universität innsbruck

Bachelor thesis

Quantum state discrimination and distillability of quantum states

Jasmin Matti

July 4, 2023

Supervised by Univ.-Prof. Dipl.-Math. Dr. Tim Netzer
Institute for Mathematics

Contents

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.
Ich erkläre mich mit der Archivierung der vorliegenden Bachelorarbeit einverstanden.

Innsbruck, am July 4, 2023 Jasmin Matti

1 Introduction

Measurements are part of our everyday life, be it to determine the length of an object with a ruler, to weigh flour while baking cakes or to stop the time for a $100\,\mathrm{m}$ run. In all these cases, the effects of the measurement apparatus are often negligible or can be minimized. Moreover, if an idealized classical measurement process is repeated many times, one would always obtain the same outcome. In quantum mechanics, however, measurements work quite differently. In quantum mechanics, the outcomes of measurements can not be predicted with certainty and if the same measurement on an object is done twice, the outcomes can vary.

POVMs (Positive Operator Valued Measures) are the most general kind of quantum measurements. They can be used to perform the task of *quantum state discrimination* and also play a fundamental role in *entanglement distillation*. Both of these aspects will be discussed in this work and play an important role in quantum information processing, such as quantum cryptography.

In quantum cryptography, two parties want to share a secret message with one another. To do so, the receiving party should determine, by suitable measurements, the state submitted by the transmitting party (and hence the intended message) out of a set of possible states, which all have been prepared with known probabilities [1]. That is exactly what quantum state discrimination describes. Quantum state discrimination is the process of distinguishing between different quantum states using measurements.

But how can the efficiency of a POVM to perform quantum state discrimination be quantified? And can POVMs be approximated by POVMs with few outcomes, what would make the task of quantum state discrimination less complicated?

Key insights to these questions can be obtained by investigating the relation between POVMs and special convex polytopes, called *zonotopes* and *zonoids*. Using fundamental results for zonotopes and zonoids from convex geometry, important insights on quantum measurements and quantum state discrimination can be gained, which is the focus of the first part of this thesis.

Further, in quantum cryptography, maximally entangled states are needed in order to guarantee secure communication. In practice, however, quantum states are usually less entangled or partially mixed, due to decoherence effects. That is where entanglement distillation comes in, which is dealt with in the second part of this thesis. With entanglement distillation, a large number of mixed entangled states can be converted into a smaller number of maximally entangled pure states using local operations (e.g. POVMs) and classical communication.

But for which states does such a distillation process work? This question is also referred to as the distillability problem and the focus of the second part of this thesis.

This thesis is structured as follows. Firstly, a mathematical introduction to Hilbert spaces and operators is given (2). Further, density matrices and important properties of quantum states are discussed (3). This is followed by a convex geometry chapter about zonotopes and zonoids (4), since there exists a connection between these geometrical objects and quantum measurements. Especially, the approximation of zonoids by zonotopes is outlined (4.2). Next, quantum measurements are discussed (5) with a specific focus on quantum state discrimination (5.2) and the relation between zonotopes and POVMs (5.2.1). Then, quantum channels and Werner states (6) are described. The main chapter in the second part of this thesis focuses on the distillation of entanglement (7). We start by outlining the dichotomy between separability and entanglement (7.1). Next, we ask ourselves, how to measure the proximity between two quantum states and introduce the concept of fidelity (7.1.1). Furthermore, distillation is defined and the distillability problem mathematically formulated (7.2). Finally, the distillability of two qubit systems is analyzed (7.2.2).

2 Notation and useful definitions

Quantum mechanics is a mathematical-physical theory that describes the physical properties of matter at the atomic and subatomic scale. Understanding the mathematical foundations is essential to grasp every physical theory. Therefore, this first chapter describes the mathematical foundations for comprehending quantum objects and the used notation. Firstly, Hilbert and Euclidean spaces are studied (2.1), followed by a description of the Dirac notation (2.2). The chapter ends with a discussion of operators and matrices (2.3).

2.1 Hilbert and Euclidean spaces

To gain a better understanding of quantum mechanics, the structure of the Hilbert space, which underlies our physical objects, is studied. We begin by examining the definitions.

Definition 2.1.1 (Pre-Hilbert space)**.** A normed vector space $(V, ||-||)$ is called pre-Hilbert space if an inner product $\langle -, - \rangle$ is defined on V, which induces the norm $||-||$, e.g., $||x||^2 = \langle x, x \rangle$ for $x \in V$.

Definition 2.1.2 (Hilbert space)**.** A normed vector space $(V, ||-||)$ is called Hilbert space, if $(V, ||-||)$ is a complete pre-Hilbert space.

In a complex Hilbert space $\mathcal{H}$, we adopt the convention that the inner product is linear in the second argument and conjugate linear in the first argument, i.e., for $\psi, \phi \in \mathcal{H}$ and $\lambda \in \mathbb{C}$ [2]:

$$\langle \lambda\psi, \phi \rangle = \bar{\lambda}\langle \psi, \phi \rangle$$
$$\langle \psi, \lambda\phi \rangle = \lambda\langle \psi, \phi \rangle.$$

Throughout this Bachelor thesis, all the considered normed spaces will be finite-dimensional.

Tensor products

Tensor products are essential for describing composite physical systems. For real or complex finite-dimensional Hilbert spaces $(\mathcal{H}_i)_{1 \leq i \leq k}$, we consider the tensor product over the real or complex field, respectively,

$$\mathcal{H} = \bigotimes_{i=1}^{k} \mathcal{H}_i = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes ... \otimes \mathcal{H}_k.$$

This tensor product is also called multipartite Hilbert space. For $k = 2$ we call it bipartite Hilbert space [2].

2.2 Bra-ket notation

As the Dirac notation, introduced by Dirac in 1939 (see [3]), is used throughout this Bachelor thesis, the following section outlines the basics of this notation.
Notation 1 A vector ψ in a Hilbert space $\mathcal{H}$ is denoted by a *ket* and represented as $|\psi\rangle$. As $\mathcal{H}$ is a vector space, it has a dual vector space $\mathcal{H}^*$, which is known as the space of linear functionals over the vector space. An element of the dual Hilbert space $\mathcal{H}^*$ is called a *bra* and represented as $\langle\phi|$. Each ket $|\psi\rangle$ has exactly one corresponding dual vector $\langle\psi|$ and vice versa. Technically, the hermitian conjugation transforms a vector into its dual vector

$$|\psi\rangle^\dagger = \langle\psi|.$$

The result of applying the bra $\langle\phi|$ to the ket $|\psi\rangle$ is referred to as the bracket (bra-ket) of the two vectors $\phi, \psi \in \mathcal{H}$. It represents the inner product of ϕ and ψ and is denoted by $\langle\phi|\psi\rangle$.
For an operator $A \in B(\mathcal{H})$ and a vector $\phi \in \mathcal{H}$, we can form the linear functional $\langle\phi|A$, i.e., the linear map $\psi \mapsto \langle\phi|A\psi\rangle$, which is generally written as

$$\langle\phi|A|\psi\rangle.$$

Notation 2 For any $\phi, \psi \in \mathcal{H}$, the expression $|\phi\rangle\langle\psi|$ denotes the linear operator (see section 2.3 for operator definition) on $\mathcal{H}$ given by

$$(|\phi\rangle\langle\psi|)(\chi) = |\phi\rangle\langle\psi|\chi\rangle = \langle\psi|\chi\rangle|\phi\rangle$$

which is also called the exterior product. In mathematical notation, $|\phi\rangle\langle\psi|$ is the operator sending χ to $\langle\psi, \chi\rangle\phi$.
If the vectors are chosen to be each other's dual vectors ($|\psi\rangle$) respectively, we get the projection operators

$$P := |\psi\rangle\langle\psi|$$

with the property $P^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = P$ [4, 5].

2.3 Operators and matrices

In quantum mechanics, linear operators on vector spaces are used to represent physical quantities and are typically defined on Hilbert spaces. This section gives a short introduction to operators and superoperators.

Definition 2.3.1 (Operator [6])**.** A map $\psi : V_1 \to V_2$ between two normed spaces is called an operator.

If the map in Definition 2.3.1 is linear, the operator is called a linear operator.
For finite-dimensional Hilbert spaces $\mathcal{H}, \mathcal{H}'$, we denote by $B(\mathcal{H}', \mathcal{H})$ the space of linear operators from $\mathcal{H}'$ to $\mathcal{H}$. Furthermore, the notation $B(\mathcal{H})$ is used for linar operators from $\mathcal{H}$ to $\mathcal{H}$.

Definition 2.3.2 (Adjoint operator [2, 7])**.** Consider an operator $A \in B(\mathcal{H}', \mathcal{H})$. The unique operator $A^\dagger \in B(\mathcal{H}, \mathcal{H}')$ which satisfies for all vectors $\psi \in \mathcal{H}, \phi \in \mathcal{H}'$

$$\langle \psi, A\phi \rangle = \langle A^\dagger \psi, \phi \rangle$$

is called the adjoint or Hermitian conjugate of the operator A.

The given definition implies $(AB)^\dagger = B^\dagger A^\dagger$. The adjoint operator $A^\dagger$ is obtained by transposing and then complex conjugating A. We denote by $B^{sa}(\mathcal{H})$ the space of self-adjoint operators satisfying $A^\dagger = A$. $B^{sa}(\mathcal{H})$ is a real vector subspace of $B(\mathcal{H})$.
Every measurable quantity in a physical experiment (observable) is associated with a self-adjoint linear operator.
Next, we want to define positive operators. Using positive operators, we will be able to define quantum states in Chapter 3. From now on, the normed spaces V_1 should be pre-Hilbert spaces.

Definition 2.3.3 (Positive operator [7])**.** An operator $A : V_1 \to V_1$ is said to be positive if $A = A^\dagger$ and for every $x \in V_1$, $\langle x, Ax \rangle = \langle x|A|x \rangle \geq 0$ holds.

A positive operator has non-negative eigenvalues.
If $\langle x|A|x \rangle$ is strictly greater than zero for all $x \neq 0$ we say that A is positive definite [7].

Superoperators

The term superoperator is used to denote linear maps acting between spaces of operators or between spaces of matrices. In order to indicate the space on which the identity map is defined, two different notations are employed [2]:

- $\mathrm{I}_{\mathcal{H}}$ is the identity operator on a Hilbert space (if $\mathcal{H} = \mathbb{C}^n$ or $\mathcal{H} = \mathbb{R}^n$ we use I_n)
- $\mathrm{Id}_{B(\mathcal{H})}$ is the identity superoperator on $B(\mathcal{H})$ (sometimes simply Id).

Trace

Similarly to the definition of the trace of a matrix as the sum of its diagonal elements, the trace of an operator A is defined as the trace of any matrix representation of A [7].

Definition 2.3.4 (Trace [8])**.** The trace of an operator A acting on a Hilbert space $\mathcal{H}$ is defined as

$$\mathrm{Tr}(A) = \sum_i \langle i|A|i \rangle$$

where $\{|i\rangle\}$ is some complete, orthonormal basis of $\mathcal{H}$.

The trace operator is linear and cyclic.

Partial trace

The partial trace is an important mathematical operation in quantum mechanics and used to only trace over a part of a bipartite system. Unlike the trace, which has a scalar as an output, the partial trace has an operator as an output which lives on a smaller Hilbert space.

Definition 2.3.5 (Partial trace [2])**.** Consider a bipartite Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. The partial trace over $\mathcal{H}_2$ is the map $\mathrm{Tr}_{\mathcal{H}_2} : B(\mathcal{H}_1) \otimes B(\mathcal{H}_2) \to B(\mathcal{H}_1)$ given by

$$Id_{B(\mathcal{H}_1)} \otimes \mathrm{Tr}.$$

The partial trace acts on product operators the following way:

$$\mathrm{Tr}_{\mathcal{H}_2}(A \otimes B) = (\mathrm{Tr}B)A$$

for $A \in B(\mathcal{H}_1), B \in B(\mathcal{H}_2)$. Analogously, one can define the partial trace with respect to $\mathcal{H}_1$.

Matrices

In the following chapters, we denote the space of $m \times n$ matrices by $\mathrm{M}_{m,n}$ and for $m = n$ by M_n. The matrix entries of $M \in \mathrm{M}_{m,n}$ get denoted by $(m_{i,j})_{1\le i\le m, 1\le j\le n}$. The Hermitian conjugate of M, i.e., $(m_{ij})^\dagger = (\bar{m}_{ji})$ gets denoted by $M^\dagger$. The subspace of M_m consisting of Hermitian (self-adjoint) matrices (matrices $M \in \mathrm{M}_m$ for which $M = M^\dagger$) will be denoted by M_m^{sa}.
Complex $m \times n$ matrices can be identified with operators from $\mathbb{C}^n$ to $\mathbb{C}^m$ and so we write $\mathrm{M}_{m,n} = B(\mathbb{C}^n, \mathbb{C}^m)$, $\mathrm{M}_n = B(\mathbb{C}^n)$ and $\mathrm{M}_n^{sa} = B^{sa}(\mathbb{C}^n)$.
As we will frequently work with matrices and operators, it is useful to have a good grasp of the Frobenius inner product, which is an inner product on the linear space of real or complex matrices.

Definition 2.3.6 (Frobenius inner product [9])**.** Consider the vector space $\mathbb{K}^{n\times n}$ over the real or complex field ($\mathbb{R}$,$\mathbb{C}$). The Frobenius inner product $\langle , \rangle_F : \mathbb{K}^{n\times n} \times \mathbb{K}^{n\times n} \mapsto \mathbb{K}$ is defined as:

$$\langle P, Q\rangle_F = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}q_{ij} = \mathrm{Tr}(P^T Q) \quad \text{for P, Q} \in \mathbb{R}^{n\times n}$$

$$\langle P, Q\rangle_F = \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{p_{ij}}q_{ij} = \mathrm{Tr}(P^\dagger Q) \quad \text{for P, Q} \in \mathbb{C}^{n\times n}.$$

3 Quantum states

As mentioned in the introduction (1), the goal of entanglement distillation is to obtain a small number of maximally entangled states from a large number of weakly entangled states. To understand entanglement distillation, it is essential to understand quantum states first. In this chapter, firstly, quantum states are defined, and the density operator language is motivated (3.1.1). Then, the Schmidt decomposition (3.1.2 gets discussed as a useful mathematical tool in quantum mechanics. Further, the partial transpose gets introduced (3.1.3), since it plays an important role in the distillability problem (see 7.2.1). At the end of this chapter, qubits are briefly explained (3.2).

3.1 Quantum states

Quantum states are used to describe observations in quantum systems and are fundamentally different from classical states. The mathematical description of quantum states is given in Definition 3.1.1.

Definition 3.1.1 (Quantum state [2])**.** A quantum state on a Hilbert space $\mathcal{H}$ is a positive self-adjoint operator of trace one. The set of states on $\mathcal{H}$ is denoted by $D(\mathcal{H})$.

This means that every quantum state can be represented as a hermitian matrix, that has trace one and non-negative eigenvalues.
If the state of a quantum system is exactly known, it is said to be in a pure state [7]. A pure state can be represented by a single vector $|\psi\rangle$ in the Hilbert space $\mathcal{H}$. If the state of a quantum system is not completely known, it is said to be in a mixed state, which is a statistical distribution of pure states. In order to describe such distributions, density operators are used.

3.1.1 Density operators

With the density operator language, quantum systems can be described, whose state is not completely known. Consider a quantum system in one of a number of pure states $|\psi_i\rangle$, where i is an index, with respective probabilities p_i. We call $\{p_i, |\psi_i\rangle\}$ an ensemble of pure states. The density operator for the system is defined by

$$\rho \equiv \sum_i p_i |\psi_i\rangle\langle\psi_i|.$$

In the density operator formulation, a pure state can be defined the following way.

Definition 3.1.2 (Pure quantum state [2])**.** A state $\rho \in D(\mathcal{H})$ is called pure if it has rank 1, i.e., if there is a unit vector $\psi \in \mathcal{H}$ such that

$$\rho = |\psi\rangle\langle\psi|.$$

Conversly, a mixed state ρ is a mixture of the different pure states in the ensemble for ρ [7].

Density operators can be characterized by the trace and positivity condition.

Theorem 3.1.1 (Characterization of the density operator [7])**.** An operator ρ is the density operator associated to an ensemble $\{p_i, |\psi_i\rangle\}$ iff it satisfies the following two conditions:

1. **Trace condition** ρ has trace equal to 1
2. **Positivity condition** ρ is a positive operator (see 2.3.3)

Proof. Suppose $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$ is a density operator, then

$$\mathrm{Tr}(\rho) = \sum_i p_i \mathrm{Tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i = 1.$$

So the first condition is satisfied.
For an arbitrary vector $|\phi\rangle$ in state space

$$\langle\phi|\rho|\phi\rangle = \sum_i p_i\langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle = \sum_i p_i|\langle\phi|\psi_i\rangle|^2 \geq 0$$

holds. So the second condition is also satisfied.
Conversely, suppose ρ is an arbitrary operator satisfying both conditions. Since ρ is positive it must have a spectral decomposition

$$\rho = \sum_j \lambda_j |j\rangle\langle j|$$

where $|j\rangle$ are orthogonal vectors and λ_j are real, non-negative eigenvalues of ρ. Using the trace condition $\sum_j \lambda_j = 1$ must hold. Thus, a system in state $|j\rangle$ with probability λ_j will have density operator ρ. □

3.1.2 Schmidt decomposition

The Schmidt decomposition is a useful mathematical tool and additionally provides insights into the nature of quantum entanglement.

Theorem 3.1.2 (Schmidt decomposition [7, 10])**.** Suppose the composite system is in a pure state $\Psi \in \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ described by the density operator $\rho = |\Psi\rangle\langle\Psi| \in D(\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B})$. Then there exists an orthonormal basis $\{|a_i\rangle \mid i = 1, 2, 3...\}$ in $\mathcal{H}_\mathrm{A}$ and $\{|b_i\rangle \mid i = 1, 2, 3...\}$ in $\mathcal{H}_\mathrm{B}$, such that

$$|\Psi\rangle = \sum_i s_i |a_i\rangle \otimes |b_i\rangle$$

or in terms of the density operator

$$\rho = \sum_{i,k} s_i s_k^* |a_i\rangle\langle a_k| \otimes |b_i\rangle\langle b_k|$$

where $\sum_i |s_i|^2 = 1$.

The proof is omitted and can be found in [10] for example. The numbers $(s_i, s_2, ..., s_d)$ are called Schmidt coefficients of $|\Psi\rangle$ and are uniquely determined if we require that $s_1 \geq s_2 \geq ... \geq s_d$ where $d = \min(d_1, d_2)$ with $d_1 = \dim(\mathcal{H}_A)$ and $d_2 = \dim(\mathcal{H}_B)$. The largest k such that $s_k > 0$ is called **Schmidt rank** of $|\Psi\rangle$. It can be easily verified that $s_1^2 + ... + s_d^2 = |\Psi|^2$.

3.1.3 Partial transpose

As we will see in section 7.2 the partial transpose of a state $\rho \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ is fundamental for characterizing the distillability of it.

Definition 3.1.3 (Partial transpose)**.** If $\mathcal{H}$ is a bipartite Hilbert space, and if T denotes the transposition on $B(\mathcal{H}_1)$ (with respect to a specified basis) and Id is the identity operator of $B(\mathcal{H}_2)$ then the partial transpose is the operation

$$\Gamma = T \otimes Id : B(\mathcal{H}_1 \otimes \mathcal{H}_2) \to B(\mathcal{H}_1 \otimes \mathcal{H}_2)$$

The partial transpose of a state $\rho \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ is denoted by $\rho^\Gamma = \Gamma(\rho)$.

Definition 3.1.4 (PPT state)**.** A state $\rho \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ has a positive partial transpose (PPT) if the operator ρ^Γ is positive.

A state $\rho \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ with a non-positive partial transpose, is called an NPT-state.

3.2 Qubits

Since the term qubit may appear in some places, it should be briefly explained.
In correspondence to a classical bit, which has either state 0 or state 1, a qubit also has a state. Two possible qubit states are the states $|0\rangle$ and $|1\rangle$ which correspond to the states 0 and 1 of a classical bit. What distinguishes qubits and bits is that a qubit can be in a state other than $|0\rangle$ or $|1\rangle$. Furthermore, linear combinations of states, called superpositions

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

can be formed, where α and β are complex numbers. Expressed differently, the state of a qubit is a vector in a two-dimensional complex vector space. The states $|0\rangle$ and $|1\rangle$ are the computational basis states, and form an orthonormal basis for this vector space [7].

4 Zonotopes and zonoids

In this chapter, we will explore some concepts of convex geometry, since there exists a connection to quantum measurements, which will be discussed in Chapter 5. Firstly, zonotopes and zonoids are introduced and the construction of zonotopes is illustrated (4.1). Further, the approximation of zonoids by zonotopes is discussed, which will be converted to a similar statement for POVMs in Chapter 5.

4.1 Zonotopes and zonoids

In this section, firstly, zonotopes are defined, and their construction is illustrated with a simple example. Then, zonoids, polar sets and extreme points are described.
Before describing zonotopes and zonoids, we introduce the Minkowski sum. The Minkowski sum of two subsets A and B of a linear space is formed by adding each element in A to each element in B, as described in Definition 4.1.1.

Definition 4.1.1 (Minkowski sum [2])**.** Given two sets $A, B \subset \mathbb{R}^n$, the Minkowski sum is defined by

$$A + B := \{x + y : x \in A, y \in B\}.$$

Using the Minkowski sum, zonotopes can be defined. Zonotopes are a special type of convex polytopes.

Definition 4.1.2 (Zonotope [2, 11])**.** A convex body $K \subset \mathbb{R}^n$ is called a zonotope if it is the Minkowski sum of finitely many segments (segments are compact one dimensional convex sets):

$$K = I_1 + I_2 + \ldots + I_m \tag{4.1}$$

where I_i is the line segment $[x_i, y_i]$ with $x_i, y_i \in \mathbb{R}^n$ for $i = 1, \ldots, m$. The line segments $[x_i, y_i]$ for $i = 1, \ldots, m$ are called the *generators* of the zonotope.

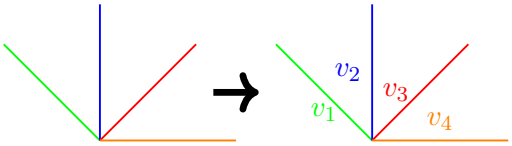
Zonotopes always have a center of symmetry, namely the sum of the centers of the segments I_j [12]. Every face of a zonotope is again a zonotope and therefore all faces of zonotopes have a center of symmetry (see [13] for proofs). In $\mathbb{R}^2$ centrally symmetric polytopes, i.e., convex polygons, are zonotopes. A simple example for a zonotope gives the cube $[-1, 1]^n$ since

$$[-1, 1]^n = [-e_1, e_1] + [-e_2, e_2] + \ldots + [-e_n, e_n],$$

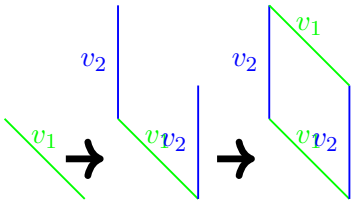
where $[-e_i, e_i]$ denotes the segment joining the ith canonical basis vector and its opposite [2]. Next, the construction of a simple zonotope is illustrated, to get some intuition for Definition 4.1.2.

Exemplary construction of zonotope

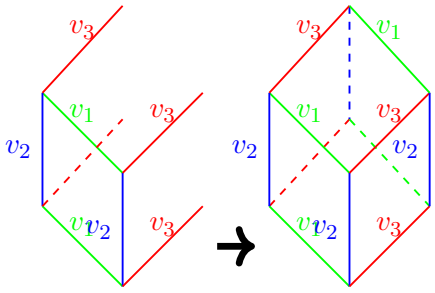
1. **We start with finitely many segments and draw them. Then we pick an order of our segments.**
In this example, we choose four line segments and pick the order $v_1 \equiv$ green, $v_2 \equiv$ blue, $v_3 \equiv$ red, $v_4 \equiv$ orange.



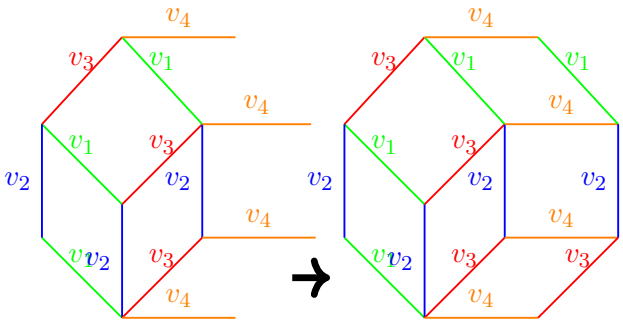
2. **Now we draw the first segment and then the second segment originating from each end of the first segment. Next, we connect the endpoints.**



3. **We repeat this process with the third segment. We draw the third segment from each vertex of the parallelogram and then connect the endpoints.**



4. **We repeat the process with the other remaining vectors.**
Note: To keep everything tidy, the invisible (dashed) edges were left out.



The boundary of the shape at the end is a zonotope [14].

Now, the Hausdorff distance is introduced as a measure for the distance of the subsets of a metric space.

Definition 4.1.3 (Hausdorff distance [15])**.** Let (M, d) be a metric space and $A, B \neq \varnothing$ be subsets of M. We define the distance between a point $x \in M$ and a non-empty subset $A \subseteq M$ as

$$d(x, A) := \inf_{a \in A} d(x, a)$$

Then the Hausdorff distance $d_H(A, B)$ between the two subsets A, B is given by

$$d_H(A, B) = \max\left\{\sup_{a \in A} d(a, B),\ \sup_{b \in B} d(b, A)\right\}. \tag{4.2}$$

With the Hausdorff distance, zonoids can be defined.

Definition 4.1.4 (Zonoid)**.** A convex body $K \subset \mathbb{R}^n$ is called a zonoid if it can be written as a limit of zonotopes in the Hausdorff distance.

Every zonoid is a zonotope and, like zonotopes, zonoids are also centrally symmetric. In Figure 4.1 a converging sequence of zonotopes is illustrated. The limit of such a sequence is a zonoid.

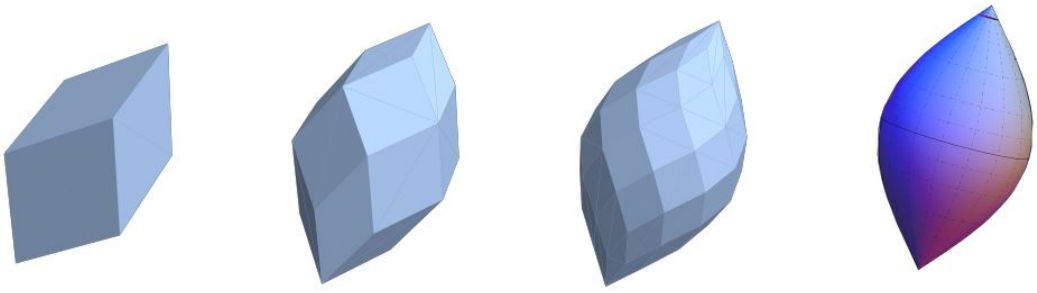


Figure 4.1: A sequence of zonotopes and its limit, a zonoid. : figure from [16].

Next, polar sets and extreme points are discussed, since they are needed for describing the relation between zonotopes and POVMs in Section 5.2.1.

Definition 4.1.5 (Minkowski functional)**.** Let K be a subset of $\mathbb{R}^n$. For $x \in \mathbb{R}^n$ the Minkowski functional of K is given by

$$||x||_K := \inf\{t \geq 0 : x \in tK\} \tag{4.3}$$

where $tK = \{tx : x \in K\}$.

Definition 4.1.6 (Polar set [2])**.** For $A \subset \mathbb{R}^n$, the polar of A is given by

$$A^\circ := \{y \in \mathbb{R}^n : \langle x, y\rangle \leq 1 \text{ for all } x \in A\}. \tag{4.4}$$

In particular, $||y||_{A^\circ} = \sup_{x \in A \cup \{0\}} \langle x, y\rangle$ holds.
An extreme point of a convex set K is a point that is interior for no interval contained in this set [17].

Definition 4.1.7 (Extreme point)**.** Let $K \subset \mathbb{R}^n$ be a convex set. A point $x \in K$ is said to be extreme if it cannot be written in a nontrivial way as a convex combination of points of K,i.e., if the equality $x = ty + (1-t)z$ for $t \in (0,1)$ and $y, z \in K$ implies that $x = y = z$.

The extreme points of a triangle are its vertices and the extreme points of a disk are the points of its boundary circle [17] as illustrated in Figure 4.2.

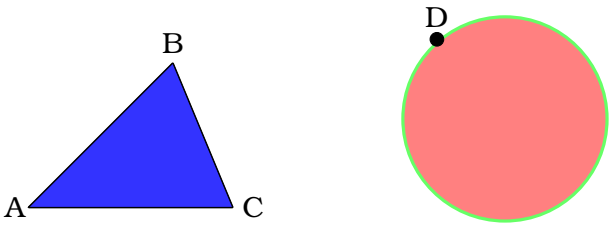


Figure 4.2: A, B, C are the extreme points of the triangle. D is an extreme point of the red circle.

4.2 Support function and approximation of zonoids by zonotopes

In this section, firstly, the support function is defined (Definition 4.2.1). Then, specifically, the support function of zonotopes is discussed. Further, signed measures and the Hausdorff measure are introduced, followed by a theorem (Theorem 4.9) which gives insight on the support function of zonoids. The section ends with a key result for the approximation of zonoids by zonotopes (4.2.2).

Definition 4.2.1 (Support function [2])**.** Given a nonempty and bounded set $K \subset \mathbb{R}^n$ and a vector $u \in \mathbb{R}^n$, we define the quantity

$$w(K,u) := \sup_{x \in K} \langle u, x \rangle \tag{4.5}$$

If $|u| = 1$, then $w(K,u)$ is called the support function of K in direction u and commonly denoted by $h(K,u)$.

If K is a convex body containing 0 in the interior, $w(K,u) = ||u||_{K^\circ}$ holds (cf. 4.1.6). Geometrically, $w(K,u)$ is the distance from the origin to the supporting hyperplane tangent to K in direction u, where u is normal to the hyperplane and outer to K, as illustrated in Figure 4.3.

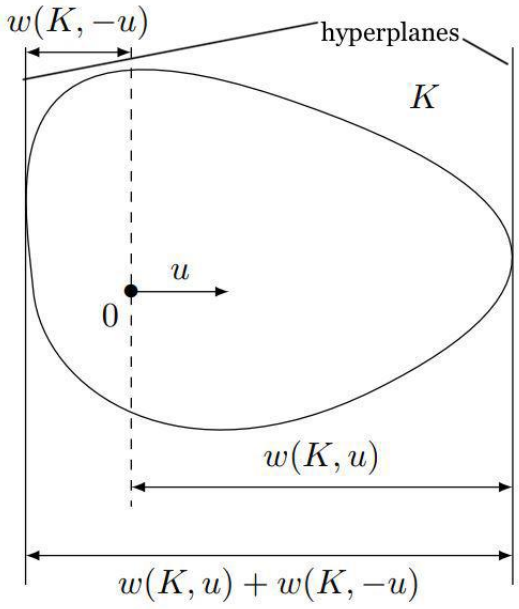


Figure 4.3: Illustration of the geometrical interpretation of $w(K,u)$ for a convex body K. For $|u| = 1$ the sum $w(K,u) + w(K,-u)$ is the width of the smallest strip in the direction orthogonal to u which contains K : figure from [2].

For a zonotope $Z = I_1 + I_2 + \ldots + I_m$ with $I_i = \text{conv}\{\alpha_i v_i, -\alpha_i v_i\}$ where $v_i \in S^{n-1}$ and $\alpha_i > 0$ for $i = 1, \ldots, m$, the support function of Z is given by

$$h(Z,\cdot) = \sum_{i=1}^{k} \alpha_i |\langle \cdot, v_i \rangle|. \tag{4.6}$$

Conversely, if a convex body Z has a support function of this form it is a zonotope with center at the origin [18].
Next, we want to generalize Equation 4.6 for zonoids (Theorem 4.2.1). To understand Theorem 4.2.1, we firstly need to introduce some measure notions.

Signed measures and the Hausdorff measure

A signed measure (measure) on S^{n-1} is a real-valued, σ–additive function on the σ–algebra $\mathcal{B}(S^{n-1})$ of Borel subsets of S^{n-1}. Signed measures and functions on S^{n-1} are called even if they are invariant under reflection in the origin [18].
The Hausdorff measure is a generalization of the Lebesgue measure to an arbitrary metric space and is needed for the proof of Theorem 4.9 [19].

Definition 4.2.2 (Hausdorff measure [19])**.** Let $0 \le s < \infty$ and $A \subset \mathbb{R}^n$. For $0 < \delta \le \infty$, define

$$\mathcal{H}^s_\delta(A) = \inf\left\{\sum_j d(C_j)^s : A \subset \bigcup_j C_j, d(C_j) < \delta, C_j \subset \mathbb{R}^n\right\} \tag{4.7}$$

where $d(C)$ is the diameter of C, i.e., $d(C) = \sup\{|x-y| : x,y \in C\}$ and $|x-y|$ is the Euclidean distance in $\mathbb{R}^n$. The s–dimensional Hausdorff measure of A is defined as

$$\mathcal{H}^s(A) = \lim_{\delta\to 0} \mathcal{H}^s_\delta(A) = \sup_{\delta>0} \mathcal{H}^s_\delta(A). \tag{4.8}$$

With these tools we can finally generalize the support function of zonotopes (Equation 4.6) for zonoids.

Theorem 4.2.1 ([18]). A convex body $K \subset \mathbb{R}^n$ is a zonoid with center at 0 if and only if its support function can be represented in the form

$$h(K,x) = \int_{S^{n-1}} |\langle x,v\rangle| d\rho_K(v) \quad \text{for } x \in \mathbb{R}^n \tag{4.9}$$

with some even measure $\rho := \rho_K$ on S^{n-1}.

Proof. "⇐"
Suppose Equation 4.9 holds. For $k \in \mathbb{N}$ we can decompose S^{n-1} into finitely many nonempty Borel sets $\Lambda^k_1, \dots, \Lambda^k_{m(k)}$ with diameter $< \frac{1}{k}$ and choose $v^k_i \in \Lambda^k_i$. Then,

$$\lim_{k\to\infty} \sum_{i=1}^{m(k)} |\langle x, v^k_i\rangle| \rho(\Lambda^k_i) = \int_{S^{n-1}} |\langle x,v\rangle| d\rho(v),$$

uniformly for $x \in S^{n-1}$. Thus, the zonotopes Z_k defined by

$$Z_k := \sum_{i=1}^{m(k)} \rho(\Lambda^k_i)\, \mathrm{conv}\{v^k_i, -v^k_i\}$$

satisfy $h(Z_k,\cdot) \to h(K,\cdot)$. Hence, $Z_k \to K$ for $k \to \infty$. According to Definition 4.1.4 K is a zonoid.
"⇒"
We can rewrite Equation 4.6 using an even measure ρ concentrated in finitely many points. Thus, we may assume that

$$h(Z_k,\cdot) = \int_{S^{n-1}} |\langle \cdot, v\rangle| d\rho_k(v) \tag{4.10}$$

with an even measure ρ_k on S^{n-1} with $k \in \mathbb{N}$. Further, $Z_k \to K$ for $k \to \infty$. Now we can integrate Equation 4.10 over S^{n-1} and receive

$$\begin{aligned}\int_{S^{n-1}} h(Z_k,u) d\mathcal{H}^{n-1}(u) &= \int_{S^{n-1}}\int_{S^{n-1}} |\langle u,v\rangle| d\rho_k(v) d\mathcal{H}^{n-1}(u)\\ &\overset{\text{Fubini}}{=} \int_{S^{n-1}}\int_{S^{n-1}} |\langle u,v\rangle| d\mathcal{H}^{n-1}(u) d\rho_k(v) = c(n)\rho_k(S^{n-1}).\end{aligned}$$

Note: $\mathcal{H}^n$ is the n-dimensional Hausdorff measure (cf. 4.2.2). Since the left-hand side converges for $k \to \infty$, $(\rho_k(S^{n-1}))_{k\in\mathbb{N}}$ is bounded. One can show that some subsequence $(\rho_{k_i})_{i\in\mathbb{N}}$ converges to a signed even measure ρ. The proof of this would exceed

the scope of this work, but for details see [18]. Thus, we get

$$h(K,x) = \lim_{i\to\infty} h(Z_{k_i},x) = \lim_{i\to\infty} \int_{S^{n-1}} |\langle x,v\rangle| d\rho_{k_i}(v)$$
$$= \int_{S^{n-1}} |\langle x,v\rangle| d\rho(v)$$

for each $x \in \mathbb{R}^n$. □

The next theorem (Theorem 4.2.2) will allow us to show that POVMs can be sparsified in Section 5.2.2. Firstly, we need to define the support of a real-valued function, e.g., a measure.

Definition 4.2.3 (Support of real-valued function [20])**.** If f is a real-valued function on a topological space, the *support* of f is the closure of the set $\{x : f(x) \neq 0\}$. Thus

$$\operatorname{supp}(f) = \overline{\{x : f(x) \neq 0\}}. \tag{4.11}$$

In the late 1980s J. Bourgain, J. Lindenstrauss, V. Milman, M. Talagrand and G. Schechtman investigated how zonoids can be approximated by zonotopes [12, 21, 22]. M. Talagrand compiled and supplemented the results (see [21]). One key result is stated in Theorem 4.2.2.

Theorem 4.2.2 ([2])**.** For any 0–symmetric zonoid $Y \subset \mathbb{R}^n$ and $\epsilon > 0$, there exists an integer $N \leq Cn\log(n)/\epsilon^2$ and vectors $x_1,\ldots,x_N \in \mathbb{R}^n$ such that $Z \subset Y \subset (1+\epsilon)Z$, where Z denotes the zonotope

$$Z = [-x_1, x_1] + \ldots + [-x_N, x_N].$$

Moreover, we can ensure that $\operatorname{supp}\rho_Z \subset \operatorname{supp}\rho_Y$, where the measures ρ_Y and ρ_Z are defined in Equation 4.9.

The proof is omitted, but can for example be found in the paper from M. Talagrand [21]. The last sentence gets clear by looking at the definition of the support of a measure (4.2.3) and by using $Z \subset Y$ (cf. construction of zonoids from zonotopes in the proof of Theorem 4.9).

5 Quantum measurements and quantum state discrimination

In the following chapter, key principles of quantum measurements are outlined (5.1). This will be especially useful in Chapter 7, since entanglement distillation relies on local operations such as measurements, for example. Firstly, the measurement postulate (Postulate 3) is stated (5.1.1), followed by a description of projective measurements (5.1.2). Further, the most general measurement formalism (POVM formalism) is introduced (5.1.3), as it will be used for the description of local filtering in Section 7.2.1. At the end of this chapter, quantum state discrimination is discussed (5.2) and the relation between POVMs and zonoids is outlined (5.2.1).

5.1 Measurements

The third postulate provides a mean for the description of measurement effects on quantum systems and will be described in Subsection 5.1.1. Projective measurements are primarily used for many applications in quantum computation and quantum information. They are a special case of the general measurement postulate (Postulate 3) and will be discussed in Section 5.1.2. Further, POVMs will be defined as a generalization of measurement (5.1.3).

5.1.1 Postulate 3 – Quantum measurement

The state space of a system refers to the mathematical space which describes all the possible states the system can be in. In quantum mechanics, the state space is described by a Hilbert space (cf. Chapters 2, 3).

Quantum measurements are described by a collection of measurement operators $\{M_m\}$, which act on the state space of the measured system. The index m refers to the measurement outcomes that could occur in the experiment. For a quantum system in a state $|\psi\rangle$ immediately before the measurement, the probability that the result m occurs is given by

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$$

and the state of the system after the measurement gets described by:

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}.$$

The measurement operators satisfy the completeness relation

$$\sum_m M_m^\dagger M_m = I$$

which expresses that the probabilities sum to one

$$1 = \sum_m p(m) = \sum_m \langle\psi|M_m^\dagger M_m|\psi\rangle.$$

5.1.2 Projective measurements

A projective measurement is described by an *observable*, M, a Hermitian operator on the state space of the system being observed. The observable has a spectral decomposition

$$M = \sum_m mP_m$$

where P_m is the projector onto the eigenspace of M with eigenvalues m. The possible measurement outcomes correspond to the eigenvalues of the observable. When measuring a state $|\psi\rangle$, the probability of receiving the measurement outcome m is given by

$$p(m) = \langle\psi|P_m|\psi\rangle.$$

When the measurement outcome is m, the state of the quantum system immediately after the measurement is

$$\frac{P_m|\psi\rangle}{\sqrt{p(m)}}.$$

Projective measurements can be understood as a special case of Postulate 3. If the measurement operators in Postulate 3, would satisfy the condition that M_m are orthogonal projectors, i.e., M_m are hermitian and $M_m M_{m'} = \delta_{m,m'} M_m$ in addition to the completeness relation, Postulate 3 would reduce to a projective measurement as defined [7].

5.1.3 Generalization of measurements

Postulate 3 describes the measurement statistics, e.g, the respective probabilities of the different possible measurement outcomes, as well as the post-measurement state of the system. For some application, the post-measurement state of the system is of little interest. Especially well adapted to the analysis of these measurements is the POVM formalism.

Definition 5.1.1 (POVM [2])**.** A Positive Operator Valued Measure (POVM) on a Hilbert space $\mathcal{H}$ is a finite family of positive operators $(M_i)_{1\leq i\leq N}$ with the property that $\sum M_i = I$.

What happens when a quantum system in a state ρ is measured with a POVM? We only focus on the case of discrete POVMs, $M = (M_i)_{1\leq i\leq N}$ here. Continuous POVMs could be treated as an approximation, then.

The POVM element M_m is associated with the measurement outcome m and the probability of obtaining m when making a measurement on the quantum state ρ is given by,

$$p(m) = \text{Tr}(\rho M_m). \tag{5.1}$$

This is also called Born's rule.
For a pure state $\rho = |\psi\rangle\langle\psi|$ this reduces to

$$p(m) = \text{Tr}(|\psi\rangle\langle\psi|M_m) = \langle\psi|M_m|\psi\rangle$$

where we used the cyclicity of the trace. This is exactly the probability we know from projective measurements [7].

5.2 Quantum state discrimination

Quantum state discrimination is the process of distinguishing between different quantum states using measurements, i.e., a quantum system is prepared in an unknown state ρ or σ and the unknown state should be found. After measuring a quantum system in an unknown state with the POVM $M = (M_i)_{1\leq i\leq N}$, outcome i occurs with probability $p_i = \text{Tr}(\rho M_i)$ (see Equation 5.1) if the unknown state is ρ and with probability $q_i = \text{Tr}(\sigma M_i)$ if the unknown state is σ. Therefore, we have the following strategy if outcome i is observed:

- if $p_i > q_i$ guess ρ
- if $q_i > p_i$ guess σ
- if $q_i = p_i$ guess ρ or σ.

The probability that the wrong state is guessed, which is also called probability of failure, is given by

$$\mathbb{P}_{\text{failure}} = \frac{1}{2}\sum_{i=1}^{N} \min(p_i, q_i) = \frac{1}{2} - \frac{1}{4}\sum_{i=1}^{N} |p_i - q_i|. \tag{5.2}$$

We can introduce a distinguishability semi-norm $||\cdot||_M$ by

$$||\Delta||_M = \sum_{i=1}^{N} |\text{Tr}(\Delta M_i)|. \tag{5.3}$$

$||\cdot||_M$ is a norm if and if only $\text{span}\{M_i : 1 \leq i \leq N\} = B^{sa}(\mathcal{H})$.
Since $\mathbb{P}_{\text{failure}} = \frac{1}{2} - \frac{1}{4}\sum_{i=1}^{N} |p_i - q_i| = \frac{1}{2} - \frac{1}{4}||\rho - \sigma||_M$ (see 5.2, 5.3) the distinguishability norm can be used to quantify the performance of POVMs for a state discrimination task. Note that for any POVM M holds $||\cdot||_M \leq ||\cdot||_1$ with $||A||_1 = \sum_{i=1}^{n} |\lambda_i|$ where λ_i is the i-th eigenvalue of A [2].
But why do we need POVMs for quantum state discrimination and can not just use projective measurements? The reason gets very clear by looking at the following example, which also illustrates the task of state discrimination.

Example: State discrimination

Suppose we want to specify if a quantum state is either $|\psi\rangle$ or $|\phi\rangle$, where $|\psi\rangle$ and $|\phi\rangle$ are non-orthogonal. With a projective measurement, we have no chance to discriminate these states, since they are non-orthogonal. Suppose for example we choose a projective measurement described by the operators $P_0 = |\psi\rangle\langle\psi|$ and $P_1 = |\psi\rangle^{\perp}\langle\psi|^{\perp}$. Then we obtain the result 0 with probability

$$p(0) = \langle\psi|P_0|\psi\rangle = \langle\psi|\psi\rangle\langle\psi|\psi\rangle = 1$$

if the unknown state is $|\psi\rangle$. However, we also obtain the result 0 with probability

$$p(0) = \langle\phi|P_0|\phi\rangle = \langle\phi|\psi\rangle\langle\psi|\phi\rangle \overset{\phi \text{ and } \psi \text{ are not orthogonal}}{=} |\langle\psi|\phi\rangle|^2$$

if the unknown state is $|\phi\rangle$. Conversely, assume a POVM measurement is performed in order to discriminate the states. Therefore, consider a POVM measurement specified by the operators $\{M_0, M_1, M_2\}$ with

$$\begin{aligned} M_0 &= a(I - |\psi\rangle\langle\psi|) \\ M_1 &= a(I - |\phi\rangle\langle\phi|) \\ M_2 &= I - M_0 - M_1 \end{aligned}$$

where $1/2 < a < 1$ and a is chosen such that $M_2 > 0$ (POVM). We receive the outcome

- 0 with probability

$$\begin{aligned} p_0 = \text{Tr}(|\psi\rangle\langle\psi|M_0) &\overset{\text{Tr cyclic}}{=} \text{Tr}(M_0|\psi\rangle\langle\psi|) = \text{Tr}((a(I - |\psi\rangle\langle\psi|))|\psi\rangle\langle\psi|) \\ &= a\text{Tr}(|\psi\rangle\langle\psi| - |\psi\rangle\langle\psi|\psi\rangle\langle\psi|) = a\text{Tr}(|\psi\rangle\langle\psi| - |\psi\rangle\langle\psi|) = 0 \end{aligned}$$

- 1 with probability

$$\begin{aligned} p_1 = \text{Tr}(|\psi\rangle\langle\psi|M_1) &\overset{\text{Tr cyclic}}{=} \text{Tr}(M_1|\psi\rangle\langle\psi|) = \text{Tr}((a(I - |\phi\rangle\langle\phi|))|\psi\rangle\langle\psi|) \\ &= a\text{Tr}(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\psi\rangle\langle\psi|) = a\text{Tr}(|\psi\rangle\langle\psi|) - a\text{Tr}(|\phi\rangle\langle\phi|\psi\rangle\langle\psi|) \\ &= a(1 - |\langle\psi|\phi\rangle|^2) \end{aligned}$$

- 2 with probability $p_2 = 1 - p_1$

if the unknown state is $|\psi\rangle$. Analogously, we receive the outcomes

- 0 with probability $q_0 = a(1 - |\langle\psi|\phi\rangle|^2)$
- 1 with probability $q_1 = 0$
- 2 with probability $q_2 = 1 - q_0$

if the unknown state is $|\phi\rangle$. Therefore, if we receive the result 1 we know with certainty that the original state was $|\psi\rangle$ and if we receive the outcome 0 we know with certainty that the original state was $|\phi\rangle$. However, if we receive the outcome 2, we can not tell with certainty what the unknown state was [23].

5.2.1 Zonotope associated to a POVM

In this subsection, we want to derive a relation between POVMs and zonotopes. Therefore, consider a POVM M. We denote by $B_M = \{||\cdot||_M \leq 1\}$ the unit ball for the distinguishability norm and by

$$K_M = (B_M)^\circ = \{A \in B^{sa}(\mathcal{H}) : \mathrm{Tr}(AB) \leq 1 \text{ whenever } ||B||_M \leq 1\}$$

its polar (cf. Definition 4.1.6). K_M is a compact convex set and has nonempty interior if and only if $||\cdot||_M$ is a norm. Using the inequality $||\cdot||_M \leq ||\cdot||_1$ it follows that K_M is always included in the unit ball for the operator norm.
The following Proposition 5.2.1 draws a connection between zonotopes and POVMs.

Proposition 5.2.1 ([2])**.** Let $K \subset B^{sa}(\mathcal{H})$ be a symmetric closed convex set. Then the following statements are equivalent.

- K is a zonotope such that $K \subset \{||\cdot||_\infty \leq 1\}$ and $\pm I \in K$.
- There exists a POVM M on $\mathcal{H}$ such that $K = K_M$.

Note: $||A||_\infty = ||A||_{\mathrm{op}} := \sup_{|x| \leq 1} |Ax|$ is the operator norm.

Proof. For a POVM $M = (M_i)_{1 \leq i \leq N}$ we firstly want to show that

$$K_M = [-M_1, M_1] + \ldots [-M_N, M_N].$$

Therefore, we denote $L := [-M_1, M_1] + \ldots [-M_N, M_N]$ and consider for every $A \in B^{sa}(\mathcal{H})$ the norm $||A||_{L^\circ}$,

$$||A||_{L^\circ} = \sup\{\mathrm{Tr}(AB) : B \in L\} = \sum_{i=1}^{N} |\mathrm{Tr}(AM_i)| = ||A||_{K_M^\circ},$$

where we used that according to Definition 4.2.1 $w(L, A) = \sup_{B \in L} \langle A, B \rangle = ||A||_{L^\circ}$ holds. Thus, $L = K_M$.
Conversely, suppose that K is a zonotope such that $K \subset \{||\cdot||_\infty \leq 1\}$ and $\pm I \in K$. Hence, I is an extreme point of K (cf. Definition 4.1.7). According to the definition of a zonotope (see Definition 4.1.2) there are operators $(M_i)_{1 \leq i \leq N}$ such that

$$K = [-M_1, M_1] + \ldots + [-M_N, M_N].$$

Any extreme point of K has the form $\pm M_1 \pm M_2 \ldots \pm M_N$ and therefore we may assume

$$I = M_1 + \ldots + M_N$$

and change M_i into $-M_i$ if necessary. Further, for every $1 \leq i \leq N$ is $I - M_i \in K$, hence $||I - M_i||_\infty \leq 1$. Thus, M_i is positive and $M = (M_i)_{i \leq i \leq N}$ is a POVM such that $K_M = K$. □

5.2.2 Sparsification of POVMs

In this section should be investigated whether POVMs can be sparsified, i.e., approximated by POVMs with few outcomes. Here, "approximation" refers to the associated distinguishability norms: a POVM M is considered to be ϵ-close to a POVM M' when their distinguishability norms satisfy

$$(1-\epsilon)||\cdot||_M \leq ||\cdot||_{M'} \leq (1+\epsilon)||\cdot||_M.$$

As a consequence of Theorem 4.2.2 about the approximation of zonotopes by zonoids a result about the sparsification of POVMs can be obtained, namely that for any given POVM M a POVM M' can be produced with relatively few outcomes, which performs the task of state discrimination almost as well as M.

Theorem 5.2.2. There is a constant C such that the following holds: for every POVM $M = (M_i)_{1\leq i\leq N}$ on $\mathbb{C}^n$ and every $\epsilon \in (0,1)$, there exists another POVM $M' = (M'_j)_{1\leq j\leq N'}$ with $N' \leq Cn^2 \log n/\epsilon^2$ outcomes such that

$$(1-\epsilon)||\cdot||_M \leq ||\cdot||_{M'}.$$

Proof. Consider the convex set $K_M \subset \mathrm{M}_n^{sa}$, which is according to Proposition 5.2.1 a zonoid. According to Theorem 4.2.2 there exists a zonotope

$$Z = [-A_1, A_1] + \ldots + [-A_{N'}, A_{N'}]$$

with positive operators A_i and $N' \leq Cn^2 \log n/\epsilon^2$, such that $(1-\epsilon)K_M \subset Z \subset K_M$. The A_i are positive since according to Theorem 4.2.2 supp $\rho_Z \subset$ supp ρ_Y. Now we define $A_0 = I - (A_1 + \ldots + A_{N'})$. A_0 is positive since $Z \subset K_M \subset S_\infty^{n,\mathrm{sa}}$, where $S_\infty^{n,\mathrm{sa}}$ is the unit ball for the operator norm. Thus, $M' := (A_0, A_1, \ldots A_{N'})$ is a POVM such that $(1-\epsilon)K_M \subset Z \subset K_{M'}$ and therefore $||\cdot||_{M'} \geq (1-\epsilon)||\cdot||_M$ as demanded. □

6 Channels and Werner states

In this chapter, quantum channels are introduced (6.1). In particular, the LOCC and the twirling channel are explained (6.1.1, 6.1.2), since they play a crucial role in entanglement distillation, what will be outlined in the next chapter 7. Further, Werner states are discussed (6.1).

6.1 Channels

A fundamental concept in quantum information theory are quantum channels, which are used to transfer quantum (or classical) information.
We start by looking at the definition of n-positivity and complete positivity first, since these definitions are necessary for defining quantum channels. Therefore, let us consider a linear map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, which is self-adjointness-preserving.

Definition 6.1.1 (Positive map)**.** The map Φ is said to be positivity preserving (short: positive) if the image of every positive operator is a positive operator.

Definition 6.1.2 (n-positivity [2])**.** The map Φ is said to be n-positive if
$\Phi \otimes Id : B^{\text{sa}}(\mathcal{H}_1 \otimes \mathbb{C}^n) \to B^{\text{sa}}(\mathcal{H}_2 \otimes \mathbb{C}^n)$ is positive.

Definition 6.1.3 (Complete-positivity [2])**.** The map Φ is said to be completely-positive if it is n-positive for every integer n.

One can observe that k-positivity is a necessary consequence of n-positivity for any $k < n$. The set of completely positive maps from $\mathcal{H}_1$ to $\mathcal{H}_2$ is denoted by $\mathbf{CP}(\mathcal{H}_1, \mathcal{H}_2)$.
Using complete positivity, we can define quantum channels.

Definition 6.1.4 (Quantum channel [2])**.** A quantum channel is a map $\Phi : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$, satisfying the following two properties:

1. Φ is completely positive
2. Φ is trace preserving

A quantum channel Φ should represent some physically realizable process, therefore states should be mapped to states. Thus, it is quite intuitive to demand from quantum channels to preserve trace, positivity and self-adjointness (compare to definition 3.1). However, demanding the preservation of positivity is not enough, since we want to apply a quantum channel to one part of a larger system and still end up with a quantum state at the end. By demanding complete positivity, we can apply our quantum channel on

one part of the system and leave the other parts unchanged and still have a quantum state as an output as demanded.
The theorem of Kraus (1971) gives a simple representation of any quantum channel.

Theorem 6.1.1 (Kraus decomposition [24])**.** For any quantum channel Φ, there exists a finite set of operators $A_1, A_2, ..., A_N \in B(\mathcal{H}_1, \mathcal{H}_2)$, such that for any $X \in B(\mathcal{H}_1)$

$$\Phi(X) = \sum_{i=1}^{N} A_i X A_i^\dagger \quad \text{with} \quad \sum_{i=1}^{N} A_i^\dagger A_i = I.$$

A proof will be omitted, but the proof can be for example found in the original article of Kraus from 1971 [25].
Using the Kraus decomposition, we can define separable maps. Therefore, assume that $\mathcal{H}_1$ and $\mathcal{H}_2$ are bipartite spaces $\mathcal{H}_1 = \mathcal{H}_{11} \otimes \mathcal{H}_{12}$ and $\mathcal{H}_2 = \mathcal{H}_{21} \otimes \mathcal{H}_{22}$.

Definition 6.1.5 (Separable maps)**.** A map $\Phi \in \mathbf{CP}(\mathcal{H}_1, \mathcal{H}_2)$ is called separable if it admits a Kraus decomposition involving product operators, i.e., if there exist operators $A_i^{(1)} : \mathcal{H}_{11} \to \mathcal{H}_{21}$ and $A_i^{(2)} : \mathcal{H}_{12} \to \mathcal{H}_{22}$ such that for any $X \in B(\mathcal{H}_1)$

$$\Phi(X) = \sum_{i=1}^{N} (A_i^{(1)} \otimes A_i^{(2)}) X (A_i^{(1)} \otimes A_i^{(2)})^\dagger.$$

6.1.1 LOCC channel

Next, two common and useful quantum channels get introduced: the LOCC channel and the twirling channel. Local operations and classical communication channels (LOCC channels) represent transformations of quantum states that can be implemented by two parties, which can classically communicate with one another and furthermore perform local operations [26, 27].

Since the mathematical structure of LOCC is complex, an exact description of LOCC would go beyond the scope of this work. Therefore, it will not be discussed in great detail. For further reading, see [28]. We denote (without proof) that

- an LOCC channel is separable (see Definition 6.1.5)
- any convex combination of product channels (of the form $\Phi_1 \otimes \Phi_2$) is an LOCC channel
- the class of LOCC channels is stable under composition
- conv{product channels} $\subset$ {LOCC channels } $\subset$ {separable channels}
- the local filtering operation is LOCC (which will be further discussed in 7.2.1) [2]

6.1.2 Twirling channel

Another well-known channel is the Werner twirling channel, defined in Definition 6.1.6.

Definition 6.1.6 (Twirling channel)**.** The quantum channel $\Upsilon : B(\mathbb{C}^2 \otimes \mathbb{C}^2) \to B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ defined as

$$\Upsilon(\rho) = \int (U \otimes U)\rho(U \otimes U)^\dagger d\eta(U) = \mathbf{E}(U \otimes U)\rho(U \otimes U)^\dagger$$

for all $\rho \in D(\mathcal{H}_A \otimes \mathcal{H}_A)$ is called the Werner twirling channel, where η denotes the Haar measure on $U(\mathcal{H}_A)$ (meaning that $U \in \mathrm{U}(2)$ is Haar - distributed).

For more information regarding the Haar measure, see [29].

6.2 Werner states

States that are invariant under a group of local unitaries play an important role in quantum mechanics. In the following section, we will focus on the class of states which are invariant under unitary transformations of the form $U \otimes U$ (e.g., under twirling). These states are called Werner states.
But before defining Werner states, we consider the symmetric and antisymmetric subspaces.

6.2.1 Flip operator, symmetric and antisymmetric subspaces

Consider the flip operator $F \in B^{sa}(\mathbb{C}^d \otimes \mathbb{C}^d)$ defined on pure tensors by $F(x \otimes y) = y \otimes x$ and extended by linearity. The eigenspaces of the flip operator are the symmetric subspace

$$\mathrm{Sym}_d = \{\psi \in \mathbb{C}^d \otimes \mathbb{C}^d : F(\psi) = \psi\}$$

and the antisymmetric subspace

$$\mathrm{Asym}_d = \{\psi \in \mathbb{C}^d \otimes \mathbb{C}^d : F(\psi) = -\psi\}$$

with the corresponding projectors $P_{\mathrm{Sym}_d} = \frac{1}{2}(I + F)$ and $P_{\mathrm{Asym}_d} = \frac{1}{2}(I - F)$. Furthermore, the symmetric and antisymmetric subspaces are irreducible for the action $U \to U \otimes U$ of the unitary group and $\dim \mathrm{Sym}_d = d(d+1)/2$ and $\dim \mathrm{Asym}_d = d(d-1)/2$. Moreover, the symmetric and antisymmetric states are defined as (see [2])

$$\pi_s = \frac{2}{d(d+1)} P_{\mathrm{Sym}_d} \quad \text{and} \quad \pi_a = \frac{2}{d(d-1)} P_{\mathrm{Asym}_d}.$$

6.2.2 Werner states: What you need to know

Werner states were introduced by R. F. Werner in 1989 (see [30]).

Definition 6.2.1 (Werner states [30])**.** Consider two Hilbert spaces $\mathcal{H}_A \cong \mathcal{H}_B \cong \mathbb{C}^d$ with equal finite dimension d. A bipartite quantum state $\rho \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ that satisfies

$$\rho = (U \otimes U)\rho(U^\dagger \otimes U^\dagger)$$

for all unitary operators U, is called Werner state.

Werner investigated which states are invariant under these transformations and showed that all such states are of the following form (see [30]):

$$w_\lambda = \frac{1}{d^2 - d\alpha} I - \alpha F \tag{6.1}$$

where $\lambda \in [0, 1]$ and

$$\alpha = \frac{1 + d(1 - 2\lambda)}{1 + d - 2\lambda} \in [-1, 1]. \tag{6.2}$$

Another equivalent expression for a Werner state is obtained as a convex combination of the symmetric and antisymmetric state

$$w_\lambda = \lambda\pi_s + (1 - \lambda)\pi_a. \tag{6.3}$$

The following proposition shows the correlation between Werner states and the twirling channel.

Proposition 6.2.1 (Werner states and twirling channel [2])**.** The twirling channel is related to Werner states according to the following statements:

1. For any $\rho \in D(\mathbb{C}^d \otimes \mathbb{C}^d)$ the Werner twirling channel (see 6.1.6) satisfies $\mathbf{E}(U \otimes U)\rho(U \otimes U)^\dagger = w_\lambda$ with $\lambda = \mathrm{Tr}(\rho P_{\mathrm{Sym}_d})$.
2. If $\psi \in S_{\mathbb{C}^d}$ is chosen uniformly at random, then $\mathbf{E}|\psi \otimes \psi\rangle\langle\psi \otimes \psi| = \pi_s$.

Proof. "1" By unitary invariance of the Haar measure

$$(V \otimes V)\Upsilon(\rho)(V \otimes V)^\dagger = \Upsilon(\rho)$$

holds for every $V \in U(d)$. Thus, $\Upsilon(\rho)$ is invariant under unitary transformations and satisfies Definition 6.2.1. Therefore, $\Upsilon(\rho)$ can be represented as

$$\Upsilon(\rho) = w_\lambda = \lambda\pi_s + (1 - \lambda)\pi_a = \lambda\frac{P_{\mathrm{Sym}_d}}{\dim P_{\mathrm{Sym}_d}} + (1 - \lambda)\frac{P_{\mathrm{Asym}_d}}{\dim P_{\mathrm{Asym}_d}}.$$

Furthermore, note that Werner states are invariant under twirling, i.e., $\Upsilon(\rho) = \rho$ for any Werner state ρ and that any quantum state ρ satisfies

$$1 = \mathrm{Tr}(\rho) = \mathrm{Tr}(\rho I) = \langle I, \rho\rangle = \langle P_{\mathrm{Sym}_d} + P_{\mathrm{Asym}_d}, \rho\rangle = \langle P_{\mathrm{Sym}_d}, \rho\rangle + \langle P_{\mathrm{Asym}_d}, \rho\rangle \tag{6.4}$$

where we used that $P_{\mathrm{Sym}_d} + P_{\mathrm{Asym}_d} = 1/2(I+F)+1/2(I-F) = I$. Moreover, we can calculate λ by using the orthogonality of P_{Sym_d} and P_{Asym_d} .

$$\langle P_{\mathrm{Sym}_d}, \Upsilon(\rho)\rangle = \langle P_{\mathrm{Sym}_d}, \lambda\frac{P_{\mathrm{Sym}_d}}{\dim P_{\mathrm{Sym}_d}} + (1-\lambda)\frac{P_{\mathrm{Asym}_d}}{\dim P_{\mathrm{Asym}_d}}\rangle = \lambda.$$

Next, we note that

$$(U\otimes U)P_{\mathrm{Sym}_d}(U\otimes U) = P_{\mathrm{Sym}_d}$$

and

$$(U\otimes U)P_{\mathrm{Asym}_d}(U\otimes U) = P_{\mathrm{Asym}_d}.$$

Therefore, we can rewrite $\lambda = \langle P_{\mathrm{Sym}_d}, \Upsilon(\rho)\rangle = \langle \Upsilon(P_{\mathrm{Sym}_d}), \rho\rangle = \langle P_{\mathrm{Sym}_d}, \rho\rangle = \mathrm{Tr}(\rho P_{\mathrm{Sym}_d})$ [31].

"2" If we apply 1 to $\rho = |x\otimes x\rangle\langle x\otimes x|$ where x is a fixed unit vector in $\mathbb{C}^d$ we receive

$$\begin{aligned} w_\lambda = \mathbf{E}((U\otimes U)|x\otimes x\rangle\langle x\otimes x|(U\otimes U)^\dagger) &= \mathbf{E}((U\otimes U)|x\rangle\otimes|x\rangle\langle x|\otimes\langle x|(U\otimes U)^\dagger) \\ &= \mathbf{E}(U|x\rangle\otimes U|x\rangle\langle x|U^\dagger\otimes\langle x|U^\dagger) \end{aligned}$$

We note that unitary transformations are norm preserving $||Ux|| = \langle Ux, Ux\rangle = x^T U^T U x = ||x||^2$) and define $|\psi\rangle := U|x\rangle \in S_{\mathbb{C}^d}$. Therefore, we get

$$w_\lambda = \mathbf{E}|\psi\otimes\psi\rangle\langle\psi\otimes\psi|.$$

Furthermore,

$$\begin{aligned} \lambda = \mathrm{Tr}(|x\otimes x\rangle\langle x\otimes x|P_{\mathrm{Sym}_d}) &= \mathrm{Tr}(|x\otimes x\rangle\langle x\otimes x|\frac{1}{2}(I+F)) \\ \overset{\text{trace cyclic, linear}}{=} \frac{1}{2}\mathrm{Tr}((I+F)|x\otimes x\rangle\langle x\otimes x|) &= \frac{1}{2}\mathrm{Tr}(I|x\otimes x\rangle\langle x\otimes x| + F|x\otimes x\rangle\langle x\otimes x|) \\ = \frac{1}{2}\mathrm{Tr}(|x\otimes x\rangle\langle x\otimes x| + |x\otimes x\rangle\langle x\otimes x|) = \frac{1}{2}\mathrm{Tr}(2|x\otimes x\rangle\langle x\otimes x|) &\overset{\text{Tr(density matrix)=1}}{=} 1 \end{aligned}$$

Thus, it follows that $w_\lambda = \lambda\pi_s + (1-\lambda)\pi_a = 1\pi_s - 0\pi_a = \pi_s$, what had to be shown. □

7 Entanglement distillation

In the last chapters, we did important preparatory work. In this chapter, we can finally talk about entanglement and entanglement distillation. As mentioned in the introduction 1, the goal of entanglement distillation is to transform weakly entangled states into maximally entangled states. But what are even entangled states, let alone highly and weakly entangled states? How can entanglement be quantified? What are distillable states and which states can be distilled? That all should be clarified within this chapter. Firstly, separability and entanglement get discussed (7.1). Then the concepts of fidelity is introduced (7.1.1) as a measure of proximity between quantum states, followed by a section about distillable states 7.2, where distillation and the distillability problem is explained.

7.1 Separability and Entanglement

In this section, the dichotomy between separability and entanglement for quantum states is studied.
Suppose $\mathcal{H}$ is a complex Hilbert space, which can be tensor decomposed

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes ... \otimes \mathcal{H}_k.$$

Definition 7.1.1 (Separable pure state [2])**.** A pure state $\rho = |\chi\rangle\langle\chi|$ on $\mathcal{H}$ is said to be *pure separable* if the unit vector χ is a product vector, i.e., if there exist unit vectors $\chi_1, ..., \chi_k$ such that $\chi = \chi_1 \otimes \chi_2 \otimes ... \otimes \chi_k$. In that case,

$$\rho = |\chi_1\rangle\langle\chi_1| \otimes |\chi_2\rangle\langle\chi_2| \otimes ... \otimes |\chi_k\rangle\langle\chi_k|$$

To formulate a corresponding separability definition for mixed states convex combinations have to be considered.

Definition 7.1.2 (Separable mixed state [2])**.** A mixed state on $\mathcal{H}$ is said to be *separable* if it can be written as a convex combination of pure separable states. We denote by Sep($\mathcal{H}$) (or simply by Sep) the set of separable states on $\mathcal{H}$. Therefore,

$$\text{Sep}(\mathcal{H}) = \text{conv}\{|\chi_1 \otimes \chi_2 \otimes ...\chi_k\rangle\langle\chi_1 \otimes \chi_2 \otimes ...\chi_k| : \chi_1 \in \mathcal{H}_1, \chi_2 \in \mathcal{H}_2..., \chi_k \in \mathcal{H}_k\}. \tag{7.1}$$

The cone of separable operators is given by $\mathcal{SEP}(\mathcal{H}) = \{\lambda\rho : \lambda \geq 0, \rho \in \text{Sep}(\mathcal{H})\}$.
Now we can finally define entangled states.

Definition 7.1.3 (Entangled state)**.** Non-separable states are called entangled.

Furthermore, one can define k-entangled states.

Definition 7.1.4 (k-entangled states)**.** A quantum state on $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ is said to be k-entangled if it can be written as a convex combination

$$\sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$$

where each unit vector $\psi_i \in \mathcal{H}_1 \otimes \mathcal{H}_2$ has Schmidt rank at most k.

Separable states are 1-entangled states.

Maximally entangled states

The goal of entanglement distillation is to achieve a state that is close to a maximally entangled state, how can such a state be mathematically described?

Definition 7.1.5 (Maximally entangled state [2])**.** A pure state ρ on $\mathbb{C}^d \otimes \mathbb{C}^d$ is called maximally entangled if it has the form $\rho = |\psi\rangle\langle\psi|$ with

$$\psi = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} e_i \otimes f_i$$

where $(e_i)_{1\leq i\leq d}$ and $(f_i)_{1\leq i\leq d}$ are two orthonormal bases in $\mathbb{C}^d$. Such a vector ψ is called a maximally entangled vector.

For systems formed by 2 qubits ($d = 2$) the maximally entangled states are called Bell states and are defined in Definition 7.1.6. The canonical basis of $\mathbb{C}^2$ is $(|0\rangle, |1\rangle)$ and often the tensor product signs are dropped (for example $|00\rangle \equiv |0\rangle \otimes |0\rangle$).

Definition 7.1.6 (Bell states)**.** The bell states are given by,

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \qquad |\phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$$
$$|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \qquad |\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

where $|00\rangle = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$ $|01\rangle = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$ $|10\rangle = \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$ $|11\rangle = \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

The family of Bell vectors $\{\phi^+, \phi^-, \psi^+, \psi^-\}$ is an orthonormal basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$.

7.1.1 Fidelity

Our goal in entanglement distillation is to get "closer " to a maximally entangled state with every round of the distillation protocol. The "closeness" of two quantum states can be mathematically formulated using the concept of fidelity. Fidelity is a measure of

distance between quantum states and can take values between 0 and 1. If the fidelity is equal to 1, both states are identical. If the fidelity is equal to 0, the states can be distinguished from one another with certainty by a quantum mechanical measurement. The fidelity can be defined as follows,

Definition 7.1.7 (Fidelity)**.** The fidelity of two states ρ and σ is defined as

$$F(\rho,\sigma) \equiv \mathrm{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}} \tag{7.2}$$

One important special case of the fidelity is the fidelity between a pure state $\rho = |\psi\rangle\langle\psi|$ and an arbitrary state σ. Using equation 7.2 we see that

$$F(\rho,\sigma) = \mathrm{Tr}\sqrt{|\psi\rangle\langle\psi|\sigma|\psi\rangle\langle\psi|} = \mathrm{Tr}\sqrt{\langle\psi|\sigma|\psi\rangle|\psi\rangle\langle\psi|} = \sqrt{\langle\psi|\sigma|\psi\rangle}\mathrm{Tr}(\sqrt{|\psi\rangle\langle\psi|}) = \sqrt{\langle\psi|\rho|\psi\rangle} \tag{7.3}$$

where we used that pure states are idempotent ($\rho^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \rho$) and $\mathrm{Tr}(\rho) = 1$.

7.2 Distillable States

After gaining an understanding of relevant concepts and definition, we get to the central question of this thesis: which categories of quantum states can be distilled and which can not?
It is relatively easy to prove that states with a positive partial transpose (PPT) are not distillable (see [32] for example). An ongoing open question is whether the converse holds, e.g., whether all bipartite states with a non-positive partial transpose are distillable. This is often referred to as the distillability problem.
In this section, we discuss some partial results, focusing on the case of two qubits. This section is based on chapter 12 of the book "Alice and Bob Meet Banach. The Interface of Asymptotic Geometric Analysis and Quantum Information Theory" by Aubrun and Szarek [2].

7.2.1 Distillability problem

We start by mathematically describing distillation. Therefore, let us consider a bipartite Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ shared between two parties, which are commonly referred to as Alice and Bob. The Hilbert space $\mathcal{H}^{\otimes n}$ for $n \in \mathbb{N}$ with $n \geq 1$ is also considered to be bipartite by identifying it with $\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n}$. In the following section separability, LOCC, partial transpose,... for states or channels on $\mathcal{H}^{\otimes n}$ are always understood relative to the A:B bipartition.

Definition 7.2.1 (Distillation [2])**.** Given two bipartite state $\rho \in D(\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B})$ and $\sigma \in D(\mathcal{H}'_\mathrm{A} \otimes \mathcal{H}'_\mathrm{B})$ we say σ can be distilled from (multiple copies of) ρ ($\rho \rightsquigarrow \sigma$) if for all $\epsilon > 0$ exists an integer $n \in \mathbb{N}$ and an LOCC quantum channel
$\Phi : B((\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B})^{\otimes n}) \to B(\mathcal{H}'_\mathrm{A} \otimes \mathcal{H}'_\mathrm{B})$ such that $||\Phi(\rho^{\otimes n}) - \sigma||_1 \leq \epsilon$.

Note: $||A||_1 = \sum_{i=1}^{n} |\lambda_i|$ where λ_i is the i-th eigenvalue of A.
Since many quantum information protocols use Bell states as a resource, the following formulation for the distillability of bipartite states is motivated:
A bipartite state $\rho \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ is said to be *distillable* if $\rho \rightsquigarrow |\psi^+\rangle\langle\psi^+|$, where $|\psi^+\rangle$ is the Bell state defined in 7.1.6.
The choice of the Bell vector ψ^+ is arbitrary, what can be easily understood by considering the following: For two maximally entangled vectors $x, y \in \mathbb{C}^d \otimes \mathbb{C}^d$ exist $U, V \in \mathrm{U}(d)$ such that $y = (U \otimes V)x$. Since the channel $\rho \mapsto (U \otimes V)\rho(U \otimes V)^\dagger$ is a product channel and therefore an LOCC channel, we receive that $|x\rangle\langle x| \rightsquigarrow |y\rangle\langle y|$. This is also called *conjugating with local unitaries*.
Therefore, the choice of the ψ^+ is arbitrary, since we can obtain any other Bell state by conjugating with local unitaries.
As mentioned in Subsection 6.1.1 the local filtering operation is LOCC.
Local filtering Given a state ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$, POVMs $(P_i)_{i\in I}$ on $\mathcal{H}_A$ and $(Q_j)_{j\in J}$ on $\mathcal{H}_B$, and $S \subset I \times J$, then $\rho \rightsquigarrow \frac{M}{\mathrm{Tr}M}$ (provided $\mathrm{Tr}M > 0$) , where

$$M = \sum_{(i,j)\in S} (P_i \otimes Q_j)\rho(P_i \otimes Q_j).$$

To enhance the intuitive comprehension, the concept of local filtering is further illustrated. Given n copies of the state ρ, Alice and Bob can successively measure copies of ρ locally using the POVMs (P_i) and (Q_j) until they obtain outcomes i and j such that $(i, j) \in S$. The post-measurement state is then given by $\frac{M}{\mathrm{Tr}M}$. If none of the n copies gives an outcome in S the protocol fails. However, as $n \to \infty$ the probability of failure $\mathbb{P}_{\text{failure}} \to 0$. If there is an outcome in S Alice and Bob verify that $(i, j) \in S$ by classically communicating.
As previously stated in the introduction to this section, the distillability problem pursues whether all non-PPT states are distillable.
Distillability problem `Is every non-PPT state distillable?`
The distillability problem is graphically illustrated in FIG. 7.1.

7.2.2 Two qubits

In this section, we will prove a very interesting proposition, namely that every entangled state on $\mathbb{C}^2 \otimes \mathbb{C}^2$ is distillable. The Peres-Horodecki criterion and the two following lemmas provide the foundation for the proof of the proposition.

Theorem 7.2.1 (Peres-Horodecki criterion [33, 34])**.** If $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ or $\mathcal{H} = \mathbb{C}^3 \otimes \mathbb{C}^2$, then every PPT state on $\mathcal{H}$ is separable.

The proof can be done using the Strømer and Woronowicz results (see [35, 36]) and is omitted as it would exceed the scope of the work, it can be found in [33, 34].
By contraposition, this statement is equivalent to the following: Every entangled (not separable) state on $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ or $\mathcal{H} = \mathbb{C}^3 \otimes \mathbb{C}^2$ is NPT.

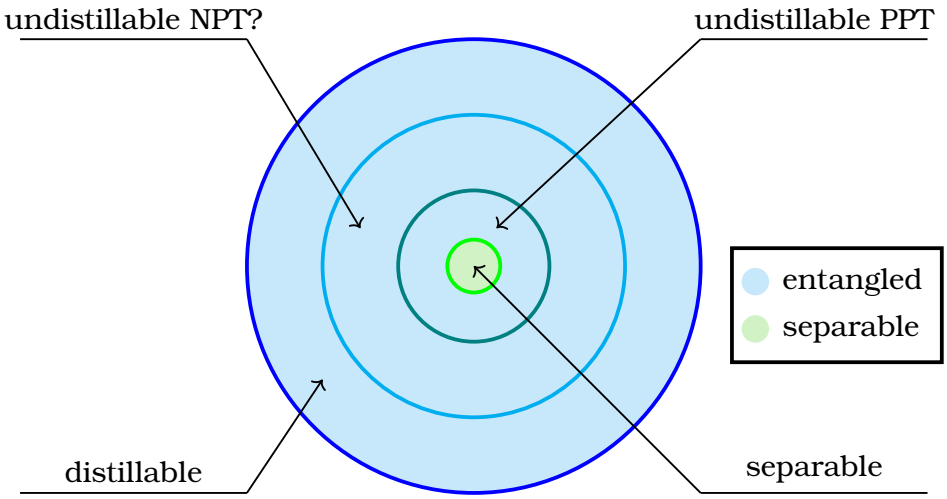


Figure 7.1: Graphical illustration of the distillability problem. The set of separable states is colored light green, the state of entangled states light blue. The big open question is if undistillable states with a non-positive partial transpose exist.

In the following two Lemmas, we will work with states that are diagonal in the basis of Bell vectors. Therefore, we denote

$$\rho_{a,b,c,d} = a|\phi^+\rangle\langle\phi^+| + b|\phi^-\rangle\langle\phi^-| + c|\psi^+\rangle\langle\psi^+| + d|\psi^-\rangle\langle\psi^-|$$

for $a, b, c, d \geq 0$ such that $a + b + c + d = 1$.
We associate the quantity

$$s(\rho) = \max_{\substack{\chi \in \mathbb{C}^2 \otimes \mathbb{C}^2 \\ \chi \text{ max. entangled}}} \{\langle\chi|\rho|\chi\rangle\}$$

to each state $\rho \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$. By comparing $s(\cdot)$ with the fidelity Definition 7.1.1 one can see that $\langle\chi|\rho|\chi\rangle$ is the square of $F(\rho, |\chi\rangle\langle\chi|)$. Therefore, the functional $s(\cdot)$ measures proximity to the set of maximally entangled states. Concretely, ρ is distillable iff there exists a sequence (σ_n) in $D(\mathbb{C}^2 \otimes \mathbb{C}^2)$ such that $s(\sigma_n) \to 1$ and that, for every n, $\rho \rightsquigarrow \sigma_n$. With this knowledge, we are well-equipped for the following lemmas.

Lemma 7.2.2 ([2]). $\rho \rightsquigarrow \rho_{s(\rho),\frac{1-s(\rho)}{3},\frac{1-s(\rho)}{3},\frac{1-s(\rho)}{3}}$

Proof. By conjugating with local unitaries $s(\rho) = \langle\psi^-|\rho|\psi^-\rangle$ can be obtained. The twirling channel defined in Definition 6.1.6 belongs to the convex hull of the set of product channels and is therefore according to 6.1.1 an LOCC channel. By using the result from Theorem 6.1, plugging in λ and expressing w_λ in form of Bell states

$$\Upsilon(\rho) = s(\rho)|\psi^-\rangle\langle\psi^-| + \frac{1 - s(\rho)}{3}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+|)$$

can be obtained. Once again, we can conjugate with local unitaries and transform ψ^- into ϕ^+. Thus $||\Upsilon(\rho) - \rho_{s(\rho),\frac{1-s(\rho)}{3},\frac{1-s(\rho)}{3},\frac{1-s(\rho)}{3}}||_1 = 0 \leq \epsilon$ and according to Definition 7.2.1 $\rho \rightsquigarrow \rho_{s(\rho),\frac{1-s(\rho)}{3},\frac{1-s(\rho)}{3},\frac{1-s(\rho)}{3}}$ holds. □

For the proof of the following lemma, quantum gates are used, which should be briefly introduced first.

Gates

A quantum gate is an elementary quantum circuit working on a small number of qubits. Unlike classical gates, quantum gates can realize superposition and entanglement [37]. We will not go into much detail discussing quantum gates and just briefly discuss the 2-qubit CNOT-gate. For more information regarding quantum gates, see [37].

CNOT-gate [7]

The controlled-NOT or CNOT gate has two inputs qubits, known as control qubit and target qubit. In FIG. 7.2 the schematic circuit for the controlled-NOT gate is shown.

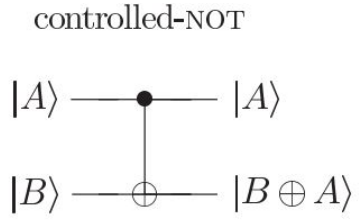


Figure 7.2: Circuit representation for the CNOT gate: The top line represents the control qubit and the bottom line the target qubit: figure from [7].

If the control qubit is set to 0, the target qubit stays the way it is. If the control qubit is 1, the target qubit is flipped, e.g.,

$$|00\rangle \to |00\rangle \quad |01\rangle \to |01\rangle \quad |10\rangle \to |11\rangle \quad |11\rangle \to |10\rangle \quad . \tag{7.4}$$

Lemma 7.2.3 ([2]). Given $a, b, c, d \geq 0$ with $a + b + c + d = 1$, denote $\alpha = \frac{a^2+b^2}{N}$, $\beta = \frac{2ab}{N}$, $\gamma = \frac{a^2+b^2}{N}$ and $\delta = \frac{2cd}{N}$, where $N = (a+b)^2 + (c+d)^2$. Then

$$\rho_{\mathrm{a,b,c,d}} \rightsquigarrow \rho_{\alpha,\beta,\gamma,\delta}.$$

Proof. In the following $\mathcal{H}_\mathrm{A}, \mathcal{H}_\mathrm{B}, \mathcal{H}'_\mathrm{A}$ and $\mathcal{H}'_\mathrm{B}$ are all equal to $\mathbb{C}^2$. Consider $\rho_{a,b,c,d}$ as a state on $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B}$ and $\rho_{a,b,c,d} \otimes \rho_{a,b,c,d}$ as a state on $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B} \otimes \mathcal{H}'_\mathrm{A} \otimes \mathcal{H}'_\mathrm{B}$. $P = |00\rangle\langle 00| + |11\rangle\langle 11|$ and $Q = I - P = |01\rangle\langle 01| + |10\rangle\langle 10|$ are rank 2 projectors acting on $\mathbb{C}^2 \otimes \mathbb{C}^2$.
We equip all operators X respectively superoperators Φ acting on $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}'_\mathrm{A}$ respectively $B(\mathcal{H}_\mathrm{A} \otimes \mathcal{H}'_\mathrm{A})$ with a subscript A (X_A, Φ_A) and all operators respectively superoperators acting on $\mathcal{H}_\mathrm{B} \otimes \mathcal{H}'_\mathrm{B}$ respectively $B(\mathcal{H}_\mathrm{B} \otimes \mathcal{H}'_\mathrm{B})$ with a subscript B (X_B, Φ_B).
Now we consider the operator $\Pi = P_A \otimes P_B + Q_A \otimes Q_B$ acting on $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B} \otimes \mathcal{H}'_\mathrm{A} \otimes \mathcal{H}'_\mathrm{B}$. Π is an orthogonal projector on the subspace spanned by the vectors

$$\phi^+ \otimes \phi^+ \quad \phi^+ \otimes \phi^- \quad \phi^- \otimes \phi^+ \quad \phi^- \otimes \phi^- \quad \psi^+ \otimes \psi^+ \quad \psi^+ \otimes \psi^- \quad \psi^- \otimes \psi^+ \quad \psi^- \otimes \psi^-$$

what can be verified by checking $\Pi^2 = \Pi$ and $\Pi^T = \Pi$ and plugging into the definition for the projection.
Let us consider a quantum channel $\Psi : B(\mathbb{C}^2 \otimes \mathbb{C}^2) \to B(\mathbb{C}^2)$ defined as

$$\rho \to \mathrm{Tr}_2 U \rho U^\dagger$$

with the partial trace over the second factor Tr_2 (see Definition 2.3.5) and the "CNOT" (controlled - NOT gate, see 7.4) unitary transformation on $\mathbb{C}^2 \otimes \mathbb{C}^2$ defined as

$$U(|00\rangle) = |00\rangle \quad U(|01\rangle) = |01\rangle \quad U(|10\rangle) = |11\rangle \quad U(|11\rangle) = |10\rangle.$$

Using this channel, for $\epsilon, \eta = \pm$ holds

$$(\Psi_A \otimes \Psi_B)(|\phi^\epsilon \otimes \phi^\eta\rangle\langle\phi^\epsilon \otimes \phi^\eta|) = |\phi^{\epsilon\eta}\rangle\langle\phi^{\epsilon\eta}|$$
$$(\Psi_A \otimes \Psi_B)(|\psi^\epsilon \otimes \psi^\eta\rangle\langle\psi^\epsilon \otimes \psi^\eta|) = |\psi^{\epsilon\eta}\rangle\langle\psi^{\epsilon\eta}|$$

which can be verified by directly calculating and using sign multiplication rules. It has to be emphasized that one has to be careful with the symbol $\otimes$ since it does not always refer to the same bipartition. By using local filtering and then the LOCC channel $\Psi_A \otimes \Psi_B$ it can be computed that

$$\rho \rightsquigarrow \frac{\Pi(\rho \otimes \rho)\Pi}{\mathrm{Tr}\Pi(\rho \otimes \rho)\Pi} \rightsquigarrow \rho_{\alpha,\beta,\gamma,\delta}.$$

□

Now we can discuss the main Proposition 7.2.4 of this subsection, which is a special case of the distillability problem.

Proposition 7.2.4 ([2])**.** Every entangled state on $\mathbb{C}^2 \otimes \mathbb{C}^2$ is distillable.

Proof. Consider an entangled state $\rho \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$, which is according to the Peres-Horodecki criterion 7.2.1 not PPT. Therefore, there exists a unit vector $x \in \mathbb{C}^2 \otimes \mathbb{C}^2$ with $\langle x|\rho^\Gamma|x\rangle < 0$. By conjugating with local unitaries, we can assume that the Schmidt decomposition (see. 3.1.2) of x is $x = a|00\rangle + b|11\rangle$. By using the operator

$$J = a|0\rangle\langle 0| + b|1\rangle\langle 1| = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

we can express x the following way: $x = \sqrt{2}(I \otimes J)|\phi^+\rangle$, which can easily be checked by plugging in J. Further, J can be one of the operators of a POVM, since $0 \le J \le I$. By applying local filtering, we receive

$$\rho \rightsquigarrow \sigma := \frac{(I \otimes J)\rho(I \otimes J)}{\mathrm{Tr}(I \otimes J)\rho(I \otimes J)}.$$

Furthermore, one can check that $\langle\phi^+|\sigma^\Gamma|\phi^+\rangle < 0$. By using $\mathrm{Tr}(A^\Gamma B) = \mathrm{Tr}(AB^\Gamma)$ we get

$$0 > \mathrm{Tr}\left(\sigma\left(|\phi^+\rangle\langle\phi^+|\right)^\Gamma\right) = \mathrm{Tr}\left(\sigma\left(\frac{1}{2}I - |\psi^-\rangle\langle\psi^-|\right)\right) = \frac{1}{2} - \langle\psi^-|\sigma|\psi^-\rangle$$
$$\Leftrightarrow \langle\psi^-|\sigma|\psi^-\rangle > \frac{1}{2} \Leftrightarrow s(\sigma) > \frac{1}{2}.$$

Therefore, the problem is reduced to show that any state σ with $s(\sigma) > 1/2$ is distillable. By successively applying Lemmas 7.2.2 and 7.2.3 on σ, we get that $\sigma \rightsquigarrow \sigma'$ for a state σ' for which $s(\sigma') \geq \phi(s(\sigma))$ with

$$\phi(t) = \frac{t^2 + \frac{1}{9}(1-t)^2}{\frac{1}{9}(1+2t)^2 + \frac{1}{9}(2-2t)^2} = \frac{1 - 2t + 10t^2}{5 - 4t + 8t^2}$$

holds. Since $\phi(t) > t$ for $t \in (1/2, 1)$ $\lim_{n\to\infty} \phi^n(s(\sigma)) = 1$. By iterating the above described process, one can receive that $\sigma \rightsquigarrow \sigma''$ for a state σ'' such that $s(\sigma'')$ is as close to 1 as demanded. □

Bibliography

[1] S. M. Barnett und S. Croke, *Quantum state discrimination* (2008). 0810.1970.

[2] G. Aubrun und S. Szarek, *Alice and Bob Meet Banach*, Mathematical Surveys and Monographs, American Mathematical Society (2017), ISBN 9781470434687. URL https://books.google.at/books?id=h28zDwAAQBAJ.

[3] P. A. M. Dirac, *A new notation for quantum mechanics*, Mathematical Proceedings of the Cambridge Philosophical Society, **35**(3), 416–418 (1939). URL http://dx.doi.org/10.1017/S0305004100021162.

[4] B. C. Hall, *Quantum theory for mathematicians*, Springer (2013).

[5] R. Bertlmann, *Theoretical Physics T2 Quantum Mechanics* (2008). Script written by Reinhold A. Bertlmann and Nicolai Friis.

[6] J. Blank, P. Exner und M. Havlicek, *Hilbert space operators in quantum physics*, Springer Science & Business Media (2008).

[7] M. A. Nielsen und I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, Cambridge University Press (2010). URL http://dx.doi.org/10.1017/CBO9780511976667.

[8] M. M. Wilde, *Quantum Information Theory*, Cambridge University Press, 2 Aufl. (2017). URL http://dx.doi.org/10.1017/9781316809976.

[9] P. Asharaf und B. K. Thomas, *A note on Frobenius inner product and the m-distance matrices of a tree*, Malaya Journal of Matematik (MJM), **8**(3, 2020), 1321 (2020).

[10] A. Ekert und P. L. Knight, *Entangled quantum systems and the Schmidt decomposition*, American Journal of Physics, **63**(5), 415 (1995). URL http://dx.doi.org/10.1119/1.17904.

[11] M. University, *Zonotopes*. URL https://www.cs.mcgill.ca/~fukuda/760B/handouts/expoly3.pdf.

[12] J. Bourgain, J. Lindenstrauss und V. Milman, *Approximation of zonoids by zonotopes*, Acta Math., **162**, 73– (1989). URL https://doi.org/10.1007/BF02392835.

[13] E. D. Bolker, *A Class of Convex Bodies*, Transactions of the American Mathematical Society, **145**, 323 (1969). URL http://www.jstor.org/stable/1995073.

[14] CJ, *What is a zonotope?* Last accessed on 2023-05-24, URL https://mitadmissions.org/blogs/entry/what-is-a-zonotope/#annotation-trigger-4.

[15] A. A. Tuzhilin, *Lectures on Hausdorff and Gromov-Hausdorff Distance Geometry* (2020). 2012.00756.

[16] L. Mathis et al., *The handbook of zonoid calculus* (2022).

[17] G. Magaril-Il'yaev und V.M.Tikhomirov, *Convex Analysis: Theory and Applications*, Translations of mathematical monographs, American Mathematical Soc. (2003), ISBN 9780821889640. URL https://books.google.at/books?id=7dHblO2nGmMC.

[18] R. Schneider, *Convex Bodies: The Brunn–Minkowski Theory*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2 Aufl. (2013). URL http://dx.doi.org/10.1017/CBO9781139003858.

[19] D. Feng., *Chapter 3* (2022). URL https://www.math.cuhk.edu.hk/course_builder/1415/math5011/MATH5011_Chapter_3.2014.pdf.

[20] H. L. Royden, *Real analysis*, Macmillan, New York, 2d ed Aufl. (1968).

[21] M. Talagrand, *Embedding Subspaces of L1 into lN 1*, Proceedings of the American Mathematical Society, **108**(2), 363 (1990). URL http://www.jstor.org/stable/2048283.

[22] G. Schechtman, *More on embedding subspaces of L_p in l_r^n*, Compositio Mathematica, **61**(2), 159 (1987).

[23] R. Yuan, *A Brief Introduction to POVM Measurement in Quantum Communications* (2022). 2201.07968.

[24] W. G. Ritter, *Quantum channels and representation theory*, Journal of Mathematical Physics, **46**(8), 082103 (2005). URL http://dx.doi.org/10.1063/1.1945768.

[25] K. Kraus, *General state changes in quantum theory*, Annals of Physics, **64**(2), 311 (1971). URL http://dx.doi.org/https://doi.org/10.1016/0003-4916(71)90108-4.

[26] J. Watrous, *The Theory of Quantum Information*, Cambridge University Press (2018). URL http://dx.doi.org/10.1017/9781316848142.

[27] L. Clarisse, *Entanglement distillation; a discourse on bound entanglement in quantum information theory*, arXiv preprint quant-ph/0612072 (2006).

[28] E. Chitambar, D. Leung, L. Mancinska, M. Ozols und A. Winter, *Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask)*, Communications in Mathematical Physics, **328** (2012). URL http://dx.doi.org/10.1007/s00220-014-1953-9.

[29] D. Cohn, *Measure Theory: Second Edition*, Birkhäuser Advanced Texts Basler Lehrbücher, Springer New York (2013), ISBN 9781461469568. URL https://books.google.at/books?id=PEC3BAAAQBAJ.

[30] R. F. Werner, *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*, Phys. Rev. A, **40**, 4277 (1989). URL http://dx.doi.org/10.1103/PhysRevA.40.4277.

[31] A. Müller-Hermes, *MAT4430 -Quantum information theory: Lecture 16: Werner states, twirling and the no-cloning theorem* (2022). URL https://www.uio.no/studier/emner/matnat/math/MAT4430/v22/lecture-notes/lecture16.pdf.

[32] D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal und A. V. Thapliyal, *Evidence for bound entangled states with negative partial transpose*, Physical Review A, **61**(6) (2000). URL http://dx.doi.org/10.1103/physreva.61.062312.

[33] A. Peres, *Separability Criterion for Density Matrices*, Phys. Rev. Lett., **77**, 1413 (1996). URL http://dx.doi.org/10.1103/PhysRevLett.77.1413.

[34] M. Horodecki, P. Horodecki und R. Horodecki, *Separability of mixed states: necessary and sufficient conditions*, Physics Letters A, **223**(1-2), 1 (1996). URL http://dx.doi.org/10.1016/s0375-9601(96)00706-2.

[35] E. Størmer, *Positive linear maps of operator algebras*, Acta Mathematica, **110**(none), 233 (1963). URL http://dx.doi.org/10.1007/BF02391860.

[36] S. Woronowicz, *Positive maps of low dimensional matrix algebras*, Reports on Mathematical Physics, **10**(2), 165 (1976). URL http://dx.doi.org/https://doi.org/10.1016/0034-4877(76)90038-0.

[37] R. Roy und A. Nath, *Introduction to Quantum Gates : Implementation of Single and Multiple Qubit Gates*, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, S. 385–392 (2021). URL http://dx.doi.org/10.32628/CSEIT217697.