SEMINAR PAPER
NONLOCAL GAMES AND THE GROTHENDIECK-TSIRELSON INEQUALITY

MAXI BRANDSTETTER, ARNE HEIMENDAHL, FELIX KIRSCHNER

ABSTRACT. Dieses Paper ist eine Ausarbeitung unseres Vortrags im Seminar "Introduction to Quantum Information and Quantum Computing", das zwischen dem 19.9. und 21.9.2018 in Köln stattgefunden hat. Es wird eine kleine Einführung in die mathematischen Methoden gegeben, mit denen die Welt der Quantenfunktion beschrieben werden kann, woraufhin wir "Nonlocal Games" einführen, die Brücke zur (semidefiniten) Optimierung schlagen und hoffentlich noch genug Zeit für die Grothendieck Ungleichungen haben. Die Grothendieck Ungleichungen finden erstaunlicherweise in einer Vielzahl an mathematischen Teilgebieten Verwendung.

CONTENTS

1. **Quantum Grundlagen**

1.1. ***Basic definitions***

In order to make sure everyone is on the same page we will have the following introduction where the necessary groundwork is done. The aim is to call a few basic definitions back to our memory so one can fluently read through this paper.
A complex matrix $A \in \mathbb{C}^{n\times n}$ is called Hermitian if $A^* = A$, where A^* denotes the conjugate transpose of A. We will denote the set of all Hermitian matrices operating on a Hilbert space $\mathscr{H}$ by $B^H(\mathscr{H})$. A complex Hermitian matrix A is called positve semidefinite (abbreviated psd.) if one of the following holds:

(i) The matrix has only real non-negative eigenvalues
(ii) There exist complex $n-$dimensional vectors $z_1, \ldots, z_n$ s.t. $A_{i,j} = \langle z_i, z_j\rangle = \sum_{k=1}^n \overline{z}_{i_k} z_{j_k}$
(iii) For every $z \in \mathbb{C}^n$ we have $z^* A z \geq 0$
(iv) There exists a complex matrix B s.t. $A = B^* B$
(v) $\mathrm{Tr}\,(A \cdot B) \geq 0$ for all positive semidefinite operators B defined in the same space.

It can be shown that (i) - (v) are in fact equivalent. The set of positive semidefinite matrices is a cone, meaning for two psd. $n \times n$ matrices A, B and $\alpha, \beta \in \mathbb{R}_+$ we have that $\alpha A + \beta B$ is also positive semidefinite. The set of real $n \times n$ positive semidefinite matrices is denoted by $\mathcal{S}_n^+$. The set of positive semidefinite operators operating on a Hilbert space $\mathscr{H}$ is denoted by $B^{psd}(\mathscr{H})$.

1.2. ***Tensor products and Dirac notation***

In case someone is not familiar with the *tensor product* a short introduction with an example or two is given here: Let $\mathcal{X} = \mathbb{C}^{n_1 \times m_1}$ 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 $\mathcal{X} \otimes \mathcal{Y} = \mathbb{C}^{n_1 n_2 \times m_1 m_2}$. The tensor product of complex matrices can be obtained as follows: Index the rows and columns of a matrix by $\mathscr{R}$ and $\mathcal{C}$ and think of the matrix as a map from $\mathscr{R} \times \mathcal{C} \to \mathbb{C}$. For two complex matrices $A : \mathscr{R}_1 \times \mathcal{C}_1 \to \mathbb{C}$ and $B : \mathscr{R}_2 \times \mathcal{C}_2 \to \mathbb{C}$ their tensor product is the matrix $A \otimes B : (\mathscr{R}_1 \times \mathscr{R}_2) \times (\mathcal{C}_1 \times \mathcal{C}_2) \to \mathbb{C}$ defined by $(A \otimes B)((r_1, r_2), (c_1, c_2)) = A(r_1, c_1)B(r_2, c_2)$. Considering a lexicographic order understand tensor products in the following way: $A \otimes B = \begin{pmatrix} a_{11}B & \ldots & a_{1n}B \\ \vdots & & \vdots \\ a_{n1}B & \ldots & a_{nn}B \end{pmatrix}$, which is the *Kronecker product* and coherent with the definition. For two complex vectors v_1, v_2 when we talk about their tensor products we will mean $v_1 \otimes v_2 = (v_{1_1} v_2, v_{1_2} v_2, \ldots, v_{1_n} v_2)^\top$ from which we can deduce that $\langle x_1 \otimes x_2, y_1 \otimes y_2\rangle = \langle x_1, x_2\rangle\langle y_1, y_2\rangle$. Also for any matrices A, B, C, D (assuming fitting dimensions) we have the following identities:

i, $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
ii, $A \otimes (B + C) = A \otimes B + A \otimes C$
iii, $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$

Also throughout this paper we will stick to the *Dirac notation*, which is the standard notation for describing quantum states. In Dirac notation $|\psi\rangle$ refers to a vector in $\mathbb{C}^n$. The conjugate transpose of this vector is written $\langle\psi|$. The non-negative integers, by convention, represent the canonical basis vectors, i.e.

$$|0\rangle = \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}, |1\rangle = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}, \ldots, |n-1\rangle = \begin{pmatrix}0\\0\\\vdots\\0\\1\end{pmatrix} \tag{1.1}$$

Usually the tensor product symbol is omitted when taking the tensor product of two vectors in Dirac notation. This means we write $|\psi\rangle|\phi\rangle$ instead of $|\psi\rangle \otimes |\phi\rangle$. We also would like to quickly remind ourselves what a Hilbert space is. Let $\mathscr{H}$ be an inner product space. Endow $\mathscr{H}$ with a norm $||x|| = \sqrt{\langle x, x\rangle}$ and a metric $d(x,y) = ||x-y||$. If every Cauchy sequence in $\mathscr{H}$ converges to an element in $\mathscr{H}$, i.e. $\mathscr{H}$ is complete, then $\mathscr{H}$ is a Hilbert space.

1.3. ***States and measurements***

Now we can define a *state*.
A state is a complex positive semidefinite matrix ρ that satisfies $\mathrm{Tr}(\rho) = 1$. The trace of a psd. matrix is equal to the sum of its eigenvalues. The spectral theorem tells us that any $n \times n$ Hermitian matrix can be decomposed as $\rho = \sum_{i=1}^{n} \lambda_i |\psi_i\rangle\langle\psi_i|$ with λ_i being its eigenvalues and $|\psi_i\rangle$ the corresponding eigenvectors. We call a state *pure* if it has rank 1, i.e. $\rho = |\psi\rangle\langle\psi|$ for some complex unit vector $|\psi\rangle$. This means every state is a convex combination of pure states. Note that complex unit vectors are often referred to as states even though states are defined as matrices. What is actually meant is the pure state $|\psi\rangle\langle\psi|$. The name state for these mathematical objects is chosen because with them the possible configurations of a quantum system can be modeled. A quantum system X is said to be *in* state ρ and is associated with a positive integer n, referred to as its dimension and a copy of $\mathbb{C}^n$. The states in $\mathbb{C}^{n\times n}$ give the possible configurations of X. A quantum system X may consist subsystems $X_1, \ldots, X_N$, where each subsystem X_i is a quantum system for itself. And X is then associated with $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_N}$ with n_i being the dimensions of the subsystems. A state ρ in which X is then is a matrix of size $n_1 \ldots n_N$.
What physicists usually do is building some mathematical model that is supposed to describe how the universe behaves and then test this model by performing experiments and comparing the results to what the model predicts. We will now define an experiment, a *measurement* of a quantum state, in a mathematical way. It shall be stated that the measurements we are talking about do not compare to a physical measurement like measuring the temperature, atmospheric pressure or any other continuous physical quantity. The outcome of measuring a temperature in Kelvin may be a real number $T \in [0, \infty)$ but we are going to assume the measurements we consider only have a finite set of outcomes $\mathscr{A}$. As before we will have an n-dimensional quantum system X and a measurement on X in state ρ with outcomes in $\mathscr{A}$ is defined by a set of psd. matrices $\{F^a\}_{a\in\mathscr{A}} \subseteq \mathbb{C}^{n\times n}$ that sum up to the identity matrix, i.e. $\sum_{a\in\mathscr{A}} F^a = I$. A *projective* measurement is defined by matrices that satisfy $F^a F^b = \delta_{ab} F^a$ for all $a, b \in \mathscr{A}$. The outcome of a measurement is a random variable χ and its probability distribution is given by $\mathbb{P}[\chi = a] = \mathrm{Tr}(\rho F^a)$. In order to be able to define an expected value for an projective measurement it is convenient to define the outcomes in $\mathscr{A}$ as real numbers.

In that case we have:

$$\mathbb{E}\left[\chi\right] = \sum_{a\in\mathcal{A}} a\mathrm{Tr}(\rho F^a) = \mathrm{Tr}(\rho(\sum_{a\in\mathcal{A}} aF^a)) \tag{1.2}$$

The sum of the matrices times their outcome value is called an *observable* associated to a projective measurement. A very simple case of this would be a $\{-1,1\}$-valued observable, where $\{-1,1\}$ is the set of outcomes. Such an observable is defined as $\sum_{a\in\{-1,1\}} aF^a = (-1)F^- + (1)F^+ = F^+ - F^-$. Since we are considering projective measurements, squaring the difference yields

$$(F^+ - F^-)^2 = \underbrace{{F^+}^2}_{=F^+} - \underbrace{F^+F^-}_{\delta_{+-}=0} + \underbrace{{F^-}^2}_{F^-} = F^+ + F^- = I \tag{1.3}$$

i.e. a $\{-1,1\}$-valued observable is both unitary and Hermitian.
Now let us take a quantum system X consisting of subsystems $X_1,\dots,X_N$ an distribute the subsystems among N parties. If X is in state ρ we say that state ρ is shared by the parties. The parties may have an arbitrary distance to each other, i.e. may be located anywhere in the universe. Every party can perform a measurement on their subsystem. This means there are N sets of psd. matrices $\{F^{a_1}\}_{a_1\in\mathcal{A}_1} \in \mathbb{C}^{n_1\times n_1},\dots,\{F^{a_N}\}_{a_N\in\mathcal{A}_N} \in \mathbb{C}^{n_N\times n_N}$ and the joint probability distribution of the N measurement outcomes $\chi_1,\dots,\chi_N$ is

$$\mathbb{P}\left[\chi_1 = a_1, \chi_2 = a_2, \dots, \chi_N = a_N\right] = \mathrm{Tr}(\rho F_1^{a_1} \otimes \dots \otimes F_N^{a_N}) \tag{1.4}$$

As we defined earlier a pure state ρ has rank 1 and in the case that the system X we have consists of subsystems there is a $|\psi\rangle \in \mathbb{C}^{n_1} \otimes \dots \otimes \mathbb{C}^{n_N}$ such that $\rho = |\psi\rangle\langle\psi|$. The state is called *product state* if it is of the form $|\psi\rangle = |\psi_1\rangle|\psi_2\rangle\dots|\psi_N\rangle$. A state that is not a product state is said to be entangled. A mixed state, i.e. a state with rank greater than 1 is said to be separable if it is a convex combination of pure states. The interesting thing and what makes quantum mechanics interesting is that entangled states can give correlated measurement outcomes. What makes this especially mind-boggling is the fact that the parties can be located anywhere in the universe. This means that the information of a measurement can travel at an instant.
In the following example we would like to show that if two players, call them Alice and Bob, share a product state, the result is in fact a product distribution, i.e. the measurement outcome do not correlate. So, let $|\psi\rangle = |\psi_A\rangle|\psi_B\rangle$ and let Alice perform a measurement $\{F^a\}_{a\in\mathcal{A}}$ on her $|\psi_A\rangle$ and let Bob perform a measurement $\{G^b\}_{b\in\mathcal{B}}$ on his $|\psi_B\rangle$. The probability of Alice getting measurement outcome $\chi_A = a$ and Bob getting $\chi_B = b$ is equal to:

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

Where the third and fourth equality follow from the fact that $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ and that the tensor product of two real numbers is equal to their ordinary product. The result is just the product of the probability of Alice measuring a and Bob measuring b, as desired.

2. Nonlocal games

In this section we will introduce nonlocal games, which are a systematic approach of studying quantum mechanics and its properties and comparing it to classical mechanics. They are called nonlocal because the players are assumed to be very far, like light-years, away from each other. But first things first. In the basic case there are three participants, two players Alice and Bob and a referee. The referee sends a piece of information to Alice and Bob. They may or may not receive the same information. Afterwards both Alice and Bob must, without communicating, send an answer to the referee, who then decides whether they both win or both lose and the game ends. Mathematically speaking this means there are four finite sets $\mathcal{A}, \mathcal{B}, \mathcal{S}, \mathcal{T}$, a joint probability distribution π over $\mathcal{S} \times \mathcal{T}$, i.e. $\pi : \mathcal{S} \times \mathcal{T} \to [0,1]$. The referee sends with probability $\pi(s,t)$ $s \in \mathcal{S}$ to Alice and $t \in \mathcal{T}$ to Bob and they both answer with an element $a \in \mathcal{A}$ and $b \in \mathcal{B}$ respectively. Whether they win or lose is determined by a map $V : \mathcal{A} \times \mathcal{B} \times \mathcal{S} \times \mathcal{T} \to \{0,1\}$. They win if $V(a,b,s,t) = 1$ and lose otherwise. All players know π and V but they they do not know what element the other player received from the referee. They may agree on a strategy beforehand but they must not communicate once the game has started. Obviously, Alice and Bob want to win the game and so they try to maximize their winning probability by choosing a promising strategy.

2.1. *Classical and entangled strategies*

When Alice and Bob use classic deterministic strategies, they both have a deterministic map $a : \mathcal{S} \to \mathcal{A}$ and $b : \mathcal{T} \to \mathcal{B}$ respectively. This means beforehand they both agree on what to answer upon what questions received. The winning probability is easily calculated:

$$\mathbb{E}_{s,t\sim\pi}\left[V(a(s), b(t), s, t)\right] \tag{2.1}$$

But, of course, we are dealing with quantum mechanics here so we are interested in entangled strategies and want to study how the availability of these influence the outcome. For an entangled strategy both Alice and Bob have a subsystem X_A, X_B of a quantum system X which is in state ρ, i.e. Alice and Bob share state ρ. If the state is entangled we know that measurements can give correlated outcomes, which means for the players that they may gain information about the other players outcome by performing a measurement. More technically, there is a positive integer n and a quantum system X consisting of two n-dimensional subsystems X_A, X_B in some entangled state ρ. Alice and Bob have measurements $\{F_s^a\}_{a\in\mathcal{A}}, \{G_t^b\}_{b\in\mathcal{B}} \subseteq \mathbb{C}^{n\times n}$. When the game starts they both get a question s and t and perform their measurement on it. They both send the outcome of their measurement as their answer to the referee. As has be established before, the probability of Alice answering with a and Bob with b is $\mathrm{Tr}(\rho F_s^a \otimes G_t^b)$. The winning probability then equals:

$$\mathbb{E}_{s,t\sim\pi}\left[\sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}} \mathrm{Tr}(\rho F_s^a \otimes G_t^b) V(a,b,s,t)\right] \tag{2.2}$$

The players want to maximize their winning probability. Since the trace function is linear and states are convex combinations of pure states, we only need to consider pure entangled states.

2.2. ***Two player XOR games***

An XOR game is a game where the set of answers $\mathcal{A}$ and $\mathcal{B}$ only consist of $\{0,1\}$ and the predicate V only depends on the exclusive-OR of the answers and the value function $f : \mathcal{S} \times \mathcal{T} \to \{0,1\}$. In the following let the spare brackets denote the 0/1 truth value of the statement in between them. Then we have $V(a,b,s,t) = [a \oplus b = f(s,t)]$. The exclusive OR returns 1 if and only if one of the inputs is 1. In a truth table:

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

For a probability distribution π and a boolean function f, $\mathcal{G} = (\pi, f)$ defines an XOR game.

Bias and violation ratio. Alice and Bob can always win an XOR with probability $\frac{1}{2}$ by flipping an unbiased coin. Interesting would be how much this can actually be increased. We define the classical bias of an XOR game to be the difference of the probability of winning and losing for an optimal classical strategy and denote it by $\beta(G)$. The bias of entangled strategies is calculated in the same way and thus is twice the amount by which the maximal winning probability is greater then $\frac{1}{2}$, since $\frac{1}{2} + \gamma - (1 - \frac{1}{2} - \gamma) = 2\gamma$, γ being the amount exceeding $\frac{1}{2}$. We denote the bias of entangled strategies by $\beta^*(G)$. The violation ratio is given by $\frac{\beta^*(G)}{\beta(G)}$.

Signs and observables. To make things a little easier regarding calculations we will use the $\{-1,1\}$-basis rather than the $\{0,1\}$-basis for boolean valued objects. If we have an XOR game (π, f) and any two classical strategies $a : \mathcal{S} \to \{0,1\}$ and $b : \mathcal{B} \to \{0,1\}$ the bias is given by the probability under π that $a(s) \oplus b(t) = f(s,t)$ minus the probability under π that $a(s) \oplus b(t) \neq f(s,t)$.

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

We define the sign matrix $\Sigma_{st} = (-1)^{f(s,t)}$ and functions $\chi(s) = (-1)^{a(s)}$ and $\psi(t) = (-1)^{b(t)}$. Thus the bias is:

$$\mathbb{E}_{(s,t)\sim\pi}\left[\chi(s)\psi(t)\Sigma_{st}\right] \tag{2.3}$$

Since in an XOR game the outcomes are $\{0,1\}$-valued the measurements Alice and Bob have are $\{F_s^0, F_s^1\}$ and $\{G_t^0, G_t^1\}$. If we consider an entangled strategy with a pure state $|\psi\rangle$ and have projective measurements the probability of Alice and Bob answering with a, b upon receiving s, t respectively is $\langle\psi|F_s^a \otimes G_t^b|\psi\rangle$. We can calculate the expected value:

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

As in (3) we define the $\{-1,1\}$-observables $F_s = F_s^0 - F_s^1$ and $G_t^0 - G_t^1$ with the property that its difference squared is the identity matrix. Using this strategy the bias becomes

$$\mathbb{E}_{(s,t)\sim\pi}\left[\langle\psi|F_s\otimes G_t|\psi\rangle\Sigma_{s,t}\right] \tag{2.4}$$

So for any XOR game the bias is defined as the difference of the probabilities of winning and loosing which is, if considering the $\{-1,1\}$ basis, the expected value and we are looking to maximize this quantity. Hence, the bias of an XOR games in classical strategies is:

$$\max\left\{\mathbb{E}_{(s,t)\sim\pi}\left[\Sigma_{st}\chi(s)\psi(t)\right] : \chi : \mathcal{S}\to\{-1,1\}, \psi : \mathcal{T}\to\{-1,1\}\right\} \tag{2.5}$$

For entangled strategies we need to use the $\sup_{n\in\mathbb{N}}$ since the winning probability might increase indefinitely with the dimension of the quantum system. The Bias of entangled strategies is

$$\sup_{n\in\mathbb{N}}\left\{\mathbb{E}_{(s,t)\sim\pi}\left[\Sigma_{st}\langle\psi|F_s\otimes G_t|\psi\rangle\right] : |\psi\rangle\in\mathbb{C}^n\otimes\mathbb{C}^n, F_s, G_t\in O(\mathbb{C}^n)\right\} \tag{2.6}$$

where $O(\mathbb{C}^n)$ denotes the set of $\{-1,1\}$-observables in $\mathbb{C}^{n\times n}$. As shown in [KNP17] we can in fact restrict ourselves to projective measurements. More general measurements like POVMs (positive operator valued (probability) measure) are not advantageous.

2.3. ***The CHSH game***

Let us consider a special instance of XOR games which leads to the result that entangled strategies actually can give a remarkable advantage over classical strategies. The game is named after four scientists Clauser, Horne, Shimony and Holt. The question set is $\{0,1\}\times\{0,1\}$ as well as the answer set. The probability distribution over the question set is the uniform distribution and the predicate $V=[a\oplus b=s\wedge t]$. Note that $s\wedge t$ only evaluates to 1 if both $s=1$ and $t=1$, which in the case of the uniform distribution happens in $\frac{1}{4}$ of the cases. The best classical strategy then would be either to always answer $a=0, b=1$ or $a=1, b=0$. Since in both cases $a\oplus b=0$ Alice and Bob win in with probability $\frac{3}{4}$. We will now study the entangled case and check how much the winning probability may be increased. Define

$$X=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, Y=\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} \tag{2.7}$$

and note that they anti-commute, i.e. $XY+YX=0$ and square to the identity matrix $X^2=Y^2=I$. For Alice define the observable for question 0 by $F_0=X$ and for question 1 by $F_1=Y$. Bobs observables are going to be $G_0=(X-Y)/\sqrt{2}$ for question 0 and $G_1=(X+Y)/\sqrt{2}$ for question 1. Define the $|\mathrm{EPR}\rangle=\frac{|0\rangle|0\rangle+|1\rangle|1\rangle}{\sqrt{2}}=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}$. The following auxiliary calculations will be helpful later:

$$\langle \mathrm{EPR}|X \otimes X|\mathrm{EPR}\rangle = \frac{1}{2}\begin{pmatrix}1 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}$$
$$= \frac{1}{2}\begin{pmatrix}1 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1\\0\\0\\1\end{pmatrix} = \frac{2}{2} = 1$$

$$\langle \mathrm{EPR}|Y \otimes Y|\mathrm{EPR}\rangle = \frac{1}{2}\begin{pmatrix}1 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & 0 & -1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ -1 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}$$
$$= \frac{1}{2}\begin{pmatrix}-1 & 0 & 0 & -1\end{pmatrix}\begin{pmatrix}1\\0\\0\\1\end{pmatrix} = -1$$

$$\langle \mathrm{EPR}|X \otimes Y|\mathrm{EPR}\rangle = \frac{1}{2}\begin{pmatrix}1 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & 0 & -i\\ 0 & 0 & i & 0\\ 0 & -i & 0 & 0\\ i & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}$$
$$= \frac{1}{2}\begin{pmatrix}i & 0 & 0 & -i\end{pmatrix}\begin{pmatrix}1\\0\\0\\1\end{pmatrix} = 0$$
$$\langle \mathrm{EPR}|Y \otimes X|\mathrm{EPR}\rangle = 0$$

Lets calculate the expected values of the sign $a \oplus b$:

- $$\begin{aligned}\langle \mathrm{EPR}|F_0 \otimes G_0|\mathrm{EPR}\rangle &= \langle \mathrm{EPR}|X \otimes \frac{1}{\sqrt{2}}(X - Y)|\mathrm{EPR}\rangle\\ &= \langle \mathrm{EPR}|X \otimes \frac{1}{\sqrt{2}}X|\mathrm{EPR}\rangle - \langle \mathrm{EPR}|X \otimes \frac{1}{\sqrt{2}}Y|\mathrm{EPR}\rangle\\ &= \frac{1}{\sqrt{2}} - 0 = \frac{1}{\sqrt{2}}\end{aligned}$$

- $$\begin{aligned}\langle \mathrm{EPR}|F_1 \otimes G_1|\mathrm{EPR}\rangle &= \langle \mathrm{EPR}|Y \otimes \frac{1}{\sqrt{2}}(X+Y)|\mathrm{EPR}\rangle \\ &= \langle \mathrm{EPR}|Y \otimes \frac{1}{\sqrt{2}}X|\mathrm{EPR}\rangle + \langle \mathrm{EPR}|Y \otimes \frac{1}{\sqrt{2}}Y|\mathrm{EPR}\rangle \\ &= 0 - \frac{1}{\sqrt{2}} = -\frac{1}{\sqrt{2}}\end{aligned}$$

- $$\begin{aligned}\langle \mathrm{EPR}|F_0 \otimes G_1|\mathrm{EPR}\rangle &= \langle \mathrm{EPR}|X \otimes \frac{1}{\sqrt{2}}(X+Y)|\mathrm{EPR}\rangle \\ &= \langle \mathrm{EPR}|X \otimes \frac{1}{\sqrt{2}}X|\mathrm{EPR}\rangle + \langle \mathrm{EPR}|X \otimes \frac{1}{\sqrt{2}}Y|\mathrm{EPR}\rangle \\ &= \frac{1}{\sqrt{2}} + 0 = \frac{1}{\sqrt{2}}\end{aligned}$$

- $$\begin{aligned}\langle \mathrm{EPR}|F_1 \otimes G_0|\mathrm{EPR}\rangle &= \langle \mathrm{EPR}|Y \otimes \frac{1}{\sqrt{2}}(X-Y)|\mathrm{EPR}\rangle \\ &= \langle \mathrm{EPR}|Y \otimes \frac{1}{\sqrt{2}}X|\mathrm{EPR}\rangle - \langle \mathrm{EPR}|Y \otimes \frac{1}{\sqrt{2}}Y|\mathrm{EPR}\rangle \\ &= 0 - (-\frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}}\end{aligned}$$

Thus, we have

$$\langle \mathrm{EPR}|F_s \otimes G_t|\mathrm{EPR}\rangle = \begin{cases} \frac{1}{\sqrt{2}}, (0,0), (1,0), (0,1) \\ -\frac{1}{\sqrt{2}}, (1.1) \end{cases} \tag{2.8}$$

which is equivalent to

$$\langle \mathrm{EPR}|F_s \otimes G_t|\mathrm{EPR}\rangle = \frac{(-1)^{s\wedge t}}{\sqrt{2}}, s,t \in \{0,1\} \tag{2.9}$$

The bias of the entangled strategy equals

$$\begin{aligned}\mathbb{E}_{(s,t)\sim\pi}\left[\Sigma_{s,t}\langle\psi|F_s \otimes G_t|\psi\rangle\right] &= \frac{1}{4}\sum_{s,t=0}^{1}(-1)^{s\wedge t}\langle \mathrm{EPR}|F_s \otimes G_t|\mathrm{EPR}\rangle \\ &= \frac{1}{4}\cdot\frac{4}{\sqrt{2}} = \frac{1}{\sqrt{2}}\end{aligned}$$

The bias is $\frac{1}{\sqrt{2}}$ from which follows that the winning probability is by definition:

$$\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{\sqrt{2}} = \cos(\pi/8) \approx 0.85\ldots \tag{2.10}$$

3. Local and Quantum correlation matrices

3.1. *Local Correlation matrices*

So far, we have been dealing with quite specific strategies. The idea is to generalize the concept of strategies into mathematical objects. Suppose the referee sends an element $s \in \mathcal{S}$ to Alice, respectively $t \in \mathcal{T}$ to Bob. Suppose Alice and Bob answer according to a deterministic strategy. We will interpret their answers as vectors $\xi \in \{-1,1\}^{\mathcal{S}}, \eta \in \{-1,1\}^{\mathcal{T}}$. Their common answer is the product $\xi_s \eta_t$. So, their strategy can be uniquely described by a matrix $\xi\eta^\top$. Instead of playing a deterministic strategy they could answer in a probabilistic way, meaning that given the input $s \in \mathcal{S}$, respectively $t \in \mathcal{T}$ their answers are determined by random variables X_s and Y_t. Then, their expected common answer is $\mathbb{E}[X_s Y_t]$. In the following definition we define the set of all such matrices encoding the common strategy of Alice and Bob.

Definition 3.1.1. Let $(X_i)_{1\leq i\leq m}$ and $(Y_j)_{1\leq j\leq n}$ be families of random variables on a common probability space such that $|X_i|, |Y_j| \leq 1$ almost surely. Then $A = (a_{ij})$ is the corresponding *classical (or local) correlation matrix* if

$$a_{ij} = \mathbb{E}[X_i Y_j]$$

for all $1 \leq i \leq m, 1 \leq j \leq n$.

As we will see in the sequel, the set of $m \times n$ correlation matrices is a polytope, denoted by $\mathrm{LC}_{m,n}$.

Lemma 3.1.2. *An alternative description of* $\mathrm{LC}_{m,n}$ *is given by*

$$\mathrm{LC}_{m,n} = \mathrm{conv}\{\xi\eta^T \,|\, \xi \in \{-1,1\}^m, \eta \in \{-1,1\}^n\}. \tag{3.1}$$

What does the lemma tell us in addition to a simples description? It states that the matrices encoding all possible strategies are the convex hull of the matrices defined by deterministic ones. We can interpret this as follows: no matter which probabilistic strategy Alice and Bob play there will be a deterministic strategy that is at least as good as theirs.

Proof. Let us denote the right hand side of 3.1 by M and let $\xi\eta^T \in M$ with $\xi \in \{-1,1\}^m, \eta \in \{-1,1\}^n$. Clearly $\xi_i, \eta_j \in \{-1,1\}$ define constant $\mathbb{R}$-valued random variables satisfying $|\xi_i|, |\eta_j| \leq 1$. Hence, it suffices to show that $\mathrm{LC}_{m,n}$ is convex since it contains the vertices of M. Therefore, consider two classical correlation matrices $a_{ij}^{(k)} = \mathbb{E}[X_i^{(k)} Y_j^{(k)}]$ for $k \in \{0,1\}$ which are defined on a common probability space such that $\left|X_i^{(k)}\right|, \left|Y_j^{(k)}\right| \leq 1$. We have to show that there exists random variables $(X_i), (Y_j)$ with $|X_i|, |Y_j| \leq 1$ almost surely such that

$$\beta a_{ij}^{(0)} + (1-\beta) a_{ij}^{(1)} = \mathbb{E}[X_i Y_j]$$

for all $\beta \in [0,1]$. Let α be a Bernoulli random variable, i.e. $\mathbb{P}(\alpha = 0) = \beta$, $\mathbb{P}(\alpha = 1) = 1 - \beta$ and set $X_i = X_i^{(\alpha)}, Y_j = Y_j^{(\alpha)}$. Clearly, it holds $|X_i|, |Y_j| \leq 1$ almost surely. Then

$$\begin{aligned}\mathbb{E}[X_iY_j] &= \mathbb{E}[X_i^{(\alpha)}Y_j^{(\alpha)}\mathbb{1}_{\{\alpha=0\}}] + \mathbb{E}[X_i^{(\alpha)}Y_j^{(\alpha)}]\mathbb{1}_{\{\alpha=1\}}] \\ &= \beta\mathbb{E}[X_i^{(0)}Y_j^{(0)}] + (1-\beta)\mathbb{E}[X_i^{(1)}Y_j^{(1)}],\end{aligned}$$

which proofs that $\mathrm{LC}_{m,n}$ is convex.

For the other inclusion, let $(a_{ij}) \in \mathrm{LC}_{m,n}$, i.e. $a_{ij} = \mathbb{E}[X_iY_j]$ for $\mathbb{R}$-valued random variables $(X_i), (Y_j)$, defined on a common probability space Ω with $|X_i|, |Y_j| \leq 1$ almost surely. We will use the characterization of the $d-$dimensional cube by its vertices, that is $[-1,1]^d = \mathrm{conv}\{\xi \,|\, \xi \in \{-1,1\}^d\}$.
If we define the random variables $X = (X_1, \dots, X_m)$ and $Y = (Y_1, \dots, Y_n)$ they satisfy $X \in [-1,1]^m$, $Y \in [-1,1]^n$ almost surely. Using the characterization of the hypercube we can define random variables $\lambda_\xi^{(X)} : \Omega^m \to [0,1]$ such that

$$X(\omega) = \sum_{\xi \in \{-1,1\}^m} \lambda_\xi^{(X)}(\omega)\xi$$

almost surely and $\sum_{\xi\in\{-1,1\}^m} \lambda_\xi^{(X)}(\omega) = 1$ for all $\omega \in \Omega$. Hence, it holds that $X_i = \sum_{\xi\in\{-1,1\}^m} \lambda_\xi^{(X)}\xi_i$ almost surely. If we proceed analogously for Y we obtain

$$\begin{aligned}a_{ij} = \mathbb{E}[X_iY_j] &= \mathbb{E}\big[\big(\sum_{\xi\in\{-1,1\}^m} \lambda_\xi^{(X)}\xi_i\big)\big(\sum_{\eta\in\{-1,1\}^n} \lambda_\eta^{(Y)}\eta_j\big)\big] \\ &= \sum_{\xi\in\{-1,1\}^m, \eta\in\{-1,1\}^n} \mathbb{E}\big[\lambda_\xi^{(X)}\lambda_\eta^{(Y)}\big]\xi_i\eta_j.\end{aligned}$$

If we now observe that $\sum_{\xi\in\{-1,1\}^m,\eta\in\{-1,1\}^n} \mathbb{E}[\lambda_\xi^{(X)}\lambda_\eta^{(Y)}] = 1$ it follows that $(a_{ij}) \in \mathrm{conv}\{\xi\eta^T \,|\, \xi \in \{-1,1\}^m, \eta \in \{-1,1\}^n\}$ which finishes the proof. □

Now we are able to count the vertices of $\mathrm{LC}_{m,n}$. Observing that $\xi\eta^T = \tilde{\xi}\tilde{\eta}^T$ if and only if $\xi = \tilde{\xi}$ and $\eta = \tilde{\eta}$ or $\xi = -\tilde{\xi}$ and $\eta = -\tilde{\eta}$ it follows that we have $2^{n+m}/2 = 2^{n+m-1}$ different matrices $\xi\eta$, hence $\mathrm{LC}_{m,n}$ has 2^{n+m-1} vertices. To analyze the facial structure of $\mathrm{LC}_{m,n}$ is rather complicated. However, we will do it later on for $n = m = 2$ and compare it to the set of quantum correlation matrices.

3.2. ***Quantum correlation matrices***

Before we introduce the quantum correlation matrices we will shortly review the framework of nonlocal games. As before, Alice and Bob share a state ρ and get inputs $s \in \mathcal{S}, t \in \mathcal{T}$ and perform measurements $\{F_s^\xi\}_{\xi=\pm1}$, respectively $\{G_t^\eta\}_{\eta=\pm1}$. As we have seen before the probability that their response is (ξ,η) for inputs (s,t) is given by $a_{st} = \mathrm{Tr}\,(\rho F_s^\xi \otimes G_t^\eta)$. Again, we can encode these information in a matrix $A = (a_{st})$. This motivates the following (slightly more general) definition.

Definition 3.2.1. Let $(X_i)_{1\leq i\leq m}$ and $(Y_j)_{1\leq j\leq n}$ be self-adjoint operators on $\mathbb{C}^{d_1}$, respectively $\mathbb{C}^{d_2}$ for some positive integers d_1, d_2, satisfying $\|X_i\|_\infty, \|Y_j\|_\infty \leq 1$. $A = (a_{ij})$ is called *quantum correlation matrix* if there exists a state $\rho \in B^{psd}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$ such that

$$a_{ij} = \mathrm{Tr}\,\rho(X_i \otimes Y_j). \tag{3.2}$$

We will write $\mathrm{QC}_{m,n}$ for the set of all $m \times n$ quantum correlation matrices. With regard to quantum information theory it is interesting to analyze the geometry of $\mathrm{LC}_{m,n}$ and $\mathrm{QC}_{m,n}$. In the following, we will proof a similar result for $\mathrm{QC}_{m,n}$, that is:

Lemma 3.2.2.

$$\mathrm{QC}_{m,n} = \{(\langle x_i, y_j \rangle)_{1\leq i \leq m, 1 \leq j \leq n} \mid x_i, y_j \in \mathbb{R}^{\min\{m,n\}}, |x_i| \leq 1, |y_j| \leq 1\}, \tag{3.3}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product.

The basic idea of the proof is to define an inner product via the right hand sight of 3.2. As before, we will denote the right hand side of equation 3.3 by M.

Proof of $\mathrm{QC}_{m,n} \subset M$. Let $(a_{ij}) \in \mathrm{QC}_{m,n}$. Then there is a sate ρ on a Hilbert space $\mathcal{H} = \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ and Hermitian operators $\{X_1, \dots X_m\} \subset B^H(\mathcal{C})$, $\{Y_1, \dots Y_n\} \subset B^H(\mathcal{H})$ satisfying $\|X_i\|_\infty, \|Y_j\|_\infty \leq 1$ such that $a_{ij} = \mathrm{Tr}\, \rho X_i \otimes Y_j$. We define a positive semi-definite symmetric bilinear form on the space of Hermitian operators on $\mathcal{H}$ by β : $\mathrm{B}^H(\mathcal{H}) \times \mathrm{B}^H(\mathcal{H}) \to \mathbb{R}$ where $\beta(S,T) = \mathrm{Re}(\mathrm{Tr}\, \rho S T)$. We have to check that it indeed satisfies the mentioned properties. Obviously, β is homogeneous in both variables due to the fact that the trace and the real part of a complex number are linear functions and thus homogeneous. We will show additivity for the first variable, the result follows analogously for the second one. It holds

$$\begin{aligned}\beta(S_1 + S_2, T) = \mathrm{Re}(\mathrm{Tr}\, \rho(S_1 + S_2)T) &= \mathrm{Re}(\mathrm{Tr}\, \rho S_1 T) + \mathrm{Re}(\mathrm{Tr}\, \rho S_2 T) \\ &= \beta(S_1, T) + \beta(S_2 + T).\end{aligned}$$

Symmetry follows form

$$\begin{aligned}\beta(S,T) &= \mathrm{ReTr}\,(\rho S T) = \mathrm{ReTr}\,(\rho S T)^* = \mathrm{Re}\ \mathrm{Tr}\,(T^* S^* \rho^*) \\ &= \mathrm{ReTr}\,(\rho^* T^* S^*) = \mathrm{ReTr}\,(\rho T S) = \beta(T,S).\end{aligned}$$

Moreover, since S^*S is a positive semidefinite operator for all complex operators S and ρ is positive semidefinite we obtain $\beta(S,S) = \mathrm{ReTr}\,(\rho S S) = \mathrm{ReTr}\,(\rho S^* S) \geq 0$, so β is positive semidefinite. Following the steps in appendix A.2 we can factorize the kernel and transform β to an inner product on the vector space $B^H(\mathcal{H}) / \ker \beta$.

Now, the crucial point is to notice that $a_{ij} = \mathrm{Tr}\,(\rho X_i Y_j) = \beta(X_i \otimes I, I \otimes Y_j)$. In order to show that these operators coincide with unit vectors with respect to β we have to show $\beta(X \otimes I, X \otimes I), \beta(I \otimes Y, I \otimes Y) \leq 1$ for all $X \in \{X_1, \dots, X_m\}$, $Y \in \{Y_1, \dots, Y_n\}$. Therefore, let $d = d_1 d_2$ and let ρ be a pure state, that is $\rho = |\psi\rangle\langle\psi|$ for $|\psi\rangle = \sum_i^d \lambda_i |\xi_i\rangle \otimes |\eta_i\rangle$, where $\{\xi_1, \dots, \xi_d\} \subset \mathbb{C}^{d_1}$ and $\{\eta_1, \dots, \eta_d\} \subset \mathbb{C}^{d_2}$ are orthonormal sets, i.e. $\langle \xi_i | \xi_j \rangle, \langle \eta_i | \eta_j \rangle = 0$ for $i \neq j$, and $\sum_{i=1}^d \lambda_i^2 = 1$. This decomposition is called *Schmidt decomposition* and a consequence of the singular value decomposition. If we can show the desired property for pure sates we can immediately conclude that it holds for general states since each state is a convex combination of pure states. Writing ρ in this form we get

$$\beta(X \otimes I, X \otimes I) = \mathrm{ReTr}\,(\rho X^2 \otimes I) = \sum_{1 \leq i,j \leq d} \lambda_i \lambda_j \mathrm{Tr}\,((|\xi_i\rangle \langle \xi_j| \otimes |\eta_i\rangle \langle \eta_j|)(X^2 \otimes I))$$

$$= \sum_{i \leq i,j \leq d} \lambda_i^2 \mathrm{Tr}\,(|\xi_i\rangle \langle \xi_i| X^2) \mathrm{Tr}\,(|\eta_i\rangle \langle \eta_i|) + \sum_{1 \leq i \neq j \leq\ d} \lambda_i \lambda_j \mathrm{Tr}\,(|\xi_i\rangle \langle \xi_j| X^2) \mathrm{Tr}\,(|\eta_i\rangle \langle \eta_j|)$$

$$= \sum_{i=1}^d \lambda_i^2 \mathrm{Tr}\,(|\xi_i\rangle \langle \xi_i| X^2).$$

We have to show that $\mathrm{Tr}\left(|\xi_i\rangle\langle\xi_i| X^2\right) = \mathrm{Tr}\left(X^2 |\xi_i\rangle\langle\xi_i|\right) \leq 1$. Note that $1 \geq \|X\|_\infty := \sup_{|y|\leq 1}|Xy|$ implies that $\left|X^2|\xi_i\rangle\right| \leq 1$. So, the problem can be reduced to $|\mathrm{Tr}\, uv^*|^2 \leq 1$ for complex vectors u, v with $|u|, |v| \leq 1$. But this holds since due to the Cauchy-Schwarz inequality $|\mathrm{Tr}\, uv^*|^2 = |\sum u_i \bar{v}_i|^2 \leq |u|^2|v|^2 \leq 1$. If we now identify the operators $X_i \otimes I$ and $I \otimes Y_j$ with vectors (x_i) and (y_j) we have found vectors that almost satisfy the desired properties but they do not have the right dimension and we do not consider the standard scalar product yet. Without loss of generality let $m \leq n$. To obtain the required dimension we will project (y_j) orthogonally onto $\mathrm{span}\{x_1, \dots, x_m\}$. Let $\{a_1, \dots, a_r\}$ be an orthonormal basis of $\mathrm{span}\{x_1, \dots, x_m\}$ with respect to β. The orthogonal projection of y_j is $\pi(y_j) := \sum_{i=1}^r \beta(a_i, y_j)a_i$ and fulfills $\beta(x_i, y_j) = \beta(x_i, \pi(y_j))$. Let x_i and $\pi(y_j)$ admit the descriptions $x_i = \sum_{k=1}^r \alpha_k^{(i)} a_k$ and $\pi(y_j) = \sum_{k=1}^r \gamma_k^{(j)} a_k$ for $\alpha^{(i)}, \gamma^{(j)} \in \mathbb{R}^r$. Then

$$a_{ij} = \beta(x, y) = \beta(x, \pi(y)) = \sum_{1\leq k,l\leq r} \alpha_k^{(i)} \gamma_l^{(j)} \beta(a_k, a_l) = \sum_{k=1}^r \alpha_k^{(i)} \gamma_k^{(j)} = \langle \alpha^{(i)}, \gamma^{(j)} \rangle.$$

Moreover, since $|\alpha^{(i)}| = |x_i| \leq 1$ and $|\gamma^{(j)}| = |y_j| = |\pi(y_j)| \leq 1$ the vectors $(\alpha^{(i)})_{1\leq i\leq m}$ and$(\gamma^{(j)})_{1\leq j\leq n}$ have all properties of the right hand side of 3.3. Additionally, we also proved that we can demand an even lower dimension for the vectors in 3.3, that is $\min\left\{\dim(\mathrm{span}\{(x_i)_{1\leq i\leq m}\}), \dim(\mathrm{span}\{(y_j)_{1lej\leq n}\})\right\}$. $\square$

In order to prove the other inclusion we will use the following proposition.

Proposition 3.2.3. *For all $n \geq 1$ there is a subspace of the $2^n \times 2^n$ Hermitian matrices where every element is the multiple of a unitary matrix.*

Proof. The proof is based on $n-$fold tensor products of the Pauli matrices which are the three matrices

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

together with the 2×2 identity matrix . They are all trace zero unitary Hermitian matrices and anti-commute pair wisely. Moreover, we define the $2^n \times 2^n$ Hermitian matrix

$$\sigma_A^i = I^{\otimes(i-1)} \otimes A \otimes I^{\otimes(n-i)}$$

for $A \in \{X, Y, Z\}$ and where I is the 2×2 identity matrix. The Hermitian property follows directly from the observation $(M \otimes N)^* = M^* \otimes N^*$. Note that σ_A^i and $\sigma_{A'}^j$ anti-commute if $i = j$ and $A \neq A'$ and commute otherwise. We use these operators in order to define

$$U_i = \sigma_X^i \prod_{k=i+1}^n \sigma_Y^k,$$

$$U_{i+n} = \sigma_Z^i \prod_{k=i+1}^n \sigma_Y^k$$

for $i = 1, \ldots, n$. Note that these operators are also trace zero Hermitian matrices and anti-commute for $i \neq j$: for $1 \leq i < j \leq n$ we have

$$\begin{aligned}
U_i U_j = (\sigma_X^i \prod_{k=i+1}^{n} \sigma_Y^k) \cdot (\sigma_X^j \prod_{k=j+1}^{n} \sigma_Y^k) &= \sigma_X^i \sigma_Y^{i+1} \cdots \sigma_Y^j (\sigma_X^j \prod_{k=j+1}^{n} \sigma_Y^k) \sigma_Y^{j+1} \cdots \sigma_Y^n \\
&= -\sigma_X^i \sigma_Y^{i+1} \cdots \sigma_Y^{j-1} (\sigma_X^j \prod_{k=j+1}^{n} \sigma_Y^k) \sigma_Y^{j+1} \cdots \sigma_Y^n \\
&= -(\sigma_X^j \prod_{k=j+1}^{n} \sigma_Y^k)(\sigma_X^i \prod_{k=i+1}^{n} \sigma_Y^k) \\
&= -U_j U_i
\end{aligned}$$

and

$$\begin{aligned}
U_i U_{n+j} = (\sigma_X^i \prod_{k=i+1}^{n} \sigma_Y^k) \cdot (\sigma_Z^j \prod_{k=j+1}^{n} \sigma_Y^k) &= \sigma_X^i \sigma_Y^{i+1} \cdots \sigma_Y^j (\sigma_Z^j \prod_{k=j+1}^{n} \sigma_Y^k) \sigma_Y^{j+1} \cdots \sigma_Y^n \\
&= -\sigma_X^i \sigma_Y^{i+1} \cdots \sigma_Y^{j-1} (\sigma_Z^j \prod_{k=j+1}^{n} \sigma_Y^k) \sigma_Y^{j+1} \cdots \sigma_Y^n \\
&= -(\sigma_Z^j \prod_{k=j+1}^{n} \sigma_Y^k)(\sigma_X^i \prod_{k=i+1}^{n} \sigma_Y^k) \\
&= -U_{n+j} U_i.
\end{aligned}$$

Moreover, the U_i's are unitary since

$$U_i U_i^* = U_i U_i = (\sigma_X^i \prod_{k=i+1}^{n} \sigma_Y^k)(\sigma_X^i \prod_{k=i+1}^{n} \sigma_Y^k) = \sigma_X^i \sigma_X^i \prod_{k=i+1} \sigma_Y^k \sigma_Y^k = I^{\otimes n}.$$

The same holds analogously for U_{n+i}. Now, if we consider a linear combination $X = \sum_{i=1}^{2n} \xi_i U_i$, we can calculate

$$\begin{aligned}
XX^* = XX &= \sum_{i=1}^{2n} \xi_i^2 I + \sum_{1 \leq i \neq j \leq 2n} \xi_i \xi_j U_i U_j \\
&= \sum_{i=1}^{2n} \xi_i^2 I + \sum_{1 \leq i < j \leq 2n} \xi_i \xi_j U_i U_j - \sum_{1 \leq i < j \leq 2n} \xi_i \xi_j U_i U_j \\
&= \sum_{i=1}^{2n} \xi_i^2 I \\
&= |\xi|^2 I.
\end{aligned}$$

So, X is a multiple of a unitary matrix and eventually, we get the desired result by taking the subspace $\operatorname{span}\{U_i \,|\, i = 1, \ldots, 2n\}$.

□

As a byproduct, if we take two linear combinations $X = \sum_{i=1}^{2n} \xi_i U_i$, $Y = \sum_{i=1}^{2n} = \eta_i U_i$ and consider the trace of the product we get

$$\operatorname{Tr}(XY) = \sum_{i=1}^{2n} \xi_i \eta_i \operatorname{Tr} I + \sum_{1 \leq i \neq j \leq \ 2n} \xi_i \eta_j \operatorname{Tr}(U_i U_j) = \sum_{i=1}^{2n} \xi_i \eta_j 2^n = 2^n \cdot \langle \xi, \eta \rangle, \tag{3.4}$$

where we used that the product U_iU_j has trace zero for $i \neq j$. We will use this in order to proceed with the other inclusion of lemma 3.2.2.

Proof of $\mathrm{LC}_{m,n} \supset M$. Let $(x_i)_{1\le i\le m}$, $(y_j)_{1\le j\le n}$ be families of vectors in $\mathbb{R}^{\min\{m,n\}}$ that satisfy $|x_i|,|y_j| \le 1$. Using the notation of the previous proposition's proof we set $X_i = \sum_{k=1}^{\min\{m,n\}} x_i(k)U_k$ and $Y_j^T = \sum_{k=1}^{\min\{m,n\}} y_j(k)U_k$ where the U_k's are $d \times d$ matrices with $d = 2^{\lceil \min\{m,n\}/2\rceil}$ Then, by equation 3.4 it follows that $\mathrm{Tr}\,(X_iY_j^T) = d \cdot \langle x_i, y_j\rangle$ and $\|X_i\|_\infty \le 1$ since $X_iX_i^* = |x_i|^2 I$ and $|x_i|^2 \le 1$. The same holds for Y since $Y_j^TY_j^T = |y_j|^2 I = |y_j|I^T = (Y_j^TY_j^T)^T = Y_jY_j$. Let $|\psi\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^d |ii\rangle$ and $\rho = |\psi\rangle\langle\psi|$. Note that we can write ρ as

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{d}\sum_{1\le k,l\le d} |kk\rangle\langle ll| = \frac{1}{d}\sum_{1\le k,l\le d} |k\rangle\langle l| \otimes |k\rangle\langle l|$$

where $(|k\rangle\langle l|)_{kl} = 1$ and $(|k\rangle\langle l|)_{ij} = 0$ for all $(i,j) \neq (k,l)$. Calculating

$$\begin{aligned}\mathrm{Tr}\,(\rho X_i \otimes Y_j) &= \frac{1}{d}\sum_{1\le k,l\le d} \mathrm{Tr}\,(\,|k\rangle\langle l|\, X_i \otimes |k\rangle\langle l|\, Y_j) = \frac{1}{d}\sum_{1\le k,l\le d} \mathrm{Tr}\,(\,|k\rangle\langle l|\, X_i)\mathrm{Tr}\,(\,|k\rangle\langle l|\, Y_j)\\ &= \frac{1}{d}\mathrm{Tr}\, X_iY_j^T = \langle x_i, y_j\rangle\end{aligned}$$

shows the desired inclusion. □

The previous lemma easily enables us to show that $\mathrm{QC}_{m,n}$ is convex. Consider $(a_{ij}), (\bar{a}_{ij}) \in \mathrm{QC}_{m,n}$ with $a_{ij} = \langle x_i, y_j\rangle$ and $\bar{a}_{ij} = \langle \bar{x}_i, \bar{y}_j\rangle$ for $x_i, y_j, \bar{x}_i, \bar{y}_j \in \mathbb{R}^{\min\{m,n\}}$ such that $|x_i|, |y_j|, |\bar{x}_i|, |\bar{y}_j| \le 1$. For $\lambda \in [0,1]$ we define vectors $\tilde{x}_i := (\sqrt{\lambda}x_i, \sqrt{1-\lambda}\bar{x}_i)$, $\tilde{y}_j := (\sqrt{\lambda}y_j, \sqrt{1-\lambda}\bar{y}_j) \in \mathbb{R}^{2\cdot\min\{m,n\}}$. They satisfy $|\tilde{x}_i| \le \lambda|(x_i,0)| + (1-\lambda)|(0,\bar{x}_i)| \le 1$. Moreover, $\langle \tilde{x}_i, \tilde{y}_j\rangle = \lambda\langle x_i, y_j\rangle + (1-\lambda)\langle \tilde{x}_i, \tilde{y}_j\rangle$. If we proceed in the same fashion as in the proof of lemma 3.2.2 we obtain vectors $\alpha^{(i)}, \gamma^{(j)}$ that satisfy $= \langle \alpha^{(i)}, \gamma^{(j)}\rangle = \langle \tilde{x}_i, \tilde{y}_j\rangle$ and have dimension smaller or equal to $\min\{m,n\}$.

3.3. *The relations between quantum correlation and local correlation matrices*

Using the descriptions of lemmata 3.1.2 and 3.2.2 we can derive some relations between the two sets. Let $\xi\eta^T$ be a vertex of $\mathrm{LC}_{m,n}$. If we just choose $x_i = \xi_i(1,0,\dots,0) \in \mathbb{R}^{\min\{m,n\}}$ and $y_j = \eta_j(1,0,\dots 0) \in \mathbb{R}^{\min\{m,n\}}$ we immediately see that x_i and y_j are unit vectors and $\xi_i\eta_j = \langle x_i, y_j\rangle$. Hence, $\xi\eta^T \in \mathrm{QC}_{m,n}$ and combined with the convexity of $\mathrm{QC}_{m,n}$ we get $\mathrm{LC}_{m,n} \subset \mathrm{QC}_{m,n}$.

However, the inclusion is strict in general. Let us consider $n = m = 2$. For $\mathrm{LC}_{2,2}$ we obtain

$$\mathrm{LC}_{2,2} = \mathrm{conv}\{\pm\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}, \pm\begin{pmatrix}-1 & -1\\ 1 & 1\end{pmatrix}, \pm\begin{pmatrix}-1 & 1\\ -1 & 1\end{pmatrix}, \pm\begin{pmatrix}-1 & 1\\ 1 & -1\end{pmatrix}\}.$$

We can easily see that $\sigma(\begin{pmatrix}-1 & 1\\ 1 & 1\end{pmatrix}) \notin \mathrm{LC}_{2,2}$ for $\sigma \in \Sigma_4$ where for $A = (a_{ij})$ we define $\sigma(A)$ by $(\sigma(A))_{ij} := a_{\sigma(1)\sigma(j)}$. We claim that

$$\mathrm{LC}_{2,2} = \{A \in \mathbb{R}^{2\times 2} \mid -1 \le \mathrm{Tr}\,(AM) \le 1 \text{ for all } M \in \mathcal{K}\}, \tag{3.5}$$

where $\mathcal{K} = \{\frac{1}{2}\sigma(\begin{pmatrix}-1 & 1\\ 1 & 1\end{pmatrix}), \sigma(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}) \mid \sigma \in \{\mathrm{id}, (1\ 2), (1\ 3), (1\ 4)\}\}$. The crucial observation is to note that $\mathrm{LC}_{2,2}$ is affinely isomorphic to the cross polytope scaled by two, i.e. $\mathrm{LC}_{2,2} \cong 2\mathrm{CP}_4 := 2\mathrm{conv}\{\pm e_i \,|\, i = 1, \ldots, 4\}$, where e_i are the vectors of the standard basis of $\mathbb{R}^4$. For example, this can be seen by interpreting the vertices of $\mathrm{LC}_{2,2}$ as elements of $\mathbb{R}^4$ and then apply the linear transformation given by the matrix

$$\begin{pmatrix} -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Since the polar dual of the cross polytope is the hypercube, i.e. $(CP_n)^o = [-1,1]^n$, the face lattice of CP_4 is isomorphic to the opposite lattice of the hypercube's face lattice which implies that the number of facets of CP_4 coincides with the number of vertices of $[-1,1]^4$ which is 2^4. Due to $\mathrm{LC}_{2,2} \cong 2\mathrm{CP}_4$ their face lattices of $\mathrm{LC}_{2,2}$ and CP_4 are isomorphic, so $\mathrm{LC}_{2,2}$ has 2^4 facets as well. Since all constraints of the right hand side of equation 3.5 clearly define non-empty proper faces of $\mathrm{LC}_{2,2}$, it suffices to show that the characterization is a non-redundant hyperplane description of $\mathrm{LC}_{2,2}$, implying that all constraints define facets of $\mathrm{LC}_{2,2}$. But this is indeed true since if we omit for example the constraint $1/2\mathrm{Tr}\,(AM) \leq 1$ for $M = \{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\}$, respectively $M = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$ the matrices $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, respectively $\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$ satisfy all other constraints. Eventually, we have all information to show that $\mathrm{LC}_{2,2}$ is a proper subset of $\mathrm{QC}_{2,2}$. Assume we want to maximize in the direction of the facet induced by $M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Note that M coincides with the matrix $\Sigma_{s,t} = (-1)^{f(s,t)}$ for the CSHS game in section 2.3. Since $\max\{\mathrm{Tr}\,(AM) \,|\, A \in \mathrm{LC}_{2,2}\} = 2$ it suffices to show that there is $A \in \mathrm{QC}_{m,n}$ achieving a better value. This is an alternative way to get the same result as in section 2.3.

For $A \in \mathrm{QC}_{2,2}$ we obtain, by lemma 3.2.2, Cauchy-Schwarz and $|y_i| \leq 1$,

$$\begin{aligned} \mathrm{Tr}\,(AM) &= \langle x_1, y_1 \rangle + \langle x_1, y_2 \rangle + \langle x_2, y_1 \rangle - \langle x_2, y_2 \rangle \\ &= \langle x_1 + x_2, y_1 \rangle + \langle x_1 - x_2, y_2 \rangle \leq |x_1 + x_2||y_1| + |x_1 - x_2||y_2| \\ &\leq |x_1 + x_2| + |x_1 - x_2|. \end{aligned}$$

Observing that for $|x_1|, |x_2| \leq 1$

$$\begin{aligned} &(|x_1 + x_2| + |x_1 - x_2|)^2 \\ &= \langle x_1 + x_2, x_1 + x_2 \rangle + \sqrt{\langle x_1 + x_2, x_1 + x_2 \rangle \langle x_1 - x_2, x_1 - x_2 \rangle} + \langle x_1 - x_2, x_1 - x_2 \rangle \\ &\leq 2|x_1|^2 + 2|x_2|^2 + \sqrt{|x_1|^4 + 2|x_1|^2|x_2|^2 - 4\langle x_1, x_2 \rangle^2 + |x_2|^4} \\ &\leq 2|x_1|^2 + 2|x_2|^2 + \sqrt{4(|x_1|^2 + |x_2|^2)^2} \\ &= 4(|x_1|^2 + |x_2|^2), \end{aligned}$$

we can give a precise upper bound for $\mathrm{Tr}\,(AM)$ by

$$|x_1 + x_2| + |x_1 - x_2| \leq 2\sqrt{|x_1|^2 + |x_2|^2} \leq 2\sqrt{2}.$$

Thus, we just have to find a matrix that satisfies this bound. A possible choice is induced by the vectors $x_1 = x_2 = \frac{1}{\sqrt{2}}(1,1)$ and $y_1 = y_2 = (1,0)$, that is $A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ yielding the value $\mathrm{Tr}\,(AM) = 2\sqrt{2}$. So as we have seen, the inclusion $\mathrm{LC}_{m,n} \subset \mathrm{QC}_{m,n}$ is strict in general. Elements in $\mathrm{QC}_{m,n} \setminus \mathrm{LC}_{m,n}$ are called *non-local*. Generally, linear functionals f mapping an $m \times n$ matrix to a real number and satisfying $f(A) \leq 1$ for all $A \in \mathrm{LC}_{m,n}$ are called *Bell correlation inequalities*. In the case of $f(A) > 1$ for some $A \in \mathrm{QC}_{m,n}$ we talk about *Bell violations*.

4. Grothendieck-Tsirelson Theorem

In the last section we have seen that $\mathrm{QC}_{m,n} \subsetneq \mathrm{LC}_{m,n}$. In this section we will prove the fact that $\mathrm{QC}_{m,n} \subset K\mathrm{LC}_{m,n}$ for some universal constant K independent of m and n. In particular, we can show that the smallest constant K so that the above inclusion holds true the Grothendieck constant $K_G \leq \pi/2 \ln(1+\sqrt{2})$ is.

4.1. *Grothendieck's Inequality*

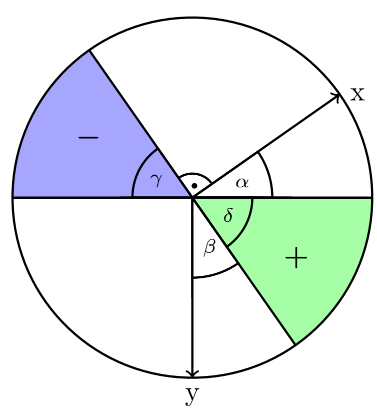
Lemma 4.1.1 (Grothendieck's identity)**.** *Let $x, y \in \mathbb{R}^d$ be unit vectors. Let $r \in \mathbb{R}^d$ be a random unit vector chosen from the probability distribution that is invariant under orthogonal transformations (rotations) on the unit sphere. Then*

i, $\mathbb{P}[\mathrm{sign}(\langle x, r\rangle) \neq \mathrm{sign}(\langle y, r\rangle)] = \frac{\arccos(\langle x,y\rangle)}{\pi}$
ii, $\mathbb{E}[\mathrm{sign}(\langle x, r\rangle)\,\mathrm{sign}(\langle y, r\rangle)] = \frac{2}{\pi}\arcsin(\langle x, y\rangle)$.

Proof. For the proof of i, assume that x and y are linearly dependent. Since both, x and y, are unit vectors, $\arccos(\langle x,y\rangle) = \arccos(1) = 0$ if $x = y$ or $\arccos(\langle x,y\rangle) = \arccos(-1) = \pi$ if $x = -y$.

Conversely assume that x and y are linearly independent, i. e. $\dim(\mathrm{span}\{x, y\}) = 2$. Now project r orthogonally on the plane spanned by x and y. This gives us a vector $s \in \mathrm{span}\{x, y\}$ with $\langle x, r\rangle = \langle x, s\rangle$ and $\langle y, r\rangle = \langle y, s\rangle$. The unit vector $n := s/|s|$ is uniformly distributed on the unit circle that occurs if we consider the intersection of the unit sphere and $\mathrm{span}\{x, y\}$ by the invariance of the probability distribution.

If n lies on the segment of the unit circle induced by the green part, the angle between x and n as well as between y and n is smaller than $\pi/2$, hence $\langle x, n\rangle$ and $\langle y, n\rangle$ are positive. Otherwise, if n lies on the segment of the unit circle induced by the blue part, the angle between both x and n as well as y and n is greater than $3\pi/2$, hence $\langle x, n\rangle$ and $\langle y, n\rangle$ are negative.



Now, if we want to calculate the probability that the signs of the two scalar products disagree, we are interested in the undyed segments of the unit circle. Thus, it is sufficient to calculate the periphery of the undyed segments of the circle. In particular on the unit circle, the angle between two vectors equals the periphery of the segment of the circle between those two vectors.

Because γ and δ are vertical angles they are both equal. Furthermore, α and β have to be equal too, since γ and δ are equal and $\alpha + \delta = \beta + \gamma = \pi/2$. With $\alpha = \arccos(\langle x, y\rangle) - \pi/2$ the first part of Lemma 4.1.1 follows:

$$\mathbb{P}[\mathrm{sign}(\langle x, n\rangle) \neq \mathrm{sign}(\langle y, n\rangle)] = 2\frac{\frac{\pi}{2} + \alpha}{2\pi} = \frac{\arccos(\langle x, y\rangle)}{\pi}.$$

We conclude with the proof of the second part of Lemma 4.1.1:

$$\begin{aligned}
&\mathbb{E}[\operatorname{sign}(\langle x,r\rangle)\operatorname{sign}(\langle y,r\rangle)]\\
&\quad= 1\cdot\mathbb{P}[\operatorname{sign}(\langle x,r\rangle)=\operatorname{sign}(\langle y,r\rangle)] - 1\cdot\mathbb{P}[\operatorname{sign}(\langle x,r\rangle)\neq\operatorname{sign}(\langle y,r\rangle)]\\
&\quad= 1-2\mathbb{P}[\operatorname{sign}(\langle x,r\rangle)\neq\operatorname{sign}(\langle y,r\rangle)]\\
&\quad= 1-2\frac{\arccos(\langle x,y\rangle)}{\pi}\\
&\quad= \frac{2}{\pi}\arcsin(\langle x,y\rangle),
\end{aligned}$$

because $\arcsin(t)+\arccos(t)=\pi/2$. □

Lemma 4.1.2 (Krivine's trick)**.** *Let $x_1,\dots,x_m,y_1,\dots,y_n\in S^{m+n-1}$ be given. Furthermore, let $r\in S^{n+m-1}$ be a random unit vector chosen form the probability distribution that is invariant under orthogonal transformations (rotations) on the unit sphere. Then there are $x'_1,\dots,x'_m,y'_1,\dots,y'_n\in S^{m+n-1}$ so that*

$$\mathbb{E}[\operatorname{sign}(\langle x'_i,r\rangle)\operatorname{sign}(\langle y'_j,r\rangle)] = \beta\langle x_i,y_j\rangle, \tag{4.1}$$

with $\beta=\frac{2}{\pi}\ln(1+\sqrt{2})$.

For the proof of 4.1.2 we need to use the k-th tensor product of $\mathbb{R}^n$. The $\mathbb{R}^n$ is an n-dimensional Euclidean space with inner product $\langle\cdot,\cdot\rangle$ and orthonormal basis $e_1,\dots,e_n$. The *k-th tensor product of* $\mathbb{R}^n$ is denoted by $(\mathbb{R}^n)^{\otimes k}$ and it is a Euclidean vector space of dimension n^k with orthonormal basis $e_{i_1}\otimes e_{i_2}\otimes\cdots\otimes e_{i_k}$, $i_l\in\{1,\dots,n\}$. In particular

$$\begin{aligned}
\langle e_{i_1}\otimes\cdots\otimes e_{i_k}, e_{j_1}\otimes\cdots\otimes e_{j_k}\rangle &= \prod_{l=1}^{k}\langle e_{i_l},e_{j_l}\rangle\\
&= \begin{cases}1 & \text{, if } i_l=j_l \text{ for all } l=1,\dots,n,\\ 0 & \text{, otherwise,}\end{cases}
\end{aligned}\tag{4.2}$$

and for $v\in\mathbb{R}^n$ with $v=v_1e_1+\cdots+v_ne_n$ we define $v^{\otimes k}\in(\mathbb{R}^n)^{\otimes k}$ by

$$v^{\otimes k} := (v_1e_1+\cdots+v_ne_n)\otimes\cdots\otimes(v_1e_1+\cdots+v_ne_n) = \sum_{i_1,\dots,i_k} v_{i_1}\cdots v_{i_k}e_{i_1}\otimes\cdots\otimes e_{i_k}, \tag{4.3}$$

where the last equality follows by the distributive law (identity ii, of the tensor product). Thus, for $v,w\in\mathbb{R}^n$

$$\begin{aligned}
\langle v^{\otimes k},w^{\otimes k}\rangle &= \sum_{i_1,\dots,i_k} v_{i_1}\cdots v_{i_k}\sum_{j_1,\dots,j_k} w_{j_1}\cdots w_{j_k}\langle e_{i_1}\otimes\cdots\otimes e_{i_k}, e_{j_1}\otimes\cdots\otimes e_{j_k}\rangle\\
&\overset{(4.2)}{=} \sum_{i_1,\dots,i_k} v_{i_1}\cdots v_{i_k}w_{i_1}\cdots w_{i_k}\\
&= (\sum_{i=1}^{n} v_iw_i)^k = \langle v,w\rangle^k.
\end{aligned}\tag{4.4}$$

Proof. Define the function $E:[-1,+1]\to[-1,+1]$ by $E(t)=\frac{2}{\pi}\arcsin(t)$. Due to Grothendieck's identity (Lemma 4.1.1):

$$\begin{aligned}
E(\langle x'_i,y'_j\rangle) &= \mathbb{E}[\operatorname{sign}(\langle x'_i,r\rangle)\operatorname{sign}(\langle y'_j,r\rangle)]\\
&\overset{!}{=} \beta\langle x_i,y_j\rangle.
\end{aligned}$$

Idea: To find β, x_i', y_j' we invert E:

$$\langle x_i', y_j' \rangle = E^{-1}(\beta \langle x_i, y_j \rangle)$$

with

$$\begin{aligned} E^{-1}(t) &= \sin(\pi/2 \cdot t) \\ &= \sum_{k=0}^{\infty} \underbrace{\frac{(-1)^{2k+1}}{(2k+1)!} \left(\frac{\pi}{2}\right)^{2k+1}}_{=: g_{2k+1}} t^{2k+1} \end{aligned}$$

which is valid for all $t \in [-1, +1]$.

Define the infinite-dimensional Hilbert space

$$H = \bigoplus_{r=0}^{\infty} (\mathbb{R}^{m+n})^{\otimes 2k+1}. \tag{4.5}$$

Define $\tilde{x}_i, \tilde{y}_j \in H$, $i = 1, \ldots, m, j = 1, \ldots, n$ componentwise:

$$(\tilde{x}_i)_k = \operatorname{sign}(g_{2k+1}) \sqrt{|g_{2k+1}| \beta^{2k+1}}\, x_i^{\otimes 2k+1} \tag{4.6}$$

$$(\tilde{y}_j)_k = \sqrt{|g_{2k+1}| \beta^{2k+1}}\, y_j^{\otimes 2k+1} \tag{4.7}$$

Then

$$\begin{aligned} \langle \tilde{x}_i, \tilde{y}_j \rangle &= \sum_{k=0}^{\infty} g_{2k+1} \beta^{2k+1} \langle x_i^{\otimes 2k+1}, y_j^{\otimes 2k+1} \rangle \\ &\overset{(4.4)}{=} \sum_{k=0}^{\infty} g_{2k+1} \beta^{2k+1} \langle x_i, y_j \rangle^{2k+1} \\ &= E^{-1}(\beta \langle x_i, y_j \rangle). \end{aligned}$$

Hence, β is defined by the condition that the vectors $\tilde{x}_1, \ldots, \tilde{x}_m, \tilde{y}_1, \ldots, \tilde{y}_n$ are unit vectors, that is

$$1 = \langle \tilde{x}_i, \tilde{x}_i \rangle = \langle \tilde{y}_j, \tilde{y}_j \rangle = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} \left(\frac{\pi}{2}\right)^{2k+1} \beta^{2k+1} = \sinh(\frac{\pi}{2}\beta).$$

Consequently

$$\beta = \frac{2}{\pi} \operatorname{arcsinh}(1) = \frac{2}{\pi} \ln(1 + \sqrt{2}),$$

since $\operatorname{arcsinh}(t) = \ln(t + \sqrt{t^2 + 1})$.

The only thing that is left to prove, is that the solution of the maximization problem yields vectors $x_1', \ldots, x_m', y_1', \ldots, y_n' \in S^{m+n-1}$, since our vectors $\tilde{x}_1, \ldots, \tilde{x}_m, \tilde{y}_1, \ldots, \tilde{y}_n$ are infinite-dimensional. For this reason consider the real matrix $G \in \mathbb{R}^{(m+n)\times(m+n)}$ given by

$$G = \begin{pmatrix} \langle \tilde{x}_1, \tilde{x}_1 \rangle & \cdots & \langle \tilde{x}_1, \tilde{x}_m \rangle & \langle \tilde{x}_1, \tilde{y}_1 \rangle & \cdots & \langle \tilde{x}_1, \tilde{y}_n \rangle \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \langle \tilde{x}_m, \tilde{x}_1 \rangle & \cdots & \langle \tilde{x}_m, \tilde{x}_m \rangle & \langle \tilde{x}_m, \tilde{y}_1 \rangle & \cdots & \langle \tilde{x}_m, \tilde{y}_n \rangle \\ \langle \tilde{y}_1, \tilde{x}_1 \rangle & \cdots & \langle \tilde{y}_1, \tilde{x}_m \rangle & \langle \tilde{y}_1, \tilde{y}_1 \rangle & \cdots & \langle \tilde{y}_1, \tilde{y}_n \rangle \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \langle \tilde{y}_n, \tilde{x}_1 \rangle & \cdots & \langle \tilde{y}_n, \tilde{x}_m \rangle & \langle \tilde{y}_n, \tilde{y}_1 \rangle & \cdots & \langle \tilde{y}_n, \tilde{y}_n \rangle \end{pmatrix} \tag{4.8}$$

called *Gram matrix.* Let $z \coloneqq (\tilde{x}_1, \dots, \tilde{x}_m, \tilde{y}_1, \dots, \tilde{y}_n)$. By the linearity of the scalar product

$$v^\top G v = \sum_{i,j} v_i G_{ij} v_j = \sum_{i,j} v_i \langle z_i, z_j \rangle v_j = \langle \sum_i v_i z_i, \sum_j v_j z_j \rangle > 0$$

for some $v \in \mathbb{R}^{m+n}$, $v \neq 0$. Hence, G is positive definite and symmetric, thus, G can be diagonalized by an orthogonal matrix. This means that there is a decomposition $G = Q\Lambda Q^\top$ with Q a real orthogonal matrix with columns that are the eigenvectors of G and Λ a real and diagonal matrix having the eigenvalues of G on the diagonal. Since the eigenvalues of a positive definite matrix are positive, $\Lambda = \Lambda^{1/2}\Lambda^{1/2}$. Thus,

$$G = (Q\Lambda^{1/2})(Q\Lambda^{1/2})^\top$$

and due to the symmetry of G likewise

$$G = (Q\Lambda^{1/2})^\top (Q\Lambda^{1/2}).$$

$A \coloneqq Q\Lambda^{1/2}$ is a real $(m+n) \times (m+n)$ matrix and its columns are the vectors we were looking for. ☐

Definition 4.1.3. For $M \in \mathbb{R}^{m \times n}$ define the quadratic program

$$\begin{aligned} \|M\|_{\infty\to 1} &= \max\left\{ \sum_{i=1}^m \sum_{j=1}^n M_{ij}\xi_i\eta_j : \xi_i^2 = 1, i = 1, \dots, m, \eta_j^2 = 1, j = 1, \dots, n \right\} \\ &= \max\left\{ \operatorname{Tr} M\eta\xi^\top : \xi \in \{-1,1\}^m, \eta \in \{-1,1\}^n \right\}. \end{aligned} \tag{4.9}$$

Computing $\|M\|_{\infty\to 1}$ is **NP**-hard (convexe optimization; reduce MaxCut in polynomial time to $\|M\|_{\infty\to 1}$).

Definition 4.1.4. The SDP relaxation of $\|M\|_{\infty\to 1}$ is given via:

$$\begin{aligned} \operatorname{sdp}_{\infty\to 1}(M) = \max & \sum_{i=1}^m \sum_{j=1}^n M_{ij} \langle x_i, y_j \rangle \\ & x_i, y_j \in \mathbb{R}^{m+n} \\ & |x_i| = 1, i = 1, \dots, m \\ & |y_j| = 1, j = 1, \dots, n \end{aligned}$$

Theorem 4.1.5 (Grothendieck's inequality)**.** *There exists a constant K such that for all $M \in \mathbb{R}^{m \times n}$:*

$$\|M\|_{\infty\to 1} \le \operatorname{sdp}_{\infty\to 1}(M) \le K \|M\|_{\infty\to 1}. \tag{4.10}$$

Proof. We will use the following approximation algorithm with randomized rounding:

Algorithm 1: Approximation algorithm with randomized rounding for $\|M\|_{\infty\to 1}$

1. Solve $\operatorname{sdp}_{\infty\to 1}(M)$. Let $x_1, \dots, x_m, y_1, \dots, y_n \in S^{m+n-1}$ be the optimal unit vectors
2. Apply Krivine's trick (Lemma 4.1.2) and use vectors x_i, y_j to create new unit vectors $x'_1, \dots, x'_m, y'_1, \dots, y'_n \in S^{m+n-1}$.
3. Choose $r \in S^{m+n-1}$ randomly

4. Round: $\xi_i = \operatorname{sign}(\langle x_i', r\rangle)$
   $\eta_j = \operatorname{sign}(\langle y_j', r\rangle)$

Expected quality of the outcome:

$$\begin{aligned}
\|M\|_{\infty\to 1} \geq \mathbb{E}\left[\sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}\xi_i\eta_j\right] \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}\mathbb{E}[\operatorname{sign}(\langle x_i', r\rangle)\operatorname{sign}(\langle y_j', r\rangle)] \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}\beta\langle x_i, y_j\rangle \\
&\overset{(4.1)}{=} \beta\,\mathrm{sdp}_{\infty\to 1}(M),
\end{aligned}$$

where $\beta = \frac{2\ln(1+\sqrt{2})}{\pi}$, thus $K \leq \beta^{-1}$. $\square$

4.2. *Tsirelson's Theorem*

Theorem 4.2.1 (Tsirelson)**.** *(Hard direction) For all positive integers n, r and any $x_1, \ldots, x_n,\ y_1, \ldots, y_n \in S^{r-1}$, there exists a positive integer $d \coloneqq d(r)$, a unit vector $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ and $\{-1,1\}$-observables $F_1, \ldots, F_n, G_1, \ldots, G_n \in \mathbb{C}^{d\times d}$, such that for every $i, j \in \{1, \ldots, n\}$, we have*

$$\langle\psi|F_i \otimes G_j|\psi\rangle = \langle x_i, y_j\rangle. \tag{4.11}$$

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

(Easy direction) Conversely, for all positive integers n, d, unit vectors $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ and $\{-1,1\}$-observables $F_1, \ldots, F_n, G_1, \ldots, G_n \in \mathbb{C}^{d\times d}$, there exist a positive integer $r \coloneqq r(d)$ and $x_1, \ldots, x_n, y_1, \ldots, y_n \in S^{r-1}$ such that for every $i, j \in \{1, \ldots, n\}$, we have

$$\langle x_i, y_j\rangle = \langle\psi|F_i \otimes G_j|\psi\rangle. \tag{4.12}$$

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

Proof. We start by proving the hard direction.

Reminder: Pauli matrices

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

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

$$S_{2l+1} = Z^{\otimes(l-1)} \otimes X \otimes I^{\otimes(\lceil r/2\rceil - l)}, \tag{4.13}$$

$$S_{2l} = Z^{\otimes(l-1)} \otimes Y \otimes I^{\otimes(\lceil r/2\rceil - l)}. \tag{4.14}$$

We will prove, that just as the Pauli matrices the Clifford matrices square to the identity matrix (of size d-by-d) and pair-wise anti-commute.

Let $l \in \{1, \ldots, \lceil r/2\rceil\}$. Then

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

by identity *iii*, of the tensor product and the properties of the Pauli matrices. The same applies for S_{2l}. Thus $d \leq 2^{\lceil r/2 \rceil}$.

Furthermore, let $k, l \in \{1, \dots, \lceil r/2 \rceil\}$, $S_k \in \{S_{2k+1}, S_{2k}\}$, $S_l \in \{S_{2l+1}, S_{2l}\}$ and corresponding $P_k, P_l \in \{X, Y\}$. W. l. o. g. let $k < l$, then

$$\begin{aligned}
S_k S_l &= (Z^{\otimes(k-1)} \otimes P_k \otimes I^{\otimes(\lceil r/2 \rceil - k)})(Z^{\otimes(l-1)} \otimes P_l \otimes I^{\otimes(\lceil r/2 \rceil - l)}) \\
&= (Z^2)^{\otimes(k-1)} \otimes (P_k Z) \otimes (IZ)^{\otimes(l-k-1)} \otimes (IP_l) \otimes I^{\otimes \lceil r/2 \rceil - l} \\
&= I^{\otimes(k-1)} \otimes (P_k Z) \otimes Z^{\otimes(l-k-1)} \otimes P_l \otimes I^{\otimes \lceil r/2 \rceil - l} \\
&= I^{\otimes(k-1)} \otimes (-Z P_k) \otimes Z^{\otimes(l-k-1)} \otimes P_l \otimes I^{\otimes \lceil r/2 \rceil - l} \\
&= -S_l S_k,
\end{aligned}$$

since P_k is a Pauli matrix.

Additionally, for every $k \neq l$, we have $\operatorname{Tr} S_k S_l = d$, if $k = l$, and $\operatorname{Tr} S_k S_l = 0$, if $k \neq l$, since $\operatorname{Tr} A \otimes B = \operatorname{Tr} A \operatorname{Tr} B$ and at least $\operatorname{Tr} ZX = \operatorname{Tr} ZY = 0$.

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

$$F_i = \sum_{k=1}^{r} (x_i)_k S_k, \tag{4.15}$$

$$G_j = \sum_{k=1}^{r} (y_j)_k S_k^\top. \tag{4.16}$$

We will start by proving that the matrices $F_1, \dots, F_n, G_1, \dots, G_n$ are $\{-1, 1\}$-observables. For this reason it is sufficient to show that $F_i^2 = G_j^2 = I$ for each $i, j \in \{1, \dots, n\}$, as this implies that the matrices have eigenvalues in $\{-1, 1\}$. To this end, consider the expansion of F_i^2,

$$\begin{aligned}
F_i^2 &= \sum_{k,l=1}^{r} (x_i)_k (x_i)_l S_k S_l \\
&= \sum_{k=l} (x_i)_k (x_i)_l \underbrace{S_k S_l}_{=I} + \underbrace{\sum_{k<l} (x_i)_k (x_i)_l S_k S_l}_{=\sum_{k>l} (x_i)_l (x_i)_k S_l S_k} + \sum_{k>l} (x_i)_k (x_i)_l S_k S_l \\
&= \underbrace{\langle x_i, x_i \rangle}_{=1} I + \sum_{k>l} (x_i)_k (x_i)_l \underbrace{(S_k S_l + S_l S_k)}_{=0} \\
&= I,
\end{aligned}$$

since x_i is a unit vector and the Clifford matrices pair-wise anti-commute. Of course, the same argument works for G_j.

We will continue with proving that for every $i, j \in \{1, \dots, n\}$, we have $\operatorname{Tr} F_i G_j^\top / d = \langle x_i, y_j \rangle$.

Fix $i, j \in \{1, \dots, n\}$. As before, consider the expansion of the product $F_i G_j^\top$,

$$F_i G_j^\top = \sum_{k,l=1}^{r} (x_i)_k (y_j)_l S_k S_l. \tag{4.17}$$

Then,

$$\operatorname{Tr} F_i G_j^\top = \operatorname{Tr} \sum_{k,l=1}^{r} (x_i)_k (y_j)_l S_k S_l = \sum_{k,l=1}^{r} (x_i)_k (y_j)_l \underbrace{\operatorname{Tr} S_k S_l}_{=0,\text{ if } k\neq l} = \sum_{k=1}^{r} (x_i)_k (y_j)_k \underbrace{\operatorname{Tr} S_k^2}_{=d}$$
$$= d\langle x_i, y_j\rangle$$

We now consider the expansion of $\operatorname{Tr} F_i G_j^\top / d$. Let $\{|1\rangle, \dots, |d\rangle\} \subseteq \mathbb{C}^d$ be an orthonormal basis for $\mathbb{C}^d$. Let

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{s=1}^{d} |s\rangle \otimes |s\rangle. \tag{4.18}$$

We have

$$\begin{aligned}
\langle\psi|F_i \otimes G_j|\psi\rangle &= \frac{1}{d}\sum_{s,t=1}^{d} \langle s| \otimes \langle s|F_i \otimes G_j|t\rangle \otimes |t\rangle \\
&= \frac{1}{d}\sum_{s,t=1}^{d} \langle s|F_i|t\rangle\langle s|G_j|t\rangle \\
&= \frac{1}{d}\operatorname{Tr} F_i G_j^\top \\
&= \langle x_i, y_j\rangle.
\end{aligned}$$

(Easy direction) Note that since $|\psi\rangle$ has norm 1 and the observables F_i and G_j are unitary operators, thus $F_i \otimes I$ and $I \otimes G_j$ are unitary operators which preserve inner products between vectors, $F_i \otimes I|\psi\rangle$ 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

$$\begin{aligned}
\langle F_i \otimes I|\psi\rangle, I \otimes G_j|\psi\rangle\rangle &= (F_i \otimes I|\psi\rangle)^*(I \otimes G_j|\psi\rangle) = (\underbrace{|\psi\rangle^*}_{=\langle\psi|}\underbrace{F_i^*}_{=F_i} \otimes I^*)(I \otimes G_j|\psi\rangle) \\
&= \langle\psi|(F_i \otimes I)(I \otimes G_j)|\psi\rangle = \langle\psi|F_i \otimes G_j|\psi\rangle \\
&= \operatorname{Tr} |\psi\rangle\langle\psi|F_i \otimes G_j
\end{aligned}$$

is a real number. Thus,

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

Define $x_i, y_j \in S^{2d^2}$ by

$$x_i = \begin{pmatrix} \operatorname{Re}(F_i \otimes I|\psi\rangle) \\ \operatorname{Im}(F_i \otimes I|\psi\rangle) \end{pmatrix}, \qquad y_j = \begin{pmatrix} \operatorname{Re}(I \otimes G_j|\psi\rangle) \\ \operatorname{Im}(I \otimes G_j|\psi\rangle) \end{pmatrix}, \tag{4.19}$$

thus $r \leq 2d^2$.

Then,

$$\begin{aligned}
\langle x_i, y_j\rangle &= \langle \operatorname{Re}(F_i \otimes I|\psi\rangle), \operatorname{Re}(I \otimes G_j|\psi\rangle)\rangle + \langle \operatorname{Im}(F_i \otimes I|\psi\rangle), \operatorname{Im}(I \otimes G_j|\psi\rangle)\rangle \\
&= \operatorname{Re}(\langle\psi|F_i \otimes G_j|\psi\rangle) \\
&= \langle\psi|F_i \otimes G_j|\psi\rangle,
\end{aligned}$$

as desired. $\square$

Since

$$\langle\psi|F_i \otimes G_j|\psi\rangle = \operatorname{Tr} \underbrace{|\psi\rangle\langle\psi|}_{=:\rho} F_i \otimes G_j, \tag{4.20}$$

where ρ is a pure state, this result can be considered similarly to the result of Lemma 3.2.2 (in fact it is a direct consequence) with the difference that $(F_i)_{1\leq i\leq n}$ and $(G_j)_{1\leq i\leq n}$ are $\{-1,1\}$-observables in contrast to $(X_i)_{1\leq i\leq m}$ and $(Y_j)_{1\leq i\leq n}$ self-adjoint with $\|X_i\|_\infty$, $\|Y_j\|_\infty \leq 1$, x_i and y_j are unit vectors and ρ is pure. But in fact, the following equality holds true:

$$\mathrm{QC}_{m,n} = \{(\langle x_i, y_j\rangle)_{1\leq i\leq m, 1\leq j\leq n} \mid x_i, y_j \in \mathbb{R}^{\min\{m,n\}}, |x_i| = 1, |y_j| = 1\} \tag{4.21}$$

Consider $(x_i)_{1\leq i\leq m}, (y_j)_{1\leq j\leq n}$ with $|x_i|, |y_j| \leq 1$. If we can find unit vectors $u, v \in \{x_i \mid 1 \leq i \leq m\}^\perp \cap \{y_j \mid 1 \leq j \leq n\}^\perp$ such that $u \perp v$, we can define unit vectors

$$x_i' = x_i + \sqrt{1 - |x_i|^2}u$$
$$y_j' = y_j + \sqrt{1 - |x_i|^2}v$$

with

$$\langle x_i', y_j'\rangle = \langle x_i, y_j\rangle.$$

If we can not find such vectors u and v we can increase the dimension of the vectors x_i', y_j' we are looking for as follows: Find $(\min\{m,n\}+2)$-dimensional unitary vectors $u, v \in \left\{\begin{pmatrix} x_i \\ 0 \\ 0 \end{pmatrix} \middle| 1 \leq i \leq m\right\}^\perp \cap \left\{\begin{pmatrix} y_j \\ 0 \\ 0 \end{pmatrix} \middle| 1 \leq j \leq n\right\}^\perp$ such that $u \perp v$ (by adding these two dimensions we ensure that we can find u and v also if for example $(x_i)_{1\leq i\leq m}$ or $(y_j)_{1\leq j\leq n}$ form a basis of the $\mathbb{R}^{\min\{m,n\}}$) and define

$$x_i' = \begin{pmatrix} x_i \\ \sqrt{1 - |x_i|^2}u \end{pmatrix}, \qquad y_j' = \begin{pmatrix} y_j \\ \sqrt{1 - |y_j|^2}v \end{pmatrix}.$$

Then,

$$\langle x_i', y_j'\rangle = \langle x_i, y_j\rangle$$

and $|x_i'|, \left|y_j'\right| = 1$. To obtain the right dimension we may a posteriori project the vectors $(x_i')_{1\leq i\leq m}, (y_j)_{1\leq j\leq n}$ onto $\operatorname{span}\{x_i : 1 \leq i \leq n\}$ or $\operatorname{span}\{y_j : 1 \leq j \leq m\}$ according to whether $m \leq n$ or $m > n$.

Furthermore, we can infer with this alternative description of $\mathrm{QC}_{m,n}$ that we can choose X_i and Y_j to be $\{-1,1\}$-observables and ρ to be a pure state and still generate a quantum correlation matrix by setting $a_{ij} = \operatorname{Tr} \rho(X_i \otimes Y_j)$.

4.3. *Grothendieck-Tsirelson*

Theorem 4.3.1 (Grothendieck-Tsirelson)**.** *There exists an absolute constant $K \geq 1$ such that, for any positive integers m, n, the following three equivalent conditions hold:*

(1) We have the inclusion

$$\mathrm{QC}_{m,n} \subset K\mathrm{LC}_{m,n}. \tag{4.22}$$

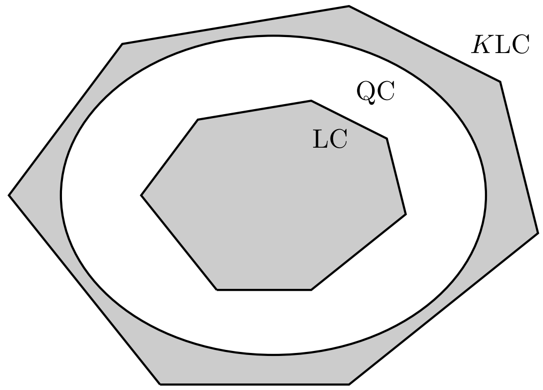


FIGURE 1. Visualization

(2) For any $M \in \mathbb{R}^{m\times n}$ and for any ρ, X_i, Y_j verifying the conditions of Definition 4.2.1 we have

$$\sum_{i,j} M_{ij} \operatorname{Tr} \rho(X_i \otimes Y_j) \leq K \max_{\xi\in\{-1,1\}^m, \eta\in\{-1,1\}^n} \sum_{i,j} M_{ij}\xi_i\eta_j \tag{4.23}$$

$$\Leftrightarrow \qquad \operatorname{Tr} MA^\top \leq \max_{\xi\in\{-1,1\}^m, \eta\in\{-1,1\}^n} \operatorname{Tr} M(\xi\eta^\top)^\top. \tag{4.24}$$

(3) For any $M \in \mathbb{R}^{m\times n}$ and for any (real) Hilbert space vectors x_i, y_j with $|x_i| \leq 1$, $|y_j| \leq 1$ we have

$$\sum_{i,j} M_{i,j}\langle x_i, y_j\rangle \leq K \max_{\xi\in\{-1,1\}^m, \eta\in\{-1,1\}^n} \operatorname{Tr} \xi^\top M\eta. \tag{4.25}$$

Proof. Since (4.25) is a direct consequence of Grothendieck's inequality the only thing left to prove is the equivalence between (1)-(3). The equivalence of (3) and (2) (the Tsirelson's bound) is a consequence of either the proof of Lemma 3.2.2 or Tsirelsons Theorem (Theorem 4.2.1).

For the implication of (2) by (1) consider figure 1. Since QC $\subset K$LC maximizing over the elements in KLC will result in a better outcome than maximizing over the elements in QC in the direction of M. The direction from (1) to (2) we will prove by contraposition. Assume that there is a quantum correlation matrix $A \in \mathrm{QC}_{m,n}$ such that $A \notin K\mathrm{LC}_{m,n}$. By the hyperplane separation theorem we can separate A from $K\mathrm{LC}_{m,n}$ by a linear functional since $\mathrm{LC}_{m,n}$ is convex. That is, we can find a matrix $M \in \mathbb{R}^{m\times n}$ and a real number c such that

$$\operatorname{Tr} MA^\top \geq c \text{ and } \operatorname{Tr} M(\xi\eta^\top)^\top \leq c$$

for all $\xi \in \{-1,1\}^m, \eta \in \{-1,1\}^n$. □

Appendix A. Bilinear forms and inner products

A.1. ***Basic definitions***

Let V and W be two vector spaces and k a field. A *bilinear form* is a map $\beta: V \times W \to k$ which is linear in both variables, that is

i) $\beta(v_1 + v_2, w) = \beta(v_1, w) + \beta(v_2, w)$
ii) $\beta(\lambda v, w) = \lambda\beta(v, w)$
iii) $\beta(v, w_1 + w_2) = \beta(v, w_1) + \beta(v, w_2)$
iv) $\beta(v, \lambda w) = \lambda\beta(v, w)$

for all $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$, $\lambda \in k$. If $V = W$, we call β *symmetric* if $\beta(v, w) = \beta(w, v)$, *positive semidefinite* if $\beta(v, v) \geq 0$ and *positive definite* if β is positive semidefinite and $\beta(v, v) = 0$ implies that $v = 0$. If $\beta: V \times V \to k$ is a symmetric positive definite bilinear form it is called an *inner product*. Note that if H is a positive semidefinite operator then $\beta(v, w) = v^T H \bar{w}$ defines an positive semidefinite bilinear form and an inner product if H is positive definite. Conversely, each positive semidefinite bilinear form β can be written as $\beta(v, w) = v^T H \bar{w}$ for an Hermitian operator H.

A.2. ***How to derive an inner product from a symmetric positive semidefinite bilinear form***

Suppose we have a $k-$vector space V equipped with symmetric positive semidefinite bilinear form $\beta: V \times V \to k$. We want to derive a vector space U that is equipped with an inner product which is induced by β. The idea is to to consider the quotient space $U := V/\ker\beta$ where $\ker\beta = \{v \in V \mid \beta(v, w) = 0 \text{ for all } w \in V\}$. Note that the Cauchy-Schwartz inequality $\beta(v, w)^2 \leq \beta(v, v)\beta(w, w)$ implies that $\ker\beta = \{v \in V \mid \beta(v, v) = 0\}$. We define $\tilde{\beta}: U \times U \to k$ by $\tilde{\beta}([v], [w]) = \beta(v, w)$ where $[v] = v + \ker\beta, [w] = w + \ker\beta$.

We have to show that $\tilde{\beta}$ is well-defined. Therefore, let $[v] = [v']$, so $v' - v \in \ker\beta$. For an arbitrary $[w] \in U$ yields

$$\beta([v], [w]) = \beta(v, w) = \beta(v, w) + \beta(v' - v, w) = \beta(v', w) = \tilde{\beta}([v'], [w]).$$

The symmetry of β combined with the observation above ensures $\tilde{\beta}([v], [w]) = \tilde{\beta}([v], [w'])$ for $[w] = [w']$.

Finally, we get the following equivalence relations:

$$\tilde{\beta}([v], [v]) = 0 \Leftrightarrow \beta(v, v) = 0 \Leftrightarrow v \in \ker\beta \Leftrightarrow [v] = \ker\beta,$$

which implies that $\tilde{\beta}$ defines an inner product on U.

We are also able to analyze the structure of U without big effort. Let $\{v_1, \ldots, v_k, \ldots v_n\}$ be a basis for V such that $\ker\beta = \operatorname{span}\{v_1, \ldots v_k\}$. So, if we take two elements $v = \sum_{i=1}^k a_i v_i + \sum_{i=k+1}^n a_i v_i$ and $w = \sum_{i=1}^k b_i v_i + \sum_{i=k+1}^n b_i v_i$ then

$$[v] = [w] \Leftrightarrow v - w \in \operatorname{span}\{v_1, ..., v_k\} \Leftrightarrow (a_{k+1}, \ldots a_n) = (b_{k+1}, \ldots b_n).$$

Hence, we can deduce that $U \cong \operatorname{span}\{v_{k+1}, ..., v_n\}$. More generally, if $V = V_1 \oplus V_2$, then $V/V_1 \cong V_2$.

References

[Au] G. Aubrun; S. J. Szarek. *Alice and Bob Meet Banach: The Interface of Asympototic Geometry Analysis and Quantum Information Theory.* Graduate Studies in Mathematics, American Mathematical Society, 2017.

[Br] J. Briët. *Grothendieck Inequalities, Nonlocal Games and Optimization.* Institute for Logic, Language and Computation, Amsterdam, 2011.