SEMINAR PAPER
UNGLEICHUNGEN UND ÄHNLICH VERWIRRENDE KONZEPTE

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. Einleitung 1
2. Quantum Grundlagen 2
2.1. Basic definitions 2
2.2. Tensor products and Dirac notation 2
2.3. States and measurements 3
3. Nonlocal games 5
3.1. Classical and entangled strategies 5
3.2. Two player XOR games 6
3.3. The CHSH game 7
4. Local and Quantum correlation matrices 9
4.1. Local Correlation matrices 9
4.2. Quantum correlation matrices 11
5. Grothendieck Inequality 16
Appendix 16
5.1. How to derive an inner product from a symmetric positive semidefinite bilinear form 17

1. **Einleitung**

Einleitung. Es geht um Grothendieck Ungleichung und so weiter.

2. Quantum Grundlagen

2.1. *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. 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, - iv, 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 $n \times n$ matrices is denoted by $\mathcal{S}_n^+$.

2.2. *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 $\mathcal{R}$ and $\mathcal{C}$ and think of the matrix as a map from $\mathcal{R} \times \mathcal{C} \to \mathbb{C}$. For two complex matrices $A : \mathcal{R}_1 \times \mathcal{C}_1 \to \mathbb{C}$ and $B : \mathcal{R}_2 \times \mathcal{C}_2 \to \mathbb{C}$ their tensor product is the matrix $A \otimes B : (\mathcal{R}_1 \times \mathcal{R}_2) \times (\mathcal{C}_1 \times \mathcal{C}_2) \to \mathbb{C}$ 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.

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

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 $\mathcal{H}$ be an inner product space. Endow $\mathcal{H}$ with a norm $||x|| = \sqrt{\langle x, x\rangle}$ and a metric $d(x, y) = ||x - y||$. If every Cauchy sequence in $\mathcal{H}$ converges to an element in $\mathcal{H}$, i.e. $\mathcal{H}$ is complete, then $\mathcal{H}$ is a Hilbert space.

2.3. ***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 $n \times n$ 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 $\mathcal{A}$. As before we will have an n-dimensional quantum system X and a measurement on X in state ρ with outcomes in $\mathcal{A}$ is defined by a set of psd. matrices $\{F^a\}_{a\in\mathcal{A}} \subseteq \mathbb{C}^{n\times n}$ that sum up to the identity matrix, i.e. $\sum_{a\in\mathcal{A}} F^a = I$. A *projective* measurement is defined by matrices that satisfy $F^a F^b = \delta_{ab} F^a$ for all $a, b \in \mathcal{A}$. 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 $\mathcal{A}$ as real numbers.

In that case we have:

$$\mathbb{E}[\chi] = \sum_{a\in\mathcal{A}} a\mathrm{Tr}(\rho F^a) = \mathrm{Tr}(\rho(\sum_{a\in\mathcal{A}} aF^a)) \tag{2}$$

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

$$(F^+ - F^-)^2 = \underbrace{F^{+^2}}_{=F^+} - \underbrace{F^+F^-}_{\delta_{+-}=0} + \underbrace{F^{-^2}}_{F^-} = F^+ + F^- = I \tag{3}$$

i.e. a $\{-1,1\}$-valued observable is both unitary and Hermitian.
Now let us take a quantum system X consisting of subsystems $X_1,\ldots,X_N$ 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 $\{F^{a_1}\}_{a_1\in\mathcal{A}_1} \in \mathbb{C}^{n_1\times n_1},\ldots,\{F^{a_N}\}_{a_N\in\mathcal{A}_N} \in \mathbb{C}^{n_N\times n_N}$ and the joint probability distribution of the N measurement outcomes $\chi_1,\ldots,\chi_N$ is

$$\mathbb{P}[\chi_1 = a_1, \chi_2 = a_2, \ldots, \chi_N = a_N] = \mathrm{Tr}(\rho F_1^{a_1}\otimes\cdots\otimes F_N^{a_N}) \tag{4}$$

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 $|\psi\rangle \in \mathbb{C}^{n_1}\otimes\cdots\otimes\mathbb{C}^{n_N}$ such that $\rho = |\psi\rangle\langle\psi|$. The state is called *product state* if it is of the form $|\psi\rangle = |\psi_1\rangle|\psi_2\rangle\ldots|\psi_N\rangle$. 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.

3. 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.

3.1. *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:

$$\mathbb{E}_{s,t\sim\pi} \left[V(a(s), b(t), s, t) \right] \tag{5}$$

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:

$$\mathbb{E}_{s,t\sim\pi} \left[\sum_{a\in\mathcal{A}} \sum_{b\in\mathcal{B}} \mathrm{Tr}(\rho F_s^a \otimes G_t^b) V(a, b, s, t) \right] \tag{6}$$

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.

3.2. *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 an 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 $b : \mathcal{B}|righta\{0,1\}$ 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:

$$\mathbb{E}_{(s,t)\sim\pi}\left[\chi(s)\psi(t)\Sigma_{st}\right] \tag{7}$$

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:

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

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

$$\mathbb{E}_{(s,t)\sim\pi}\left[\langle\psi|F_s\otimes G_t|\psi\rangle\right] \tag{8}$$

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:

$$\max\left\{\mathbb{E}_{(s,t)\sim\pi}\left[\Sigma_{st}\chi(s)\psi(t)\right]:\chi:\mathcal{S}\to\{-1,1\},\psi:\mathcal{T}\to\{-1,1\}\right\} \tag{9}$$

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

$$\sup_{n\in\mathbb{N}}\left\{\mathbb{E}_{(s,t)\sim\pi}\left[\Sigma_{st}\langle\psi|F_s\otimes G_t|\psi\rangle\right]:|\psi\rangle\in\mathbb{C}^n\otimes\mathbb{C}^n,F_s,G_t\in O(\mathbb{C}^n)\right\} \tag{10}$$

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.

3.3. *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

$$X=\begin{bmatrix}0&1\\1&0\end{bmatrix},Y=\begin{bmatrix}0&-i\\i&0\end{bmatrix} \tag{11}$$

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:

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

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

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

$$\langle \mathrm{EPR}|Y\otimes X|\mathrm{EPR}\rangle = 0$$

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

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

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

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

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

Thus, we have

$$\langle \mathrm{EPR}|F_s \otimes G_t|\mathrm{EPR}\rangle = \begin{cases} \frac{1}{\sqrt{2}}, (0,0), (1,0), (0,1) \\ -\frac{1}{\sqrt{2}}, (1.1) \end{cases} \tag{12}$$

which is equivalent to

$$\langle \mathrm{EPR}|F_s \otimes G_t|\mathrm{EPR}\rangle = \frac{(-1)^{s\wedge t}}{\sqrt{2}}, s,t \in \{0,1\} \tag{13}$$

The bias of the entangled strategy equals

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

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

$$\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{\sqrt{2}} = \cos(\pi/8) \approx 0.85\ldots \tag{14}$$

4. Local and Quantum correlation matrices

4.1. *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 $a \in [-1,1]^{\mathcal{S}}, b \in [-1,1]^{\mathcal{T}}$. Their common answer is the product $a_s b_t$. So, their strategy can be uniquely described by a matrix $ab^\top$. 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 4.1.1. Let $(X_i)_{1\le i\le m}$ and $(Y_j)_{1\le j\le n}$ be families of random variables on a common probability space such that $|X_i|, |Y_j| \le 1$ 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 $1 \le i \le m, 1 \le j \le n$.

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

As we will see in the following two lemmata, both sets can be described in a more useful way.

Lemma 4.1.2. *An alternative description of* $\mathrm{LC}_{m,n}$ *is given by*

$$\mathrm{LC}_{m,n} = \mathrm{conv}\{\xi\eta^T \,|\, \xi \in \{-1,1\}^m, \eta \in \{-1,1\}^n\}. \tag{15}$$

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 15 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 $|X_i|^{(k)}, |Y|_j^{(k)} \leq 1$ We have to show that there exists random variables $(X_i), (Y_j)$ with $\|X_i\|, \|Y_j\| \leq 1$ almost surely such that

$$\beta a_{ij}^{(0)} + (1-\beta) a_{ij}^{(1)} = \mathbb{E}[X_i Y_j] \tag{16}$$

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 $X_i = X_i^{(\alpha)}, Y_j = Y_j^{(\alpha)}$. Then

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

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

For the other inclusion, let $(a_{ij}) \in \mathrm{LC}_{m,n}$, i.e. $a_{ij} = \mathbb{E}[X_i Y_j]$ 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 $[-1,1]^d = \mathrm{conv}\{\xi \,|\, \xi \in \{-1,1\}^d\}$ (this can be proved by induction). If we define the random variables $X = (X_1, ..., X_m)$ and $Y = (Y_1, ..., Y_n)$ 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 $\sum_{\xi\in\{-1,1\}^m}\lambda_\xi^{(X)}(\omega)=1$. We can deduce that $X_i=\sum_{\xi\in\{-1,1\}^m}\lambda_\xi^{(X)}\xi_i$ almost surely. If we proceed analogously for Y we obtain

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

where we used that $\lambda_\xi^{(X)}$ and $\lambda_\eta^{(Y)}$ are independent. Since $\sum_{\xi\in\{-1,1\}^m,\eta\in\{-1,1\}^n}\mathbb{E}[\lambda_\xi^{(X)}]\mathbb{E}[\lambda_\eta^{(Y)}]=1$ it follows that $(a_{ij})\in\operatorname{conv}\{\xi\eta^T\,|\,\xi\in\{-1,1\}^m,\eta\in\{-1,1\}^n\}$ 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 $\mathrm{QC}_{m,n}$.

4.2. *Quantum correlation matrices*

Definition 4.2.1. Let $(X_i)_{1\le i\le m}$ and $(Y_j)_{1\le j\le n}$ be self-adjoint operators on $\mathbb{C}^{d_1}$, respectively $\mathbb{C}^{d_2}$ for some positive integers d_1,d_2, satisfying $\|X_i\|_\infty,\|Y_j\|_\infty\le 1$. $A=(a_{ij})$ is called *quantum correlation matrix* if there exists a state **Introduce a symbol zo define operators form one space to another** $\rho\in D(\mathbb{C}\otimes\mathbb{C})$ such that

$$a_{ij}=\operatorname{Tr}\rho(X_i\otimes Y_j). \tag{17}$$

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 4.2.2.

$$QC_{m,n}=\{(\langle x_i,y_j\rangle)_{1\le 1\le m,1\le j\le n}\,|\,x_i,y_j\in\mathbb{R}^{\min\{m,n\}},|x_i|\le 1,|y_j|\le 1\},$$

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

In order to proof this we have to review some definitions and introduce a special class of matrices, namely the *Pauli matrices*.

For the first inclusion we review the definition of an inner product. The basic idea is to define an inner product via the definition of the a_{ij} in 17. Let V and W be two vector spaces and k a field. A *bilinear form* is a map $\beta: V\times W\to k$ that is linear in both variables, that is

(1) $\beta(v_1+v_2,w)=\beta(v_1,w)+\beta(v_2,w)$
(2) $\beta(\lambda v,w)=\lambda\beta(v,w)$
(3) $\beta(v,w_1+w_2)=\beta(v,w_1)+\beta(v,w_2)$
(4) $\beta(v,\lambda w)=\lambda\beta(v,w)$

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* and usually denoted by $\langle \cdot\,, \cdot \rangle$. Again, we will write M for the right hand side of equation 18.

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 $(X_i)_{1 \geq m}$, $(Y_j)_{1 \geq n}$ on $\mathbb{C}^{d_1}$, respectively $\mathbb{C}^{d_2}$ 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 semidefinite symmetric bilinear form on the space of Hermitian operators on $\mathcal{H}$ by $\beta : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ where $\beta(S,T) = \mathrm{Re}(\mathrm{Tr}\,\rho ST)$. 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

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

Perhaps explanation

Moreover, since S^*S is a positive semidefinite operator for all complex operators S we obtain $\beta(S,S) = \mathrm{ReTr}\,\rho SS = \mathrm{ReTr}\,\rho S^* S \geq 0$ and β is positive semidefinite.

Following the steps in appendix **XXX** we can factorize the kernel and transform β to an inner product on the vector space of self-adjoint operators modulo the kernel of β. Equipped with an inner product, we can regard $\mathrm{B}^{sa}(\mathcal{H})$ **Introduce notation** as a real Euclidean space. We immediately get that $a_{ij} = \mathrm{Tr}\,\rho X_i Y_j = \beta(X_i \otimes I, I \otimes Y_j)$. In order to show that the norms of our vectors are bounded by one we have to show $\beta(X \otimes I, X \otimes I), \beta(I \otimes Y, I \otimes Y) \leq 1$ for all $X \in \{X_1, ..., Y_m\}$, $Y \in \{Y_1, ..., Y_n\}$. Therefore, let $d = d_1 d_2$ and $\rho = |\phi\rangle\langle\phi|$ for $|\phi\rangle = \sum_i^d \lambda_i \xi_i \otimes \eta_i$, where $\{\xi_1, ..., x_d\} \subset \mathbb{C}^{d_1}$ and $\{\eta_1, ..., \eta_d\} \subset \mathbb{C}^{d_2}$ 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. Writing ρ in this form we get

$$\begin{aligned}\beta(X \otimes I, X \otimes I) = \mathrm{ReTr}\,\rho X^2 \otimes I &= \sum_{i=1}^d \lambda_i^2 \mathrm{Tr}\,(|\xi_i\rangle\langle\xi_i| \otimes |\eta_i\rangle\langle\eta_i|)(X^2 \otimes I) \\ &= \sum_{i=1}^d \lambda_i^2 \mathrm{Tr}\,(|\xi_i\rangle\langle\xi_i| X^2) \mathrm{Tr}\,(|\eta_i\rangle\langle\eta_i|) = \sum_{i=1}^d \lambda_i^2 \mathrm{Tr}\,(|\xi_i\rangle\langle\xi_i| X^2).\end{aligned}$$

In order to get the desired result we have to show that $\mathrm{Tr}\,(|\xi_i\rangle\langle\xi_i| X^2) = \mathrm{Tr}\,(X^2 |\xi_i\rangle\langle\xi_i|) \leq 1$. Note that $1 \geq \|X\|_\infty := \sup_{|y| \leq 1} |Xy|$ implies that $|X^2 |\xi_i\rangle| \leq 1$. 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 $\text{span}\{x_1, ..., x_m\}$. Let $\{a_1, ..., a_r\}$ be an orthonormal basis of $\text{span}\{x_1, ..., x_m\}$ with respect to β. The orthogonal projection of y_j is $\pi(y) := \sum_{i=1}^{r} \beta(a_i, y_j)a_i$ and fulfills $\beta(x_i, y_j) = \beta(x_i, \pi(y_j))$. **perhaps explanation but standard calculation** 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

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

Moreover, since $|x_i|, |y_j| = |\pi(y_j)| \leq 1$ we get $\langle \alpha^{(i)}, \alpha^{(i)} \rangle, \langle \gamma^{(j)}, \gamma^{(j)} \rangle \leq 1$ and thus the vectors $(\alpha^{(i)})$ and$(\gamma^{(j)})$ all required properties. Additionally, we also proved that we can take an even lower dimension for our vectors, precisely $\min\{\dim(\text{span}\{(x_i)\}), \dim(\text{span}\{(y_j)\})\}$. □

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

Proposition 1. *For all $n \geq 1$ there is a subspace of the $2^n \times 2^n$ Hermitian matrices where every vector 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 0 unitary Hermitian matrices and anti-commute pairwise. 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 $A = A'$ 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 $i = 1, ..., n$. Note that these operators are also traceless 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

$$U_iU_{n+j} = (\sigma_X^i \prod_{k=i+1}^{n} \sigma_Y^k) \cdot (\sigma_Z^j \prod_{k=j+1}^{n} \sigma_Y^k) = \sigma_X^i \sigma_Y^{i+1} \cdots \sigma_Y^j (\sigma_Z^j \prod_{k=j+1}^{n} \sigma_Y^k) \sigma_Y^{j+1} \cdots \sigma_Y^n$$
$$= -\sigma_X^i \sigma_Y^{i+1} \cdots \sigma_Y^{j-1} (\sigma_Z^j \prod_{k=j+1}^{n} \sigma_Y^k) \sigma_Y^{j+1} \cdots \sigma_Y^n$$
$$= -(\sigma_Z^j \prod_{k=j+1}^{n} \sigma_Y^k)(\sigma_X^i \prod_{k=i+1}^{n} \sigma_Y^k)$$
$$= -U_jU_i.$$

Since $U_iU_i^* = U_iU_i = I$ they are also unitary. Moreover, taking the product of two linear combinations $X = \sum_{i=1}^{2n} \xi_i U_i$, $Y = \sum_{i=1}^{2n} = \eta_i U_i$ we can calculate

$$XY = \sum_{i=1}^{2n} \xi_i\eta_i I + \sum_{1\leq i,j\leq\leq 2n} \xi_i\eta_j U_iU_j = \sum_{i=1}^{2n} \xi_i\eta_i I + \sum_{1\leq i<j\leq\leq 2n} \xi_i\eta_j U_iU_j - \sum_{1\leq i<j\leq\leq 2n} U_iU_j$$
$$= \sum_{i=1}^{2n} \xi_i\eta_i I$$
$$= \langle \xi, \eta \rangle I.$$

So, if we set $Y = X$ we get the desired result by taking the subspace $\text{span}\{U_i \,|\, i = 1, ..., 2n\}$

□

We are now ready to prove the other inclusion of lemma 4.2.2.

Proof of $\text{LC}_{m,n} \supset M$. Let $(x_i)_{1\leq i\leq m}$, $(y_j)_{1\leq j\leq n}$ be vectors in $\mathbb{R}^{\min\{m,n\}}$ that satisfy $|x_i|,|y_j| \leq 1$. Using the notation of the previous proposition's proof we set $X_i = \sum_{k=1}^{\min\{m,n\}} x_i(k)U_i$ and $Y_j^T = \sum_{k=1}^{\min\{m,n\}} y_j(k)U_k$ where the U_i's are $d \times d$ matrices with $d = 2^{\lceil \min\{m,n\}/2 \rceil}$ Then $\text{Tr}\,(XY^T) = d \cdot \langle x, y \rangle$ and $\|X\|_\infty \leq 1$ since $X_iX_i^* = |x_i|^2 I$ and $|x_i|^2 \leq 1$. The same holds for Y^T which also implies $\|Y\|_\infty \leq 1$. **Elaborate** Let $|\phi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle$ and $\rho = |\phi\rangle \langle\rho|$. Note that we can write ρ as

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

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

Then we get

(18)

$$\text{Tr}\,(\rho X_i \otimes Y_j) = \frac{1}{d} \sum_{1\leq k,l\leq d} \text{Tr}\,(|k\rangle \langle l| X_i \otimes |k\rangle \langle l| Y_j) = \frac{1}{d} \sum_{1\leq k,l\leq d} \text{Tr}\,(|k\rangle \langle l| X_i) \text{Tr}\,(|k\rangle \langle l| Y_j)$$

(19)

$$= \frac{1}{d} \text{Tr}\, X_iY_j^T = \langle x, y \rangle.$$

□

We can easily see that $\text{QC}_{m,n}$ is convex. Consider $(a_{ij}), (\bar{a}_{ij}) \in \text{QC}_{m,n}$ 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 $|x_i|, |y_j|, |\bar{x}_i|, |\bar{y}_j| \leq 1$. For $\lambda \in [0,1]$ we define vectors $\tilde{x}_i := (\sqrt{\lambda}x_i, \sqrt{1-\lambda}\bar{x}_i)$, $\tilde{y}_j := (\sqrt{\lambda}y_j, \sqrt{1-\lambda}\bar{y}_j)$ and due to $|\tilde{x}_i| \leq \lambda|(x_i, 0)| + (1-\lambda)|(0, \bar{x}_i)| \leq 1$ they are unit vectors. 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 4.2.2 we obtain vectors $\alpha^{(i)}, \gamma^{(j)}$ that satisfy $= \langle \alpha^i, \gamma^j\rangle = \langle \tilde{x}_i, \tilde{y}_j$ and have dimension smaller or equal to $\min\{m, n\}$.

Using these both descriptions 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 $x_i = \xi_i |0\rangle$ and $y_j = \eta_j |0\rangle$ we immediately see that $\xi_i\eta_j = \langle x_i, y_j\rangle$. 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

$$\mathrm{QC}_{2,2} = \{A \in \mathbb{R}^{2\times 2} \mid -1 \leq \mathrm{Tr}\, AM \leq 1 \text{ for all } M \in \mathcal{K}\}, \tag{20}$$

where $\mathcal{K} = \{\frac{1}{2}\sigma(\begin{pmatrix}-1 & 1\\ 1 & 1\end{pmatrix}), \sigma(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}) \mid \sigma \in \{(1\ 2), (1\ 3), (1\ 4)\}\}$. The crucial observation is to note that $\mathrm{LC}_{2,2}$ is affinely isomorphic to the cross polytope scaled by two, i.e. $\mathrm{LC}_{2,2} \cong 2\mathrm{CP}_4 := \mathrm{conv}\{\pm e_i \mid i = 1, ..., 4\}$, 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

$$\frac{1}{2}\begin{pmatrix}-1 & -1 & 1 & 1\\ -1 & 1 & -1 & 1\\ -1 & 1 & 1 & -1\\ 1 & 1 & 1 & 1\end{pmatrix}.$$

An easy analysis shows that $(\mathrm{CP}_4)^o = [-1, 1]^4$ **maybe introduce polar dual**, hence the face lattice of CP_4 is isomorphic to the inverse **look up the correct name** lattice of the hypercube 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 in 20 clearly define non-empty 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 $1/2M = \begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$. Since $\max\{1/2\mathrm{Tr}\, AM \mid A \in \mathrm{LC}_{2,2}\} = 1$ it suffices to show that there is $A \in \mathrm{QC}_{m,n}$ achieving a better value. For $A \in \mathrm{QC}_{2,2}$ we obtain, by lemma 4.2.2, 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 $\operatorname{Tr} AM$ 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 $\operatorname{Tr} AM = 2\sqrt{2}$.

5. Grothendieck Inequality

Lemma 5.0.1 (Grothendieck's identity)**.** *Let $u, v \in \mathbb{R}^d$ be unit vectors. Let $r \in \mathbb{R}^d$ be a random unit vector chosen from $O(d)$-invariant probability distribution on the unit sphere. Then*

i, $\mathbb{P}[\operatorname{sign}(u^\top r) \neq \operatorname{sign}(v^\top r)] = \frac{\arccos(u^\top v)}{\pi}$
ii, $\mathbb{E}[\operatorname{sign}(u^\top r) \operatorname{sign}(v^\top r)] = \frac{2}{\pi} \arcsin(u^\top v)$.

Proof. Assume that u and v are linearly dependent. Since both, u and v, are unit vectors, that is, either $u = v$, then $\arccos(u^\top v) = \arccos(1) = 0$ or $u = -v$, then $\arccos(u^\top v) = \arccos(-1) = \pi$. Conversely assume that u and v are linearly independent, i. e. $\dim(\operatorname{span}\{u, v\}) = 2$. Now project r orthogonally on the plane spanned by u and v. This gives us a vector $s \in \operatorname{span}\{u, v\}$ such that $u^\top r = u^\top s$, $v^\top r = v^\top s$. The unit vector $s/\|s\|$ is uniformly distributed on the unit circle by the $O(d)$-invariance of the probability distribution. □

Appendix

Dinge, die definiert werden sollten.

(1) Injective tensor product
(2) norms
(3) Notation, operators of norms
(4) perhaps what a state is

5.1. *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 β defines an inner product on U.