608–608. Springer Berlin / Heidelberg, 2005.

[BB84] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, pages 175–179. IEEE, 1984.

[BBC04] N. Bansal, A. Blum, and S. Chawla. Correlation clustering. *Machine Learning, Special Issue on Clustering*, 56:89–113, 2004. doi:10.1023/B:MACH.0000033116.57574.95.

[BBLV09] J. Briët, H. Buhrman, T. Lee, and T. Vidick. Multiplayer XOR games and quantum communication complexity with clique-wise entanglement. 2009. ArXiv:0911.4007.

[BBLV11] J. Briët, H. Buhrman, T. Lee, and T. Vidick. All Schatten spaces endowed with the Schur product are Q-algebras. 2011. Submitted.

[BBT08] N. Bansal, S. Bravyi, and B. Terhal. Classical approximation schemes for the ground-state energy of quantum and classical ising spin hamiltonians on planar graphs, 2008. Available at arXiv:0705.1115v4.

[BBT11] J. Briët, H. Buhrman, and B. Toner. A generalized grothendieck inequality and nonlocal correlations that require high entanglement. *Comm. Math. Phys.*, 305(3):1–17, August 2011. ISSN 0010-3616. 10.1007/s00220-011-1280-3.

[BCD01] H. Burhman, R. Cleve, and W. v. Dam. Quantum entanglement and communication complexity. *SIAM Journal on Computing*, 30(6):1829–1841, 2001.

[Bel64] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[BFG06] S. Bravyi, D. Fattal, and D. Gottesman. GHZ extraction yield for multipartite stabilizer states. *J. Math. Phys.*, 47(062106), 2006.

[BGJR88] F. Barahona, M. Grötschel, M. Jünger, and G. Reinelt. An application of combinatorial optimization to statistical physics and circuit layout design. *Operations Research*, 36(3):493–513, 1988. ISSN 0030-364X.

[Bha07] R. Bhatia. *Positive definite matrices*. Princeton University Press, Princeton, New Jersey, USA, 2007.

[BHH+08] B. Barak, M. Hardt, I. Haviv, A. Rao, O. Regev, and D. Steurer. Rounding parallel repetitions of unique games. In *Proceedings of the 49th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2008)*, pages 374–383. IEEE Computer Society, 2008.

[BHK05] J. Barrett, L. Hardy, and A. Kent. No signaling and quantum key distribution. *Phys. Rev. Lett.*, 95:010503, 2005. Quant-ph/0405101.

[BKCD02] R. Blume-Kohout, C. Caves, and I. Deutsch. Climbing mount scalable: Physical resource requirements for a scalable quantum computer. *Foundations of Physics*, 32(11):1641–1670, 2002. ISSN 0015-9018.

[BL76] J. Bergh and J. Löfström. *Interpolation spaces: An introduction*. Springer-Verlag, 1976.

[Ble79] R. C. Blei. Multidimensional extensions of the Grothendieck inequality and applications. *Arkiv fur Matematik*, 17:51–68, December 1979. doi:10.1007/BF02385457.

[BMMN11] M. Braverman, K. Makarychev, Y. Makarychev, and A. Naor. The Grothendieck constant is strictly smaller than Krivine's bound. In *Proceedings of the 52nd Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2011)*. IEEE Computer Society, 2011. Available at arXiv:1103.6161.

[BOFV10a] J. Briët, F. M. d. Oliveira Filho, and F. Vallentin. The positive semidefinite grothendieck problem with rank constraint. In S. Abramsky, C. Gavoille, C. Kirchner, F. Meyer auf der Heide, and P. Spirakis, editors, *Automata, Languages and Programming*, volume 6198 of *Lecture Notes in Computer Science*, pages 31–42. Springer Berlin / Heidelberg, 2010.

[BOFV10b] J. Briët, F. M. d. Olivera-Filho, and F. Vallentin. Grothendieck inequalities for semidefinite programs with rank constraint, November 2010. Submitted. Available at arXiv:1011.1754.

[BPA+08] N. Brunner, S. Pironio, A. Acin, N. Gisin, A. A. Méthot, and V. Scarani. Testing the dimension of Hilbert spaces. *Phys. Rev. Lett.*, 100(21):210503, 2008.

[BPD+99] D. Bouwmeester, J. Pan, M. Daniell, H. Weinfurter, and A. Zeilinger. Observation of three-photon Greenberger-Horne-Zeilinger entanglement. *Phys. Rev. Lett.*, 82(7):1345–1349, 1999. ISSN 1079-7114.

[BRSW10] H. Buhrman, O. Regev, G. Scarpa, and R. d. Wolf. Near-Optimal and Explicit Bell Inequality Violations. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity (CCC 2011)*. IEEE Computer Society, 2010.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004. Available online at http://www.stanford.edu/~boyd/cvxbook/.

[Car80] T. K. Carne. Banach Lattices and Extensions of Grothendieck's Inequality. *J. London Math. Soc.*, s2-21(3):496–516, 1980. doi: 10.1112/jlms/s2-21.3.496.

[CB97] R. Cleve and H. Buhrman. Substituting quantum entanglement for communication. *Phys. Rev. A*, 56(2):1201, 1997.

[CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, 1969.

[CHTW04] R. Cleve, P. Høyer, B. Toner, and J. Watrous. Consequences and limits of nonlocal strategies. In *Proceedings of the 19th Annual IEEE Conference on Computational Complexity (CCC 2004)*, pages 236–249. IEEE Computer Society, Amherst, Massachusetts, 2004. ArXiv: 0404076 [quant-ph].

[CSUU08] R. Cleve, W. Slofstra, F. Unger, and S. Upadhyay. Perfect parallel repetition theorem for quantum XOR proof systems. *Computational Complexity*, 17(2):282–299, 2008.

[CW04] M. Charikar and A. Wirth. Maximizing quadratic programs: extending Grothendieck's inequality. In *Proceedings of the 45th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2004)*, pages 54–60. IEEE Computer Society Press, 2004.

[Dav73] A. Davie. Quotient algebras of uniform algebras. *J. London Math. Soc.*, 7:31–40, 1973.

[Dav84] A. M. Davie. Lower bound for K_G, 1984. Unpublished note.

[Dav85] A. M. Davie. Matrix norms related to grothendieck's inequality. *Banach Spaces*, pages 22–26, 1985.

[DH03] W. v. Dam and P. Hayden. Universal entanglement transformations without communication. *Phys. Rev. A*, 67:060302(R), 2003. doi:doi:10.1103/PhysRevA.67.060302.

[DJT95] J. Diestel, H. Jarchow, and A. Tonge. *Absolutely summing operators*. Number 43 in Cambridge Studies in Advanced Mathematics. Cambrige University Press, New York, NY, 1995. ISBN 0521431689.

[Eke91] A. K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67(6):661–663, 1991. doi: 10.1103/PhysRevLett.67.661.

[EMOT54] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. Tricomi. *Tables of integral transforms, Vol. 1*. McGraw-Hill, New York, 1954.

[EPR35] A. Einstein, P. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, 1935.

[Fei91] U. Feige. On the success probability of the two provers in one-round proof systems. In *Proceedings of the Sixth Annual IEEE Structure in Complexity Theory Conference*, pages 116–123. IEEE Computer Society, 1991. ISBN 0818622555.

[FS02] U. Feige and G. Schechtman. On the optimality of the random hyperplane rounding technique for MAX CUT. *Random Struct. Algorithms*, 20(3):403–440, May 2002.

[GHZ89] D. M. Greenberger, M. A. Horne, and A. Zeilinger. Going beyond Bell's theorem. *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, pages 73–76, 1989. Available at arXiv:0712.0921 [quant-ph].

[GJ76] M. R. Garey and L. Johnson. Some simplified NP-complete graph problems. *Theoretical computer science*, 1(3):237–267, 1976.

[GKP94] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, second edition, 1994.

[GLS93] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, second edition, 1993. ISBN 3-540-56740-2.

[Got97] D. Gottesman. *Stabilizer codes and quantum error correction*. Ph.D. thesis, California Institute of Technology, 1997. Available at arXiv:9705052 [quant-ph].

[Gro53] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques (French). *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953.

[GTHB05] O. Gühne, G. Tóth, P. Hyllus, and H. J. Briegel. Bell inequalities for graph states. *Phys. Rev. Lett.*, 95(12):120405, Sep 2005. doi: 10.1103/PhysRevLett.95.120405.

[GW94] M. X. Goemans and D. P. Williamson. .878-approximation algorithms for MAX CUT and MAX 2SAT. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing (STOC 1994)*, pages 422–431. ACM Press, New York, 1994.

[GW95] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, 1995. ISSN 0004-5411.

[Haa87] U. Haagerup. A new upper bound for the complex Grothendieck constant. *Israel Journal of Mathematics*, 60(2):199–224, 1987. ISSN 0021-2172.

[Hås99] J. Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182:105–142, 1999.

[Hås01] J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.

[Hel99] S. Helgason. *The Radon Transform*, volume 5 of *Progress in Mathematics*. Birkhauser, 1999. ISBN 0817641092.

[Hol07] T. Holenstein. Parallel repetition: simplifications and no-signaling case. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing (STOC 2007)*. ACM, 2007.

[HV04] J. Håstad and S. Venkatesh. On the advantage over a random assignment. *Random Structures and Algorithms*, 25(2):117–149, 2004. doi:10.1002/rsa.20031.

[JL01] W. Johnson and J. Lindenstrauss. Basic concepts in the geometry of Banach spaces. *Handbook of the geometry of Banach spaces*, 1:1–84, 2001. ISSN 1874-5849.

[JP11] M. Junge and C. Palazuelos. Large Violation of Bell Inequalities with Low Entanglement. *Comm. Math. Phys.*, pages 1–52, 2011. ISSN 0010-3616. 10.1007/s00220-011-1296-8.

[Kar72] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, New York, 1972.

[Kar96] H. Karloff. How good is the Goemans-Williamson MAX CUT algorithm? In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC 1996)*, pages 427–434. ACM, 1996. ISBN 0897917855.

[Kar05] G. Karakostas. A better approximation ratio for the Vertex Cover problem. In *Proceedings of the 32nd International Conference on Automata, Languages, and Programming (ICALP 2005)*, volume 3580 of *Lecture Notes in Computer Science*, pages 1043–1050. Springer, 2005. doi:10.1007/11523468_84.

[Kho02] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC 2002)*, pages 767–775. ACM Press, 2002.

[Kho10] S. Khot. Inapproximability of np-complete problems, discrete fourier analysis, and geometry. In *Proc. the International Congress of Mathematicians*. 2010.

[KKMO04] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other two-variable CSPs? In *Proceedings of the 45th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2004)*, pages 146–154. IEEE Computer Society, 2004.

[KN08] S. Khot and A. Naor. Linear equations modulo two and the L_1 diameter of convex bodies. *SIAM Journal on Computing*, 38(4):1448–1463, 2008.

[KN09] S. Khot and A. Naor. Approximate kernel clustering. *Mathematika*, 55(1-2):129–165, 2009. ISSN 0025-5793.

[KN10] S. Khot and A. Naor. Sharp kernel clustering algorithms and their associated grothendieck inequalities. In *Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 664–683. Society for Industrial and Applied Mathematics, 2010.

[KN11] S. Khot and A. Naor. Grothendieck-type inequalities in combinatorial optimization, August 2011. Available at arXiv: 1108.2464.

[KNS10] G. Kindler, A. Naor, and G. Schechtman. The UGC hardness threshold of the Lp Grothendieck problem. *Mathematics of Operations Research*, 35(2):267–283, 2010. ISSN 0364-765X.

[KO06] S. Khot and R. O'Donnell. SDP gaps and UGC-hardness for MAX-CUTGAIN. In *Proceedings of the 47th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2006)*, pages 217–226. IEEE Computer Society, 2006.

[KO08] S. Khot and R. O'Donnell. SDP gaps and UGC-hardness for Max-Cut-Gain. *Theory of Computing*, 0:1–35, 2008.

[Kön91] H. König. On the complex Grothendieck constant in the n-dimensional case. In P. F. X. Müller and W. Schachermayer, editors, *Geometry of Banach spaces: proceedings of the conference held in Strobl, Austria, 1989*, pages 181–198. Cambridge University Press, 1991. ISBN 0521408504.

[Kön01] H. König. On an extremal problem originating in questions of unconditional convergence. In W. Haussmann, K. Jetter, and M. Reimer, editors, *Recent progress in multivariate approximation: 4th international conference, Witten-Bommerholtz (Germany), September 2000*, volume 137 of *International series of numerical mathematics*, pages 185–192. Birkhauser, 2001. ISBN 3764365056.

[KR08] S. Khot and O. Regev. Vertex cover might be hard to approximate to within $2-\epsilon$. *Journal of Computer and System Sciences*, 74:335–349, 2008. Preliminary version in *Proc. CCC'03*.

[KR10] J. Kempe and O. Regev. No strong parallel repetition with entangled and non-signaling provers. In *Proceedings of the 25th Annual*

IEEE Conference on Computational Complexity (CCC 2010), pages 7–15. IEEE, 2010. ISSN 1093-0159.

[KR11] B. Klartag and O. Regev. Quantum One-Way Communication is Exponentially Stronger Than Classical Communication. In *Proceedings of the 43th Annual ACM Symposium on Theory of Computing (STOC 2011)*. 2011. Available at arXiv:1009.3640.

[Kri77] J. L. Krivine. Sur la complexification des opérateurs de L^∞ dans L^1. *C. R. Acad. Sci., Ser. A*, 284:377–379, 1977.

[Kri79] J. L. Krivine. Constantes de Grothendieck et fonctions de type positif sur les sphères. *Adv. Math.*, 31:16–30, 1979.

[KRT08] J. Kempe, O. Regev, and B. Toner. Unique games with entangled provers are easy. In *Proceedings of the 49th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2008)*, pages 457–466. IEEE Computer Society, 2008. ISSN 0272-5428.

[KV11] J. Kempe and T. Vidick. Parallel Repetition of Entangled Games. In *Proceedings of the 43th Annual ACM Symposium on Theory of Computing (STOC 2011)*. 2011. Available at arXiv:1012.4728.

[KVR90] D. Knuth, A. Vardi, and R. Richberg. The asymptotic expansion of the middle binomial coefficient. *The American Mathematica Monthly*, 97(7):626–630, August–Septmber 1990.

[Lit30] J. E. Littlewood. On bounded bilinear forms in an infinite number of variables. *Q J Math*, os-1(1):164–174, 1930. doi: 10.1093/qmath/os-1.1.164.

[LM98] C. Le Merdy. The Schatten space S_4 is a Q-algebra. *Proc. Amer. Math. Soc.*, 126:715–719, 1998.

[LO94] R. Latala and K. Oleszkiewicz. On the best constant in the Khintchine-Kahane inequality. *Studia Math*, 109(1):101–104, 1994.

[Lov79] L. Lovász. On the Shannon capacity of a graph. *Information Theory, IEEE Transactions on*, 25(1):1–7, 1979.

[LP68] J. Lindenstrauss and A. Pełczyński. Absolutely summing operators in L_p-spaces and their applications. *Studia Math.*, 29:275–326, 1968. ISSN 0039-3223.

[LR05] M. Laurent and F. Rendl. Semidefinite Programming and Integer Programming. In K. Aardal, G. Nemhauser, and R. Weismantel, editors, *Handbook on Discrete Optimization*, pages 393–514. Elsevier, December 2005.

[LS09] N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures and Algorithms*, 34:368–394, 2009.

[LSS09] T. Lee, G. Schechtman, and A. Shraibman. Lower bounds on quantum multiparty communication complexity. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity (CCC 2009)*, pages 254–262. IEEE Computer Society, 2009.

[LV11] M. Laurent and A. Varvistiotis. Computing the Grothendieck constant for some graphs, 2011.

[Mat76] T. Matsunawa. Some inequalities based on inverse factorial series. *Annals of the Institute of Statistical Mathematics*, 28(1):291–305, 1976. ISSN 0020-3157.

[Mat99] P. Mattila. *Geometry of sets and measures in Euclidean spaces: Fractals and rectifiability*. Cambridge University Press, 1999.

[Meg01] A. Megretski. Relaxations of quadratic programs in operator theory and system analysis. In *Systems, Approximation, Singular Integral Operators, and Related Topics: International Workshop on Operator Theory and Applications*. Birkhauser, 2001.

[Mer90] N. D. Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65(15):1838–1840, Oct 1990.

[MF53] P. M. Morse and H. Feshbach. *Methods of Theoretical Physics, Part I*. McGraw-Hill, New York, 1953.

[MR95] S. Mahajan and H. Ramesh. Derandomizing semidefinite programming based approximation algorithms. In *Proceedings of the 36th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 1995)*, page 162. Published by the IEEE Computer Society, 1995.

[MS86] V. D. Milman and G. Schechtman. Asymptotic theory of finite dimensional normed spaces. volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, 1986.

[MS07] A. A. Methot and V. Scarani. An anomaly of non-locality. *QIC*, 7:157–170, January 2007.

[MST99] G. A. Munoz, Y. Sarantopoulos, and A. Tonge. Complexifications of real Banach spaces, polynomials and multilinear maps. *Studia Math.*, 134:1–33, 1999.

[MT80] A. Mantero and A. Tonge. The Schur multiplication in tensor algebras. *Studia Math*, 68(1):1–24, 1980.

[Mui82] R. Muirhead. *Aspects of multivariate statistical theory*. Wiley series in probability and mathematical statistics. Wiley, 1982.

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, New York, 2000.

[Nes97] Y. Nesterov. Quality of semidefinite relaxation for nonconvex quadratic optimization. Technical report, CORE, Université catholique de Louvain, 1997.

[Nes98] Y. E. Nesterov. Semidefinite relaxations and nonconvex quadratic optimization. *Optimization Methods and Software*, (9):141–160, 1998.

[Nes05] M. v. Nest. *Local equivalence of stabilizer states and codes*. Ph.D. thesis, Katholieke Universiteit Leuven, 2005.

[NRT99] A. Nemirovski, C. Roos, and T. Terlaky. On maximization of quadratic form over intersection of ellipsoids with common center. *Math. Program.*, 86(3):463–473, 1999. doi: 10.1007/s101070050100.

[OF09] F. M. d. Oliveira Filho. *New bounds for geometric packing and coloring via harmonic analysis and optimization*. Ph.D. thesis, University of Amsterdam, 2009.

[Oli10] M. d. O. Oliveira. Embezzlement States are Universal for Non-Local Strategies. 2010. ArXiv:1009.0771.

[PBD+00] J. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger. Experimental test of quantum nonlocality in three-photon Greenberger–Horne–Zeilinger entanglement. *Nature*, 403(6769):515–519, 2000. ISSN 0028-0836.

[PG06] D. Pérez-Garcıa. The trace class is a q-algebra. *Annales Academiæ Scientiarum Fennicæ*, 31:287–295, 2006.

[PGWP+08] D. Perez-Garcia, M. Wolf, C. Palazuelos, I. Villanueva, and M. Junge. Unbounded violation of tripartite Bell inequalities. *Comm. Math. Phys.*, 279:455, 2008.

[Pie72] A. Pietsch. *Nuclear locally convex spaces*. Springer, Berlin, 1972. ISBN 3540056440. (Translated from German by W. H. Ruckle).

[Pis99] G. Pisier. *The volume of convex bodies and Banach space geometry*. Cambridge University Press, 1999. ISBN 052166635X.

[Pis11] G. Pisier. Grothendieck's Theorem, past and present. *Arxiv preprint arXiv:1101.4195*, 2011.

[PT68] A. Pietsch and H. Triebel. Interpolationstheorie für Banachideale von beschränkten linearen Operatoren. *Studia Math.*, 31:95–109, 1968.

[PV08] K. Pál and T. Vértesi. Efficiency of higher-dimensional Hilbert spaces for the violation of Bell inequalities. *Phys. Rev. A*, 77(4), 2008.

[Rag08] P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proc. 40th STOC*, pages 245–254. ACM Press, 2008.

[Raz98] R. Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27:763–803, 1998.

[Raz08] R. Raz. A counterexample to strong parallel repetition. In *Proceedings of the 49th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2008)*, pages 369–373. IEEE, 2008.

[Ree91] J. A. Reeds. A new lower bound on the real Grothendieck constant, 1991. Unpublished note, available at http://www.dtc.umn.edu/ reedsj/bound2.dvi.

[Reg11] O. Regev. Bell Violations through Independent Bases Games, 2011. Arxiv preprint arXiv:1101.0576.

[Rie74] R. Rietz. A proof of the Grothendieck inequality. *Israel Journal of Mathematics*, 19(3):271–276, 1974.

[Rob55] H. Robbins. A remark on Stirling's formula. *Amer. Math. Monthly*, 62(1):26–29, 1955.

[RS72] M. Reed and B. Simon. *Methods of modern mathematical physics*, volume 1. Academic Press, Inc., New York, NY, USA, 1972.

[RS09] P. Raghavendra and D. Steurer. Towards computing the Grothendieck constant. In *Proceedings of the 20st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2009)*, pages 525–534. 2009.

[RT09] O. Regev and B. Toner. Simulating quantum correlations with finite communication. *SIAM J. Comput.*, 39:1562–1580, December 2009. ISSN 0097-5397. doi:http://dx.doi.org/10.1137/080723909.

[Rud86] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, New York, NY, USA, 1986.

[RW08] R. Renner and S. Wolf. Towards characterizing the non-locality of entangled quantum states. *Phys. Rev. Lett*, 101:050403, 2008.

[Sch42] I. J. Schoenberg. Positive definite functions on spheres. *Duke Math. J.*, 9:96–108, 1942. ISSN 0012-7094.

[Sch46] R. Schatten. The cross-space of linear transformations. *The Annals of Mathematics*, 47(1):73–84, 1946. ISSN 0003-486X.

[Sim05] B. Simon. Trace ideals and their applications. volume 120 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, Rhode Island, 2nd edition, 2005.

[Slo10] W. Slofstra. Lower bounds on the entanglement needed to play XOR non-local games, 2010. Arxiv preprint arXiv:1007.2248.

[SSGB07] L. Song, A. Smola, A. Gretton, and K. Borgwardt. A dependence maximization view of clustering. In *Proceedings of the 24th international conference on Machine learning*, pages 815–822. ACM, 2007.

[ST09] A. Samorodnitsky and L. Trevisan. Gowers uniformity, influence of variables, and PCPs. *SIAM Journal on Computing*, 39:323, 2009.

[Sta68] H. E. Stanley. Spherical model as the limit of infinite spin dimensionality. *Physical Review*, 176(2):718–722, 1968.

[SvN46] R. Schatten and J. von Neumann. The cross-space of linear transformations. II. *The Annals of Mathematics*, 47(3):608–630, 1946. ISSN 0003-486X.

[SvN48] R. Schatten and J. von Neumann. The cross-space of linear transformations. iii. *The Annals of Mathematics*, 49(3):557–582, 1948. ISSN 0003-486X.

[Sza76] S. J. Szarek. On the best constants in the Khinchin inequality. *Studia Math.*, 58(2):1s97–208, 1976. ISSN 0039-3223.

[Sze75] G. Szegö. *Orthogonal polynomials*, volume XXIII of *Americal Mathematical Society Colloquim Publications*. American Mathematical Society, Providence, Rhode Island, fourth edition edition, 1975.

[Tal03] M. Talagrand. *Spin glasses: a challenge for mathematicians: cavity and mean field models*. Springer Verlag, 2003. ISBN 3540003568.

[TGB06] G. Tóth, O. Gühne, and H. Briegel. Two-setting Bell inequalities for graph states. *Phys. Rev. A*, 73(2), 2006.

[Ton78] A. Tonge. The von Neumann inequality for polynomials in several Hilbert-Schmidt operators. *J. London Math. (2)*, 18:519–526, 1978.

[Ton09] B. Toner. Monogamy of non-local quantum correlations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 465(2101):59–69, 2009. doi:10.1098/rspa.2008.0149.

[Tsi87] B. S. Tsirelson. Quantum analogues of the Bell inequalities. The case of two spatially separated domains. *J. Soviet Math.*, 36:557–570, 1987.

[TV06] B. Toner and F. Verstraete. Monogamy of Bell correlations and Tsirelson's bound, 2006. Available at arXiv:0611001 [quant-ph].

[Ung09] F. Unger. A probabilistic inequality with applications to threshold direct-product theorems. In *Proceedings of the 50th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2009)*. IEEE, 2009.

[Var72] N. Varopoulos. Some remarks on Q-algebras. *Ann. Inst. Fourier, Grenoble*, 22(2):1–11, 1972.

[Var74] N. Varopoulos. On an inequality of von Neumann and an application of the metric theory of tensor products to operators theory. *Journal of Functional Analysis*, 16(1):83–100, 1974. ISSN 0022-1236.

[Var75] N. T. Varopoulos. A theorem on operator algebras. *Math. Scand.*, 37(1):173–182, 1975.

[VW10] T. Vidick and S. Wehner. More non-locality with less entanglement. 2010. ArXiv:1011.52068.

[Wat08] J. Watrous. Lectures notes on theory of quantum information, 2008.

[WCD08] S. Wehner, M. Christandl, and A. Doherty. Lower bound on the dimension of a quantum system given measured data. *Physical Review A*, 78(6):062112, 2008.

[Wer69] J. Wermer. Quotient algebras of uniform algebras, Symposium on function algebras and rational approximation, University of Michigan, 1969.

[Zuk93] M. Zukowski. Bell theorem involving all settings of measuring apparatus. *Physics Letters A*, 177(4-5):290 – 296, 1993. ISSN 0375-9601. doi:DOI: 10.1016/0375-9601(93)90002-H.

Index

List of symbols

$\mathbb{N}$: The positive integers $1, 2, 3, \ldots$.
$[n]$: The set $\{1, \ldots, n\}$.
$[n]^N$: The cartesian product of $[n]$ with itself N times.
$\mathbb{Z}$: The integers $\ldots, -2, -1, 0, 1, 2, \ldots$.
$\mathbb{R}$: The real numbers.
$\mathbb{R}_+$: The real nonnegative numbers.
$\mathbb{C}$: The complex numbers.

$\ell_2(\mathbb{R})$: The Hilbert space of real square-summable sequences.
$\ell_2(\mathbb{C})$: The Hilbert space of complex square-summable sequences.
S^{n-1}: The real n-dimensional unit sphere.
S^∞: The unit sphere of $\ell_2(\mathbb{R})$.
$S_{\mathbb{C}}^{n-1}$: The complex n-dimensional unit sphere.
$B_{\mathcal{V}}$: The unit ball of normed vector space $\mathcal{V}$.
$\mathcal{S}_n^+$: Cone of n-by-n positive semidefinite matrices
$\mathcal{O}(\mathcal{H})$: The set of $\{-1, 1\}$-valued observables on Hilbert-space $\mathcal{H}$.

ϑ: The Lovász theta number

$\mathrm{SDP}_r(A)$: See Definition 2.1.
$\mathrm{SDP}_r(G, A)$: See Definition 2.3.4.
OPT: See Definition 2.6.
GIP: See Definition 2.7.

K_G: The (real) Grothendieck constant.
$K_G^{\mathbb{C}}$: The complex Grothendieck constant.
$K_G(q \mapsto r)$: See Definition 2.3.1
$K_G^{\succeq}(q \mapsto r)$: See Defintion 2.3.2
$K_G^L(q \mapsto r)$: See Definition 2.3.3
$K(G)$: The Grothendieck constant of graph G (See Section 2.3.2).
$K(r, G)$: The rank-r Grothendieck constant of graph G (See Section 2.3.2).
$K(q \mapsto r, G)$: See Definition 2.3.5.

$\mathcal{G}$: Nonlocal game
$\beta(\pi, \Sigma)$: The classical bias of XOR game $\mathcal{G} = (\pi, \Sigma)$
$\beta^*(\pi, \Sigma)$: The entangled bias of XOR game $\mathcal{G} = (\pi, \Sigma)$
$\beta^*_{|\psi\rangle}(\pi, \Sigma)$: The entangled bias of XOR game $\mathcal{G} = (\pi, \Sigma)$ where the players share state $|\psi\rangle$

$\circ$: The entry-wise multiplication for matrices and tensors.
$\sim$: "Distributed according to"
$\langle x_1, x_2, \cdots, x_N \rangle$: The generalized inner product of $x_1, \ldots, x_N \in \mathbb{C}^d$. See Section 2.3.4.

Samenvatting

Gemotiveerd door toepassingen in de kwantuminformatietheorie en optimalisatie introduceren we nieuwe varianten van de beroemde *Grothendieck ongelijkheid*. In de kwantuminformatietheorie passen we deze wiskundige gereedschappen toe in de studie van de meest verrassende en merkwaardige voorspelling van de kwantummechanica: *verstrengeling*. In optimalisatie gebruiken we ze om de nauwkeurigheid te bepalen van efficiente approximatie algoritmen voor geometrische problemen die op natuurlijk wijze voortkomen uit de studie van verstrengeling en uit modellen voor interacterende deeltjes die beschouwd worden in de klassieke statistische fysica.

In dit proefschrift wordt verstrengeling bestudeert met behulp van *nonlokale spellen*. Een nonlokaal spel wordt gespeeld door twee of meer deelnemers die niet met elkaar mogen communiceren, maar wel in contact staan met een scheidsrechter. Als het spel begint vraagt de scheidsrechter aan elke deelnemer een vraag, waarna ze hem elk een antwoord terugsturen. De scheidsrechter bepaalt vervolgens of de deelnemers winnen of verliezen op basis van enkel de gestelde vragen en verkregen antwoorden. De deelnemers weten van te voren welke antwoorden nodig zijn om het spel te winnen; dat is natuurlijk het doel. Het probleem is dat een deelnemer alleen de vraag kent die direct aan hem gesteld is en niet de vragen die aan de andere deelnemers gesteld zijn. De deelnemers spelen dus niet tegen elkaar, maar moeten juist proberen hun strategieën te coördineren.

In een wereld waar de wetten van de klassieke mechanica gelden is de beste strategie voor een nonlokaal spel altijd de meest voor de hand liggende: bepaal vooraf de antwoorden op alle mogelijke vragen. In een kwantummechanische wereld daarentegen, kunnen meer ingewikkelde strategieën soms een beter

resultaat geven. Elke deelnemer kan zijn antwoord laten afhangen van de uitkomst van een natuurkundig experiment. De onderscheidende eigenschap van een dergelijke handelwijze is dat het de deelnemers kunnen produceren die *gecorreleerd* zijn op een manier die onmogelijk is in een klassieke wereld. In dat geval zijn de deelnemers *verstrengeld*.

Het feit dat de kwantummechanica het bestaan van zo'n fenomeen voorspelt, werd in 1935 door Einstein, Podolski en Rosen gebruikt om te beargumenteren dat deze theorie niet compleet zou kunnen zijn. Volgens hen zou verstrengeling geen deel uit moeten maken van een redelijke beschrijving van de natuur. Verrassend genoeg gaven experimenten van Aspect el al. uit de jaren '80 overtuigend bewijs dat de wereld waarin wij leven wel degelijk zulke effecten toestaat!

Optimalisatie betekent het doorzoeken van een doorgaans grote verzameling met als doel een element met de beste eigenschappen te vinden. Een voorbeeld daarvan is het vinden van een strategie voor een nonlokaal spel waarmee de deelnemers de grootste kans hebben om te winnen. Een ander voorbeeld is het orienteren van de magnetische velden van interacterende deeltjes, zodat de energie van het systeem dat deze vormen minimaal is.

De optimalisatieproblemen die het meest bestudeerd worden zijn van een combinatorisch type. Voorbeelden zijn het vinden van een optimale klassieke strategie voor een nonlokaal spel, maar ook het minimaliseren van de energie van een verzameling deeltjes in het Ising-model uit de klassieke statistische mechanica. Beide vergen een zoektocht over een discrete verzameling mogelijkheden. In dit proefschrift beschouwen we optimalisatieproblemen van een meer *geometrisch* type. Een typisch voorbeeld hiervan is het zoeken van een optimale verdeling van een eindig aantal punten op het oppervlak van een driedimensionale bal. Deze geometrische optimalisatieproblemen vloeien op natuurlijke wijze voort uit de studie van verstrengeling wanneer men de mate van verstrengeling beperkt die gebruikt mag worden door deelnemers van een nonlokaal spel. Deze problemen komen ook voort uit het Heisenberg-model van interacterende deeltjes, welk model ook gebruikt wordt in klassieke statistische fysica.

De meeste van die hiervoor beschreven problemen kunnen waarschijnlijk door geen enkele computer binnen een redelijke hoeveelheid tijd precies worden opgelost. Als tijd een belangrijke rol speelt, dan is het beste alternatief om te zoeken naar een zo goed mogelijke oplossing die snel gevonden kan worden. We gebruiken nieuwe varianten van Grothendiecks ongelijkheid in de

analyse van algoritmen voor de hiervoor beschreven geometrische problemen die precies zo een alternatief bieden.

Abstract

Motivated by applications in quantum information theory and optimization we introduce new variants of a celebrated inequality known as *Grothendieck's Inequality*. In quantum information theory we apply these mathematical tools to study of one of the most surprising and counter-intuitive predictions of Quantum Mechanics: *entanglement*. In optimization we use them to determine the precision of efficient approximation algorithms for geometric problems that arise naturally from the study of entanglement and from models of interacting particles considered in classical statistical physics.

In this thesis we study entanglement by using *nonlocal games*. A nonlocal game involves two or more players who are not allowed to communicate with each other, but do interact with an extra party usually referred to as the referee. At the start of the game the referee asks each of the players a question, upon which they each reply to him with some answer. Then, the referee decides if the players win or lose based only on the questions he asked and the answers he received. The players know in advance what set of answers would cause them to win, which of course is their objective. The catch is that they only know the question that was aimed directly at them and not any of the other players' questions. The players thus don't play against each other, but should somehow coordinate their strategies to win.

The best course of action for players who live in a world described by Classical Mechanics is the simplest kind imaginable: just fix in advance what to answer to each question. In a Quantum Mechanical world, more sophisticated strategies sometimes give better results. Each player can base their answer on the outcome of an experiment done on some private physical system. The key feature of such strategies is that they can cause the players to produce answers

that are *correlated* in ways that are impossible in a classical world. In this case the players are said to be *entangled*.

The fact that Quantum Mechanics predicts such phenomena was used by Einstein, Podolski and Rosen in 1935 to argue that this theory must be incomplete, as surely entanglement could not be part of a reasonable description of Nature. Surprisingly, experiments done by Aspect et al. in the 1980's gave convincing evidence that the world we live does in fact allow for this!

Entanglement is usually mathematically described by a vector in a Hilbert space. Such a vector is referred to as a *state*. We prove that for a large class of states the advantage gained by using them over classical strategies in the simplest nonlocal games involving three or more players is severely limited. As a bonus, the proof of this result can also be used to resolve a 35-year-old open problem posed by Varopoulos in an area of mathematics called Banach Space Theory.

Optimization means searching over a huge collection to find some element with the best characteristics. One example of such a problem is finding a strategy for a nonlocal game that maximizes the players' winning probability. An second example is to optimize the directions of the magnetic fields of interacting particles so as to minimize the total energy of the system.

The most-studied optimization problems usually have a combinatorial nature. For example, finding an optimal classical strategy for a nonlocal game or minimizing the energy of interacting particles described by the celebrated Ising model amounts to searching over a discrete set of possibilities. In this thesis we consider problems with a more *geometric* flavor. To picture this, imagine searching for some optimal configuration of a finite number of points on a three-dimensional sphere. Such problems arise naturally from the study of entanglement when one restricts the amount of entanglement players are allowed to use in nonlocal games, and from the Heisenberg model of interacting particles in classical statistical physics.

Unfortunately, most problems like the ones described above likely can't be solved exactly by any computer in a reasonable amount of time. If time is of the essence, then the next-best thing is to search for any solution that is near-optimal, but can be found in a reasonable amount of time. We will use new variants of Grothendieck's Inequality to analyze algorithms that offer exactly such an alternative for the geometric optimization problems mentioned above.

Titles in the ILLC Dissertation Series:

ILLC DS-2006-01: **Troy Lee**
Kolmogorov complexity and formula size lower bounds

ILLC DS-2006-02: **Nick Bezhanishvili**
Lattices of intermediate and cylindric modal logics

ILLC DS-2006-03: **Clemens Kupke**
Finitary coalgebraic logics

ILLC DS-2006-04: **Robert Špalek**
Quantum Algorithms, Lower Bounds, and Time-Space Tradeoffs

ILLC DS-2006-05: **Aline Honingh**
The Origin and Well-Formedness of Tonal Pitch Structures

ILLC DS-2006-06: **Merlijn Sevenster**
Branches of imperfect information: logic, games, and computation

ILLC DS-2006-07: **Marie Nilsenova**
Rises and Falls. Studies in the Semantics and Pragmatics of Intonation

ILLC DS-2006-08: **Darko Sarenac**
Products of Topological Modal Logics

ILLC DS-2007-01: **Rudi Cilibrasi**
Statistical Inference Through Data Compression

ILLC DS-2007-02: **Neta Spiro**
What contributes to the perception of musical phrases in western classical music?

ILLC DS-2007-03: **Darrin Hindsill**
It's a Process and an Event: Perspectives in Event Semantics

ILLC DS-2007-04: **Katrin Schulz**
Minimal Models in Semantics and Pragmatics: Free Choice, Exhaustivity, and Conditionals

ILLC DS-2007-05: **Yoav Seginer**
Learning Syntactic Structure

ILLC DS-2008-01: **Stephanie Wehner**
Cryptography in a Quantum World

ILLC DS-2008-02: **Fenrong Liu**
Changing for the Better: Preference Dynamics and Agent Diversity

ILLC DS-2008-03: **Olivier Roy**
Thinking before Acting: Intentions, Logic, Rational Choice

ILLC DS-2008-04: **Patrick Girard**
Modal Logic for Belief and Preference Change

ILLC DS-2008-05: **Erik Rietveld**
Unreflective Action: A Philosophical Contribution to Integrative Neuroscience

ILLC DS-2008-06: **Falk Unger**
Noise in Quantum and Classical Computation and Non-locality

ILLC DS-2008-07: **Steven de Rooij**
Minimum Description Length Model Selection: Problems and Extensions

ILLC DS-2008-08: **Fabrice Nauze**
Modality in Typological Perspective

ILLC DS-2008-09: **Floris Roelofsen**
Anaphora Resolved

ILLC DS-2008-10: **Marian Counihan**
Looking for logic in all the wrong places: an investigation of language, literacy and logic in reasoning

ILLC DS-2009-01: **Jakub Szymanik**
Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language

ILLC DS-2009-02: **Hartmut Fitz**
Neural Syntax

ILLC DS-2009-03: **Brian Thomas Semmes**
A Game for the Borel Functions

ILLC DS-2009-04: **Sara L. Uckelman**
Modalities in Medieval Logic

ILLC DS-2009-05: **Andreas Witzel**
Knowledge and Games: Theory and Implementation

ILLC DS-2009-06: **Chantal Bax**
Subjectivity after Wittgenstein. Wittgenstein's embodied and embedded subject and the debate about the death of man.

ILLC DS-2009-07: **Kata Balogh**
Theme with Variations. A Context-based Analysis of Focus

ILLC DS-2009-08: **Tomohiro Hoshi**
Epistemic Dynamics and Protocol Information

ILLC DS-2009-09: **Olivia Ladinig**
Temporal expectations and their violations

ILLC DS-2009-10: **Tikitu de Jager**
"Now that you mention it, I wonder…": Awareness, Attention, Assumption

ILLC DS-2009-11: **Michael Franke**
Signal to Act: Game Theory in Pragmatics

ILLC DS-2009-12: **Joel Uckelman**
More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains

ILLC DS-2009-13: **Stefan Bold**
Cardinals as Ultrapowers. A Canonical Measure Analysis under the Axiom of Determinacy.

ILLC DS-2010-01: **Reut Tsarfaty**
Relational-Realizational Parsing

ILLC DS-2010-02: **Jonathan Zvesper**
Playing with Information

ILLC DS-2010-03: **Cédric Dégremont**
The Temporal Mind. Observations on the logic of belief change in interactive systems

ILLC DS-2010-04: **Daisuke Ikegami**
Games in Set Theory and Logic

ILLC DS-2010-05: **Jarmo Kontinen**
Coherence and Complexity in Fragments of Dependence Logic

ILLC DS-2010-06: **Yanjing Wang**
Epistemic Modelling and Protocol Dynamics

ILLC DS-2010-07: **Marc Staudacher**
Use theories of meaning between conventions and social norms

ILLC DS-2010-08: **Amélie Gheerbrant**
Fixed-Point Logics on Trees

ILLC DS-2010-09: **Gaëlle Fontaine**
Modal Fixpoint Logic: Some Model Theoretic Questions

ILLC DS-2010-10: **Jacob Vosmaer**
Logic, Algebra and Topology. Investigations into canonical extensions, duality theory and point-free topology.

ILLC DS-2010-11: **Nina Gierasimczuk**
Knowing One's Limits. Logical Analysis of Inductive Inference

ILLC DS-2011-01: **Wouter M. Koolen**
Combining Strategies Efficiently: High-Quality Decisions from Conflicting Advice

ILLC DS-2011-02: **Fernando Raymundo Velazquez-Quesada**
Small steps in dynamics of information

ILLC DS-2011-03: **Marijn Koolen**
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

ILLC DS-2011-04: **Junte Zhang**
System Evaluation of Archival Description and Access

ILLC DS-2011-05: **Lauri Keskinen**
Characterizing All Models in Infinite Cardinalities

ILLC DS-2011-06: **Rianne Kaptein**
Effective Focused Retrieval by Exploiting Query Context and Document Structure

ILLC DS-2011-07: **Jop Briët**
Grothendieck Inequalities, Nonlocal Games and Optimization