Entanglement as a resource in quantum information theory

Guillaume Aubrun

We present a short introduction to some mathematical aspects of quantum entanglement. These notes are based on lectures given during the workshop "from quantum to classical", held in CIRM in April 2019. We limit ourselves on purpose to a sparse bibliography. For more details and references, we refer the reader to [1] and [6].

1 What is entanglement?

1.1 Quantum states

Let $\mathcal{H}$ be a finite-dimensional complex Hilbert space, and d its dimension. We write $B(\mathcal{H})$ for the algebra of linear operators on $\mathcal{H}$, and $B^{\mathrm{sa}}(\mathcal{H})$ for the real subspace of Hermitian operators. Choosing an orthonormal basis in $\mathcal{H}$ allows to identify $B(\mathcal{H})$ with the matrix algebra M_d, and $B^{\mathrm{sa}}(\mathcal{H})$ with the subspace $\mathsf{M}_d^{\mathrm{sa}}$ of self-adjoint matrices.

The set of (quantum) states on $\mathcal{H}$ is defined as

$$\mathrm{D}(\mathcal{H}) = \{\rho \in B^{\mathrm{sa}}(\mathcal{H}) \,:\, \rho \geqslant 0,\ \mathrm{Tr}\,\rho = 1\}.$$

Here the letter D stands for "density matrices". The set $\mathrm{D}(\mathcal{H})$ is a convex compact set whose extreme points are rank 1 orthogonal projectors. We use Dirac bra-ket notation: for a unit vector $\psi \in \mathcal{H}$ (sometimes denoted by $|\psi\rangle$), the orthogonal projection onto $\mathbf{C}\psi$ is denoted by $|\psi\rangle\langle\psi|$. A state of the form $|\psi\rangle\langle\psi|$ is called a pure state. A general state (sometimes called a mixed state) is a convex combination (=a mixture) of pure states.

Exercise 1 shows that $\mathrm{D}(\mathbf{C}^2)$ is geometrically a 3-dimensional Euclidean ball. This is specific to the 2-dimensional case. In larger dimensions, it is more accurate to think of $\mathrm{D}(\mathbf{C}^d)$ as a non-commutative analogue of the probability simplex

$$\Delta_d = \{(x_1, \cdots, x_d) \in \mathbf{R}^d \ \ x_i \geqslant 0,\ x_1 + \cdots + x_d = 1\}.$$

Exercise 1 (The Bloch ball)**.** *Show that the set of quantum states on* $\mathbf{C}^2$ *can be described as*

$$\{A \in \mathsf{M}_2^{\mathrm{sa}} \quad \mathrm{Tr}(A) = 1,\ \mathrm{Tr}(A - \mathrm{Id}/2)^2 \leqslant 1/2\}$$

which is geometrically a 3*-dimensional Euclidean ball.*

A natural metric on $\mathrm{D}(\mathbf{C}^2)$ is induced by the trace norm, defined on $B(\mathcal{H})$ by $\|A\|_1 = \mathrm{Tr}\,|A|$, where $|A| = (AA^*)^{1/2}$. The trace norm has an operational interpretation, which we now explain. Consider two quantum states ρ_1, ρ_2, and suppose that you are given a state ρ chosen uniformly at random among $\{\rho_1, \rho_2\}$. Can you infer whether $\rho = \rho_1$ or $\rho = \rho_2$ from observation? We know from the axioms of quantum mechanics that when measuring a state ρ in an orthonormal basis (ψ_k), we observe the output k with probability $\langle\psi_k|\rho|\psi_k\rangle$. By the principle of maximum likelihood, when observing k, one should infer that $\rho = \rho_1$ if $\langle\psi_k|\rho_1|\psi_k\rangle \geqslant \langle\psi_k|\rho_2|\psi_k\rangle$, and $\rho = \rho_2$ otherwise. Analyzing the probability of error for the optimal choice of basis involves $\|\rho_1 - \rho_2\|_1$ (see Exercise 2).

Exercise 2 (Holevo–Helström theorem)**.** *Show that in the above scenario, the probability of error equals*

$$\frac{1}{2} - \frac{1}{4}\sum_k |\langle\psi_k|\rho_1 - \rho_2|\psi_k\rangle|. \tag{1}$$

Then show that the infimum of (1) *over orthonormal bases* (ψ_k) *equals* $\frac{1}{2} - \frac{1}{4}\|\rho_1 - \rho_2\|_1$.

Exercise 3 shows that the trace norm of the difference between two pure states depends only on the inner product between the corresponding vectors.

Exercise 3 (Trace norm and pure states)**.** *Show the formula, for unit vectors* ψ, φ

$$\Big\||\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi|\Big\|_1 = 2\sqrt{1 - |\langle\psi,\varphi\rangle|^2}.$$

1.2 Entanglement of bipartite states

To define entanglement, we need a tensor product structure on the Hilbert space $\mathcal{H}$. Assume that $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ for Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$ (for simplicity we restrict ourselves to bipartite Hilbert spaces). The interpretation from quantum mechanics is that $\mathcal{H}_A$ and $\mathcal{H}_B$ describe the state of two subsystems A and B, while $\mathcal{H}$ describes the joint state of their union. It is traditional in quantum information theory to think of the systems A and B as held respectively by Alice and Bob, who are distant observers.

A unit vector $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ is called product is it has the form $\psi_A \otimes \psi_B$, and otherwise is called entangled. We now come to the main definition of these notes.

Definition. *The set of* separable states *on* $\mathcal{H}_A \otimes \mathcal{H}_B$ *is the subset of* $\mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ *defined as*

$$\begin{aligned}
\mathrm{Sep}(\mathcal{H}) &= \mathrm{conv}\{|\psi\rangle\langle\psi| : \psi \text{ product unit vector in } \mathcal{H}_A \otimes \mathcal{H}_B\} \\
&= \mathrm{conv}\{|\psi_A\rangle\langle\psi_A| \otimes |\psi_B\rangle\langle\psi_B|\} \\
&= \mathrm{conv}\{\rho_A \otimes \rho_B : \rho_A \in \mathrm{D}(\mathcal{H}_A),\ \rho_B \in \mathrm{D}(\mathcal{H}_B)\}
\end{aligned}$$

A non-separable quantum state is called entangled.

It is easily checked that a pure state $|\psi\rangle\langle\psi| \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is entangled if and only if ψ is an entangled vector. One gets an idea of the geometry of Sep by figuring out what the convex hull of the curvilinear oval on a standard tennis ball is.

1.3 Schmidt decomposition

The Schmidt decomposition of tensors is a reformulation of the singular value decomposition for matrices.

Proposition 1 (Schmidt decomposition)**.** *Let $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a unit vector, and $d = \min(\dim \mathcal{H}_A, \dim \mathcal{H}_B)$. There exist orthonormal families $(\psi_i^A)_{1 \leqslant i \leqslant d}$ in $\mathcal{H}_A$ and $(\psi_i^B)_{1 \leqslant i \leqslant d}$ in $\mathcal{H}_B$, and nonnegative numbers $(\lambda_i)_{1 \leqslant i \leqslant d}$ such that*

$$\psi = \sum_{i=1}^{d} \sqrt{\lambda_i}\, \psi_i^A \otimes \psi_i^B.$$

Numbers (λ_i) from Proposition 1 are called the Schmidt coefficients of ψ, and denoted by $\lambda_\psi = (\lambda_1, \ldots, \lambda_d)$. Schmidt coefficients are uniquely determined if we require that $\lambda_1 \geqslant \cdots \geqslant \lambda_d$. Note also that $\lambda_1 + \cdots + \lambda_d = 1$, so that λ_ψ belongs to the probability simplex Δ_d.

2 How to use entanglement?

We here present quantum teleportation, as an example of a quantum information protocol which uses entanglement as a resource.

We set some notation: the canonical basis of the Hilbert space $\mathbf{C}^2$ is denoted by $(|0\rangle, |1\rangle)$. We often drop the tensor product sign: $|01\rangle$ should be understood as $|0\rangle \otimes |1\rangle$. We consider the following entangled vectors in $\mathbf{C}^2 \otimes \mathbf{C}^2$.

$$|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle \pm |11\rangle \right),$$

$$|\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}} \left(|01\rangle \pm |10\rangle \right).$$

Note that the 4 vectors $|\Phi^+\rangle$, $|\Phi^-\rangle$, $|\Psi^+\rangle$, $|\Psi^-\rangle$ form an orthonormal basis of $\mathbf{C}^2 \otimes \mathbf{C}^2$. The pure state $|\Phi^+\rangle\langle\Phi^+|$ is often called a *Bell state*.

The teleportation protocol involves three Hilbert spaces $\mathcal{H}_{A'}$, $\mathcal{H}_A$ and $\mathcal{H}_B$, all isomorphic to $\mathbf{C}^2$, where Alice has access to $\mathcal{H}_{A'} \otimes \mathcal{H}_A$. Assume that Alice and Bob "share a Bell state", and that Alice holds an unknown pure state $|\psi\rangle\langle\psi|$ for some unit vector $\psi \in \mathcal{H}_{A'}$. The meaning of this sentence is that the initial state of the system is the pure state corresponding to the vector $|\psi\rangle_{A'} \otimes |\Phi^+\rangle_{AB}$.

We have $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ for some complex numbers α, β with $|\alpha|^2 + |\beta|^2 = 1$. It is a routine exercise to check that

$$\begin{aligned}(\alpha|0\rangle + \beta|1\rangle)_{A'}|\Phi^+\rangle_{AB} &= \frac{1}{2}\Big[|\Phi^+\rangle_{A'A}(\alpha|0\rangle + \beta|1\rangle)_B + |\Phi^-\rangle_{A'A}(\alpha|0\rangle - \beta|1\rangle)_B \\ &+ |\Psi^+\rangle_{A'A}(\alpha|1\rangle + \beta|0\rangle)_B + |\Psi^-\rangle_{A'A}(\alpha|1\rangle - \beta|0\rangle)_B\Big]. \end{aligned} \quad (2)$$

Denote the right-hand side of (2) by

$$\frac{1}{2}\Big[|\Phi^+\rangle_{A'A}|\phi_1\rangle_B + |\Phi^-\rangle_{A'A}|\phi_2\rangle_B + |\Phi^+\rangle_{A'A}|\phi_3\rangle_B + |\Phi^-\rangle_{A'A}|\phi_4\rangle_B\Big].$$

We now describe the quantum teleportation protocol, which involves only pure states. First, Alice measures the system in the orthonormal basis $(|\Phi^\pm\rangle, |\Psi^\pm\rangle)$. By the axioms of quantum mechanics, as a result of the measurement, Alice observes one of the four possible outcomes (each occurring with probability $\frac{1}{4}$), and the system collapses to one of the four summands in (2). Next, Alice communicates to Bob which outcome $i \in \{1, 2, 3, 4\}$ she observed: this information can be encoded into 2 classical bits. According to the information $i \in \{1, 2, 3, 4\}$ he received, Bob applies a unitary matrix $U_i \in \mathsf{U}(2)$ with the property that $U_i|\phi_i\rangle = |\psi\rangle$ (it is easy to find U_i having this property for all values of α, β). After this step, the system in a state of the form $|?\rangle_{A'A}|\psi\rangle_B$ (the value of $|?\rangle_{A'A}$ can be computed but is irrelevant): the vector $|\psi\rangle$ has "traveled" from A' to B.

This protocol can be interpreted as the transmission of 1 qubit of quantum information (the vector $\psi \in \mathbf{C}^2$). To achieve this, it used two kinds of resources:

a) 2 bits of classical communication,

b) 1 "bit of entanglement" (sometimes called an ebit) in the form of the Bell state $|\Phi^+\rangle_{AB}$.

Note that entanglement has effectively been consumed: at the end of the protocol, the state has a product form.

3 How to quantify entanglement?

3.1 The LOCC paradigm

We explain how to quantify entanglement, to guarantee that one state is "more entangled" than another one. This is achieved by the paradigm of LOCC transformations, which use Local Operations and Classical Communication. The class of LOCC transformations can be precisely defined (see [6], Section 6.1.2) but we prefer to give an informal definition. LOCC transformations are maps (in fact, quantum channels) from $\mathcal{H}_A \otimes \mathcal{H}_B$ to $\mathcal{H}'_A \otimes \mathcal{H}'_B$ corresponding to protocols of the following nature

a) Alice applies local operations (e.g. a measurement) on her part of the system, and then communicates the outcomes to Bob.

b) Depending on the information he received, Bob applies local operations on his part of the system, and then communicates the outcomes to Alice.

c) Repeat a) and b) arbitrarily many times.

For example, the teleportation protocol is a LOCC protocol. For states $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $\sigma \in \mathrm{D}(\mathcal{H}'_A \otimes \mathcal{H}'_B)$, we write $\rho \xrightarrow{\text{LOCC}} \sigma$ if there is a LOCC protocol which maps σ to ρ. This defines a partial order on the set of bipartite states of arbitrary dimensions.

3.2 Quantifying pure state entanglement

In the case of pure states, the order $\xrightarrow{\text{LOCC}}$ is perfectly understood and is connected to the notion of majorization.

For a probability vector $x \in \Delta_d$, denote by $x^{\downarrow}$ the nonincreasing rearrangement of x, i.e. the vector satisfying $x_1^{\downarrow} \geqslant x_2^{\downarrow} \geqslant \cdots \geqslant x_d^{\downarrow}$ obtained from permuting the coordinates of x.

Definition. *Let $x \in \Delta_{d_1}$, $y \in \Delta_{d_2}$ be probability vectors. We say y majorizes x, and write $x \prec y$, if*

$$x_1^{\downarrow} + \cdots + x_k^{\downarrow} \leqslant y_1^{\downarrow} + \cdots + y_k^{\downarrow}$$

for every $k \in \{1, \ldots, \min(d_1, d_2)\}$.

Exercise 4 (Equivalent definitions for majorization)**.** *Let x, $y \in \Delta_d$. Show that the following are equivalent.*

1. *$x \prec y$.*
2. *There is a bistochastic matrix B such that $x = By$. (A $d \times d$ matrix is bistochastic if it has nonnegative entries, and the sum of entries in every row and every column equals 1.)*
3. *For every convex function $\Phi : \mathbf{R} \to \mathbf{R}$, we have*

$$\Phi(x_1) + \cdots + \Phi(x_d) \leqslant \Phi(y_1) + \cdots + \Phi(y_d).$$

The following theorem characterizes exactly LOCC convertibility of pure states.

Theorem 2 (Nielsen)**.** *Let $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ and $\phi \in \mathcal{H}'_A \otimes \mathcal{H}'_B$ be unit vectors with respective Schmidt coefficients λ_ψ and λ_ϕ. Then*

$$|\phi\rangle\langle\phi| \xrightarrow{\text{LOCC}} |\psi\rangle\langle\psi| \quad \text{if and only if} \quad \lambda_\phi \prec \lambda_\psi.$$

We now would like to quantify the amount of entanglement contained in a given state. The idea is to use the Bell state as a gold standard, and to consider conversion rates between Bell states and an arbitrary state $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Informally, we can define

a) the entanglement cost of ρ, the rate of LOCC conversion from $|\Phi^+\rangle$ to ρ,

b) the entanglement of distillation of ρ, the rate of LOCC conversion from ρ to $|\Phi^+\rangle$.

As often in information theory, the rate should be understood in the asymptotic limit of many-copy transformations. Moreover, it is acceptable to allow errors (measured with in the trace norm) with vanish in this asymptotic limit. The formal definition of the entanglement cost $E_C(\rho)$ and the entanglement of distillation $E_D(\rho)$ is as follows

$$E_C(\rho) = \inf\left\{ r > 0 \,:\, \lim_{n\to\infty} \inf_{A \text{ LOCC}} \left\| A\left(|\Phi^+\rangle\langle\Phi^+|^{\otimes \lfloor rn \rfloor}\right) - \rho^{\otimes n} \right\|_1 = 0 \right\}$$

$$E_D(\rho) = \sup\left\{ r > 0 \,:\, \lim_{n\to\infty} \inf_{A \text{ LOCC}} \left\| A\left(\rho^{\otimes n}\right) - |\Phi^+\rangle\langle\Phi^+|^{\otimes \lfloor rn \rfloor} \right\|_1 = 0 \right\}$$

Intuitively, since LOCC transformations cannot create entanglement, one expects that $E_D(\rho) \leqslant E_C(\rho)$ (this is not hard to establish rigorously, see Proposition 6.37 in [6]). For a mixed state ρ, computing the values of $E_D(\rho)$ and $E_C(\rho)$ is often very hard. However, for a pure state, the entanglement cost and the entanglement of distillation both coincide with a more familiar notion of entropy, which we now introduce.

Let $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a unit vector, and $\lambda_1, \ldots, \lambda_d$ its Schmidt coefficients. The *entropy of entanglement* of ψ is

$$E(\psi) = -\sum_{i=1}^{d} \lambda_i \log_2 \lambda_i.$$

Theorem 3. *Let $\rho = |\psi\rangle\langle\psi| \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a pure state. Then*

$$E_C(\rho) = E_D(\rho) = E(\psi).$$

Proof. For an integer n, we have $\rho^{\otimes n} = |\psi^{\otimes n}\rangle\langle\psi^{\otimes n}|$, and the Schmidt coefficients of $\psi^{\otimes n}$ are given by the family

$$(\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_n})_{1\leqslant i_1,\ldots,i_n\leqslant d}\,.$$

Fix $\varepsilon > 0$, and denote by $I_{n,\varepsilon}$ the interval $[2^{-n(E(\psi)+\varepsilon)}, 2^{-n(E(\psi)-\varepsilon)}]$. This is the "typical set" for the Schmidt coefficients of $\psi^{\otimes n}$, in the sense that, as n tends to infinity

$$\sum_{\substack{i_1,\ldots,i_n\in\{1,\ldots,d\} \\ \lambda_{i_1}\cdots\lambda_{i_n}\in I_{n,\varepsilon}}} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_n} = 1 - o(1). \tag{3}$$

Exercise 5. *Show* (3) *by applying the weak law of large numbers to a random variable taking the value* $\log\lambda_i$ *with probability* λ_i.

It follows that for every n, we can find a vector $\psi_n \in \mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n}$ with all Schmidt coefficients inside $I_{n,\varepsilon}$, such that $\|\psi_n - \psi^{\otimes n}\| = o(1)$. This implies $\||\psi_n\rangle\langle\psi_n| - |\psi^{\otimes n}\rangle\langle\psi^{\otimes n}|\|_1 = o(1)$ by Exercise 3. If $r \geqslant E(\psi) + \varepsilon$ and rn is an integer, it is easily checked from the definition of majorization that

$$\lambda_{|\Psi^+\rangle^{\otimes rn}} = \underbrace{(2^{-rn}, \cdots, 2^{-rn})}_{2^{rn} \text{ times}} \prec \lambda_{\psi_n}.$$

By Nielsen's theorem, this implies that

$$|\Psi^+\rangle\langle\Psi^+|^{\otimes rn} \xrightarrow{\text{LOCC}} |\psi_n\rangle\langle\psi_n| \approx |\psi^{\otimes n}\rangle\langle\psi^{\otimes n}|$$

and we proved that $E_C(\rho) \leqslant E(\psi) + \varepsilon$ for every $\varepsilon > 0$, i.e. $E_C(\rho) \leqslant E(\psi)$. A similar argument can be used to prove that $E_D(\rho) \geqslant E(\psi)$, and the theorem follows from the general inequality $E_D(\rho) \leqslant E_C(\rho)$. □

It follows from Theorem 3 that the manipulation of pure state entanglement is *reversible* in the many-copy limit. As we will see later, the situation for mixed states is much more complicated, and extreme forms of irreversibility appear.

3.3 Partial transposition

How to detect if a given state $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is separable or entangled? In the case of pure states, this is an easy task: it reduces to the computation of Schmidt coefficients, and even to the largest Schmidt coefficient only. For mixed states, the problem is known to be NP-hard, but a very useful sufficient condition for entanglement arises from the notion of partial transposition.

Define $T : B(\mathcal{H}_B) \to B(\mathcal{H}_B)$ as the transposition with respect to a given basis, i.e. the operation which maps (a_{ij}) to (a_{ji}) when expressing operators as matrices in that basis. One checks that $T(|\psi\rangle\langle\psi|) = |\overline{\psi}\rangle\langle\overline{\psi}|$, where $\overline{\psi}$ denotes the coordinatewise complex conjugation of a unit vector $\psi \in \mathcal{H}_B$.

Consider now the map $\Gamma = \mathrm{Id} \otimes T : B(\mathcal{H}_A \otimes \mathcal{H}_B) \to B(\mathcal{H}_A \otimes \mathcal{H}_B)$, which is called the *partial transposition*. (Note that $B(\mathcal{H}_A \otimes \mathcal{H}_B)$ is canonically isomorphic to $B(\mathcal{H}_A) \otimes B(\mathcal{H}_B)$.) If we represent $M \in B(\mathcal{H}_A \otimes \mathcal{H}_B)$ as a block-matrix, $\Gamma(M)$ is the matrix obtained by transposing inside each block of M.

Proposition 4. *If $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a separable state, then $\Gamma(\rho) \geqslant 0$. Equivalently, is $\Gamma(\rho)$ is not positive semi-definite, then ρ is entangled.*

Proof. A separable state ρ has the form $\sum_i \mu_i |\psi_i^A\rangle\langle\psi_i^A| \otimes |\psi_i^B\rangle\langle\psi_i^B|$, and therefore $\Gamma(\rho) = \sum_i \mu_i |\psi_i^A\rangle\langle\psi_i^A| \otimes |\overline{\psi_i^B}\rangle\langle\overline{\psi_i^B}|$ is positive. □

While the operator $\Gamma(\rho)$ depends on the choice of an orthonormal basis in $\mathcal{H}_B$, the statement "$\Gamma(\rho)$ is positive" is basis-independent!

Exercise 6. *Show that a pure state is entangled if and only if its partial transposition is not positive.*

Remark 1. *It is a non-trivial result that the converse to Proposition 4 holds true in small dimensions: for $2 \otimes 2$ and $2 \otimes 3$ system, a state is separable if and only if its partial transposition is positive.*

Exercise 7. *For which values of the parameter t is the state*

$$t\frac{\mathrm{Id}}{4} + (1-t)|\Phi^+\rangle\langle\Phi^+|$$

on $\mathbf{C}^2 \otimes \mathbf{C}^2$ entangled? In the easy mode, you are allowed to use the statement from Remark 1. In the hard mode, you should prove everything by yourself.

One can check that the partial transposition essentially commutes with the LOCC requirement. It follows that $\rho \xrightarrow{\text{LOCC}} \sigma$ and $\Gamma(\rho) \geqslant 0$ together imply that $\Gamma(\sigma) \geqslant 0$. Since the partial transposition of a Bell state is not positive (see Exercise 6), this means that a state ρ with positive transposition cannot be converted into Bell states. Its distillable entanglement vanishes, i.e. $E_D(\rho) = 0$.

It is a theorem that any entangled state ρ satisfies $E_C(\rho) > 0$ ([7], this is easy to believe, since LOCC transformations cannot produce entanglement out of nothing). The more surprising fact is that there are *bound entangled* states, i.e. entangled states with zero distillable entanglement. Such states are entangled, but useless for the teleportation protocol presented here.

An important open problem is the distillability problem: is $E_D(\rho) = 0$ equivalent to the positivity of $\Gamma(\rho)$? This is unknown already for states on $\mathbf{C}^3 \otimes \mathbf{C}^3$.

Computing the value of E_C and E_D is a hard task. This is related to the fact the decomposition of a mixed state into pure states is highly non-unique. Suppose that $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ admits a convex decomposition into pure states

$$\rho = \sum_i \mu_i |\psi_i\rangle\langle\psi_i|. \tag{4}$$

This can be turned into a LOCC protocol to create ρ from Bell states with rate $\sum_i \mu_i E(\psi_i)$. The optimization over decompositions (4) yields to the definition of the entanglement of formation

$$E_F(\rho) = \inf\left\{ \sum_i \mu_i E(\psi_i) \ : \ \sum_i \mu_i |\psi_i\rangle\langle\psi_i| = \rho, \ \mu_i \geqslant 0, \ \sum \mu_i = 1 \right\}.$$

The entanglement cost can then be retrieved as the "regularization" of the entanglement of formation

$$E_C(\rho) = \lim_{n\to\infty} \frac{1}{n} E_F(\rho^{\otimes n}).$$

If true, the additivity of the entanglement of formation, i.e. the formula $E_F(\rho\otimes\sigma) = E_F(\rho)+E_F(\sigma)$, would simplify greatly the theory since it would imply $E_F = E_C$. The additivity was a famous conjecture, which was disproved by Hastings [4] using the probabilistic method. The situation is still poorly understood, and the only known counterexamples to additivity are random constructions in extremely high dimensions.

4 Entanglement and correlations

A $m \times n$ real-valued matrix $A = (a_{ij})$ is called a classical correlation matrix if there exist random variables $(X_i)_{1\leqslant i\leqslant m}$ and $(Y_j)_{1\leqslant j\leqslant n}$ defined on some probability space, satisfying $|X_i| \leqslant 1$, $|Y_j| \leqslant 1$, and $a_{ij} = \mathbf{E}[X_i Y_j]$. We denote by $\mathsf{C}_{m,n}$ the set of $m \times n$ classical correlation matrices.

Exercise 8 (Classical correlation matrices as a polytope)**.** *Show that*

$$\mathsf{C}_{m,n} = \operatorname{conv}\left\{(\xi_i\eta_j)_{ij} \,:\, \xi \in \{-1,1\}^m,\ \eta \in \{-1,1\}^n\right\}.$$

A $m \times n$ real-valued matrix $A = (A_{ij})$ is called a quantum correlation matrix if there are Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, a state $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, operators $(X_i)_{1\leqslant i\leqslant m} \subset B^{\mathrm{sa}}(\mathcal{H}_A)$ and $(Y_j)_{1\leqslant j\leqslant n} \subset B^{\mathrm{sa}}(\mathcal{H}_B)$ satisfying $\|X_i\|_\infty \leqslant 1$, $\|Y_j\|_\infty \leqslant 1$ and

$$a_{ij} = \operatorname{Tr}(\rho(X_i \otimes Y_j)). \tag{5}$$

We denote by $\mathsf{Q}_{m,n}$ the set of $m \times n$ quantum correlation matrices. It is easy to check the inclusion $\mathsf{C}_{m,n} \subset \mathsf{Q}_{m,n}$.

If ρ is a separable state, then the matrix (a_{ij}) defined by (5) is actually a classical correlation matrix. Indeed, assuming the decomposition $\rho = \sum \mu_k |\psi_k^A \otimes \psi_k^B\rangle\langle\psi_k^A \otimes \psi_k^B|$, we have

$$a_{ij} = \sum_{k=1}^{K} \mu_k \langle\psi_k^A|X_i|\psi_k^A\rangle\langle\psi_k^B|Y_j|\psi_k^B\rangle$$

and one can interpret μ_k as a probability distribution on $\{1,\dots,K\}$, and $\langle\psi_k^A|X_i|\psi_k^A\rangle$ and $\langle\psi_k^B|Y_j|\psi_k^B\rangle$ as random variables. Therefore, non-classical correlations are a feature of entanglement.

Proposition 5 (Tsirelson)**.** *The set $\mathsf{Q}_{m,n}$ is convex and can be described as*

$$\left\{(\langle x_i, y_j\rangle)_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n} \,:\, x_i, y_j \in \mathcal{H}, ||x_i|| \leqslant 1, ||y_j|| \leqslant 1\right\}, \tag{6}$$

where $(\mathcal{H}, \|\cdot\|)$ is a real Hilbert space.

Remark 2. *A priori the Hilbert space $\mathcal{H}$ in the previous proposition is arbitrary, possibly infinite-dimensional. However by considering* $\operatorname{span}(x_i, y_j)$ *one can reduce to the situation where* $\dim\mathcal{H} = m + n$.

Proof. Consider an element $(a_{ij}) \in \mathsf{Q}_{m,n}$, of the form $a_{ij} = \operatorname{Tr}[\rho(X_i \otimes Y_j)]$ for a state $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Define a bilinear form on $B^{\mathrm{sa}}(\mathcal{H}_A \otimes \mathcal{H}_B)$ by

$$\beta(S,T) = \operatorname{Re}\operatorname{Tr}[\rho ST].$$

It can be checked (do it!) that β is positive semi-definite, and therefore after passing to a quotient, makes $B^{\mathrm{sa}}(\mathcal{H}_A \otimes \mathcal{H}_B)$ into a real Euclidean space. Note that $a_{ij} = \beta(X_i \otimes \mathrm{Id}, \mathrm{Id} \otimes Y_j)$. Since $\beta(X_i \otimes \mathrm{Id}, X_i \otimes \mathrm{Id}) = \operatorname{Tr}[\rho(X_i^2 \otimes \mathrm{Id})] \leqslant \|X_i\|_\infty^2 \leqslant 1$ and similarly $\beta(\mathrm{Id} \otimes Y_j, \mathrm{Id} \otimes Y_j) \leqslant 1$, it follows that (a_{ij}) belongs to the set described by (6).

To prove the converse inclusion, we use the following algebraic fact: for $d = 2^k$ there is a real $2k$-dimensional subspace of $\mathsf{M}_d^{\mathrm{sa}}(\mathbf{C})$ in which every matrix is a multiple of a self-adjoint unitary matrix. This can be constructed as follows: start from the Pauli matrices

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

and consider $U_1, \ldots, U_{2k}$ to be the $2k$ unitary matrices of size 2^k the form

$$\mathrm{Id} \otimes \mathrm{Id} \otimes \cdots \otimes \mathrm{Id} \otimes X \otimes Y \otimes Y \otimes \cdots \otimes Y,$$

$$\mathrm{Id} \otimes \mathrm{Id} \otimes \cdots \otimes \mathrm{Id} \otimes Z \otimes Y \otimes Y \otimes \cdots \otimes Y.$$

Theses matrices have the following property: if $V = \sum \xi_i U_i$ and $W = \sum \eta_i U_i$ for real scalars (ξ_i), (η_i), then

$$VV^* = \Big(\sum \xi_i^2\Big)\,\mathrm{Id} \quad \text{and} \quad \operatorname{Tr}(VW) = 2^k \sum_{i=1}^{2k} \xi_i \eta_i.$$

As a consequence, for any unit vectors (x_i), (y_j) in $\mathbf{R}^{2k}$, there are unitary matrices (A_i), (B_j) in $\mathbf{C}^d$ $(d = 2^k)$ such that

$$\langle x_i, y_j \rangle = \frac{1}{d} \operatorname{Tr}(A_i \overline{B_j}) = \operatorname{Tr}(\rho A_i \otimes B_j) \tag{7}$$

where $\rho = |\psi\rangle\langle\psi|$, for $\psi = \frac{1}{\sqrt{d}} \sum_{j=1}^d |e_j\rangle \otimes |e_j\rangle$ the maximally entangled vector in $\mathbf{C}^d \otimes \mathbf{C}^d$ (we let the reader check the last equality in (7)). □

A Bell inequality is a linear inequality (for $M \in \mathrm{M}_{n,m}$ and $a \in \mathbf{R}$) of the form

$$\operatorname{Tr}(AM) \leqslant a \tag{8}$$

which is valid for every $A \in \mathsf{C}_{m,n}$.

Exercise 9 (CHSH inequality)**.** *Show that for every classical correlation matrix* $(a_{ij}) \in \mathsf{C}_{2,2}$*,*

$$a_{11} + a_{21} + a_{12} - a_{22} \leqslant 2.$$

A Bell inequality of the form (8) is usually not valid for $A \in \mathsf{Q}_{m,n}$. This is because the inclusion $\mathsf{C}_{m,n} \subset \mathsf{Q}_{m,n}$ is strict in general. We illustrate this for the CHSH inequality given in Exercise 9. Consider a Bell state $\rho = |\Phi^+\rangle\langle\Phi^+|$. As in the proof of Proposition 5 we have $\mathrm{Tr}(\rho A \otimes B) = \frac{1}{2}\,\mathrm{Tr}(A\overline{B})$ for every $A, B \in \mathsf{M}_2(\mathbf{C})$. Consider now A_1, A_2, B_1, B_2 to be axial symmetries with respect to lines $\mathbf{C}x_1$, $\mathbf{C}x_1$, $\mathbf{C}y_1$, $\mathbf{C}y_2$, where the vectors x_1, y_1, x_2, y_2 (in this order) are chosen with consecutive angles $\pi/8$ (see Figure 1).

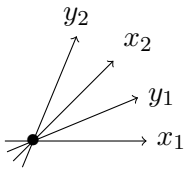

Figure 1: Angles between vectors equal $\pi/8$.

Using the result from Exercise 10, we conclude that

$$(\mathrm{Tr}(\rho(A_i \otimes B_j))_{1\leqslant i,j\leqslant 2} = \begin{pmatrix} \cos(\pi/4) & \cos(\pi/4) \\ \cos(\pi/4) & \cos(3\pi/4) \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix} \in \mathsf{Q}_{2,2}$$

If we denote by (a_{ij}) this matrix, we have $a_{11}+a_{21}+a_{12}-a_{22} = 2\sqrt{2}$. The CHSH inequality is violated by a factor of $\sqrt{2}$ for quantum correlation matrices.

Exercise 10. *Show that if R and R' are two axial symmetries in $\mathbf{C}^2$, with angle θ between both axes, then* $\mathrm{Tr}(RR') = 2\cos(2\theta)$.

A fundamental result, due to Grothendieck [3] and Tsirelson [5], states that the violations of Bell inequalities are bounded by a universal constant—the Grothendieck constant (defined to be the smallest K which makes the following theorem true).

Theorem 6 (Grothendieck–Tsirelson)**.** *There is a universal constant K such that the inclusion $\mathsf{Q}_{m,n} \subset K\mathsf{C}_{m,n}$ holds for every m, n.*

Equivalently, whenever an inequality $\mathrm{Tr}(AM) \leqslant a$ is valid for every $A \in \mathsf{C}_{m,n}$, we have $\mathrm{Tr}(AM) \leqslant Ka$ for every $A \in \mathsf{Q}_{m,n}$.

Proof. We prove the theorem in the following form, which in view of Proposition 5 is equivalent, and usually called the Grothendieck inequality : for every matrix $(m_{ij}) \in \mathsf{M}_{m,n}$,

$$\sup_{\|x_i\|\leqslant 1, \|y_j\|\leqslant 1} \sum_{i,j} m_{ij}\langle x_i, y_j\rangle \leqslant K \sup_{\xi_i,\eta_j=\pm 1} \sum_{i,j} m_{ij}\xi_i\eta_j, \tag{9}$$

where on the left-hand side the supremum is taken over vectors (x_i), (y_j) in a Hilbert space.

We rely on a couple of lemmas.

Lemma 7. *Let u, v be unit vectors in the Euclidean space $\mathbf{R}^N$, and G a standard Gaussian vector in $\mathbf{R}^N$ (i.e., the coordinates of G in any orthonormal basis are independent random variable with a Gaussian distribution of mean 0 and variance 1). Then*

$$\mathbf{E}\left[\mathrm{sign}(\langle u, G\rangle)\mathrm{sign}(\langle v, G\rangle)\right] = \frac{2}{\pi}\arcsin\langle u, v\rangle.$$

Exercise 11. *Prove Lemma 7 via 2-dimensional geometric considerations.*

Lemma 8. *Let (x_i), (y_j) be unit vectors in a real Hilbert space. Then there exists (X_i), (Y_j) unit vectors in a real Hilbert space such that, for any i, j*

$$\arcsin\langle X_i, Y_j\rangle = c\langle x_i, y_j\rangle$$

where $c = \sinh^{-1}(1) = \ln(1+\sqrt{2})$.

Assuming Lemma 8 for the moment, we show (9). By convexity the supremum in the left-hand side of (9) is achieved for unit vectors. Given unit vectors (x_i), (y_j), let (X_i), (Y_i) as in Lemma 8 and write

$$\begin{aligned}
\sum_{i,j} m_{ij}\langle x_i, y_j\rangle &= \frac{1}{c}\sum_{i,j} m_{ij}\arcsin\langle X_i, Y_j\rangle \\
&= \frac{\pi}{2c}\sum_{i,j} m_{ij}\,\mathbf{E}\left[\mathrm{sign}(\langle X_i, G\rangle)\cdot\mathrm{sign}(\langle Y_j, G\rangle)\right] \\
&= \frac{\pi}{2c}\,\mathbf{E}\left[\sum_{i,j} m_{ij}\cdot\mathrm{sign}(\langle X_i, G\rangle)\cdot\mathrm{sign}(\langle Y_j, G\rangle)\right] \\
&\leqslant \frac{\pi}{2c}\sup_{\xi_i,\eta_j=\pm 1}\left[\sum_{i,j} m_{ij}\xi_i\eta_j\right],
\end{aligned}$$

which proves Theorem 6 with $K = \pi/2c$. Incidentally, this value (proved by Krivine) was conjectured to be the optimal value, but this conjecture turned out to be wrong [2]. □

Proof of Lemma 8. Suppose that (x_i), (y_j), belong to $\mathbf{R}^N$, and let $\mathcal{H} = \bigoplus_{k\geqslant 0}(\mathbf{R}^N)^{\otimes(2k+1)}$, where the direct sum and tensor products are understood in the Hilbert space category. Define

$$X_i = \sum_{k\geqslant 0}\sqrt{\frac{c^k}{(2k+1)!}}\,x_i^{\otimes(2k+1)},$$

$$Y_j = \sum_{k\geqslant 0}(-1)^k\sqrt{\frac{c^k}{(2k+1)!}}\,y_j^{\otimes(2k+1)}.$$

One checks that $\|X_i\|^2 = \|Y_j\|^2 = \sum_k \frac{c^k}{(2k+1)!} = \sinh(c) = 1$, while

$$\langle X_i, Y_j \rangle = \sum_{k \geqslant 0} (-1)^k \frac{c^k}{(2k+1)!} \langle x_i, y_j \rangle^{2k+1} = \sin(c\langle x_i, y_j \rangle)$$

as needed. □

References

[1] G. Aubrun and S.J. Szarek. *Alice and Bob meet Banach*, volume 223 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2017. The interface of asymptotic geometric analysis and quantum information theory.

[2] Mark Braverman, Konstantin Makarychev, Yury Makarychev, and Assaf Naor. The Grothendieck constant is strictly smaller than Krivine's bound. *Forum Math. Pi*, 1:e4, 42, 2013.

[3] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953.

[4] Matthew B Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5(4):255–257, 2009.

[5] B. S. Tsirelson. Quantum analogues of Bell's inequalities. The case of two spatially divided domains. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 142:174–194, 200, 1985. Problems of the theory of probability distributions, IX.

[6] John Watrous. *The Theory of Quantum Information.* Cambridge University Press, 2018.

[7] Dong Yang, Michał Horodecki, Ryszard Horodecki, and Barbara Synak-Radtke. Irreversibility for all bound entangled states. *Phys. Rev. Lett.*, 95:190501, Oct 2005.