Lecture notes "Quantum Information & Dynamics" – Part 2

Cécilia Lancien

November 5th 2025

Chapter 1

Reminder/Notation

1.1 Matrix p-norms

Let $\mathcal{H} \equiv \mathbb{C}^d$ be a finite-dimensional complex Hilbert space. We denote by $B(\mathcal{H})$ the set of (bounded) linear operators on $\mathcal{H}$, i.e. of $d \times d$ complex matrices. Given $X \in B(\mathcal{H})$, its Schatten p-norm is defined as

$$\|X\|_p = \left(\mathrm{Tr}\left(|X|^p\right)\right)^{1/p} \text{ for } p \in \mathbb{N} \text{ and } \|X\|_\infty = \lim_{p\to\infty} \|X\|_p,$$

where $|X| = (X^*X)^{1/2}$. Equivalently, denoting by $s_1(X) \geqslant \cdots \geqslant s_d(X) \geqslant 0$ the singular values of X (i.e. the eigenvalues of $|X|$), we have

$$\|X\|_p = \left(\sum_{i=1}^{d} s_i(X)^p\right)^{1/p} \text{ for } p \in \mathbb{N} \text{ and } \|X\|_\infty = s_1(X).$$

The most notable cases are $p = 1$ (trace norm), $p = 2$ (Hilbert-Schmidt norm), $p = \infty$ (operator norm). The 2-norm is a Euclidean norm, associated to the scalar product

$$\forall\ X, Y \in B(\mathcal{H}),\ \langle X, Y\rangle = \mathrm{Tr}(X^*Y). \tag{1.1}$$

Important properties of these Schatten p-norms include:

1. Ordering: for all $1 \leqslant p \leqslant q \leqslant \infty$, $\|\cdot\|_q \leqslant \|\cdot\|_p$.
2. Duality (with respect to the scalar product introduced in equation (1.1)): for all $1 \leqslant p \leqslant \infty$, $\|X\|_p = \sup\{|\mathrm{Tr}(X^*Y)| : \|Y\|_{p'} = 1\}$, where $1/p + 1/p' = 1$. In particular, $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are dual to one another, while $\|\cdot\|_2$ is self-dual.

1.2 Prerequisites on the mathematical formalism of quantum physics

A system is described by a complex Hilbert space $\mathcal{H}$. We will only consider finite-dimensional systems, so that $\mathcal{H}$ can be identified with $\mathbb{C}^d$ for some $d \in \mathbb{N}$.

A state of system $\mathcal{H}$ is described by a density operator, i.e. a Hermitian, positive semi-definite operator on $\mathcal{H}$ that has trace 1. We denote by $\mathrm{D}(\mathcal{H})$ the set of density operators on $\mathcal{H}$, i.e.

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

Equivalently, $\mathrm{D}(\mathcal{H})$ is the convex hull of rank 1 projectors on $\mathcal{H}$, i.e.

$$\mathrm{D}(\mathcal{H}) = \mathrm{conv}\left\{|\varphi\rangle\langle\varphi| : \varphi \in \mathcal{H},\ \|\varphi\| = 1\right\}.$$

States of the form $|\varphi\rangle\langle\varphi|$, for $\varphi \in \mathcal{H}$ a unit vector, are called pure states and are exactly the extreme points of $\mathrm{D}(\mathcal{H})$.

An observable of system $\mathcal{H}$ is described by a Hermitian operator on $\mathcal{H}$, i.e. $X \in B(\mathcal{H})$ such that $X^* = X$. Given an observable X, its expectation value when measured on a system in state ρ is

$$\langle X\rangle_\rho = \mathrm{Tr}(A\rho).$$

A measurement of system $\mathcal{H}$ is described by a collection of Hermitian positive semi-definite operators on $\mathcal{H}$ that sum up to the identity, i.e. $\mathrm{M} = \{M_i\}_{i\in I}$ such that, for all $i \in I$, $M_i^* = M_i$ and $M_i \geqslant 0$, and $\sum_{i\in I} M_i = \mathrm{I}$. Given a measurement $\mathrm{M} = \{M_i\}_{i\in I}$, the probability that outcome $i \in I$ is obtained when it is measured on a system in state ρ is

$$\mathbb{P}_\rho(i) = \mathrm{Tr}(M_i\rho).$$

It is easy to check that $\{\mathbb{P}_\rho(i)\}_{i\in I}$ is indeed a probability distribution: for all $i \in I$, $\mathbb{P}_\rho(i) = \mathrm{Tr}(M_i\rho) \geqslant 0$ because $M_i, \rho \geqslant 0$, and $\sum_{i\in I} \mathbb{P}_\rho(i) = \sum_{i\in I} \mathrm{Tr}(M_i\rho) = \mathrm{Tr}((\sum_{i\in I} M_i)\rho) = 1$ because $\sum_{i\in I} M_i = \mathrm{I}$ and $\mathrm{Tr}(\rho) = 1$.

1.3 Quantum information quantities

Let $\mathcal{H} \equiv \mathbb{C}^d$ be a finite-dimensional complex Hilbert space.

1.3.1 Quantum entropy and relative entropy

Given $\rho \in \mathrm{D}(\mathcal{H})$, its eigenvalues $\lambda_1(\rho), \ldots, \lambda_d(\rho)$ are non-negative and summing up to 1, so $\lambda(\rho) = \{\lambda_1(\rho), \ldots, \lambda_d(\rho)\}$ can be viewed as a finite probability distribution. The von Neumann entropy of ρ is defined as

$$S(\rho) = -\mathrm{Tr}(\rho\log(\rho)).$$

It can be seen as the Shannon entropy of $\lambda(\rho)$. Indeed, we have

$$S(\rho) = -\mathrm{Tr}(\rho\log(\rho)) = -\sum_{i=1}^{d} \lambda_i(\rho)\log(\lambda_i(\rho)) = H(\lambda(\rho)).$$

It is easy to see that we always have

$$0 \leqslant S(\rho) \leqslant \log(d),$$

with equality in the first inequality iff ρ is pure and in the second inequality iff $\rho = \mathrm{I}/d$ is maximally mixed.

Given $\rho, \sigma \in \mathrm{D}(\mathcal{H})$, their relative entropy is defined as

$$S(\rho\|\sigma) = \mathrm{Tr}(\rho(\log(\rho) - \log(\sigma))).$$

$S(\rho\|\sigma)$ is always non-negative. And if $\ker(\sigma) \nsubseteq \ker(\rho)$, then $S(\rho\|\sigma)$ is infinite. The following lower bound on the relative entropy between two states, in terms of their 1-norm distance, is known as Pinsker inequality. As a direct consequence we have that $S(\rho\|\sigma) = 0$ iff $\rho = \sigma$. So $S(\rho\|\sigma)$ can be seen as a distance measure between two states ρ and σ.

Proposition 1.3.1. *For any $\rho, \sigma \in \mathrm{D}(\mathcal{H})$, we have*

$$S(\rho\|\sigma) \geqslant \frac{1}{2\log(2)}\|\rho - \sigma\|_1^2.$$

As a corollary of Proposition 1.3.1 we have that the von Neumann entropy is strongly concave.

Corollary 1.3.2. *For any $\rho, \rho' \in \mathrm{D}(\mathcal{H})$ and any $0 \leqslant \lambda \leqslant 1$, we have*

$$S(\lambda\rho + (1-\lambda)\rho') \geqslant \lambda S(\rho) + (1-\lambda)S(\rho') + \frac{\lambda(1-\lambda)}{2\log(2)}\|\rho - \rho'\|_1^2.$$

Proof. To prove Corollary 1.3.2, we just have to observe that, by definition of the relative entropy, we have the identity

$$S(\lambda\rho + (1-\lambda)\rho') - \lambda S(\rho) - (1-\lambda)S(\rho') = \lambda S(\rho\|\lambda\rho + (1-\lambda)\rho') + (1-\lambda)S(\rho'\|\lambda\rho + (1-\lambda)\rho').$$

Now, by Proposition 1.3.1, we know that the right-hand side of the above equality is lower bounded by

$$\frac{\lambda}{2\log(2)}\|\rho - (\lambda\rho + (1-\lambda)\rho')\|_1^2 + \frac{1-\lambda}{2\log(2)}\|\rho' - (\lambda\rho + (1-\lambda)\rho')\|_1^2 = \frac{\lambda(1-\lambda)}{2\log(2)}\|\rho - \rho'\|_1^2,$$

and this concludes the proof. □

Another important property of relative entropy is its joint convexity, as explained below.

Proposition 1.3.3. *For any $\rho, \rho', \sigma, \sigma' \in \mathrm{D}(\mathcal{H})$ and any $0 \leqslant \lambda \leqslant 1$, we have*

$$S(\lambda\rho + (1-\lambda)\rho'\|\lambda\sigma + (1-\lambda)\sigma') \leqslant \lambda S(\rho\|\sigma) + (1-\lambda)S(\rho'\|\sigma').$$

1.3.2 Quantum fidelity

Given $\rho, \sigma \in \mathrm{D}(\mathcal{H})$, their fidelity is defined as

$$F(\rho, \sigma) = \mathrm{Tr}\left(\left(\rho^{1/2}\sigma\rho^{1/2}\right)^{1/2}\right) = \left\|\rho^{1/2}\sigma^{1/2}\right\|_1 .$$

$F(\rho, \sigma)$ takes values between 0 and 1, with $F(\rho, \sigma) = 0$ iff ρ and σ are orthogonal and $F(\rho, \sigma) = 1$ iff $\rho = \sigma$. The fidelity between two states is related to their 1-norm distance through so-called Fuchs / van de Graaf inequalities.

Proposition 1.3.4. *For any $\rho, \sigma \in \mathrm{D}(\mathcal{H})$, we have*

$$1 - F(\rho, \sigma) \leqslant \frac{1}{2}\|\rho - \sigma\|_1 \leqslant \left(1 - F(\rho, \sigma)^2\right)^{1/2} .$$

Chapter 2

Multipartite quantum systems

Main references for this chapter:

- [1] Chapter 0 Section 0.4 and Chapter 2 Section 2.2.
- [5] Chapter 1 Section 1.5.

2.1 Tensor product of Hilbert spaces

2.1.1 First definitions

Let $\mathcal{H}_1, \ldots, \mathcal{H}_n$ be complex finite-dimensional Hilbert spaces. For each $1 \leqslant k \leqslant n$, denote by $\{e_1^{(k)}, \ldots, e_{d_k}^{(k)}\}$ an orthonormal basis of $\mathcal{H}_k$, where $d_k = \dim(\mathcal{H}_k)$. Then, the tensor product of $\mathcal{H}_1, \ldots, \mathcal{H}_n$ is defined as the complex finite-dimensional vector space $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ having as orthonormal basis

$$\left\{ e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_n}^{(n)},\ 1 \leqslant i_1 \leqslant d_1, \ldots, 1 \leqslant i_n \leqslant d_n \right\}.$$

This means that any vector $\varphi \in \mathcal{H}$ can be written as

$$\varphi = \sum_{1 \leqslant i_1 \leqslant d_1, \ldots, 1 \leqslant i_n \leqslant d_n} \varphi_{i_1 \cdots i_n} e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_n}^{(n)},$$

where the $\varphi_{i_1 \cdots i_n}$'s are complex numbers. Observe also that, by definition,

$$\dim(\mathcal{H}) = \dim(\mathcal{H}_1) \times \cdots \times \dim(\mathcal{H}_n).$$

Given vectors $\varphi_k = \sum_{i_k=1}^{d_k} \varphi_{i_k}^{(k)} e_{i_k}^{(k)} \in \mathcal{H}_k$, for each $1 \leqslant k \leqslant n$, their tensor product vector $\varphi_1 \otimes \cdots \otimes \varphi_n \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ is defined as

$$\varphi_1 \otimes \cdots \otimes \varphi_n = \sum_{1 \leqslant i_1 \leqslant d_1, \ldots, 1 \leqslant i_n \leqslant d_n} \varphi_{i_1}^{(1)} \ldots \varphi_{i_n}^{(n)} e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_n}^{(n)}.$$

In other words, we have a multilinear map

$$(\varphi_1, \ldots, \varphi_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \mapsto \varphi_1 \otimes \cdots \otimes \varphi_n \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n,$$

and $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ can be viewed as the linear span (over the complex field) of product vectors, i.e.

$$\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n = \operatorname{span}\left\{ \varphi_1 \otimes \cdots \otimes \varphi_n : \varphi_1 \in \mathcal{H}_1, \ldots, \varphi_n \in \mathcal{H}_n \right\}.$$

$\mathcal{H}$ carries a natural Hilbert space structure, inherited from the Hilbert space structure of $\mathcal{H}_1, \ldots, \mathcal{H}_n$. It is given by the inner product defined for product vectors by

$$\forall\ \varphi_1, \chi_1 \in \mathcal{H}_1, \ldots, \varphi_n, \chi_n \in \mathcal{H}_n,\ \langle \varphi_1 \otimes \cdots \otimes \varphi_n | \chi_1 \otimes \cdots \otimes \chi_n \rangle_{\mathcal{H}} = \langle \varphi_1 | \chi_1 \rangle_{\mathcal{H}_1} \times \cdots \times \langle \varphi_n | \chi_n \rangle_{\mathcal{H}_n},$$

and extended to $\mathcal{H}$ by sesquilinearity. $\mathcal{H}$ is often called a multipartite, or more precisely n-partite, Hilbert space (or bipartite Hilbert space when $n = 2$).

Remark 2.1.1. *Let us illustrate the above definitions in the bipartite case, i.e. $n = 2$. Denoting by $\{e_1, \dots, e_{d_1}\}$ and $\{f_1, \dots, f_{d_2}\}$ orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$, $\{e_{i_1} \otimes f_{i_2} : 1 \leqslant i_1 \leqslant d_1,\, 1 \leqslant i_2 \leqslant d_2\}$ is an orthonormal basis of $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. This means that any vector $\varphi \in \mathcal{H}$ can be written as*

$$\varphi = \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \varphi_{i_1 i_2} e_{i_1} \otimes f_{i_2}.$$

If we have two vectors $\chi = \sum_{i=1}^{d_1} \chi_i e_i \in \mathcal{H}_1$ and $\omega = \sum_{i=1}^{d_2} \omega_i f_i \in \mathcal{H}_2$, their tensor product is simply

$$\chi \otimes \omega = \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \chi_{i_1} \omega_{i_2} e_{i_1} \otimes f_{i_2}.$$

Example 2.1.2. *Let us write things down even more concretely in the simplest case where $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}^2$. We then have $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2 \equiv \mathbb{C}^4$, but instead of labeling its orthonormal basis by the indices $1, 2, 3, 4$, we label it by the indices $(1,1), (1,2), (2,1), (2,2)$. Hence, we write $\varphi \in \mathbb{C}^2 \otimes \mathbb{C}^2$ as*

$$\varphi = \begin{pmatrix} \varphi_{11} \\ \varphi_{12} \\ \varphi_{21} \\ \varphi_{22} \end{pmatrix}.$$

We can identify the set of bounded operators on $\mathcal{H}$ with the tensor product of the sets of bounded operators on the $\mathcal{H}_k$'s, i.e.

$$B\left(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n\right) \equiv B(\mathcal{H}_1) \otimes \cdots \otimes B(\mathcal{H}_n).$$

Indeed, the vector space on the right-hand side is clearly included in the vector space on the left-hand side, and both have the same dimension $d_1^2 \times \cdots \times d_n^2$ (over the complex field). Concretely, we can identify their orthonormal bases: for all $1 \leqslant i_1, j_1 \leqslant d_1, \dots, 1 \leqslant i_n, j_n \leqslant d_n$,

$$|e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_n}^{(n)}\rangle\langle e_{j_1}^{(1)} \otimes \cdots \otimes e_{j_n}^{(n)}| \equiv |e_{i_1}^{(1)}\rangle\langle e_{j_1}^{(1)}| \otimes \cdots \otimes |e_{i_n}^{(n)}\rangle\langle e_{j_n}^{(n)}|.$$

Remark 2.1.3. *Let us again illustrate things better in the bipartite case. Denoting by $\{e_1, \dots, e_{d_1}\}$ and $\{f_1, \dots, f_{d_2}\}$ orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$, $\{|e_{i_1}\rangle\langle e_{j_1}| \otimes |f_{i_2}\rangle\langle f_{j_2}| : 1 \leqslant i_1, j_1 \leqslant d_1,\, 1 \leqslant i_2, j_2 \leqslant d_2\}$ is an orthonormal basis of $B(\mathcal{H}) = B(\mathcal{H}_1 \otimes \mathcal{H}_2)$. This means that any operator $X \in B(\mathcal{H})$ can be written as*

$$X = \sum_{i_1, j_1=1}^{d_1} \sum_{i_2, j_2=1}^{d_2} X_{i_1 i_2 j_1 j_2} |e_{i_1}\rangle\langle e_{j_1}| \otimes |f_{i_2}\rangle\langle f_{j_2}|.$$

If we have two operators $Y = \sum_{i,j=1}^{d_1} Y_{ij} |e_i\rangle\langle e_j| \in B(\mathcal{H}_1)$ and $Z = \sum_{i,j=1}^{d_2} Z_{ij} |f_i\rangle\langle f_j| \in B(\mathcal{H}_2)$, their tensor product is simply

$$Y \otimes Z = \sum_{i_1, j_1=1}^{d_1} \sum_{i_2, j_2=1}^{d_2} Y_{i_1 j_1} Z_{i_2 j_2} |e_{i_1}\rangle\langle e_{j_1}| \otimes |f_{i_2}\rangle\langle f_{j_2}|.$$

A sometimes convenient way of writing a bipartite matrix $X \in \mathcal{M}_{d_1 d_2}(\mathbb{C})$ is as a block matrix $X = (X_{ij})_{1 \leqslant i,j \leqslant d_1}$, where $X_{ij} \in \mathcal{M}_{d_2}(\mathbb{C})$ for each $1 \leqslant i, j \leqslant d_1$. Concretely, we have

$$X = \begin{pmatrix} X_{11} & \cdots & X_{1 d_1} \\ \vdots & & \vdots \\ X_{d_1 1} & \cdots & X_{d_1 d_1} \end{pmatrix},$$

where, for each $1 \leqslant i, j \leqslant d_1$,

$$X_{ij} = \begin{pmatrix} X_{i1j1} & \cdots & X_{i1jd_2} \\ \vdots & & \vdots \\ X_{id_2j1} & \cdots & X_{id_2jd_2} \end{pmatrix}.$$

Example 2.1.4. *In the simplest case where $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}^2$, we have $B(\mathcal{H}) = B(\mathbb{C}^2 \otimes \mathbb{C}^2) \equiv \mathcal{M}_4(\mathbb{C})$. And we write $X \in B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ as*

$$X = \begin{pmatrix} X_{1111} & X_{1112} & X_{1121} & X_{1122} \\ X_{1211} & X_{1212} & X_{1221} & X_{1222} \\ X_{2111} & X_{2112} & X_{2121} & X_{2122} \\ X_{2211} & X_{2212} & X_{2221} & X_{2222} \end{pmatrix} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}.$$

Note that, given operators $X_1 \in B(\mathcal{H}_1), \ldots, X_n \in B(\mathcal{H}_n)$, the action of the product operator $X_1 \otimes \cdots \otimes X_n$ on product vectors factorizes, i.e. for any vectors $\varphi_1 \in \mathcal{H}_1, \ldots, \varphi_n \in \mathcal{H}_n$,

$$(X_1 \otimes \cdots \otimes X_n)(\varphi_1 \otimes \cdots \otimes \varphi_n) = (X_1\varphi_1) \otimes \cdots \otimes (X_n\varphi_n).$$

We also have that the product of two product operators $X_1 \otimes \cdots \otimes X_n$ and $Y_1 \otimes \cdots \otimes Y_n$ factorizes as

$$(X_1 \otimes \cdots \otimes X_n)(Y_1 \otimes \cdots \otimes Y_n) = (X_1Y_1) \otimes \cdots \otimes (X_nY_n).$$

What is more, one can check that the eigenvalues, resp. singular values, of a product operator are the product of the eigenvalues, resp. singular values, of the individual operators. As a consequence, the tensor product satisfies the following important properties.

Proposition 2.1.5. *First, the tensor product of Hermitian, resp. positive semi-definite, operators, is a Hermitian, resp. positive semi-definite, operator. Second, the Schatten p-norms and the trace of a product operator factorize. Concretely, this means that, for any $X_1 \in B(\mathcal{H}_1), \ldots, X_n \in B(\mathcal{H}_n)$,*

1. *$(X_1 \otimes \cdots \otimes X_n)^* = X_1^* \otimes \cdots \otimes X_n^*$, so that $X_1^* = X_1, \ldots, X_n^* = X_n$ implies $(X_1 \otimes \cdots \otimes X_n)^* = X_1 \otimes \cdots \otimes X_n$;*
2. *$X_1 \geqslant 0, \ldots, X_n \geqslant 0$ implies $X_1 \otimes \cdots \otimes X_n \geqslant 0$;*
3. *for all $1 \leqslant p \leqslant \infty$, $\|X_1 \otimes \cdots \otimes X_n\|_p = \|X_1\|_p \cdots \|X_n\|_p$;*
4. *$\mathrm{Tr}(X_1 \otimes \cdots \otimes X_n) = \mathrm{Tr}(X_1) \cdots \mathrm{Tr}(X_n)$.*

2.1.2 Partial trace

A fundamental concept when dealing with multipartite Hilbert spaces is that of partial trace. Given a Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ and an index $1 \leqslant k \leqslant n$, the partial trace over $\mathcal{H}_k$, denoted $\mathrm{Tr}_{\mathcal{H}_k}$, is the linear map from $B(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$ to $B(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_{k-1} \otimes \mathcal{H}_{k+1} \otimes \cdots \otimes \mathcal{H}_n)$ that acts as the trace on $B(\mathcal{H}_k)$ and the identity on the other tensor factors, i.e.

$$\mathrm{Tr}_{\mathcal{H}_k} = \mathrm{id}_{B(\mathcal{H}_1)} \otimes \cdots \otimes \mathrm{id}_{B(\mathcal{H}_{k-1})} \otimes \mathrm{Tr}_{B(\mathcal{H}_k)} \otimes \mathrm{id}_{B(\mathcal{H}_{k+1})} \otimes \cdots \otimes \mathrm{id}_{B(\mathcal{H}_n)}.$$

For simplicity of notation, let us focus on the bipartite case, i.e. $n = 2$. The partial trace over (say) $\mathcal{H}_2$ is the linear map $\mathrm{Tr}_{\mathcal{H}_2} : B(\mathcal{H}_1 \otimes \mathcal{H}_2) \to B(\mathcal{H}_1)$ that acts on product operators as

$$\mathrm{Tr}_{\mathcal{H}_2}(X_1 \otimes X_2) = \mathrm{Tr}(X_2)\, X_1,$$

and is extended to $B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ by linearity. It can be equivalently characterized as being the unique linear map from $B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ to $B(\mathcal{H}_1)$ satisfying

$$\forall\, X \in B(\mathcal{H}_1 \otimes \mathcal{H}_2), Y \in B(\mathcal{H}_1),\ \mathrm{Tr}(X(Y \otimes \mathrm{I})) = \mathrm{Tr}\left(\mathrm{Tr}_{\mathcal{H}_2}(X)Y\right). \tag{2.1}$$

Concretely, denoting by $\{e_1, \ldots, e_{d_1}\}$ and $\{f_1, \ldots, f_{d_2}\}$ orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$, an operator $X \in B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ can always be written as

$$X = \sum_{i_1, j_1 = 1}^{d_1} \sum_{i_2, j_2 = 1}^{d_2} X_{i_1 i_2 j_1 j_2} |e_{i_1}\rangle\langle e_{j_1}| \otimes |f_{i_2}\rangle\langle f_{j_2}|,$$

and we have

$$\mathrm{Tr}_{\mathcal{H}_2}(X) = \sum_{i_1, j_1 = 1}^{d_1} \left(\sum_{i_2 = 1}^{d_2} X_{i_1 i_2 j_1 i_2} \right) |e_{i_1}\rangle\langle e_{j_1}| = \sum_{i_2 = 1}^{d_2} \langle f_{i_2}|X|f_{i_2}\rangle.$$

However, $\mathrm{Tr}_{\mathcal{H}_2}$ is independent of the orthonormal basis $\{f_1, \ldots, f_{d_2}\}$ of $\mathcal{H}_2$ that is chosen to perform the trace.

Remark 2.1.6. *The partial trace can be easily described in the block representation introduced in Remark 2.1.3. Indeed, if we write $Z \in \mathcal{M}_{d_1 d_2}(\mathbb{C})$ as a block matrix $Z = (Z_{ij})_{1 \leqslant i,j \leqslant d_1}$, where $Z_{ij} \in \mathcal{M}_{d_2}(\mathbb{C})$ for each $1 \leqslant i, j \leqslant d_1$, we have $\mathrm{Tr}_{\mathcal{H}_2}(Z) = (\mathrm{Tr}(Z_{ij}))_{1 \leqslant i,j \leqslant d_1} \in \mathcal{M}_{d_1}(\mathbb{C})$.*

Example 2.1.7. *If we look at the case where $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}^2$, we see that, given $X \in B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ as in Example 2.1.4, $\mathrm{Tr}_{\mathcal{H}_2}(X) \in B(\mathbb{C}^2) \equiv \mathcal{M}_2(\mathbb{C})$ is simply*

$$\mathrm{Tr}_{\mathcal{H}_2}(X) = \begin{pmatrix} X_{1111} + X_{1212} & X_{1121} + X_{1222} \\ X_{2111} + X_{2212} & X_{2121} + X_{2222} \end{pmatrix} = \begin{pmatrix} \mathrm{Tr}(X_{11}) & \mathrm{Tr}(X_{12}) \\ \mathrm{Tr}(X_{21}) & \mathrm{Tr}(X_{22}) \end{pmatrix}.$$

Proposition 2.1.8. *The partial trace preserves the trace, hermiticity and positivity. Concretely, this means that, for any $1 \leqslant k \leqslant n$, we have, for any $X \in B(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$,*

1. $\mathrm{Tr}\left(\mathrm{Tr}_{\mathcal{H}_k}(X)\right) = \mathrm{Tr}(X)$;
2. $\left(\mathrm{Tr}_{\mathcal{H}_k}(X)\right)^* = \mathrm{Tr}_{\mathcal{H}_k}(X^*)$, *so that* $X^* = X$ *implies* $\left(\mathrm{Tr}_{\mathcal{H}_k}(X)\right)^* = \mathrm{Tr}_{\mathcal{H}_k}(X)$;
3. $X \geqslant 0$ *implies* $\mathrm{Tr}_{\mathcal{H}_k}(X) \geqslant 0$.

Proof. We can suppose without loss of generality that $n = 2$, i.e. $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, and $k = 2$, i.e. the partial trace is performed over $\mathcal{H}_2$ (otherwise, we simply have to set $\mathcal{H}'_1 = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_{k-1} \otimes \mathcal{H}_{k+1} \otimes \cdots \otimes \mathcal{H}_n$, $\mathcal{H}'_2 = \mathcal{H}_k$ and observe that $\mathcal{H} = \mathcal{H}'_1 \otimes \mathcal{H}'_2$). Points (1) and (2) are easy to check writing things in coordinates, as done above. As for point (3), we will use the characterization of the partial trace provided by equation (2.1). Suppose that $X \in B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is positive semi-definite. Equivalently, this means that, for any $Y \in B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ that is positive semi-definite, $\mathrm{Tr}(XY) \geqslant 0$. In particular, for any $Z \in B(\mathcal{H}_1)$ that is positive semi-definite, $Z \otimes \mathrm{I} \in B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is positive semi-definite, and hence $\mathrm{Tr}(X(Z \otimes \mathrm{I})) \geqslant 0$. Now, by equation (2.1), $\mathrm{Tr}(X(Z \otimes \mathrm{I})) = \mathrm{Tr}(\mathrm{Tr}_{\mathcal{H}_2}(X)Z)$. So we have shown that, for any $Z \in B(\mathcal{H}_1)$ that is positive semi-definite, $\mathrm{Tr}(\mathrm{Tr}_{\mathcal{H}_2}(X)Z) \geqslant 0$, which equivalently means that $\mathrm{Tr}_{\mathcal{H}_2}(X) \in B(\mathcal{H}_1)$ is positive semi-definite. □

2.1.3 Schmidt decomposition of a bipartite vector

We focus on the case where $n = 2$, i.e. $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \equiv \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ is a bipartite Hilbert space.

As we have already explained, denoting by $\{e_1, \ldots, e_{d_1}\}$ and $\{f_1, \ldots, f_{d_2}\}$ orthonormal bases of $\mathbb{C}^{d_1}$ and $\mathbb{C}^{d_2}$, a vector $\varphi \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ can always be written as

$$\varphi = \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \varphi_{i_1 i_2} e_{i_1} \otimes f_{i_2}.$$

We can thus identify φ with a matrix $M_\varphi \in \mathcal{M}_{d_1,d_2}(\mathbb{C})$, defined as

$$M_\varphi = \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \varphi_{i_1 i_2} |e_{i_1}\rangle\langle f_{i_2}|.$$

We recall the notion of singular value decomposition of a matrix $M \in \mathcal{M}_{d_1,d_2}(\mathbb{C})$: Let $d = \min(d_1, d_2)$. There exist (uniquely determined) non-negative numbers $s_1 \geqslant \cdots \geqslant s_d$ as well as orthonormal families $\{u_1, \ldots, u_d\}$ and $\{v_1, \ldots, v_d\}$ in $\mathbb{C}^{d_1}$ and $\mathbb{C}^{d_2}$ such that

$$M = \sum_{i=1}^{d} s_i |u_i\rangle\langle v_i|.$$

The s_i's are called the singular values of M and are in fact the eigenvalues of $(M^*M)^{1/2}$ if $d_1 \geqslant d_2$ and $(MM^*)^{1/2}$ if $d_2 \geqslant d_1$. Translating this to the corresponding vector $\varphi_M \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$, we obtain its so-called Schmidt decomposition, i.e. a decomposition of the form

$$\varphi_M = \sum_{i=1}^{d} s_i u_i \otimes v_i,$$

where the s_i's are called the Schmidt coefficients of φ_M.

Given a vector $\varphi \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$, with Schmidt decomposition

$$\varphi = \sum_{i=1}^{d} s_i u_i \otimes v_i,$$

the smallest $r \leqslant d$ such that, for all $i > r$, $s_i = 0$, is called the Schmidt rank of φ, and denoted $\mathrm{SR}(\varphi)$. It is equal to the rank, as an operator, of the corresponding matrix M_φ.

Note that this identification between elements of $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ and $\mathcal{M}_{d_1,d_2}(\mathbb{C})$ is an isomorphism between Hilbert spaces, that preserves the scalar product, i.e.

$$\forall\ \varphi, \chi \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2},\ \langle \varphi | \chi \rangle = \mathrm{Tr}\left(M_\varphi^* M_\chi\right).$$

The Euclidean norms of a vector $\varphi \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ and its matrix version $M_\varphi \in \mathcal{M}_{d_1,d_2}(\mathbb{C})$ are thus the same, i.e.

$$\|\varphi\| = \|M_\varphi\|_2 = \left(\sum_{i=1}^{d} s_i^2\right)^{1/2}.$$

Besides, the operator norm of M_φ corresponds to the so-called maximal overlap with product tensors of φ, i.e.

$$\sup\left\{\langle \varphi | \chi_1 \otimes \chi_2 \rangle : \chi_1 \in \mathbb{C}^{d_1},\ \chi_2 \in \mathbb{C}^{d_2},\ \|\chi_1\| = \|\chi_2\| = 1\right\} = \|M_\varphi\|_\infty = s_1.$$

Example 2.1.9. *If we look at the case where $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}^2$, we see that, given $\varphi \in \mathbb{C}^2 \otimes \mathbb{C}^2$ as in Example 2.1.2, $M_\varphi \in \mathcal{M}_2(\mathbb{C})$ is simply*

$$M_\varphi = \begin{pmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{pmatrix}.$$

2.1.4 Mathematical description of multipartite quantum systems

If a quantum system is composed of n distinguished subsystems, that are individually described by Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$, then it is globally described by the tensor product Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$. Here 'distinguished subsystems' might refer to either spatially separated parts of a system (e.g. two particles) or to different degrees of freedom of a system (e.g. spin and position).

A state of such multipartite quantum system $\mathcal{H}$ is described by a density operator ρ on $\mathcal{H}$, i.e. $\rho \in B(\mathcal{H})$ such that $\rho^* = \rho$, $\rho \geqslant 0$ and $\mathrm{Tr}(\rho) = 1$. Given $1 \leqslant k \leqslant n$, the state of subsystem k alone, discarding all other subsystems, is given by the so-called reduced state ρ_k on $\mathcal{H}_k$, i.e.

$$\rho_k = \mathrm{Tr}_{\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_{k-1} \otimes \mathcal{H}_{k+1} \otimes \cdots \otimes \mathcal{H}_n}(\rho).$$

By Proposition 2.1.8, we know that ρ_k is indeed a valid density operator, since the partial trace of a Hermitian, positive semi-definite trace 1 operator remains a Hermitian, positive semi-definite trace 1 operator. More generally, given a subset of indices $K \subset \{1, \ldots, n\}$, the state of subsystems whose index is contained in K, discarding all subsystems whose index is not contained in K, is given by the reduced state ρ_K on $\bigotimes_{k \in K} \mathcal{H}_k$, i.e.

$$\rho_K = \mathrm{Tr}_{\bigotimes_{k \notin K} \mathcal{H}_k}(\rho).$$

Remark 2.1.10. *Let us make a brief analogy between quantum and classical states, i.e. between density operators and probability distributions. The analog of a bipartite density operator ρ, on a product Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$, is the joint probability distribution P of a pair of random variables (X_1, X_2), taking values in a product space $\{1, \ldots, d_1\} \times \{1, \ldots, d_2\}$. In the latter setting, the probability distribution of X_1 alone is the marginal probability distribution P_1 obtained by summing P over all values of X_2, i.e.*

$$\forall\ 1 \leqslant x_1 \leqslant d_1,\ P_1(x_1) = \sum_{x_2=1}^{d_2} P(x_1, x_2).$$

This corresponds to taking the partial trace over $\mathcal{H}_2$ of ρ in order to obtain the density operator ρ_1 on $\mathcal{H}_1$ alone.

2.2 States on tensor product Hilbert spaces

Let $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ be a multipartite Hilbert space.

2.2.1 Entanglement vs separability

A pure state $\rho = |\varphi\rangle\langle\varphi|$ on $\mathcal{H}$ is called separable if the unit vector $\varphi \in \mathcal{H}$ is a product vector, i.e. if there exist unit vectors $\varphi_1 \in \mathcal{H}_1, \ldots, \varphi_n \in \mathcal{H}_n$ such that $\varphi = \varphi_1 \otimes \cdots \otimes \varphi_n$. In this case, we have

$$\rho = |\varphi_1\rangle\langle\varphi_1| \otimes \cdots \otimes |\varphi_n\rangle\langle\varphi_n|.$$

A mixed state ρ on $\mathcal{H}$ is called separable if it can be written as a convex combination of pure separable states, i.e. if there exist unit vectors $\varphi_1^{(i)} \in \mathcal{H}_1, \ldots, \varphi_n^{(i)} \in \mathcal{H}_n$, $1 \leqslant i \leqslant r$, and non-negative numbers $\lambda_1, \ldots, \lambda_r$ summing up to 1 such that

$$\rho = \sum_{i=1}^{r} \lambda_i |\varphi_1^{(i)}\rangle\langle\varphi_1^{(i)}| \otimes \cdots \otimes |\varphi_n^{(i)}\rangle\langle\varphi_n^{(i)}|.$$

We denote by $\mathrm{S}(\mathcal{H}) \subset \mathrm{D}(\mathcal{H})$ the set of separable states on $\mathcal{H}$. We thus have by definition

$$\mathrm{S}(\mathcal{H}) = \mathrm{conv}\left\{|\varphi_1\rangle\langle\varphi_1| \otimes \cdots \otimes |\varphi_n\rangle\langle\varphi_n| : \varphi_1 \in \mathcal{H}_1, \ldots, \varphi_n \in \mathcal{H}_n,\ \|\varphi_1\| = \cdots = \|\varphi_n\| = 1\right\},$$

and pure product states are exactly the extreme points of the convex set $\mathrm{S}(\mathcal{H})$ (as pure states are exactly the extreme points of the convex set $\mathrm{D}(\mathcal{H})$). Since $\mathrm{D}(\mathcal{H}_k) = \mathrm{conv}\{|\varphi_k\rangle\langle\varphi_k| : \varphi_k \in \mathcal{H}_k,\ \|\varphi_k\| = 1\}$, for each $1 \leqslant k \leqslant n$, $\mathrm{S}(\mathcal{H})$ can in fact be equivalently described as

$$\mathrm{S}(\mathcal{H}) = \mathrm{conv}\left\{\rho_1 \otimes \cdots \otimes \rho_n : \rho_1 \in \mathrm{D}(\mathcal{H}_1), \ldots, \rho_n \in \mathrm{D}(\mathcal{H}_n)\right\}.$$

Definition 2.2.1. *A state ρ on $\mathcal{H}$ is called separable if it can be written as a convex combination of product states, i.e. if there exist states $\rho_1^{(i)} \in \mathrm{D}(\mathcal{H}_1), \ldots, \rho_n^{(i)} \in \mathrm{D}(\mathcal{H}_n)$, $1 \leqslant i \leqslant r$, and non-negative numbers $\lambda_1, \ldots, \lambda_r$ summing up to 1 such that*

$$\rho = \sum_{i=1}^{r} \lambda_i \rho_1^{(i)} \otimes \cdots \otimes \rho_n^{(i)}. \tag{2.2}$$

A state ρ that is not separable (i.e. that cannot be written under the form (2.2)) is called entangled.

Characterizing whether a given multipartite state ρ is entangled or separable is a fundamental question. Indeed, if ρ is separable, it means that the correlations between its subsystems are only classical: ρ is a probabilistic mixture of states of the form $\rho_1 \otimes \cdots \otimes \rho_n$, that have no correlation between subsystems. On the contrary, if ρ is entangled, it means that it exhibits genuinely quantum correlations between its subsystems. We will comment more precisely on this later.

2.2.2 Entanglement witnesses

We explained before that being able to guarantee that a given multipartite state is entangled is a key problem in practice. One way of achieving this is described in the following easy but important statement.

Theorem 2.2.2. *A state ρ on $\mathcal{H}$ is entangled iff there exists $W \in B(\mathcal{H})$ such that $W^* = W$ satisfying*

1. *for all separable state σ on $\mathcal{H}$, $\mathrm{Tr}(W\sigma) \geqslant 0$;*
2. *$\mathrm{Tr}(W\rho) < 0$.*

Proof. We consider the space of Hermitian operators on $\mathcal{H}$ as a real Hilbert space, when equipped with the so-called Hilbert-Schmidt inner product $(X, Y) \mapsto \mathrm{Tr}(XY)$. In this space, the subset $\mathrm{S}(\mathcal{H})$ of separable states is compact and convex. Hence, by Hahn-Banach separation theorem, any Hermitian operator ρ on $\mathcal{H}$ that does not belong to $\mathrm{S}(\mathcal{H})$ (in particular any $\rho \in \mathrm{D}(\mathcal{H}) \setminus \mathrm{S}(\mathcal{H})$) can be separated from $\mathrm{S}(\mathcal{H})$ by a hyperplane, i.e. there exist a Hermitian operator W' on $\mathcal{H}$ and some constant $c \in \mathbb{R}$ such that $\mathrm{Tr}(W'\rho) < c$ while $\mathrm{Tr}(W'\sigma) \geqslant c$ for all $\sigma \in \mathrm{S}(\mathcal{H})$. We then just have to take $W = W' - c\mathrm{I}$ to prove the 'only if' part of the statement. While the 'if' part of the statement is clear. □

The Hermitian operator W appearing in Theorem 2.2.2 is often referred to as an entanglement witness for ρ, as it allows to discriminate ρ from the set of separable states. More generally, a Hermitian operator satisfying condition (1) is usually called block-positive. This terminology comes from the observation that W satisfying condition (1) is equivalent to W being such that, for all unit vectors $\varphi_1 \in \mathcal{H}_1, \ldots, \varphi_n \in \mathcal{H}_n$, $\langle\varphi_1 \otimes \cdots \otimes \varphi_n|W|\varphi_1 \otimes \cdots \otimes \varphi_n\rangle \geqslant 0$, i.e. for all $1 \leqslant k \leqslant n$, for all unit vectors $\varphi_1 \in \mathcal{H}_1, \ldots, \varphi_{k-1} \in \mathcal{H}_{k-1}, \varphi_{k+1} \in \mathcal{H}_{k+1}, \ldots, \varphi_n \in \mathcal{H}_n$, $\hat{W} = \langle\varphi_1 \otimes \cdots \otimes \varphi_{k-1} \otimes \varphi_{k+1} \otimes \cdots \otimes \varphi_n|W|\varphi_1 \otimes \cdots \otimes \varphi_{k-1} \otimes \varphi_{k+1} \otimes \cdots \otimes \varphi_n\rangle \in B(\mathcal{H}_k)$ is positive semi-definite.

2.3 Bipartite case

We focus on the case where $n = 2$, i.e. $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ is a bipartite Hilbert space.

2.3.1 Entanglement of bipartite pure states

We recall that a unit vector $\varphi \in \mathcal{H}$ can always be written in its Schmidt decomposition

$$\varphi = \sum_{i=1}^{r} \sqrt{\lambda_i} u_i \otimes v_i, \tag{2.3}$$

where $r = \mathrm{SR}(\varphi) \leqslant d = \min(d_1, d_2)$, $1 \geqslant \lambda_1 \geqslant \cdots \geqslant \lambda_r > 0$ are positive numbers summing up to 1 and $\{u_1, \ldots, u_r\}, \{v_1, \ldots, v_r\}$ are orthonormal families in $\mathcal{H}_1, \mathcal{H}_2$.

The corresponding pure state $\rho = |\varphi\rangle\langle\varphi|$ on $\mathcal{H}$ is thus separable iff the unit vector φ has Schmidt rank 1. On the contrary, ρ is called maximally entangled if the unit vector φ has Schmidt rank d and Schmidt coefficients all equal to $1/\sqrt{d}$, i.e.

$$\varphi = \sum_{i=1}^{d} \frac{1}{\sqrt{d}} u_i \otimes v_i.$$

Example 2.3.1. *On $\mathbb{C}^2 \otimes \mathbb{C}^2$, the maximally entangled state in the canonical orthonormal basis is usually called the Bell pair state or EPR (Einstein-Podolsky-Rosen) pair state. It corresponds to the unit vector*

$$\psi = \frac{1}{\sqrt{2}}(e_1 \otimes e_1 + e_2 \otimes e_2).$$

It is easy to see that, if the unit vector $\varphi \in \mathcal{H}$ has Schmidt decomposition (2.3), then the reduced states of $\rho = |\varphi\rangle\langle\varphi|$ on $\mathcal{H}_1$ and $\mathcal{H}_2$ are

$$\rho_1 = \mathrm{Tr}_{\mathcal{H}_2}(\rho) = \sum_{i=1}^{r} \lambda_i |u_i\rangle\langle u_i| \quad \text{and} \quad \rho_2 = \mathrm{Tr}_{\mathcal{H}_1}(\rho) = \sum_{i=1}^{r} \lambda_i |v_i\rangle\langle v_i|.$$

This is because

$$\rho = \sum_{i,j=1}^{r} \sqrt{\lambda_i}\sqrt{\lambda_j} |u_i\rangle\langle u_j| \otimes |v_i\rangle\langle v_j|.$$

So both ρ_1 and ρ_2 have the same rank r and the same spectrum $\{\lambda_1, \ldots, \lambda_r\}$, which is usually referred to as the entanglement spectrum of ρ. Conversely, given a state ρ_1 on some Hilbert space $\mathcal{H}_1$ that has rank r, there always exist a Hilbert space $\mathcal{H}_2$ with $\dim(\mathcal{H}_2) = r$ and a pure state $\rho = |\varphi\rangle\langle\varphi|$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ such that $\rho_1 = \mathrm{Tr}_{\mathcal{H}_2}(\rho)$. Indeed, writing $\rho_1 = \sum_{i=1}^{r} \lambda_i |u_i\rangle\langle u_i|$ in its spectral decomposition, we just have to set $\varphi = \sum_{i=1}^{r} \sqrt{\lambda_i} u_i \otimes v_i$, where $\{v_1, \ldots, v_r\}$ is an orthonormal basis of $\mathcal{H}_2$. Such pure state ρ is called a purification of ρ_1. It is not unique, as any unit vector of the form $(\mathrm{I} \otimes V)\varphi$, for V an isometry, would also provide a purification of ρ_1. So we see from this discussion that two states ρ_1 on $\mathcal{H}_1$ and ρ_2 on $\mathcal{H}_2$ can arise as the two reduced states of one common pure state $\rho = |\varphi\rangle\langle\varphi|$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ iff they have the same spectrum.

It is also interesting to note that the pure state ρ is separable iff its reduced states are pure, while ρ is maximally entangled iff its reduced state on the smallest subsystem is maximally mixed. This observation allows to interpret mixed states on a system of interest as emerging from pure entangled states on a larger system, composed of the system of interest coupled with an 'environment' or 'ancillary' system, which is traced out. Indeed, a mixed state ρ_1 on the single system $\mathcal{H}_1$ can be seen as the reduced state on $\mathcal{H}_1$ of a pure entangled state $\rho = |\varphi\rangle\langle\varphi|$ on the composed system $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Definition 2.3.2. *Given a pure state $\rho = |\varphi\rangle\langle\varphi|$ on $\mathcal{H}$, its entanglement entropy $E(\rho)$ is defined as the entropy of its reduced state ρ_1 on $\mathcal{H}_1$ (or equivalently of its reduced state ρ_2 on $\mathcal{H}_2$). Denoting by $\{\lambda_1, \ldots, \lambda_r\}$ the entanglement spectrum of ρ, as defined above, we have*

$$E(\rho) = S(\rho_1) = S(\rho_2) = -\sum_{i=1}^{r} \lambda_i \log(\lambda_i).$$

Proposition 2.3.3. *Let $d = \min(d_1, d_2)$. For any pure state $\rho = |\varphi\rangle\langle\varphi|$ on $\mathcal{H}$, we have*

$$0 \leqslant E(\rho) \leqslant \log(d),$$

with equality in the first inequality iff ρ is separable and in the second inequality iff ρ is maximally entangled.

Proof. It is immediate from the discussion above. Indeed, $E(\rho) = 0$ is equivalent to the reduced states of ρ being pure, i.e. to ρ being separable. And $E(\rho) = \log(d)$ is equivalent to the reduced state of ρ on the smallest subsystem being maximally mixed, i.e. to ρ being maximally entangled. □

The entanglement entropy quantifies the amount of entanglement present in a given bipartite pure state, and is thus called an entanglement measure. It has the property of being faithful, i.e. equal to 0 iff the bipartite pure state is separable.

2.3.2 Entanglement measures for bipartite mixed states

The entanglement entropy, introduced in Definition 2.3.2 above, allows to quantify how entangled a given bipartite pure state is. We would now like to be able to do the same with bipartite mixed states. But it is clear that simply looking at how mixed reduced states are is not enough anymore. Indeed, a separable mixed state, contrary to a separable pure state, can have very mixed reduced states. For instance, $\rho = \rho_1 \otimes \rho_2$ with $\rho_1 = I/d_1$ and $\rho_2 = I/d_2$ is a product state that has maximally mixed reduced states.

Definition 2.3.4. *Given a state ρ on $\mathcal{H}$, its entanglement of formation $E_F(\rho)$ is defined as*

$$E_F(\rho) = \min\left\{\sum_{i=1}^{r} \lambda_i E(|\varphi_i\rangle\langle\varphi_i|) : \rho = \sum_{i=1}^{r} \lambda_i |\varphi_i\rangle\langle\varphi_i|,\ \lambda_i \geqslant 0,\ \sum_{i=1}^{r} \lambda_i = 1,\ \varphi_i \in \mathcal{H},\ \|\varphi_i\| = 1\right\},$$

where $E(|\varphi_i\rangle\langle\varphi_i|)$ is the entanglement entropy of the pure state $|\varphi_i\rangle\langle\varphi_i|$ on $\mathcal{H}$, as defined in Definition 2.3.2.

Proposition 2.3.5. *Let $d = \min(d_1, d_2)$. For any state ρ on $\mathcal{H}$, we have*

$$0 \leqslant E_F(\rho) \leqslant \log(d),$$

with equality in the first inequality iff ρ is a separable state and in the second inequality iff ρ is a convex combination of maximally entangled states. What is more, E_F is convex, and if ρ is a pure state, then $E_F(\rho) = E(\rho)$.

Proof. From the definition of E_F, we see that $E_F(\rho) = 0$ is equivalent to the existence of positive numbers λ_i's summing up to 1 and unit vectors φ_i's such that $\rho = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$ and $\sum_i \lambda_i E(|\varphi_i\rangle\langle\varphi_i|) = 0$. The latter is satisfied iff $E(|\varphi_i\rangle\langle\varphi_i|) = 0$ for all i, i.e. $|\varphi_i\rangle\langle\varphi_i|$ separable for all i. Hence, $E_F(\rho) = 0$ is equivalent to the existence of a decomposition of ρ into a convex combination of separable pure states, i.e. to ρ being separable.

Similarly $E_F(\rho) = \log(d)$ is equivalent to the existence of positive numbers λ_i's summing up to 1 and unit vectors φ_i's such that $\rho = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$ and $\sum_i \lambda_i E(|\varphi_i\rangle\langle\varphi_i|) = \log(d)$. The latter is satisfied iff $E(|\varphi_i\rangle\langle\varphi_i|) = \log(d)$ for all i, i.e. $|\varphi_i\rangle\langle\varphi_i|$ maximally entangled for all i. Hence, $E_F(\rho) = \log(d)$ is equivalent to the existence of a decomposition of ρ into a convex combination of maximally entangled states.

Finally, let us show that E_F is convex. Define $\sigma = p\rho + (1-p)\rho'$, for $\rho, \rho' \in \mathrm{D}(\mathcal{H})$ and $0 \leqslant p \leqslant 1$. Let $\rho = \sum_i \mu_i |\chi_i\rangle\langle\chi_i|$ and $\rho' = \sum_i \mu'_i |\chi'_i\rangle\langle\chi'_i|$ be pure state decompositions of ρ and ρ' such that $E_F(\rho) = \sum_i \mu_i E(|\chi_i\rangle\langle\chi_i|)$ and $E_F(\rho') = \sum_i \mu'_i E(|\chi'_i\rangle\langle\chi'_i|)$. Then, $\sigma = p\sum_i \mu_i |\chi_i\rangle\langle\chi_i| + (1-p)\sum_i \mu'_i |\chi'_i\rangle\langle\chi'_i|$ is a pure state decomposition of σ. Hence,

$$\begin{aligned} E_F(\sigma) &= \min\left\{\sum_i \lambda_i E(|\varphi_i\rangle\langle\varphi_i|) : \sigma = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|\right\} \\ &\leqslant p\sum_i \mu_i E(|\chi_i\rangle\langle\chi_i|) + (1-p)\sum_i \mu'_i E(|\chi'_i\rangle\langle\chi'_i|) \\ &= pE_F(\rho) + (1-p)E_F(\rho'), \end{aligned}$$

which proves that E_F is convex, as wanted. □

The entanglement of formation is an entanglement measure that can be seen as the most natural generalization of the entanglement entropy to bipartite mixed states. Indeed, it is defined as the entanglement entropy on pure states, and as its so-called convex-roof extension on mixed states. Just as the entanglement entropy, it is a faithful entanglement measure. We mention that convex-roof extension is a common way of generalizing to mixed states definitions that are initially suited only for pure states.

Definition 2.3.6. *Given a state ρ on $\mathcal{H}$, its relative entropy of entanglement $E_R(\rho)$ is defined as*

$$E_R(\rho) = \min\left\{S(\rho\|\sigma) : \sigma \in \mathrm{S}(\mathcal{H})\right\},$$

where $S(\rho\|\sigma) = \mathrm{Tr}(\rho(\log(\rho) - \log(\sigma)))$ is the relative entropy between ρ and σ.

Proposition 2.3.7. *Let $d = \min(d_1, d_2)$. For any state ρ on $\mathcal{H}$, we have*

$$0 \leqslant E_R(\rho) \leqslant \log(d),$$

with equality in the first inequality iff ρ is a separable state and in the second inequality iff ρ is a maximally entangled state. What is more, E_R is convex, and if ρ is a pure state, then $E_R(\rho) = E(\rho)$.

Proof. We know that $S(\rho\|\sigma) = 0$ iff $\sigma = \rho$. Hence,

$$\min\left\{S(\rho\|\sigma) : \sigma \in \mathrm{S}(\mathcal{H})\right\} = 0 \iff \exists\ \sigma \in \mathrm{S}(\mathcal{H}) : S(\rho\|\sigma) = 0 \iff \rho \in \mathrm{S}(\mathcal{H}).$$

This proves that $E_R(\rho) = 0$ iff ρ is separable.

The fact that E_R is convex follows easily from the convexity of the set of separable states and the joint convexity of the relative entropy. Indeed, let $\rho, \rho' \in \mathrm{D}(\mathcal{H})$ and $\tau, \tau' \in \mathrm{S}(\mathcal{H})$ such that $E_R(\rho) = S(\rho\|\tau), E_R(\rho') = S(\rho'\|\tau')$. Then, for any $0 \leqslant p \leqslant 1$, $p\tau + (1-p)\tau' \in \mathrm{S}(\mathcal{H})$ by convexity of $\mathrm{S}(\mathcal{H})$, and thus

$$E_R(p\rho + (1-p)\rho') \leqslant S(p\rho + (1-p)\rho'\|p\tau + (1-p)\tau') \leqslant pS(\rho\|\tau) + (1-p)S(\rho'\|\tau') = pE_R(\rho) + (1-p)E_R(\rho'),$$

where the second inequality is by joint convexity of $S(\cdot\|\cdot)$, as recalled in Proposition 1.3.3.

We will next admit the following statement: if $\rho = |\varphi\rangle\langle\varphi|$ is a pure state, with the unit vector φ having Schmidt decomposition $\varphi = \sum_{i=1}^{d} \sqrt{\lambda_i} u_i \otimes v_i$, then the state $\tau \in \mathrm{S}(\mathcal{H})$ such that $E_R(\rho) = S(\rho\|\tau)$ is $\tau = \sum_{i=1}^{d} \lambda_i |u_i\rangle\langle u_i| \otimes |v_i\rangle\langle v_i|$. We thus have

$$E_R(\rho) = S(\rho\|\tau) = S(\tau) = S(\rho_1) = E(\rho),$$

and the latter quantity is always at most $\log(d)$, with equality iff ρ is maximally entangled.

Finally, since we have shown that E_R is convex and upper bounded by $\log(d)$ on pure states, it is clear that it is upper bounded by $\log(d)$ on all states, which can always be written as convex combinations of pure states. □

The relative entropy of entanglement is one example of a distance measure between a given state and its closest separable state. Another natural way of measuring such distance is through norms, e.g. one could define, for any $p \in \mathbb{N}$, the p-norm distance of a given $\rho \in \mathrm{D}(\mathcal{H})$ to $\mathrm{S}(\mathcal{H})$ as

$$d_p\left(\rho, \mathrm{S}(\mathcal{H})\right) = \min\left\{\|\rho - \sigma\|_p : \sigma \in \mathrm{S}(\mathcal{H})\right\},$$

which is equal to 0 iff $\rho \in \mathrm{S}(\mathcal{H})$. The most natural case to consider is that of the 1-norm distance to separable states. Indeed, the latter is related to the relative entropy of entanglement, through Pinsker inequality (see Proposition 1.3.1), by

$$\forall\ \rho \in \mathrm{D}(\mathcal{H}),\ E_R(\rho) \geqslant \frac{1}{2\log(2)} d_1\left(\rho, \mathrm{S}(\mathcal{H})\right).$$

One can also define the maximal fidelity of a given $\rho \in \mathrm{D}(\mathcal{H})$ to $\mathrm{S}(\mathcal{H})$ as

$$F\left(\rho, \mathrm{S}(\mathcal{H})\right) = \max\left\{F(\rho, \sigma) : \sigma \in \mathrm{S}(\mathcal{H})\right\},$$

which is equal to 1 iff $\rho \in \mathrm{S}(\mathcal{H})$. The 1-norm distance to separable states is also related to the maximal fidelity to separable states, through Fuchs / van de Graff inequalities (see Proposition 1.3.4), by

$$\forall\ \rho \in \mathrm{D}(\mathcal{H}),\ 1 - F\left(\rho, \mathrm{S}(\mathcal{H})\right) \leqslant \frac{1}{2} d_1\left(\rho, \mathrm{S}(\mathcal{H})\right) \leqslant \left(1 - F\left(\rho, \mathrm{S}(\mathcal{H})\right)^2\right)^{1/2}.$$

2.3.3 Entanglement criteria for bipartite states

Deciding whether a given bipartite mixed state is entangled or separable (i.e. deciding whether or not there exists a convex decomposition of a given bipartite density matrix into product density matrices) is in general a computationally hard problem. One solution in practice is to find sufficient conditions for entanglement that are easier to check than entanglement itself.

PPT

We recall that, given an orthonormal basis $\{e_1, \dots, e_{d_1}\}$ of $\mathcal{H}_1$, the transposition on $B(\mathcal{H}_1)$ with respect to this basis is the linear map $T : B(\mathcal{H}_1) \to B(\mathcal{H}_1)$ that acts as

$$T\left(\sum_{i,j=1}^{d_1} X_{ij}|e_i\rangle\langle e_j|\right) = \sum_{i,j=1}^{d_1} X_{ij}|e_j\rangle\langle e_i|.$$

We then define the partial transposition on $B(\mathcal{H})$ as the linear map $\Gamma = T \otimes \mathrm{id} : B(\mathcal{H}_1 \otimes \mathcal{H}_2) \to B(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Concretely, it acts as

$$\Gamma\left(\sum_{i_1,j_1=1}^{d_1}\sum_{i_2,j_2=1}^{d_2} X_{i_1 j_1 i_2 j_2}|e_{i_1}\rangle\langle e_{j_1}| \otimes |f_{i_2}\rangle\langle f_{j_2}|\right) = \sum_{i_1,j_1=1}^{d_1}\sum_{i_2,j_2=1}^{d_2} X_{i_1 j_1 i_2 j_2}|e_{j_1}\rangle\langle e_{i_1}| \otimes |f_{i_2}\rangle\langle f_{j_2}|.$$

The (partial) transposition depends on the choice of basis, but the spectrum of the (partial) transposition of a matrix does not. One crucial difference between transposition and partial transposition, though, is that transposition preserves the spectrum (i.e. for all $X \in B(\mathcal{H}_1)$, $\mathrm{spec}(T(X)) = \mathrm{spec}(X)$) while partial transposition generally does not (i.e. there exists $X \in B(\mathcal{H})$, $\mathrm{spec}(\Gamma(X)) \neq \mathrm{spec}(X)$).

Definition 2.3.8. *A state ρ on $\mathcal{H}$ is called positive under partial transposition (PPT) if $\Gamma(\rho) \geqslant 0$. We denote by $\mathrm{P}(\mathcal{H}) \subset \mathrm{D}(\mathcal{H})$ the set of PPT states on $\mathcal{H}$.*

Remark 2.3.9. *It is easy to see that the partial transposition preserves the trace (i.e. for all $X \in B(\mathcal{H})$, $\mathrm{Tr}(\Gamma(X)) = \mathrm{Tr}(X)$). Therefore, a state ρ on $\mathcal{H}$ is PPT iff its partial transpose $\Gamma(\rho)$ is a state on $\mathcal{H}$. Hence an alternative way of describing the set of PPT states is as the intersection between the set of all states and its image under partial transposition, i.e.*

$$\mathrm{P}(\mathcal{H}) = \mathrm{D}(\mathcal{H}) \cap \Gamma(\mathrm{D}(\mathcal{H})).$$

Theorem 2.3.10. *Let ρ be a state on $\mathcal{H}$. If ρ is separable, then it is PPT. In other words, we have the inclusion*

$$\mathrm{S}(\mathcal{H}) \subset \mathrm{P}(\mathcal{H}).$$

Proof. If ρ is a product state, i.e. $\rho = \rho_1 \otimes \rho_2$ with $\rho_1 \in \mathrm{D}(\mathcal{H}_1)$ and $\rho_2 \in \mathrm{D}(\mathcal{H}_2)$, we have $\Gamma(\rho) = T(\rho_1) \otimes \rho_2$. Now, since transposition preserves the spectrum, $T(\rho_1) \geqslant 0$, and thus $\Gamma(\rho) \geqslant 0$. More generally, if $\rho = \sum_{i=1}^r \lambda_i \rho_1^{(i)} \otimes \rho_2^{(i)}$ is a convex combination of product states, then by linearity of the partial transposition, $\Gamma(\rho) = \sum_{i=1}^r \lambda_i T(\rho_1^{(i)}) \otimes \rho_2^{(i)}$ is a convex combination of positive semi-definite operators, and is therefore positive semi-definite. □

Theorem 2.3.10 is usually used in its contrapositive form, as a so-called entanglement criterion, namely: if ρ is not PPT, then it is entangled. The interest is that checking whether or not a given state is PPT is in general much easier than checking whether or not it is separable. So entanglement can potentially be guaranteed by a condition that is much simpler to check.

Theorem 2.3.11. *Let $\rho = |\varphi\rangle\langle\varphi|$ be a pure state on $\mathcal{H}$. ρ is separable iff it is PPT.*

Proof. We have already proven in Theorem 2.3.10 that ρ being separable always implies ρ being PPT (even for mixed states).

To prove the other implication, let us write $\varphi = \sum_{i=1}^r \sqrt{\lambda_i} u_i \otimes v_i$ in its Schmidt decomposition, as defined by equation (2.3). We then have

$$\rho = \sum_{i,j=1}^r \sqrt{\lambda_i}\sqrt{\lambda_j}|u_i\rangle\langle u_j| \otimes |v_i\rangle\langle v_j|$$

Completing the orthonormal family $\{u_1, \dots, u_r\}$ into an orthonormal basis $\{u_1, \dots, u_{d_1}\}$, the partial transposition of ρ with respect to this basis is

$$\Gamma(\rho) = \sum_{i,j=1}^{r} \sqrt{\lambda_i}\sqrt{\lambda_j} |u_j\rangle\langle u_i| \otimes |v_i\rangle\langle v_j|.$$

Suppose that $r \geqslant 2$, so that $\lambda_1, \lambda_2 > 0$. Then, it is easy to check that the restriction of $\Gamma(\rho)$ to $\mathrm{span}\{u_1 \otimes v_2, u_2 \otimes v_1\}$ is not positive semi-definite. We have thus shown that, if $r > 1$, i.e. if ρ is not separable, then ρ is not PPT. Equivalently, this means that ρ being PPT implies ρ being separable. □

Remark 2.3.12. *More precisely, we can show that, given a pure state $\rho = |\varphi\rangle\langle\varphi|$ on $\mathcal{H}$, with $\varphi \in \mathcal{H}$ having Schmidt coefficients $\{\sqrt{\lambda_i} : 1 \leqslant i \leqslant r\}$, the spectrum of $\Gamma(\rho)$ is $\{\lambda_i : 1 \leqslant i \leqslant r\} \cup \{\pm\sqrt{\lambda_i\lambda_j} : 1 \leqslant i < j \leqslant r\}$.*

Theorem 2.3.13. *On $\mathbb{C}^2 \otimes \mathbb{C}^2$, $\mathbb{C}^2 \otimes \mathbb{C}^3$ and $\mathbb{C}^3 \otimes \mathbb{C}^2$, being separable is equivalent to being PPT. But for all other dimensions d_1, d_2, there exist PPT states on $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ which are not separable. In other words, for all $(d_1, d_2) \notin \{(2,2), (2,3), (3,2)\}$, we have the strict inclusion*

$$\mathrm{S}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}) \subsetneq \mathrm{P}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}).$$

Realignment criterion

Given orthonormal bases $\{e_1, \dots, e_{d_1}\}$ and $\{f_1, \dots, f_{d_2}\}$ of $\mathcal{H}_1$ and $\mathcal{H}_2$, the realignment with respect to theses bases is the linear map $R : B(\mathcal{H}_1 \otimes \mathcal{H}_2) \to B(\mathcal{H}_2 \otimes \mathcal{H}_2, \mathcal{H}_1 \otimes \mathcal{H}_1)$ that acts as

$$R\left(\sum_{i_1,j_1=1}^{d_1}\sum_{i_2,j_2=1}^{d_2} X_{i_1 j_1 i_2 j_2} |e_{i_1}\rangle\langle e_{j_1}| \otimes |f_{i_2}\rangle\langle f_{j_2}|\right) = \sum_{i_1,j_1=1}^{d_1}\sum_{i_2,j_2=1}^{d_2} X_{i_1 j_1 i_2 j_2} |e_{i_1}\rangle\langle f_{i_2}| \otimes |e_{j_1}\rangle\langle f_{j_2}|.$$

Theorem 2.3.14. *Let ρ be a state on $\mathcal{H}$. If ρ is separable, then*

$$\|R(\rho)\|_1 \leqslant 1.$$

Proof. If ρ is a product state, i.e. $\rho = \rho_1 \otimes \rho_2$ with $\rho_1 \in \mathrm{D}(\mathcal{H}_1)$ and $\rho_2 \in \mathrm{D}(\mathcal{H}_2)$, we have $R(\rho) = |\varphi_1\rangle\langle\varphi_2|$, where $\varphi_1 = \varphi_{\rho_1} \in \mathcal{H}_1 \otimes \mathcal{H}_1$, resp. $\varphi_2 = \varphi_{\rho_2} \in \mathcal{H}_2 \otimes \mathcal{H}_2$, is the vector version of ρ_1, resp. ρ_2 (as defined in Section 2.1.3). Hence,

$$\|R(\rho)\|_1 = \||\varphi_1\rangle\langle\varphi_2|\|_1 = \|\varphi_1\|\|\varphi_1\| = \|\rho_1\|_2 \|\rho_2\|_2 \leqslant \|\rho_1\|_1 \|\rho_2\|_1 = 1,$$

and this concludes the case where ρ is a product state. More generally, if $\rho = \sum_{i=1}^{r} \lambda_i \rho_1^{(i)} \otimes \rho_2^{(i)}$ is a convex combination of product states, then by linearity of the realignment and convexity of the 1-norm we have

$$\|R(\rho)\|_1 = \left\|\sum_{i=1}^{r} \lambda_i R\left(\rho_1^{(i)} \otimes \rho_2^{(i)}\right)\right\|_1 \leqslant \sum_{i=1}^{r} \lambda_i \left\|R\left(\rho_1^{(i)} \otimes \rho_2^{(i)}\right)\right\|_1 \leqslant 1,$$

and this concludes the proof. □

Theorem 2.3.14 is usually used in its contrapositive form, as a so-called entanglement criterion, namely: if ρ is such that $\|R(\rho)\|_1 > 1$, then it is entangled. The interest is that computing the 1-norm of the realignment of a given state is in general much easier than checking whether or not it is separable. So, with the realignment criterion, entanglement can potentially be guaranteed by a condition that is much simpler to check (just as with the PPT criterion).

Theorem 2.3.15. *Let $\rho = |\varphi\rangle\langle\varphi|$ be a pure state on $\mathcal{H}$. ρ is separable iff $\|R(\rho)\|_1 \leqslant 1$.*

Proof. Let us write $\varphi = \sum_{i=1}^{r} \sqrt{\lambda_i} u_i \otimes v_i$ in its Schmidt decomposition, as defined by equation (2.3). We then have

$$\rho = \sum_{i,j=1}^{r} \sqrt{\lambda_i}\sqrt{\lambda_j} |u_i\rangle\langle u_j| \otimes |v_i\rangle\langle v_j|$$

Completing the orthonormal families $\{u_1, \dots, u_r\}$ and $\{v_1, \dots, v_r\}$ into orthonormal bases $\{u_1, \dots, u_{d_1}\}$ and $\{v_1, \dots, v_{d_2}\}$, the realignment of ρ with respect to these bases is

$$R(\rho) = \sum_{i,j=1}^{r} \sqrt{\lambda_i}\sqrt{\lambda_j} |u_i\rangle\langle v_i| \otimes |u_j\rangle\langle v_j| = M_\varphi \otimes M_\varphi,$$

where $M_\varphi \in B(\mathcal{H}_2, \mathcal{H}_1)$ is the matrix version of φ (as defined in Section 2.1.3). Hence,

$$\|R(\rho)\|_1 = \|M_\varphi \otimes M_\varphi\|_1 = \|M_\varphi\|_1^2 = \left(\sum_{i=1}^{r} \sqrt{\lambda_i}\right)^2.$$

And the latter quantity is at most (in fact equal to) 1 iff $r = 1$. We have thus shown that $r = 1$, i.e. ρ being separable, is equivalent to ρ satisfying $\|R(\rho)\|_1 \leqslant 1$. □

2.3.4 Important examples of bipartite states

We consider the case of a balanced bipartite system, i.e. $\mathcal{H}_1 \equiv \mathcal{H}_2 \equiv \mathbb{C}^d$.

Isotropic states

Isotropic states on $\mathbb{C}^d \otimes \mathbb{C}^d$ are states which are convex (actually affine) combinations of the maximally mixed state I/d^2 and the maximally entangled state $|\psi\rangle\langle\psi|$. They have the form

$$\rho_\alpha = \alpha|\psi\rangle\langle\psi| + (1-\alpha)\frac{\mathrm{I}}{d^2}, \tag{2.4}$$

for $-1/(d^2-1) \leqslant \alpha \leqslant 1$.

Proposition 2.3.16. *For $-1/(d^2-1) \leqslant \alpha \leqslant 1$, let ρ_α be the corresponding isotropic state, as defined by equation (2.4). The following statements are equivalent:*

1. *ρ_α is separable;*
2. *ρ_α is PPT;*
3. *$-1/(d^2-1) \leqslant \alpha \leqslant 1/(d+1)$.*

Werner states

The symmetric and anti-symmetric subspaces of $\mathbb{C}^d \otimes \mathbb{C}^d$ are defined as the $+1$ and -1 eigenspaces of the so-called flip operator F on $\mathbb{C}^d \otimes \mathbb{C}^d$, whose action on product vectors is described by $F(\varphi_1 \otimes \varphi_2) = \varphi_2 \otimes \varphi_1$ and extended by linearity. We thus have

$$S_{\mathbb{C}^d \otimes \mathbb{C}^d} = \left\{\varphi \in \mathbb{C}^d \otimes \mathbb{C}^d : F\varphi = \varphi\right\} \quad \text{and} \quad A_{\mathbb{C}^d \otimes \mathbb{C}^d} = \left\{\varphi \in \mathbb{C}^d \otimes \mathbb{C}^d : F\varphi = -\varphi\right\}.$$

We denote by Π_S and Π_A the orthogonal projectors onto $S_{\mathbb{C}^d \otimes \mathbb{C}^d}$ and $A_{\mathbb{C}^d \otimes \mathbb{C}^d}$. Those can be written as

$$\Pi_S = \frac{1}{2}(\mathrm{I} + F) \quad \text{and} \quad \Pi_A = \frac{1}{2}(\mathrm{I} - F).$$

The symmetric and anti-symmetric states on $\mathbb{C}^d \otimes \mathbb{C}^d$ are then defined as $\pi_S = \Pi_S / \operatorname{Tr}(\Pi_S)$ and $\pi_A = \Pi_A / \operatorname{Tr}(\Pi_A)$. Since $\dim(S_{\mathbb{C}^d \otimes \mathbb{C}^d}) = d(d+1)/2$ and $\dim(A_{\mathbb{C}^d \otimes \mathbb{C}^d}) = d(d-1)/2$, we have

$$\pi_S = \frac{2}{d(d+1)}(\mathrm{I} + F) \quad \text{and} \quad \pi_A = \frac{2}{d(d-1)}(\mathrm{I} - F).$$

Werner states on $\mathbb{C}^d \otimes \mathbb{C}^d$ are states which are convex combinations of the symmetric and anti-symmetric states π_S and π_A. They have the form

$$\sigma_\lambda = \lambda\pi_S + (1-\lambda)\pi_A, \tag{2.5}$$

for $0 \leqslant \lambda \leqslant 1$.

Proposition 2.3.17. *For $0 \leqslant \lambda \leqslant 1$, let σ_λ be the corresponding Werner state, as defined by equation (2.5). The following statements are equivalent:*

1. *σ_λ is separable;*
2. *σ_λ is PPT;*
3. *$1/2 \leqslant \lambda \leqslant 1$.*

2.4 Observables and measurements on tensor product Hilbert spaces

Let $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ be a multipartite Hilbert space.

2.4.1 Local and separable observables

Definition 2.4.1. *A Hermitian operator X on $\mathcal{H}$ is called local if it can be written as a tensor product of Hermitian operators on $\mathcal{H}_1, \ldots, \mathcal{H}_n$, i.e. if there exist Hermitian operators $X_1, \ldots, X_n$ on $\mathcal{H}_1, \ldots, \mathcal{H}_n$ such that*

$$X = X_1 \otimes \cdots \otimes X_n.$$

Observe that a local observable $X = X_1 \otimes \cdots \otimes X_n$ can be written as

$$X = (X_1 \otimes \mathrm{I}_2 \otimes \cdots \otimes \mathrm{I}_n)(\mathrm{I}_1 \otimes X_2 \otimes \mathrm{I}_3 \otimes \cdots \mathrm{I}_n) \cdots (\mathrm{I}_1 \otimes \cdots \otimes \mathrm{I}_{n-2} \otimes X_{n-1} \otimes \mathrm{I}_n)(\mathrm{I}_1 \otimes \cdots \otimes \mathrm{I}_{n-1} \otimes X_n),$$

where the product can actually be taken in any order since the involved operators commute. So performing the observable X on a composite system $\mathcal{H}$ can be seen as each local observer k performing the observable X_k on its sub-system $\mathcal{H}_k$. This is where the denomination 'local' comes from.

For observables that are positive semi-definite, one can straightforwardly define the subset of separable observables, which are nothing else than non-negative linear combinations of local positive semi-definite observables.

Definition 2.4.2. *A Hermitian positive semi-definite operator X on $\mathcal{H}$ is called separable if it can be written as a non-negative linear combination of local Hermitian positive semi-definite operators on $\mathcal{H}$, i.e. if there exist Hermitian positive semi-definite operators $X_1^{(i)}, \ldots, X_n^{(i)}$ on $\mathcal{H}_1, \ldots, \mathcal{H}_n$, $1 \leqslant i \leqslant r$, non-negative numbers $\mu_1, \ldots, \mu_r$ such that*

$$X = \sum_{i=1}^{r} \mu_i X_1^{(i)} \otimes \cdots \otimes X_n^{(i)}.$$

An equivalent way of characterizing that a Hermitian positive semi-definite operator on $\mathcal{H}$ is separable is by saying that it belongs to the cone generated by the set of separable states on $\mathcal{H}$. That is, given $X \in B(\mathcal{H})$ such that $X^* = X$ and $X \geqslant 0$, it is called separable if there exist $\mu \geqslant 0$ and $\rho \in \mathrm{S}(\mathcal{H})$ such that $X = \mu\rho$. In fact, a similar observation can be made for the set of all Hermitian positive semi-definite operators on $\mathcal{H}$: they can be seen as belonging to the cone generated by the set of all states on $\mathcal{H}$. That is, given $X \in B(\mathcal{H})$ such that $X^* = X$ and $X \geqslant 0$, there exist $\mu \geqslant 0$ and $\rho \in \mathrm{D}(\mathcal{H})$ such that $X = \mu\rho$.

2.4.2 Local and separable measurements

Definition 2.4.3. *A measurement $\mathrm{M} = (M_i)_{i \in I}$ on $\mathcal{H}$ is called local if it can be written as a tensor product of measurements on $\mathcal{H}_1, \ldots, \mathcal{H}_n$, i.e. if there exist $I_1, \ldots, I_n$ such that $I = I_1 \times \cdots \times I_n$ and measurements $\mathrm{M}^{(1)} = (M_{i_1}^{(1)})_{i_1 \in I_1}, \ldots, \mathrm{M}^{(n)} = (M_{i_n}^{(n)})_{i_n \in I_n}$ on $\mathcal{H}_1, \ldots, \mathcal{H}_n$ such that*

$$\forall\ i_1 \in I_1, \ldots, i_n \in I_n,\ M_i = M_{i_1 \cdots i_n} = M_{i_1}^{(1)} \otimes \cdots \otimes M_{i_n}^{(n)}.$$

Just as for local observables, local measurements on composite systems can be seen as a sequence (in any order) of measurements performed by local observers on their sub-system. If the system $\mathcal{H}$ is in a global state ρ, then the probability that observer 1 obtains outcome i_1, observer 2 obtains outcome i_2, etc, is given by

$$\mathbb{P}_\rho(i_1, \ldots, i_n) = \mathrm{Tr}\left(\left(M_{i_1}^{(1)} \otimes \cdots \otimes M_{i_n}^{(n)}\right)\rho\right).$$

And, denoting by ρ_k the reduced state of ρ on sub-system $\mathcal{H}_k$, the probability that observer k obtains outcome i_k (whatever the outcomes of the other observers) is given by

$$\begin{aligned}
\mathbb{P}_\rho(i_k) &= \sum_{i_1 \in I_1, \ldots, i_{k-1} \in I_{k-1}, i_{k+1} \in I_{k+1}, \ldots, i_n \in I_n} \mathbb{P}_\rho(i_1, \ldots, i_{k-1}, i_k, i_{k+1}, \ldots, i_n) \\
&= \sum_{i_1 \in I_1, \ldots, i_{k-1} \in I_{k-1}, i_{k+1} \in I_{k+1}, \ldots, i_n \in I_n} \mathrm{Tr}\left(\left(M_{i_1}^{(1)} \otimes \cdots \otimes M_{i_{k-1}}^{(k-1)} \otimes M_{i_k}^{(k)} \otimes M_{i_{k+1}}^{(k+1)} \otimes \cdots \otimes M_{i_n}^{(n)}\right)\rho\right) \\
&= \mathrm{Tr}\left(\left(\mathrm{I}_1 \otimes \cdots \otimes \mathrm{I}_{k-1} \otimes M_{i_k}^{(k)} \otimes \mathrm{I}_{k+1} \otimes \cdots \otimes \mathrm{I}_n\right)\rho\right) \\
&= \mathrm{Tr}\left(M_{i_k}^{(k)} \rho_k\right) \\
&= \mathbb{P}_{\rho_k}(i_k),
\end{aligned}$$

where the third equality is because $\sum_{i_l \in I_l} M_{i_l}^{(l)} = \mathrm{I}_l$ for each $1 \leqslant l \leqslant n$.

Note that, if $\rho = \rho_1 \otimes \cdots \otimes \rho_n$ for some states $\rho_1, \ldots, \rho_n$ on $\mathcal{H}_1, \ldots, \mathcal{H}_n$, then

$$\forall\ i_1 \in I_1, \ldots, i_n \in I_n,\ \mathrm{Tr}\left(\left(M_{i_1}^{(1)} \otimes \cdots \otimes M_{i_n}^{(n)}\right)\rho\right) = \mathrm{Tr}\left(M_{i_1}^{(1)}\rho_1\right) \times \cdots \times \mathrm{Tr}\left(M_{i_n}^{(n)}\rho_n\right),$$

which means that

$$\forall\ i_1 \in I_1, \ldots, i_n \in I_n,\ \mathbb{P}_\rho(i_1, \ldots, i_n) = \mathbb{P}_{\rho_1}(i_1) \times \cdots \times \mathbb{P}_{\rho_n}(i_n).$$

This shows that, if the system is in a product state, then the measurement outcomes of the n observers are independent from one another.

Remark 2.4.4. *We see from the above discussion that the outcomes of local measurements are correlated as soon as the state ρ is not product, even if it is separable. So the probability distribution of outcomes $\mathbb{P}_\rho$ reflects both the classical and the quantum correlations present in ρ.*

Example 2.4.5. *Let us look at an example in the simplest case where $n = 2$ and $\mathcal{H}_1 \equiv \mathcal{H}_2 \equiv \mathbb{C}^2$. Suppose that both local observers perform the same measurement $(|e_1\rangle\langle e_1|, |e_2\rangle\langle e_2|)$ on a global system in the Bell pair state $\rho = |\psi\rangle\langle\psi|$, as introduced in Example 2.3.1, i.e.*

$$\rho = \frac{1}{2}\left(|e_1\rangle\langle e_1| \otimes |e_1\rangle\langle e_1| + |e_1\rangle\langle e_2| \otimes |e_1\rangle\langle e_2| + |e_2\rangle\langle e_1| \otimes |e_2\rangle\langle e_1| + |e_2\rangle\langle e_2| \otimes |e_2\rangle\langle e_2|\right).$$

It is easy to check that observer 1 has equal probabilities $1/2$ of obtaining outcomes 1 and 2. And if they obtain outcome 1, then the post-measurement state is $|e_1\rangle\langle e_1| \otimes |e_1\rangle\langle e_1|$, so observer 2 then obtains outcome 1 with probability 1. While conversely, if they obtain outcome 2, then the post-measurement state is $|e_2\rangle\langle e_2| \otimes |e_2\rangle\langle e_2|$, so observer 2 then obtains outcome 2 with probability 1. Another way of putting it is to say that it cannot be that observers 1 and 2 obtain different outcomes: they can obtain either both outcome 1 or both outcome 2, with equal probabilities $1/2$. This illustrates that on an entangled state, local measurement outcomes might be (highly) correlated.

A slight generalization of local measurements are so-called LOCC measurements, where LOCC stands for local operations and classical communication. The setting is the following: local observer k_1 performs a measurement on their sub-system and communicates the outcome to the others, based on this outcome local observer k_2 decides which measurement to perform on their sub-system and communicates the outcome to the others, etc. Giving a mathematical description of the full set of LOCC measurements on multipartite systems, without any limitation on the number of communication rounds, is quite cumbersome. We thus limit ourselves to formally defining the set of so-called 1-way LOCC measurements on bipartite systems. The latter setting is when there is only one round of communication from observer 1 to observer 2, i.e. observer 1 performs a local measurement, communicates the outcome to observer 2, who decides accordingly which measurement to perform.

Definition 2.4.6. *A measurement $\mathrm{M} = (M_i)_{i \in I}$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is called 1-way LOCC if there exist I_1, I_2 such that $I = I_1 \times I_2$ and measurements $\mathrm{M}^{(1)} = (M_{i_1}^{(1)})_{i_1 \in I_1}$ on $\mathcal{H}_1$, $\mathrm{M}^{(2,i_1)} = (M_{i_2}^{(2,i_1)})_{i_2 \in I_2}$ on $\mathcal{H}_2$, $i_1 \in I_1$, such that*

$$\forall\ i_1 \in I_1, i_2 \in I_2,\ M_i = M_{i_1 i_2} = M_{i_1}^{(1)} \otimes M_{i_2}^{(2,i_1)}.$$

LOCC measurements are considered the most general kind of measurements that local observers can implement. Since this set of measurements is mathematically hard to describe, we also introduce the larger set of separable measurements, whose definition is straightforward. The interest in practice is the following: if one can show that a given task cannot be achieved with separable measurements, then it necessarily cannot be achieved with LOCC measurements neither.

Definition 2.4.7. *A measurement $\mathrm{M} = (M_i)_{i \in I}$ on $\mathcal{H}$ is called separable if, for each $i \in I$, the Hermitian positive semi-definite operator M_i is separable.*

2.5 Entropy and correlations in multipartite systems

In this section we slightly change our notation. We denote bipartite systems as $\mathcal{H}_A \otimes \mathcal{H}_B$ instead of $\mathcal{H}_1 \otimes \mathcal{H}_2$ and tripartite systems as $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ instead of $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$. The reason is that we are going to introduce information theoretic notions, in a quantum setting, and we want to stick to the notation that is commonly used in classical information theory.

2.5.1 Entropic inequalities

Lemma 2.5.1. *The relative entropy satisfies the following properties:*

1. *for all state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$, $S(\rho_{AB}\|\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB})$;*
2. *for all states ρ_A, σ_A on $\mathcal{H}_A$ and ρ_B, σ_B on $\mathcal{H}_B$, $S(\rho_A \otimes \rho_B\|\sigma_A \otimes \sigma_B) = S(\rho_A\|\sigma_A) + S(\rho_B\|\sigma_B)$.*

Corollary 2.5.2. *The von Neumann entropy is subadditive. This means that, for any state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$,*

$$S(\rho_{AB}) \leqslant S(\rho_A) + S(\rho_B).$$

And there is equality in the above inequality iff ρ_{AB} is a product state, i.e. $\rho_{AB} = \rho_A \otimes \rho_B$.

Proof. By Lemma 2.5.1, we see that $S(\rho_A) + S(\rho_B) - S(\rho_{AB}) = S(\rho_{AB}\|\rho_A \otimes \rho_B)$. So by non-negativity of the relative entropy, $S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \geqslant 0$. □

Definition 2.5.3. *Given a state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$, its mutual information is defined as*

$$I(A:B)_\rho = S(\rho_A) + S(\rho_B) - S(\rho_{AB}).$$

Proposition 2.5.4. *Let $d = \min(d_A, d_B)$. For any state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$, we have*

$$0 \leqslant I(A:B)_\rho \leqslant 2\log(d),$$

with equality in the first inequality iff ρ_{AB} is a product state and in the second inequality iff ρ_{AB} is a maximally entangled state.

Proof. Let us start with the lower bound. By Lemma 2.5.1, we see that we can rewrite the mutual information as $I(A:B)_\rho = S(\rho_{AB}\|\rho_A \otimes \rho_B)$. So by non-negativity and faithfulness of the relative entropy, $I(A:B)_\rho \geqslant 0$, with equality iff $\rho_{AB} = \rho_A \otimes \rho_B$.

Let us now move on to the upper bound. If ρ_A, ρ_B are reduced states of a joint state ρ_{AB}, then $S(\rho_A), S(\rho_B) \leqslant \log(d)$. Hence $I(A:B)_\rho = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \leqslant S(\rho_A) + S(\rho_B) \leqslant 2\log(d)$. And there is equality in the second inequality iff ρ_{AB} is maximally entangled (which guarantees that there is also equality in the first inequality). □

We see from the above proof that the mutual information of ρ_{AB} can be equivalently expressed as

$$I(A:B)_\rho = S(\rho_{AB}\|\rho_A \otimes \rho_B).$$

It can thus be seen as a distance measure between ρ_{AB} and $\rho_A \otimes \rho_B$, i.e. as quantifying the amount of correlations, both classical and quantum, in ρ_{AB}. Indeed, contrary to an entanglement measure, that is equal to 0 as soon as ρ_{AB} is separable (i.e. has no quantum correlations), the mutual information is not equal to 0 on separable states that are not product (i.e. that have classical correlations).

Lemma 2.5.5. *The relative entropy is invariant under unitary conjugations and non-increasing under the action of quantum channels. This means that for any states ρ, σ on $\mathcal{H}$, we have, for any unitary $U \in B(\mathcal{H})$,*

$$S\left(U\rho U^* \| U\sigma U^*\right) = S\left(\rho\|\sigma\right),$$

while for any completely positive and trace-preserving linear map $\Phi : B(\mathcal{H}) \to B(\mathcal{K})$,

$$S\left(\Phi(\rho)\|\Phi(\sigma)\right) \leqslant S\left(\rho\|\sigma\right).$$

The physical interpretation of Lemma 2.5.5 is that irreversible transformations make quantum states less distinguishable from one another. A particular case of interest is when $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, $\mathcal{K} = \mathcal{H}_A$ and $\Phi = \mathrm{Tr}_{\mathcal{H}_B}$ is the partial trace over $\mathcal{H}_B$. We then get the following inequality, relating the relative entropy between two bipartite states and the relative entropy between their reduced states: for all states ρ_{AB}, σ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$,

$$S(\rho_A\|\sigma_A) \leqslant S(\rho_{AB}\|\sigma_{AB}).$$

Corollary 2.5.6. *The von Neumann entropy is strongly subadditive. This means that, for any state ρ_{ABC} on $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$,*

$$S(\rho_{ABC}) + S(\rho_C) \leqslant S(\rho_{AC}) + S(\rho_{BC}).$$

And there is equality in the above inequality iff ρ_{ABC} is a Markov state, i.e. there exist a direct sum decomposition of $\mathcal{H}_C$ as $\mathcal{H}_C = \bigotimes_{i=1}^r \mathcal{H}_{C_i^A} \otimes \mathcal{H}_{C_i^B}$ and states $\rho_{AC_i^A}, \rho_{BC_i^B}$ on $\mathcal{H}_{C_i^A}, \mathcal{H}_{C_i^B}$ such that ρ_{ABC} can be written as a convex combination $\rho_{ABC} = \sum_{i=1}^r p_i \rho_{AC_i^A} \otimes \rho_{BC_i^B}$.

2.5.2 Correlations vs mutual information

Definition 2.5.7. *Let ρ_{AB} be a state on $\mathcal{H}_A \otimes \mathcal{H}_B$ and X_A, Y_B be observables on $\mathcal{H}_A, \mathcal{H}_B$. We define the correlation function*

$$\gamma_\rho(X_A, Y_B) = \left|\mathrm{Tr}\left((X_A \otimes Y_B)\,\rho_{AB}\right) - \mathrm{Tr}\left(X_A \rho_A\right)\mathrm{Tr}\left(Y_B \rho_B\right)\right| = \left|\langle X_A \otimes Y_B\rangle_{\rho_{AB}} - \langle X_A\rangle_{\rho_A}\langle Y_B\rangle_{\rho_B}\right|.$$

The quantity $\gamma_\rho(X_A, Y_B)$ is nothing else than the difference (in absolute value) between the expectation value of the observable $X_A \otimes Y_B$, when computed on the state ρ_{AB}, and the product of the expectation values of the observables X_A and Y_B, when computed on the states ρ_A and ρ_B respectively. So it measures the amount of correlations that are created between the expectation values of the observables X_A and Y_B when they are computed jointly on the bipartite state ρ_{AB}, compared to when they are computed separately on ρ_A and ρ_B. If ρ_{AB} is a product state, we clearly have $\gamma_\rho(X_A, Y_B) = 0$, i.e. no such correlation is created. More generally, the following result upper bounds $\gamma_\rho(X_A, Y_B)$ in terms of the mutual information of ρ_{AB}, i.e. in terms of its distance to being product.

Theorem 2.5.8. *Let ρ_{AB} be a state on $\mathcal{H}_A \otimes \mathcal{H}_B$. For any observables X_A, Y_B on $\mathcal{H}_A, \mathcal{H}_B$, satisfying $\|X_A\|_\infty, \|Y_B\|_\infty \leqslant 1$, we have*

$$\gamma_\rho(X_A, Y_B) \leqslant \sqrt{2\log(2)\, I(A:B)_\rho}.$$

Proof. Note that we can re-write $\gamma_\rho(X_A, Y_B)$ as

$$\gamma_\rho(X_A, Y_B) = \left|\mathrm{Tr}\left(\left(\rho_{AB} - \rho_A \otimes \rho_B\right) X_A \otimes Y_B\right)\right|.$$

And since $\|X_A \otimes Y_B\|_\infty \leqslant 1$, we have by duality between $\|\cdot\|_1$ and $\|\cdot\|_\infty$,

$$\left|\mathrm{Tr}\left(\left(\rho_{AB} - \rho_A \otimes \rho_B\right) X_A \otimes Y_B\right)\right| \leqslant \left\|\rho_{AB} - \rho_A \otimes \rho_B\right\|_1.$$

Now, we know by Pinsker inequality that

$$\left\|\rho_{AB} - \rho_A \otimes \rho_B\right\|_1 \leqslant \sqrt{2\log(2)\, S\left(\rho_{AB} \| \rho_A \otimes \rho_B\right)}.$$

So we just have to recall that $S(\rho_{AB}\|\rho_A \otimes \rho_B) = I(A:B)_\rho$ to conclude. □

Chapter 3

Quantum channels

Main references for this chapter:

- [1] Chapter 2 Section 2.3.
- [4] Chapter 2, Chapter 3 and Chapter 6.
- [5] Chapter 1 Section 1.7.

3.1 Definitions and first properties

A transformation of a quantum system $\mathcal{H}_1$ into another quantum system $\mathcal{H}_2$ is described by a linear map Φ from $B(\mathcal{H}_1)$ to $B(\mathcal{H}_2)$, i.e. an element of $B(B(\mathcal{H}_1), B(\mathcal{H}_2))$. In order for the linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ to be a valid quantum transformation, we need it to have the property that it maps quantum states on $\mathcal{H}_1$ to quantum states on $\mathcal{H}_2$, i.e. that $\Phi(\mathrm{D}(\mathcal{H}_1)) \subset \mathrm{D}(\mathcal{H}_2)$. For the latter to hold, we have to impose that Φ is

- Hermiticity-preserving and positivity-preserving (or positive in brief), i.e. for all $X \in B(\mathcal{H}_1)$ such that $X^* = X$ and $X \geqslant 0$, $\Phi(X) \in B(\mathcal{H}_2)$ is such that $\Phi(X)^* = \Phi(X)$ and $\Phi(X) \geqslant 0$;
- trace-preserving, i.e. for all $X \in B(\mathcal{H}_1)$, $\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X)$.

Indeed, if this is the case, then Φ sends Hermitian positive semi-definite operators with trace 1 on $\mathcal{H}_1$ to Hermitian positive semi-definite operators with trace 1 on $\mathcal{H}_2$.

However, the above conditions are not quite enough. We actually also need that, for any Hilbert space $\mathcal{K}$, the linear map $\Phi \otimes \mathrm{id} : B(\mathcal{H}_1 \otimes \mathcal{K}) \to B(\mathcal{H}_2 \otimes \mathcal{K})$ is a valid quantum transformation, i.e. that it preserves Hermiticity, positivity and trace. Indeed, we need to take into account the fact that the system $\mathcal{H}_1$, that we are acting on with the linear map Φ, could be coupled with an ancillary system $\mathcal{K}$, to which we do not have access.

Definition 3.1.1. *A transformation of a quantum system $\mathcal{H}_1$ into another quantum system $\mathcal{H}_2$ is described by a linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, called a quantum channel, which is*

- *completely positive, i.e. such that for any Hilbert space $\mathcal{K}$, the linear map $\Phi \otimes \mathrm{id} : B(\mathcal{H}_1 \otimes \mathcal{K}) \to B(\mathcal{H}_2 \otimes \mathcal{K})$ is positive;*
- *trace-preserving.*

Remark 3.1.2. *Being completely positive is a strictly stronger constraint than being positive (in the sense that there exist linear maps that are positive but not completely positive, as we will discuss in more details in Section 3.4). To check that a linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is completely positive, it is actually enough to check that the linear map $\Phi \otimes \mathrm{id} : B(\mathcal{H}_1 \otimes \mathcal{K}) \to B(\mathcal{H}_2 \otimes \mathcal{K})$ is positive for $\mathcal{K}$ such that $\dim(\mathcal{K}) = \dim(\mathcal{H}_1)$.*

In what follows, the linear maps we will consider will always be assumed to be Hermiticity-preserving (even if we do not explicitly specify it). It is easy to check that a linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is Hermiticity-preserving iff, for all $X \in B(\mathcal{H}_1)$, $\Phi(X)^* = \Phi(X^*)$.

Given a (Hermiticity-preserving) linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, its dual (Hermiticity-preserving) linear map $\Phi^* : B(\mathcal{H}_2) \to B(\mathcal{H}_1)$, for the Hilbert-Schmidt inner product, is defined by

$$\forall\ X_1 \in B(\mathcal{H}_1), X_2 \in B(\mathcal{H}_2),\ \mathrm{Tr}\left(\Phi^*(X_2)X_1\right) = \mathrm{Tr}\left(X_2\Phi(X_1)\right). \tag{3.1}$$

Theorem 3.1.3. *Given a linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ and its dual linear map $\Phi^* : B(\mathcal{H}_2) \to B(\mathcal{H}_1)$, we have the following equivalences:*

1. *Φ is positive iff Φ^* is positive;*
2. *Φ is completely positive iff Φ^* is completely positive;*
3. *Φ is trace-preserving iff Φ^* is unital, i.e. such that $\Phi^*(\mathrm{I}_{\mathcal{H}_2}) = \mathrm{I}_{\mathcal{H}_1}$.*

Proof. To prove point (1), we just have to observe that Φ being positive is equivalent to $\Phi(X_1) \geqslant 0$ for all $X_1 \geqslant 0$, i.e. $\mathrm{Tr}\,(X_2\Phi(X_1)) \geqslant 0$ for all $X_1, X_2 \geqslant 0$, while Φ^* being positive is equivalent to $\Phi^*(X_2) \geqslant 0$ for all $X_2 \geqslant 0$, i.e. $\mathrm{Tr}\,(X_1\Phi^*(X_2)) \geqslant 0$ for all $X_1, X_2 \geqslant 0$. Now, by equation (3.1), we know that $\mathrm{Tr}\,(X_1\Phi^*(X_2)) = \mathrm{Tr}\,(X_2\Phi(X_1))$ for all X_1, X_2, so the two conditions are indeed equivalent.

To prove point (2) we apply the same reasoning as to prove point (1) to $\Phi \otimes \mathrm{id}$, simply noting that $(\Phi \otimes \mathrm{id})^* = \Phi^* \otimes \mathrm{id}$.

Finally, taking $X_2 = \mathrm{I}$ in equation (3.1), we get that $\mathrm{Tr}(\Phi(X_1)\mathrm{I}) = \mathrm{Tr}(\Phi^*(\mathrm{I})X_1)$ for all X_1. Substracting $\mathrm{Tr}(X_1)$ on both sides, we thus have that $\mathrm{Tr}(\Phi(X_1)) - \mathrm{Tr}(X_1) = \mathrm{Tr}((\Phi^*(\mathrm{I}) - \mathrm{I})X_1)$ for all X_1. Now, the left-hand side of the latter equality being equal to 0 is equivalent to Φ being trace-preserving, while its right-hand side being equal to 0 is equivalent to Φ^* being unital. So we have indeed proven the equivalence claimed in point (3). □

Let $1 \leqslant p, q \leqslant \infty$, and let $1 \leqslant q' \leqslant \infty$ be such that $1/q + 1/q' + 1$. Given a linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, its p-to-q norm is defined by

$$\begin{aligned}\|\Phi\|_{p\to q} &= \sup\{\|\Phi(X_1)\|_q : X_1 \in B(\mathcal{H}_1),\, \|X_1\|_p \leqslant 1\} \\ &= \sup\{|\mathrm{Tr}\,(\Phi(X_1)X_2)| : X_1 \in B(\mathcal{H}_1),\, \|X_1\|_p \leqslant 1,\, X_2 \in B(\mathcal{H}_2),\, \|X_2\|_{q'} \leqslant 1\}\,.\end{aligned}$$

Proposition 3.1.4. *Given a linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ and its dual linear map $\Phi^* : B(\mathcal{H}_2) \to B(\mathcal{H}_1)$, we have*

$$\|\Phi\|_{1\to 1} = \|\Phi^*\|_{\infty\to\infty}.$$

Furthermore, if Φ is positive, then both norms are equal to $\|\Phi^*(\mathrm{I})\|_\infty$.

Proof. The first claim follows directly from the definition of the 1-to-1 and ∞-to-∞ norms, together with the observation that 1 and ∞ are dual indices. We thus have

$$\begin{aligned}\|\Phi\|_{1\to 1} &= \sup\{|\mathrm{Tr}\,(\Phi(X_1)X_2)| : X_1 \in B(\mathcal{H}_1),\, \|X_1\|_1 \leqslant 1,\, X_2 \in B(\mathcal{H}_2),\, \|X_2\|_\infty \leqslant 1\}\,, \\ \|\Phi^*\|_{\infty\to\infty} &= \sup\{|\mathrm{Tr}\,(\Phi^*(X_2)X_1)| : X_2 \in B(\mathcal{H}_2),\, \|X_2\|_\infty \leqslant 1,\, X_1 \in B(\mathcal{H}_1),\, \|X_1\|_1 \leqslant 1\}\,.\end{aligned}$$

Now, by equation (3.1), we know that the right-hand sides of the two equalities above are equal to one another, and hence the left-hand sides as well.

In order to prove the second claim, we will admit the following fact: if a linear map Ψ is positive, then $\|\Psi(X)\|_\infty \leqslant \|\Psi(\mathrm{I})\|_\infty \|X\|_\infty$ for any X, i.e. $\|\Psi\|_{\infty\to\infty} \leqslant \|\Psi(\mathrm{I})\|_\infty$, and there is in fact equality since $\|\mathrm{I}\|_\infty = 1$. We now just have to apply this result to Φ^*, which is positive if Φ is positive, by Theorem 3.1.3. □

As an immediate consequence of Proposition 3.1.4 we have that, if $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is a positive and trace-preserving map, then Φ, resp. Φ^*, is contracting for the 1-norm, resp. ∞-norm. This means that

$$\forall\, X_1 \in B(\mathcal{H}_1),\ \|\Phi(X_1)\|_1 \leqslant \|X_1\|_1 \quad \text{and} \quad \forall\, X_2 \in B(\mathcal{H}_2),\ \|\Phi(X_2)\|_\infty \leqslant \|X_2\|_\infty.$$

Indeed, in such case we have the chain of equalities

$$\|\Phi\|_{1\to 1} = \|\Phi^*\|_{\infty\to\infty} = \|\Phi^*(\mathrm{I})\|_\infty = \|\mathrm{I}\|_\infty = 1.$$

3.2 Representations of quantum channels

3.2.1 Choi representation

Linear maps between spaces of operators $B(\mathcal{H}_1)$ and $B(\mathcal{H}_2)$ are sometimes referred to as super operators. We first observe that there is a one-to-one correspondence between such super operators and operators on $\mathcal{H}_2 \otimes \mathcal{H}_1$. Indeed, fixing $\{e_1, \ldots, e_{d_1}\}$ an orthonormal basis of $\mathcal{H}_1$, we have the following isomorphism:

$$C : \Phi \in B(B(\mathcal{H}_1), B(\mathcal{H}_2)) \mapsto \sum_{i,j=1}^{d_1} \Phi(|e_i\rangle\langle e_j|) \otimes |e_i\rangle\langle e_j| \in B(\mathcal{H}_2 \otimes \mathcal{H}_1). \tag{3.2}$$

We call $C(\Phi)$ the Choi matrix of Φ. One can recover Φ from the knowledge of $C(\Phi)$ through

$$\forall\ X_1 \in B(\mathcal{H}_1), X_2 \in B(\mathcal{H}_2),\ \mathrm{Tr}\left(X_2 \Phi(X_1)\right) = \mathrm{Tr}\left(C(\Phi)\, X_2 \otimes X_1^T\right), \tag{3.3}$$

where we used the short-hand notation $X_1^T = T(X_1)$ for the transposition with respect to the orthonormal basis $\{e_1, \ldots, e_{d_1}\}$.

Let us make the following simple but useful observation. Denoting by $|\psi\rangle\langle\psi|$ the maximally entangled state on $\mathcal{H}_1 \otimes \mathcal{H}_1$ (in the orthonormal basis $\{e_1, \ldots, e_{d_1}\}$), i.e.

$$|\psi\rangle\langle\psi| = \frac{1}{d_1} \sum_{i,j=1}^{d_1} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|,$$

we can rewrite $C(\Phi)$ as

$$C(\Phi) = d_1 \times \Phi \otimes \mathrm{id}(|\psi\rangle\langle\psi|). \tag{3.4}$$

The definition of the Choi matrix depends on the choice of orthonormal basis, but the following properties of the Choi matrix do not.

Theorem 3.2.1. *Given a linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ with associated Choi matrix $C(\Phi) \in B(\mathcal{H}_2 \otimes \mathcal{H}_1)$, we have the following equivalences:*

1. *Φ is completely positive iff $C(\Phi) \geqslant 0$;*
2. *Φ is trace-preserving iff $\mathrm{Tr}_{\mathcal{H}_2}(C(\Phi)) = \mathrm{I}_{\mathcal{H}_1}$;*
3. *Φ is unital iff $\mathrm{Tr}_{\mathcal{H}_1}(C(\Phi)) = \mathrm{I}_{\mathcal{H}_2}$.*

Before we go into the proof, let us point out the following useful fact. Given a unit vector $\varphi \in \mathcal{H}_1 \otimes \mathcal{H}_1$, it can always be written as $\varphi = (\mathrm{I} \otimes Z)\psi$ for some maximally entangled unit vector $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_1$ and some operator $Z \in B(\mathcal{H}_1)$. Indeed, if φ has Schmidt decomposition

$$\varphi = \sum_{i=1}^{d_1} \sqrt{\lambda_i}\, u_i \otimes v_i,$$

we just have to take

$$\psi = \frac{1}{\sqrt{d_1}} \sum_{i=1}^{d_1} u_i \otimes u_i \quad \text{and} \quad Z = \sum_{i=1}^{d_1} \sqrt{d_1 \lambda_i}\, |v_i\rangle\langle u_i|.$$

Proof. Let us start with proving point (1). Suppose that Φ is completely positive. Then by definition $\Phi \otimes \mathrm{id}(X) \geqslant 0$ for any $X \geqslant 0$, so in particular $\Phi \otimes \mathrm{id}(|\psi\rangle\langle\psi|) \geqslant 0$ for $|\psi\rangle\langle\psi|$ the maximally entangled state. From the observation made in equation (3.4) that $C(\Phi) = d_1 \times \Phi \otimes \mathrm{id}(|\psi\rangle\langle\psi|)$, we have proved ($\Rightarrow$). Suppose now that $C(\Phi) \geqslant 0$, i.e. by the same observation as before that $\Phi \otimes \mathrm{id}(|\psi\rangle\langle\psi|) \geqslant 0$. Given any unit vector φ, we have just seen that it can always be written as $\varphi = (\mathrm{I} \otimes Z)\psi$ for some maximally entangled unit vector ψ and some operator Z. Hence,

$$\Phi \otimes \mathrm{id}(|\varphi\rangle\langle\varphi|) = \Phi \otimes \mathrm{id}((\mathrm{I} \otimes Z)|\psi\rangle\langle\psi|(\mathrm{I} \otimes Z^*)) = (\mathrm{I} \otimes Z)\, \Phi \otimes \mathrm{id}(|\psi\rangle\langle\psi|)\, (\mathrm{I} \otimes Z^*) \geqslant 0,$$

where the last inequality is because, for $M \geqslant 0$, $YMY^* \geqslant 0$ for all Y. Consequently, given any $X \geqslant 0$, which can always be decomposed as $X = \sum_{i=1}^{r} \mu_i |\varphi_i\rangle\langle\varphi_i|$ for some positive numbers μ_i and some unit vectors φ_i, we have by linearity that $\Phi \otimes \mathrm{id}(X) \geqslant 0$.

Let us now turn to proving point (2). By the definition (3.2) of $C(\Phi)$, we have

$$\mathrm{Tr}_{\mathcal{H}_2}(C(\Phi)) = \sum_{i,j=1}^{d_1} \mathrm{Tr}\left(\Phi(|e_i\rangle\langle e_j|)\right) |e_i\rangle\langle e_j|.$$

It is thus clear that $\mathrm{Tr}_{\mathcal{H}_2}(C(\Phi)) = \mathrm{I}_{\mathcal{H}_1}$ is equivalent to

$$\forall\ 1 \leqslant i, j \leqslant d_1,\ \mathrm{Tr}\left(\Phi(|e_i\rangle\langle e_j|)\right) = \delta_{ij} = \mathrm{Tr}\left(|e_i\rangle\langle e_j|\right),$$

which is itself equivalent to

$$\forall\ X \in B(\mathcal{H}_1),\ \mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X),$$

i.e. to Φ being TP.

We can finally prove point (3) in a similar fashion. By the definition (3.2) of $C(\Phi)$, we have

$$\mathrm{Tr}_{\mathcal{H}_1}(C(\Phi)) = \sum_{i=1}^{d_1} \Phi(|e_i\rangle\langle e_i|) = \Phi\left(\sum_{i=1}^{d_1} |e_i\rangle\langle e_i|\right) = \Phi(\mathrm{I}_{\mathcal{H}_1}).$$

It is thus clear that $\mathrm{Tr}_{\mathcal{H}_1}(C(\Phi)) = \mathrm{I}_{\mathcal{H}_2}$ is equivalent to $\Phi(\mathrm{I}_{\mathcal{H}_1}) = \mathrm{I}_{\mathcal{H}_2}$, i.e. to Φ being unital.

□

What point (1) of Theorem 3.2.1 tells us is that, in order to check that the linear map $\Phi \otimes \mathrm{id}$ preserves positivity, we do not have to check that $\Phi \otimes \mathrm{id}(X) \geqslant 0$ for all $X \geqslant 0$: it is enough to check that $\Phi \otimes \mathrm{id}(|\psi\rangle\langle\psi|) \geqslant 0$.

3.2.2 Kraus representation

Theorem 3.2.2. *A linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is completely positive iff there exist $r \in \mathbb{N}$ and $K_1, \ldots, K_r \in B(\mathcal{H}_1, \mathcal{H}_2)$ such that*

$$\forall\ X \in B(\mathcal{H}_1),\ \Phi(X) = \sum_{i=1}^{r} K_i X K_i^*. \tag{3.5}$$

The fact that Φ is trace-preserving, resp. unital, is equivalent to $\sum_{i=1}^r K_i^* K_i = \mathrm{I}_{\mathcal{H}_1}$, resp. $\sum_{i=1}^r K_i K_i^* = \mathrm{I}_{\mathcal{H}_2}$.

Expression (3.5) is called a Kraus representation of Φ, with Kraus operators $K_1, \ldots, K_r$. Such Kraus representation is not unique. For quantum channels, the possible ambiguity is well-identified: Two sets of operators $\{K_1, \ldots, K_r\}$ and $\{L_1, \ldots, L_r\}$ are Kraus operators associated to the same trace-preserving completely positive linear map iff there exists a unitary $U \in \mathcal{M}_r(\mathbb{C})$ such that, for all $1 \leqslant i \leqslant r$, $L_i = \sum_{j=1}^r U_{ij} K_j$.

Proof. By Theorem 3.2.1, we know that Φ being completely positive is equivalent to $C(\Phi) \geqslant 0$, i.e. to the fact that it can be written as $C(\Phi) = \sum_{i=1}^r |v_i\rangle\langle v_i|$, for some vectors v_i. Next, we recall that, as explained in Section 2.1.3, elements of $\mathcal{H}_2 \otimes \mathcal{H}_1$ can be identified with elements of $B(\mathcal{H}_1, \mathcal{H}_2)$. With this in mind, we will now make the following observation: If $C(\Phi) = |v\rangle\langle v|$, for some vector v, then $\Phi : X \mapsto M_v X M_v^*$, where M_v is the matrix version of v. And conversely, if $\Phi : X \mapsto KXK^*$, then $C(\Phi) = |\varphi_K\rangle\langle\varphi_K|$, where φ_K is the vector version of K. Hence, if $C(\Phi) = \sum_{i=1}^r |v_i\rangle\langle v_i|$, we get an expression of the form (3.5) for Φ by defining $K_i = M_{v_i}$, $1 \leqslant i \leqslant r$. And conversely, if Φ admits an expression of the form (3.5), we get that $C(\Phi) = \sum_{i=1}^r |\varphi_{K_i}\rangle\langle\varphi_{K_i}|$.

Next, suppose that Φ has Kraus operators $K_1, \ldots, K_r$. First, Φ being unital is equivalent to $\Phi(\mathrm{I}_{\mathcal{H}_1}) = \mathrm{I}_{\mathcal{H}_2}$, i.e. by taking $X = \mathrm{I}_{\mathcal{H}_1}$ in equation (3.5), $\sum_{i=1}^r K_i K_i^* = \mathrm{I}_{\mathcal{H}_2}$. Second, Φ being trace-preserving is equivalent to Φ^* being unital. It is easy to check that Φ^* has Kraus operators $K_1^*, \ldots, K_r^*$. So by the same argument as before the latter is equivalent to $\sum_{i=1}^r K_i^* K_i = \mathrm{I}_{\mathcal{H}_1}$. □

3.2.3 Stinespring representation

Theorem 3.2.3. *A linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is completely positive iff there exist $r \in \mathbb{N}$, a Hilbert space $\mathcal{K}$ with $\dim(\mathcal{K}) = r$ and an embedding $V : \mathcal{H}_1 \to \mathcal{H}_2 \otimes \mathcal{K}$ such that*

$$\forall\ X \in B(\mathcal{H}_1),\ \Phi(X) = \mathrm{Tr}_{\mathcal{K}}\left(VXV^*\right). \tag{3.6}$$

The fact that Φ is trace-preserving, resp. unital, is equivalent to $V^*V = \mathrm{I}_{\mathcal{H}_1}$ (i.e. to V being an isometry), resp. $\mathrm{Tr}_{\mathcal{K}}(VV^*) = \mathrm{I}_{\mathcal{H}_2}$.

Expression (3.6) is called a Stinespring representation of Φ, with Stinespring embedding V.

Proof. Assume that Φ is completely positive. By Theorem 3.2.2, this means that there exists a Kraus representation of the form (3.5) for Φ. We now set $\mathcal{K} = \mathbb{C}^r$, let $\{e_1, \ldots, e_r\}$ be an orthonormal basis of $\mathcal{K}$, and define $V : \mathcal{H}_1 \to \mathcal{H}_2 \otimes \mathcal{K}$ by

$$\forall\ \varphi \in \mathcal{H}_1,\ V|\varphi\rangle = \sum_{i=1}^{r} K_i|\varphi\rangle \otimes |e_i\rangle \in \mathcal{H}_2 \otimes \mathcal{K}.$$

It is easy to check that we then have

$$\forall\ X \in B(\mathcal{H}_1),\ VXV^* = \sum_{i,j=1}^{r} K_i X K_j^* \otimes |e_i\rangle\langle e_j| \in B(\mathcal{H}_2 \otimes \mathcal{K}),$$

and hence

$$\forall\ X \in B(\mathcal{H}_1),\ \mathrm{Tr}_{\mathcal{K}}(VXV^*) = \sum_{i=1}^{r} K_i X K_i^* = \Phi(X).$$

Conversely, assume that Φ can be written in the form (3.6). Then, setting $K_i = (\mathrm{I}_{\mathcal{H}_2} \otimes \langle e_i|)V \in B(\mathcal{H}_1, \mathcal{H}_2)$ for $1 \leqslant i \leqslant r$, we have

$$\forall\ X \in B(\mathcal{H}_1),\ \Phi(X) = \sum_{i=1}^{r} \langle e_i|VXV^*|e_i\rangle = \sum_{i=1}^{r} K_i X K_i^*.$$

By Theorem 3.2.2, this means that Φ is completely positive.

Next, suppose that Φ has Stinespring embedding V. First, Φ being unital is equivalent to $\Phi(\mathrm{I}_{\mathcal{H}_1}) = \mathrm{I}_{\mathcal{H}_2}$, i.e. by taking $X = \mathrm{I}_{\mathcal{H}_1}$ in equation (3.6), $\mathrm{Tr}_{\mathcal{K}}(VV^*) = \mathrm{I}_{\mathcal{H}_2}$. Second, Φ being trace-preserving is equivalent to Φ^* being unital. It is easy to check that the action of Φ^* can be described by $\Phi^*(X) = V^*(X \otimes \mathrm{I})V$. So by the same argument as before, Φ^* being unital is equivalent to $V^*V = \mathrm{I}_{\mathcal{H}_1}$. □

In the case where $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$, the Stinespring representation of a quantum channel (as defined in Theorem 3.2.3 above) can be reformulated as a so-called open system representation (as defined in Corollary 3.2.4 below). The latter is obtained by lifting the Stinespring isometry V from $\mathcal{H}$ to $\mathcal{H} \otimes \mathcal{K}$ to a unitary U on $\mathcal{H} \otimes \mathcal{K}$ together with some resource state $|\varphi\rangle\langle\varphi|$ on $\mathcal{K}$.

Corollary 3.2.4. *Let $\Phi : B(\mathcal{H}) \to B(\mathcal{H})$ be a completely positive trace-preserving linear map. There exist $r \in \mathbb{N}$, a Hilbert space $\mathcal{K}$ with $\dim(\mathcal{K}) = r$, a unitary $U \in B(\mathcal{H} \otimes \mathcal{K})$ and a unit vector $\varphi \in \mathcal{K}$ such that*

$$\forall\ X \in B(\mathcal{H}),\ \Phi(X) = \mathrm{Tr}_{\mathcal{K}}\left(U\left(X \otimes |\varphi\rangle\langle\varphi|\right)U^*\right). \tag{3.7}$$

Proof. We know from Theorem 3.2.3 that the action of Φ can be described by $\Phi : X \mapsto \mathrm{Tr}_{\mathcal{K}}(VXV^*)$ for some isometry $V : \mathcal{H} \to \mathcal{H} \otimes \mathcal{K}$. Now, pick a unit vector $\varphi \in \mathcal{K}$, and define the isometry $\hat{U} : \mathcal{H} \otimes \mathrm{span}\{\varphi\} \to \mathcal{H} \otimes \mathcal{K}$ by $\hat{U} : \chi \otimes \varphi \mapsto V\chi$. Since $\mathcal{H} \otimes \mathrm{span}\{\varphi\}$ is a subspace of $\mathcal{H} \otimes \mathcal{K}$, $\hat{U}$ can be extended to a unitary U on $\mathcal{H} \otimes \mathcal{K}$. By the way it is defined, such unitary U indeed satisfies, for all $X \in B(\mathcal{H})$, $\mathrm{Tr}_{\mathcal{K}}(VXV^*) = \mathrm{Tr}_{\mathcal{K}}(U\left(X \otimes |\varphi\rangle\langle\varphi|\right)U^*)$. □

Remark 3.2.5. *The smallest r such that a Kraus decomposition of Φ with r Kraus operators exists is the same as the smallest r such that a Stinespring representation of Φ with $\dim(\mathcal{K}) = r$ exists. This number is called the Kraus rank of Φ. It is also equal to $\mathrm{rank}(C(\Phi))$, the rank of the $d_1d_2 \times d_1d_2$ matrix $C(\Phi)$, and is therefore always at most d_1d_2.*

3.3 Important examples of quantum channels

3.3.1 Unitary and mixture of unitary channels

The simplest type of quantum channels are unitary channels, which are of the form:

$$\Phi : X \in B(\mathcal{H}) \mapsto UXU^* \in B(\mathcal{H}),$$

for some unitary operator U on $\mathcal{H}$. Such channel is unital and has Kraus rank 1. It is also reversible. Indeed, we have: $\Phi^{-1} = \Phi^* : X \in B(\mathcal{H}) \mapsto U^*XU \in B(\mathcal{H})$.

A generalization of the above are unitary mixture channels, which are of the form:

$$\Phi : X \in B(\mathcal{H}) \mapsto \sum_{i=1}^{r} \lambda_i U_i X U_i^* \in B(\mathcal{H}),$$

for some unitary operators $U_1, \ldots, U_r$ on $\mathcal{H}$ and some non-negative numbers $\lambda_1, \ldots, \lambda_r$ summing up to 1. Such channel is unital and has Kraus rank at most r. As soon as $r > 1$ (and $U_1 \neq U_2$), it is not reversible. The Kraus rank can thus be seen as measuring the 'amount of irreversibility'.

3.3.2 Depolarizing and dephasing channels

The fully depolarizing channel Π on $\mathcal{H} \equiv \mathbb{C}^d$ is the channel that maps any input state to the maximally mixed state I/d. It is thus defined as

$$\Pi : X \in B(\mathcal{H}) \mapsto \mathrm{Tr}(X)\frac{\mathrm{I}}{d} \in B(\mathcal{H}). \tag{3.8}$$

Such channel is unital and has maximal Kraus rank d^2. It can indeed be seen as a channel that is 'maximally irreversible', since all information on the input state is lost.

A generalization of the above is the family of depolarizing channels $\{\Pi_\lambda\}_{0\leqslant\lambda\leqslant 1}$ on $\mathcal{H} \equiv \mathbb{C}^d$, which are mixtures of the identity channel id and the fully depolarizing channel Π, i.e.

$$\Pi_\lambda : X \in B(\mathcal{H}) \mapsto \lambda\,\mathrm{id}(X) + (1-\lambda)\Pi(X) = \lambda X + (1-\lambda)\,\mathrm{Tr}(X)\frac{\mathrm{I}}{d} \in B(\mathcal{H}). \tag{3.9}$$

Such channel is unital and has maximal Kraus rank d^2 (unless $\lambda = 1$).

Another generalization of the fully depolarizing channel is the channel Π_σ on $\mathcal{H} \equiv \mathbb{C}^d$ that maps any input state to a given state σ on $\mathbb{C}^d$. It is thus defined as

$$\Pi_\sigma : X \in B(\mathcal{H}) \mapsto \mathrm{Tr}(X)\sigma \in B(\mathcal{H}).$$

Such channel has Kraus rank $d \times \mathrm{rank}(\sigma)$.

The fully dephasing channel Δ on $\mathcal{H} \equiv \mathbb{C}^d$ is the channel that maps any input state to its diagonal part, with respect to a fixed orthonormal basis $\{e_1, \ldots, e_d\}$ of $\mathbb{C}^d$. It is thus defined as

$$\Delta : X \in B(\mathcal{H}) \mapsto \sum_{i=1}^{d} \langle e_i|X|e_i\rangle |e_i\rangle\langle e_i| \in B(\mathcal{H}). \tag{3.10}$$

Such channel is unital and has Kraus rank d.

3.3.3 Quantum-to-classical and classical-to-quantum channels

A measurement procedure can also be seen as a quantum channel, often referred to as a quantum-classical channel. Indeed, given a measurement $\mathrm{M} = \{M_i\}_{1\leqslant i\leqslant n}$ on $\mathcal{H}$ (i.e. $M_i \geqslant 0$, $1 \leqslant i \leqslant n$, and $\sum_{i=1}^n M_i = \mathrm{I}_{\mathcal{H}}$), we can define the associated channel $\Phi_{\mathrm{M}} : B(\mathcal{H}) \to B(\mathbb{C}^n)$ as

$$\Phi_{\mathrm{M}} : X \in B(\mathcal{H}) \mapsto \sum_{i=1}^{n} \mathrm{Tr}(M_i X)|e_i\rangle\langle e_i| \in B(\mathbb{C}^n), \tag{3.11}$$

where $\{e_1, \ldots, e_n\}$ is a fixed orthonormal basis of $\mathbb{C}^n$. Such channel sends a quantum state ρ on $\mathbb{C}^d$ on a classical probability distribution $p = (\mathrm{Tr}(M_1\rho), \ldots, \mathrm{Tr}(M_n\rho))$ in $\mathbb{R}^n$ (which is encoded in the diagonal of an $n \times n$ matrix).

The dual concept is that of a so-called classical-quantum channel, which implements a preparation procedure. Given a collection of states $\Sigma = \{\sigma_i\}_{1\leqslant i\leqslant n}$ on $\mathcal{H}$, we can define the associated channel $\Phi_\Sigma : B(\mathbb{C}^n) \to B(\mathcal{H})$ as

$$\Phi_\Sigma : X \in B(\mathbb{C}^n) \mapsto \sum_{i=1}^{n} \langle e_i|X|e_i\rangle \sigma_i \in B(\mathcal{H}), \tag{3.12}$$

where $\{e_1, \ldots, e_n\}$ is a fixed orthonormal basis of $\mathbb{C}^n$. Such channel sends a classical probability distribution p in $\mathbb{R}^n$ (which is encoded in the diagonal of an $n \times n$ matrix) on a quantum state $\rho = \sum_{i=1}^n p_i\sigma_i$ on $\mathbb{C}^d$.

Remark 3.3.1. *The claim that the concepts of quantum-to-classical and classical-to-quantum channels are dual to one another can actually be made more precise. In order for the dual Φ^* of a quantum channel Φ to be itself a quantum channel, we need Φ to be unital. Now, for a quantum-classical channel, of the form given by equation* (3.11)*, to be unital, we need to impose that $n = d$ and that there exists $\{u_1, \ldots, u_d\}$ an orthonormal basis of $\mathbb{C}^d$ such that $M_i = |u_i\rangle\langle u_i|$ for all $1 \leqslant i \leqslant d$. This means that we actually need the quantum-classical channel to be a fully dephasing channel, as introduced in equation* (3.10)*. If this is so, then its dual is indeed a classical-quantum channel, of the form given by equation* (3.12)*, with $n = d$ and $\sigma_i = |u_i\rangle\langle u_i|$ for all $1 \leqslant i \leqslant d$.*

3.3.4 Entanglement-breaking and PPT-binding channels

A quantum channel $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is called entanglement-breaking if, for any Hilbert space $\mathcal{K}$, the quantum channel $\Phi \otimes \mathrm{id} : B(\mathcal{H}_1 \otimes \mathcal{K}) \to B(\mathcal{H}_2 \otimes \mathcal{K})$ maps any input state to a separable state, i.e.

$$\forall\ \rho \in \mathrm{D}(\mathcal{H}_1 \otimes \mathcal{K}),\ \Phi \otimes \mathrm{id}(\rho) \in \mathrm{S}(\mathcal{H}_2 \otimes \mathcal{K}).$$

This is actually equivalent to $C(\Phi)$ being a separable operator on $\mathcal{H}_2 \otimes \mathcal{H}_1$ (i.e. a non-negative multiple of a separable state).

Entanglement-breaking channels are sometimes referred to as quantum-classical-quantum channels. Indeed, a channel is entanglement-breaking iff it can be written as the composition of a quantum-classical channel and a classical-quantum channel. The latter is itself equivalent to having a Kraus decomposition where all Kraus operators have rank 1.

Similarly, a quantum channel $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is called PPT-binding if, for any Hilbert space $\mathcal{K}$, the quantum channel $\Phi \otimes \mathrm{id} : B(\mathcal{H}_1 \otimes \mathcal{K}) \to B(\mathcal{H}_2 \otimes \mathcal{K})$ maps any input state to a PPT state, i.e.

$$\forall\ \rho \in \mathrm{D}(\mathcal{H}_1 \otimes \mathcal{K}),\ \Phi \otimes \mathrm{id}(\rho) \in \mathrm{P}(\mathcal{H}_2 \otimes \mathcal{K}).$$

This is actually equivalent to $C(\Phi)$ being a PPT operator on $\mathcal{H}_2 \otimes \mathcal{H}_1$ (i.e. a non-negative multiple of a PPT state).

Remark 3.3.2. *We see from the above discussion that, in order to check that a quantum channel Φ is entanglement-breaking, resp. PPT-binding, it is not necessary to check that $\Phi \otimes \mathrm{id}(\rho)$ is separable, resp. PPT, for any input state ρ: it is enough to check it for $\rho = |\psi\rangle\langle\psi|$ a maximally entangled state.*

3.4 Positive but not completely positive maps

3.4.1 Entanglement detection

Linear maps between spaces of operators that are positive but not completely positive do not correspond to physical transformations. However, we will see that they are important tools for detecting the entanglement of bipartite states.

On a system $\mathcal{H}_1 \otimes \mathcal{H}_2$, if a state ρ is separable, then for any positive (but not necessarily completely positive) map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, $\Phi \otimes \mathrm{id}(\rho) \in B(\mathcal{H}_2 \otimes \mathcal{H}_2)$ is positive semi-definite. Indeed, it is clear that the latter is true if $\rho = \rho_1 \otimes \rho_2$ is product, as $\Phi \otimes \mathrm{id}(\rho_1 \otimes \rho_2) = \Phi(\rho_1) \otimes \rho_2 \geqslant 0$. And positivite semi-definiteness is preserved under convex combination. It turns out that this actually constitutes a characterization of separable states.

Theorem 3.4.1. *A state ρ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is separable iff for any positive linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, $\Phi \otimes \mathrm{id}(\rho) \in B(\mathcal{H}_2 \otimes \mathcal{H}_2)$ is positive semi-definite.*

Proof. The 'only if' part of the statement is clear, and explained just above. As for the 'if' part, it is a direct consequence of Theorem 2.2.2. Indeed, assume that a state ρ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is entangled. Then by Theorem 2.2.2, we know that there exists a Hermitian operator W on $\mathcal{H}_1 \otimes \mathcal{H}_2$ such that $\mathrm{Tr}(W\rho) < 0$ while $\mathrm{Tr}(W\sigma) \geqslant 0$ for all separable state σ on $\mathcal{H}_1 \otimes \mathcal{H}_2$. By the Choi isomorphism, as described by equation (3.2), we know that there exists a linear map $\Phi^* : B(\mathcal{H}_2) \to B(\mathcal{H}_1)$ such that $W = C(\Phi^*)$. Such Φ^* satisfies, for all $X_1 \in B(\mathcal{H}_1), X_2 \in B(\mathcal{H}_2)$ with $X_1, X_2 \geqslant 0$,

$$\mathrm{Tr}\left(X_1 \Phi^*(X_2)\right) = \mathrm{Tr}\left(C(\Phi^*)\, X_1 \otimes X_2^T\right) = \mathrm{Tr}\left(W\, X_1 \otimes X_2^T\right) \geqslant 0,$$

where the first equality is by equation (3.3) and the last inequality is by assumption on W (since $X_1, X_2^T \geqslant 0$ and thus $X_1 \otimes X_2^T$ is a non-negative multiple of a separable state). This shows that $\Phi^* : B(\mathcal{H}_2) \to B(\mathcal{H}_1)$, and hence $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ as well, is a positive linear map. What is more, denoting by $|\psi\rangle\langle\psi|$ the maximally entangled state on $\mathcal{H}_2 \otimes \mathcal{H}_2$, we have

$$\mathrm{Tr}\left(|\psi\rangle\langle\psi|\, \Phi \otimes \mathrm{id}(\rho)\right) = \mathrm{Tr}\left(\Phi^* \otimes \mathrm{id}(|\psi\rangle\langle\psi|)\, \rho\right) = \frac{1}{d_2}\, \mathrm{Tr}\left(C(\Phi^*)\rho\right) = \frac{1}{d_2}\, \mathrm{Tr}\left(W\rho\right) < 0.$$

This shows that $\Phi \otimes \mathrm{id}(\rho) \in B(\mathcal{H}_2 \otimes \mathcal{H}_2)$ is not positive semi-definite. □

The content of Theorem 3.4.1 can be rephrased as the following entanglement detection statement: If a state ρ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is entangled, then there exists a positive (but not completely positive) map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ such that $\Phi \otimes \mathrm{id}(\rho) \in B(\mathcal{H}_2 \otimes \mathcal{H}_2)$ is not positive semi-definite. Such map Φ is said to detect the entanglement of ρ (since on the contrary, for any separable state σ on $\mathcal{H}_1 \otimes \mathcal{H}_2$, $\Phi \otimes \mathrm{id}(\sigma) \in B(\mathcal{H}_2 \otimes \mathcal{H}_2)$ is positive semi-definite).

The most obvious example of a map that is positive but not completely positive is the transposition $T : B(\mathcal{H}_1) \to B(\mathcal{H}_1)$. And indeed, transposition detects some entangled states, namely those that are not PPT. But we also know that it does not detect all entangled states, as there exist states that are PPT and nevertheless entangled. What Theorem 3.4.1 tells us is that looking at all positive maps rather than just a specific one actually allows to detect all entangled states.

3.4.2 n-positivity

Definition 3.4.2. *Let $n \in \mathbb{N}$. A linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is called n-positive if the linear map $\Phi \otimes \mathrm{id} : B(\mathcal{H}_1 \otimes \mathbb{C}^n) \to B(\mathcal{H}_2 \otimes \mathbb{C}^n)$ is positive. We denote by $\mathbf{P}_n(\mathcal{H}_1, \mathcal{H}_2)$ the set of n-positive linear maps from $B(\mathcal{H}_1)$ to $B(\mathcal{H}_2)$.*

Clearly, if $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is n-positive, then it is automatically m-positive for all $m \leqslant n$. What is more, a 1-positive map is nothing else than a positive map, while for all $n \geqslant d_1$ an n-positive map is nothing else than a completely positive map. Denoting by $\mathbf{P}(\mathcal{H}_1, \mathcal{H}_2)$, resp. $\mathbf{CP}(\mathcal{H}_1, \mathcal{H}_2)$, the set of positive, resp. completely positive, linear maps from $B(\mathcal{H}_1)$ to $B(\mathcal{H}_2)$, we thus have the chain of inclusions:

$$\mathbf{CP}(\mathcal{H}_1, \mathcal{H}_2) = \mathbf{P}_{d_1}(\mathcal{H}_1, \mathcal{H}_2) \subset \mathbf{P}_{d_1-1}(\mathcal{H}_1, \mathcal{H}_2) \subset \cdots \subset \mathbf{P}_2(\mathcal{H}_1, \mathcal{H}_2) \subset \mathbf{P}_1(\mathcal{H}_1, \mathcal{H}_2) = \mathbf{P}(\mathcal{H}_1, \mathcal{H}_2).$$

Theorem 3.4.3. *Given a linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ with associated Choi matrix $C(\Phi) \in B(\mathcal{H}_2 \otimes \mathcal{H}_1)$, we have the following equivalence: For any $n \leqslant d_1$, Φ is n-positive iff $\langle\varphi|C(\Phi)|\varphi\rangle \geqslant 0$ for all $\varphi \in \mathcal{H}_2 \otimes \mathcal{H}_1$ with Schmidt rank at most n.*

Proof. By definition, Φ being n-positive is equivalent to $\Phi \otimes \mathrm{id}(|\varphi\rangle\langle\varphi|) \geqslant 0$ for all $\varphi \in \mathcal{H}_1 \otimes \mathbb{C}^n$. Embedding $\mathbb{C}^n$ into $\mathcal{H}_1$, the latter is equivalent to $\Phi \otimes \mathrm{id}(|\varphi\rangle\langle\varphi|) \geqslant 0$ for all $\varphi \in \mathcal{H}_1 \otimes \mathcal{H}_1$ with Schmidt rank at most n. Now, such $\varphi \in \mathcal{H}_1 \otimes \mathcal{H}_1$ can always be written as $|\varphi\rangle = \mathrm{I} \otimes Z|\psi\rangle$ for some $Z \in B(\mathcal{H}_1)$ with rank at most n and $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_1$ maximally entangled. We thus have that Φ being n-positive is equivalent to $(\mathrm{I} \otimes Z)\Phi \otimes \mathrm{id}(|\psi\rangle\langle\psi|)(\mathrm{I} \otimes Z^*) = (\mathrm{I} \otimes Z)C(\Phi)(\mathrm{I} \otimes Z^*) \geqslant 0$ for all $Z \in B(\mathcal{H}_1)$ with rank at most n. Now, the latter is equivalent to $\langle\chi|(\mathrm{I} \otimes Z)C(\Phi)(\mathrm{I} \otimes Z^*)|\chi\rangle \geqslant 0$ for all $Z \in B(\mathcal{H}_1)$ with rank at most n and all $\chi \in \mathcal{H}_2 \otimes \mathcal{H}_1$. And since any $\varphi' \in \mathcal{H}_2 \otimes \mathcal{H}_1$ with Schmidt rank at most n can be written as $|\varphi'\rangle = (\mathrm{I} \otimes Z^*)|\chi\rangle$ for some $Z \in B(\mathcal{H}_1)$ with rank at most n and some $\chi \in \mathcal{H}_2 \otimes \mathcal{H}_1$, we have actually proven the desired equivalence. □

3.5 Outputs of quantum channels

3.5.1 Spectral properties

If a linear map has same input and output spaces, we can assign a spectrum to it. Given a linear map $\Phi : B(\mathcal{H}) \to B(\mathcal{H})$, its spectrum $\mathrm{spec}(\Phi)$ is defined as

$$\mathrm{spec}(\Phi) = \{\lambda \in \mathbb{C} : \exists\, X \in B(\mathcal{H}) : \Phi(X) = \lambda X\}.$$

If $\lambda \in \mathbb{C}$ is such that $\Phi(X) = \lambda X$ for some $X \in B(\mathcal{H})$, we call it an eigenvalue of Φ. If $X \in B(\mathcal{H})$ is such that $\Phi(X) = \lambda X$ for some $\lambda \in \mathbb{C}$, we call it an eigenvector of Φ (even though it is an operator). We further define the spectral radius of Φ, which we denote $\varrho(\Phi)$, as

$$\varrho(\Phi) = \sup\{|\lambda| : \lambda \in \mathrm{spec}(\Phi)\}.$$

We have already explained, in Section 2.1.3, that linear maps on $\mathbb{C}^d$ can be identified with $\mathbb{C}^d \otimes \mathbb{C}^d$. Similarly, linear maps on $\mathcal{M}_d(\mathbb{C})$ can be identified with $\mathcal{M}_d(\mathbb{C}) \otimes \mathcal{M}_d(\mathbb{C})$. Hence, to any linear map $\Phi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$, we can associate its matrix version $M_\Phi \in \mathcal{M}_d(\mathbb{C}) \otimes \mathcal{M}_d(\mathbb{C}) \equiv \mathcal{M}_{d^2}(\mathbb{C})$. And the spectrum of Φ is nothing else than the spectrum of the $d^2 \times d^2$ complex matrix M_Φ.

Theorem 3.5.1. *Let $\Phi : B(\mathcal{H}) \to B(\mathcal{H})$ be a positive linear map. Its spectral radius always satisfies*

$$\varrho(\Phi) \leqslant \|\Phi(\mathrm{I})\|_\infty.$$

If Φ is additionally trace-preserving or unital, then $\varrho(\Phi) = 1$ and $1 \in \mathrm{spec}(\Phi)$.

Proof. We will admit the following fact: if Φ is positive, then for any $X \in B(\mathcal{H})$, $\|\Phi(X)\|_\infty \leqslant \|\Phi(\mathrm{I})\|_\infty \|X\|_\infty$. Therefore, for any $\lambda \in \mathrm{spec}(\Phi)$, letting $X \in B(\mathcal{H})$ be such that $\Phi(X) = \lambda X$, we have

$$|\lambda| \|X\|_\infty = \|\Phi(X)\|_\infty \leqslant \|\Phi(\mathrm{I})\|_\infty \|X\|_\infty.$$

This proves that $\varrho(\Phi) \leqslant \|\Phi(\mathrm{I})\|_\infty$.

If Φ is additionally unital, then $\Phi(\mathrm{I}) = \mathrm{I}$, which proves both that $\varrho(\Phi) \leqslant \|\Phi(\mathrm{I})\|_\infty = \|\mathrm{I}\|_\infty = 1$ and that 1 is actually an eigenvalue of Φ (with associated eigenvector I). While if Φ is additionally trace-preserving, we just have to apply what precedes to Φ^*, which is unital and has the same spectrum as Φ. □

What Theorem 3.5.1 tells us is that a positive linear map $\Phi : B(\mathcal{H}) \to B(\mathcal{H})$ that is either trace-preserving or unital has all its eigenvalues lying in the complex unit disc and a fixed point, i.e. there exists $X \in B(\mathcal{H})$ such that $\Phi(X) = X$. In particular, any quantum channel $\Phi : B(\mathcal{H}) \to B(\mathcal{H})$ satisfies this.

Theorem 3.5.2. *Let $\Phi : B(\mathcal{H}) \to B(\mathcal{H})$ be a positive linear map, and denote by $\varrho = \varrho(\Phi)$ its spectral radius. Then, ϱ is an eigenvalue of Φ, and it has a positive semi-definite associated eigenvector, i.e.*

$$\exists\ X \in B(\mathcal{H}) : X \geqslant 0 \text{ and } \Phi(X) = \varrho X.$$

Combining Theorems 3.5.1 and 3.5.2 we get that a positive linear map $\Phi : B(\mathcal{H}) \to B(\mathcal{H})$ that is either trace-preserving or unital always has a positive semi-definite fixed point, i.e. there exists $X \in B(\mathcal{H})$ such that $X \geqslant 0$ and $\Phi(X) = X$. This is in particular the case for any quantum channel $\Phi : B(\mathcal{H}) \to B(\mathcal{H})$. Simply dividing by the trace, we can further impose that the positive semi-definite fixed point has trace 1, i.e. is a quantum state. Hence summarizing, we have that, for any quantum channel $\Phi : B(\mathcal{H}) \to B(\mathcal{H})$, there exists $\rho_* \in \mathrm{D}(\mathcal{H})$ such that $\Phi(\rho_*) = \rho_*$.

Remark 3.5.3. *The latter statement, that quantum channels always have a fixed state, is in fact a direct consequence of Brower's fixed point theorem. Indeed, a quantum channel Φ is a continuous map from the non-empty compact convex set $\mathrm{D}(\mathcal{H})$ (which is a subset of the real vector space of Hermitian operators on $\mathcal{H}$), to itself. So there must exist a fixed point of Φ, i.e. $\rho_* \in \mathrm{D}(\mathcal{H})$ such that $\Phi(\rho_*) = \rho_*$.*

3.5.2 Output entropies

Definition 3.5.4. *Let $p > 1$. Given a state ρ on $\mathcal{H}$, its p-Rényi entropy is defined as*

$$S_p(\rho) = \frac{1}{1-p} \log\left(\mathrm{Tr}\left(\rho^p\right)\right).$$

This definition can be extended to the case $p = 1$: the 1-Rényi entropy of ρ is the limit as p goes to 1 of its p-Rényi entropy, which actually coincides with its von Neumann entropy, i.e.

$$S_1(\rho) = \lim_{p \to 1} S_p(\rho) = S(\rho) = -\mathrm{Tr}\left(\rho \log(\rho)\right).$$

Note that, for $p > 1$, the p-Rényi entropy can be expressed in terms of the Schatten p-norm as

$$S_p(\rho) = \frac{p}{1-p} \log\left(\|\rho\|_p\right).$$

We see from this observation that the definition of the p-Rényi entropy can also be extended to the case $p = \infty$ as

$$S_\infty(\rho) = \lim_{p \to \infty} S_p(\rho) = -\log\left(\|\rho\|_\infty\right).$$

The p-Rényi entropies, for $p \in [1, \infty]$, share many properties of the von Neumann entropy. In particular,

1. S_p is concave.
2. Given a state ρ on $\mathcal{H} \equiv \mathbb{C}^d$, $0 \leqslant S_p(\rho) \leqslant \log(d)$, with equality in the first inequality iff ρ is pure and in the second inequality iff ρ is maximally mixed.
3. S_p is invariant under unitary conjugation, i.e. given a state ρ on $\mathcal{H}$, for all unitary U on $\mathcal{H}$, $S_p(U\rho U^*) = S_p(\rho)$.
4. S_p is additive on tensor products, i.e. given states ρ, σ on $\mathcal{H}_1, \mathcal{H}_2$, $S_p(\rho \otimes \sigma) = S_p(\rho) + S_p(\sigma)$.

Definition 3.5.5. *Let $p \in [1,\infty]$. Given a quantum channel $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, its minimum output p-Rényi entropy is defined as*

$$S_p^{min}(\Phi) = \min\left\{S_p(\Phi(\rho)) : \rho \in \mathrm{D}(\mathcal{H}_1)\right\}.$$

In the case $p = 1$, we usually write S^{min} instead of S_1^{min} for the minimum output von Neumann entropy.

For $p \in]1,\infty]$, the minimum output p-Rényi entropy of a quantum channel Φ is related to its 1-to-p norm. Indeed, observe that

$$\min\left\{S_p(\Phi(\rho)) : \rho \in \mathrm{D}(\mathcal{H}_1)\right\} = \frac{p}{1-p}\log\left(\max\left\{\|\Phi(\rho)\|_p : \rho \in \mathrm{D}(\mathcal{H}_1)\right\}\right),$$

and we thus have

$$S_p^{min}(\Phi) = \frac{p}{1-p}\log\left(\|\Phi\|_{1\to p}\right).$$

For a long time, a central open question in quantum information theory has been to decide whether the quantity S_p^{min}, for $p \in [1,\infty]$, is additive under tensor product, i.e. whether, for any quantum channels $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ and $\Phi' : B(\mathcal{H}_1') \to B(\mathcal{H}_2')$,

$$S_p^{min}(\Phi \otimes \Phi') = S_p^{min}(\Phi) + S_p^{min}(\Phi').$$

It is clear that S_p^{min} is sub-additive. Indeed,

$$\begin{aligned}
S_p^{min}(\Phi \otimes \Phi') &= \min\left\{S_p(\Phi \otimes \Phi'(\sigma)) : \sigma \in \mathrm{D}(\mathcal{H}_1 \otimes \mathcal{H}_1')\right\} \\
&\leqslant \min\left\{S_p(\Phi \otimes \Phi'(\rho \otimes \rho')) : \rho \in \mathrm{D}(\mathcal{H}_1),\, \rho' \in \mathrm{D}(\mathcal{H}_1')\right\} \\
&= \min\left\{S_p(\Phi(\rho)) : \rho \in \mathrm{D}(\mathcal{H}_1)\right\} + \min\left\{S_p(\Phi'(\rho')) : \rho' \in \mathrm{D}(\mathcal{H}_1')\right\} \\
&= S_p^{min}(\Phi) + S_p^{min}(\Phi'),
\end{aligned}$$

where the first inequality is because $\mathrm{D}(\mathcal{H}_1 \otimes \mathcal{H}_1') \supset \{\rho \otimes \rho' : \rho \in \mathrm{D}(\mathcal{H}_1),\, \rho' \in \mathrm{D}(\mathcal{H}_1')\}$ and the second equality is because $\Phi \otimes \Phi'(\rho \otimes \rho') = \Phi(\rho) \otimes \Phi'(\rho')$ and $S_p(\tau \otimes \tau') = S_p(\tau) + S_p(\tau')$. So the question is, more precisely, whether there exist quantum channels $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ and $\Phi' : B(\mathcal{H}_1') \to B(\mathcal{H}_2')$ such that

$$S_p^{min}(\Phi \otimes \Phi') < S_p^{min}(\Phi) + S_p^{min}(\Phi').$$

In other words, we are wondering if minimizing the output entropy of a product channel $\Phi \otimes \Phi'$ over product or non-product input states is the same. In fact, it is easy to see that, for any separable state σ on $\mathcal{H}_1 \otimes \mathcal{H}_1'$, $S_p(\Phi \otimes \Phi'(\sigma)) \geqslant S_p^{min}(\Phi) + S_p^{min}(\Phi')$. So the question is whether entangled inputs allow to attain a strictly smaller output entropy.

Note that, for $p \in]1,\infty]$, this problem is equivalent to deciding whether or not the 1-to-p norm is multiplicative under tensor product, i.e. whether, for any quantum channels $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ and $\Phi' : B(\mathcal{H}_1') \to B(\mathcal{H}_2')$,

$$\|\Phi \otimes \Phi'\|_{1\to p} = \|\Phi\|_{1\to p} \times \|\Phi'\|_{1\to p},$$

or there exist quantum channels $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ and $\Phi' : B(\mathcal{H}_1') \to B(\mathcal{H}_2')$ such that

$$\|\Phi \otimes \Phi'\|_{1\to p} > \|\Phi\|_{1\to p} \times \|\Phi'\|_{1\to p}.$$

The case $p \in]1,\infty]$ turns out to be easier to handle than the case $p = 1$, precisely thanks to this norm formulation of the problem. Counter-examples to the conjecture that S_p^{min} is additive were therefore first found for $p \in]1,\infty]$, and only later for $p = 1$.

Theorem 3.5.6. *Let $p \in]1,\infty]$. There exist $d_0 \in \mathbb{N}$ and $c > 0$ such that, for all $d \geqslant d_0$, there is a quantum channel $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, with $d_1 = cd^{1+1/p}$ and $d_2 = d$, satisfying*

$$S_p^{min}(\Phi \otimes \overline{\Phi}) < S_p^{min}(\Phi) + S_p^{min}(\overline{\Phi}).$$

Theorem 3.5.7. *There exist $d_0 \in \mathbb{N}$ and $c > 0$ such that, for all $d \geqslant d_0$, there is a quantum channel $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, with $d_1 = cd$ and $d_2 = d$, satisfying*

$$S^{min}(\Phi \otimes \overline{\Phi}) < S^{min}(\Phi) + S^{min}(\overline{\Phi}).$$

Bibliography

[1] G. Aubrun, S.J. Szarek. *Alice and Bob meet Banach: The interface of asymptotic geometric analysis and quantum information theory.* American Mathematical Society, Mathematical Surveys and Monographs 223, 2017. Pdf available at http://math.univ-lyon1.fr/ aubrun/ABMB/ABMB.pdf.

[2] J. Watrous. *The Theory of Quantum Information.* Cambridge University Press, 2018. Pdf available at https://cs.uwaterloo.ca/ watrous/TQI/TQI.pdf.

[3] M.M. Wilde. *Quantum Information Theory.* Cambridge University Press, 2013. Pdf available at https://markwilde.com/qit-notes.pdf.

[4] M.M. Wolf. *Quantum channels & Operations – Guided tour.* Lecture notes available at https://mediatum.ub.tum.de/doc/1701036/1701036.pdf.

[5] M.M. Wolf. *Mathematical Introduction to Quantum Information Processing.* Lecture notes available at https://mediatum.ub.tum.de/doc/1706981/1706981.pdf.