ECOLE
POLYTECHNIQUE
DE BRUXELLES

ULB

UNIVERSITÉ LIBRE DE BRUXELLES

Majorization relations and quantum entanglement

Mémoire présenté en vue de l'obtention du diplôme
d'Ingénieur Civil Physicien à finalité spécialisée

Serge Deside

Directeur
Professeur Nicolas Cerf

Superviseur
Matthieu Arnhem

Service
QuIC

Année académique
2021 - 2022

Abstract

"Majorization relations and quantum entanglement" by Serge Deside, for the degree of *Master of Science in Physical Engineering*, 2021-2022.

The theory of majorization has long been recognized as a powerful formalism in probability theory, providing appropriate mathematical tools in order to compare random variables. More recently, it has been realized to have equally prominent applications in quantum information sciences, especially in relation with quantum entanglement theory. Notably, in 1999, Nielsen exhibited the relevance of majorization theory for comparing quantum entangled states. He proved that the deterministic conversion between two bipartite pure states using only local operations and classical communication is linked to a majorization relation between the Schmidt coefficients of both states. While this finding has led to numerous ramifications over the last 20 years, the quantum significance of several more refined notions of majorization theory has remained mostly unexplored as of today. The objective of this Ms thesis is to fill this gap and explore in particular the notions of weak (sub or super) majorization as well as majorization lattice in the context of entanglement theory. First, we have conjectured (and partly proven) a necessary condition on the deterministic conversion between two bipartite pure states based on a weak majorization relation connecting the operator Schmidt decomposition of their corresponding density operators. Then, we have shown that this weak majorization relation also yields a sufficient condition when considering certain probabilistic (instead of deterministic) conversion protocols. Such protocols led us to understand the role of a weak majorization condition between bipartite mixed states of a specific form, which opens the way to a possible extension of Nielsen's theorem to mixed states by exploiting the operator Schmidt decomposition. Next, we have investigated the potential quantum applications of the lattice structure induced by majorization displayed by a collection of random variables. Turning to the quantum counterpart of this notion, we have analyzed the optimal common resource and common product entangled states, elaborating on the so-called "join" and "meet" probabilities in the lattice. Finally, we have exploited the optimal common resource state to define a new protocol for bipartite state conversions that performs as well as the optimal probabilistic protocol. This strongly suggests that the majorization lattice formalism is a most promising avenue for better understanding the resource of quantum entanglement.

Keywords: Weak majorization, quantum entanglement, operator Schmidt decomposition, deterministic and probabilistic conversions, majorization lattice.

Remerciements

D’emblée, j’aimerais sincèrement remercier Monsieur Nicolas Cerf pour m’avoir donné l’opportunité de travailler sur ce mémoire ainsi que pour sa disponibilité et ses conseils tout au long de l’année.

Ensuite, je souhaite remercier Matthieu Arnhem et Célia Griffet pour l’attention qu’ils ont accordé à mes questions. Les réunions régulières m’ont indéniablement aidé à structurer mon travail sur ce mémoire.

Je tiens également à remercier mes camarades, particulièrement Ludéric, Raphaël et Nicolas, sans qui ces années d’études n’auraient pas été aussi réjouissantes.

Enfin, je remercie ma famille, mes amis, Jeremy et Ophélie pour m’avoir soutenu durant ces dernières années.

Contents

Introduction

Quantum entanglement has troubled physicists for almost a century. From the seminal paper of Einstein, Podolsky and Rosen [1] concluding that the quantum mechanical description of physical reality could not be considered complete, to the confirmation by Bell [2] that non-local properties of quantum mechanics were necessary for the theory to hold, the notion of entanglement has been at the center of many philosophical debates about Nature and, as such, still constitutes one of the most puzzling features of quantum mechanics.

On the other hand, in quantum information theory, entanglement is seen as a resource rather than an anomaly. In fact, many significant protocols, such as quantum teleportation [3] or dense coding [4], rely on the use of maximally-entangled states. It is in this context that, in 1999, Nielsen developed a necessary and sufficient condition for the entanglement transformation of bipartite pure states under the constraint of using only local operations and classical communication (LOCC) for the two remote parties [5]. Interestingly enough, such condition is simply expressed in terms of a majorization relation between the eigenvalues of the initial and final reduced states. Soon after, Vidal generalized Nielsen's theorem to the class of probabilistic LOCC transformations [6].

The theory of majorization is a well-known mathematical field [7] which gives a preorder on probability distributions. In an informational context, the majorization relation is connected to a strong measure of disorder as it implies an inequality on the Shannon entropy [8] as well as on all Rényi entropies [9] of both distributions under comparison.

Furthermore, in 2002, Cicalese and Vaccaro combined majorization with the concept of lattice [10]. Applied to quantum information [11, 12], the majorization lattice enables the definition of an optimal common resource (OCR), i.e., a state that can be used as a "minimal entanglement" resource for a given set of states, and its counterpart, an optimal common product (OCP), i.e., a state with "maximum entanglement" that can be produced from any one of the states in the set.

In this Ms thesis, we investigate quantum information applications of some more refined notions of majorization, namely weak majorization and the majorization lattice. Especially, we address the possibility of extending Nielsen's theorem to mixed states. Such states being ubiquitous in realistic conditions, this problem is of great importance. However, finding a necessary and sufficient condition for the conversion of bipartite mixed states is known to be hard, hence we don't propose a full solution but rather explore some leads in that direction. At first, we exploit the operator Schmidt (OS) decomposition of bipartite mixed states, recently studied at QuIC [13], to the set of bipartite pure states. Indeed, pure states being a special case of mixed states, we begin by exploring how Nielsen's theorem would translate in the OS decomposition language. It turns out that, by doing so, distributions associated to pure states are now non-normalized, hence we have to make use of an extension of majorization, namely weak majorization. Overall, the relation we propose is studied in regards to Nielsen's and Vidal's theorems and also in the context of the majorization lattice. Finally, we explore quantum information applications of the majorization lattice, notably by proposing a protocol for the probabilistic conversion of bipartite pure states equivalent to the optimal one presented by Vidal in [6] and exploiting the majorization lattice notion of OCR.

This Ms thesis is divided in two parts. The first one (Chapters 1-3) introduces the theoretical background, while the second one (Chapters 4-6) presents original results. In Chapter 1, we review the main quantum mechanical ingredients used in the rest of the Ms thesis. In Chapter 2, we present the mathematical theory of majorization, which is then applied, in Chapter 3, to entanglement transformations for Nielsen's and Vidal's theorems. Thereafter, in Chapter 4, we propose a conjecture linking weak majorization on the OS coefficients and Nielsen's theorem, i.e., deterministic LOCC transformations. Similarly, in Chapter 5, we prove a theorem linking weak majorization on the OS coefficients and Vidal's theorem, i.e., probabilistic LOCC transformations. Moreover, at the end of the chapter, we make a first incursion into the realm of mixed states by proposing a criterion for specific LOCC transformations, linked to Vidal's theorem, involving an initial pure state and a final mixed state. Finally, in Chapter 6, we introduce the theory of majorization lattice, allowing us to showcase in a unified picture the conjecture and the theorem coming from Chapters 4 and 5. Then, we present some applications of majorization lattice to entanglement transformations. Elaborating on this, we conclude by proposing an alternative protocol to Vidal's [6] making use of the OCR.

Part I

Theoretical Background

Chapter 1

Basics of quantum mechanics

This first chapter introduces the fundamental notions and notations of quantum mechanics that will be used throughout the entirety of this report.

In Section 1.1, we introduce the two classes of states used to describe quantum systems, the pure states and the mixed states. Then, in Section 1.2, we present the theory of bipartite systems, where states are shared between two, possibly distant, parties. Finally, in Section 1.3, we give a brief overview of the concept of measurement in quantum mechanics.

Unless otherwise stated, the content of this chapter follows [14] and [15].

1.1 States in quantum mechanics

The quantum mechanical description of states can be classified into two parts, the pure states and the mixed states. On the one hand, a pure state constitutes the ideal case, with the state of the system well defined. On the other hand, a mixed state constitutes the realistic case, the system being considered as a statistical mixture of pure states. Obviously, pure states arise as special cases of mixed states, when the system contains no statistical uncertainty.

1.1.1 Pure states

In the Dirac notation, a pure state is described as an element of a complex vector space, called the Hilbert space, which we denote

$$|\psi\rangle \in \mathcal{H}. \tag{1.1}$$

An important feature of quantum mechanics, which distinguishes it from classical mechanics, is the superposition principle. This means that the state of a system can be constructed as the sum of other states, e.g., $|\psi\rangle = a_1 |\psi_1\rangle + a_2 |\psi_2\rangle$[1], with $a_1, a_2 \in \mathbb{C}$. This peculiarity implies that, although the state is fixed, an appropriate measurement will lead to either measuring the system in the state $|\psi_1\rangle$ or in the state $|\psi_2\rangle$. Hence, uncertainties lie at the heart of quantum mechanics.

In order to describe elements of a vector space, a basis must be chosen. In quantum information, analogously to the classical bits, quantum bits (qubits) are defined as

[1]States must be normalized, hence $|a_1|^2 + |a_2|^2 = 1$.

elements of a 2-dimensional Hilbert space and the often used basis is the set $\{|0\rangle, |1\rangle\}$, called the computational basis. A general state under this basis can be written

$$|\psi\rangle = a_0 |0\rangle + a_1 |1\rangle = a_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}, \tag{1.2}$$

where the two last equalities are used to emphasize the vector representation of states.

1.1.2 Mixed states

We have seen that quantum mechanical states exhibit fundamental uncertainties but, in a more general case, classical uncertainties are often also part of the system. When there are, we say that the system is in a statistical mixture of pure states or simply, in a mixed state.

In the Dirac notation, a mixed state is described by means of an operator, the so-called density operator[2].

A density operator has to fulfill three conditions :

- $\hat{\rho} = \hat{\rho}^\dagger$, i.e., the operator is hermitian (thus, its eigenvalues are real).
- $\hat{\rho} \geq 0$, i.e., the operator is positive semi-definite (thus, its eigenvalues are non-negative).
- $\mathrm{Tr}(\hat{\rho}) = 1$, meaning that the state is normalized (its eigenvalues sum to 1).

Because pure states are a special case of mixed states, the density operator formalism can also be used to describe such states, hence we begin by their description.

1.1.2.a Density operator for pure states

Let $|\psi\rangle$ be the state of the system, the density operator derived from $|\psi\rangle$ is

$$\hat{\rho} = |\psi\rangle \langle\psi|, \tag{1.3}$$

where $\langle\psi|$ is an element of the dual space $\mathcal{H}^*$. As such, it can be seen as the complex conjugate transpose of $|\psi\rangle$ in a vector representation.

For qubits, the general form of a density operator for a pure state using the computational basis (1.2) is as follows

$$\hat{\rho} = (a_0 |0\rangle + a_1 |1\rangle)(a_0^* \langle 0| + a_1^* \langle 1|), \tag{1.4}$$

$$= |a_0|^2 |0\rangle\langle 0| + a_0 a_1^* |0\rangle\langle 1| + a_1 a_0^* |1\rangle\langle 0| + |a_1|^2 |1\rangle\langle 1|, \tag{1.5}$$

$$= \begin{pmatrix} |a_0|^2 & a_0 a_1^* \\ a_1 a_0^* & |a_1|^2 \end{pmatrix}, \tag{1.6}$$

which is obviously a specific form of density operator that not all density operators fulfill, e.g., $\hat{\rho} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$. This last state corresponds in fact to a maximally mixed state.

[2]While it could seem much more complex to use operators, it turns out to be one of the simpler ways to treat such states.

1.1.2.b Density operator for mixed states

Let a system be described by an ensemble of pure states $\{p_i, |\psi_i\rangle\}$. This means that the system is in state $|\psi_i\rangle$ with a probability p_i. The density operator of the system can be written

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle \langle\psi_i| \, . \tag{1.7}$$

Note the important remark that the states $|\psi_i\rangle$ need not be orthogonal to one another. Furthermore, there exist several realizations[3] of $\hat{\rho}$ depending on the basis used. An important one is the eigenbasis decomposition $\{\lambda_i, |\chi_i\rangle\}$, with λ_i the eigenvalues of $\hat{\rho}$ and $|\chi_i\rangle$ the corresponding eigenvectors:

$$\hat{\rho} = \sum_i \lambda_i |\chi_i\rangle \langle\chi_i| \, . \tag{1.8}$$

Obviously, because we used the eigenbasis for this realization, the states are now orthogonal to one another. Note that Equation 1.8 is deduced from the spectral theorem which states that any hermitian operator is diagonal in its eigenbasis.

1.1.2.c Distinguishing pure states from mixed states

As the density operator formalism allows to treat both classes of states, a criterion to distinguish them is of great importance.

We first define the notion of trace of an operator in the Dirac notation. This is essentially the sum of the diagonal elements of the matrix representation of the operator.

Definition 1.1 (Trace). *Let $\hat{A}$ be an operator, then the trace of $\hat{A}$ is defined as*

$$\mathrm{Tr}\left(\hat{A}\right) = \sum_i \langle i|\hat{A}|i\rangle \, , \tag{1.9}$$

where $\{|i\rangle\}$ is any orthonormal basis.

Therefore, we can introduce the concept of purity, allowing to distinguish pure states from mixed states.

Definition 1.2 (Purity). *The purity of a state $\hat{\rho}$ is defined as $\gamma = \mathrm{Tr}(\hat{\rho}^2)$.*

The following theorem allows, when given a state $\hat{\rho}$, to determine whether it is pure or not.

Theorem 1.1. *Let $\hat{\rho}$ be a density operator, $\hat{\rho}$ is a pure state if and only if its purity equals one.*

Note that the purity of a mixed state is always strictly smaller than one.

[3] Synonym of ensemble of pure states.

1.2 Composite systems

Communication, which is central in this report, needs two or more distant parties interacting with each other, hence the need to formalise the description of composite quantum systems. In our case, we will consider communication between two parties. As it is commonly used in quantum information, the first system will be associated to Alice's laboratory, while the second to Bob's.

Conventionally, a bipartite system is described by means of a tensor product. The state space of the whole system ($\mathcal{H}_{\mathcal{AB}}$) is built up from the state spaces of Alice and Bob ($\mathcal{H}_{\mathcal{A}}$ and $\mathcal{H}_{\mathcal{B}}$ respectively), denoted

$$\mathcal{H}_{\mathcal{AB}} = \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}. \tag{1.10}$$

1.2.1 Bipartite pure states

Suppose that Alice possesses a qubit in the state $|\psi\rangle_A = a_0 |0\rangle + a_1 |1\rangle$ and Bob a qubit in the state $|\psi\rangle_B = b_0 |0\rangle + b_1 |1\rangle$. The state of the total system can be written

$$\begin{aligned} |\psi\rangle_{AB} &= |\psi\rangle_A \otimes |\psi\rangle_B\,, && (1.11)\\ &= (a_0 |0\rangle + a_1 |1\rangle) \otimes (b_0 |0\rangle + b_1 |1\rangle), && (1.12)\\ &= a_0 b_0 |0\rangle \otimes |0\rangle + a_0 b_1 |0\rangle \otimes |1\rangle + a_1 b_0 |1\rangle \otimes |0\rangle + a_1 b_1 |1\rangle \otimes |1\rangle\,, && (1.13)\\ &= a_0 b_0 |00\rangle + a_0 b_1 |01\rangle + a_1 b_0 |10\rangle + a_1 b_1 |11\rangle\,, && (1.14) \end{aligned}$$

where the last line is a common convention to compress notation.

This simple calculation has been performed to show that not all bipartite pure states can be written in the form $|\psi\rangle_A \otimes |\psi\rangle_B$, for example the state $|\phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ cannot be written in the form of (1.14). Such states are nonetheless physical, they are said to be entangled and constitute one of the most peculiar feature of quantum mechanics which we will extensively exploit throughout the rest of this report.

We define hereafter the notion of separability for pure states, which is the antonymical notion of the entanglement.

Definition 1.3 (**Separability for pure states**). *A bipartite pure state is separable if and only if it can be written in the form*

$$|\psi\rangle_{AB} = |\psi\rangle_A \otimes |\psi\rangle_B\,, \tag{1.15}$$

otherwise it is entangled.

Separable bipartite pure states are also called product states because they are written as a tensor product between the state of each party.

1.2.2 Bipartite mixed state

Suppose that Alice's qubit is described by the density operator $\hat{\rho}^A$ and Bob's qubit by $\hat{\rho}^B$. Therefore, the total system can be written

$$\hat{\rho}^{AB} = \hat{\rho}^A \otimes \hat{\rho}^B. \tag{1.16}$$

However, such as for pure states, this decomposition is not general for bipartite mixed states and there exist states that cannot be written in that form.

As for pure states, the particular form of (1.16) is called a product state. However, it is a special case of separable mixed states. In general, separability for mixed states can be defined as follows.

Definition 1.4 (Separability for mixed states). *A bipartite mixed state is separable if and only if it can be written*

$$\hat{\rho}^{AB} = \sum_i p_i \hat{\rho}_i^A \otimes \hat{\rho}_i^B, \qquad \textit{with } \{p_i\} \textit{ a probability distribution,} \tag{1.17}$$

otherwise it is entangled.

1.2.3 Reduced density operator

We have seen that it is not always possible to describe the total state of a bipartite system as a product state, hence we may wonder how to describe the state of Alice's or Bob's qubit. It turns out that the density operator formalism allows to do so for both pure and mixed bipartite systems by means of the so-called reduced density operator. We simply define the local state of Alice as

$$\hat{\rho}^A = \mathrm{Tr}_B(\hat{\rho}^{AB}), \tag{1.18}$$

where Tr_B is the partial trace over Bob's system defined as

$$\mathrm{Tr}_B(\hat{\rho}^{AB}) = \sum_{i_B} \langle i_B | \hat{\rho}^{AB} | i_B \rangle \,, \tag{1.19}$$

with $\{i_B\}$ any orthonormal basis of Bob's Hilbert space.

To illustrate this new tool, we take again the pure state $|\phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. The state of Alice's system is described by $\hat{\rho}^A = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$, which is a mixed state[4]! This means that, while the state of the total system is pure, locally, each subsystem[5] is described by a statistical mixture of pure states. This remarkable result emphasizes again the peculiarity of entanglement in quantum mechanics.

1.2.4 Schmidt decomposition

Now that we have described both single party and bipartite systems of qubits, we generalize the latter to qudits, which are systems living in a d-dimensional Hilbert space and whose basis can be written $\{|0\rangle, ..., |d-1\rangle\}$.

A very convenient and usual way of writing pure states of bipartite systems is the Schmidt decomposition, which can be stated as follows.

[4] It is easy to show that $\mathrm{Tr}\big((\hat{\rho}^A)^2\big) = 1/2 < 1$.

[5] Bob's system is also described by $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$.

Definition 1.5 (**Schmidt decomposition**). *Let* $|\psi\rangle$[6] *be a bipartite pure state of a* $(d \times d)$ *system shared between Alice and Bob. Then, there exists an orthonormal basis of Alice's system, written* $\{|i_A\rangle\}$*, and an orthonormal basis of Bob's system, written* $\{|i_B\rangle\}$*, such that,*

$$|\psi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i}\, |i_A\rangle\, |i_B\rangle\,, \tag{1.20}$$

with $\sum_{i=1}^{d} \lambda_i = 1$ *and* $\lambda_i \geq 0, \forall i$. *The* λ_i *are unique and are called the Schmidt coefficients (SC) of* $|\psi\rangle$.

Because the SC are unique for a given state, we often characterize the state only by stating its vector of SC. Furthermore, the number of non-zero SC is called the Schmidt rank.

Example 1:

Let us consider the state $|\phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)$, which we can rewrite

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle\,, \tag{1.21}$$

showing by the way that the state is separable because it is in the form of (1.15).

By choosing the basis $\{|+\rangle, |-\rangle\} = \{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}$ for Alice and the computational basis for Bob, the Schmidt decomposition of the state is finally given by

$$|\phi\rangle = |+\rangle\, |0\rangle\,. \tag{1.22}$$

In this case, there is only one SC, which is 1, thus the Schmidt rank of $|\phi\rangle$ equals one.

The following property relates the SC to the eigenvalues of the reduced density operators.

Property 1.1. *Let* $|\psi\rangle$ *have the Schmidt decomposition* $|\psi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i}\, |i_A\rangle\, |i_B\rangle$*, then* $\{\lambda_i\}$ *is the set of eigenvalues of the reduced density operator of both Alice and Bob.*

Proof. If $|\psi\rangle_{AB} = \sum_{i=1}^{d} \sqrt{\lambda_i}\, |i_A\rangle\, |i_B\rangle$ is a bipartite pure state, its density matrix can be written

$$\hat{\rho}^{AB} = \left(\sum_{i=1}^{d} \sqrt{\lambda_i}\, |i_A\rangle \otimes |i_B\rangle\right)\left(\sum_{j=1}^{d} \sqrt{\lambda_j}\, \langle j_A| \otimes \langle j_B|\right), \tag{1.23}$$

$$= \sum_{i=1}^{d}\sum_{j=1}^{d} \sqrt{\lambda_i \lambda_j}\, |i_A\rangle\langle j_A| \otimes |i_B\rangle\langle j_B|\,. \tag{1.24}$$

[6]Because this report will always consider bipartite states, we will omit the subscript AB from now on.

By tracing out Bob's system, the partial system of Alice can be written

$$\hat{\rho}^A = \mathrm{Tr}_B(\hat{\rho}^{AB}), \tag{1.25}$$

$$= \sum_{i=1}^{d}\sum_{j=1}^{d} \sqrt{\lambda_i\lambda_j}\, |i_A\rangle\langle j_A| \underbrace{\mathrm{Tr}(|i_B\rangle\langle j_B|)}_{\delta_{ij}}, \tag{1.26}$$

$$= \sum_{i=1}^{d} \lambda_i |i_A\rangle\langle i_A|, \tag{1.27}$$

and the same for the partial system of Bob

$$\hat{\rho}^B = \mathrm{Tr}_A(\hat{\rho}^{AB}), \tag{1.28}$$

$$= \sum_{i=1}^{d}\sum_{j=1}^{d} \sqrt{\lambda_i\lambda_j}\, \underbrace{\mathrm{Tr}(|i_A\rangle\langle j_A|)}_{\delta_{ij}} |i_B\rangle\langle j_B|, \tag{1.29}$$

$$= \sum_{i=1}^{d} \lambda_i |i_B\rangle\langle i_B|, \tag{1.30}$$

which ends the proof. Thus, we can observe that both partial states have same purity. □

1.2.4.a Separability

Finally, we express the separability definition for pure states using the Schmidt decomposition.

Theorem 1.2. *A bipartite pure state is separable if and only if its Schmidt rank is equal to one, otherwise it is entangled.*

This is quite obvious because states having a single non-zero SC are product states.

1.2.5 Operator Schmidt decomposition

There exists an equivalent to the Schmidt decomposition for mixed states, the so-called operator Schmidt (OS) decomposition [16, 13], which can be defined as

Definition 1.6 (Operator Schmidt decomposition). *Let ρ be a bipartite mixed state of a $(d \times d)$ system shared between Alice and Bob. Then, there exists an orthonormal basis of the operator space of Alice's system $\{\hat{A}_i\}$ (orthogonality meaning that $\mathrm{Tr}\left(\hat{A}_i^\dagger \hat{A}_{i'}\right) = \delta_{ii'}$), and an orthonormal basis of the operator space of Bob's system $\{\hat{B}_i\}$, such that,*

$$\hat{\rho} = \sum_{i=1}^{d^2} \lambda_i^{OS} \hat{A}_i \otimes \hat{B}_i, \tag{1.31}$$

with $\lambda_i^{OS} \geq 0, \forall i$. The λ_i^{OS} are unique and are called the operator Schmidt coefficients (OSC) of $\hat{\rho}$.

Notice a major difference with the pure state case because now there is no condition on $\sum_{i=1}^{d^2} \lambda_i^{OS}$. As for pure states, the vector of OSC completely characterizes the state.

1.2.5.a Operator Schmidt decomposition of a pure state

Let $|\psi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i} |i_A\rangle |i_B\rangle$ be a bipartite pure state. As before, we write its density matrix

$$\hat{\rho} = \sum_{i=1}^{d} \sum_{j=1}^{d} \sqrt{\lambda_i \lambda_j} |i_A\rangle \langle j_A| \otimes |i_B\rangle \langle j_B| . \tag{1.32}$$

This expression can be compressed using the following substitutions $\lambda_{ij}^{OS} = \sqrt{\lambda_i \lambda_j}$, $\hat{A}_{ij} = |i_A\rangle \langle j_A|$ and $\hat{B}_{ij} = |i_B\rangle \langle j_B|$ that give

$$\hat{\rho} = \sum_{i=1}^{d} \sum_{j=1}^{d} \lambda_{ij}^{OS} \hat{A}_{ij} \otimes \hat{B}_{ij}. \tag{1.33}$$

Note that $\hat{A}_{ij}$ satisfies $\mathrm{Tr}\left(\hat{A}_{ij}\right) = \delta_{ij}$ and $\mathrm{Tr}\left(\hat{A}_{ij}^\dagger \hat{A}_{i'j'}\right) = \delta_{ii'}\delta_{jj'}$, the same for $\hat{B}_{ij}$.

Now, rather than summing on i from 1 to d and on j from 1 to d, we sum on k from 1 to d^2 and obtain, by labelling $k = (i, j)$,

$$\hat{\rho} = \sum_{k=1}^{d^2} \lambda_k^{OS} \hat{A}_k \otimes \hat{B}_k. \tag{1.34}$$

By changing the indices, we have made implicit the relation $\mathrm{Tr}\left(\hat{A}_{ij}\right) = \mathrm{Tr}\left(\hat{B}_{ij}\right) = \delta_{ij}$. However, the relation $\mathrm{Tr}\left(\hat{A}_{ij}^\dagger \hat{A}_{i'j'}\right) = \mathrm{Tr}\left(\hat{B}_{ij}^\dagger \hat{B}_{i'j'}\right) = \delta_{ii'}\delta_{jj'}$ now becomes $\mathrm{Tr}\left(\hat{A}_k^\dagger \hat{A}_k'\right) = \mathrm{Tr}\left(\hat{B}_k^\dagger \hat{B}_k'\right) = \delta_{kk'}$, which ensures that (1.34) is a proper OS decomposition for pure states.

For a general bipartite pure state of a d-dimensional system, its vector of OSC can be written in terms of the SC as

$$(\lambda_1, \sqrt{\lambda_1\lambda_2}, \sqrt{\lambda_2\lambda_1}, \lambda_2, ..., \sqrt{\lambda_i\lambda_j}, ..., \lambda_d), \text{ where } 1 \leq i \leq d \text{ and } 1 \leq j \leq d. \tag{1.35}$$

Example 2:

Let us consider the following state $|\phi\rangle = \frac{1}{\sqrt{3}}(\sqrt{2}\,|00\rangle + |11\rangle)$, its vector of SC is $(\lambda_1, \lambda_2) = (2/3, 1/3)$. Using (1.32), its density matrix can be expressed as

$$\hat{\rho} = \frac{2}{3} |00\rangle \langle 00| + \frac{\sqrt{2}}{3} |00\rangle \langle 11| + \frac{\sqrt{2}}{3} |11\rangle \langle 00| + \frac{1}{3} |11\rangle \langle 11| , \tag{1.36}$$

$$= \frac{2}{3}\hat{A}_1 \otimes \hat{B}_1 + \frac{\sqrt{2}}{3}\hat{A}_2 \otimes \hat{B}_2 + \frac{\sqrt{2}}{3}\hat{A}_3 \otimes \hat{B}_3 + \frac{1}{3}\hat{A}_4 \otimes \hat{B}_4. \tag{1.37}$$

Its vector of OSC is

$$\left(\lambda_1, \sqrt{\lambda_1\lambda_2}, \sqrt{\lambda_1\lambda_2}, \lambda_2\right) = \left(\frac{2}{3}, \frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{3}, \frac{1}{3}\right). \tag{1.38}$$

We can verify that the state is pure by calculating its purity

$$\mathrm{Tr}\big(\hat{\rho}^2\big) = \left(\frac{2}{3}\right)^2 + \left(\frac{\sqrt{2}}{3}\right)^2 + \left(\frac{\sqrt{2}}{3}\right)^2 + \left(\frac{1}{3}\right)^2, \tag{1.39}$$

$$= \frac{4}{9} + \frac{2}{9} + \frac{2}{9} + \frac{1}{9} = 1. \tag{1.40}$$

1.2.5.b Separability criterion

Finally, although rather weak, the following simple criterion allows to get closer to a separability criterion using the OS decomposition [13].

Theorem 1.3. *If the sum of the OSC of a state is strictly greater than one, then the state is entangled.*

As can be observed from the general form of the vector of OSC for a pure state (1.35), we have

$$\sum_{i=1}^{d^2} \lambda_i^{OS} \geq \sum_{i=1}^{d} \lambda_i = 1, \tag{1.41}$$

with $\sum_{i=1}^{d^2} \lambda_i^{OS} > 1$ if and only if the Schmidt rank of the pure state is strictly greater than one. This allows to make the connection with Theorem 1.2.

1.3 Measurements

One of the major differences between quantum mechanics and classical physics lies in the measurement of a system. While classically the measure of a system doesn't affect its state, quantum mechanically, the state is generally modified after the measurement. Furthermore, in classical physics, measurements give deterministic results, in contrast with quantum mechanics where random results are generally obtained[7].

In order to describe measurements in quantum mechanics, we make use of operators that are applied to the quantum state of the system we want to measure. Let $\{\hat{M}_m\}$ be a set of measurement operators, where the index m is a label of the measurement outcome, and $|\psi\rangle$ a pure state being measured. The probability to obtain outcome m is given by

$$p(m) = \langle\psi|\hat{M}_m^\dagger \hat{M}_m|\psi\rangle\,, \tag{1.42}$$

and the system is left in the state

$$|\psi_m\rangle = \frac{\hat{M}_m\,|\psi\rangle}{\sqrt{p(m)}}, \tag{1.43}$$

where the denominator is necessary for the state to be normalized.

The only constraint a measurement set needs to obey is the following completeness relation

$$\sum_m \hat{M}_m^\dagger \hat{M}_m = \hat{I}, \tag{1.44}$$

[7]In the sense that if several copies of the same state are measured, different results can generally be obtained.

where $\hat{I}$ is the identity operator. This relation essentially reflects that probabilities in (1.42) sum to one.

Now, if the measured state $\hat{\rho}$ is a mixed state, the probability to obtain outcome m is given by

$$p(m) = \mathrm{Tr}\Big(\hat{M}_m^\dagger \hat{M}_m \hat{\rho}\Big), \tag{1.45}$$

and the system is left in the state

$$\hat{\rho}_m = \frac{\hat{M}_m \hat{\rho} \hat{M}_m^\dagger}{p(m)}, \tag{1.46}$$

where the denominator is also used for normalization.

However, if we imagine for instance that we lost track of the result of the measurement, we can view the system as in a statistical mixture of the possible post-measurement states $\hat{\rho}_m$, with corresponding probabilities $p(m)$, i.e.,

$$\hat{\rho} = \sum_m p(m)\hat{\rho}_m, \tag{1.47}$$

$$= \sum_m \hat{M}_m \hat{\rho} \hat{M}_m^\dagger. \tag{1.48}$$

The choice of one or the other post-measurement mixed state of the system depends on the application.

1.3.1 Projective measurements

Projective measurements correspond to a special class of measurements. As their name suggests, the operators are now projectors. An operator $\hat{P}_m$ is a projector if and only if it is both idem-potent ($\hat{P}_m = \hat{P}_m^2$) and hermitian ($\hat{P}_m = \hat{P}_m^\dagger$). Therefore, all statements of the previous section can be simplified by making use of these properties.

For example, the probability to obtain outcome m when measuring a pure state $|\psi\rangle$ (1.42) using the set of projectors $\{\hat{P}_m\}$ is now $p(m) = \langle\psi|\hat{P}_m|\psi\rangle$, with the completeness relation (1.44) that becomes $\sum_m \hat{P}_m = \hat{I}$.

Example 3:

Let us consider an arbitrary pure state of a qubit $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$ (1.2) and a set of projective measurements onto the computational basis, i.e., $\{P_0 = |0\rangle\langle 0|, P_1 = |1\rangle\langle 1|\}$. With the restatement of the probability of outcome (1.42) for projective measurements, we get the following probabilities of measuring 0 or 1,

$$p(0) = \langle\psi|\hat{P}_0|\psi\rangle = \Big(a_0^*\langle 0| + a_1^*\langle 1|\Big)|0\rangle\langle 0|\Big(a_0|0\rangle + a_1|1\rangle\Big) = a_0^* a_0 = |a_0|^2, \tag{1.49}$$

$$p(1) = \langle\psi|\hat{P}_1|\psi\rangle = \Big(a_0^*\langle 0| + a_1^*\langle 1|\Big)|1\rangle\langle 1|\Big(a_0|0\rangle + a_1|1\rangle\Big) = a_1^* a_1 = |a_1|^2. \tag{1.50}$$

Because we measure either 0 or 1, $p(0) + p(1) = 1$. This explains why the state must be normalized ($|a_0|^2 + |a_1|^2 = 1$).

1.3.2 POVM measurements

When the post-measurement state is not useful for a certain application, the general measurement formalism seen above can be simplified. This is the purpose of the POVM[8] formalism.

In this picture, we simply define

$$\hat{E}_m = \hat{M}_m^\dagger \hat{M}_m, \tag{1.51}$$

and say that the $\{\hat{E}_m\}$ constitute a set of POVM elements.

From the definition of $\hat{E}_m$, it is obvious that the probability to obtain the outcome m is simplified to $p(m) = \langle\psi|\hat{E}_m|\psi\rangle$ for pure states (1.42) and to $p(m) = \mathrm{Tr}\left(\hat{E}_m\hat{\rho}\right)$ for mixed states (1.45). However, the resulting state ((1.43) and (1.46)) cannot be re-expressed in terms of the $\hat{E}_m$ only, hence the usefulness of the formalism when the post-measurement state is not needed.

[8] Positive Operator-Valued Measure.

Chapter 2

Theory of majorization

The theory of majorization is a well known mathematical theory, of which [7] provides a thorough description. In the field of quantum information sciences, majorization is used for numerous different applications, e.g., in the study of bosonic systems [17, 18], separability criteria [19], uncertainty relations [20] or even in the field of quantum thermodynamics [21]. In our case, we will study majorization in regards to the already broad field of quantum entanglement transformations.

In Section 2.1, we introduce the (strong) majorization[1] framework, which allows to compare vectors of the same absolute norm, hence probability distributions. Then, in Section 2.2, we present an extended framework allowing to compare non-normalized distributions, namely weak majorization.

In the present report, we only focus on discrete distributions. However, the theory can be extended to continuous distributions [7, 22].

2.1 Strong majorization relations

As a first introduction to majorization, we choose to present it in an intuitive approach. The following example comes from the field of econometrics, emphasizing that this tool applies to numerous domains.

Consider 2 given income distributions $\boldsymbol{x} = (\frac{1}{d}, ..., \frac{1}{d})$ and $\boldsymbol{y} = (\frac{1}{2}, \frac{1}{4}, ..., (\frac{1}{2})^d)$ over a certain population of size d, majorization relations give an order on the most even (or spread out) distribution between $\boldsymbol{x}$ and $\boldsymbol{y}$, which is obviously $\boldsymbol{x}$. Hence, we say that $\boldsymbol{x}$ is majorized by $\boldsymbol{y}$, written $\boldsymbol{x} \prec \boldsymbol{y}$. We can observe that the left-hand side is a "more even, or disordered, distribution" than the right-hand side. Therefore, majorization relations are often used to quantify disorder.

2.1.1 Definition with inequalities

There exist many ways to define majorization. One of the most intuitive is by using a set of inequalities, hence we will mainly use it in this report and begin this chapter by introducing it.

[1] Often simply called majorization.

2.1.1.a Cumulative sums

The definition of majorization relations involves reordered probability vectors, either in decreasing or increasing order. Hence, we define the cumulative sums, one for each of the two choices of ordering, as follows.

Definition 2.1 (**Cumulative sums**). *The decreasing cumulative sum $S_k^{\downarrow}(\boldsymbol{x})$ of a vector $\boldsymbol{x} \in \mathbb{R}^d$ represents the sum of the k largest components of $\boldsymbol{x}$, i.e.,*

$$S_k^{\downarrow}(\boldsymbol{x}) = \sum_{i=1}^{k} x_i^{\downarrow}, \tag{2.1}$$

where $x_i^{\downarrow}$ indicates the i^{th} largest component of $\boldsymbol{x}$.

The increasing cumulative sum $S_k^{\uparrow}(\boldsymbol{x})$ of a vector $\boldsymbol{x} \in \mathbb{R}^d$ represents the sum of the k smallest components of $\boldsymbol{x}$, i.e.,

$$S_k^{\uparrow}(\boldsymbol{x}) = \sum_{i=1}^{k} x_i^{\uparrow}, \tag{2.2}$$

where $x_i^{\uparrow}$ indicates the i^{th} smallest component of $\boldsymbol{x}$.

In other words, in the following, we will always consider reordered vectors. When computing $S_k^{\downarrow}(\boldsymbol{x})$, the vector's components are reordered in decreasing order, whereas for $S_k^{\uparrow}(\boldsymbol{x})$, the vector's components are reordered in increasing order.

2.1.1.b Majorization inequalities

The definitions of cumulative sums allow us to define majorization inequalities [7] in a compact way.

Definition 2.2 (**Majorization**). *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two d-dimensional[2] vectors. We say that $\boldsymbol{x}$ is majorized by $\boldsymbol{y}$, written $\boldsymbol{x} \prec \boldsymbol{y}$, if and only if :*

$$\begin{cases} S_k^{\downarrow}(\boldsymbol{x}) \leq S_k^{\downarrow}(\boldsymbol{y}) & \text{for } k = 1, ..., d-1, \\ \\ S_d^{\downarrow}(\boldsymbol{x}) = S_d^{\downarrow}(\boldsymbol{y}). \end{cases} \tag{2.3}$$

First, remark that the equality of the last relation means that the sum of all components of $\boldsymbol{x}$ is equal to the sum of all components of $\boldsymbol{y}$, which is obviously necessary when comparing probability distributions. Second, because we will always compare probability distributions in later chapters, all vector's components will be considered non-negative from now on.

Even though majorization relations are usually expressed with the vector's components in decreasing order, it is equivalent to write them with the vector's components in increasing order as follows.

[2] If the two vectors don't have the same dimension, we simply append zeros at the end of the one with the smaller dimension.

Definition 2.3 (Majorization). *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two d-dimensional vectors. We say that $\boldsymbol{x}$ is majorized by $\boldsymbol{y}$, written $\boldsymbol{x} \prec \boldsymbol{y}$, if and only if :*

$$\begin{cases} S_k^{\uparrow}(\boldsymbol{x}) \geq S_k^{\uparrow}(\boldsymbol{y}) & \textit{for } k = 1, ..., d-1, \\ \\ S_d^{\uparrow}(\boldsymbol{x}) = S_d^{\uparrow}(\boldsymbol{y}). \end{cases} \tag{2.4}$$

These last inequalities are evidently complementary to the ones of (2.3). It is also meaningful to note that the number of non-zero components of $\boldsymbol{x}$ should always be greater or equal to the number of non-zero components of $\boldsymbol{y}$ to satisfy the above relations. Indeed, otherwise the first majorization relation $S_1^{\uparrow}(\boldsymbol{x}) \geq S_1^{\uparrow}(\boldsymbol{y})$ would already be unsatisfied because $0 \ngeq y_1$, if $y_1 \neq 0$. The importance of this remark will become clearer when applying majorization to entanglement transformations of states in Chapter 3.

In the rest of this report, we will always arrange vectors in decreasing order, hence use (2.3) to unfold majorization relations.

Remark that for any probability distribution $\boldsymbol{x}$ of size d, we have

$$\left(\frac{1}{d}, ..., \frac{1}{d}\right) \prec \boldsymbol{x} \prec (1, 0, ..., 0), \tag{2.5}$$

which can be intuitively understood because $\left(\frac{1}{d}, ..., \frac{1}{d}\right)$ is the most disordered distribution and $(1, 0, ..., 0)$ the most ordered one.

2.1.1.c Lorenz curves

In order to visualize the majorization inequalities of (2.3), we present a convenient graphical representation of the cumulative sums, the so-called Lorenz curves [7, 23]. Such curves are obtained by linear interpolation between the points of the set $\left\{\left(k, S_k^{\downarrow}(\boldsymbol{x})\right)\right\}_{k=0}^{d}$ (or $\left\{\left(k, S_k^{\uparrow}(\boldsymbol{x})\right)\right\}_{k=0}^{d}$), with the convention that $S_0^{\downarrow}(\boldsymbol{x}) = 0$, and they are denoted $L_{\boldsymbol{x}}^{\downarrow}(\omega)$ (or $L_{\boldsymbol{x}}^{\uparrow}(\omega)$), where $\omega \in [0, d]$.

For example, Figure 2.1 shows the Lorenz curves $L_{\boldsymbol{x}}^{\downarrow}(\omega)$ and $L_{\boldsymbol{y}}^{\downarrow}(\omega)$ of the two extreme distributions $\boldsymbol{x} = \left(\frac{1}{d}, ..., \frac{1}{d}\right)$ and $\boldsymbol{y} = (1, 0, ..., 0)$ for $d = 5$. We see that the most disordered distribution corresponds to a straight line and the most ordered one almost to a step.

Note already that $L_{\boldsymbol{x}}^{\downarrow}(\omega)$ corresponds to a concave polygonal curve, whereas $L_{\boldsymbol{x}}^{\uparrow}(\omega)$ corresponds to a convex polygonal curve[3]. In the following, we will omit the term "polygonal" and say that Lorenz curves $L_{\boldsymbol{x}}^{\downarrow}(\omega)$ are concave curves.

2.1.1.d Examples

In this section, we take the time to showcase examples of majorization relations between probability distributions.

As a first example, we present the typical case of a majorization relation between two distributions[4], as well as their Lorenz curves.

[3]See Section 2.1.3 of this chapter for definitions of convex and concave functions.

[4]Example 5 shows that it is not always the case.

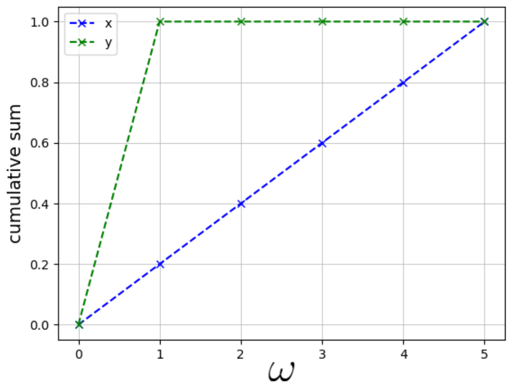


Figure 2.1: Lorenz curves $L^{\downarrow}_{\boldsymbol{x}}(\omega)$ and $L^{\downarrow}_{\boldsymbol{y}}(\omega)$ associated to $\boldsymbol{x} = (1/5,\ 1/5,\ 1/5,\ 1/5,\ 1/5)$ and $\boldsymbol{y} = (1,\ 0,\ 0,\ 0,\ 0)$. Any other distribution possesses a Lorenz curve in between those two curves. Note that the vertical axis is called "cumulative sum" to recall that integer values of ω correspond to values of the (in this case decreasing) cumulative sums.

Example 4:

Let $\boldsymbol{x} = (0.5, 0.3, 0.2)$ and $\boldsymbol{y} = (0.7, 0.2, 0.1)$, it is easy to show that $\boldsymbol{x} \prec \boldsymbol{y}$:

$$\begin{cases} 0.5 \leq 0.7, \\ 0.5 + 0.3 \leq 0.7 + 0.2, \\ 0.5 + 0.3 + 0.2 = 0.7 + 0.2 + 0.1. \end{cases}$$

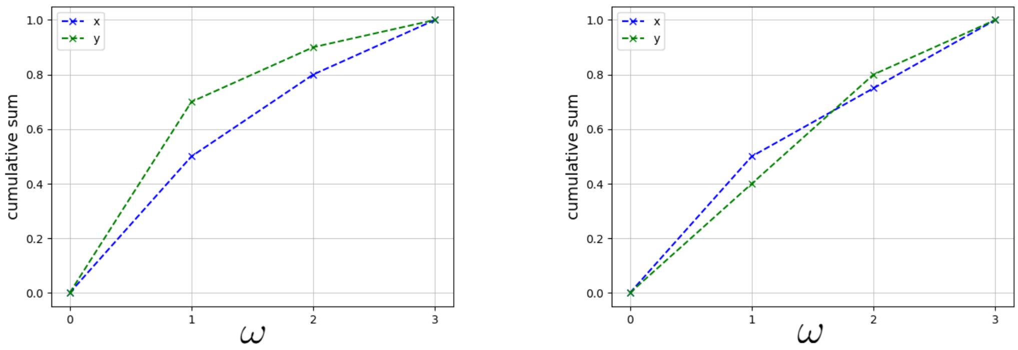
Moreover, in order to understand these relations better, we make use of the Lorenz curves. As can be seen on Figure 2.2a, $\boldsymbol{x}$ is majorized by $\boldsymbol{y}$ means that $L^{\downarrow}_{\boldsymbol{x}}(\omega)$ is always below $L^{\downarrow}_{\boldsymbol{y}}(\omega)$. They can overlap[a], as long as $L^{\downarrow}_{\boldsymbol{x}}(\omega)$ is never strictly above $L^{\downarrow}_{\boldsymbol{y}}(\omega)$.

[a] Which they must at $\omega = d$ because they reach the same value (see last relation of (2.3)).

It is important to stress that majorization relations only define a partial order on reordered probability distributions, hence some pairs of distributions are incomparable under majorization. A simple example of this remark is given by the next example.

Example 5:

Let $\boldsymbol{x} = (0.5, 0.25, 0.25)$ and $\boldsymbol{y} = (0.4, 0.4, 0.2)$. It is easy to show that $\boldsymbol{x}$ and $\boldsymbol{y}$ are incomparable under majorization. In terms of Lorenz curves, it means that the two curves intersect as can be seen on Figure 2.2b.



(a) Lorenz curves associated to $\boldsymbol{x} = (0.5, 0.3, 0.2)$ and $\boldsymbol{y} = (0.7, 0.2, 0.1)$, $\boldsymbol{x} \prec \boldsymbol{y}$ means that the graph of $L^{\downarrow}_{\boldsymbol{x}}(\omega)$ is below the graph of $L^{\downarrow}_{\boldsymbol{y}}(\omega)$, except at $\omega = 3$ where they both equal 1 because of the equality of their absolute norms.

(b) Lorenz curves associated to $\boldsymbol{x} = (0.5, 0.25, 0.25)$ and $\boldsymbol{y} = (0.4, 0.4, 0.2)$, $\boldsymbol{x} \not\prec \boldsymbol{y}$ and $\boldsymbol{x} \not\succ \boldsymbol{y}$ means that the graph of $L^{\downarrow}_{\boldsymbol{x}}(\omega)$ and the graph of $L^{\downarrow}_{\boldsymbol{y}}(\omega)$ intersect.

Figure 2.2: Lorenz curves for Examples 4 and 5.

2.1.2 Definition with matrices

For some applications, majorization relations as stated in (2.2) and (2.3) can be somewhat heavy. We present here another definition based on matrices rather than inequalities.

We first define the notion of a permutation matrix, which essentially permutes the components of a vector.

Definition 2.4 (Permutation matrix). *A square matrix $P \in \mathbb{R}^{d\times d}$ is a permutation matrix if and only if it contains exactly one 1 per line and exactly one 1 per column, all other coefficients being equal to 0.*

Now, the central notion needed when defining majorization with matrices is a more general family of matrices called doubly stochastic matrices [7], which can be defined as follows

Definition 2.5 (Doubly stochastic matrix). *A square matrix $D \in \mathbb{R}^{d\times d}$ is doubly stochastic (or bistochastic) if and only if all its entries are non-negative, each row sums to 1 and each column sums to 1, i.e.,*

$$D_{ij} \geq 0, \quad \sum_{k=1}^{d} D_{kj} = \sum_{l=1}^{d} D_{il} = 1, \quad \text{where } 1 \leq i, j \leq d. \tag{2.6}$$

Permutation matrices and doubly stochastic matrices are linked by the following theorem [7].

Theorem 2.1. *A square matrix $D \in \mathbb{R}^{d\times d}$ is doubly stochastic if and only if D can be written as a convex combination of permutation matrices P_j, i.e.,*

$$D = \sum_{j} p_j P_j, \tag{2.7}$$

with $\{p_j\}$ a probability distribution.

In its core, Theorem 2.1 expresses the fact that permutation matrices are the extremal points of a convex set whose internal points are doubly stochastic matrices.

We finally make the link with majorization by means of the following simple theorem [7].

Theorem 2.2. *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two d-dimensional vectors, $\boldsymbol{x} \prec \boldsymbol{y}$ if and only if there exists a doubly stochastic matrix $D \in \mathbb{R}^{d\times d}$ such that $\boldsymbol{x} = D\boldsymbol{y}$.*

This theorem, with the help of Theorem 2.1, means that $\boldsymbol{x} \prec \boldsymbol{y}$ is equivalent to saying that $\boldsymbol{x}$ can be obtained by stochastically[5] permuting components of $\boldsymbol{y}$.

Finally, we close this section by presenting an example of application of Theorem 2.2.

Example 6:

Let $\boldsymbol{x} = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ and $\boldsymbol{y} = (\frac{3}{4}, \frac{1}{8}, \frac{1}{8})$. We can easily check that $\boldsymbol{x} \prec \boldsymbol{y}$ using (2.3). Hence, by means of Theorem 2.2, there exists a doubly stochastic matrix D such that $\boldsymbol{x} = D\boldsymbol{y}$. For example,

$$\boldsymbol{x} = D\boldsymbol{y} \Leftrightarrow \begin{pmatrix} 1/2 \\ 1/4 \\ 1/4 \end{pmatrix} = \begin{pmatrix} 3/5 & -3/5 & 1 \\ 1/5 & 4/5 & 0 \\ 1/5 & 4/5 & 0 \end{pmatrix} \begin{pmatrix} 3/4 \\ 1/8 \\ 1/8 \end{pmatrix}. \tag{2.8}$$

However, in general, this matrix is not unique. In our case, another doubly stochastic candidate could be the following matrix

$$D^{'} = \begin{pmatrix} 3/5 & 2/5 & 0 \\ 1/5 & -1/5 & 1 \\ 1/5 & 4/5 & 0 \end{pmatrix}. \tag{2.9}$$

2.1.3 Majorization order preserving functions

In this section, we present classes of functions that preserve the (partial) order defined by majorization relations, namely Schur-convex and Schur-concave functions.

We give first the definition of a convex set [24].

Definition 2.6 (Convex set). *A subset $\mathcal{A} \subseteq \mathbb{R}^d$ is said to be convex if and only if for all $x, y \in \mathcal{A}$ and for all $t \in [0,1]$, we have*

$$tx + (1-t)y \in \mathcal{A}, \tag{2.10}$$

which means that a straight line linking any x and y of $\mathcal{A}$ is entirely contained in $\mathcal{A}$. We can now define the notion of convex (and concave) functions [7].

Definition 2.7 (Convex function). *Let $\mathcal{A} \subseteq \mathbb{R}^d$ be a convex set. A function $f : \mathcal{A} \to \mathbb{R}$ is said to be convex if and only if for all $x, y \in \mathcal{A}$ and for all $t \in [0,1]$, we have*

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y). \tag{2.11}$$

[5] In the sense that each permutation P_i acts with a certain weight p_i.

On the contrary, a function f is said to be concave on $\mathcal{A}$ if and only if $(-f)$ is convex on $\mathcal{A}$, hence the inequality in (2.11) would be reversed.

We are now able to give the definition of the majorization order preserving functions [7], that are more general cases of convex functions as will be seen on Lemma 2.1.

Definition 2.8 (Schur-convex function). *Let $\mathcal{A} \subseteq \mathbb{R}^d$ be a convex set. A function $\phi : \mathcal{A} \subseteq \mathbb{R}^d \to \mathbb{R}$ is said to be Schur-convex on $\mathcal{A}$ if and only if*

$$\boldsymbol{x} \prec \boldsymbol{y} \text{ on } \mathcal{A} \Rightarrow \phi(\boldsymbol{x}) \leq \phi(\boldsymbol{y}). \tag{2.12}$$

Conversely, a function ϕ is said to be Schur-concave on $\mathcal{A}$ if and only if $(-\phi)$ is Schur-convex on $\mathcal{A}$, which means that the inequality in (2.12) would be reversed.

A notable example of a Schur-concave function is the informational notion of entropy defined by Shannon [8] as follows

$$H(\boldsymbol{p}) = -\sum_{i=1}^{d} p_i \log_2(p_i), \tag{2.13}$$

where $\boldsymbol{p}$ is a d-dimensional probability distribution.

To show that this function is indeed Schur-concave, we make use of the following lemma [10].

Lemma 2.1. *If a function $\phi : \mathcal{A} \subseteq \mathbb{R}^d \to \mathbb{R}$ is both symmetric (invariant regarding any permutation of its inputs) and convex (resp. concave), then ϕ is Schur-convex (resp. Schur-concave) on $\mathcal{A}$.*

Note that the condition is only sufficient. Indeed, there exist Schur-convex (resp. Schur-concave) functions that are not convex (resp. concave). Nevertheless, all must be symmetric. In the case of the Shannon entropy, it is obviously symmetric and concavity is proven in [25], hence it is Schur-concave.

In essence, Shannon entropy is a measure of the uncertainty content of a random variable distributed along the probability $\boldsymbol{p}$. Intuitively, the most uncertain random variable is one that is uniformly distributed, whereas the less uncertain one is distributed along $(1, 0, ..., 0)$.

Being Schur-concave, Shannon entropy fulfills

$$\boldsymbol{x} \prec \boldsymbol{y} \Rightarrow H(\boldsymbol{x}) \geq H(\boldsymbol{y}). \tag{2.14}$$

This allows to understand what we meant when we said that majorization relations give a certain measure of disorder, in the sense that "$\boldsymbol{x}$ is more disordered than $\boldsymbol{y}$" is a necessary condition of $\boldsymbol{x} \prec \boldsymbol{y}$, as (2.14) states.

The notion of Shannon entropy constitutes one the foundations of information theory and is therefore a very useful resource in quantum information theory (as we will see in Chapter 3).

2.2 Weak majorization relations

By replacing the equality condition in the majorization relation (see (2.3) and (2.4)) by an inequality, we obtain the framework of weak majorization. Those relations are split into two different weaker versions of majorization, namely sub-majorization and super-majorization. However, because we compare now vectors of different norms, it is not possible anymore to intuitively understand those relations in terms of disorder.

We will present this section with the same structure as the one on strong majorization, making links between the two easier to follow for the reader.

2.2.1 Definition with inequalities

As for strong majorization, there exist multiple ways to define weak majorization relations. However, using inequalities is again quite easy both conceptually and practically, hence it will also be the main definition used in this report.

2.2.1.a Sub-majorization

First, we present sub-majorization, also called weak majorization from below. The relation "$\boldsymbol{x}$ is sub-majorized by $\boldsymbol{y}$" holds the constraint that $\boldsymbol{x}$ has a smaller absolute norm than $\boldsymbol{y}$[6]. Formally, sub-majorization is defined as follows [7].

Definition 2.9 (Sub-majorization). *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two d-dimensional vectors. We say that $\boldsymbol{x}$ is sub-majorized by $\boldsymbol{y}$, written $\boldsymbol{x} \prec_w \boldsymbol{y}$, if and only if :*

$$S_k^{\downarrow}(\boldsymbol{x}) \leq S_k^{\downarrow}(\boldsymbol{y}) \qquad \text{for } k = 1, ..., d. \tag{2.15}$$

Let us emphasize the difference between sub-majorization and strong majorization. In fact, the last relation of (2.3), an equality, has been replaced by an inequality, hence we understand why sub-majorization relations are weaker than strong majorization relations.

2.2.1.b Super-majorization

Second, we present super-majorization, also called weak majorization from above. The relation "$\boldsymbol{x}$ is super-majorized by $\boldsymbol{y}$" holds this time the constraint that $\boldsymbol{x}$ has a greater absolute norm than $\boldsymbol{y}$[7]. Formally, super-majorization is defined as follows [7].

Definition 2.10 (Super-majorization). *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two d-dimensional vectors. We say that $\boldsymbol{x}$ is super-majorized by $\boldsymbol{y}$, written $\boldsymbol{x} \prec^w \boldsymbol{y}$, if and only if :*

$$S_k^{\uparrow}(\boldsymbol{x}) \geq S_k^{\uparrow}(\boldsymbol{y}) \qquad \text{for } k = 1, ..., d. \tag{2.16}$$

Similarly as for sub-majorization, the equality of (2.4) has been replaced by an inequality, allowing us to understand why super-majorization is weaker than strong majorization. However, note that this time, contrary to sub-majorization that uses decreasing cumulative sums, we make use of increasing cumulative sums in the definition.

[6] Which explains why we call it "weak majorization from below".
[7] Which explains why we call it "weak majorization from above".

Let us now state a few remarks. First, if both $\boldsymbol{x} \prec_w \boldsymbol{y}$ and $\boldsymbol{x} \prec^w \boldsymbol{y}$ relations hold, we recover the strong majorization relation $\boldsymbol{x} \prec \boldsymbol{y}$. This can be understood by looking at the d-th inequality of (2.15) and (2.16). Indeed, if we have both $S_d^{\downarrow}(\boldsymbol{x}) \leq S_d^{\downarrow}(\boldsymbol{y})$ and $S_d^{\uparrow}(\boldsymbol{x}) \geq S_d^{\uparrow}(\boldsymbol{y})$, then it must be that $S_d(\boldsymbol{x}) = S_d(\boldsymbol{y})$[8]. Second, even if both vectors under comparison must not have the same absolute norm anymore, two vectors can be incomparable in terms of sub-majorization and/or in terms of super-majorization.

2.2.1.c Examples

In this section, we showcase an example of both sub- and super-majorization relations as well as their intuitive graphical approach by means of the Lorenz curves already presented in Section 2.1.1.c.

Example 7:

Let $\boldsymbol{x} = (0.6, 0.5, 0.3, 0.1)$ and $\boldsymbol{y} = (0.8, 0.6, 0.3, 0.3)$. With (2.15), we show easily that $\boldsymbol{x} \prec_w \boldsymbol{y}$:

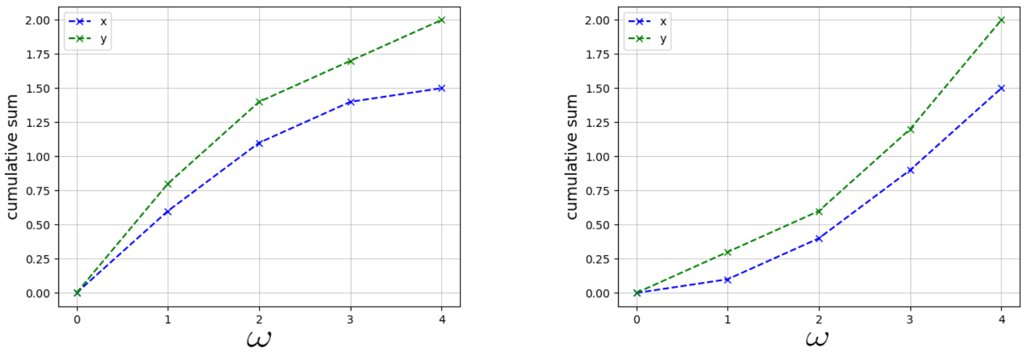
$$\begin{cases} 0.6 \leq 0.8, \\ 0.6 + 0.5 \leq 0.8 + 0.6, \\ 0.6 + 0.5 + 0.3 \leq 0.8 + 0.6 + 0.3, \\ 0.6 + 0.5 + 0.3 + 0.1 \leq 0.8 + 0.6 + 0.3 + 0.3. \end{cases}$$

Furthermore, we also show with (2.16) that $\boldsymbol{y} \prec^w \boldsymbol{x}$:

$$\begin{cases} 0.3 \geq 0.1, \\ 0.3 + 0.3 \geq 0.1 + 0.3, \\ 0.3 + 0.3 + 0.6 \geq 0.1 + 0.3 + 0.5, \\ 0.3 + 0.3 + 0.6 + 0.8 \geq 0.1 + 0.3 + 0.5 + 0.6. \end{cases}$$

As previously, both relations can be understood by means of Lorenz curves (see Figure 2.3). The only difference between weak and strong majorization lies in the (possibly) different absolute norms of the vectors compared. In terms of Lorenz curves, it translates into the two curves not reaching the same value at $\omega = d$ (see Figure 2.3).

[8] We have dropped the arrow sign because, when summing all vector components, the order doesn't matter anymore, hence $S_d^{\downarrow}(\boldsymbol{x}) = S_d^{\uparrow}(\boldsymbol{x}) = S_d(\boldsymbol{x})$.



(a) Lorenz curves associated to $\boldsymbol{x} = (0.6, 0.5, 0.3, 0.1)$ and $\boldsymbol{y} = (0.8, 0.6, 0.3, 0.3)$, $\boldsymbol{x} \prec_w \boldsymbol{y}$ means that the graph of $L^{\downarrow}_{\boldsymbol{x}}(\omega)$ is below the graph of $L^{\downarrow}_{\boldsymbol{y}}(\omega)$.

(b) Lorenz curves associated with $\boldsymbol{x} = (0.1, 0.3, 0.5, 0.6)$ and $\boldsymbol{y} = (0.3, 0.3, 0.6, 0.8)$, $\boldsymbol{y} \prec^w \boldsymbol{x}$ means that the graph of $L^{\uparrow}_{\boldsymbol{y}}(\omega)$ is above the graph of $L^{\uparrow}_{\boldsymbol{x}}(\omega)$.

Figure 2.3: Lorenz curves of Example 7.

2.2.2 Definition with matrices

Again, as for strong majorization, weak majorization relations can be formulated using matrices [7], hence we begin this section by defining two types of matrices that will be used to formulate sub- and super-majorization relations.

Definition 2.11 (**Doubly sub-stochastic matrix**). *A square matrix $P \in \mathbb{R}^{d\times d}$ is doubly sub-stochastic if and only if there exists a doubly stochastic matrix $D \in \mathbb{R}^{d\times d}$ such that, for all $i, j \in [1, d]$, $0 \leq P_{ij} \leq D_{ij}$.*

Definition 2.12 (**Doubly super-stochastic matrix**). *A square matrix $P \in \mathbb{R}^{d\times d}$ is doubly super-stochastic if and only if there exists a doubly stochastic matrix $D \in \mathbb{R}^{d\times d}$ such that, for all $i, j \in [1, d]$, $P_{ij} \geq D_{ij}$.*

With these definitions, it is clear that any doubly stochastic matrix D can be used to generate sub- and super-stochastic matrices, provided the matrix elements of D are multiplied by certain parameters $\alpha_{ij} \in [0, 1]$ for doubly sub-stochastic matrices or $\beta_{ij} \in [1, \infty)$ for doubly super-stochastic matrices.

Finally, in a very similar way to Theorem 2.2, the next theorem links weak majorization to sub- and super-stochastic matrices.

Theorem 2.3. *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two d-dimensional vectors, $\boldsymbol{x} \prec_w \boldsymbol{y}$ (resp. $\boldsymbol{x} \prec^w \boldsymbol{y}$) if and only if there exists a doubly sub-stochastic (resp. super-stochastic) matrix $D \in \mathbb{R}^{d\times d}$ such that $\boldsymbol{x} = D\boldsymbol{y}$.*

We choose to call this last statement and Theorem 2.2 "theorems" because the basic definitions of majorization relations in this report are stated using inequalities. However, they could very well be called "definitions" as is done in [7] for example.

Chapter 3

Entanglement transformations

In this chapter, we present one of the major applications of majorization theory to quantum information theory, namely the field of entanglement transformations between bipartite systems. For a full review of the use of entanglement in quantum information, see [26].

In Section 3.1, we study the type of state transformations allowed in all the protocols we will consider, called Local Operations and Classical Communication (LOCC) [27]. Then, in Section 3.2, we come to the central theorem of this report, which states that bipartite pure state transformations are feasible with certainty using only LOCC if and only if a specific majorization relation between their vectors of SC is obeyed [5]. Finally, we study probabilistic bipartite pure state transformations in Section 3.3.

3.1 Local Operations and Classical Communication

Quantum information communication protocols include two different cases, namely classical communication and quantum communication. By classical communication, we mean the transmission of classical bits of information between different parties. Whereas by quantum communication, we mean the transmission of qudits[1] between them. However, quantum communication is quite hard to perform in practice (notably because of the decoherence time of quantum states [28], or due to loss in the quantum channel used [29]) and classical communication, if possible, is often preferable.

In this report, we restrict ourselves to protocols involving only two parties (Alice and Bob) that perform local operations (i.e., each party acts only on its subsystem) and use classical communication [30]. We call this the class of LOCC transformations. For example, a famous protocol of quantum information involving only LOCC is the so-called quantum teleportation [3].

We use LOCC to transform the state of a bipartite system into another state, which we denote

$$|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle \,, \tag{3.1}$$

if $|\psi\rangle$ is the initial state and $|\phi\rangle$ the final state.

A general protocol to apply the transformation $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ is composed of the following cycle [27]. First, Alice performs a measurement on her qudit. Then she

[1]For the sake of generality, we will always refer to d-dimensional systems, rather than the usual qubits which are states of a 2-dimensional system.

sends classical information (generally depending on the result she obtained) to Bob who performs a measurement on his qudit and sends classical information back to Alice who performs a measurement on her qudit, etc. They can also apply unitary operations on their respective qudits at any point of the protocol. Note that the number of rounds during the protocol depends on the two states and is of order d.

However, the following result allows to show that this long two-way communication protocol can be simulated by a much more simple protocol involving only one-way communication protocol [27].

Proposition 3.1. *Let $|\psi\rangle$ and $|\phi\rangle$ be two bipartite pure states shared between Alice and Bob and suppose $|\psi\rangle \xrightarrow{LOCC} |\phi\rangle$, then the next protocol allows to apply such transformation. First, Alice performs a single measurement described by measurement operators M_j^A, then she can perform a unitary transformation described by the unitary operators U_j^A conditional on the result she obtained. Second, she sends her result (j) to Bob who can also perform a unitary transformation described by the unitary operators U_j^B conditional on the result he received.*

In other words, any measurement performed by Bob can be simulated by a measurement performed by Alice and both, eventually, applying unitary operations.

Example 8:

This example is inspired from [14]. Let us suppose that Alice and Bob initially share the state $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ which they want to transform into any state $|\phi\rangle = \sqrt{a}\,|00\rangle + \sqrt{1-a}\,|11\rangle$ with $a \geq 1/2$[a].

First, Alice performs a measurement on her qubit using the general measurement $\{\hat{M}_1, \hat{M}_2\}$ with

$$\hat{M}_1 = \sqrt{a}\,|0\rangle\langle 0| + \sqrt{1-a}\,|1\rangle\langle 1|\,, \tag{3.2}$$

$$\hat{M}_2 = \sqrt{1-a}\,|0\rangle\langle 0| + \sqrt{a}\,|1\rangle\langle 1|\,. \tag{3.3}$$

One can easily check that the completeness relation (1.44) is fulfilled, hence it constitutes a proper measurement set.

The state after the measurement is either $|\psi_1\rangle = \sqrt{a}\,|00\rangle + \sqrt{1-a}\,|11\rangle$ or $|\psi_2\rangle = \sqrt{1-a}\,|00\rangle + \sqrt{a}\,|11\rangle$ by using (1.43). Hence, if she measured $\hat{M}_1$, she applies no unitary operator[b], communicates classically her result to Bob who also applies no unitary because they already have obtained the desired state $|\phi\rangle$.

However, if the state after Alice's measurement is $|\psi_2\rangle$, Alice applies the unitary $\hat{U} = |0\rangle\langle 1| + |1\rangle\langle 0|$ (a NOT gate), hence she obtains $|\psi_2'\rangle = (\hat{U}\otimes\hat{I})\,|\psi_2\rangle = \sqrt{1-a}\,|10\rangle + \sqrt{a}\,|01\rangle$. She communicates it to Bob, who can in turn apply the same unitary on his qubit to get the desired state $|\phi\rangle = (\hat{I}\otimes\hat{U})\,|\psi_2'\rangle = \sqrt{1-a}\,|11\rangle + \sqrt{a}\,|00\rangle$.

[a]Note that this condition includes all states. Indeed, for example $|\phi_1\rangle = \sqrt{0.3}\,|00\rangle + \sqrt{0.7}\,|11\rangle$ is equivalent (up to unitary transformations) to $|\phi_2\rangle = \sqrt{0.7}\,|00\rangle + \sqrt{0.3}\,|11\rangle$ because $|\phi_1\rangle = (\hat{U}\otimes\hat{U})\,|\phi_2\rangle$, where $\hat{U} = |0\rangle\langle 1| + |1\rangle\langle 0|$.

[b]Or we could say the identity operator.

3.2 Deterministic convertibility using LOCC

3.2.1 Nielsen's theorem

In this section, we review the theorem at the origin of the motivation of our contribution. This theorem is due to Nielsen [5, 14] and links the theory of majorization to the transformation of bipartite pure states using only LOCC.

Theorem 3.1 (Nielsen). [2] *Let $|\psi\rangle$ and $|\phi\rangle$ be two bipartite pure states, with $\boldsymbol{\lambda}$ the vector of Schmidt coefficients (SC) of $|\psi\rangle$ and $\boldsymbol{\eta}$ the vector of SC of $|\phi\rangle$. Then, $|\psi\rangle$ is convertible with certainty into $|\phi\rangle$ using only LOCC if and only if the vector of SC of $|\psi\rangle$ is majorized by that of $|\phi\rangle$, i.e.,*

$$|\psi\rangle \xrightarrow{LOCC} |\phi\rangle \Longleftrightarrow \boldsymbol{\lambda} \prec \boldsymbol{\eta}. \tag{3.4}$$

This is a very strong statement because the feasibility of a LOCC transformation can be checked only by verifying a rather simple mathematical relation, the majorization relation (2.3), between the vectors of SC (Definition 1.5) of both states. We understand now why we announced, in Chapter 2, that we would only consider probability distributions for strong majorization relations. Indeed, as seen on Chapter 1, the vector of SC of a state is only made of non-negative values which sum to one.

Example 9:

This is a simple application of Example 4 to the conversion of bipartite pure states. Let us consider the Schmidt decompositions of two states $|\psi\rangle$ and $|\phi\rangle$ to be $|\psi\rangle = \sum_{i=1}^{3} \sqrt{\lambda_i}\,|i_A\rangle\,|i_B\rangle = \sqrt{0.5}\,|00\rangle + \sqrt{0.3}\,|11\rangle + \sqrt{0.2}\,|22\rangle$ and $|\phi\rangle = \sum_{i=1}^{3} \sqrt{\eta_i}\,|i_A\rangle\,|i_B\rangle = \sqrt{0.7}\,|00\rangle + \sqrt{0.2}\,|11\rangle + \sqrt{0.1}\,|22\rangle$.

By Theorem 3.1, because we have $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$, this implies that $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ is possible, i.e., there exists a protocol[a] transforming with certainty $|\psi\rangle$ into $|\phi\rangle$ using only LOCC.

[a]Such protocol is described in the next section.

In the rest of this report, unless otherwise stated, all vectors of SC are arranged in decreasing order, meaning that they are all of the form

$$\boldsymbol{v} = (v_1, v_2, ..., v_d) \text{ with } v_1 \geq v_2 \geq ... \geq v_d. \tag{3.5}$$

This is in no way restrictive because permutations (special case of unitary operations, hence allowed during LOCC transformations) can be applied in order to reorder the SC of the state.

3.2.2 Nielsen's protocol

In this section, we detail the protocol that Alice and Bob can follow in order to apply the transformation $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ (as introduced in [31]).

[2]Proof can be found in Appendix A.

Let $|\psi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i} |i_A\rangle |i_B\rangle$ be the initial bipartite pure state shared between Alice and Bob that they want to transform into $|\phi\rangle = \sum_{i=1}^{d} \sqrt{\eta_i} |i_A\rangle |i_B\rangle$. Obviously, we suppose that Nielsen's theorem (3.4) is obeyed, i.e., $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$. It can be shown [31] that, if $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$, we can write the final state

$$|\phi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i} |\psi_{iA}\rangle |i_B\rangle , \tag{3.6}$$

with $\{|\psi_{iA}\rangle\}$ a set of (possibly) non-orthogonal states of Alice's system[3].

We define the following operators on Alice's side

$$\hat{M}_j = \sum_{i=1}^{d} \omega_d^{ij} |\psi_i\rangle \langle i| , \qquad \text{with } \omega_d = \frac{e^{\frac{2\pi i}{d}}}{\sqrt{d}}. \tag{3.7}$$

To correspond to general measurement operators, the operators have to fulfill the completeness relation (1.44). We quickly show that this is the case.

$$\sum_j \hat{M}_j^\dagger \hat{M}_j = \sum_j \sum_{ii'} \omega_d^{-ij} |i\rangle \langle\psi_i| \omega_d^{i'j} |\psi_{i'}\rangle \langle i'| , \tag{3.8}$$

$$= \sum_{ii'} \underbrace{\sum_j \omega_d^{(i'-i)j}}_{\delta_{ii'}} \langle\psi_i|\psi_{i'}\rangle |i'\rangle \langle i| , \tag{3.9}$$

$$= \sum_i |i\rangle \langle i| , \tag{3.10}$$

$$= \hat{I}. \tag{3.11}$$

Hence, after measuring using the operators $\{\hat{M}_j\}$, by (1.43), the system is left in the state

$$|\psi'\rangle = (\hat{M}_j \otimes \hat{I}) |\psi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i} \omega_d^{ij} |\psi_i\rangle |i\rangle . \tag{3.12}$$

Now, to obtain the desired state (3.6), the phase ω_d^{ij} of (3.12) needs to disappear, which Alice cannot do. Hence, Bob must apply the following unitary operator $\hat{U}_j = \sum_{i=1}^{d} \omega_d^{-ij} |i\rangle \langle i|$ to his qudit, resulting finally in $|\phi\rangle$,

$$(\hat{I} \otimes \hat{U}_j) |\psi'\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i} \omega_d^{ij} |\psi_i\rangle \omega^{-ij} |i\rangle , \tag{3.13}$$

$$= \sum_{i=1}^{d} \sqrt{\lambda_i} |\psi_i\rangle |i\rangle , \tag{3.14}$$

$$= |\phi\rangle . \tag{3.15}$$

3.2.3 Resource theory of entanglement

Until now, we have not referred to the term "entanglement" of the title of this chapter. Nonetheless, we have exploited it implicitly from the beginning. Indeed, any pure

[3]We omit the subscripts A and B in the following.

state having more than one non-zero SC is entangled (1.2), hence all non-trivial state transformations involve entanglement. This gives the intuition why entanglement can be seen as a resource in quantum information. We formalise this notion of resource in the following.

Originating from economic principles, a resource theory [32] comprises two central notions, free states and free operations. The former representing the states where no resource can be extracted from, in our case separable states (i.e., states with Schmidt rank equal to one), the latter representing physical operations that do not create any resource[4], LOCC transformations in our case. Then, one defines resource monotones, which are functions that measure the amount of resource contained in a state.

It turns out that the Shannon entropy (2.13) of the vector of SC of a state corresponds to such resource monotone because it decreases during an LOCC transformation. Indeed, being a Schur-concave function (see Section 2.1.3 of Chapter 2), we can write

$$|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle \Rightarrow H(\boldsymbol{\lambda}) \geq H(\boldsymbol{\eta}). \tag{3.16}$$

This means that, for a state transformation to be possible under LOCC, it is necessary that the Shannon entropy of the vector of SC of the initial state is greater than the one of the final state.

This is in correspondence with the name "maximally-entangled states" for d-dimensional states having vectors of SC of the form $(\frac{1}{d}, ..., \frac{1}{d})$. Indeed such vectors have a maximum Shannon entropy and hence contain the most entanglement resource. Therefore, they can be transformed using LOCC into any other d-dimensional[5] state (e.g., in Example 8, the initial state is a 2-dimensional maximally-entangled state, hence why it can be transformed into any other 2-dimensional state).

3.2.4 Catalytic convertibility using LOCC

Sometimes, when two states are incomparable under majorization, meaning that Nielsen's theorem does not apply, a LOCC transformation is nonetheless possible by appealing to a so-called catalyst state [33]. We call it a catalyst because, as in chemistry, the state is not consumed during the transformation. This broader range of state transformations is called ELOCC, for entanglement-assisted LOCC, in the sense that the catalyst state constitutes an entanglement resource that has assisted the transformation.

More formally, let us suppose that there exists a catalyst $|c\rangle$ allowing the transformation of $|\psi\rangle$ into $|\phi\rangle$ to happen, i.e.,

$$|\psi\rangle \overset{\text{LOCC}}{\not\longrightarrow} |\phi\rangle \;\text{ but }\; |\psi\rangle \otimes |c\rangle \xrightarrow{\text{LOCC}} |\phi\rangle \otimes |c\rangle\,. \tag{3.17}$$

Then, we can write

$$|\psi\rangle \xrightarrow{\text{ELOCC}} |\phi\rangle\,. \tag{3.18}$$

It is important to stress that this kind of transformation can only be performed if both states $|\psi\rangle$ and $|\phi\rangle$ are incomparable.

Now, practically, imagine Alice and Bob cannot perform a certain LOCC transformation because the initial and final states are incomparable under majorization. It may happen that there exists some catalyst state they can borrow to a "bank state",

[4] They can decrease the resource.

[5] Or less than d.

each party taking one qudit of the entangled state, in order to apply the desired LOCC transformation. Then, they can return the catalyst state[6] to the "bank state" to complete the protocol.

3.3 Probabilistic convertibility using LOCC

In this section, we present an extension of Nielsen's theorem (Theorem 3.1) to probabilistic conversions between bipartite pure states.

3.3.1 Vidal's theorem

The following theorem is due to Vidal [6, 31] and allows to check if, given an initial bipartite pure state $|\psi\rangle$, a final state $|\phi\rangle$ and a certain probability of conversion p, the probabilistic transformation from $|\psi\rangle$ to $|\phi\rangle$ is possible or not, using only LOCC.

Theorem 3.2 (Vidal). *Let $|\psi\rangle$ and $|\phi\rangle$ be two bipartite pure states, with $\boldsymbol{\lambda}$ the vector of Schmidt coefficients (SC) of $|\psi\rangle$ and $\boldsymbol{\eta}$ the vector of SC of $|\phi\rangle$. Then, $|\psi\rangle$ is convertible with probability p into $|\phi\rangle$ using only LOCC if and only if $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$.*

Note that the use of a super-majorization relation can be understood from the fact that, by multiplying the second vector by $p < 1$, the two vectors do not have the same absolute norm anymore (the absolute norm of the second one being smaller).

The following theorem [6] allows us to answer the question "Given two bipartite pure states, what is the probability to convert one into the other by means of LOCC transformations only ?".

Theorem 3.3. *Let $|\psi\rangle$ and $|\phi\rangle$ be two bipartite pure states, with $\boldsymbol{\lambda}$ the vector of Schmidt coefficients (SC) of $|\psi\rangle$ and $\boldsymbol{\eta}$ the vector of SC of $|\phi\rangle$ (both are d-dimensional vectors). Then, the maximal probability with which the conversion $|\psi\rangle \xrightarrow{LOCC} |\phi\rangle$ is possible is*

$$p_{\max} = \min_{l\in[1,d]} \frac{S_l^{\uparrow}(\boldsymbol{\lambda})}{S_l^{\uparrow}(\boldsymbol{\eta})}. \tag{3.19}$$

This theorem is a direct consequence of Theorem 3.2. Indeed, let us suppose that $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$ for a maximum probability $p_{\max}$ and let us find this probability. We can develop the super-majorization relation as follows (2.16)

$$\begin{cases} S_1^{\uparrow}(\boldsymbol{\lambda}) \geq pS_1^{\uparrow}(\boldsymbol{\eta}), \\ \qquad\quad ... \\ S_l^{\uparrow}(\boldsymbol{\lambda}) \geq pS_l^{\uparrow}(\boldsymbol{\eta}), \\ \qquad\quad ... \\ S_d^{\uparrow}(\boldsymbol{\lambda}) \geq pS_d^{\uparrow}(\boldsymbol{\eta}), \end{cases}$$

which is true for all values of p obeying $p \leq \frac{S_l^{\uparrow}(\boldsymbol{\lambda})}{S_l^{\uparrow}(\boldsymbol{\eta})}, \forall l \in [1, d]$. Hence if we take the largest value such that all these relations are obeyed, we get the maximal probability $p_{\max}$ (3.19).

[6]The catalyst state only helped perform the conversion and remained untouched, hence its name.

Example 10:

Let us analyze qubits examples.

- Separable state to entangled state via LOCC

  Let us find the probability with which a separable state with a vector of SC $(1, 0)$ can be transformed into any state with a vector of SC $(a, 1 - a)$ with $a \geq 1/2$ via LOCC. In some sense, this is the complementary problem to Example 8 where we saw that maximally-entangled states could be transformed into any other state.

  Now, by making use of (3.19), we find that $p_{\max} = 0$, hence it will never be possible to achieve this transformation.

  This is in correspondence with the fact that separable states are free states (which do not contain any entanglement resource) and LOCC transformations, free operations (which can never create entanglement). Mathematically, this translates into the fact that any (possibly probabilistic) LOCC transformation is achievable only if the Schmidt rank of the initial state is greater (or equal) to the Schmidt rank of the final state.

- Less entangled state to more entangled state via LOCC

  Let us find now the probability with which a state with a vector of SC $(3/4, 1/4)$ can be transformed into a maximally-entangled state, hence with a vector of SC $(1/2, 1/2)$. Surprisingly, the transformation is possible with a maximum probability $p_{\max} = 1/2$.

  While it could seem to negate what we said about LOCC transformations that cannot increase the entanglement, it should be stressed that it is *deterministic* LOCC transformations that are concerned, whereas no restriction is put on *probabilistic* LOCC transformations.

3.3.2 Vidal's protocol

In order to understand Vidal's probabilistic protocol [6], we give the intuition behind it, while the explicit conversion strategy can be found in Appendix B. Note that the protocol presented here is shown to be optimal for probabilistic conversions.

The probabilistic transformation of a bipartite pure state $|\psi\rangle$ into any other bipartite pure state $|\phi\rangle$ goes as follows. First, Alice and Bob transform their shared state into a specific[7] intermediary state, which we will call $|\chi\rangle$, using Nielsen's theorem, i.e., $|\psi\rangle \xrightarrow{LOCC} |\chi\rangle$. Then, Alice (or Bob) performs a specific[7] two-outcome measurement outputting state $|\phi\rangle$ with probability $p_{\max}$ (3.19) and outputting another bipartite pure state, let us call it $|\xi\rangle$, with probability $(1 - p_{\max})$.

In the special case where the vector of SC of $|\psi\rangle$ is majorized by that of $|\phi\rangle$, the probability to perform the transformation is 1 and the intermediary state is already $|\phi\rangle$, completing the protocol. Hence, this probabilistic conversion scheme is a more general one than the deterministic scheme seen on Section 3.2.

[7]Its construction is detailed in Appendix B.

Part II

Results

Chapter 4

Weak majorization and deterministic convertibility

One of the objectives of this report is to explore the possibility of extending Nielsen's theorem for bipartite pure states (Theorem 3.1) to bipartite mixed states. As a first step in that direction, we have chosen to exploit the operator Schmidt (OS) decomposition (see Definition 1.6) because it can be used to describe both pure and mixed states.

In fact, this allows us to present a (conjectured) necessary condition for the convertibility of bipartite pure states now based on the OS decomposition rather than on the Schmidt decomposition (as Theorem 3.1 does). We do so because, in order for a condition to be valid for mixed states, it is first necessary to be valid for pure states.

In Section 4.1, we state the (conjectured) OS based condition. Then, the rest of the section is devoted to the proof of the condition for qubits (Section 4.1.1), the explanation of the difficulties of a complete proof (Section 4.1.2), the partial proof for qudits (Section 4.1.3). Finally, in Section 4.2, we present numerical evidences that strongly suggest the validity of the conjectured condition.

4.1 Condition based on the OS decomposition

The following conjecture links majorization over the Schmidt decomposition to weak majorization, more specifically super-majorization, over the OS decomposition for bipartite pure states.

Conjecture 4.1. *Let $|\psi\rangle$ and $|\phi\rangle$ be two bipartite pure states, with $\boldsymbol{\lambda}$ (resp. $\boldsymbol{\eta}$) the vector of Schmidt coefficients (SC) of $|\psi\rangle$ (resp. $|\phi\rangle$) and $\boldsymbol{\lambda}^{OS}$ (resp. $\boldsymbol{\eta}^{OS}$) the vector of operator Schmidt coefficients (OSC) of $|\psi\rangle\langle\psi|$ (resp. $|\phi\rangle\langle\phi|$). Then,*

$$\boldsymbol{\lambda} \prec \boldsymbol{\eta} \Rightarrow \boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}. \tag{4.1}$$

Remember that if a vector of SC $\boldsymbol{\lambda}$ has a number of non-zero components equal to d, its corresponding vector of OSC $\boldsymbol{\lambda}^{OS}$ has a number of non-zero components equal to d^2, hence this conjecture is not trivial as $\boldsymbol{\lambda} \prec \boldsymbol{\eta} \Rightarrow \boldsymbol{\lambda} \prec^{w} \boldsymbol{\eta}$ would be.

In the following, if not mentioned, the vector of SC of the initial state is denoted by $\boldsymbol{\lambda}$ while the vector of SC of the final state is denoted by $\boldsymbol{\eta}$. Furthermore, as before, all vectors are arranged in decreasing order.

First, we give two examples to understand better Conjecture 4.1, the first one shows an application of it, the second one shows that the implication sign of (4.1) cannot be an equivalence sign[1].

Example 11:

- Majorization on the SC implies super-majorization on the OSC

  Let $\boldsymbol{\lambda} = (1/3,\ 1/3,\ 1/3)$ (vector of SC of a 3-dimensional maximally-entangled state) and $\boldsymbol{\eta} = (5/10,\ 3/10,\ 2/10)$. One can easily check using (2.3)[a] that $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$.

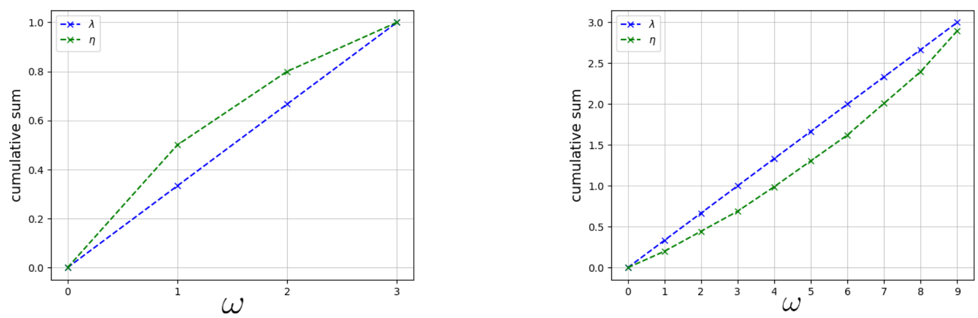
  Now, according to (1.35), we have $\boldsymbol{\lambda}^{OS} = \underbrace{(1/3, ..., 1/3)}_{9 \text{ times}}$ and $\boldsymbol{\eta}^{OS} = (5/10,\ \sqrt{15}/10,\ \sqrt{15}/10,\ \sqrt{10}/10,\ \sqrt{10}/10,\ 3/10,\ \sqrt{6}/10,\ \sqrt{6}/10,\ 2/10)$. Again, one can check, using (2.16) now, that we have $\boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$. To see this more immediately, Lorenz curves corresponding to both relations ($\boldsymbol{\lambda} \prec \boldsymbol{\eta}$ and $\boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$) are drawn in Figure 4.1.

- Super-majorization on the OSC does not imply majorization on the SC

  Let $\boldsymbol{\lambda} = (5/10,\ 3/10,\ 2/10)$ and $\boldsymbol{\eta} = (13/30,\ 13/30,\ 4/30)$. Again, by using (1.35), we can find the vector of OSC of both states and, by means of (2.3) and (2.16), it can be shown that $\boldsymbol{\lambda} \not\prec \boldsymbol{\eta}$ while $\boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$. The Lorenz curves of both relations are drawn in Figure 4.2.

  This counter-example proves that the relation $\boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$ is not sufficient to deduce $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$, hence the implication sign can only go one way in (4.1).

[a] Or just by using the fact that $\boldsymbol{\lambda}$ corresponds to a maximally-entangled state, i.e., a resource state that can be transformed into any other qudit, with $d \leq 3$.

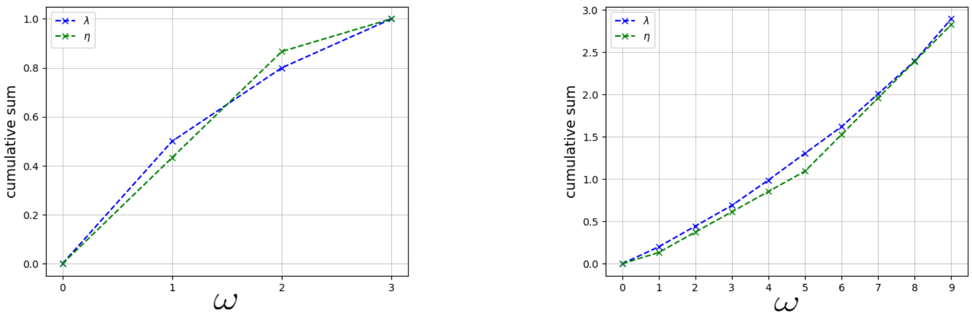


(a) Lorenz curves associated to $\boldsymbol{\lambda} = (1/3,\ 1/3,\ 1/3)$ and $\boldsymbol{\eta} = (0.5,\ 0.3,\ 0.2)$, the graph of $L^{\downarrow}_{\boldsymbol{\lambda}}(\omega)$ is below the graph of $L^{\downarrow}_{\boldsymbol{\eta}}(\omega)$, hence $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$.

(b) Lorenz curves associated to the vector of OSC of $\boldsymbol{\lambda} = (1/3,\ 1/3,\ 1/3)$ and to the vector of OSC of $\boldsymbol{\eta} = (0.5,\ 0.3,\ 0.2)$, the graph of $L^{\uparrow}_{\boldsymbol{\lambda}^{OS}}(\omega)$ is above the graph of $L^{\uparrow}_{\boldsymbol{\eta}^{OS}}(\omega)$, hence $\boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$.

Figure 4.1: Lorenz curves for the first point of Example 11.

[1] Except for qubits, see Section 4.1.1.



(a) Lorenz curves associated to $\boldsymbol{\lambda} = (0.5, 0.3, 0.2)$ and $\boldsymbol{\eta} = (13/30, 13/30, 4/30)$, the graph of $L^{\downarrow}_{\boldsymbol{\lambda}}(\omega)$ intersects the graph of $L^{\downarrow}_{\boldsymbol{\eta}}(\omega)$, hence $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ are incomparable under majorization.

(b) Lorenz curves associated to the vector of OSC of $\boldsymbol{\lambda} = (0.5, 0.3, 0.2)$ and to the vector of OSC of $\boldsymbol{\eta} = (13/30, 13/30, 4/30)$, the graph of $L^{\uparrow}_{\boldsymbol{\lambda}^{OS}}(\omega)$ is above the graph of $L^{\uparrow}_{\boldsymbol{\eta}^{OS}}(\omega)$, hence $\boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$.

Figure 4.2: Lorenz curves for the second point of Example 11.

4.1.1 Proof for qubits

In this section, we prove Conjecture 4.1 for qubits, i.e., for 2-dimensional bipartite systems. Moreover, we show that the implication goes in both direction in that case, i.e., $\boldsymbol{\lambda} \prec \boldsymbol{\eta} \Leftrightarrow \boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$.

Proof. Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ and $\boldsymbol{\eta} = (\eta_1, \eta_2)$ be such that $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$. Now, let us write the corresponding vectors of OSC as $\boldsymbol{\lambda}^{OS} = (\lambda_1, \sqrt{\lambda_1\lambda_2}, \sqrt{\lambda_1\lambda_2}, \lambda_2)$ and $\boldsymbol{\eta}^{OS} = (\eta_1, \sqrt{\eta_1\eta_2}, \sqrt{\eta_1\eta_2}, \eta_2)$ and show that we have indeed $\boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$.

To do so, let us expand both majorization relations. First, by means of (2.3), $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$ can be written

$$\begin{cases} \lambda_1 \leq \eta_1, & (4.2a) \\ \lambda_1 + \lambda_2 = \eta_1 + \eta_2. & (4.2b) \end{cases}$$

Second, using (2.16), $\boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$ can be written

$$\begin{cases} \lambda_2 \geq \eta_2, & (4.3a) \\ \lambda_2 + \sqrt{\lambda_1\lambda_2} \geq \eta_2 + \sqrt{\eta_1\eta_2}, & (4.3b) \\ \lambda_2 + 2\sqrt{\lambda_1\lambda_2} \geq \eta_2 + 2\sqrt{\eta_1\eta_2}, & (4.3c) \\ \lambda_2 + 2\sqrt{\lambda_1\lambda_2} + \lambda_1 \geq \eta_2 + 2\sqrt{\eta_1\eta_2} + \eta_1. & (4.3d) \end{cases}$$

Let us now prove that each inequality of (4.3) is implied by (4.2).

- 4.3a : This inequality is trivially verified because

$$\lambda_1 \leq \eta_1 \quad \text{and} \quad \lambda_1 + \lambda_2 = \eta_1 + \eta_2 \quad \Longleftrightarrow \quad \lambda_2 \geq \eta_2. \tag{4.4}$$

- 4.3d : We first prove the last inequality because, as we shall see, the proofs of (4.3b) and (4.3c) will therefore be straightforward.

Let us first rewrite (4.3d), by making use of the equality (4.2b), in a simpler form

$$\lambda_2 + 2\sqrt{\lambda_1\lambda_2} + \lambda_1 \geq \eta_2 + 2\sqrt{\eta_1\eta_2} + \eta_1 \Leftrightarrow \sqrt{\lambda_1\lambda_2} \geq \sqrt{\eta_1\eta_2}, \tag{4.5}$$

$$\Leftrightarrow \sqrt{\lambda_1(1-\lambda_1)} \geq \sqrt{\eta_1(1-\eta_1)}, \tag{4.6}$$

$$\Rightarrow \lambda_1(1-\lambda_1) \geq \eta_1(1-\eta_1). \tag{4.7}$$

Now, because of (4.2a) and the fact that $\lambda_1 + \lambda_2 = \eta_1 + \eta_2 = 1$ due to the definition of the Schmidt decomposition and because both $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ are arranged in decreasing order, we have $1/2 \leq \lambda_1 \leq \eta_1 \leq 1$. Hence, it suffices to show that the function $f(x) = x(1-x)$ is decreasing in the interval $[1/2, 1]$ to prove (4.7). This is indeed the case because $\frac{\mathrm{d}}{\mathrm{d}x}(f(x)) = 1 - 2x$, thus we have proved (4.3d).

- <u>(4.3b) and (4.3c)</u> : Having proved (4.3a) and the second inequality of (4.5) immediately proves both relations.

Now, to prove the converse statement, i.e., that the inequalities of (4.3) imply the relations of (4.2), we simply make use of the definition of the Schmidt decomposition (the sum of all SC equals one) to prove that $\lambda_1 + \lambda_2 = \eta_1 + \eta_2$. Then, by also making use of (4.3a), (4.2a) is immediately proved, completing the proof of this converse statement.

□

4.1.2 Difficulties encountered for qudits ($d > 2$)

One of the main challenges in proving Conjecture 4.1 for qudits (with $d > 2$) lies in the non-fixed ordering of the OSC. Indeed, for a qubit with a vector of SC $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ with the usual ordering $\lambda_1 \geq \lambda_2$, it was clear that the decreasingly ordered vector of OSC should always be written $\boldsymbol{\lambda}^{OS} = (\lambda_1, \sqrt{\lambda_1\lambda_2}, \sqrt{\lambda_1\lambda_2}, \lambda_2)$, because $\lambda_1 \geq \sqrt{\lambda_1\lambda_2} \geq \lambda_2$ then. Therefore, the order was fixed for the vector of OSC of a qubit.

However, it is not the case for $d > 2$. For example, let us consider a qutrit ($d = 3$) with a vector of SC $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ with the usual ordering $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Its decreasingly ordered vector of OSC can be written in 2 different ways, depending on the values of λ_1, λ_2 and λ_3, i.e., by means of(1.35),

- <u>if $\lambda_2 \geq \sqrt{\lambda_1\lambda_3}$</u>

$$\boldsymbol{\lambda}^{OS} = (\lambda_1, \sqrt{\lambda_1\lambda_2}, \sqrt{\lambda_1\lambda_2}, \lambda_2, \sqrt{\lambda_1\lambda_3}, \sqrt{\lambda_1\lambda_3}, \sqrt{\lambda_2\lambda_3}, \sqrt{\lambda_2\lambda_3}, \lambda_3). \tag{4.8}$$

- <u>if $\sqrt{\lambda_1\lambda_3} \geq \lambda_2$</u>

$$\boldsymbol{\lambda}^{OS} = (\lambda_1, \sqrt{\lambda_1\lambda_2}, \sqrt{\lambda_1\lambda_2}, \sqrt{\lambda_1\lambda_3}, \sqrt{\lambda_1\lambda_3}, \lambda_2, \sqrt{\lambda_2\lambda_3}, \sqrt{\lambda_2\lambda_3}, \lambda_3). \tag{4.9}$$

Furthermore, if we wanted to make a proof similar to the one of Section 4.1.1 for qutrits, it would require to prove nine ($= d^2$) super-majorization inequalities with four different possible orderings, two for the ordering of $\boldsymbol{\lambda}$ and two for the ordering of $\boldsymbol{\eta}$. We clearly see that this method is unpractical for arbitrary high values of d, hence we need to find a different approach.

4.1.2.a Majorization using matrices

A possible idea would be to use the definitions of majorization and super-majorization in terms of matrices (see Sections 2.1.2 and 2.2.2 of Chapter 2). However, although possible again for qubits, it turns out to be hard to construct systematically the matrix elements of a $d^2 \times d^2$ doubly super-stochastic matrix based on the knowledge of a $d \times d$ doubly stochastic matrix. Therefore, this direction has not been pursued further.

4.1.2.b Catalytic convertibility

Another way that could be explored is the ELOCC convertibility (see Section 3.2.4 of Chapter 3). Indeed, the idea would be that the super-majorization over the OSC could be weaker because it encompasses the pairs of states convertible under ELOCC. However, the following lemma [33] immediately forbids us to pursue in that direction.

Lemma 4.1. *Let $|\psi\rangle$ and $|\phi\rangle$ be two d-dimensional bipartite pure states, with $\boldsymbol{\lambda}$ (resp. $\boldsymbol{\eta}$) the vector of SC of $|\psi\rangle$ (resp. $|\phi\rangle$). Then, $|\psi\rangle \xrightarrow{ELOCC} |\phi\rangle$ only if both*

$$\lambda_1 \leq \eta_1 \text{ and } \lambda_d \geq \eta_d. \tag{4.10}$$

For example, the second point of Example 11 showcases a pair of states such that $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS}$ while $\boldsymbol{\lambda} \not\prec \boldsymbol{\eta}$. Nonetheless, one cannot hope to transform $|\psi\rangle$ into $|\phi\rangle$ using ELOCC because $\lambda_1 = 5/10 \geq \eta_1 = 13/30$ in that case.

4.1.3 Partial proof for qudits

In the preceding section, we have seen that the non-fixed ordering of the OSC for qudits with $d > 2$ greatly complicates the analytical treatment of the proof in terms of majorization inequalities. However, the first and the last inequality of the super-majorization relation over the OSC are always of the same form because they don't depend on the specific ordering, hence they can be more easily proved.

The first super-majorization inequality (see Definition 2.16 of Chapter 2) writes

$$\lambda_d \geq \eta_d. \tag{4.11}$$

This is obviously true because of the definition of strong majorization in increasing order (see Definition 2.4 of Chapter 2 where the first inequality for d-dimensional vectors is already simply (4.11)).

Now, the last super-majorization inequality consists in the sum of all the OSC for both vectors (1.35) and can be formulated as

$$\left(\sum_{i=1}^{d} \sqrt{\lambda_i}\right)^2 \geq \left(\sum_{i=1}^{d} \sqrt{\eta_i}\right)^2. \tag{4.12}$$

Proving this inequality is more tedious and requires the notion of Schur-concavity (see Section 2.1.3 of Chapter 2). Indeed, if we show that the function $f : \mathbb{R}^d \to \mathbb{R} : \boldsymbol{x} \mapsto \left(\sum_{i=1}^d \sqrt{x_i}\right)^2$ is Schur-concave, then, by definition of the Schur-concavity, we have

$$\boldsymbol{\lambda} \prec \boldsymbol{\eta} \Rightarrow f(\boldsymbol{\lambda}) \geq f(\boldsymbol{\eta}), \tag{4.13}$$

$$\Rightarrow \left(\sum_{i=1}^d \sqrt{\lambda_i}\right)^2 \geq \left(\sum_{i=1}^d \sqrt{\eta_i}\right)^2, \tag{4.14}$$

hence the last super-majorization inequality over the OSC is proved because $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$ is an hypothesis.

Theorem 4.1. *The function* $f : \mathbb{R}^d \to \mathbb{R} : \boldsymbol{x} \mapsto \left(\sum_{i=1}^d \sqrt{x_i}\right)^2$ *is Schur-concave.*

Proof. To prove that f is Schur-concave, we make use of Lemma 2.1 of Chapter 2, hence we need to prove that f is both symmetric and concave.

The invariance of f regarding any permutation of its input is obvious, thus we only need to prove that f is concave.

We first recall the definition of a concave function (see Definition 2.11 of Chapter 2), f is concave if and only if

$$f(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \geq tf(\boldsymbol{x}) + (1-t)f(\boldsymbol{y}), \quad \forall t \in [0,1]. \tag{4.15}$$

With our definition of f, let us show that (4.15) is true. First, expanding both terms gives

$$\left(\sqrt{tx_1 + (1-t)y_1} + ... + \sqrt{tx_d + (1-t)y_d}\right)^2 \geq t\left(\sqrt{x_1} + ... + \sqrt{x_d}\right)^2 + (1-t)\left(\sqrt{y_1} + ... + \sqrt{y_d}\right)^2. \tag{4.16}$$

By developing the squares in both terms, we obtain

$$t(x_1 + ... + x_d) + (1-t)(y_1 + ... + y_d) + 2\sum_{i=1}^d \sum_{j<i} \sqrt{(tx_i + (1-t)y_i)(tx_j + (1-t)y_j)}$$
$$\geq t(x_1 + ... + x_d) + (1-t)(y_1 + ... + y_d) + 2\sum_{i=1}^d \sum_{j<i} (t\sqrt{x_ix_j} + (1-t)\sqrt{y_iy_j}), \tag{4.17}$$

which simplifies into

$$\sum_{i=1}^d \sum_{j<i} \sqrt{(tx_i + (1-t)y_i)(tx_j + (1-t)y_j)} \geq \sum_{i=1}^d \sum_{j<i} (t\sqrt{x_ix_j} + (1-t)\sqrt{y_iy_j}). \tag{4.18}$$

Finally, by doubly distributing under the square root of the first term, we get

$$\sum_{i=1}^d \sum_{j<i} \sqrt{t^2x_ix_j + (1-t)^2y_iy_j + t(1-t)(x_iy_j + x_jy_i)} \geq \sum_{i=1}^d \sum_{j<i} (t\sqrt{x_ix_j} + (1-t)\sqrt{y_iy_j}). \tag{4.19}$$

We will show now that

$$\sqrt{t^2 x_i x_j + (1-t)^2 y_i y_j + t(1-t)(x_i y_j + x_j y_i)} \geq (t\sqrt{x_i x_j} + (1-t)\sqrt{y_i y_j}) \quad (4.20)$$

is true for all i, j, directly proving that (4.19) is true. First, we square both positive terms of (4.20)

$$t^2 x_i x_j + (1-t)^2 y_i y_j + t(1-t)(x_i y_j + x_j y_i) \geq t^2 x_i x_j + (1-t)^2 y_i y_j + 2t(1-t)\sqrt{x_i x_j y_i y_j}, \quad (4.21)$$

which simplifies into[2]

$$(x_i y_j + x_j y_i) \geq 2\sqrt{x_i x_j y_i y_j}. \quad (4.22)$$

Again, both terms being positive, we square them to obtain

$$x_i^2 y_j^2 + x_j^2 y_i^2 + 2x_i x_j y_i y_j \geq 4 x_i x_j y_i y_j, \quad (4.23)$$

giving the following remarkable identity

$$(x_i y_j - x_j y_i)^2 \geq 0, \quad (4.24)$$

which is obviously always true. Therefore, f is concave. Moreover, because of its symmetry, it is Schur-concave. □

Note that this proof has an important implication, namely that sub-majorization (see Definition 2.15 of Chapter 2) over the OSC is never[3] possible when $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$. Indeed, the last sub-majorization inequality would simply be

$$\left(\sum_{i=1}^{d} \sqrt{\lambda_i}\right)^2 \leq \left(\sum_{i=1}^{d} \sqrt{\eta_i}\right)^2, \quad (4.25)$$

which we have proved is never[3] possible if $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$ because of Theorem 4.1.

4.2 Numerical evidences

Although no analytical proof of Conjecture 4.1 has been found, numerical evidences strongly seem to indicate that the conjecture is true.

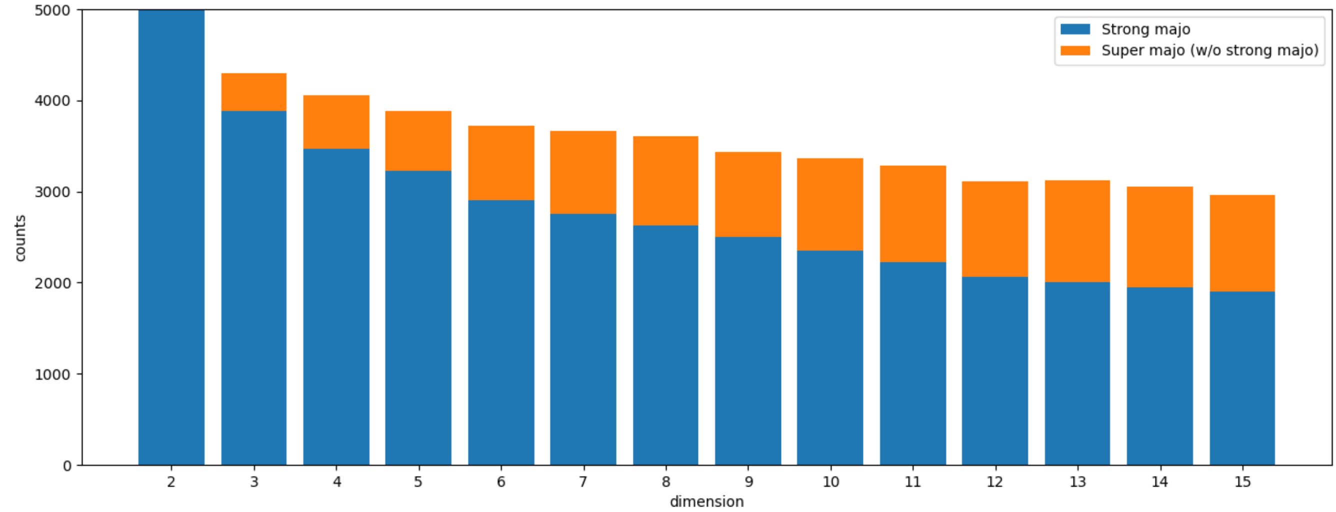
Numerical tests for qudits are performed in the following way

- Two d-dimensional vectors of SC, $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$, are randomly generated.
- $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ are compared under strong majorization and a counter, let us call it `majo_cnt`, is incremented if they are comparable
- The associated vectors of OSC, $\boldsymbol{\lambda}^{OS}$ and $\boldsymbol{\eta}^{OS}$, are deduced for both vectors.
- $\boldsymbol{\lambda}^{OS}$ and $\boldsymbol{\eta}^{OS}$ are compared under super- and sub-majorization. If they are comparable under super-majorization, a counter `super_majo_cnt` is incremented. If they are comparable under sub-majorization, a counter `sub_majo_cnt` is incremented.

[2]We suppose that $t \neq 0$ and 1, otherwise (4.21) is already trivially true.

[3]Unless (4.12) is an equality, which would be a trivial case.

In the following figure, for each dimension $d \in [2, 15]$, 5000 loops of the preceding procedure have been performed.



	2	3	4	5	6	7	8	9	10	11	12	13	14	15
majo_cnt	5000	3888	3467	3223	2909	2754	2627	2497	2347	2223	2062	2011	1945	1900
super_majo_cnt	5000	4294	4055	3883	3725	3665	3611	3429	3366	3281	3115	3120	3050	2958
sub_majo_cnt	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 4.3: Numerical evidence of Conjecture 4.1. In blue, the number of (randomly generated) vectors of SC that are comparable under strong majorization. In (orange+blue), the number of associated vectors of OSC that are comparable under super-majorization.

First, remark that for $d = 2$, all vectors of SC are comparable under majorization, which can be easily analytically shown[4]. Furthermore, still for $d = 2$, all vectors of OSC are also comparable under super-majorization, this is in correspondence with the proof of Section 4.1.1 where we have shown that $\boldsymbol{\lambda} \prec \boldsymbol{\eta} \Leftrightarrow \boldsymbol{\lambda}^{OS} \prec^{w} \boldsymbol{\eta}^{OS}$.

Second, for $d > 2$, not all vectors of SC are comparable under majorization, which is consistent with Example 5 and some vectors of OSC are comparable under super-majorization while the corresponding vectors of SC are not comparable under strong majorization. Note the important remark that, while not made explicit in Figure 4.3, all counts of `majo_cnt` are well present in `super_majo_cnt`.

Third, as proved at the end of Section 4.1.3, vectors of OSC can never be comparable under sub-majorization if their corresponding vectors of SC are comparable under majorization. Moreover, we observe that no vectors of OSC are comparable under sub-majorization, even if their corresponding vectors of SC are incomparable under majorization.

[4]Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ and $\boldsymbol{\eta} = (\eta_1, \eta_2)$. If $\lambda_1 \leq \eta_1$, then $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$. Else, if $\eta_1 \leq \lambda_1$, then $\boldsymbol{\eta} \prec \boldsymbol{\lambda}$, hence they are always comparable under strong majorization.

Chapter 5

Weak majorization and probabilistic convertibility

In the preceding chapter, we have conjectured (and partly proved) a necessary condition for the deterministic conversion of bipartite pure states involving the OS decomposition.

In this chapter, we prove a sufficient condition for the probabilistic conversion of bipartite pure states using the OS decomposition in Section 5.1. Then, in Section 5.2, we propose an abstract representation to summarize the (conjectured) result of the previous chapter with the result of Section 5.1 of the present chapter.

Finally, in Section 5.3, we make a first incursion into the realm of mixed states by exploring a connection between the super-majorization relation and the probabilistic protocol of Vidal (see Section 3.3.2 of Chapter 3). Note that this is quite important because we finally use the OS decomposition for mixed states.

5.1 Condition based on the OS decomposition

In this section, we present a theorem allowing one to know with what probability a conversion between two bipartite pure states whose OSC are comparable under super-majorization is, at least, possible. Although rather weak, because the condition is only sufficient, it is nonetheless a link made between probabilistic conversion and super-majorization on the OSC.

Theorem 5.1. *Let $|\psi\rangle$ and $|\phi\rangle$ be two d-dimensional bipartite pure states, with $\boldsymbol{\lambda}$ (resp. $\boldsymbol{\eta}$) the vector of Schmidt coefficients (SC) of $|\psi\rangle$ (resp. $|\phi\rangle$) and $\boldsymbol{\lambda}^{OS}$ (resp. $\boldsymbol{\eta}^{OS}$) the vector of operator Schmidt coefficients (OSC) of $|\psi\rangle\langle\psi|$ (resp. $|\phi\rangle\langle\phi|$).*

If $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS}$, then $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$, with

$$p \leq p_{super} = \min\left\{1, \min_{k\in[1,d-1]}\left\{1 - \frac{2}{\sum_{i=k}^{d}\eta_i}\sum_{\substack{i\neq j\\ i+j\geq 2k}}(\sqrt{\lambda_i\lambda_j} - \sqrt{\eta_i\eta_j})\right\}\right\}, \tag{5.1}$$

which implies, by means of Theorem 3.2, that $|\psi\rangle$ is convertible into $|\phi\rangle$ using only LOCC with probability $p \leq p_{super}$.

Proof. We prove the theorem for qutrits, the generalization following naturally afterwards.

On the one hand, using (2.16), we first develop[1] the super-majorization relation $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS}$.

$$\lambda_3 \geq \eta_3, \tag{5.2}$$
$$\lambda_3 + \sqrt{\lambda_2\lambda_3} \geq \eta_3 + \sqrt{\eta_2\eta_3}, \tag{5.3}$$
$$\lambda_3 + 2\sqrt{\lambda_2\lambda_3} \geq \eta_3 + 2\sqrt{\eta_2\eta_3}, \tag{5.4}$$
$$\lambda_3 + 2\sqrt{\lambda_2\lambda_3} + ? \geq \eta_3 + 2\sqrt{\eta_2\eta_3} + ?, \tag{5.5}$$
$$\lambda_3 + 2\sqrt{\lambda_2\lambda_3} + ? \geq \eta_3 + 2\sqrt{\eta_2\eta_3} + ?, \tag{5.6}$$
$$\lambda_3 + 2\sqrt{\lambda_2\lambda_3} + 2\sqrt{\lambda_1\lambda_3} + \lambda_2 \geq \eta_3 + 2\sqrt{\eta_2\eta_3} + 2\sqrt{\eta_1\eta_3} + \eta_2, \tag{5.7}$$
$$\lambda_3 + 2\sqrt{\lambda_2\lambda_3} + 2\sqrt{\lambda_1\lambda_3} + \lambda_2 + ? \geq \eta_3 + 2\sqrt{\eta_2\eta_3} + 2\sqrt{\eta_1\eta_3} + \eta_2 + ?, \tag{5.8}$$
$$\lambda_3 + 2\sqrt{\lambda_2\lambda_3} + 2\sqrt{\lambda_1\lambda_3} + \lambda_2 + ? \geq \eta_3 + 2\sqrt{\eta_2\eta_3} + 2\sqrt{\eta_1\eta_3} + \eta_2 + ?, \tag{5.9}$$
$$\lambda_3 + 2\sqrt{\lambda_2\lambda_3} + 2\sqrt{\lambda_1\lambda_3} + \lambda_2 + 2\sqrt{\lambda_1\lambda_2} + \lambda_1 \geq \eta_3 + 2\sqrt{\eta_2\eta_3} + 2\sqrt{\eta_1\eta_3} + \eta_2 + 2\sqrt{\eta_1\eta_2} + \eta_1. \tag{5.10}$$

Remember that, as seen on Section 4.1.2 of Chapter 4, the non-fixed ordering of the OSC depending on the values of the SC implies that there is no general form for (5.5), (5.6), (5.8) and (5.9), whence the questions marks in those inequalities.

However, blue inequalities are present whatever the variable ordering in the decreasingly ordered vectors of OSC, thus we will make use of some inequalities belonging to the blue subset of inequalities, namely those where the simple terms ($\lambda_1, \lambda_2, \lambda_3$ and η_1, η_2, η_3) appear for the first time, i.e., (5.2), (5.7) and (5.10) in our case.

Those three inequalities can be rewritten in the following way[2]

$$\frac{\lambda_3}{\eta_3} \geq 1, \tag{5.11}$$
$$\frac{\lambda_2 + \lambda_3}{\eta_2 + \eta_3} \geq 1 - \frac{1}{\eta_2 + \eta_3}\left(2\sqrt{\lambda_2\lambda_3} + 2\sqrt{\lambda_1\lambda_3} - 2\sqrt{\eta_2\eta_3} - 2\sqrt{\eta_1\eta_3}\right), \tag{5.12}$$
$$\frac{\lambda_1 + \lambda_2 + \lambda_3}{\eta_1 + \eta_2 + \eta_3} \geq 1 - \frac{1}{\eta_1 + \eta_2 + \eta_3}\Big(2\sqrt{\lambda_2\lambda_3} + 2\sqrt{\lambda_1\lambda_3} + 2\sqrt{\lambda_1\lambda_2} - 2\sqrt{\eta_2\eta_3} - 2\sqrt{\eta_1\eta_3} - 2\sqrt{\eta_1\eta_2}\Big). \tag{5.13}$$

On the other hand, using (2.16), we develop now the super-majorization relation $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$

$$\lambda_3 \geq p\eta_3, \tag{5.14}$$
$$\lambda_2 + \lambda_3 \geq p(\eta_2 + \eta_3), \tag{5.15}$$
$$\lambda_1 + \lambda_2 + \lambda_3 \geq p(\eta_1 + \eta_2 + \eta_3), \tag{5.16}$$
$$\tag{5.17}$$

[1] As always, vectors are decreasingly ordered, hence $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and $\eta_1 \geq \eta_2 \geq \eta_3$.

[2] We suppose that η_3 and η_2 are different from 0, otherwise the inequalities (5.2) and (5.7) are trivially satisfied.

which can be rewritten in a similar way to (5.11), (5.12) and (5.13)[3]

$$\frac{\lambda_3}{\eta_3} \geq p, \tag{5.18}$$

$$\frac{\lambda_2+\lambda_3}{\eta_2+\eta_3} \geq p, \tag{5.19}$$

$$\frac{\lambda_1+\lambda_2+\lambda_3}{\eta_1+\eta_2+\eta_3} \geq p. \tag{5.20}$$

Thus, we see that, if $p \leq p_{\text{super}} = \min\Big\{1, 1 - \frac{1}{\eta_2+\eta_3}\Big(2\sqrt{\lambda_2\lambda_3} + 2\sqrt{\lambda_1\lambda_3} - 2\sqrt{\eta_2\eta_3} - 2\sqrt{\eta_1\eta_3}\Big), 1 - \frac{1}{\eta_1+\eta_2+\eta_3}\Big(2\sqrt{\lambda_2\lambda_3} + 2\sqrt{\lambda_1\lambda_3} + 2\sqrt{\lambda_1\lambda_2} - 2\sqrt{\eta_2\eta_3} - 2\sqrt{\eta_1\eta_3} - 2\sqrt{\eta_1\eta_2}\Big)\Big\}$, then, inequalities (5.18), (5.19) and (5.20) are all satisfied, meaning that $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$ for $p \leq p_{\text{super}}$.

□

However, in general, there may exist higher values of p such that the conversion is still possible. This can be understood because of the following inequality that can easily be shown,

$$p_{\text{super}} \leq p_{\max}, \tag{5.21}$$

where $p_{\max}$ corresponds to the optimal conversion probability (see Theorem 3.3).

The following two examples showcase (i) a typical application of Theorem 5.1 and (ii) an example why the theorem can only be sufficient.

Example 12:

- $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS} \Rightarrow \boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$, with $p \leq p_{\text{super}}$

  Let us reuse the second point of Example 11, with $\boldsymbol{\lambda} = (5/10,\ 3/10,\ 2/10)$ and $\boldsymbol{\eta} = (13/30,\ 13/30,\ 4/30)$. We had shown that $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS}$, while $\boldsymbol{\lambda} \not\prec \boldsymbol{\eta}$. Now, by Theorem 5.1, we can calculate, from the components of $\boldsymbol{\lambda}^{OS}$ and $\boldsymbol{\eta}^{OS}$, the value of p_{super}

  $$p_{\text{super}} \simeq \min\{1, 0.716, 0.872\} = 0.716, \tag{5.22}$$

  which implies that $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$ for $p \leq 0.716$.

  In order to see that both super-majorization relations are fulfilled, we draw the associated Lorenz curves in Figure 5.1.

  Note that, by (3.19), $p_{\max} = 15/17 \simeq 0.882$ which is higher than p_{super} as expected.

- $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$, with $p \leq p_{\text{super}} \not\Rightarrow \boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS}$

  Let us choose now $\boldsymbol{\lambda} = (5/10,\ 3/10,\ 2/10)$ and $\boldsymbol{\eta} = (44/100,\ 39/100,\ 17/100)$. If we calculate p_{super} with the use of (5.1), we have

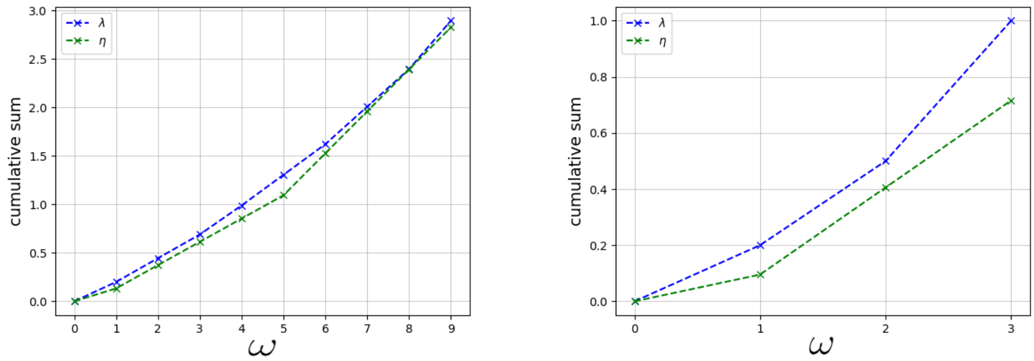
  $$p_{\text{super}} \simeq \min\{1, 0.892, 0.994\} = 0.892, \tag{5.23}$$

[3] Again, assuming that η_3 and η_2 are different from 0.

which we show allows $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$ for $p \leq 0.892$ (see Figure 5.2a).

However, we can show that $\boldsymbol{\lambda}^{OS} \not\prec^w \boldsymbol{\eta}^{OS}$, implying that the implication sign of Theorem 5.1 can only go one way.

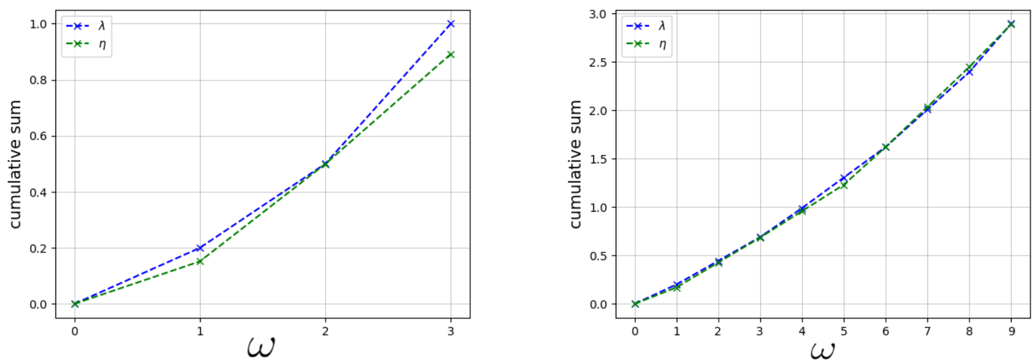
Again, we make use of Lorenz curves to verify that $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$ for $p \leq 0.892$, while $\boldsymbol{\lambda}^{OS} \not\prec^w \boldsymbol{\eta}^{OS}$ (see Figure 5.2).



(a) Lorenz curves associated to the vector of OSC of $\boldsymbol{\lambda} = (0.5,\ 0.3,\ 0.2)$ and to the vector of OSC of $\boldsymbol{\eta} = (13/30,\ 13/30,\ 4/30)$, the graph of $L^{\uparrow}_{\boldsymbol{\lambda}^{OS}}(\omega)$ is above the graph of $L^{\uparrow}_{\boldsymbol{\eta}^{OS}}(\omega)$, hence $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS}$.

(b) Lorenz curves associated to $\boldsymbol{\lambda} = (0.5,\ 0.3,\ 0.2)$ and $p_{\text{super}}\boldsymbol{\eta} = (0.716\times 13/30,\ 0.716\times 13/30,\ 0.716\times 4/30)$, the graph of $L^{\uparrow}_{\boldsymbol{\lambda}}(\omega)$ is above the graph of $L^{\uparrow}_{p_{\text{super}}\boldsymbol{\eta}}(\omega)$, hence $\boldsymbol{\lambda} \prec^w p_{\text{super}}\boldsymbol{\eta}$.

Figure 5.1: Lorenz curves for the first point of Example 12.



(a) Lorenz curves associated to $\boldsymbol{\lambda} = (0.5,\ 0.3,\ 0.2)$ and $p_{\text{super}}\boldsymbol{\eta} = (0.892\times 44/100,\ 0.892\times 39/100,\ 0.892\times 17/100)$, the graph of $L^{\uparrow}_{\boldsymbol{\lambda}}(\omega)$ is above the graph of $L^{\uparrow}_{p_{\text{super}}\boldsymbol{\eta}}(\omega)$, hence $\boldsymbol{\lambda} \prec^w p_{\text{super}}\boldsymbol{\eta}$.

(b) Lorenz curves associated to the vector of OSC of $\boldsymbol{\lambda} = (0.5,\ 0.3,\ 0.2)$ and to the vector of OSC of $\boldsymbol{\eta} = (44/100,\ 39/100,\ 17/100)$, the graph of $L^{\uparrow}_{\boldsymbol{\lambda}^{OS}}(\omega)$ intersects the graph of $L^{\uparrow}_{\boldsymbol{\eta}^{OS}}(\omega)$ (see the interval $\omega \in [5, 8]$), hence $\boldsymbol{\lambda}^{OS} \not\prec^w \boldsymbol{\eta}^{OS}$.

Figure 5.2: Lorenz curves for the second point of Example 12.

5.2 Diagrammatic representation

In this section, we bring together Conjecture 4.1 of the previous chapter and Theorem 5.1 of this present chapter. This allows to better identify where the super-majorization relation on the OSC stands compared to the majorization relation on the SC of Nielsen's theorem (3.4) and the super-majorization relation on the SC of Theorem 3.2.

The following figure displays, for a fixed vector of SC $\boldsymbol{\lambda}$, the vectors of SC $\boldsymbol{\eta}$ such that, $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$ (inside of the green ellipse), $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS}$ (inside of the red circle) and $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$ with $p \in [0, p_{\text{super}}]$ (inside of the blue ellipse).

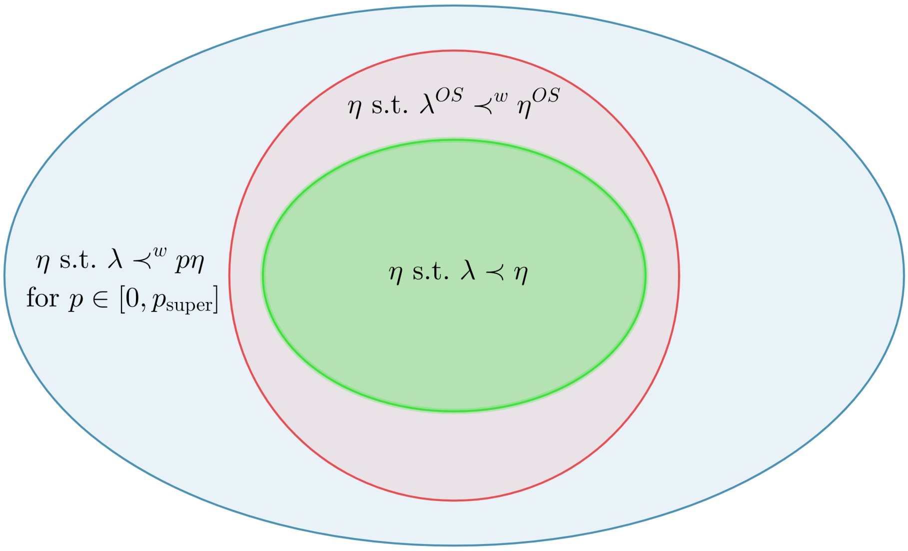


Figure 5.3: For a fixed vector of SC $\boldsymbol{\lambda}$, diagrammatic representation of the sets of vectors of SC $\boldsymbol{\eta}$ such that, (i) inside the green ellipse, $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$, i.e., by Theorem 3.1, the states with vectors of SC $\boldsymbol{\eta}$ that can be deterministically transformed to, by LOCC, from the state of SC $\boldsymbol{\lambda}$, (ii) inside the red circle, $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS}$, (iii), inside the blue ellipse, $\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta}$ with $p \in [0, p_{\text{super}}]$, i.e., by Theorem 3.2, the states with vectors of SC $\boldsymbol{\eta}$ that can be probabilistically transformed to, by LOCC, from the state of SC $\boldsymbol{\lambda}$, with probability belonging at least to $[0, p_{\text{super}}]$, where p_{super} can be calculated from (5.1).

Firstly, the inclusion of the green ellipse in the red circle reflects the result of Conjecture 4.1, i.e., $\boldsymbol{\lambda} \prec \boldsymbol{\eta} \Rightarrow \boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\eta}^{OS}$. Secondly, the inclusion of the red circle in the blue ellipse reflects the result of Theorem 5.1.

Finally, unless we consider qubits[4] or specific values of $\boldsymbol{\lambda}$ such as $(1/d, ..., 1/d)$ or $(1, 0, ..., 0)$, the inclusions between the three sets are strict. This has been shown by the counter-examples presented in Example 11 (see its second point) and in Example 12 (see its second point).

[4]In this case the green and red ensembles have been shown to be equal in Section 4.1.1 of Chapter 4. Furthermore, the blue set can also be shown to be equal to the green and red ones.

5.3 Application of super-majorization to mixed states

In this section, we make a first connection between weak majorization, more specifically super-majorization, on the OSC and mixed states.

Let us consider Vidal's protocol (see Section 3.3.2 of Chapter 3), the goal being to transform a bipartite pure state $|\psi\rangle$ into another bipartite pure state $|\phi\rangle$ (possibly) probabilistically using only LOCC.

Recall that, if the conversion cannot be made deterministically, an intermediate state $|\chi\rangle$ is first reached by deterministic conversion. Then, by performing a proper measurement, this state is transformed into $|\phi\rangle$ with probability $p_{\max}$ (3.19). Otherwise, it is transformed into another state $|\xi\rangle$ with probability $1 - p_{\max}$.

However, we can choose to look at the final state of the system rather in a statistical mixture of the two possible outcomes (1.47), i.e.,

$$\hat{\sigma} = p_{\max} |\phi\rangle \langle\phi| + (1 - p_{\max}) |\xi\rangle \langle\xi| . \tag{5.24}$$

In the rest of this chapter, we will prove for qubits, and conjecture for qudits, that the following relation is obeyed between the vector of OSC $\boldsymbol{\gamma}^{OS}$ of the intermediary state $|\chi\rangle$ and the vector of OSC $\boldsymbol{\varsigma}^{OS}$ of $\hat{\sigma}$,

$$\boldsymbol{\gamma}^{OS} \prec^{w} \boldsymbol{\varsigma}^{OS}, \tag{5.25}$$

where $\boldsymbol{\varsigma}^{OS}$ is now the vector of OSC of a mixed state.

5.3.1 Case of qubits

First, let us consider the case of qubits and assume that $|\psi\rangle$ cannot be transformed deterministically into $|\phi\rangle$, otherwise the problem would be trivial because we would have $p_{\max} = 1$. Because of these two assumptions, we can show[5] that the intermediary state is the initial state $|\psi\rangle$, meaning that no LOCC is performed before the measurement performed by Alice (or Bob).

By construction, we can show that $|\xi\rangle$, the unwanted outcome state of the measurement, has a strictly lower Schmidt rank than $|\phi\rangle$[5]. Hence, because we consider qubits, it has a Schmidt rank equal to 1, meaning that $|\xi\rangle = |00\rangle$. This allows to simply write the vector of OSC of $\hat{\sigma}$ as follows[6]

$$\boldsymbol{\varsigma}^{OS} = p_{\max} \begin{pmatrix} \eta_1 \\ \sqrt{\eta_1\eta_2} \\ \sqrt{\eta_1\eta_2} \\ \eta_2 \end{pmatrix} + (1 - p_{\max}) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 - p_{\max}\eta_2 \\ p_{\max}\sqrt{\eta_1\eta_2} \\ p_{\max}\sqrt{\eta_1\eta_2} \\ p_{\max}\eta_2 \end{pmatrix}, \tag{5.26}$$

where we made use of the fact that $\eta_2 = 1 - \eta_1$.

We can replace $p_{\max}$ by its value in terms of the SC of $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ (3.19), i.e., $p_{\max} = \lambda_2/\eta_2$, allowing to rewrite $\boldsymbol{\varsigma}^{OS}$ as $(\lambda_1,\ \lambda_2\sqrt{\eta_1/\eta_2},\ \lambda_2\sqrt{\eta_1/\eta_2},\ \lambda_2)$.

[5]See Appendix B.

[6]As always, $\boldsymbol{\lambda}$ is the vector of SC of $|\psi\rangle$ and $\boldsymbol{\eta}$ the vector of SC of $|\phi\rangle$, both vectors being decreasingly reordered.

Now, we will show that we have $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\varsigma}^{OS}$. By unraveling the super-majorization relation using (2.16), we get

$$\begin{cases} \lambda_2 \geq \lambda_2, \\ \lambda_2 + \sqrt{\lambda_1 \lambda_2} \geq \lambda_2 + \lambda_2 \sqrt{\frac{\eta_1}{\eta_2}}, \\ \lambda_2 + 2\sqrt{\lambda_1 \lambda_2} \geq \lambda_2 + 2\lambda_2 \sqrt{\frac{\eta_1}{\eta_2}}, \\ \lambda_2 + 2\sqrt{\lambda_1 \lambda_2} + \lambda_1 \geq \lambda_2 + 2\lambda_2 \sqrt{\frac{\eta_1}{\eta_2}} + \lambda_1, \end{cases}$$

hence it suffices to show that $\sqrt{\lambda_1 \lambda_2} \geq \lambda_2 \sqrt{\frac{\eta_1}{\eta_2}}$ to prove that the set of inequality is satisfied. First, let us rewrite the latter inequality as $\frac{\lambda_1}{\eta_1} \geq \frac{\lambda_2}{\eta_2}$. To show that it is true, recall that we have assumed that $\boldsymbol{\lambda} \not\prec \boldsymbol{\eta}$. Therefore, for qubits, $\boldsymbol{\eta} \prec \boldsymbol{\lambda}$, i.e., $\eta_1 \leq \lambda_1$ and $\eta_2 \geq \lambda_2$, implying immediately that $\frac{\lambda_1}{\eta_1} \geq \frac{\lambda_2}{\eta_2}$.

This allows us to conclude that we indeed have $\boldsymbol{\lambda}^{OS} \prec^w \boldsymbol{\varsigma}^{OS}$.

5.3.2 Case of qudits

The generalization of the previous proof to any dimension d is quite cumbersome.

Indeed, the intermediary state $|\chi\rangle$ is now, in general, different from $|\psi\rangle$ and its structure depends on the values of the SC of $|\psi\rangle$ and $|\phi\rangle$[7]. As a matter of fact, the analytical proof brings up several different cases, whose number is growing exponentially with d.

Furthermore, in each case, we have to prove a super-majorization relation, i.e., d^2 inequalities, which have already proved to be difficult to treat analytically in another context (see Section 4.1.2 of Chapter 4).

Nonetheless, once again, a large number of numerical tests have been performed[8] and no counter-example has been encountered, giving us the strong intuition that the super-majorization between the vector of OSC of $|\chi\rangle$ and the one of $\hat{\sigma}$ should hold.

[7]See Appendix B.

[8]For values of $d < 16$.

Chapter 6

Majorization lattice in quantum entanglement theory

This last chapter is devoted to the particular subject of the majorization lattice, first introduced in [10]. As of now, this mathematical field remains rarely applied to quantum information and only few recent papers exploit it, see e.g., [12, 34]. However, it turns out to be a very convenient tool in entanglement transformations as we will see throughout this chapter.

At first, Section 6.1 exposes the theory of majorization lattice by introducing mathematical concepts and terminology. Then, Section 6.2 presents Conjecture 4.1 and Theorem 5.1 in a unified picture, as was done in Section 5.2 of Chapter 5, using the majorization lattice formalism. Next, in Section 6.3, we interpret and apply the notions of majorization lattice to quantum information. Finally, in Section 6.4, we explore a link between the majorization lattice and Vidal's protocol (see Section 3.3.2 of Chapter 3), emphasizing the possible applications of majorization lattice notions to quantum information protocols.

6.1 Theory of majorization lattice

The majorization lattice is an application of the order theory notion of "lattice" to the theory of majorization. In this report, we focus only on the majorization lattice definition while the general concept of lattice is detailed in-depth in [35].

6.1.1 Definition

The majorization lattice is defined in [10] as a quadruple $\langle \mathcal{P}_d, \prec, \wedge, \vee \rangle$, where

- $\mathcal{P}_d$ is the set of all decreasingly ordered d-dimensional (or less) probability distributions, i.e., $\mathcal{P}_d = \{\boldsymbol{p} = (p_1, ..., p_d)$ such that $p_1 \geq ... \geq p_d \geq 0$ and $\sum_{i=1}^{d} p_i = 1\}$.
- $\prec$ is the (partial) order relation of majorization.
- $\wedge$ defines a unique greatest lower bound, called the meet.
- $\vee$ defines a unique least upper bound, called the join.

As such, $\mathcal{P}_d$ constitutes the set of the lattice elements and $\prec$ the relation in which they are ordered.

Now, the meet of $\boldsymbol{\lambda}$, $\boldsymbol{\eta} \in \mathcal{P}_d$ is denoted $\boldsymbol{\lambda} \wedge \boldsymbol{\eta}$ and is defined as the unique element of $\mathcal{P}_d$ such that

$$\forall \boldsymbol{\alpha} \in \mathcal{P}_d \text{ with } \boldsymbol{\alpha} \prec \boldsymbol{\lambda} \text{ and } \boldsymbol{\alpha} \prec \boldsymbol{\eta}, \text{ we have } \boldsymbol{\alpha} \prec \boldsymbol{\lambda} \wedge \boldsymbol{\eta}, \tag{6.1}$$

which implies obviously that $\boldsymbol{\lambda} \wedge \boldsymbol{\eta} \prec \boldsymbol{\lambda}$ and $\boldsymbol{\lambda} \wedge \boldsymbol{\eta} \prec \boldsymbol{\eta}$. Therefore, the meet can be understood as the only distribution majorized by $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ that majorizes all distributions majorized by both $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$, hence the name greatest lower bound.

Analogously, the join of $\boldsymbol{\lambda}$, $\boldsymbol{\eta} \in \mathcal{P}_d$ is denoted $\boldsymbol{\lambda} \vee \boldsymbol{\eta}$ and is defined as the unique element of $\mathcal{P}_d$ such that

$$\forall \boldsymbol{\beta} \in \mathcal{P}_d \text{ with } \boldsymbol{\lambda} \prec \boldsymbol{\beta} \text{ and } \boldsymbol{\eta} \prec \boldsymbol{\beta}, \text{ we have } \boldsymbol{\lambda} \vee \boldsymbol{\eta} \prec \boldsymbol{\beta}, \tag{6.2}$$

which implies obviously that $\boldsymbol{\lambda} \prec \boldsymbol{\lambda} \vee \boldsymbol{\eta}$ and $\boldsymbol{\eta} \prec \boldsymbol{\lambda} \vee \boldsymbol{\eta}$. Therefore, the join can be understood as the only distribution majorizing $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ that is majorized by all distributions majorizing both $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$, hence the name least upper bound.

Note already that these notions of meet and join are of real interest only if both distributions are incomparable under majorization. Indeed, if they are comparable, e.g., if $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$, then the meet $\boldsymbol{\lambda} \wedge \boldsymbol{\eta} = \boldsymbol{\lambda}$ and the join $\boldsymbol{\lambda} \vee \boldsymbol{\eta} = \boldsymbol{\eta}$. This remark will be clearer when representing graphically the majorization lattice in Section 6.1.2 and when using Lorenz curves to describe the meet and the join in Section 6.1.3.

6.1.2 Abstract representation

Before detailing how the meet and the join can be determined analytically, we present an abstract representation [36] of the majorization lattice in Figure 6.1. This representation should be kept in mind all along the rest of this chapter.

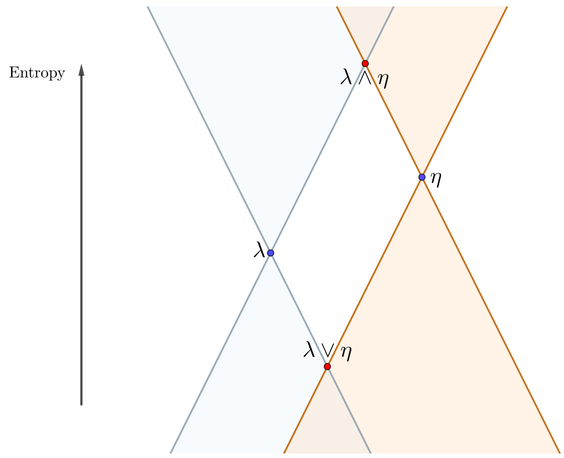


Figure 6.1: Abstract representation of the majorization lattice. $\boldsymbol{\lambda}, \boldsymbol{\eta} \in \mathcal{P}_d$ are incomparable under majorization. The meet of $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$, $\boldsymbol{\lambda} \wedge \boldsymbol{\eta}$, is situated at the intersection of the upper cones of both distributions, while the join of $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$, $\boldsymbol{\lambda} \vee \boldsymbol{\eta}$, is situated at the intersection of the lower cones of both distributions.

Let us now describe the elements of Figure 6.1. First, all elements of $\mathcal{P}_d$ are represented by points.

Second, by convention, the higher a distribution is in the lattice, the more it possesses entropy (2.13). This implies that the highest distribution is the uniform distribution, corresponding to the vector of SC of the maximally-entangled state, while the lowest distribution is (1, 0, ..., 0), corresponding to the vector of SC of the separable state.

Third, each distribution of the lattice, e.g., $\boldsymbol{\lambda}$, possesses an upper and a lower cone[1]. The upper cone of $\boldsymbol{\lambda}$ corresponds to all distributions majorized by $\boldsymbol{\lambda}$, i.e., all $\boldsymbol{\gamma} \in \mathcal{P}_d$ such that $\boldsymbol{\gamma} \prec \boldsymbol{\lambda}$, while the lower cone of $\boldsymbol{\lambda}$ corresponds to all distributions majorizing $\boldsymbol{\lambda}$, i.e., all $\boldsymbol{\delta} \in \mathcal{P}_d$ such that $\boldsymbol{\lambda} \prec \boldsymbol{\delta}$.

Finally, when two upper cones intersect, it corresponds to the meet $\boldsymbol{\lambda} \wedge \boldsymbol{\eta}$ of both distributions, i.e., the distribution with the smallest entropy that is majorized by both distributions. Conversely, when two lower cones intersect, it corresponds to the join $\boldsymbol{\lambda} \vee \boldsymbol{\eta}$ of both distributions, i.e., the distribution with the greatest entropy that majorizes both distributions.

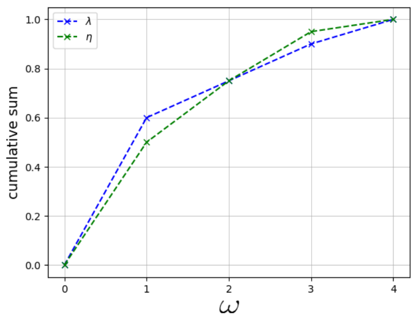
In the special case where the two distributions are comparable under majorization, it is obvious from Figure 6.1 that the intersection of the upper cones, the meet, happens immediately on the uppermost distribution on the lattice and conversely for the join.

6.1.3 Determination of the meet and the join

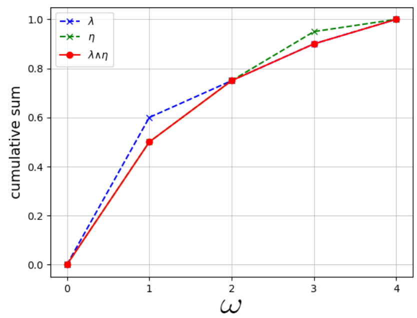
The analytical construction of the meet and the join is based on Lorenz curves (see Section 2.1.1.c of Chapter 2). Thus, for the analytical result to be clearer, we first construct Lorenz curves for both and then detail their analytical form.

6.1.3.a Meet

In order to construct the Lorenz curve of the meet of two distributions, it suffices to build the unique curve that is below the Lorenz curves of both distributions and such that all other Lorenz curves below both distributions are also below the Lorenz curve of the meet. Figure 6.2 displays such example.



(a) Let $\boldsymbol{\lambda} = (0.6, 0.15, 0.15, 0.1)$ and $\boldsymbol{\eta} = (0.5, 0.25, 0.20, 0.05)$ be two distributions $\in \mathcal{P}_4$. Because the two Lorenz curves $L^{\downarrow}_{\boldsymbol{\lambda}}(\omega)$ and $L^{\downarrow}_{\boldsymbol{\eta}}(\omega)$ intersect, the two distributions are incomparable under majorization, making this example a non-trivial one.



(b) The Lorenz curve of the meet of $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$, $L^{\downarrow}_{\boldsymbol{\lambda}\wedge\boldsymbol{\eta}}(\omega)$, consists in the union of all the lowest segments of $L^{\downarrow}_{\boldsymbol{\lambda}}(\omega)$ and $L^{\downarrow}_{\boldsymbol{\eta}}(\omega)$. It can be calculated that $\boldsymbol{\lambda} \wedge \boldsymbol{\eta}$ = (0.5, 0.25, 0.15, 0.1).

Figure 6.2: Example of the Lorenz curve for the meet of two distributions.

[1]Except the two extreme points, i.e., the uniform distribution and the peaked distribution.

Now, we can detail the analytical construction of the meet of two distributions [10] using decreasing cumulative sums (2.1).

Theorem 6.1 (Meet). *For all $\boldsymbol{\lambda}, \boldsymbol{\eta} \in \mathcal{P}_d$, the meet $\boldsymbol{\lambda} \wedge \boldsymbol{\eta} = (a_1, a_2, ..., a_d)$ is such that,*

$$a_k = \min\left\{S_k^{\downarrow}(\boldsymbol{\lambda}), S_k^{\downarrow}(\boldsymbol{\eta})\right\} - \min\left\{S_{k-1}^{\downarrow}(\boldsymbol{\lambda}), S_{k-1}^{\downarrow}(\boldsymbol{\eta})\right\} \qquad \text{for } k \in [1, d], \tag{6.3}$$

with the convention that $S_0^{\downarrow}(\boldsymbol{\lambda}) = S_0^{\downarrow}(\boldsymbol{\eta}) = 0$.

This theorem essentially translates analytically the construction procedure based on Lorenz curves that we have exhibited in Figure 6.2.

6.1.3.b Join

Contrary to the notion of meet, the join of two distributions is more complicated to determine (a full calculation can be found in [10]). This originates from the fact that if we build a curve that is just above the Lorenz curves of both distributions, analogously to what we did for the meet, this curve will not be concave anymore (see Figure 6.3a), meaning that it is not a decreasingly ordered distribution, thus not $\in \mathcal{P}_d$ and therefore not the join.

At first sight, one could think that by reordering the distribution the join can be obtained, let us call this new distribution the postulated join. In fact, now the curve is concave by construction. However, this postulated join is not the join because, in general, there exist other concave polygonal curves above the Lorenz curves of both distributions but below the Lorenz curve of this postulated join, hence it cannot be the join.

The solution is therefore to perform a certain algorithm to straighten the non-concave curve of Figure 6.3a "just enough" so that the curve is concave and such that no other Lorenz curve above the Lorenz curves of both distributions is below this straightened curve. Figure 6.3 displays such procedure where we have reused the example already presented in Figure 6.2a.

6.1.3.c Extension to more than two distributions

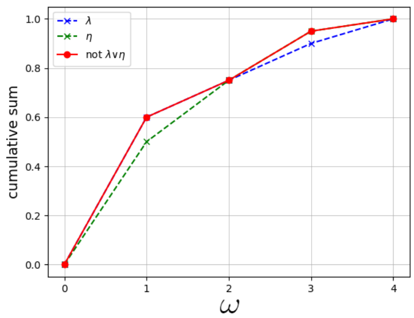
The notions of meet and join can be extended to more than two distributions [23]. As a matter of fact, because of the algebraic properties of the majorization lattice, the meet of n elements $\{\boldsymbol{\lambda}^j\}_{j=1}^n$ belonging to $\mathcal{P}_d$ is defined as

$$\wedge\{\boldsymbol{\lambda}^j\}_{j=1}^n = \boldsymbol{\lambda}^1 \wedge \boldsymbol{\lambda}^2 \wedge ... \wedge \boldsymbol{\lambda}^{n-1} \wedge \boldsymbol{\lambda}^n = \boldsymbol{\lambda}^1 \wedge (\boldsymbol{\lambda}^2 \wedge ... \wedge (\boldsymbol{\lambda}^{n-1} \wedge \boldsymbol{\lambda}^n \underbrace{)...)}_{n-2}, \tag{6.4}$$

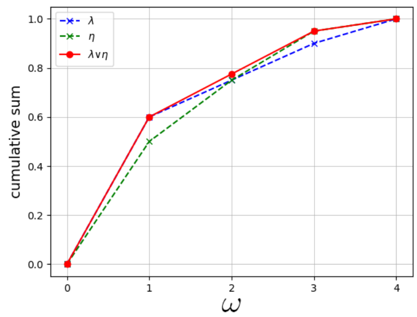
hence, it suffices to calculate the meet between $\boldsymbol{\lambda}^{n-1}$ and $\boldsymbol{\lambda}^n$, then between $\boldsymbol{\lambda}^{n-2}$ and $(\boldsymbol{\lambda}^{n-1} \wedge \boldsymbol{\lambda}^n)$, and so on.

Analogously, the join of n elements $\{\boldsymbol{\lambda}^j\}_{j=1}^n$ belonging to $\mathcal{P}_d$ is defined as

$$\vee\{\boldsymbol{\lambda}^j\}_{j=1}^n = \boldsymbol{\lambda}^1 \vee \boldsymbol{\lambda}^2 \vee ... \vee \boldsymbol{\lambda}^{n-1} \vee \boldsymbol{\lambda}^n = \boldsymbol{\lambda}^1 \vee (\boldsymbol{\lambda}^2 \vee ... \vee (\boldsymbol{\lambda}^{n-1} \vee \boldsymbol{\lambda}^n \underbrace{)...)}_{n-2}. \tag{6.5}$$



(a) Let $\boldsymbol{\lambda} = (0.6, 0.15, 0.15, 0.1)$ and $\boldsymbol{\eta} = (0.5, 0.25, 0.20, 0.05)$ be two distributions $\in \mathcal{P}_4$. The union of all the highest segments of $L^{\downarrow}_{\boldsymbol{\lambda}}(\omega)$ and $L^{\downarrow}_{\boldsymbol{\eta}}(\omega)$ does not constitute a concave polygonal curve, hence the red curve is not the join of $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$.



(b) The Lorenz curve of the join of $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$, $L^{\downarrow}_{\boldsymbol{\lambda}\vee\boldsymbol{\eta}}(\omega)$, consists in the straightening of the red curve of Figure 6.3a such that the curve is now a concave polygonal curve. It can be calculated that $\boldsymbol{\lambda} \vee \boldsymbol{\eta} = (0.6,\ 0.175,\ 0.175,\ 0.05)$.

Figure 6.3: Example of the difficulty encountered when searching the Lorenz curve for the join of two distributions.

6.2 Weak majorization and 3-dimensional majorization lattice

In this section, we introduce two new lattices involving weak majorization, in particular, super-majorization. The first one is based on Conjecture 4.1 and is denoted $\langle \mathcal{P}_d, \text{“} \prec^w \text{OS”}, \wedge, \vee \rangle$, where “ $\prec^w$ OS” is a shortcut for the “super-majorization on the OSC” ordering relation. The second one is based on Theorem 5.1 and is denoted $\langle \mathcal{P}_d, \text{“} \prec^w p_{\text{super}} \text{”}, \wedge, \vee \rangle$, where “ $\prec^w p_{\text{super}}$” is a shortcut for the “super-majorization on the SC with factor p_{super} (5.1) on the right-hand side” ordering relation. The notions of $\wedge$ and $\vee$ are defined the same way as in Section 6.1.1 for each order relation.

In order to make a graphical representation of the three majorization lattices, the first non-trivial set to consider is $\mathcal{P}_3$[2]. Therefore, we display in Figure 6.4 those three majorization lattices for 3-dimensional distributions. However, to be able to represent the three lattices in one graph is tedious, hence we only represent the upper and lower cones for one fixed distribution ($\boldsymbol{\lambda} = (0.5,\ 0.3,\ 0.2)$) associated to each of the three lattices. Moreover, a 2-dimensional graph is enough to fully represent 3-dimensional distributions because of the constraint on the absolute norm ($\lambda_1 + \lambda_2 + \lambda_3 = 1$), thus we make use of the following convenient parametrization

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3) = \left(\frac{1}{3} + y + x, y, x\right), \tag{6.6}$$

where $y = 1/3 - \lambda_2$ and $x = 1/3 - \lambda_3$ are the axes of Figure 6.4.

Note that dots represent 3-dimensional distributions in the graph and their colour denotes which of the three majorization lattices they belong to. As a matter of fact, this graph is simply the reproduction of the abstract Figure 5.3 using the majorization lattice formalism for 3-dimensional lattices.

Let us now analyze each part of the graph.

[2] Because all distributions are comparable in $\mathcal{P}_2$.

First, at the leftmost edge of the graph is situated the uniform distribution, i.e., the distribution associated to a maximally-entangled state (1/3, 1/3, 1/3), whereas at the rightmost upper edge is situated the distribution associated to a separable state (1, 0, 0).

Second, the upper and lower cones of the usual majorization lattice $\langle \mathcal{P}_3, \prec, \wedge, \vee \rangle$ contain all the green dots. Remark that the lower cone of $\boldsymbol{\lambda}$, i.e., the distributions $\boldsymbol{\eta} \in \mathcal{P}_3$ such that $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$, is situated to the right of $\boldsymbol{\lambda}$, while the upper cone of $\boldsymbol{\lambda}$, i.e., the distributions $\boldsymbol{\delta} \in \mathcal{P}_3$ such that $\boldsymbol{\delta} \prec \boldsymbol{\lambda}$, is situated to its left.

Third, the upper and lower cones (resp. at the left and at the right of $\boldsymbol{\lambda}$) of $\langle \mathcal{P}_3, \text{“} \prec^w \text{OS”}, \wedge, \vee \rangle$ contain all the red and green dots, which makes explicit why Conjecture 4.1 is a sufficient (but not necessary) condition for the super-majorization on the OSC. Indeed, if a distribution is green, we know that super-majorization on the SC is fulfilled while the converse is not true.

Fourth, the upper and lower cones (resp. at the left and at the right of $\boldsymbol{\lambda}$) of $\langle \mathcal{P}_3, \text{“} \prec^w p_{\text{super}} \text{”}, \wedge, \vee \rangle$ contain all the blue, red and green dots, which makes explicit why Theorem 5.1 is a necessary (but not sufficient) condition for the super-majorization on the OSC.

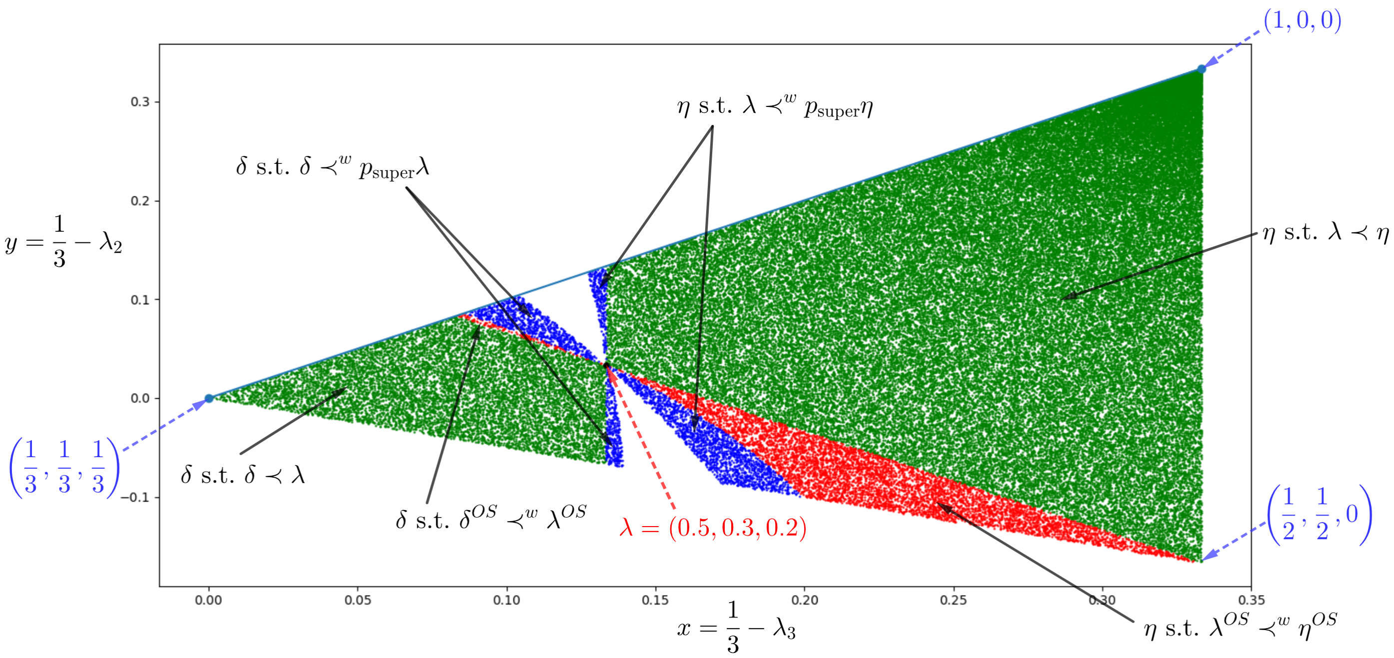


Figure 6.4: For a loop on 100,000 random 3-dimensional distributions, in blue, all distributions comparable with $\boldsymbol{\lambda} = (0.5, 0.3, 0.2)$ under “ $\prec^w p_{\text{super}}$”, in red, all distributions comparable with $\boldsymbol{\lambda}$ under “ $\prec^w p_{\text{super}}$” and “ $\prec^w$ OS” and, in green, all distributions comparable with $\boldsymbol{\lambda}$ under “ $\prec^w p_{\text{super}}$”, “ $\prec^w$ OS” and $\prec$.

6.3 Entanglement transformations

In the context of quantum information, and in our case in the study of entanglement transformations, distributions in the lattice correspond obviously to the vectors of SC of d-dimensional bipartite pure states. In the following, we will make the abuse of language that elements of the lattice are bipartite pure states, whereas, rigorously speaking, the elements of the lattice are the vectors of SC of such states[3].

In this section, we present the quantum informational notion of the meet, called *optimal common resource* (OCR), first introduced in [11]. Similarly, we propose a quantum informational notion of the join, which we call *optimal common product* (OCP), never introduced before to our knowledge. Building on those concepts, we propose a theoretical scenario making use of both OCR and OCP of multiple states.

6.3.1 Optimal common resource/product

In this context, the meet of two bipartite pure states $|\psi\rangle$ and $|\phi\rangle$ is the state $|\psi \wedge \phi\rangle$ such that both following conditions are fulfilled

- It can be deterministically transformed using LOCC (because of Theorem 3.1) into any of the two states $|\psi\rangle$ or $|\phi\rangle$.
- Among all possible states from the first point, its entropy[4] is minimal, hence it has a minimal amount of resource of entanglement.

In other words, the meet is the state with the minimal entanglement resource that can still nonetheless be transformed into both states using LOCC. As such, it can be seen as an optimal common resource [11, 23].

Naturally, the notion of join of two bipartite pure states $|\psi\rangle$ and $|\phi\rangle$ is defined analogously. The join $|\psi \vee \phi\rangle$ is the state such that both following conditions are fulfilled

- It can be deterministically obtained, from $|\psi\rangle$ or $|\phi\rangle$, using LOCC (because of Theorem 3.1).
- Among all possible states from the first point, its entropy[4] is maximal, hence it has a maximal amount of resource of entanglement.

In other words, the join is the state with the maximal entanglement resource that any of both states can be transformed into using LOCC. As such, it can be seen as an optimal common product.

Note that both definitions can be understood in the light of Figure 6.1. Indeed, the amount of entanglement resource is higher for states with more entropy[4], e.g., the maximally-entangled state is the highest one in the lattice, hence the amount of entanglement increases while going up in the lattice. Therefore, LOCC transformations are only allowed if going down in the lattice.

Finally, the notion of OCR and OCP can be extended to a set of n states, as was done for distributions in Section 6.1.3.c of this chapter.

---

[3]Because there is a one-to-one correspondence, up to unitary operations, between the vector of SC of a state and the state itself, this abuse of language seems reasonable.

[4]More precisely, the von Neumann entropy of its reduced density matrix, or equivalently, the Shannon entropy of its vector of SC.

6.3.2 Scenario involving multiple states

Let us now consider the following quantum information scenario making use of the OCR and the OCP.

Assume that Alice and Bob are far apart and wish to perform later on a specific quantum information protocol involving one of n different bipartite entangled pure states, i.e., a state of $\{|\psi^i\rangle\}_{i=1}^n$. However, at this time, we suppose that they don't know what state they will need. Furthermore, the only moment they have access to quantum communication is now, meaning that, when the protocol will begin, no quantum communication will be allowed for them and they will only be able to apply operations on their respective subsystems and communicate classically, hence to apply LOCC transformations.

Facing this problem, Alice and Bob decide to make use of the notions of OCR and OCP they recently discovered reading this report. To do so, each of them chooses to enter in communication with a "bank state". Such bank owns bipartite entangled pure states and can lend one qubit to each of the two requesting parties and, after some time, needs to recover each qubit, even if their state has changed. Furthermore, the loan price, that has to be paid to the bank, depends on the entanglement resource of the requested state and on the entanglement resource of the returned state, in the sense that the more entangled a requested state is, the more expensive its loan is and, conversely, the less entangled a returned state is, the more expensive its loan was. In order for the bank to be as organized as possible, they need the loaners to tell them, at the time of the loan, what state they want to loan and what state they will return.

In the case we are considering here, Alice and Bob request to the bank the OCR of the n different bipartite pure states because they wish to acquire a state that can be transformed using LOCC into any of those n states while possessing as less as possible entanglement, so that they minimize their loan price. Conversely, they tell the bank that they will return the OCP of the m different final[5] bipartite pure states $\{|\phi^i\rangle\}_{i=1}^m$ because they are sure that they will be able to transform their state into this OCP and they want the state they return to be as entangled as possible, in order, again, to minimize their loan price.

6.4 Probabilistic conversion of entanglement

In this section, we propose an alternative protocol to Vidal's optimal protocol (see Section 3.3.2 of Chapter 3) for converting probabilistically a state $|\psi\rangle$ into a state $|\phi\rangle$ using LOCC. Our protocol involves the OCR between $|\psi\rangle$ and $|\phi\rangle$ and is shown, in Theorem 6.2, to be equivalent in terms of probability to Vidal's protocol.[6]

6.4.1 Alternative protocol involving the OCR state

Before describing our protocol, we present how Vidal's protocol to convert $|\psi\rangle$ into $|\phi\rangle$ can be represented on the majorization lattice (see the green arrows in Figure 6.5). Let us consider the non-trivial case of $|\psi\rangle$ and $|\phi\rangle$ with vectors of SC incomparable

[5]We suppose that they know, in advance, what state they will recover after using one of the n different states $|\psi^i\rangle$ for the protocol, that for each of the n states.

[6]On a side note, we have found a criterion for the maximum conversion probability of a given state into any other state. Statement and proof can be found in Appendix C and won't be used in the following.

under majorization, hence with a conversion probability strictly less than one. $|\psi\rangle$ is first deterministically converted into $|\chi^1\rangle$[7] using LOCC. Note that the state $|\chi^1\rangle$ is situated in the lower cone of $|\psi\rangle$ because we have seen that deterministic conversions are only allowed to states situated in the lower cone of the initial state[8]. Moreover, it is proved in [34] that $|\chi^1\rangle$ must be in the lower cone of the join state $|\psi \vee \phi\rangle$. Now, from $|\chi^1\rangle$, Alice (or Bob) can perform a specific two-outcome measurement leading with probability $p_{\max}^{(1)}$ (3.19) to $|\phi\rangle$, completing the protocol.

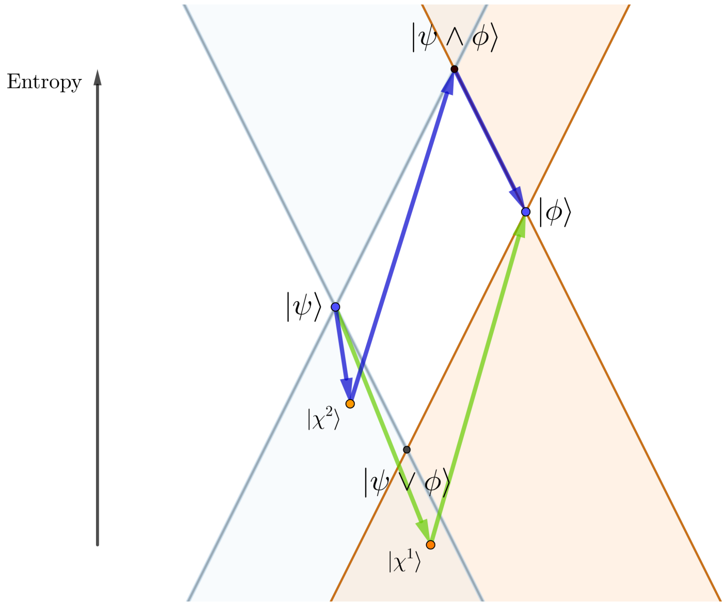


Figure 6.5: For two incomparable states under majorization $|\psi\rangle$ and $|\phi\rangle$, lattice representation of the conversions involved in Vidal's protocol (green arrows) and the conversions involved in the protocol using the OCR $|\psi \wedge \phi\rangle$ (blue arrows).

The protocol we propose (see the blue arrows in Figure 6.5) involves, as a subroutine, Vidal's protocol but this time between $|\psi\rangle$ and $|\psi \wedge \phi\rangle$. Indeed, we show in Theorem 6.2 that the probability $p_{\max}^{(2)}$ to perform this conversion is equal to the conversion probability between $|\psi\rangle$ and $|\phi\rangle$, $p_{\max}^{(1)}$.

Let us now describe our protocol. First, $|\psi\rangle$ is deterministically converted into $|\chi^2\rangle$. Note that, this time, $|\chi^2\rangle$ has the only constraint to be situated in the lower cone of $|\psi\rangle$ [34] because the join of $|\psi\rangle$ and $|\psi \wedge \phi\rangle$ is immediately $|\psi\rangle$. Then, Alice (or Bob) performs a specific two-outcome measurement leading, with probability $p_{\max}^{(2)}$, to state $|\psi \wedge \phi\rangle$. Finally, to complete the protocol, it suffices to perform a deterministic conversion from $|\psi \wedge \phi\rangle$ to $|\phi\rangle$, which is allowed because $|\phi\rangle$ is situated in the lower cone of $|\psi \wedge \phi\rangle$.

Although the protocol we propose may seem useless compared to Vidal's protocol, its usefulness resides in the fact that it is the first protocol, to our knowledge, involving the OCR between states. It is also quite surprising to be equally probable to reach an incomparable state ($|\phi\rangle$ in Figure 6.5) than to reach a state majorizing the initial state ($|\psi \wedge \phi\rangle$ in Figure 6.5), i.e., a state situated in the upper cone of the initial state.

[7] $|\chi^1\rangle \neq |\psi\rangle$ and $|\chi^1\rangle \neq |\phi\rangle$ because of the incomparability assumption.

[8] Because the Shannon entropy of the vector of SC must decrease during deterministic LOCC transformations (3.16).

6.4.2 Equivalence with Vidal's protocol

Finally, we present the theorem showing that it is equivalent, in terms of probability, to convert a state $|\psi\rangle$ into a state $|\phi\rangle$ using Vidal's protocol or using the protocol we propose involving the OCR.

Theorem 6.2. *Let $|\psi\rangle$ and $|\phi\rangle$ be two d-dimensional bipartite pure states. Let us call $p^{(1)}_{\max}$ the optimal probability (3.19) with which $|\psi\rangle$ can be transformed using LOCC into $|\phi\rangle$ and let us call $p^{(2)}_{\max}$ the optimal probability with which $|\psi\rangle$ can be transformed using LOCC into $|\psi \wedge \phi\rangle$. Then, both transformations are equiprobable, i.e.,*

$$p^{(1)}_{\max} = p^{(2)}_{\max}. \tag{6.7}$$

Proof. As always, let $\boldsymbol{\lambda}$ be the vector of SC of $|\psi\rangle$ and $\boldsymbol{\eta}$ the vector of SC of $|\phi\rangle$.

By means of (3.19), the expressions of $p^{(1)}_{\max}$ and $p^{(2)}_{\max}$ are

$$p^{(1)}_{\max} = \min_{i\in[1,d]} \frac{S^{\uparrow}_i(\boldsymbol{\lambda})}{S^{\uparrow}_i(\boldsymbol{\eta})} \quad \text{and} \quad p^{(2)}_{\max} = \min_{i\in[1,d]} \frac{S^{\uparrow}_i(\boldsymbol{\lambda})}{S^{\uparrow}_i(\boldsymbol{\lambda}\wedge\boldsymbol{\eta})}, \tag{6.8}$$

where $\boldsymbol{\lambda}\wedge\boldsymbol{\eta}$ is the vector of SC of $|\psi\wedge\phi\rangle$, whose components will be written $(a_1, a_2, ..., a_d)$ with $a_1 \geq a_2 \geq ... \geq a_d$.

The term $S^{\uparrow}_i(\boldsymbol{\lambda}\wedge\boldsymbol{\eta})$ of $p^{(2)}_{\max}$ is the sum of the i^{th} smallest components of $\boldsymbol{\lambda}\wedge\boldsymbol{\eta}$, i.e.,

$$S^{\uparrow}_i(\boldsymbol{\lambda}\wedge\boldsymbol{\eta}) = a_d + a_{d-1} + ... + a_{d-i+1}. \tag{6.9}$$

By using Theorem 6.1, the components of the meet $\boldsymbol{\lambda}\wedge\boldsymbol{\eta}$ can be expended as

$$\begin{aligned} S^{\uparrow}_i(\boldsymbol{\lambda}\wedge\boldsymbol{\eta}) = &\min\left\{S^{\downarrow}_d(\boldsymbol{\lambda}), S^{\downarrow}_d(\boldsymbol{\eta})\right\} - \min\left\{S^{\downarrow}_{d-1}(\boldsymbol{\lambda}), S^{\downarrow}_{d-1}(\boldsymbol{\eta})\right\} \\ &+ \min\left\{S^{\downarrow}_{d-1}(\boldsymbol{\lambda}), S^{\downarrow}_{d-1}(\boldsymbol{\eta})\right\} - \min\left\{S^{\downarrow}_{d-2}(\boldsymbol{\lambda}), S^{\downarrow}_{d-2}(\boldsymbol{\eta})\right\} \\ &+ ... \\ &+ \min\left\{S^{\downarrow}_{d-i+1}(\boldsymbol{\lambda}), S^{\downarrow}_{d-i+1}(\boldsymbol{\eta})\right\} - \min\left\{S^{\downarrow}_{d-i}(\boldsymbol{\lambda}), S^{\downarrow}_{d-i}(\boldsymbol{\eta})\right\}, \end{aligned} \tag{6.10}$$

hence, all terms cancel out two by two except the first term and the last term, i.e.,

$$S^{\uparrow}_i(\boldsymbol{\lambda}\wedge\boldsymbol{\eta}) = \min\left\{S^{\downarrow}_d(\boldsymbol{\lambda}), S^{\downarrow}_d(\boldsymbol{\eta})\right\} - \min\left\{S^{\downarrow}_{d-i}(\boldsymbol{\lambda}), S^{\downarrow}_{d-i}(\boldsymbol{\eta})\right\}. \tag{6.11}$$

Now, because $S^{\downarrow}_d(\boldsymbol{\lambda})$ (resp. $S^{\downarrow}_d(\boldsymbol{\eta})$) is the sum of all components of $\boldsymbol{\lambda}$ (resp. $\boldsymbol{\eta}$), it is equal to 1 (see Definition 1.5), giving

$$S^{\uparrow}_i(\boldsymbol{\lambda}\wedge\boldsymbol{\eta}) = 1 - \min\left\{S^{\downarrow}_{d-i}(\boldsymbol{\lambda}), S^{\downarrow}_{d-i}(\boldsymbol{\eta})\right\}. \tag{6.12}$$

This last relation is obviously equivalent to the following relation

$$S^{\uparrow}_i(\boldsymbol{\lambda}\wedge\boldsymbol{\eta}) = \max\left\{S^{\uparrow}_i(\boldsymbol{\lambda}), S^{\uparrow}_i(\boldsymbol{\eta})\right\}, \tag{6.13}$$

where we now make use of increasing cumulative sums.

Therefore, $p_{\max}^{(2)}$ (6.8) can be rewritten

$$p_{\max}^{(2)} = \min_{i\in[1,d]} \frac{S_i^{\uparrow}(\boldsymbol{\lambda})}{\max\left\{S_i^{\uparrow}(\boldsymbol{\lambda}), S_i^{\uparrow}(\boldsymbol{\eta})\right\}}. \tag{6.14}$$

To prove now that $p_{\max}^{(1)} = p_{\max}^{(2)}$, we first rewrite both probabilities ((6.8) and (6.14)) as $p_{\max}^{(1)} = \min_{i\in[1,d]} p_{\max,i}^{(1)}$ and $p_{\max}^{(2)} = \min_{i\in[1,d]} p_{\max,i}^{(2)}$ and analyze each of the two possible cases, either $S_i^{\uparrow}(\boldsymbol{\lambda}) \geq S_i^{\uparrow}(\boldsymbol{\eta})$ or $S_i^{\uparrow}(\boldsymbol{\lambda}) \leq S_i^{\uparrow}(\boldsymbol{\eta})$, for each $i \in [1, d]$.

- For $i \in [1, d]$, if $S_i^{\uparrow}(\boldsymbol{\lambda}) \geq S_i^{\uparrow}(\boldsymbol{\eta})$:

$$p_{\max,i}^{(1)} = \frac{S_i^{\uparrow}(\boldsymbol{\lambda})}{S_i^{\uparrow}(\boldsymbol{\eta})} \geq 1, \tag{6.15}$$

  hence, it does not have any influence on $p_{\max}^{(1)}$ because $p_{\max,d}^{(1)} = 1 \leq p_{\max,i}^{(1)}$.

  For $p_{\max}^{(2)}$, we have $\max\left\{S_i^{\uparrow}(\boldsymbol{\lambda}), S_i^{\uparrow}(\boldsymbol{\eta})\right\} = S_i^{\uparrow}(\boldsymbol{\lambda})$, implying that

$$p_{\max,i}^{(2)} = \frac{S_i^{\uparrow}(\boldsymbol{\lambda})}{S_i^{\uparrow}(\boldsymbol{\lambda})} = 1, \tag{6.16}$$

  which, again, does not have any influence on $p_{\max}^{(2)}$ because $p_{\max,d}^{(2)} = 1 = p_{\max,i}^{(2)}$.

- For $i \in [1, d]$, if $S_i^{\uparrow}(\boldsymbol{\lambda}) \leq S_i^{\uparrow}(\boldsymbol{\eta})$:

$$p_{\max,i}^{(1)} = \frac{S_i^{\uparrow}(\boldsymbol{\lambda})}{S_i^{\uparrow}(\boldsymbol{\eta})} \leq 1. \tag{6.17}$$

  For $p_{\max}^{(2)}$, we have $\max\left\{S_i^{\uparrow}(\boldsymbol{\lambda}), S_i^{\uparrow}(\boldsymbol{\eta})\right\} = S_i^{\uparrow}(\boldsymbol{\eta})$, implying that

$$p_{\max,i}^{(2)} = \frac{S_i^{\uparrow}(\boldsymbol{\lambda})}{S_i^{\uparrow}(\boldsymbol{\eta})} \leq 1, \tag{6.18}$$

  which is the same expression as $p_{\max,i}^{(1)}$.

Therefore, by taking the minimum over i for both $p_{\max,i}^{(1)}$ and $p_{\max,i}^{(2)}$, we show that both probabilities $p_{\max}^{(1)}$ and $p_{\max}^{(2)}$ are equal, completing the proof.

□

Consequently, we have found another optimal probabilistic protocol than the one Vidal proposed in [6], this time involving the OCR between states.

To sum up, in this chapter we have used the theory of majorization lattice to interpret results obtained in previous chapters and we have proposed a protocol making use of lattice notions, which we have proved to be equivalent to the seminal optimal probabilistic protocol of Vidal.

Conclusion

In this Ms thesis, we have mainly investigated a way of extending Nielsen's theorem to mixed states (a problem known to be hard) based on weak majorization and the operator Schmidt (OS) decomposition, an approach which had never been conducted until now to our knowledge. Let alone a brief use of mixed states in the end of Chapter 5, we only focused on pure states because, in order for a general criterion to be valid for mixed states, it is first necessary to hold for the special case of pure states. Nevertheless, some interesting results lead us to believe that the study of these concepts should be pursued further.

First, we conjectured in Chapter 4 a necessary condition for the deterministic conversion between bipartite pure states using weak-majorization, more specifically super-majorization, on the OS coefficients of both states. The first part of the chapter was devoted to the proof for qubits and the partial proof for qudits, as well as the exposition of difficulties encountered in the attempt of proving such relation. Then, the second part of the chapter was dedicated to numerical evidences strongly suggesting the validity of the above conjecture.

Building on this, we proposed in Chapter 5, in a similar approach to Chapter 4, a sufficient condition for probabilistic conversion still based on the OS decomposition. This led us to better identify the position of our super-majorization relation on the OS coefficients for pure states in regards to conditions on deterministic and probabilistic conversions already described by Nielsen's and Vidal's theorems. At the end of the chapter, we first used the OS decomposition for mixed states involved in Vidal's probabilistic protocol and showed the usefulness of super-majorization in this context.

Finally, in Chapter 6, we introduced the mathematical notion of majorization lattice which generates a graphical representation of distributions according to how they compare with one another under majorization. Thereafter, we presented in a unified picture the conjecture proposed in Chapter 4 and the theorem proposed in Chapter 5 by means of a numerical representation of the majorization lattice for qutrits. Then, the important lattice concepts of meet and join have been connected to entanglement transformations by interpreting the meet as an optimal common resource (OCR) state and the join as an optimal common product (OCP) state for such transformations. This enabled us to propose a theoretical quantum information scenario involving both states in order to understand one of many possible applications they could have. Lastly, we presented a specific protocol for bipartite state conversions making use of the OCR between two states and proved its equivalence with Vidal's probabilistic protocol, leading us to believe that the OCR and OCP have deep properties still to be found in entanglement transformations.

In conclusion, we have seen in this Ms thesis applications of weak majorization and majorization lattice to entanglement transformations, hinting at us that using the theory of majorization in entanglement theory remains relevant for future research.

Bibliography

[1] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935.

[2] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195–200, 1964.

[3] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, 1993.

[4] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69:2881–2884, 1992.

[5] M. A. Nielsen. Conditions for a Class of Entanglement Transformations. *Phys. Rev. Lett.*, 83(2):436–439, 1999.

[6] G. Vidal. Entanglement of pure states for a single copy. *Phys. Rev. Lett.*, 83:1046–1049, 1999.

[7] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. Springer New York, 2011.

[8] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[9] A. Rényi. On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1*, pages 547–561, 1961.

[10] F. Cicalese and U. Vaccaro. Supermodularity and subadditivity properties of the entropy on the majorization lattice. *IEEE Transactions on Information Theory*, 48:933–938, 2002.

[11] C. Guo, E. Chitambar, and R. Duan. Common Resource State for Preparing Multipartite Quantum Systems via Local Operations and Classical Communication. 2016. arXiv:1601.06220 [quant-ph].

[12] G. M. Bosyk, H. Freytes, G. Bellomo, and G. Sergioli. The lattice of trumping majorization for 4D probability vectors and 2D catalysts. *Sci Rep*, 8(1):3671, 2018.

[13] A. Hertz, M. Arnhem, A. Asadian, and N. J. Cerf. Realignment separability criterion assisted with filtration for detecting continuous-variable entanglement. *Phys. Rev. A*, 104(2):022427, 2021.

[14] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[15] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Mécanique Quantique - Tome 1*. CNRS/EDP Sciences, 2018.

[16] O. Gittsovich, O. Gühne, P. Hyllus, and J. Eisert. Unifying several separability conditions using the covariance matrix criterion. *Phys. Rev. A*, 78(5):052319, 2008.

[17] R. Garcia-Patron, C. Navarrete-Benlloch, S. Lloyd, J. Shapiro, and N. J. Cerf. Majorization theory approach to the gaussian channel minimum entropy conjecture. *Phys. Rev. Lett.*, 108:110505, 2012.

[18] G. De Palma, D. Trevisan, and V. Giovannetti. Passive states optimize the output of bosonic gaussian quantum channels. *IEEE Transactions on Information Theory*, 62:2895, 2015.

[19] M. A. Nielsen and J. Kempe. Separable states are more disordered globally than locally. *Phys. Rev. Lett.*, 86:5184–5187, 2001.

[20] Z. Puchała, Ł. Rudnicki, and K. Życzkowski. Majorization entropic uncertainty relations. *Journal of Physics A: Mathematical and Theoretical*, 46(27):272002, 2013.

[21] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner. The second laws of quantum thermodynamics. *Proceedings of the National Academy of Sciences*, 112(11):3275–3279, 2015.

[22] Z. Van Herstraeten, M. G. Jabbour, and N. J. Cerf. Continuous majorization in quantum phase space. 2021. arXiv:2108.09167 [quant-ph].

[23] G. M. Bosyk, G. Bellomo, F. Holik, H. Freytes, and G. Sergioli. Optimal common resource in majorization-based resource theories. *New J. Phys.*, 21(8):083028, 2019.

[24] D. P. Bertsekas. *Convex Optimization Theory*. Athena Scientific, 2009.

[25] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 2006.

[26] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, 2009.

[27] H.-K. Lo and S. Popescu. Concentrating entanglement by local actions: Beyond mean values. *Phys. Rev. A*, 63(2):022301, 2001.

[28] H. Aschauer and H. J. Briegel. *Quantum Communication and Decoherence*. Springer Berlin Heidelberg, 2002.

[29] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller. Quantum repeaters: The role of imperfect local operations in quantum communication. *Phys. Rev. Lett.*, 81:5932–5935, 1998.

[30] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, 1996.

[31] M. A. Nielsen and G. Vidal. Majorization and the interconversion of bipartite states. *Quantum Information & Computation*, 1:76–93, 2001.

[32] E. Chitambar and G. Gour. Quantum resource theories. *Rev. Mod. Phys.*, 91(2):025001, 2019.

[33] D. Jonathan and M. B. Plenio. Entanglement-Assisted Local Manipulation of Pure Quantum States. *Phys. Rev. Lett.*, 83(17):3566–3569, 1999.

[34] G. M. Bosyk, G. Sergioli, H. Freytes, F. Holik, and G. Bellomo. Approximate transformations of bipartite pure-state entanglement from the majorization lattice. *Physica A: Statistical Mechanics and its Applications*, 473:403–411, 2017.

[35] G. Grätzer. *Lattice Theory: First Concepts and Distributive Lattices.* W. H. Freeman, 1971.

[36] K. Korzekwa. Structure of the thermodynamic arrow of time in classical and quantum theories. *Phys. Rev. A*, 95(5):052318, 2017.

Appendix A

Proof of Nielsen's theorem

This proof is based on [14]. First, we recall the statement of Theorem 3.1.

Theorem A.1 (**Nielsen**). *Let $|\psi\rangle$ and $|\phi\rangle$ be two bipartite pure states, with $\boldsymbol{\lambda}$ the vector of Schmidt coefficients of $|\psi\rangle$ and $\boldsymbol{\eta}$ the vector of Schmidt coefficients of $|\phi\rangle$. Then, $|\psi\rangle$ is convertible with certainty into $|\phi\rangle$ using only LOCC if and only if the vector of Schmidt coefficients of $|\psi\rangle$ is majorized by that of $|\phi\rangle$, i.e.*

$$|\psi\rangle \xrightarrow{LOCC} |\phi\rangle \Longleftrightarrow \boldsymbol{\lambda} \prec \boldsymbol{\eta}. \tag{A.1}$$

Proof. To begin, assume $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$. Therefore, $|\psi\rangle$ can be transformed into $|\phi\rangle$ by letting first Alice perform a measurement on her system described by the measurement operators M_j^A. Then, by transmitting to Bob the result j she obtained, he can apply a unitary operation described by operators U_j^B to complete the transformation.

Now, from Alice's point of view, her system was first in state $\hat{\rho}_\psi^A$ and finally in state $\hat{\rho}_\phi^A$ because of her measurement. Therefore, by making use of the postulate on the measure (1.46), we write

$$\hat{\rho}_\phi^A = \frac{\hat{M}_j^A \hat{\rho}_\psi^A \hat{M}_j^{A\dagger}}{\mathrm{Tr}\Big(\hat{M}_m^{A\dagger} \hat{M}_m^A \hat{\rho}_\psi^A\Big)}, \tag{A.2}$$

where $\mathrm{Tr}\Big(\hat{M}_m^{A\dagger} \hat{M}_m^A \hat{\rho}_\psi^A\Big)$ is simply the probability of the outcome j, which we will denote as p_j. Furthermore, because all the reasoning is made on Alice's side, we will omit the superscript A in the following. Thus, we rewrite

$$\hat{\rho}_\phi = \frac{\hat{M}_j \hat{\rho}_\psi \hat{M}_j^\dagger}{p_j}. \tag{A.3}$$

Next, we make use of the polar decomposition of a matrix which allows to write, for any matrix A,

$$A = \sqrt{AA^\dagger} U, \tag{A.4}$$

where U is some unitary matrix.

Polar decomposing $\hat{M}_j \sqrt{\hat{\rho}_\psi}$, we get

$$\hat{M}_j \sqrt{\hat{\rho}_\psi} = \sqrt{\hat{M}_j \sqrt{\hat{\rho}_\psi} \sqrt{\hat{\rho}_\psi} \hat{M}_j^\dagger} \hat{V}_j, \tag{A.5}$$

$$= \sqrt{\hat{M}_j \hat{\rho}_\psi \hat{M}_j^\dagger} \hat{V}_j, \tag{A.6}$$

$$\overset{A.3}{=} \sqrt{p_j \hat{\rho}_\phi} \hat{V}_j, \tag{A.7}$$

where $\hat{V}_j$ is some unitary operator.

Now, by left multiplying the last equation by its adjoint, we get

$$\sqrt{\hat{\rho}_\psi}\hat{M}_j^\dagger \hat{M}_j \sqrt{\hat{\rho}_\psi} = p_j \hat{V}_j^\dagger \hat{\rho}_\phi \hat{V}_j. \tag{A.8}$$

Finally, summing on j and using the completeness relation for measurement operators $\sum_j \hat{M}_j^\dagger \hat{M}_j = \hat{I}$ (1.44),

$$\hat{\rho}_\psi = \sum_j p_j \hat{V}_j^\dagger \hat{\rho}_\phi \hat{V}_j, \tag{A.9}$$

which is nothing but $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$ because of the following theorem [14] and Property (1.1).

Theorem A.2 (Uhlmann). *$A \prec B$ if and only if there exists a set of unitary matrices $\{U_i\}$ and probabilities p_i such that $A = \sum_i p_i U_i B U_i^\dagger$.*

Let us now prove the converse statement. Assume $\boldsymbol{\lambda} \prec \boldsymbol{\eta}$, hence $\hat{\rho}_\psi \prec \hat{\rho}_\phi$. Again, using Theorem A.2, we can say that there exist probabilities p_j and unitary operators $\hat{U}_j$ such that $\hat{\rho}_\psi = \sum_j p_j \hat{U}_j^\dagger \hat{\rho}_\phi \hat{U}_j$.

Then, we define Alice's measurement operators $\hat{M}_j$ such that

$$\hat{M}_j \sqrt{\hat{\rho}_\psi} = \sqrt{p_j \hat{\rho}_\phi} \hat{U}_j. \tag{A.10}$$

To check that these operators are measurement operators, we verify the completeness relation (1.44). Let us first left multiply the previous equation by its adjoint, then sum on j to obtain

$$\sum_j \sqrt{\hat{\rho}_\psi}\hat{M}_j^\dagger \hat{M}_j \sqrt{\hat{\rho}_\psi} = \sum_j p_j \hat{U}_j \hat{\rho}_\phi \hat{U}_j^\dagger = \hat{\rho}_\psi, \tag{A.11}$$

which is obviously true if and only if $\sum_j \hat{M}_j^\dagger \hat{M}_j = \hat{I}$.

If now Alice performs a measurement with those operators $\hat{M}_j$, the outcome she gets is

$$\hat{\rho}_j \propto \hat{M}_j \hat{\rho}_\psi \hat{M}_j^\dagger \overset{A.10}{=} p_j \hat{\rho}_\phi, \tag{A.12}$$

implying that $\hat{\rho}_j = \hat{\rho}_\phi$ because states are normalized.

Finally, by means of the following lemma [14], we find that Bob can retrieve state $|\phi\rangle$ by applying a certain unitary operator $\hat{V}_j$.

Lemma A.1. *Let $|AR_1\rangle = \sum_i \sqrt{p_i}\,|i_A\rangle\,|i_{R_1}\rangle$ and $|AR_2\rangle = \sum_i \sqrt{p_i}\,|i_A\rangle\,|i_{R_2}\rangle$ be two purifications of a state $\hat{\rho}^A = \sum_i p_i\,|i_A\rangle\,\langle i_A|$ to a composite system AR. Then, there exists a unitary transformation $\hat{U}_R$ acting on system R such that*

$$|AR_1\rangle = (\hat{I} \otimes \hat{U}_R)\,|AR_2\rangle\,. \tag{A.13}$$

We can show that $\hat{U}_R = \sum_i |i_{R_1}\rangle\,\langle i_{R_2}|\,.$

□

Appendix B

Vidal's explicit conversion strategy

In this appendix, we detail Vidal's probabilistic protocol [6].

The strategy of probabilistically transforming $|\psi\rangle$ into $|\phi\rangle$ via LOCC transformations can be split into two stages. First, the deterministic conversion of $|\psi\rangle$ to an intermediary state $|\chi\rangle$. Then, the local two-outcome measurement of $|\chi\rangle$ leading to $|\phi\rangle$ with a certain probability $p_{\max}$ (3.19).

Without loss of generality, let us suppose that $|\psi\rangle$ is a bipartite state (with a vector of SC $\boldsymbol{\lambda}$) living in a $(d \times d)$-dimensional system and that $|\phi\rangle$ is a bipartite state (with a vector of SC $\boldsymbol{\eta}$) living in a $(d' \times d')$-dimensional system with $d \geq d'$. Indeed, it can easily be shown with (3.19) that, if $d < d'$, then the probability of success of the conversion is 0.

B.1 Intermediate state

We first describe the construction of the intermediary state $|\chi\rangle$. The strategy goes iteratively as follows.

Let $t_0 = 0$ and construct some values r_i ($i = 1$ in the first iteration),

$$r_i = \min_{t \in [t_{i-1}+1, d]} \frac{S_t^{\uparrow}(\boldsymbol{\lambda})}{S_t^{\uparrow}(\boldsymbol{\eta})} = \frac{S_{t_i}^{\uparrow}(\boldsymbol{\lambda})}{S_{t_i}^{\uparrow}(\boldsymbol{\eta})}. \tag{B.1}$$

As long $t_i \neq d$, repeat the above procedure for $i \leftarrow i + 1$. Finally, when $t_i = d$, call this last i, "k". Thus, we have obtained a $k + 1$ series $0 = t_0 < t_1 < ... < t_k = d$ and a k series $0 < r_1 < ... < r_k$.

Hence, we can now construct the following vector

$$\boldsymbol{\gamma} = \begin{bmatrix} r_k \begin{pmatrix} \lambda^{\uparrow}_{t_k} \\ \vdots \\ \lambda^{\uparrow}_{t_{k-1}+1} \end{pmatrix} \\ \vdots \\ r_2 \begin{pmatrix} \lambda^{\uparrow}_{t_2} \\ \vdots \\ \lambda^{\uparrow}_{t_1+1} \end{pmatrix} \\ r_1 \begin{pmatrix} \lambda^{\uparrow}_{t_1} \\ \vdots \\ \lambda^{\uparrow}_{t_0+1} \end{pmatrix} \end{bmatrix} . \tag{B.2}$$

This will be the vector of SC of the intermediary state $|\chi\rangle$. One can check that, by construction, $\boldsymbol{\lambda} \prec \boldsymbol{\gamma}$, hence $|\psi\rangle \xrightarrow{LOCC} |\chi\rangle$ by Theorem 3.1.

B.2 Measurement

Now that Alice and Bob possess the state $|\chi\rangle$, one of them will perform a two-outcome measurement described by operators $\{\hat{M}, \hat{N}\}$. We will construct $\hat{M}$ such that if it is measured, the state is left in the desired state $|\phi\rangle$.

$\hat{M}$ is structured as follows

$$\hat{M} = \begin{bmatrix} \hat{M}_k & & & \\ & \ddots & & \\ & & \hat{M}_2 & \\ & & & \hat{M}_1 \end{bmatrix} = \hat{M}^{\dagger}, \tag{B.3}$$

where

$$\hat{M}_j = \sqrt{\frac{r_1}{r_j}} \hat{I}_{[t_j - t_{j-1}]} \text{ with } 1 \leq j \leq k, \tag{B.4}$$

is an operator proportional to a $[t_j - t_{j-1}]$-dimensional identity operator, hence $\hat{M}$ is diagonal.

Then, $\hat{N}$ is defined as $\hat{N} = \sqrt{1 - \hat{M}^2}$ such that the two operators $\{\hat{M}, \hat{N}\}$ form a generalized measurement set[1].

Finally, one can show that if one party, let's say Alice, measures $\hat{M}$, the final state is indeed $|\phi\rangle$ with probability r_1, i.e.

$$\frac{\hat{M} \otimes \hat{I}}{\sqrt{r_1}} = |\phi\rangle . \tag{B.5}$$

Furthermore, this allows to prove the value of the conversion probability of Theorem 3.3 because $r_1 = p_{\max}$.

[1]by construction of $\hat{N}$, the completeness relation (1.44) is obeyed.

One last observation that can be made is the fact that if, on the contrary, Alice measures $\hat{N}$, the final state, let us call it $|\xi\rangle$, has a Schmidt rank strictly lower than $|\phi\rangle$. Therefore, one cannot hope to find a way to use LOCC to transform $|\xi\rangle$ into $|\phi\rangle$, hence the conversion has failed.

Appendix C

Optimal probability for any conversion

The following theorem allows one to determine the optimal conversion probability with which a fixed initial d-dimensional bipartite pure state can be converted any d-dimensional (or less) bipartite pure state.

Theorem C.1. *Let $|\psi\rangle$ be a d-dimensional bipartite pure state, with $\boldsymbol{\lambda}$ its vector of Schmidt coefficients (SC). Then, $|\psi\rangle$ is convertible using only LOCC into any other d-dimensional bipartite pure state $|\phi\rangle$ (vector of SC $\boldsymbol{\eta}$) with maximum probability $p_{\max} = d\lambda_d$, i.e.*

$$\boldsymbol{\lambda} \prec^w p\boldsymbol{\eta},\ \forall \boldsymbol{\eta} \iff p \leq p_{\max} = d\lambda_d. \tag{C.1}$$

Proof. In order for $|\psi\rangle$ to be able to be LOCC transformed into any other d-dimensional state $|\phi\rangle$, it is necessary and sufficient that it can be transformed into the maximally-entangled state $|\alpha\rangle$, with vector of SC $\boldsymbol{\alpha} = (\frac{1}{d}, ..., \frac{1}{d})$. Hence, we show what is the maximum probability for which the probabilistic conversion of $|\psi\rangle$ into $|\alpha\rangle$ using LOCC is possible, i.e. what is the maximum value p may take in $\boldsymbol{\lambda} \prec^w p\boldsymbol{\alpha}$.

The latter super-majorization relation can be developed using (2.16),

$$S_i^{\uparrow}(\boldsymbol{\lambda}) \geq pi\frac{1}{d}, \quad \forall i \in [1, d], \tag{C.2}$$

which can be rewritten

$$p \leq \frac{dS_i^{\uparrow}(\boldsymbol{\lambda})}{i}, \quad \forall i \in [1, d]. \tag{C.3}$$

Because $\lambda_d \leq \lambda_{d-1} \leq ... \leq \lambda_1$, we obtain

$$p \leq dS_1^{\uparrow}(\boldsymbol{\lambda}) = d\lambda_d = \frac{d}{2}(\lambda_d + \lambda_d) \leq \frac{d}{2}(\lambda_d + \lambda_{d-1}) = \frac{dS_2^{\uparrow}(\boldsymbol{\lambda})}{2}, \tag{C.4}$$

meaning that satisfying (C.3) for $i = 1$ immediately satisfies (C.3) for $i = 2$.

This can be generalized to any other value $i \in [1, d]$

$$p \leq dS_1^{\uparrow}(\boldsymbol{\lambda}) = d\lambda_d = \frac{d}{i}\underbrace{(\lambda_d + ... + \lambda_d)}_{\text{i times}} \leq \frac{d}{i}(\lambda_d + \lambda_{d-1} + ... + \lambda_{d-i+1}) = \frac{dS_i^{\uparrow}(\boldsymbol{\lambda})}{i}, \tag{C.5}$$

completing the proof by taking the value $p_{\max} = d\lambda_d$. □