UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

Determinación variacional del entrelazamiento geométrico de estados puros multi-qubit en computadores cuánticos
Variational determination of pure multi-qubit geometrical entanglement in NISQ computers

Por: Andrés Damián Muñoz Moller

Tesis presentada a la Facultad de Ciencias Físicas de la Universidad de Concepción para optar al grado académico de Magíster en Ciencias con Mención en Física

Marzo 2022

Concepción, Chile

Profesor Guía: Dr. Aldo Delgado Hidalgo

AGRADECIMIENTOS

Quiero agradecer a mi madre y a mi padre por el infinito apoyo que me han brindado durante toda mi vida, por la estupenda educación que me otorgaron, por estimular mi curiosidad desde temprana edad y por siempre velar por mi felicidad. Similarmente, quiero agradecer a mis hermanas por el continuo apoyo y cariño incondicional.

Quiero agradecer también a mi tutor, Dr. Aldo Delgado, tanto por su excelente disposición y ayuda para el desarrollo de este trabajo, como también por su apoyo durante mi formación como científico. Quiero agradecer a mis compañeros de estudios que me han acompañado por el arduo pero apasionante camino que significa estudiar física, y en particular quiero agradecer a Luciano Pereira y Leonardo Zambrano por el tiempo que dedicaron para enseñarme y por la ayuda entregada durante mi magíster.

Finalmente, quiero agradecer a mis amigos, que me han acompañado en todas las etapas de mi vida; desde mi infancia en Valdivia hasta mi vida universitaria en Concepción. En especial quiero agradecer a mis hermanos de la Burschenschaft Montania, por los inolvidables recuerdos que hemos compartido y por todas las enseñanzas de vida que esta institución me ha otorgado.

Contents

AGRADECIMIENTOS ii

List of Figures v

Resumen vi

Abstract vii

1 Introduction 1

2 Quantum Mechanics 4
2.1 Probability Theory . . . 4
2.1.1 Events and Probabilities . . . 4
2.1.2 Discrete Random Variables . . . 5
2.1.3 Continuous Random Variables . . . 7
2.2 Hilbert space and linear operators . . . 10
2.3 Postulates of Quantum Mechanics . . . 16
2.4 Density Matrix . . . 19
2.5 Schmidt Decomposition and purification . . . 22

3 Quantum Entanglement 26
3.1 Entanglement . . . 26
3.2 Distance between Quantum States . . . 27
3.2.1 Trace distance . . . 28
3.2.2 Fidelity . . . 29
3.3 Quantum and LOCC operations . . . 31
3.3.1 LOCC (Local Operations and classical communication) operations . . . 32
3.3.2 Classes of equivalence . . . 34
3.4 Axioms on entanglement measures . . . 36
3.4.1 Convex Roof Construction . . . 37
3.5 Bipartite Entanglement . . . 39
3.5.1 Entanglement Cost and Distillable Entanglement . . . 39
3.5.2 Entanglement of formation and Concurrence . . . 40
3.5.3 Distance based measures of entanglement . . . 41
3.6 Multipartite entanglement . . . 44

3.6.1 Three qubits entanglement 44
3.6.2 Entanglement classes for the general case 46
3.6.3 Symmetric states 48
3.6.4 Multipartite Entanglement Measures 48
3.6.4.1 Schmidt measure 50
3.6.4.2 Multipartie relative entropy of entanglement . . . 50
3.6.4.3 Geometric measure of entanglement (GME) . . . 50

4 Variational Algorithm for Geometric Entanglement Measure 52
4.1 Complex Simultaneous Perturbation Stochastic Approximation (CSPSA) 52
4.2 Variational quantum algorithms 53
4.3 Variational determination of geometrical entanglement 56
4.4 Numerical simulations 59
4.5 Experimental results 64

5 Conclusion 68

A Monte-Carlo Simulations 70

B Haar Measure 73

C Matrix Product States (MPS) 75

Bibliography 78

List of Figures

4.3.1 VDGE quantum circuit example for an arbitrary three-qubit state. 57

4.3.2 Flowchart of the VDGE algorithm. 59

4.4.1 Geometric measure of entanglement of the state $|\mathrm{GW}(s,\varphi)\rangle$ as a function of s for $\varphi = 0, \pi/4, \pi/2, \pi$. 60

4.4.2 Difference between the GME E_G obtained with the Basin-hopping optimization method and the estimated GME $\hat{E}_G$ obtained with VDGE versus the number of iterations k using 5 different initial states. 61

4.4.3 Difference between the GME E_G obtained with the Basin-hopping optimization method and the estimated GME $\hat{E}_G$ obtained with VDGE versus the number of iterations k using 20 different initial states. 61

4.4.4 Difference between the GME E_G obtained with the Basin-hopping optimization method and the estimated GME $\hat{E}_G$ obtained with VDGE versus the number of qubits. 62

4.4.5 Difference between the GME E_G obtained with the Basin-hopping optimization method and the estimated GME $\hat{E}_G$ obtained with VDGE in a simulation based on matrix product states techniques versus the number of iterations. 64

4.5.1 Experimental GME for a GHZ state of three qubits versus the number of iterations. 66

4.5.2 Experimental GME for a GHZ state of four qubits versus the number of iterations. 66

4.5.3 Experimental GME for a GHZ state of five qubits versus the number of iterations. 67

Resumen

Los algoritmos cuánticos, como el algoritmo de Shor o el algoritmo de Grover, prometen una mejora significativa de la complejitud temporal por sobre sus equivalentes clásicos. Los aparatos cuánticos actuales, conocidos como computadores NISQ (*Noisy intermediate-scale quantum*), se caracterizan por presentar altos niveles de ruido en la realización fisica de qubits y operaciones cuánticas. Por otro lado, el entrelazamiento es reconocido como uno de los principales recursos en el estudio de la computación cuántica y la información cuántica. Sin embargo, las actuales limitaciones técnicas dificultan la aplicación de métodos convencionales para caracterizar entrelazamiento. En esta tesis se propone un algoritmo variacional para estimar la medida geométrica de entrelazamiento de estados puros multi-qubit. El algoritmo es robusto ante ruido y requiere solo de mediciones y compuertas de un qubit, por lo que es adecuado para aparatos NISQ. Esto es demostrado aplicando el método con éxito en los computadores de IBM Quantum para estados Greenberger-Horne-Zeilinger de 3, 4, y 5 qubits. Simulaciones numéricas con estados aleatorios muestran robustez y precisión del método. La escalabilidad del protocolo es demostrada numéricamente mediante técnicas de *matrix product states* (MPS) hasta 25 qubits.

Abstract

Quantum algorithms, like Shor's algorithm or Grover's algorithm, promise a significant improvement in time complexity over their classical equivalents. Current quantum devices, known as NISQ (*Noisy intermediate-scale quantum*), are characterized for presenting high levels of noise in the physical realizations of qubits and quantum operations. On the other hand, entanglement is recognized as one of the core resources in the study of quantum computing and quantum information. Nonetheless, current technical limitations difficult the applicability of conventional methods to characterize entanglement. In this thesis, a variational algorithm to estimate the geometric measure of entanglement of multi-qubit pure states is proposed. The algorithm is robust against noise and requires only single-qubit gates and measurements, so it is well suited for NISQ devices. This is demonstrated by successfully implementing the method on IBM Quantum devices for Greenberger-Horne-Zeilinger states of 3, 4, and 5 qubits. Numerical simulations with random states show the robustness and accuracy of the method. The scalability of the protocol is numerically demonstrated via matrix product states (MPS) techniques up to 25 qubits.

Chapter 1

Introduction

Entangled quantum states [1], which were first discussed as an argument against the completeness of Quantum mechanics [2], are considered nowadays a distinguishing feature of this theory. Entanglement is also widely recognized as one of the core resources in the study of quantum computing and quantum information [3]; the success of quantum algorithms such as Shor's algorithm [4] or Grover's algorithm [5] is tied to proper implementations of non-local gates in quantum computers, while protocols like quantum teleportation [6], superdense coding [7] and quantum key distribution [8] rely on entangled states between two or more parties. Entanglement is also needed for both quantum sensing [9] and quantum metrology [10, 11].

Due to the role played by entanglement in the understanding of quantum mechanics and its many feasible applications, several functions have been developed to quantify it, known as entanglement measures. Nonetheless, they are generally difficult to compute even for the bipartite case, and cannot be easily determined experimentally. Some popular examples include the concurrence [12] and the relative entropy of entanglement [13, 14]. A proper entanglement monotone is the the geometric measure of entanglement (GME), first introduced by Shimony [15] for bipartite pure states and later generalized by Barnum and Linden [16] to the multipartite case. This measure characterizes the entanglement of a pure state $|\psi\rangle$ as the distance to the nearest separable pure state $|\phi\rangle$. This entanglement measure has been used, for example, to quantify the difficulty to distinguish multipartite quantum states using local operations [17], to study quantum phase transitions in

spin models [18] and to quantify how well a state serves as the input to Grover's search algorithm [19].

The GME of a state $|\psi\rangle$ can be obtained by maximizing the fidelity $F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2$ in the set $\{|\phi\rangle\}$ of separable pure states. This is an experimentally accessible quantity since its evaluation only requires local measurements. In particular, the fidelity can be efficiently measured in state-of-the-art quantum devices, known as noisy intermediate-scale quantum (NISQ) [20] computers. This class of devices is characterized by a low number of available qubits, limited connectivity between qubits, low coherence times, and noisy entangling gates that restrict the depth of the circuits. Within this adverse scenario, hybrid quantum-classical variational algorithms are among the most popular [21, 22] strategies to achieve quantum advantage, showing promising results in quantum chemistry to find large Hamiltonian eigenvalues [23, 24, 25, 26], solving tasks of quantum metrology [27, 28] and being applied in quantum machine learning [29].

In this thesis, the problem of experimentally measuring the amount of entanglement of multi-qubit pure states in quantum hardware is addressed, proposing an algorithm for the variational determination of geometrical entanglement (VDGE). This variational quantum algorithm employs a quantum device to estimate the fidelity between the target state $|\psi\rangle$ and a trial state, which is used to optimize a parameterized quantum circuit with a classical optimization method: the complex simultaneous perturbation stochastic approximation (CSPSA) algorithm.

The algorithm is shown to correctly reproduce the value of the GME for arbitrary superpositions of Greenberger-Horne-Zeilinger (GHZ) and W states. The overall accuracy achieved by the method is studied using Monte Carlo numerical experiments for states of $n = 2, 3, 4, 5, 6$ qubits. The feasibility of the VDGE approach to characterize the entanglement of states of larger numbers of qubits is also studied by means of matrix product state (MPS) techniques. Finally, the experimental results are obtained by applying the method to the measurement of the GME of a GHZ state in IBM Quantum Falcon Processors [30] `ibmq_lima` and `ibmq_bogota` for $n = 3, 4, 5$ qubits.

In chapter 2, the necessary mathematical tools and the foundations of quantum mechanics are presented. In chapter 3, the basic definitions of entanglement and measures of entanglement are presented, along with some common examples.

In chapter 4, the proposed algorithm to measure the geometric measure of entanglement is explained, with the results of both the numerical simulations and the experiments in quantum hardware. In chapter 5, the conclusions of this work are presented, with possible extensions for future study.

Chapter 2

Quantum Mechanics

The following chapter presents the fundamental tools to formulate the theory of quantum mechanics, which are going to be relevant for the correct understanding of this work.

2.1 Probability Theory

The theory of probability is essential for quantum mechanics, so the following section is dedicated to present the basic notions of probability and random variables [31, 32].

2.1.1 Events and Probabilities

A probability model is constructed according to a particular situation or *experiment*, whose results are known as *outcomes* and denote as ω. The set of all possible outcomes of an experiment is known as the *sample space* Ω. The outcomes can be either finite, countably infinite, or uncountably infinite. If Ω is countable, we call it a *discrete sample space*; if it is uncountable, we call it continuous sample space.

An *event* E corresponds to a subset of Ω. The set of events $\{E_i\}_{i=1}^{n}$ is *mutually exclusive* if any pair of events have no element in common, that is, if $E_i \cap E_j = \emptyset$ for every i, j satisfying $i \neq j$.

The *probability* P is a measure on the sample space that assigns to each event E the likelihood for it to occur, and must satisfy the following axioms:

1. $P(E) \geq 0, \quad \forall E.$
2. If $E = \Omega$, then $P(E) = 1$.
3. If the events $\{E_i\}_{i=1}^{n}$ are mutually exclusive, then

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P\left(E_i\right). \tag{2.1.1}$$

The axioms of probability theorem imply a number of rules that are useful for calculating probabilities, for example:

1. For any event E,

$$P(E) = 1 - P\left(E^c\right) \tag{2.1.2}$$

   where the event E^c consists of all outcomes that are not in E.

2. For any two events A and B,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{2.1.3}$$

In some experiments, the probability of an event E_1 depends of the outcome of another event E_2 that occurs before. In this case, if $P(E_2) > 0$ then the *conditional probability* that E_1 occurs given that E_2 occurs is defined to be

$$P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)}. \tag{2.1.4}$$

Two events E_1 and E_2 are said to be *independent* if $P(E_1 \cap E_2) = P(E_1)P(E_2)$. A relationship between the conditional probabilities $P\left(E_1|E_2\right)$ and $P\left(E_2|E_1\right)$ can be obtained using the *Bayes' rule*:

$$P\left(E_1|E_2\right) = \frac{P\left(E_2|E_1\right) P\left(E_1\right)}{P\left(E_2\right)}. \tag{2.1.5}$$

2.1.2 Discrete Random Variables

A *random variable* X is a function that assigns a numerical value $X(\omega)$ to each element ω of the sample space, and it is said to be *discrete* if its set possible values is either finite or countably infinite. The set of possible values of X is called

the *range* of X and is denoted by I. The *probability mass function* of a discrete random variable X is defined by $P(X = x)$ for $x \in I$, where

$$P(X = x) = P(\{\omega : X(\omega) = x\}). \tag{2.1.6}$$

Let X and Y be two discrete random variables defined on a same sample space with probability measure P. The *joint probability mass function* of X and Y is noted by $P(X = x, Y = y\ )$, and corresponds to the probability assigned by P to the intersection of the two sets $A = \{\omega : X(\omega) = x\}$ and $B = \{\omega : Y(\omega) = y\}$.

The *marginal probability mass function* can be obtained from the joint probability mass function by

$$p_X(x) = \sum_y P(X = x, Y = y), \quad p_Y(y) = \sum_x P(X = x, Y = y). \tag{2.1.7}$$

The *conditional probability mass function* of X given that $Y = y$ is defined by

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \tag{2.1.8}$$

for any fixed y with $P(Y = y) > 0$.

For discrete random variables X and Y, the unconditional probability $P(X = a)$ can be calculated from

$$P(X = a) = \sum_b P(X = a \mid Y = b)P(Y = b). \tag{2.1.9}$$

This rule is called the *law of conditional probability*. On the other hand, two discrete random variables X and Y are said to be independent if and only if

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y) \tag{2.1.10}$$

for any two real numbers x and y, where $P(X \leq x, Y \leq y)$ represents the probability of occurrence of both event $\{X \leq x\}$ and event $\{Y \leq y\}$.

The *expectation value* of a discrete random variable X corresponds to a weighted average of its range I, where the weights are given by the probability mass function,

and it is defined as

$$\mathbb{E}(X) = \sum_{x \in I} xP(X = x), \tag{2.1.11}$$

and satisfy the following rules:

1. Linearity: For any two discrete random variables X and Y,

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y), \tag{2.1.12}$$

   for any constants a, b provided that $\mathbb{E}(X)$ and $\mathbb{E}(Y)$ exist and are finite.

2. For any two random variables X and Y, the conditional expectation of X given that $Y = y$ is defined by

$$\mathbb{E}(X|Y = b) = \sum_{x} xP(X = x \mid Y = y) \tag{2.1.13}$$

   for each y with $P(Y = y) > 0$ (assuming that the sum is well-defined).

3. For any two independent discrete random variables X and Y,

$$\mathbb{E}(X \cap Y) = \mathbb{E}(X)\mathbb{E}(Y), \tag{2.1.14}$$

   provided that $\mathbb{E}(X)$ and $\mathbb{E}(Y)$ exist and are finite.

4. Substitution rule: For any function g of the discrete random variable X,

$$\mathbb{E}[g(X)] = \sum_{x \in I} g(x)P(X = x) \tag{2.1.15}$$

   provided that $\sum_{x \in I} |g(x)|P(X = x) < \infty$.

2.1.3 Continuous Random Variables

The random variable X is said to be *continuous* if a function $f(x)$ exists such that

$$P(X \leq a) = \int_{-\infty}^{a} f(x)dx \quad \text{for each real number } a, \tag{2.1.16}$$

where the function $f(x)$ is known as the *probability density function* of X and satisfies

$$f(x) \geq 0 \quad \text{for all } x \quad \text{and} \quad \int_{-\infty}^{\infty} f(x)dx = 1. \tag{2.1.17}$$

The continuous random variables X and Y are said to have a *joint probability density function* $f(x, y)$ if the joint cumulative probability distribution function $P(X \leq a, Y \leq b)$ allows for the representation

$$P(X \leq a, Y \leq b) = \int_{x=-\infty}^{a} \int_{y=-\infty}^{b} f(x, y)dxdy, \quad -\infty < a, b < \infty, \tag{2.1.18}$$

where the function $f(x, y)$ satisfies

$$f(x, y) \geq 0 \quad \text{for all } x, y \quad \text{and} \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)dxdy = 1. \tag{2.1.19}$$

From this, it is possible to define the *marginal probability density functions* of X and Y as

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy, \quad -\infty < x < \infty, \tag{2.1.20}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx, \quad -\infty < y < \infty. \tag{2.1.21}$$

Using the previous definitions, the *conditional probability density function* of X given that $Y = y$ is defined by

$$f_X(x|y) = \frac{f(x, y)}{f_Y(y)}, \quad -\infty < x < \infty \tag{2.1.22}$$

for any fixed y with $f_Y(y) > 0$. The law of conditional probability for the random variables X and Y that are continuously distributed with a joint density function $f(x, y)$ states that

$$P(X \leq a) = \int_{-\infty}^{\infty} P(X \leq a|Y = y) f_Y(y)dy \tag{2.1.23}$$

where $f_Y(y)$ is the marginal density function of Y. On the other hand, two continuous random variables X and Y are said to be independent if and only if

$$f(x, y) = f_X(x) f_Y(y) \quad \text{for all } x, y. \tag{2.1.24}$$

The expectation value of a continuous random variable X with probability density function $f(x)$ is defined by

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx \tag{2.1.25}$$

provided that the integral $\int_{-\infty}^{\infty} |x| f(x) dx$ is finite, and satisfy the following rules:

1. Linearity: For any two continuous random variables X and Y,

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y) \tag{2.1.26}$$

   for any constants a, b provided that $\mathbb{E}(X)$ and $\mathbb{E}(Y)$ exist.

2. For any two random variables X and Y continuously distributed with joint probability density function $f(x, y)$, the conditional expectation of X given that $Y = y$ is defined by

$$\mathbb{E}(X|Y = y) = \int_{-\infty}^{\infty} x f_X(x|y) dx. \tag{2.1.27}$$

   for each y with $f_Y(y) > 0$ (assuming that the integral is well-defined).

3. For any two independent continuous random variables X and Y,

$$\mathbb{E}(X \cap Y) = \mathbb{E}(X)\mathbb{E}(Y), \tag{2.1.28}$$

   provided that $\mathbb{E}(X)$ and $\mathbb{E}(Y)$ exist and are finite.

4. Substitution rule: For any function g of the continuous random variable X,

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx, \tag{2.1.29}$$

   provided that the integral exists.

5. Two-dimensional substitution rule: If the random variables X and Y have a joint probability density function $f(x, y)$, then

$$\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy \tag{2.1.30}$$

   for any function $g(x, y)$ provided that the integral is well-defined.

2.2 Hilbert space and linear operators

In quantum mechanics a physical state is represented by a normalized state vector in a complex vector space called *Hilbert space* [33]. Both the mathematical properties and structure of Hilbert spaces and the understanding of linear operators are essential for a proper understanding of the formalism of quantum mechanics [34, 35, 36]. In this section and the rest of this work we will employ Dirac's notation [37].

A *linear vector space* consists of a set of vectors V and scalar field F (that can be either real or complex), with a given addition $(+)$ and multiplication rule $(\cdot)$.

Let $|a\rangle, |b\rangle, |c\rangle \in V$ and $\alpha, \beta, \gamma \in F$. The addition rule has the following properties:

- Closure under addition: $|a\rangle + |b\rangle \in V$.
- Commutativity: $|a\rangle + |b\rangle = |b\rangle + |a\rangle$.
- Associativity: $(|a\rangle + |b\rangle) + |c\rangle = |a\rangle + (|b\rangle + |c\rangle)$.
- Existence of a zero vector: $|0\rangle + |a\rangle = |a\rangle + |0\rangle = |a\rangle$.
- Existence of an inverse vector: $|a\rangle + |-a\rangle = |-a\rangle + |a\rangle = |0\rangle$.

Similarly, the multiplication of vectors by scalars satisfy the following properties:

- Closure under scalar multiplication: $\alpha\,|a\rangle \in V$.
- Distributivity: $\alpha(|a\rangle + |b\rangle) = \alpha\,|a\rangle + \alpha\,|b\rangle\,, \quad (\alpha + \beta)\,|a\rangle = \alpha\,|a\rangle + \beta\,|a\rangle$.
- Associativity: $\alpha(\beta\,|a\rangle) = (ab)\,|a\rangle$.
- Existence of an identity scalar: $1\,|a\rangle = |a\rangle\,1 = |a\rangle$
- Existence of a zero scalar: $0\,|a\rangle = |a\rangle\,0 = |0\rangle$

where for simplicity we omitted the dot in the scalar multiplication.

A Hilbert space $\mathcal{H}$ is a linear vector space with a defined *inner scalar product* $(\cdot,\cdot) : \mathcal{H} \times \mathcal{H} \longrightarrow \mathbb{C}$, which satisfy the following properties:

- Hermiticity: $(|a\rangle, |b\rangle) = (|b\rangle, |a\rangle)^*$.
- Linearity in the first argument: $(|c\rangle, \alpha\,|a\rangle + \beta\,|b\rangle) = \alpha(|c\rangle, |a\rangle) + \beta(|c\rangle, |b\rangle)$.

- Positivity: $(|a\rangle, |a\rangle) \geq 0$, where the equality holds only for $|a\rangle = |0\rangle$.

The inner scalar product in the Dirac's notation is $(|a\rangle, |b\rangle) \equiv \langle a|b\rangle$, where $\langle b| = |b\rangle^\dagger$ is the *dual vector* of the vector $|b\rangle$ in the *dual space* $\mathcal{H}^*$. The *norm* of a vector $|a\rangle$ is defined as:

$$\| |a\rangle \| \equiv \sqrt{\langle a|a\rangle} \tag{2.2.1}$$

and satisfies the *Cauchy-Schwarz inequality*

$$|\langle a|b\rangle| \leq \| |a\rangle \| \, \| |b\rangle \| \tag{2.2.2}$$

for any vectors $|a\rangle, |b\rangle \in \mathcal{H}$.

Finally, a Hilbert space is *complete*, that is, every Cauchy sequence $\{|\psi_i\rangle\}_{i=1}^n \in \mathcal{H}$ converges to an element of $\mathcal{H}$. For any $|\psi_n\rangle$, the relation

$$\lim_{n,m\to\infty} \| |\psi_n\rangle - |\psi_m\rangle \| = 0, \tag{2.2.3}$$

defines a unique limit $|\psi\rangle$ of $\mathcal{H}$ such that

$$\lim_{n\to\infty} \| |\psi\rangle - |\psi_n\rangle \| = 0 \tag{2.2.4}$$

Given a Hilbert space $\mathcal{H}$, a spanning set $S = \{|e_i\rangle\}_i$ is a set of vectors such that each $|\psi\rangle \in \mathcal{H}$ can be written as a linear combination of elements of S,

$$|\psi\rangle = \sum_i \alpha_i |e_i\rangle \tag{2.2.5}$$

If the set S is also *linearly independent*, that is, for any set of n non-zero complex numbers $\{\alpha_i\}$ it satisfy that

$$\sum_i \alpha_i |e_i\rangle = 0, \tag{2.2.6}$$

then the set S is called a *basis* for $\mathcal{H}$. The total number d of basis vectors is the dimension of the Hilbert space. This dimension can be finite or infinite, while the index i can be discrete or continuous. A basis is called *orthonormal* if each

element is a unit vector, and distinct vectors in the basis are orthogonal, that is, $\langle e_i|e_j\rangle = \delta_{ij}$, where i and j are both chosen from the index set.

A *linear operator* between Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ is defined to be any function $A \in \mathcal{L}(\mathcal{H}, \mathcal{H}') : \mathcal{H} \to \mathcal{H}'$ which is linear in its inputs, that is,

$$A\left(\sum_i \alpha_i |e_i\rangle\right) = \sum_i \alpha_i A |e_i\rangle . \tag{2.2.7}$$

A linear operator $A \in \mathcal{L}(\mathcal{H})$ is said to be defined on a Hilbert space $\mathcal{H}$, if it maps $\mathcal{H}$ into itself.

The space of *bounded linear operators* from $\mathcal{H}$ to $\mathcal{H}'$ is denoted by $\mathcal{B}(\mathcal{H}, \mathcal{H}')$. A linear operator $A \in \mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H}, \mathcal{H})$ is said to be bounded if

$$\|A|\psi\rangle\| \leq c\||f\rangle\|, \quad \forall|\psi\rangle \in \mathcal{H}, \quad c \in \mathbb{R}. \tag{2.2.8}$$

Let's suppose that $\{|e_i\rangle\}_i^m$ is a basis for $\mathcal{H}$ and $\{|g_i\rangle\}_i^n$ is a basis for $\mathcal{H}'$. Then for each j in the range $1, \ldots, m$, there exist complex numbers A_{1j} through A_{nj} such that

$$A |e_j\rangle = \sum_i A_{ij} |g_i\rangle \tag{2.2.9}$$

The matrix whose entries are the values A_{ij} is said to form a *matrix representation* of the operator A. This matrix can also be represented using the *outer product representation*. For two vectors $|\psi\rangle = [a_1 a_2 \ldots a_n]^T$ and $|\phi\rangle = [b_1 b_2 \ldots b_n]^T$, the outer product is denoted by $|\psi\rangle\langle\phi|$ and represented by a matrix given by

$$|\psi\rangle\langle\phi| = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1^* & b_2^* & \ldots & b_n^* \end{bmatrix} = \begin{bmatrix} a_1 b_1^* & a_1 b_2^* & \ldots & a_1 b_n^* \\ a_2 b_1^* & a_2 b_2^* & \ldots & a_2 b_n^* \\ \vdots & \vdots & \ldots & \vdots \\ a_n b_1^* & a_n b_2^* & \ldots & a_n b_n^* \end{bmatrix} \tag{2.2.10}$$

A linear operator A with matrix elements A_{ij} is the expansion

$$A \equiv \sum_{i,j} A_{ij} |e_i\rangle\langle e_j| \tag{2.2.11}$$

with matrix elements $A_{ij} = \langle e_i|A|e_j\rangle$. An important result of the outer product

representation is the *completeness relation*

$$\sum_i |e_i\rangle\langle e_i| = I \tag{2.2.12}$$

where I is the identity operator.

An *eigenvector* of a linear operator A on a Hilbert space is a non-zero vector $|\psi\rangle$ such that

$$A|\psi\rangle = \lambda|\psi\rangle, \tag{2.2.13}$$

where λ is a complex number known as the *eigenvalue* of A corresponding to $|\psi\rangle$. The eigenvalues can be obtained with the *characteristic function* defined as

$$c(\lambda) \equiv \det|A - \lambda I|. \tag{2.2.14}$$

The solutions of the characteristic equation $c(\lambda) = 0$ are the eigenvalues of the operator A. The *eigenspace* corresponding to an eigenvalue λ is the set of vectors which have eigenvalue λ, and is *degenerate* if it is more than one dimensional.

A *diagonal representation* for an operator A on a Hilbert space is a representation $A = \sum_i \lambda_i |e_i\rangle\langle e_i|$, where the vectors $|e_i\rangle$ form an orthonormal set of eigenvectors for A, with corresponding eigenvalues λ_i. An operator is said to be diagonalizable if it has a diagonal representation.

For a linear operator A on the Hilbert space $\mathcal{H}$, there exist a unique linear operator $A^\dagger$ known as the *adjoint* operator of A, which satisfies that

$$(|\psi\rangle, A|\phi\rangle) = \langle\psi|A|\phi\rangle = \langle\phi|A^\dagger|\psi\rangle = \left(A^\dagger|\psi\rangle, |\phi\rangle\right), \quad \forall|\psi\rangle, |\phi\rangle \in \mathcal{H}. \tag{2.2.15}$$

An operator is called *Hermitian* or *self-adjoint* if

$$A = A^\dagger \tag{2.2.16}$$

The eigenvalues of these operators are always real. This is an important property in quantum mechanics, because physical *observables* are represented by self-adjoint operators.

A special subclass of self-adjoint operators is the *positive semidefinite operators*. A positive operator A is called positive semidefinite if it satisfies that

$$\langle\psi|A|\psi\rangle \geq 0, \quad \forall|\psi\rangle \neq 0 \tag{2.2.17}$$

which implies that A have non-negative real eigenvalues.

Another important type of self-adjoint operators are the *projection operators*, whose action is to project the state along another state $|a\rangle$:

$$\mathbb{P}_a = |a\rangle\langle a|. \tag{2.2.18}$$

The projection operator satisfy the following properties:

1. Idempotency: $\mathbb{P}_a^2 = \mathbb{P}_a$.
2. Mutual orthogonality: $\mathbb{P}_a\mathbb{P}_b = \delta_{ab}\mathbb{P}_b$.
3. Completeness: $\sum_a \mathbb{P}_a = \sum_a |a\rangle\langle a| = I$

An operator is called *normal* if

$$NN^\dagger = N^\dagger N. \tag{2.2.19}$$

The concept of normal operators is very useful because they have a spectral decomposition, according to the following theorem:

Theorem 1. *(Spectral decomposition) Any normal operator N on a Hilbert space $\mathcal{H}$ is diagonal with respect to some orthonormal basis for $\mathcal{H}$, that is, it can be written as*

$$N = \sum_i \lambda_i \left|e_i\right\rangle \left\langle e_i\right| \tag{2.2.20}$$

where $\{|e_i\rangle\}_i$ are the orthogonal eigenvectors of N with eigenvalues $\{\lambda_i\}_i$.

An operator U is called *unitary* if

$$UU^\dagger = U^\dagger U = I. \tag{2.2.21}$$

An important feature of unitary operators is that they preserve inner products between vectors. Let $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, then

$$(U|\psi\rangle, U|\phi\rangle) = \langle\psi|U^\dagger U|\phi\rangle = \langle\psi|I|\phi\rangle = \langle\psi|\phi\rangle = (|\psi\rangle, |\phi\rangle). \tag{2.2.22}$$

Let A be a normal operator with spectral decomposition $A = \sum_i \lambda_i |e_i\rangle \langle e_i|$, where λ_i are the eigenvalues of A and $|e_i\rangle$ the corresponding eigenvectors, and f a function from the complex numbers to the complex numbers. The function of the normal operator $f(A)$ is defined as

$$f(A) = \sum_i f(\lambda_i) |e_i\rangle \langle e_i| . \tag{2.2.23}$$

An important matrix function is the *trace* of a matrix. The trace of A is defined to be the sum of its diagonal elements,

$$\mathrm{tr}(A) \equiv \sum_i A_{ii} = \sum_i \langle e_i|A|e_i\rangle \tag{2.2.24}$$

Let A, B, C be arbitrary matrices, the trace function satisfy the following properties:

1. Invariance under cyclic permutations: $\mathrm{tr}(ABC) = \mathrm{tr}(CAB) = \mathrm{tr}(BCA)$
2. Linearity: $\mathrm{tr}(\alpha A \beta B]) = \alpha \,\mathrm{tr}(A) + \beta \,\mathrm{tr}(B)$, where α, β are complex numbers.
3. Invariance under the unitary similarity transformations $A \to UAU^\dagger$:

$$\mathrm{tr}\left(UAU^\dagger\right) = \mathrm{tr}\left(U^\dagger UA\right) = \mathrm{tr}(A) \tag{2.2.25}$$

The *commutator* between two operators A and B is defined to be

$$[A, B] \equiv AB - BA \tag{2.2.26}$$

If $[A, B] = 0$, that is, $AB = BA$, then we say A *commutes* with B. If two self-adjoint operators A and B commute, it is possible to *simultaneously diagonalize* both, that is, to write

$$A = \sum_i a_i |e_i\rangle\langle e_i| , \quad B = \sum_i b_i |e_i\rangle\langle e_i| \tag{2.2.27}$$

where $|e_i\rangle$ is some common orthonormal set of eigenvectors for A and B.

Finally, we present two useful decompositions for operators: the *polar* and *singular*

value decomposition.

Theorem 2. *(Polar decomposition) Let A be a linear operator on a Hilbert space $\mathcal{H}$. Then there exists unitary U and positive operators J and K such that*

$$A = UJ = KU, \tag{2.2.28}$$

where the unique positive operators J and K satisfying these equations are defined by $J \equiv \sqrt{A^\dagger A}$ and $K \equiv \sqrt{AA^\dagger}$. Moreover, if A is invertible then U is unique.

The expression $A = UJ$ is known as the left polar decomposition of A, and $A = KU$ the right polar decomposition of A.

Corollary 2.1. *(Singular value decomposition) Let $A \in \mathbb{R}^{m\times n}$ be any matrix. Then there exist unitary matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$, and a diagonal matrix $D \in \mathbb{R}^{m\times n}$ non-negative real numbers on the diagonal such that*

$$A = UDV \tag{2.2.29}$$

The diagonal elements D_{ii} of D are called the singular values of A.

2.3 Postulates of Quantum Mechanics

The following section presents the postulates required for the construction of the quantum theory [35, 38, 39].

Postulate 1. *The state of a closed quantum mechanical system corresponds to a unit vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$ known as the state space.*

A closed quantum system corresponds to a system which is not interacting in any way with other systems i.e it doesn't interchange information with another system.

We can write a general d-dimensional quantum state i.e. a *qudit* as

$$|\psi\rangle = \sum_{i=1}^{d} \alpha_i |e_i\rangle = \sum_{i=1}^{d} \langle e_i|\psi\rangle \, |e_i\rangle \tag{2.3.1}$$

where the complex coefficients α_i are interpreted as *probability amplitudes* for the state to be in the corresponding basis state $|e_i\rangle$.

As Hilbert spaces correspond to linear vector spaces, quantum states satisfy the *superposition principle*: a quantum particle can be in a linear combination state of any two other allowable states. Therefore, the state $|\psi\rangle$ is said to be in a *coherent superposition* of the basis states.

Because the state $|\psi\rangle$ is unitary, and since probabilities must add to 1, we get what is called the *normalization condition*

$$\langle\psi|\psi\rangle = \sum_{i=1}^{d} |\alpha_i|^2 = 1. \tag{2.3.2}$$

The simplest quantum mechanical system is the *qubit*, which has a two-dimensional state space. Let $|0\rangle$ and $|1\rangle$ form an orthonormal basis for that state space known as the *computational basis*. Then an arbitrary state vector in the state space can be written as

$$|\psi\rangle = a|0\rangle + b|1\rangle. \tag{2.3.3}$$

Postulate 2. *The evolution in time of a closed quantum system is described by a unitary transformation. A state $|\psi(t_0)\rangle$ in an initial time t_0 is related to the state $|\psi(t)\rangle$ at a time t by an unitary operator U,*

$$|\psi(t)\rangle = U(t_0, t)\,|\psi(t_0)\rangle\,. \tag{2.3.4}$$

A closed quantum system does not interchange information (i.e. energy and/or matter) with another system.

Postulate 3. *Quantum measurements are described by a collection $\{M_m\}$ of measurement operators acting on the state space. The index m refers to possible measurement outcome. If the state of the system is $|\psi\rangle$ before the measurement, the probability that the result m occurs is given by*

$$p(m) = \left\langle \psi \left| M_m^\dagger M_m \right| \psi \right\rangle, \tag{2.3.5}$$

and the state of the system after the measurement $|\psi'\rangle$ is

$$|\psi'\rangle = \frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}. \tag{2.3.6}$$

The measurement operators satisfy the completeness equation,

$$\sum M_m^\dagger M_m = I. \tag{2.3.7}$$

An important class of measurements is the *projective measurements*. A projective measurement is described by an observable, O on the state space of the system being observed, with spectral decomposition,

$$O = \sum_m m \left|\phi_m\right\rangle\!\left\langle\phi_m\right| = \sum_m m\mathbb{P}_m, \tag{2.3.8}$$

where $\mathbb{P}_m = \left|\phi_m\right\rangle\!\left\langle\phi_m\right|$ is the projector with eigenstate $|\phi_m\rangle$ onto the eigenspace of O with eigenvalue m. The probability of getting result m upon measuring the state $|\psi\rangle$, is given by

$$p(m) = \langle\psi|\mathbb{P}_m|\psi\rangle\,, \tag{2.3.9}$$

and the state of the system after the measurement $|\psi'\rangle$ is

$$|\psi'\rangle = \frac{\mathbb{P}_m|\psi\rangle}{\sqrt{p(m)}}. \tag{2.3.10}$$

The expected value of the observable O is defined as

$$\begin{aligned}
\mathbb{E}(O) &= \sum_m mp(m) && (2.3.11)\\
&= \sum_m m \left\langle\psi\right|\left|\mathbb{P}_m\right|\left.\psi\right\rangle && (2.3.12)\\
&= \langle\psi|\left(\sum_m m\mathbb{P}_m\right)|\psi\rangle && (2.3.13)\\
&= \langle\psi|O|\psi\rangle. && (2.3.14)
\end{aligned}$$

On a different topic, let's consider the state $|\psi'\rangle = e^{i\theta}|\psi\rangle$, where $|\psi\rangle$ is a state vector, and θ is a real number. We say that the state $|\psi'\rangle$ is equal to $|\psi\rangle$, up to the *global phase* factor $e^{i\theta}$. It is easy to prove that the statistics of measurement predicted for these two states are the same. Indeed, let M_m be a measurement operator associated to some quantum measurement, the respective probabilities

for outcome m occurring are

$$\langle\psi'|M_m^\dagger M_m|\psi'\rangle = \langle\psi|e^{-i\theta}M_m^\dagger M_m e^{i\theta}|\psi\rangle = \langle\psi|M_m^\dagger M_m|\psi\rangle\,. \tag{2.3.15}$$

The set of all states differing by a global phase is called a *ray* in Hilbert space. Thus the space of quantum states of a system is the space of rays in Hilbert space, also called the *projective Hilbert space*.

On the other hand, two states are said to differ by a *relative phase* in some basis if each of the probability amplitudes in that basis is related by a phase factor $e^{i\theta}$.

Postulate 4. *The Hilbert space of a composite system $\mathcal{S}$ is the tensor product of Hilbert spaces of the component physical systems $A, B, C \ldots$.*

$$\mathcal{H}_\mathcal{S} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \cdots\,. \tag{2.3.16}$$

If we have systems numbered 1 through n, and system number i is prepared in the state $|\psi_i\rangle$, then the joint state of the total system $|\Psi\rangle \in \mathcal{S}$ is

$$|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle = |\psi_1\rangle\,|\psi_2\rangle \cdots |\psi_n\rangle\,. \tag{2.3.17}$$

States of $\mathcal{S}$ that can be represented in this form are called *separable states* or *product states*, while those that can't be written as a product of states of its component systems are called *entangled states*. The study of entanglement is tackled in the next chapter.

2.4 Density Matrix

A quantum system whose state $|\psi\rangle \in \mathcal{H}$ is known exactly is said to be in a pure state. The density matrix formulation provides a convenient description of quantum systems whose state is not completely known. Let $\{p_i, |\psi_i\rangle\}$ be an ensemble of pure states, that is, a probability distribution in which a state $|\psi_i\rangle$ occurs with probability p_i. The *density matrix* or *density operator* $\rho \in \mathcal{E}(\mathcal{H})$ for this system is defined as:

$$\rho = \sum_n p_n \left|\psi_n\right\rangle\!\left\langle\psi_n\right|, \tag{2.4.1}$$

where $\mathcal{E}(\mathcal{H})$ corresponds to the convex set of density operators supported by $\mathcal{H}$. The sum over states in this expression looks similar to the superposition of states, but it is considered an *incoherent superposition*, which comes from the fact that the relative phases of the states $|\psi_n\rangle$ are not available to us.

The density operator on a Hilbert space satisfies the following properties:

1. ρ is Hermitian.
2. ρ is a positive operator, that is, its eigenvalues are non-negative.
3. ρ has a trace equal to one.
4. Convexity: A set of operators $\{\rho_i\}$ form a convex set if

$$\rho = \lambda\rho_1 + (1-\lambda)\rho_2, \quad 0 < \lambda < 1, \tag{2.4.2}$$

   for every pair $\rho_1, \rho_2 \in \{\rho_i\}$.

The density operator of a pure state is simply $\rho = |\psi_n\rangle\!\langle\psi_n|$. A pure state satisfies $\mathrm{tr}\,(\rho^2) = 1$, while a mixed state satisfies $\mathrm{tr}\,(\rho^2) < 1$. The quantity $\mathrm{tr}\,(\rho^2)$ is known as the purity of the state ρ.

The maximally mixed state π is the density operator corresponding to a uniform ensemble of orthogonal states $\left\{\frac{1}{d}, |\psi\rangle\right\}$, where d is the dimension of the Hilbert space, and it is equal to

$$\pi \equiv \frac{1}{d} I. \tag{2.4.3}$$

In general, the eigenvectors and eigenvalues of a density matrix just indicate one of many possible ensembles that may give rise to a specific density matrix. For a pure state, there is a unique decomposition of ρ in the form of (2.4.1), and it consists of only one term.

On the other hand, a mixed state has no unique decomposition in terms of pure states. According to the convexity property, a given density operator ρ can be expressed as (2.4.2) in infinitely many ways, so that it is impossible to identify any unique component density operators ρ_1 and ρ_2.

The expectation value of the observable O in the density matrix representation can be obtained using the trace operation:

$$\begin{aligned}
\mathrm{tr}(\rho O) &= \sum_i \langle e_i|\rho O|e_i\rangle && (2.4.4)\\
&= \sum_i \langle e_i| \left\{ \sum_n p_n |\psi_n\rangle\langle\psi_n| O \right\} |e_i\rangle && (2.4.5)\\
&= \sum_{i,n} p_n \langle\psi_n|O|e_i\rangle \langle e_i|\psi_i\rangle && (2.4.6)\\
&= \sum_n p_n \langle\psi_n|O|\psi_n\rangle && (2.4.7)\\
&= \langle O\rangle_\rho. && (2.4.8)
\end{aligned}$$

We can reformulate the postulates defined earlier in terms of the density operator formalism.

Postulate 5. *The state of a closed quantum mechanical system is described by a density operator in $\mathcal{E}(\mathcal{H})$ i.e. a positive Hermitian operator with unit trace.*

Postulate 6. *The evolution in time of a closed quantum system is described by a unitary transformation. A state $\rho(t_0)$ in an initial time t_0 is related to the state $\rho(t)$ at a time t by an unitary operator U,*

$$\rho(t) = U(t_0, t)\rho(t_0)U^\dagger(t_0, t). \tag{2.4.9}$$

Postulate 7. *Quantum measurements are described by a collection $\{M_m\}$ of measurement operators acting on the state space. If the state of the system is ρ before the measurement, the probability that the result m occurs is given by*

$$p(m) = \mathrm{tr}\left(M_m^\dagger M_m \rho\right), \tag{2.4.10}$$

and the state of the system after the measurement ρ' is

$$\rho' = \frac{M_m \rho M_m^\dagger}{\mathrm{tr}\left(M_m^\dagger M_m \rho\right)}. \tag{2.4.11}$$

The measurement operators satisfy the completeness equation,

$$\sum_m M_m^\dagger M_m = I. \tag{2.4.12}$$

Postulate 8. *If we have systems numbered 1 through n, and system number i is prepared in the state ρ_i, then the joint state of the total system is $\rho_1 \otimes \rho_2 \otimes \cdots \rho_n$.*

Suppose we have physical systems A and B, whose state is described by a density operator ρ^{AB}. If subsystems A and B are given by Hilbert spaces spanned by the bases $\left\{\left|e_i^A\right\rangle\right\}$ and $\left\{\left|e_j^B\right\rangle\right\}$ respectively, we define the *partial trace* of ρ^{AB} with respect to subsystem A as

$$\rho^A = \mathrm{tr}_A\, \rho^{AB} = \sum_i \left\langle e_i^A \right| \rho^{AB} \left| e_i^A \right\rangle, \tag{2.4.13}$$

which will be an operator on the Hilbert space of subsystem B alone. The density operator ρ^A is known as the *reduced density operator* for system A. If the quantum system is a bipartite state i.e. is in the product state $\rho^{AB} = \rho \otimes \sigma$, where ρ is a density operator for system A, and σ is a density operator for system B. Then

$$\rho^A = \mathrm{tr}_B(\rho \otimes \sigma) = \rho\, \mathrm{tr}(\sigma) = \rho, \quad \rho^B = \sigma. \tag{2.4.14}$$

The partial trace can be understood as a way to obtain a marginal density function in quantum mechanics: the density matrix ρ^{AB} describes the probability distribution $p_{A,B}$ of the combined systems A and B, while the density matrix ρ^A describes the marginal density function p_A of system A alone.

2.5 Schmidt Decomposition and purification

Two useful tools for the study of composite systems are the *Schmidt decomposition* and *purification*.

Theorem 3. *Every bipartite pure state $|\psi\rangle$ in the Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ with $N_A = \dim \mathcal{H}_A$ and $N_B = \dim \mathcal{H}_B$ can be expressed in the form:*

$$|\psi\rangle = \sum_{i=1}^{N} \sqrt{\lambda_i}\, |e_i\rangle\, |f_i\rangle, \tag{2.5.1}$$

where $\{|e_i\rangle\}_{i=1}^{N}$ *is an orthonormal basis for* $\mathcal{H}_A$, $\{|f_i\rangle\}_{i=1}^{N}$ *is an orthonormal basis for* $\mathcal{H}_B$, *and* $N \leq \min\{N_A, N_B\}$.

The real numbers λ_i that appear in the Schmidt decomposition are known as the Schmidt coefficients, and they satisfy:

$$\sum_i \lambda_i = 1, \quad \lambda_i \geq 0. \tag{2.5.2}$$

The number of non-vanishing λ_i is called the Schmidt rank of the state $|\psi\rangle$, and it is equal to one if the state is separable. To obtain the Schmidt decomposition, it is easier to express the pure state in the form:

$$|\psi\rangle = \sum_{i=1}^{N_A}\sum_{j=1}^{N_B} C_{ij}\,|\tilde{e}_i\rangle\ |\tilde{f}_j\rangle\,, \tag{2.5.3}$$

where C is some complex valued matrix and the local bases are orthonormal bases for the Hilbert's space $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. Let the singular values of C be $\sqrt{\lambda_i}$. Following the singular value decomposition, there exist an unitary $N_A \times N_A$ matrix U and an unitary $N_B \times N_B$ matrix V such that:

$$C = UDV, \tag{2.5.4}$$

where D is a diagonal $N_A \times N_B$ matrix with elements $\sqrt{\lambda_i}$. The matrix elements of C can then be expressed as:

$$C_{ij} = \sum_{k=1}^{N_A}\sum_{l=1}^{N_B} U_{ik}\sqrt{\lambda_k}\delta_{kl}V_{lj} = \sum_{l=1}^{N} U_{il}\sqrt{\lambda_l}V_{lj}. \tag{2.5.5}$$

Replacing these elements in equation (2.5.3), we have:

$$|\psi\rangle = \sum_{i=1}^{N_A}\sum_{j=1}^{N_B}\sum_{l=1}^{N} U_{il}\sqrt{\lambda_l}V_{lj}\,|\tilde{e}_i\rangle\ |\tilde{f}_j\rangle\,. \tag{2.5.6}$$

Using U and V to effect changes of the bases in $\mathcal{H}_A$ and $\mathcal{H}_B$, we recover the Schmidt decomposition. Indeed, by redefining the new basis for each subsystem as:

$$|e_l\rangle \equiv \sum_{i=1}^{N_A} U_{il} \, |\tilde{e}_i\rangle \, , \quad |f_l\rangle \equiv \sum_{j=1}^{N_B} V_{lj} \, |\tilde{f}_j\rangle \tag{2.5.7}$$

we obtain:

$$|\psi\rangle = \sum_{l=1}^{N} \sqrt{\lambda_l} \, |e_l\rangle \, |f_l\rangle \, . \tag{2.5.8}$$

An important result of the Schmidt decomposition is at follows: Let $|\psi\rangle$ be a pure state of a composite system, AB. Then by the Schmidt decomposition $\rho^A = \sum_i \lambda_i \left|e_i^A\right\rangle\!\left\langle e_i^A\right|$ and $\rho^B = \sum_i \lambda_i \left|f_i^B\right\rangle\!\left\langle f_i^B\right|$, so the eigenvalues of ρ^A and ρ^B are identical, namely λ_i for both density operators.

The process known as purification consists of the following: given a density matrix ρ^A for a mixed state of a system A, one can construct a bigger system AB, such that $\left|\psi^{AB}\right\rangle$ is a pure state, and

$$\rho^A = \mathrm{tr}_B \left|\psi^{AB}\right\rangle\!\left\langle\psi^{AB}\right| . \tag{2.5.9}$$

The proof is as follows. Let $\left\{\left|e_i^A\right\rangle\right\}$ be an orthonormal basis of ρ^A such that

$$\rho^A = \sum_i p_i \left|e_i^A\right\rangle\!\left\langle e_i^A\right| . \tag{2.5.10}$$

Then, we introduce a system B which has the same state space as system A, with orthonormal basis $\left\{\left|e_i^B\right\rangle\right\}$, and define the pure state for the combined system

$$|AB\rangle = \sum_i \sqrt{p_i} \left|e_i^A\right\rangle \left|e_i^B\right\rangle . \tag{2.5.11}$$

The reduced density matrix of the pure state $\rho^{AB} = |AB\rangle\langle AB|$ will give ρ^A.

Indeed:

$$\operatorname{tr}_B \rho^{AB} = \operatorname{tr}_B \sum_{ij} \sqrt{p_i}\sqrt{p_j}\left(\left|e_i^A\right\rangle\!\left\langle e_j^A\right|\right)\left(\left|e_i^B\right\rangle\!\left\langle e_j^B\right|\right) \tag{2.5.12}$$

$$= \sum_{ij} \sqrt{p_i}\sqrt{p_j}\left|e_i^A\right\rangle\!\left\langle e_j^A\right|\left\langle e_i^B \middle| e_j^B\right\rangle \tag{2.5.13}$$

$$= \sum_{i} p_i \left|e_i^A\right\rangle\!\left\langle e_i^A\right| = \rho^A. \tag{2.5.14}$$

If there exist two purifications $|AB_1\rangle$ and $|AB_2\rangle$ for the system A, then B_1 and B_2 are related by a unitary transformation U_2:

$$|AB_1\rangle = I \otimes U_2 \,|AB_2\rangle\,. \tag{2.5.15}$$

Chapter 3

Quantum Entanglement

The following chapter introduces the basis of quantum entanglement, focusing on both its definition and characterization according to entanglement measures for the bipartite and multipartite cases [35, 40, 41, 42, 43, 44, 45], focusing beforehand in the notion of distance between states and the quantum operations formalism.

3.1 Entanglement

Let's consider a system of n states. In the classical description, the total state space of the system is described by the product state of the n separate systems. On the other hand, in the quantum formalism the total Hilbert space $\mathcal{H}$ is a tensor product of the subsystem spaces $\mathcal{H} = \otimes_{i=1}^{n} \mathcal{H}_i$. Following the superposition principle, the total state of the system can be written as

$$|\psi\rangle = \sum_{i_1, i_2 \ldots i_n} c_{i_1, i_2 \ldots i_n} |i_1\rangle |i_2\rangle \cdots |i_n\rangle , \tag{3.1.1}$$

In general, such states cannot always be described as a product of states of individual subsystems $|\psi\rangle \neq |\psi_1\rangle |\psi_2\rangle \cdots |\psi_n\rangle$. These kind of pure states are known as entangled states.

Similarly, a mixed state of n systems is entangled if it cannot be written as a convex combination of product states [46]

$$\rho \neq \sum_i p_i \rho_i^1 \otimes \cdots \otimes \rho_i^n. \tag{3.1.2}$$

For bipartite systems, the *maximally entangled state* is given by:

$$|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |e_i\rangle |f_i\rangle \tag{3.1.3}$$

where $\{|e_i\rangle\}_{i=1}^d$ and $\{|f_i\rangle\}_{i=1}^d$ are orthonormal families in $\mathcal{H}_A$ and $\mathcal{H}_B$ and $d = \min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\}$. These states satisfy that both the reduced states ρ_A and ρ_B are proportional to the identity matrix i.e. the maximally mixed state π.

The family of Bell states $\left|\phi_d^+\right\rangle$ are special cases of bipartite maximally entangled states, which satisfy that

$$\left|\phi_d^+\right\rangle = U_A \otimes U_B |\Psi\rangle, \tag{3.1.4}$$

where $U_A \otimes U_B$ is a local unitary transformation. For $d = 2$, it is useful to define a unit of bipartite entanglement called *e-bit*, which corresponds to the amount of entanglement contained in a Bell state of two qubits $|\phi^+\rangle$.

Maximally entangled pure states are one of the key resources for quantum computational tasks and for quantum communication. However, the coupling with the environment produces decoherence, transforming these states into non-maximally entangled mixed states. The following sections will establish a basic ground to quantify the entanglement for arbitrary states.

3.2 Distance between Quantum States

Before reviewing the different ways to identify and measure the entanglement between two states, it is necessary to establish the notion of distance or closeness between quantum states, as it will become an important tool to quantify entanglement. The following subsection presents the basic notion of a metric and two common measures for quantum states.

A *metric* or *distance function* [47] on a set X is a function $\mathcal{D}$ that takes any pair

(p, q) of points of a set X into the real line

$$\mathcal{D} : X \times X \to \mathbb{R}_0^+ \tag{3.2.1}$$

and satisfies the following properties for all $x, y, z \in X$:

1. $\mathcal{D}(x, y) = 0 \Leftrightarrow x = y$.
2. Symmetry: $\mathcal{D}(x, y) = \mathcal{D}(y, x)$.
3. Triangle inequality: $\mathcal{D}(x, y) \leq \mathcal{D}(x, z) + \mathcal{D}(z, y)$.

3.2.1 Trace distance

For two quantum states ρ and σ, the *trace distance* is defined as

$$D_T(\rho, \sigma) \equiv \frac{1}{2} \operatorname{tr} |\rho - \sigma| \tag{3.2.2}$$

where $|A| \equiv \sqrt{A^\dagger A}$ is the positive square root of $A^\dagger A$. The trace distance for any two density matrices satisfy the following bounds:

$$0 \leq D_T(\rho, \sigma) \leq 1. \tag{3.2.3}$$

For pure states, that is, if $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\phi\rangle\langle\phi|$, the trace distance simplifies to:

$$D_T(\rho, \sigma) = \sqrt{1 - |\langle\psi|\phi\rangle|^2} \tag{3.2.4}$$

The trace distance is invariant under unitary transformations

$$D_T(U\rho U^\dagger, U\sigma U^\dagger) = D_T(\rho, \sigma). \tag{3.2.5}$$

If Λ is a quantum channel[1], the trace distance satisfies that

$$D_T(\Lambda(\rho), \Lambda(\sigma)) \leq D_T(\rho, \sigma). \tag{3.2.6}$$

This property states that a quantum channel makes two quantum states less distinguishable from each other.

[1]Quantum channels are presented in the section about Quantum and LOCC operations

The trace distance satisfy also the *strong convexity property*. That is, for two ensembles $\{p_i, \rho_i\}$ and $\{q_i, \sigma_i\}$, where p_i and q_i are probability distributions, the following inequality holds

$$D_T\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \leq \frac{1}{2}\sum_i |p_i - q_i| + \sum_i p_i D_T(\rho_i, \sigma_i). \tag{3.2.7}$$

3.2.2 Fidelity

The *fidelity* is not a distance measure, but it can be used to define a metric from it, as we will soon see. For two quantum states ρ and σ, the fidelity is defined as

$$F(\rho, \sigma) \equiv \left(\mathrm{tr}\left\{\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right\}\right)^2. \tag{3.2.8}$$

The fidelity for any two density matrices satisfy the following bounds:

$$0 \leq F(\rho, \sigma) \leq 1. \tag{3.2.9}$$

If $\rho = |\psi\rangle\langle\psi|$ is a pure state, the fidelity becomes then:

$$F(\rho, \sigma) = \langle\psi|\sigma|\psi\rangle\,. \tag{3.2.10}$$

More so, if $\sigma = |\phi\rangle\langle\phi|$ is also a pure state, it is easy to see that the fidelity is given by:

$$F(\rho, \sigma) = \langle\psi|\phi\rangle\,\langle\phi|\psi\rangle = |\langle\psi|\phi\rangle|^2. \tag{3.2.11}$$

The fidelity is invariant under unitary transformations

$$F(U\rho U^\dagger, U\sigma U^\dagger) = F(\rho, \sigma). \tag{3.2.12}$$

If Λ is a quantum channel, the fidelity satisfies that

$$F(\Lambda(\rho), \Lambda(\sigma)) \geq F(\rho, \sigma). \tag{3.2.13}$$

The fidelity behaves the opposite as the trace distance; as the quantum channel makes states less distinguishable, the fidelity increases.

The *root fidelity* is given by

$$\sqrt{F} = \mathrm{tr}\left\{\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right\} \tag{3.2.14}$$

and satisfies the *strong concavity property*. That is, for two ensembles $\{p_i, \rho_i\}$ and $\{q_i, \sigma_i\}$, where p_i and q_i are probability distributions, the following inequality holds:

$$\sqrt{F}\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \geq \sum_i \sqrt{p_i q_i}\sqrt{F}\left(\rho_i, \sigma_i\right). \tag{3.2.15}$$

The following bound applies to the trace distance and the fidelity between two quantum states ρ, σ:

$$1 - \sqrt{F(\rho,\sigma)} \leq D_T(\rho,\sigma) \leq \sqrt{1 - F(\rho,\sigma)} \tag{3.2.16}$$

Although the fidelity satisfies the symmetry property, it is not a distance measure as it violates the triangle inequality and $F(\rho,\rho) = 1$, whereas a distance measure should be equal to zero when two states are equal.

Nonetheless, it is possible to build a metric from the fidelity. Indeed, the following three functions

$$D_{\mathrm{BA}}(\rho,\sigma) := \arccos\sqrt{F(\rho,\sigma)} \tag{3.2.17}$$

$$D_{\mathrm{BM}}(\rho,\sigma) := \sqrt{2 - 2\sqrt{F(\rho,\sigma)}} \tag{3.2.18}$$

$$D_{\sin}(\rho,\sigma) := \sqrt{1 - F(\rho,\sigma)} \tag{3.2.19}$$

correspond to metric on the space of density matrices, and are known in the literature as the *Bures angle*, the *Bures metric*, and the *Sine metric*, respectively [48].

3.3 Quantum and LOCC operations

The quantum operations formalism corresponds to a general tool for describing the evolution of quantum systems in a wide variety of circumstances, for example a unitary time evolution, the post-measurement effect or the state interactions with an environment.

The most general quantum operation Λ is a map from the set of density operators of the input space $\mathcal{H}$ to the set of density operators for the output space $\mathcal{H}'$,

$$\rho \to \rho' = \frac{\Lambda(\rho)}{\mathrm{tr}[\Lambda(\rho)]}, \quad \rho \in \mathcal{E}(\mathcal{H}), \rho' \in \mathcal{E}(\mathcal{H}') \tag{3.3.1}$$

which satisfies the following axioms:

1. For the initial state ρ, the probability that the process represented by Λ occurs is given by $\mathrm{tr}[\Lambda(\rho)]$. Thus, $0 \leq \mathrm{tr}[\Lambda(\rho)] \leq 1$ for any state ρ.

2. Λ is a linear map on the set of density matrices, that is, for probabilities $\{p_i\}$,
$$\Lambda\left(\sum_i p_i \rho_i\right) = \sum_i p_i \Lambda(\rho_i) \tag{3.3.2}$$

3. Λ is a *completely positive* (CP) map. That is, if Λ maps density operators of system $\mathcal{H}$ to density operators of system $\mathcal{H}'$, then $\Lambda(A)$ must be positive for any positive operator A. If an extra system R of arbitrary dimensionality is introduced, it must be true that $(I \otimes \Lambda)(A)$ is positive for any positive operator A on the combined system.

Any quantum operation can be expressed in the form:

$$\Lambda(\rho) = \sum_i V_i(\rho)V_i^\dagger, \quad \sum_i V_i^\dagger V_i \leqslant I \tag{3.3.3}$$

for some set of operators $\{V_i\}$ called *Kraus operators*.

The probability $\mathrm{tr}[\Lambda(\rho)]$ is equal to 1 if and only if the CP map Λ is trace-preserving, which corresponds to $\sum_i V_i^\dagger V_i = I$. In this case, Λ is called a *quantum channel*.

3.3.1 LOCC (Local Operations and classical communication) operations

An important type of quantum operations in quantum information is the *LOCC (Local Operations and classical communication)* operations or LOCC protocols, where a product of local operations are made in a subspace of the system, and the results are communicated classically to another party which usually realize local operations conditioned to the received information.

Let's consider an entangled state ρ shared by two observers Alice and Bob, which can realize any quantum operations in their respective subsystems A and B:

$$\Lambda_A : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}'_A) \tag{3.3.4}$$

$$\Lambda_B : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}'_B) \tag{3.3.5}$$

The final spaces $\mathcal{H}'_A$ and $\mathcal{H}'_B$ can include local ancillas i.e. be in a product state with the environment, or can correspond to some subspace of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Both Alice and Bob can communicate their measurement results through classic communication, which can increase the classical correlations between the subsystems A and B, but not the entanglement between both. This process can be repeated several times, that is, a LOCC protocol can proceed in several rounds.

There exist several classes of LOCC operations:

1. Class of local operations (C_1): In this case no communication between Alice and Bob is allowed. The mathematical structure of the map is elementary $\Lambda^{\varnothing}_{AB} = \Lambda_A \otimes \Lambda_B$ with Λ_A, Λ_B being both quantum channels.

2. Class of *one-way* LOCC operations (C_2): These operations allow classical communication from Alice to Bob (or similarly from Bob to Alice). In this case, Alice makes measurements over the subspace A and Bob realizes local operations depending on the i result of Alice. The map has the form:

$$\Lambda^{\rightarrow}_{AB}(\rho) = \sum_i I_A \otimes \Lambda^{(i)}_B \Big[\Big(A_i \otimes I_B \Big) \rho \Big(A_i^\dagger \otimes I_B \Big) \Big] \tag{3.3.6}$$

$$= \sum_i A_i \otimes I_B \Big[\Big(I_A \otimes \Lambda^{(i)}_B \Big) (\rho) \Big] A_i^\dagger \otimes I_B. \tag{3.3.7}$$

3. Class of *two-way* LOCC operations (C_3): In this case, both observers are allowed to send classical information to each other, and are composed of local operations and the following maps:

$$\Lambda^1_{AB}(\rho) = \sum_i \left(A_i \otimes I_B \rho A_i^\dagger \otimes I_B\right) \otimes |e_i\rangle\langle e_i|_{B'}\,, \quad \sum_i A_i^\dagger A_i = I_A \tag{3.3.8}$$

$$\Lambda^2_{AB}(\rho) = \sum_i \left(I_A \otimes B_i \rho I_A \otimes B_i^\dagger\right) \otimes |f_i\rangle\langle f_i|_{A'}\,, \quad \sum_i B_i^\dagger B_i = I_B \tag{3.3.9}$$

where $\{|e_i\rangle\}$ (respectively $\{|f_i\rangle\}$) is an orthonormal base for the ancilla state of Bob (respectively the ancilla state of Alice).

4. Class of separable operations (C_4): Corresponds to the set of all the separable quantum operations, that is, operations with product Kraus operators $A_i \otimes B_i$, and are given by:

$$\Lambda^{\text{sep}}_{AB}(\rho) = \sum_i A_i \otimes B_i \rho A_i^\dagger \otimes B_i^\dagger, \quad \sum_i A_i^\dagger A_i \otimes B_i^\dagger B_i = I_A \otimes I_B \tag{3.3.10}$$

There is an order of inclusions $C_1 \subset C_2 \subset C_3 \subset C_4$, where all inclusions are strict.

In cases where the LOCC operation are not achievable deterministically, but rather only with some arbitrary probability, they are considered *stochastic local operations* and denoted as SLOCC. Among the class of separable operations, this implies that the Kraus operators satisfy that $\sum_i A_i^\dagger A_i \otimes B_i^\dagger B_i \leq I_A \otimes I_B$.

The set $\mathcal{S}_{\text{AB}}$ of bipartite separable states is invariant under separable operations, and every separable state can be converted into any other separable state by a separable operation. More so, any separable state $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i|$ can be obtained from the classical state $\rho_c = \sum_{jk} p_{jk} |e_j\rangle\langle e_j| \otimes |e_k\rangle\langle e_k|$. Indeed, let the Kraus operators be $A_{ijk} = \sqrt{\eta_i}\,|\psi_i\rangle\langle j|$ and $B_{ijk} = |\phi_i\rangle\langle k|$, then the separable operation $\Lambda^{\text{sep}}_{AB}(\rho)$ returns:

$$\Lambda_{AB}^{\text{sep}}(\rho_c) = \sum_i A_i \otimes B_i \rho_c A_i^\dagger \otimes B_i^\dagger \tag{3.3.11}$$

$$= \sum_{ijk} p_{jk}\eta_i (|\psi_i\rangle\langle e_j| \otimes |\phi_i\rangle\langle e_k|)(|e_j\rangle\langle e_j| \otimes |e_k\rangle\langle e_k|)(|e_j\rangle\langle \psi_i| \otimes |e_k\rangle\langle \phi_i|) \tag{3.3.12}$$

$$= \sum_{ijk} p_{jk}\eta_i \, |\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i| = \sum_i \eta_i \, |\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i| = \rho \tag{3.3.13}$$

Furthermore, an arbitrary state ρ can be transformed into a classical state ρ_c by a measurement in the product basis $\{|e_j\rangle|e_k\rangle\}$, which is a local operation.

3.3.2 Classes of equivalence

Depending on the class of operations considered, it is possible to group entangled states into different classes of equivalent entanglement. For the first class of local operations C_1, it is possible to establish the *local unitary (LU) equivalence*, where two pure states $|\psi\rangle$ and $|\phi\rangle$ are considered equivalently entangled if they differ only by a local unitary basis change:

$$|\psi\rangle \sim_{LU} |\phi\rangle \quad \Leftrightarrow \quad |\psi\rangle = (U_1 \otimes \cdots \otimes U_N)\,|\phi\rangle \tag{3.3.14}$$

for suitable $(d_i \times d_i)$ unitary matrices U_i, with $d_i = \dim \mathcal{H}_i$. Up to a global phase, the unitary matrices can be chosen with determinant set to unity and the set of states can be divided into orbits of the product group $G_U = SU(d)^{\otimes N}$, for N d-dimensional quantum states i.e. N qudits.

More generally, two states are said to be LOCC-*equivalent* if they can be converted into each other by a LOCC protocol. Similarly, two states are said to be LOCC_r-*equivalent* if they can be converted into each other using an LOCC protocol with no more than r rounds. Finally, they are $\overline{\text{LOCC}}$-*equivalent* if, starting from any of them, it is possible to approximate the other one to arbitrary precision, as the number of rounds r tends to infinity. For the bipartite case, Nielsen's theorem provides a simple criterion for the equivalence of bipartite pure states under LOCC:

Theorem 4. *Let $|\psi\rangle$ and $|\phi\rangle$ be two pure states of the bipartite system AB and*

$\boldsymbol{\lambda}_\psi = \left(\lambda_1^\psi, \ldots, \lambda_n^\psi\right)$ and $\boldsymbol{\lambda}_\phi = \left(\lambda_1^\phi, \ldots, \lambda_n^\psi\right)$ the vectors formed by the Schmidt coefficients of $|\psi\rangle$ and $|\phi\rangle$, respectively. Then, $|\psi\rangle$ can be transformed into $|\phi\rangle$ by a LOCC operation if and only if $\boldsymbol{\lambda}_\psi$ is majorized by $\boldsymbol{\lambda}_\phi$ ($\boldsymbol{\lambda}_\psi \prec \boldsymbol{\lambda}_\phi$), that is, if for all $1 \leq l < n$ we have that

$$\sum_{i=1}^{l} \lambda_i^\psi \leq \sum_{i=1}^{l} \lambda_i^\phi, \quad \sum_{i=1}^{n} \lambda_i^\psi = \sum_{i=1}^{n} \lambda_i^\phi \tag{3.3.15}$$

where n denotes the number of non-zero Schmidt coefficients.

A direct consequence of these theorem is that there are *incomparable* states, i.e. pairs of states such that neither can be converted into the other with certainty. Even if the state transformation $|\psi\rangle \to |\phi\rangle$ is not possible with a LOCC operations, the transformation $|\psi\rangle|\eta\rangle \to |\phi\rangle|\eta\rangle$ may be possible using a suitable ancilla state $|\eta\rangle$. This phenomenon is called *entanglement catalysis*, as the state $|\eta\rangle$ is returned unchanged after the transformation.

The above result doesn't work for multipartite pure states, as the Schmidt decomposition exists only for the bipartite case, therefore Nielsen's theorem cannot be easily generalized. In fact, no tractable mathematical description of LOCC-equivalence in the multipartite case has been identified so far. But it is possible to treat states as equivalent, if with some non-zero probability they can be transformed into each other by LOCC, that is, using a SLOCC protocol.

Let's consider a SLOCC transformation acting on a N-partite state with Kraus operators A_i:

$$|\psi\rangle \mapsto (A_1 \otimes \cdots \otimes A_N)\,|\psi\rangle. \tag{3.3.16}$$

Then, two states $|\psi\rangle$ and $|\phi\rangle$ are SLOCC-*equivalent* up to a normalization if and only if the matrices are invertible, that is

$$|\psi\rangle \sim_{\text{SLOCC}} |\phi\rangle = L_1 \otimes L_2 \otimes \cdots \otimes L_N |\psi\rangle, \tag{3.3.17}$$

where the matrices have unit determinant $\det L_i = 1$. Thus the group that governs the SLOCC equivalence for N d-dimensional quantum states is the special linear group composed with itself N times, $G_L = SL(d, \mathbb{C})^{\otimes N}$ [49, 50].

3.4 Axioms on entanglement measures

A measure of entanglement of a bipartite system AB is a function $E : \mathcal{E}(\mathcal{H}_{AB}) \to \mathbb{R}$ such that:

1. $E(\rho) = 0$ if and only if ρ is separable.
2. E is convex.
3. *Monotonicity*: E cannot increase under LOCC operations. If Λ is a LOCC operation, then:
$$E(\Lambda(\rho)) \leq E(\rho) \tag{3.4.1}$$

As any two separable states can be transformed into each others using LOCC operations, the monotonicity property implies that E is constant in the set of separable states S_{AB}, and can be set to zero without loss of generality.

As any state ρ can be transformed into a separable state via a LOCC operation, $E(\rho)$ is then minimum for separable states and $E(\rho) \geq 0$.

The second condition is justified as it captures the notion of the loss of information i.e. describing the process of going from a selection of identifiable states ρ_i that appear with probabilities p_i to a mixture of these states of the form $\rho = \sum p_i \rho_i$. Let's suppose that two parties (Alice and Bob) share m pairs of particles in the states $\rho_1 \dots \rho_m$. By classical communication, they can agree to keep the i-th pair with probability p_i, preparing thus the ensemble $\{\rho_i, p_i\}_{i=1}^m$.

By erasing the information about which state ρ_i was kept, the state becomes $\rho = \sum_i p_i \rho_i$, and the convexity axiom establishes that this local loss of information does not increase the average entanglement:

$$E(\rho) = E\left(\sum_i p_i \rho_i\right) \leq \sum_i p_i E(\rho_i) \tag{3.4.2}$$

Nonetheless, the convexity condition is not strictly necessary, it is a merely a convenient mathematical property, as many entanglement measures are convex by construction.

The monotonicity condition implies also that the entanglement measures are invariant under conjugations by local unitaries i.e. $E\Big(U_A \otimes U_B \rho U_A^\dagger \otimes U_B^\dagger\Big) =$

$E(\rho)$.

For pure states $|\psi\rangle$, this implies that $E(|\psi\rangle)$ only depends on the Schmidt coefficients λ_i of $|\psi\rangle$. Consequently, $E(|\psi\rangle) = f(\rho_A)$ is a unitary-invariant function of the reduced state $\rho^A = \mathrm{tr}_B(|\psi\rangle\langle\psi|)$ (or, equivalently, of $\rho^B = \mathrm{tr}_A(|\psi\rangle\langle\psi|)$). Given that a pure state is separable if and only if it has a single non-vanishing Schmidt coefficient, one deduces from the first axiom that $f\left(\rho^A\right)$ vanishes if and only if ρ^A is of rank one.

From the above, a stronger monotonicity condition can be proposed [51]. Let $f : \mathcal{E}(\mathcal{H}_A) \to \mathbb{R}$ be a concave[2], unitary invariant function, such that $f\left(\rho^A\right) = 0$ if and only if ρ^A is a pure state. Then,

$$E_f(|\psi\rangle) = f\left(\rho^A\right) \tag{3.4.3}$$

defines an entanglement measure on the set of pure states of pure state AB, which satisfies the stronger monotinicity condition:

4. $\sum_i p_i E(|\Phi_i\rangle) \leq E(|\psi\rangle)$, where $p_i = \|A_i \otimes B_i|\psi\rangle\|^2$ and $|\Phi_i\rangle = p_i^{-1/2} A_i \otimes B_i|\psi\rangle$ correspond the probabilities and conditional states of a separable measurement with Kraus operators $A_i \otimes B_i$.

That way, there exist a correspondence between all the possible entanglement measures for pure states characterized by E_f. The stronger condition is equivalent to saying that the entanglement does not increase on average under LOCC operations.

3.4.1 Convex Roof Construction

In general, it is easier to define entanglement measures for pure states, as the generalization of such measures for mixed states are not straight-forward. Even so, the *convex roof construction* allows to extend an entanglement measure for mixed sates. More generally, one is often in the position to know a quantity for pure states of a quantum system g, without a definite meaning in classical physics. The quantity G denotes then the *extension* of g to all mixed states [52].

[2] A concave function satisfies that

$$f(\rho) \geqslant \lambda f(\rho_1) + (1-\lambda) f(\rho_2),$$

for any $\lambda \in [0,1]$ and any pair of density matrices ρ_1, ρ_2 such that $\rho = \lambda\rho_1 + (1-\lambda)\rho_2$.

Mathematically speaking, let G be a real valued function on the compact convex set $\mathcal{E}(\mathcal{H})$ of density operators supported by $\mathcal{H}$. The density operator $\omega \in \mathcal{E}(\mathcal{H})$ is called a *roof point* of G, if there is at least one extremal convex decomposition

$$\omega = \sum p_j \pi_j, \quad \pi_j \in \mathcal{E}(\mathcal{H})^{\text{pure}} \tag{3.4.4}$$

such that

$$G(\omega) = \sum p_j G(\pi_j) \tag{3.4.5}$$

where $\mathcal{E}(\mathcal{H})^{\text{pure}}$ is the set of density operators of pure states. G is said to be a *roof* if every $\omega \in \mathcal{E}(\mathcal{H})$ is a roof point of G.

Let g be a real function acting on the set $\mathcal{E}(\mathcal{H})^{\text{pure}}$. A roof G is called a *roof extension* of g, if $G(\pi) = g(\pi)$ for pure states. The objective of a roof extension is to extend g "as linearly as possible" or, more correctly, "as affine as possible" to all mixed states.

If G is convex and coincides on $\mathcal{E}(\mathcal{H})^{\text{pure}}$ with g, then G is called a *convex extension* of g. Between any convex or roof extension of a function g the following inequality is valid

$$G^{\text{convex}} \leq G^{\text{roof}}. \tag{3.4.6}$$

Indeed, with an optimal decomposition (3.4.4) for G^{roof} one obtains that

$$G^{\text{convex}}(\omega) \leq \sum p_j G^{\text{convex}}(\pi_j) = \sum p_j G^{\text{roof}}(\pi_j) = G^{\text{roof}}(\omega) \tag{3.4.7}$$

This property is also valid pointwise; let G^{convex} be a convex and G any extension of g. If ω is a roof point of G, then

$$G^{\text{convex}}(\omega) \leq G(\omega). \tag{3.4.8}$$

The largest convex extension of g to $\mathcal{E}(\mathcal{H})$ is denoted as $g^{\cup}$,

$$g^{\cup} = \text{ largest convex extension of } g \text{ from } \mathcal{E}(\mathcal{H})^{\text{pure}} \text{ to } \mathcal{E}(\mathcal{H}) \tag{3.4.9}$$

Finally, let G be an extension of g and ω one of its roof points. If G is convex, then $G(\omega) = g^{\cup}(\omega)$. Indeed, because G is convex, $G \leq g^{\cup}$. Because ω is a roof

point, (3.4.8) asserts $G(\omega) \geq g^{\cup}(\omega)$. Therefore, $g^{\cup}(\omega)$ is unique and it is known as a *convex roof extension.*

How can one be sure that a convex extension for a given g exists at all? If there is one then there is also a largest one, i.e. $g^{\cup}$ exists. The answer to the question is affirmative and has been given in [53] by a variational characterization which is well known in quantum information theory as a recipe to construct entanglement measures:

$$g^{\cup}(\omega) = \inf \left\{ \sum p_j g\left(\pi_j\right) \right\} \tag{3.4.10}$$

where the infimum is running over all extremal convex decompositions of ω

$$\omega = \sum p_j \pi_j, \quad \pi_j \in \mathcal{E}(\mathcal{H})^{\text{pure}} . \tag{3.4.11}$$

3.5 Bipartite Entanglement

3.5.1 Entanglement Cost and Distillable Entanglement

For a given state ρ, the *entanglement cost* $E_C(\rho)$ quantifies the maximal possible rate r at which one can convert blocks of maximally entangled two qubits states i.e. e-bits into output states that approximate n copies of ρ, such that the approximations become vanishingly small in the limit of large block sizes $n \to \infty$. Let Λ be a general trace preserving LOCC operation and $\Phi_{2^{rn}}^{+} = \left(|\phi_+\rangle\langle\phi_+|\right)^{\otimes rn}$, then the entanglement cost may be defined as

$$E_C(\rho) = \inf \left\{ r : \lim_{n\to\infty} \left[\inf_{\Lambda} \operatorname{tr} \left| \rho^{\otimes n} - \Lambda\left(\Phi_{2^{rn}}^{+}\right) \right| \right] = 0 \right\} . \tag{3.5.1}$$

On the other hand, there exist a dual measure to the $E_C(\rho)$ known as the *distillable entanglement* $E_D(\rho)$, which quantifies the rate at which one can convert copies of ρ into e-bits with LOCC operations, and it is mathematically defined as

$$E_D(\rho) = \sup \left\{ r : \lim_{n\to\infty} \left[\inf_{\Lambda} \operatorname{tr} \left| \Lambda\left(\rho^{\otimes n}\right) - \Phi_{2^{rn}}^{+} \right| \right] = 0 \right\} . \tag{3.5.2}$$

For pure states, $E_C = E_D$, i.e. the entanglement transformations become reversible in the asymptotic limit for pure states. Both of these entanglement measures are monotone.

3.5.2 Entanglement of formation and Concurrence

The von Neumann entropy of a quantum state is defined as

$$S(\rho) = -\operatorname{tr}(\rho \ln \rho) = -\sum_i \mu_i \ln \mu_i \tag{3.5.3}$$

where μ_i are the eigenvalues of ρ, and satisfies the following properties:

1. Non-negative: $S(\rho) \geq 0$, where the equality only holds if the state is pure.
2. Invariant under local unitary operations: $S(U\rho U^\dagger) = S(\rho)$
3. Additive: $S(\rho \otimes \sigma) = S(\rho) + S(\sigma)$
4. Concavity [54]: Let ρ and σ be two density matrices, then

$$S(\lambda\rho + (1-\lambda)\sigma) \geq \lambda S(\rho) + (1-\lambda)S(\sigma), \quad 0 \leq \lambda \leq 1 \tag{3.5.4}$$

5. If ρ^A and ρ^B are the reduced density operators of the pure state $|\psi\rangle$ of the composite system AB, then

$$S\left(\rho^A\right) = S\left(\rho^B\right) \tag{3.5.5}$$

These properties allows to define an entanglement measure that uniquely describes the entanglement for bipartite pure states $|\psi\rangle$ around the von Neumann entropy and that also satisfies the stronger monotinity conditions previously defined, as follows

$$E_S(|\psi\rangle) = S\big(\rho^A\big) = S\big(\rho^B\big) = -\sum_i \lambda_i \ln \lambda_i \tag{3.5.6}$$

where λ_i are the Schmidt coefficients of $|\psi\rangle$. This entanglement measure is known as the *entropy of entanglement*. It can be shown that, for pure states, the entropy of entanglement is equivalent to both the entanglement cost and the distillable entanglement in the asymptotic case ($n \to \infty$): $E_S/\ln 2 = E_C = E_D$ [55]. This measure is not applicable for mixed states, since the von Neumann entropy of a subsystem can be non-zero even if the states are not entangled. Nonetheless, it is easy to extend the above measure to mixed states via the convex roof construction.

The *entanglement of formation* of a mixed state $\rho \in \mathcal{E}(\mathcal{H}_{AB})$ is given by

$$E_F(\rho) = \inf_{\{p_i, |\psi_i\rangle\}} \left\{ \sum_i p_i E_S(|\psi_i\rangle) \right\}, \tag{3.5.7}$$

where the infimum is taken over all ensembles $\{p_i, |\psi_i\rangle\}$ for which $\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i|$.

In practice, finding the best possible ensemble that minimizes equation (3.5.7) is a difficult tasks. However, it is possible to compute the entanglement of formation using *concurrence* for the two qubits case [12].

Let θ be the antiunitary transformation defined as

$$\theta |\psi\rangle = \sigma_y \otimes \sigma_y |\psi\rangle^* \tag{3.5.8}$$

where $|\psi\rangle^*$ is the complex conjugate of $|\psi\rangle$ in the computational basis. The concurrence for a two qubit pure space can be represented as:

$$C(|\psi\rangle) = \langle\psi|\theta|\psi\rangle\,. \tag{3.5.9}$$

Defining $\tilde{\rho} = \theta\rho\theta$ and the operator $\omega = \sqrt{\rho}\sqrt{\tilde{\rho}}$, the concurrence of a two qubit mixed state is defined as:

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \tag{3.5.10}$$

where $\lambda_1, \ldots, \lambda_4$ are the square roots of the eigenvalues of ω in decreasing order. Then, it can be showed that

$$E_F(\rho) = H\left(\frac{1 + \sqrt{1 - C^2(\rho)}}{2}\right) \tag{3.5.11}$$

where H is the function $H(x) = -x \ln x - (1 - x) \ln(1 - x)$.

3.5.3 Distance based measures of entanglement

Another convenient way to quantify entanglement is using some distance measure between the bipartite state ρ and the closest state in the set of separable states

$\mathcal{S}$, as a well defined metric satisfies by definition the first axiom in entanglement measures, needing only to satisfy convexity and monotonicity.

Using the trace distance as metric, it is possible to define the following measure of entanglement:

$$E_T(\rho) = \inf_{\sigma\in\mathcal{S}} D_T(\rho,\sigma) = \frac{1}{2}\inf_{\sigma\in\mathcal{S}} \operatorname{tr}|\rho-\sigma|. \tag{3.5.12}$$

Similarly, another useful measure of entanglement can be constructed using either the Bures metric or the Sine metric previously defined:

$$E_{\mathrm{BM}}(\rho) = \inf_{\sigma\in\mathcal{S}} D_{\mathrm{BM}}(\rho,\sigma)^2 = 2-2\sqrt{F_s(\rho)}, \tag{3.5.13}$$

$$E_{\sin^2}(\rho) = \inf_{\sigma\in\mathcal{S}} D_{\sin}(\rho,\sigma)^2 = 1-F_s(\rho), \tag{3.5.14}$$

where $F_s(\rho)$ corresponds to the maximal fidelity between ρ and a separable state

$$F_s(\rho) = \sup_{\sigma_\in\mathcal{S}} \{F(\rho,\sigma)\}. \tag{3.5.15}$$

If we consider a pure bipartite state $|\psi\rangle$ with Schmidt decomposition $|\psi\rangle = \sum_{i=1}\sqrt{\lambda_i}\,|e_i\rangle\,|f_i\rangle$, then it is possible to obtain $E_{\mathrm{BM}}(|\psi\rangle)$ and $E_{\sin^2}(|\psi\rangle)$ as

$$E_{\mathrm{Bu}}(|\psi\rangle) = 2\left(1-\sqrt{\lambda_{\max}}\right) \tag{3.5.16}$$

$$E_{\sin^2}(|\psi\rangle) = 1-\lambda_{\max}, \tag{3.5.17}$$

where $\lambda_{\max} = \max\{\lambda_i\}$ is the largest Schmidt coefficient of $|\psi\rangle$ [56], which can easily be obtained as the largest eigenvalue of the reduced density matrix.

Neither of these measures are a convex roof, but it is possible to establish a relationship between the Bures metric (and also the Sine metric) to another measure of entanglement called the *geometric measure of entanglement* [15], which can be extended to mixed states via the convex roof extension. The geometric measure of entanglement is defined for pure states $|\psi\rangle$ as:

$$E_G(|\psi\rangle) = 1-\sup_{|\phi\rangle\in\mathcal{S}} |\langle\phi|\psi\rangle|^2 = E_{\sin^2}(|\psi\rangle), \tag{3.5.18}$$

and it is convex roof extension for mixed states ρ corresponds to:

$$E_G(\rho) = \inf_{\{p_i, |\psi_i\rangle\}} \left\{ \sum_i p_i E_G\left(|\psi_i\rangle\right) \right\}, \tag{3.5.19}$$

where the infimum is taken over all ensembles $\{p_i, |\psi_i\rangle\}$ for which $\rho = \sum_i p_i \left|\psi_i\right\rangle \left\langle\psi_i\right|$.

The geometric measure of entanglement for a two-qubit state can also be related to the concurrence as follows:

$$E_G(\rho) = \frac{1}{2}\left(1 - \sqrt{1 - C(\rho)^2}\right). \tag{3.5.20}$$

For the next measure of entanglement, let's present first the *quantum relative entropy* between two states ρ and σ:

$$S(\rho\|\sigma) \equiv \operatorname{tr}[\rho(\ln\rho - \ln\sigma)] = -\operatorname{tr}\rho\ln\sigma - S(\rho). \tag{3.5.21}$$

This function satisfies the following important properties:

1. Positivity: $S(\rho\|\sigma) \geq 0$, where the equality only holds if $\rho = \sigma$.
2. Invariant under local unitary operations: $S(U\rho U^\dagger\|U\sigma U^\dagger) = S(\rho\|\sigma)$
3. Additivity for composite systems: $S\left(\rho^A \otimes \rho^B \| \sigma^A \otimes \sigma^B\right) = S\left(\rho^A\|\sigma^A\right) + S\left(\rho^B\|\sigma^B\right)$.
4. Joint convexity: Let ρ_0, ρ_1 and σ_0, σ_1 be two pairs of density matrices and $0 \leq \lambda \leq 1$, then

$$S\left((1-\lambda)\rho_0 + \lambda\rho_1 \| (1-\lambda)\sigma_0 + \lambda\sigma_1\right) \leq (1-\lambda)S\left(\rho_0\|\sigma_0\right) + \lambda S\left(\rho_1\|\sigma_1\right). \tag{3.5.22}$$

5. Monotonicity under trace-preserving CP maps: For any quantum operation $\Lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}\left(\mathcal{H}'\right)$ one has $S(\rho\|\sigma) \geq S(\Lambda(\rho)\|\Lambda(\sigma))$ for all states $\rho, \sigma \in \mathcal{E}(\mathcal{H})$.

It is important to note that the quantum relative entropy is not a true metric, as it is not symmetric and does not satisfy the triangle inequality, but rather a contrast

function between two states. Nonetheless it may also be used to characterize entanglement using the *relative entropy of entanglement*, defined as:

$$E_R(\rho) = \inf_{\sigma \in \mathcal{S}} S(\rho \| \sigma) \tag{3.5.23}$$

This measure has an appealing interpretation as distinguishability of ρ from the closest separable state. For pure states it coincides with the entropy of entanglement, and for mixed states ρ it is bounded from above by the entanglement of formation, that is,

$$E_R(\rho) \leq E_F(\rho). \tag{3.5.24}$$

A lower bound on the relative entropy of entanglement of a pure state $|\psi\rangle$ can be obtained using a related quantity to the geometric measure of entanglement [57]

$$E_R(|\psi\rangle) \geq -2 \ln \sup_{|\phi\rangle \in \mathcal{S}} |\langle \phi | \psi \rangle|. \tag{3.5.25}$$

3.6 Multipartite entanglement

Although many of the axiomatic measures can be extended to the multipartite case, the entanglement between more than two parties is far richer and more complex than the bipartite entanglement, as the mere way of defining separable states changes, and the notion of maximally entangled states ceases to be general.

3.6.1 Three qubits entanglement

Before generalizing for the n-qubit case, let's consider a system of pure three-qubit states. Any pure three-qubit state can be transformed by local unitary operations to the state:

$$|\psi\rangle = \lambda_0 |000\rangle + \lambda_1 e^{i\theta} |100\rangle + \lambda_2 |101\rangle + \lambda_3 |110\rangle + \lambda_4 |111\rangle, \tag{3.6.1}$$

where $\lambda_i \geq 0, \sum_i \lambda_i^2 = 1$ and $\theta \in [0, \pi]$. Considering the normalization condition, five real parameters are necessary to characterize the non-local properties of a pure state. This corresponds to a generalization of the Schmidt decomposition to three qubits states [58].

There are two different types of separability:

1. *Fully separable states*, which can be written as

$$\left|\psi^{\text{fs}}\right\rangle_{A|B|C} = |\alpha\rangle_A \otimes |\beta\rangle_B \otimes |\gamma\rangle_C. \tag{3.6.2}$$

2. *Biseparable states*, which can be written as a product state in the bipartite system, that is, if two of the three qubits are grouped together to one party. There are three possibilities of grouping two qubits together, hence there are three classes of biseparable states:

$$\left|\psi^{\text{bs}}\right\rangle_{A|BC} = |\alpha\rangle_A \otimes |\delta\rangle_{BC}, \tag{3.6.3}$$

$$\left|\psi^{\text{bs}}\right\rangle_{B|AC} = |\beta\rangle_B \otimes |\delta\rangle_{AC}, \tag{3.6.4}$$

$$\left|\psi^{\text{bs}}\right\rangle_{C|AB} = |\gamma\rangle_C \otimes |\delta\rangle_{AB}, \tag{3.6.5}$$

where $|\delta\rangle$ denotes a two-party state that might be entangled.

A pure state is called *genuine tripartite entangled* if it is neither fully separable nor biseparable. The genuine entangled three-qubit states can be further divided into two inequivalent classes, that is, states of each equivalence classes of genuine tripartite entangled states cannot be transformed into another by SLOCC protocols: the class of Greenberger-Horne-Zeilinger (GHZ) states $|\text{GHZ}\rangle$

$$|\text{GHZ}\rangle = \lambda_0|000\rangle + \lambda_4|111\rangle, \tag{3.6.6}$$

and the class of W states $|\text{W}\rangle$

$$|\text{W}\rangle = \lambda_1|100\rangle + \lambda_2|010\rangle + \lambda_3|001\rangle. \tag{3.6.7}$$

The GHZ state is a maximally entangled states and a generalization of the Bell states of two qubits, and have the appealing property that entanglement across any bipartite cut assume the largest possible value of 1 ebit. On the other hand, the entanglement of W states is said to be more *robust* against particle losses.

Indeed, considering the state $\rho^{\text{GHZ}} = |\text{GHZ}\rangle\langle\text{GHZ}|$, and taking the partial trace on the subsystem C, one gets:

$$\rho_{AB}^{\mathrm{GHZ}} = \mathrm{tr}_C\left(\rho^{\mathrm{GHZ}}\right) = |\lambda_0|^2 |00\rangle\langle 00| + |\lambda_4|^2 |11\rangle\langle 11| = \sum_{i=0}^{1} p_i |i\rangle\langle i| \otimes |i\rangle\langle i|, \quad (3.6.8)$$

where $p_0 = |\lambda_0|^2$ and $p_1 = |\lambda_4|^2$. The reduced state ρ_{AB}^{GHZ} corresponds to an unentangled mixed state.

On the other hand, considering the state $\rho^{\mathrm{W}} = |\mathrm{W}\rangle\langle\mathrm{W}|$, and taking the partial trace on the subsystem C, one gets:

$$\rho_{AB}^{\mathrm{W}} = \mathrm{tr}_C\left(\rho^{\mathrm{W}}\right) = |\lambda_3|^2 |00\rangle\langle 00| + \left|\phi_{12}^+\right\rangle\left\langle\phi_{12}^+\right|, \quad (3.6.9)$$

where $\left|\phi_{12}^+\right\rangle = \lambda_1 |10\rangle + \lambda_2 |01\rangle$ corresponds to the two qubits Bell state when $\lambda_1 = \lambda_2 = \frac{1}{\sqrt{3}}$. Therefore, the reduced state ρ_{AB}^{W} corresponds to a mixed entangled bipartite state.

The five SLOCC-inequivalent subsets of three-qubit pure states are composed of the class of fully separable states, the three inequivalent classes of bisebarable states, and the two inequivalent classes of genuine entangled states [59].

3.6.2 Entanglement classes for the general case

Similarly, the classification of entanglement for the general multipartite case follows a similar logic, distinguishing different types of entanglement for pure states.

Let $|\psi\rangle$ be a pure N-partite state. This state is *fully separable* if it is a product state of all parties, that is, if it can be written as

$$|\psi\rangle = \bigotimes_{i=1}^{N} |\phi_i\rangle, \quad (3.6.10)$$

where $|\phi_i\rangle$ belongs to the i-th partition.

A mixed state is called fully separable if it can be written as a convex combination of pure fully separable states, that is, if it can be written as

$$\rho = \sum_i p_i \rho_i^1 \otimes \cdots \otimes \rho_i^n. \quad (3.6.11)$$

If a pure state is not fully separable, it must contain some degree of entanglement, although it doesn't necessarily have to be true N-partite entanglement. Let $\alpha_m = (P_1, \ldots, P_m)$ denote a partition of $\{1, \ldots, N\}$ into m disjoint nonempty subsets, with $1 < m < N$. Such a partition corresponds to a division of the system into m distinct subsystems, also called a *m-partite split*. A pure state is called *m-separable* if it can be written as

$$|\psi\rangle = \bigotimes_{i=1}^{m} |\phi_i\rangle^{P_i} . \tag{3.6.12}$$

There are $m^N/m!$ possible partitions of the N parties into m parts. For mixed states, a quantum state ρ is *m-separable under a specific m-partite split* α_m if and only if it is fully separable in terms of the m subsystems in this split, that is, if it can be written as

$$\rho = \sum_i p_i \bigotimes_{n=1}^{m} \rho_i^{P_n}, \tag{3.6.13}$$

where ρ^{P_n} is a state of subsystem corresponding to P_n in the split α_m.

More generally, a state ρ is called *m-separable* [60] if it can be written as a convex combinations of pure m-separable states, which might belong to different partitions $\alpha_m^{(j)}$, such that

$$\rho = \sum_j p_j \bigotimes_{n=1}^{m} \rho^{P_n^{(j)}}, \tag{3.6.14}$$

where each state $\bigotimes_{n=1}^{m} \rho^{P_n^{(j)}}$ is a tensor product of m density matrices of the subsystems corresponding to some such partition $\alpha_m^{(j)}$, i.e., it factorizes under this split $\alpha_m^{(j)}$.

Finally, a state is called *truly N-partite entangled* when it is neither fully separable, nor m-separable, for any $m > 1$. Although there is a classification via equivalence classes under SLOCC for three qubits, it has been shown that already for four qubits there are infinitely many equivalence classes under SLOCC [61, 62]. Therefore, a classification of the pure truly N-partite entangled states is not straightforward.

3.6.3 Symmetric states

Because of the high complexity of multipartite entanglement, it is useful to study families of interesting multi-qubit states with an arbitrary number of qubits.

For example, a family of states arise if we restrict ourselves to the symmetric subspace $\mathcal{H}_{\text{sym}}^{\otimes N}$ of the full Hilbert space $\mathcal{H}^{\otimes N}$. For N qubits this subspace is in itself a Hilbert space of dimension $N+1$, and many states of interest such as the GHZ and W states, belong to it.

The symmetric subspace admits an orthonormal basis consisting of the *symmetric states*

$$|S(k,N)\rangle = \left(\begin{array}{c} N \\ k \end{array} \right)^{-\frac{1}{2}} \sum_j P_j \left\{ |1\rangle^{\otimes k} \otimes |0\rangle^{\otimes N-k} \right\}, \tag{3.6.15}$$

where $\sum_j P_j\{\ldots\}$ denotes the sum over all possible permutations of the qubits. The basis states $|S(k,N)\rangle$ can be identified with the *symmetric Dicke states*, the angular momentum eigenstates with maximal eigenvalue $\frac{N}{2}\left(\frac{N}{2}+1\right)$.

The N-qubit W state is an example of a symmetric state:

$$|\mathrm{W}_N\rangle = \frac{1}{\sqrt{N}}(|100\ldots0\rangle + |010\ldots0\rangle + \ldots + |00\ldots01\rangle) = |S(1,N)\rangle\,. \tag{3.6.16}$$

Similarly, the N-qubit GHZ state, also known as *Schrödinger cat state*, corresponds to a superposition of symmetric states

$$|\mathrm{GHZ}_N\rangle = \frac{1}{\sqrt{2}}(|S(0,N)\rangle + |S(N,N)\rangle) = \frac{1}{\sqrt{2}}\left(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}\right). \tag{3.6.17}$$

3.6.4 Multipartite Entanglement Measures

For the multipartite case, it is not possible to define an analogue for the entanglement cost and distillable entanglement. Although the GHZ states are considered maximally entangled states, not every state can be obtained from the GHZ state using LOCC alone, as is the case for W states. One may ask if in the asymptotic setting of arbitrarily many identically prepared states, the same can be said. To answer this question, is important to give a precise definition to the

concept of *asymptotic reversibility* [63].

One says that $|\psi\rangle^{\otimes x}$ is asymptotically reducible to $|\phi\rangle^{\otimes y}$ under LOCC, if for all $\delta, \varepsilon > 0$ there exist natural n, m such that

$$\left|\frac{n}{m} - \frac{x}{y}\right| < \delta, \quad \operatorname{tr}\left|\Lambda\left(|\psi\rangle\langle\psi|^{\otimes n}\right) - |\phi\rangle\langle\phi|^{\otimes m}\right| \geq 1 - \varepsilon, \tag{3.6.18}$$

where Λ is a LOCC operation. If both $|\psi\rangle^{\otimes x}$ can be transformed into $|\phi\rangle^{\otimes y}$ as well as $|\phi\rangle^{\otimes y}$ into $|\psi\rangle^{\otimes x}$, the transformation is asymptotically reversible. In the bipartite case, it is always true that any $|\psi\rangle$ can asymptotically be transformed into $|\phi^+\rangle^{\otimes E_S(|\psi\rangle)}$.

In the multi-partite setting, there is no single state to which any other state can be asymptotically reversibly transformed. The same can be said about the entanglement of formation, as the von Neumann entropy is a reliable measure only of bipartite entanglement.

Let X denote a nontrivial subset of the parties of an N-partite pure state $|\psi\rangle$, and let $\bar{X}$ be the set of remaining parties. The reduced density matrix of subset X of the parties is defined as

$$\rho_x = \operatorname{tr}_{\bar{X}}(|\psi\rangle\langle\psi|). \tag{3.6.19}$$

The *partial entropy* of subset X is the von Neumann entropy

$$S_X(|\psi\rangle) = -\operatorname{tr}\left(\rho_x \ln \rho_x\right). \tag{3.6.20}$$

The partial entropies of a multipartite state of more than two components can be unequal. Because partial entropies are conserved by asymptotically reversible LOCC operations, and can therefore no longer be viewed as absolute entanglement measures beyond the bipartite case, in which there is only one way of partitioning the composite system.

3.6.4.1 Schmidt measure

In general, any pure state vector can be written in the form

$$|\psi\rangle = \sum_{i=1}^{R} c_i \, |\psi_i^{(1)}\rangle \otimes \ldots \otimes \, |\psi_i^{(N)}\rangle \, , \tag{3.6.21}$$

where, unlike the Schmidt decomposition, the vectors $\left\{ |\psi_i^{(j)}\rangle \right\}_{i=1}^{R}$ are not required to be orthogonal for each subsystem. The *tensor rank* $R_{\min}(|\psi\rangle)$ correspond to the minimal number of product terms needed to express $|\psi\rangle$. For the bipartite case, it is equivalent to the matrix rank.

The *Schmidt measure* is the logarithm of the minimal number of terms in a product decomposition

$$E_S(|\psi\rangle) = \log R_{\min}(|\psi\rangle) \tag{3.6.22}$$

and corresponds to an entanglement monotone [64]. In the bipartite case, this measure reduces to the *Schmidt rank*, i.e., the rank of either reduced density matrix. The measure is zero if and only if the state is a full product. Therefore it cannot distinguish true multipartite entanglement from bipartite entanglement.

3.6.4.2 Multipartie relative entropy of entanglement

The relative entropy of entanglement can be generalized to the multipartite case following the same equation (3.5.23) as in the bipartite case. However, it is possible to define the multipartite relative entropy of entanglement over the set $\mathcal{S}_m^N$ of m-separable states of an N-partite system as

$$E_R^m(\rho) = \inf_{\sigma \in \mathcal{S}_m^N} S(\rho \| \sigma). \tag{3.6.23}$$

For the single copy setting, then it is clear that the set set $\mathcal{S}_m^N$ does not increase under LOCC, making $E_R^m(\rho)$ a suitable entanglement monotone.

3.6.4.3 Geometric measure of entanglement (GME)

The geometric measure of entanglement was originally generalized to the multipartite setting via projection operators of various ranks in [16]. Either

way, distance of the pure state $|\psi\rangle$ can be minimized either using the set of fully separable states or the set $\mathcal{S}_m^N$, such that it becomes

$$E_G^m(|\psi\rangle) = 1 - \sup_{\phi \in \mathcal{S}_m^N} |\langle\phi|\psi\rangle|^2. \tag{3.6.24}$$

Chapter 4

Variational Algorithm for Geometric Entanglement Measure

The previous chapter presented a global view of the properties of entanglement, along with the different ways to characterize and measure entanglement for bipartite and multipartite cases, using the axiomatic approach for entanglement measures as basis. The following chapter presents an algorithm to measure the geometric measure of entanglement (GME) using the fidelity as measure of closeness as presented in [65], using a similar approach as the self-guided tomographics method presented in [66, 67]. The main results of this work are presented here, with numerical simulations for random states and higher dimension matrix product states, and experimental results using IBM Quantum Falcon Processors [30].

4.1 Complex Simultaneous Perturbation Stochastic Approximation (CSPSA)

The complex simultaneous perturbation stochastic approximation (CSPSA) is an algorithmic method for stochastic optimization of multivariate systems, introduced in [67], being of the same family as the commonly used simultaneous perturbation stochastic approximation (SPSA) algorithm [68, 69], but with the advantage that it can be applied to real-valued target functions of complex variables. CSPSA is a gradient-descent approximation method, which requires only two measurements

of the target function regardless of the dimension of the optimization problem. This feature allows for a significant decrease in the cost of optimization, especially in problems with a large number of variables to be optimized, where the gradient of the target function is not directly available.

Let f be a real function on complex variables, $f : \mathbf{z} \in \mathbb{C}^p \to \mathbb{R}$. It is clear that any nontrivial function f cannot accomplish the Cauchy-Riemmann conditions, and has a formal dependence upon $\mathbf{z}^* \in \mathbb{C}^p$, where $\mathbf{z}^*$ denotes the complex conjugate of $\mathbf{z}$ [70]. The dependence of the non-holomorphic function f on $\mathbf{z}$ can be studied using the Wirtinger calculus by considering $\mu = (\mathbf{z}, \mathbf{z}^*)^T \in \mathbb{C}^{2p}$ rather than only $\mathbf{z}$. In practice however, for the CSPSA method it is enough to consider as variable $\mathbf{z}$, establishing the following iterative rule

$$\mathbf{z}_{k+1} = \mathbf{z}_k - a_k \mathbf{g}_k(\mathbf{z}_k), \tag{4.1.1}$$

where $a_k = a/(k+A)^s$ and the gradient estimator at iteration k is vector given by

$$\mathbf{g}_k(\mathbf{z}) = \frac{f(\mathbf{z} + b_k \boldsymbol{\Delta}_k) - f(\mathbf{z} - b_k \boldsymbol{\Delta}_k)}{2b_k} \begin{pmatrix} 1/\Delta^*_{k,1} \\ \vdots \\ 1/\Delta^*_{k,p} \end{pmatrix}, \tag{4.1.2}$$

with $b_k = b/k^t$, and the k-th perturbation vector $\boldsymbol{\Delta}_k$ is made up from p random elements taken from the set $\{\pm 1, \pm i\}$, to span any possible direction in the complex space. The values of a, A, s, b and t are known as the gain parameters, and they are adjusted to optimize the rate of convergence and depend on the target function. Two common gains set of gain parameters are the standard gains with $a = 3$, $b = 0.1$, $A = 0$, $s = 0.602$, and $t = 0.101$, and the asymptotic gains with $a = 3$, $b = 0.1$, $A = 0$, $s = 1$, and $t = 0.166$.

4.2 Variational quantum algorithms

To tackle the limitations of classical computing, quantum computing was suggested as a potential alternative to simulate quantum systems, under the assumption that a more natural simulation could be achieved using quantum resources like entanglement or quantum superpositions [71]. Several algorithms has been proposed, like Shor's algorithm [4] for prime numbers factorization or Grover's

algorithm [5] for the unstructured search problem, showing substantial speedups to their classical counterparts.

However, fault-tolerant quantum computers are not currently available. The state-of-the-art quantum computers are known as *Noisy Intermidiate-Scale Quantum* (NISQ) [20], which are characterized by noisy entangling gates, short coherence times, limited connectivity between qubits and large sampling error.

The leading strategy to obtain quantum advantage on NISQ devices are the variational quantum algorithms (VQA) [21, 22], which rely on classical optimization algorithms to train a parameterized quantum circuit. In general, VQAs have three basic elements: (1) a *cost function* $C(\boldsymbol{\theta})$ that codifies the solution of the problem, (2) a quantum circuit that prepares the state that best meets the problem's objective called *ansatz*, that depends on the set of parameters $\boldsymbol{\theta}$ and (3) a *classical optimization method* in charge of minimizing the cost function, that is, solving the following optimization task:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} C(\boldsymbol{\theta}). \tag{4.2.1}$$

The cost function maps the values of a set of trainable parameters to a real number. This cost defines a hyper-surface known as cost landscape such that the task of the optimizer is to navigate through the landscape and find the global minima. Without loss of generality, a cost function can be expressed as

$$C(\boldsymbol{\theta}) = \sum_k f_k \left(\mathrm{Tr} \left[O_k U(\boldsymbol{\theta}) \rho_k U^\dagger(\boldsymbol{\theta}) \right] \right), \tag{4.2.2}$$

where f_k is a set of function that codifies a given task, $U(\boldsymbol{\theta})$ is a parameterized unitary operator, $\boldsymbol{\theta}$ can either be continuous or discrete parameters, $\{\rho_k\}$ are the inputs states from a training set and $\{O_k\}$ is a set of observables.

For a VQA to truly have a real advantage over classical alternatives, the cost function must satisfy the following criteria. First, the minimum of $C(\boldsymbol{\theta})$ must correspond to the solution of the problem. Second, one must be able to efficiently estimate the cost using a quantum computer and possibly perform classical post-processing, under the assumption that such cost should not be easily computable using classical hardware. A third requirement is for the cost function to be

operationally meaningful, that is, lower values of the cost indicate a better solution quality. Finally, the cost function needs to be trainable, meaning that it is possible to optimize efficiently the parameters $\boldsymbol{\theta}$ in a suitable computation time lapse.

The ansatz is constructed using a parameterized quantum circuit as

$$|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})\,|\psi_0\rangle\,, \tag{4.2.3}$$

where $\boldsymbol{\theta}$ are the variational parameters and $|\psi_0\rangle$ is some initial state, which typically correspond to a product state with all qubits in the $|0\rangle$ state, i.e. $|00\cdots0\rangle = |0\rangle^{\otimes n}$, where n is the number of qubits.

The form of the ansatz dictates what kind of parameters $\boldsymbol{\theta}$ are used, and how they can be trained to minimize the cost. In general, there are two types of ansätze. The *problem-inspired ansätze* employ known information about the underlying physics of the problem, for example the *Suzuki-Trotter expansion* [72] or the *Unitary Coupled Cluster* (UCC) [73] ansatz, commonly used in quantum simulation and quantum chemistry respectively.

On the other hand, the *hardware-efficient ansätze* [26] are constructed taking in consideration a limited set of quantum gates and a particular qubit connection topology, such that it is ad-hoc to the available quantum device. The gate set usually consists of a two-qubit entangling gate and up to three single qubit gates. The ansatz is then constructed from *layers* of single-qubit gates and entangling gates, which are applied to multiple or all qubits in parallel.

The quantum circuit of a hardware-efficient ansatz with L layers is usually given by

$$U(\boldsymbol{\theta}) = \prod_{k=1}^{L} U_k\left(\boldsymbol{\theta}_k\right) W_k = \prod_{k=1}^{L} e^{-i\boldsymbol{\theta}_k H_k} W_k, \tag{4.2.4}$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_L)$ are the variational parameters, W_k is an unparametrized unitary and H_k is an hermitian operator.

The success of a VQA depends on the efficiency and reliability of the classical optimization used. An effective optimizer should try to minimize the number of measurements or the cost function evaluations. The optimizer must also be

resilient to noisy data coming from the high amounts of noise sources on NISQ computers and the limited number of shots in the measurement that introduce statistical errors. The optimizers can be grouped into two classes: the *gradient descent methods* and *gradient free methods*.

The gradient descent methods optimize the cost function via its gradient, i.e. the change of the function with respect to a variation of its parameters. The gradient indicates the direction in which the objective function shows the greatest change. This include for example methods that rely on the parameter shift rule [74, 75], or *stochastic gradient descent* (SGD) methods like SPSA and CSPSA.

The gradient free methods, as it names implies, don't rely on gradients measured on the quantum computer, and are typically methods already studied in machine learning, like *reinforced learning* (RL) [76] and the *sequential minimal optimization* (SMO) method [77].

The principal advantage of VQAs is that they grant a general framework that can be used to solve a wide arrange of problems. Some examples include *Variational quantum eigensolvers* (VQE) [23, 24, 25, 26], which find the ground and excited states of a given Hamiltonian, dynamical quantum simulation [78], solving tasks of quantum metrology [27, 28] and applications in quantum machine learning [29].

4.3 Variational determination of geometrical entanglement

Let us consider a n-qubit system described by the pure state

$$|\psi\rangle = \sum_{i_1,i_2\ldots i_n} c_{i_1,i_2\ldots i_n} |i_1\rangle |i_2\rangle \cdots |i_n\rangle , \tag{4.3.1}$$

The geometric measure of entanglement (GME) can be obtained solving the optimization problem

$$E_G(|\psi\rangle) = 1 - \sup_{|\phi\rangle\in\mathcal{S}} |\langle\phi|\psi\rangle|^2 = 1 - \sup_{|\phi\rangle\in\mathcal{S}} F(|\psi\rangle , |\phi\rangle) = 1 - \Lambda^2_{\max}, \tag{4.3.2}$$

where $\mathcal{S}$ is the set of fully separable states and $\Lambda^2_{\max}$ is known as the *entanglement eigenvalue*.

This optimization problem can be solved following a similar variational approach as the self-guided quantum tomography [66, 67]. The variational determination of geometrical entanglement (VDGE) requires two basic quantum steps:

1. A variational ansatz for separable states, which in this case is constructed with n single-qubit unitary transformations acting on the separable state $|0\rangle^{\otimes n}$, that is,
$$|\phi(\boldsymbol{\theta})\rangle = U_1(\boldsymbol{\theta}_1) \otimes \cdots \otimes U_n(\boldsymbol{\theta}_n) |0\rangle^{\otimes n}, \tag{4.3.3}$$
where the vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)$ contains the parameters defining the action of the unitary transformations.

2. Measuring the fidelity $F(\boldsymbol{\theta})$ as a function of $\boldsymbol{\theta}$.

Both steps can be naturally implemented in a quantum computer. Here, each unitary $U_i(\boldsymbol{\theta}_i)$ is generated as a sequence of local quantum gates acting on the i-th qubit and the fidelity is obtained applying the operator $U_1^\dagger(\boldsymbol{\theta}_1) \otimes \cdots \otimes U_n^\dagger(\boldsymbol{\theta}_n)$ onto the entangled state $|\psi\rangle$, followed by a projection onto the computational basis $\{|0\rangle, |1\rangle\}^{\otimes n}$, as is shown in Fig. 4.3.1. Then, the fidelity $F(\boldsymbol{\theta})$ is estimated as
$$F(\boldsymbol{\theta}) = |\langle\phi(\boldsymbol{\theta})|\psi\rangle|^2 \approx \frac{n_0}{N}, \tag{4.3.4}$$
where n_0 is the number of counts obtained when projecting onto the state $|0\rangle^{\otimes n}$ and N is the total number of copies of $|\psi\rangle$ employed in the projective measurements.

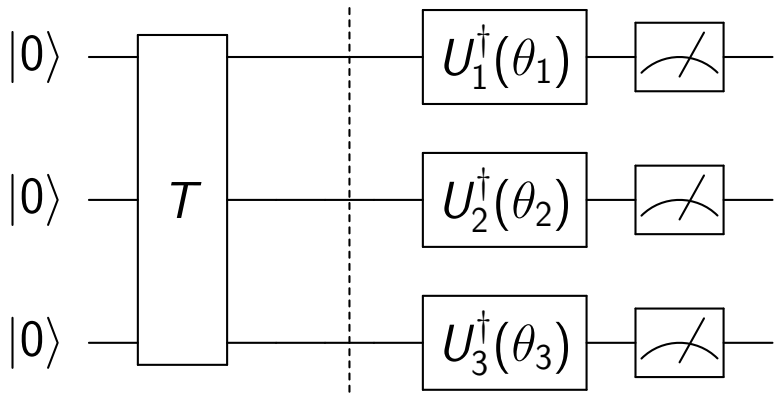

Figure 4.3.1: VDGE quantum circuit example for an arbitrary three-qubit state. The target state $|\psi\rangle$ is prepared applying a unitary gate T onto the separable state $|0\rangle^{\otimes 3}$. The fidelity $F(\boldsymbol{\theta}) = |\langle\phi(\boldsymbol{\theta})|\psi\rangle|^2$ is obtained applying local unitary gates $U_i^\dagger(\boldsymbol{\theta}_i)$ and measuring all qubits in the computational basis.

Thereby, the calculation of the GME reduces to running an optimization loop in which a classical optimizer is used to maximize $F(\boldsymbol{\theta})$ by varying the ansatz's parameters $\boldsymbol{\theta}$, or equivalently to minimize the GME as cost function of the variational quantum algorithm. Once convergence is reached, the optima $\hat{\boldsymbol{\theta}}$ is used to estimate the GME as $\hat{E}_G = 1 - \Lambda^2_{\max}(\hat{\boldsymbol{\theta}})$.

Since the method uses experimental estimates of the fidelity in each step, it is affected by several sources of noise. These include errors due to state preparation and measurement (SPAM), decoherence, and stochastic fluctuations produced by finite sampling. It has been proven that CSPSA is robust to noise and it only requires two evaluations of the cost function per iteration, making it a suitable classical optimization algorithm for this problem. Given that CSPSA works in the field of complex variables, $\boldsymbol{\theta}$ is set to be a vector of $2n$ complex numbers, where each pair of parameters defines a single qubit state. This is a linear scaling in the number of parameters $O(n)$, which makes the method well suited for large numbers of qubits. CSPSA provides a sequence of estimates $\boldsymbol{\theta}_k$ that converges to the minimizer $\hat{\boldsymbol{\theta}}$ of the fidelity. At a given iteration k, a new estimate $\boldsymbol{\theta}_{k+1}$ is generated from the fidelity values $F(\boldsymbol{\theta}_{k,\pm})$ at vectors $\boldsymbol{\theta}_{k,\pm}$, which are generated from the previous estimate $\boldsymbol{\theta}_k$ following the iterative rule (4.1.1). At the last iteration the fidelity $F(\hat{\boldsymbol{\theta}})$ is measured, which leads to the estimate $\hat{E}_G$ of the GME. The optimization loop is presented in Fig. 4.3.2.

A drawback present in the evaluation of the GME is its landscape, which may contain several local maxima. This can cause the optimization algorithm to be trapped in a local maximum. Also, the maximum may not be unique. For example, both states $|00\rangle$ and $|11\rangle$ maximize the entanglement eigenvalue for the Bell state $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, giving the same value for the GME. A small modification in the parameters of this state gives us a landscape with local maxima in which the optimization algorithm can get trapped. To overcome this problem, a multi-start strategy is employed, where the algorithm is repeated several times. Since CSPSA is a stochastic optimization method, it approaches $\hat{\boldsymbol{\theta}}$ from different paths in search space, which leads to a set of estimates $\{\hat{E}^j_G\}$. The highest value in $\{\hat{E}^j_G\}$ is selected as the final estimate $\hat{E}_G$.

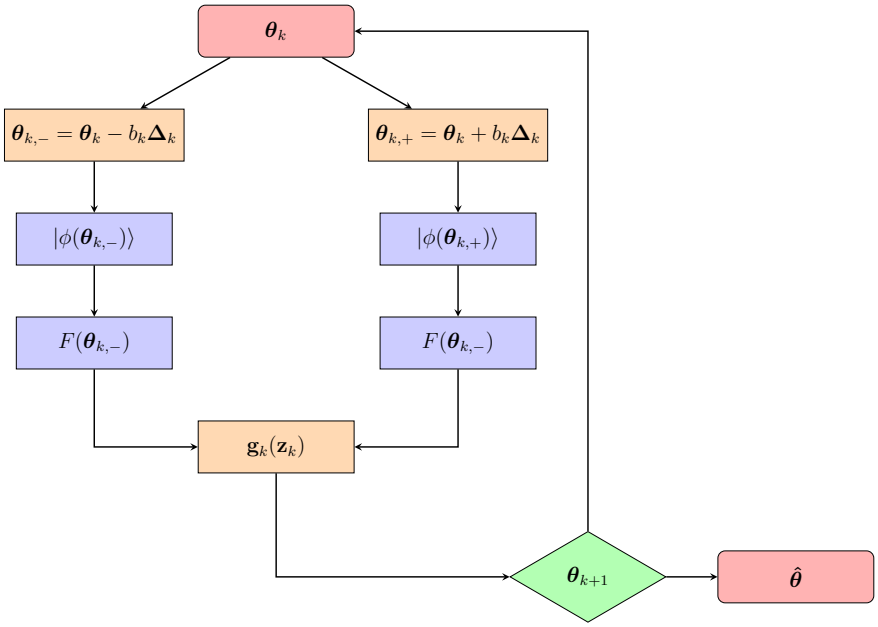

Figure 4.3.2: Flowchart of the VDGE algorithm. At iteration k, the parameters $\boldsymbol{\theta}_k$ are used to prepare the separable states $|\phi(\boldsymbol{\theta}_{k,\pm})\rangle$ and estimate $F(\boldsymbol{\theta}_{k,\pm})$. The gradient $\mathbf{g}_k(\mathbf{z}_k)$ is then constructed from the fidelities following (4.1.2), to generate a new estimate $\boldsymbol{\theta}_{k+1}$. The optimization loop is repeated until the optima $\hat{\boldsymbol{\theta}}$ is obtained. Blue squares represent the steps that are performed on the quantum computer, that is, the state preparation and the measurement of the fidelity, while the rest of the algorithm is performed on a classical computer.

4.4 Numerical simulations

First, the performance of the method is tested with states of $n = 3$ qubits. Let $|\mathrm{GW}(s,\varphi)\rangle$ be a superposition of GHZ and W states given by [65]

$$|\mathrm{GW}(s,\varphi)\rangle = \sqrt{s}\,|\mathrm{GHZ}\rangle + e^{i\varphi}\sqrt{1-s}\,|\mathrm{W}\rangle\,, \tag{4.4.1}$$

where $s \in [0,1]$, φ is a relative phase.

For each one of the values $\varphi = 0, \pi/4, \pi/2, \pi$, we generated 31 equally spaced values of s, which leads to a total of 124 $|\mathrm{GW}(s,\varphi)\rangle$ states. The GME of each one of these states is calculated 100 times with the VDGE algorithm, where each initial condition is chosen according to a Haar-uniform distribution. Each repetition of the algorithm consists of 150 iterations of CSPSA, where the values of the fidelity are simulated employing a sample of $N = 2^{13}$ shots, a number commonly available in open hardware like IBM Quantum.

In order to estimate errors, the *Bootstrap method* [79] is employed as follows. For each simulated state, a bootstrap sample is obtained by randomly sampling 5 times, with replacement, from the original set of 100 VDGE repetitions. This process is repeated 10000 times, to calculate the median and interquartile range for each simulated state. These are summarized in Fig. 4.4.1, where the value of the GME achieved by VDGE as a function of s for the values $\varphi = 0, \pi/4, \pi/2, \pi$ from bottom to top is shown. In this figure, points represent the median value of the GME and bars correspond to the interquartile range. Solid lines correspond to the theoretical solution of the optimization problem obtained using the Basin-hopping global optimization algorithm [80]. As is apparent from this figure, VDGE generates median values that are almost identical to the theoretical solutions and interquartile ranges that are very narrow, with mean errors of $0.00011, 0.00091, 0.00022$ and 0.00124 for the respective values of $\varphi = 0, \pi/4, \pi/2, \pi$. Thereby, these simulations indicate that VDGE generates accurate values of the GME for all the simulated states.

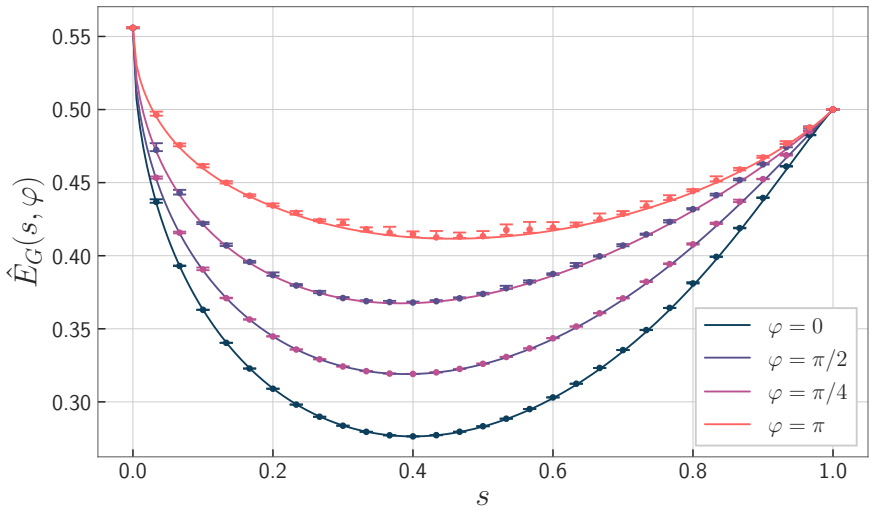

Figure 4.4.1: Geometric measure of entanglement $\hat{E}_G(s, \varphi)$ of the state $|\mathrm{GW}(s, \varphi)\rangle$ as a function of s for $\varphi = 0, \pi/4, \pi/2, \pi$, from bottom to top. Solid lines correspond to the theoretical solution of the optimization problem using a classical optimization algorithm i.e. Basin-hopping. Dots represent the median value of the GME obtained using VDGE with 5 repetitions and choosing the best result, while error bars correspond to the interquartile range.

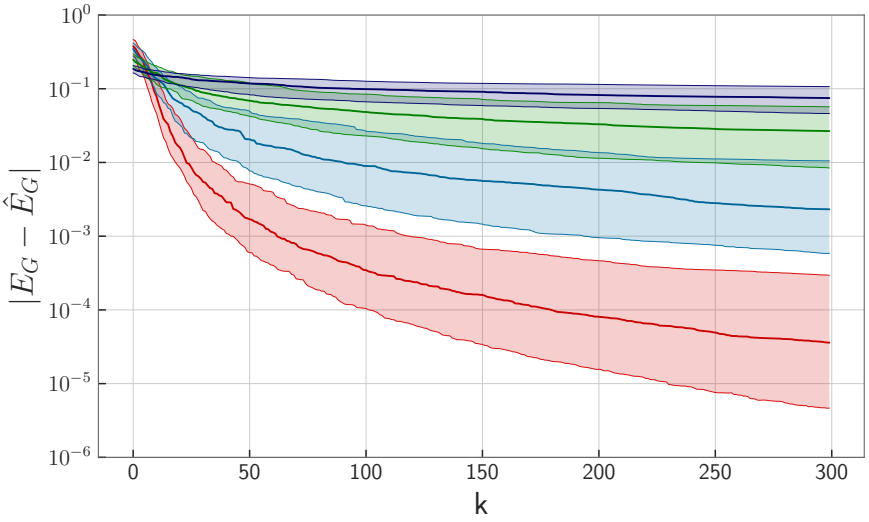

Figure 4.4.2: Difference between the GME E_G obtained with the Basin-hopping optimization method and the estimated GME $\hat{E}_G$ obtained with VDGE for 100 random pure states versus the number of iterations k. The curves show the results for 3, 4, 5, and 6 qubits from bottom to top, using in each case a sample size of 5 different initial states. Solid lines denote the median difference, while shaded areas represent the corresponding interquartile range

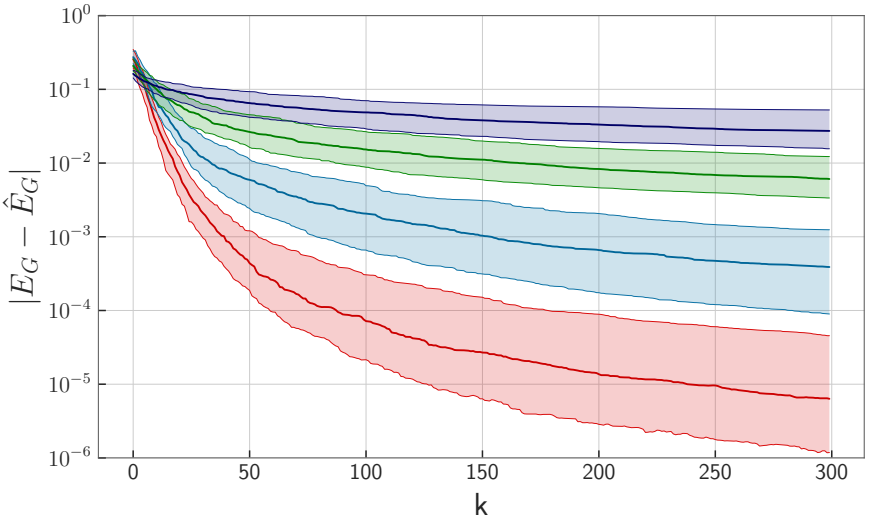

Figure 4.4.3: Difference between the GME E_G obtained with the Basin-hopping optimization method and the estimated GME $\hat{E}_G$ obtained with VDGE for 100 random pure states versus the number of iterations k. The curves show the results for 3, 4, 5, and 6 qubits from bottom to top, using in each case a sample size of 20 different initial states. Solid lines denote the median difference, while shaded areas represent the corresponding interquartile range

The method was tested with randomly generated states in various dimensions.

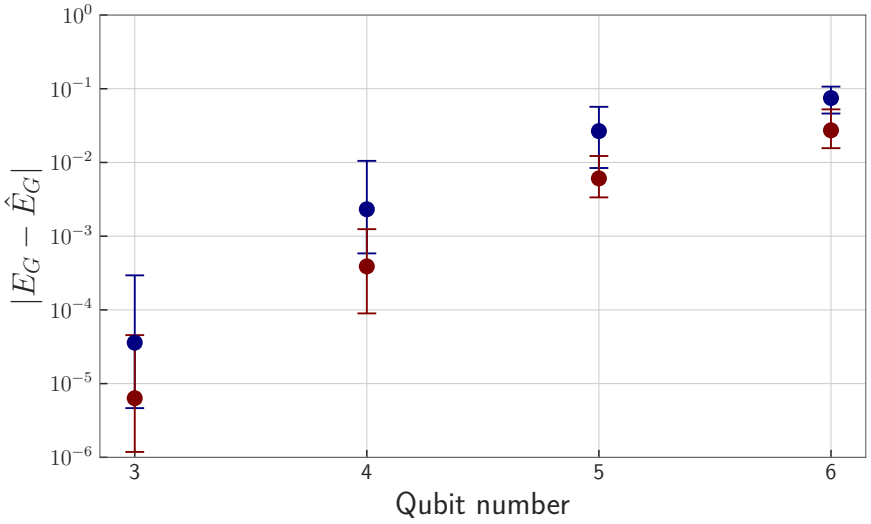

Figure 4.4.4: Difference between the GME E_G obtained with the Basin-hopping optimization and the estimated GME $\hat{E}_G$ obtained with VDGE versus the number of qubits, using a sample size of 5 initial states (blue) and 20 initial states (red). Dots correspond to the median difference of the last value obtained with VDGE and the theoretical value, while error bars represent the corresponding interquartile ranges.

Similar to the previous simulation, for 3, 4, 5, and 6 qubits a set of 100 random pure states is obtained according to a Haar-uniform distribution. For each state, the GME is selected as the maximum value obtained over 5 and 20 repetitions of the algorithm using $N = 2^{13}$ shots to simulate the measurement of the fidelity. These results are compared with the value obtained using Basin-hopping. In particular, the comparison is established using the median of $|E_G - \hat{E}_G|$ between the value E_G of the GME obtained via the Basin-hopping algorithm and the value $\hat{E}_G$ of the GME estimated via VDGE. The results of these simulations are presented in Figs. 4.4.2 and 4.4.3. These show the median of $|E_G - \hat{E}_G|$ (solid lines) as a function of the number of iterations of CSPSA, using 5 and 20 repetitions, respectively, for 3, 4, 5 and 6 qubits from bottom to top. Shaded areas represent the interquartile range. As is apparent from Figs. 4.4.2 and 4.4.3, the error in the estimation provided by VDGE decreases as the number of iterations of the optimization algorithm increases. Each curve exhibits a rapid decrease within the first tens iterations followed by approximately linear asymptotic behavior. This holds for all simulated qubit numbers. However, the quality of the estimation decreases as the number of qubits increases. This is depicted in Fig. 4.4.4, which

shows the error in estimating the GME as a function of the qubit number for 5 and 20 repetitions. Here, the result at iteration $k = 150$ of the optimization process is compared to the theoretical value of the GME. This behavior is to be expected because the dimension increases exponentially with the number of qubits and the measurement of the fidelity is simulated with a constant number $N = 2^{13}$ of shots. The main difference between Figs. 4.4.2 and 4.4.3 is that errors and dispersion are slightly less pronounced for 20 repetitions, but in both cases, the algorithm converges successfully when tested with random pure states.

The previous simulations are restricted to small qubit numbers. Tensor network algorithms [81] are employed in order to extend the simulations to higher qubit numbers. These have proven to be a useful tool for performing numerical simulations in many-body quantum systems [82, 83, 84, 85, 86, 87]. In particular, the simulations are carried out over matrix product states (MPS), that is, states that can be written as

$$|\psi\rangle = \sum_{i_1,\ldots,i_n} A^{i_1}_{[1]} A^{i_2}_{[2]} \ldots A^{i_{n-1}}_{[n-1]} A^{i_n}_{[n]} |i_1\rangle |i_2\rangle \cdots |i_n\rangle . \tag{4.4.2}$$

where $A_{[m]}$ are rank-3 tensors, with $m = 1, \ldots, n$. Algorithms have been proposed to compute efficiently the GME of MPS [88, 89]. The study is focused on states in the neighborhood of GHZ and W states, which have an efficient MPS representation. To generate the probe MPSs, tensors $A_{[m]}$ of GHZ and W states are perturbed with random matrices generated by a Gaussian distribution of null mean and variance λ, and then normalized. Figure 4.4.5 shows the median of $|E_G - \hat{E}_G|$ on a set of 10^3 perturbed GHZ and W states of $n = 25$ qubits, with $\lambda = 0.1$. VDGE is executed with 10^4 iterations of CSPSA simulating the fidelity with a sample of $N = 2^{13}$ shots. The optimums of GHZ and W states without perturbation are used as initial conditions, because the dimension of the space is too large, and therefore an unattainable number of shots is required to converge when a random initial condition is used. Figure 4.4.5 shows that with only 2×10^4 fidelity evaluations, the average error in estimating the GME can be reduced by about half of an order of magnitude. This number of evaluations is a thousand times smaller than the dimension of the system $d = 2^{25} \approx 3 \times 10^7$.

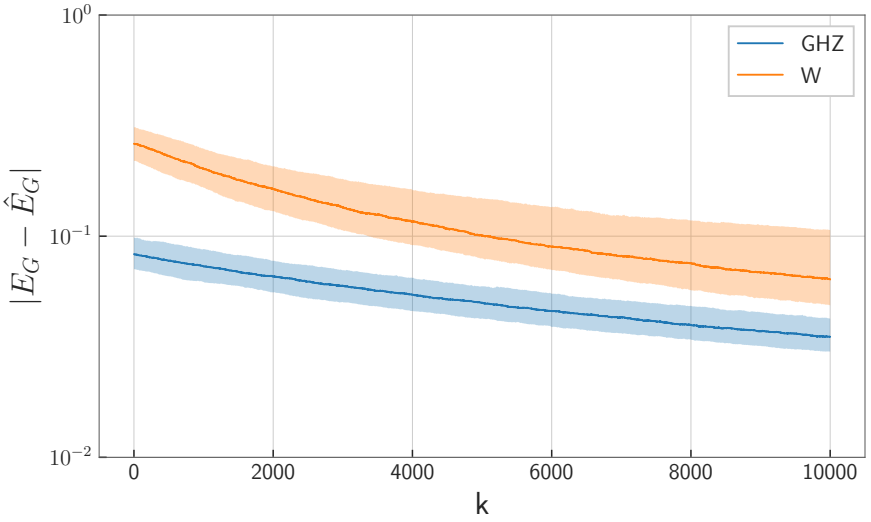

Figure 4.4.5: Difference between the GME E_G obtained with the Basin-hopping optimization method and the estimated GME $\hat{E}_G$ obtained with VDGE in a simulation based on matrix product states techniques for 10^3 perturbed W state (upper orange curve) and GHZ state (lower blue curve) for 25 qubits, as a function of the number of iterations k. Solid lines denote the median difference, while shaded areas represent the interquartile range.

4.5 Experimental results

The experimental demonstration of VDGE are performed using IBM Quantum systems `ibmq_lima` and `ibmq_bogota`. In particular, the geometric measure of entanglement of a n-qubit GHZ state with $n = 3, 4, 5$ is measured. VDGE is repeated 10 times for each state considering 150 iterations and $N = 2^{13}$ shots for measuring the fidelity at each iteration.

The results of the experiment are illustrated in Figs. 4.5.1, 4.5.2 and 4.5.3 for 3, 4, and 5 qubits, respectively. In each inset are depicted the values $\hat{E}_G(\boldsymbol{\theta}_{k,+})$ (blue dots) and $\hat{E}_G(\boldsymbol{\theta}_{k,-})$ (red dots) as functions of the number of iterations k. The value of the final estimate is indicated as a green square at the last iteration. The values of $\hat{E}_G(\boldsymbol{\theta}_{k,\pm})$ exhibit a rapid decrease within the first tens iterations followed by approximately linear asymptotic behavior[1]. Also, the values of $\hat{E}_G(\boldsymbol{\theta}_{k,\pm})$ become very close as the number of iterations increases, which shows the convergence of

[1]The terms $\hat{E}_G(\boldsymbol{\theta}_{k,\pm})$ correspond to the evaluations of cost function i.e. the fidelity with the perturbations $\boldsymbol{\theta}_{k,\pm} = \boldsymbol{\theta}_k \pm b_k \boldsymbol{\Delta}_k$, which follows from the definition of the gradient estimator. The estimated GME at each iteration $\hat{E}_G(\boldsymbol{\theta}_k)$ is not calculated, as it would require an additional experiment on the quantum computer, and it is not necessary for the optimization loop.

the algorithm toward the maximum. The final estimates $\hat{E}_G^\star$ are 0.5029, 0.5184, and 0.5640 for $n = 3, 4, 5$, respectively. These figures can first be compared with the value $E_G = 0.5$, which is the theoretical value of the GME for a GHZ state, independently of the number of qubits. VDGE provides a value of the GME which is close to the theoretical value with relative errors of 0.0058, 0.0368, and 0.128, respectively. The relative error increases as the number of qubits increases, which can be explained by the increase in the dimension of the search space. However, as the algorithm enters and stay into the linear regime it is possible to reduce the relative error by increasing the number of iterations. The GHZ state is generated by concatenating several Control-Not gates. This gate has an error larger than local gates and its impact on the generated state and the GME value can be significant. Therefore, the generated state are reconstructed via standard quantum tomography and forced purity and calculate its GME via Basin-hopping. This leads to the GME values 0.48214, 0.48996, and 0.4868 for the generated states for $n = 3, 4, 5$ qubits, respectively, which in Figs. 4.5.1, 4.5.2 and 4.5.3 are indicated with continuous black lines. With respect to these values, VDGE has the relative errors of 0.0431, 0.0581, and 0.1586, correspondingly.

Overall, VDGE exhibits convergence and provides in this experiment errors in the order of 10^{-2}. This is larger than what was observed in the numerical simulations in Figs. 4.4.2 and 4.4.3. This can be explained by the single-qubit gate error and readout assignment error, which are in the order of 10^{-4} and 10^{-2}, respectively. Since VDGE relies solely on single-qubit gates, its implementation in current NISQ computers is feasible.

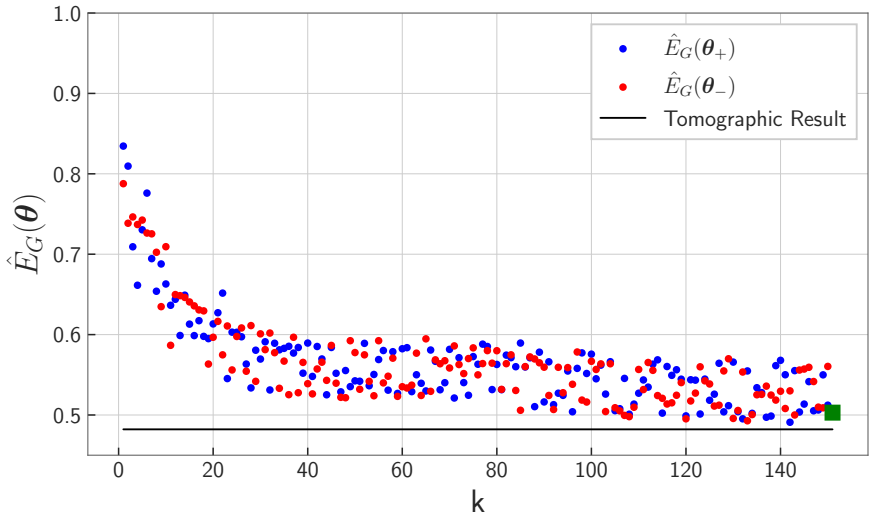

Figure 4.5.1: Values of $\hat{E}_G(\boldsymbol{\theta}_{k,+})$ (blue dots) and $\hat{E}_G(\boldsymbol{\theta}_{k,-})$ (red dots) as functions of the number of iterations k for a GHZ state of three qubits. The final estimate $\hat{E}_G^{\star}$ of the GME at $k = 150$ obtained via VDGE is indicated with a solid green square. Solid black lines indicate the value of the GME obtained by reconstructing the generated states with standard quantum tomography and then solving the optimization problem via Basin-hopping.

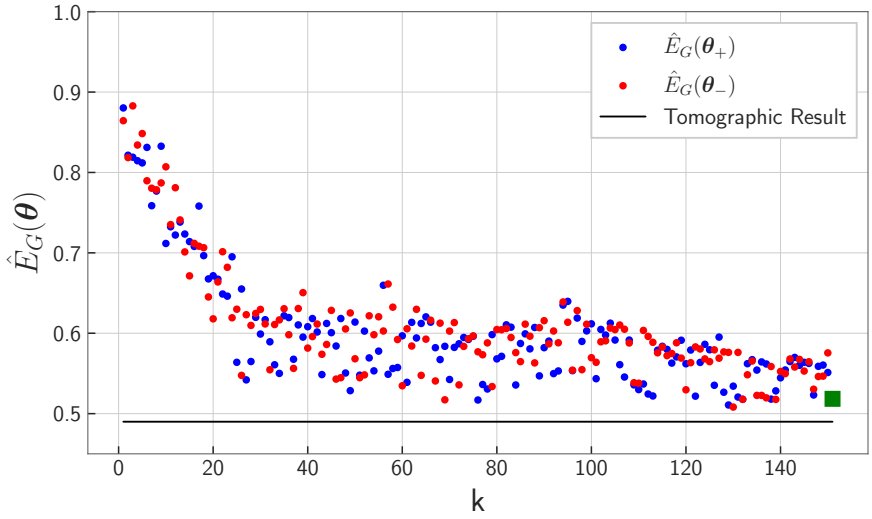

Figure 4.5.2: Values of $\hat{E}_G(\boldsymbol{\theta}_{k,+})$ (blue dots) and $\hat{E}_G(\boldsymbol{\theta}_{k,-})$ (red dots) as functions of the number of iterations k for a GHZ state of four qubits. The final estimate $\hat{E}_G^{\star}$ of the GME at $k = 150$ obtained via VDGE is indicated with a solid green square. Solid black lines indicate the value of the GME obtained by reconstructing the generated states with standard quantum tomography and then solving the optimization problem via Basin-hopping.

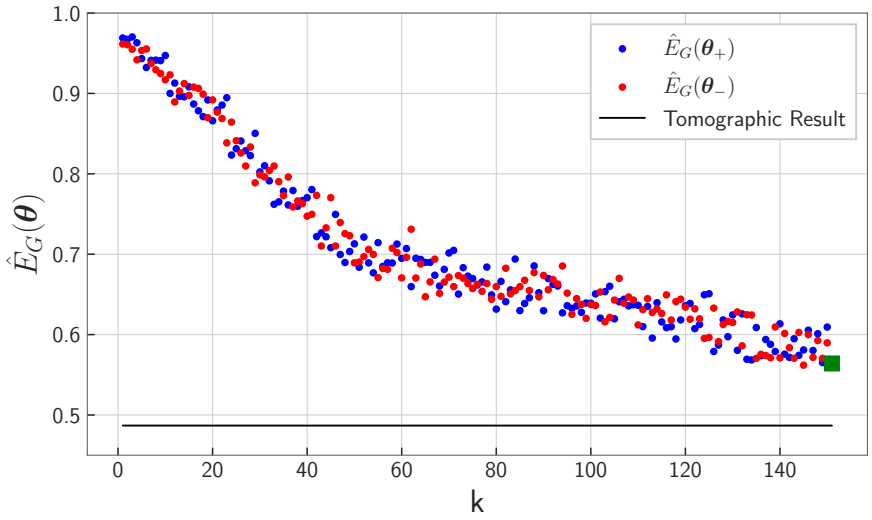

Figure 4.5.3: Values of $\hat{E}_G(\boldsymbol{\theta}_{k,+})$ (blue dots) and $\hat{E}(\boldsymbol{\theta}_{k,-})$ (red dots) as functions of the number of iterations k for a GHZ state of four qubits. The final estimate $\hat{E}_G^{\star}$ of the GME at $k = 150$ obtained via VDGE is indicated with a solid green square. Solid black lines indicate the value of the GME obtained by reconstructing the generated states with standard quantum tomography and then solving the optimization problem via Basin-hopping.

Chapter 5

Conclusion

Multipartite entanglement plays a key role in quantum information and quantum communication. However, the theoretical characterization of entanglement, as well as its experimental determination and certification, appear to be difficult problems. This is mainly due to the fact that most entanglement measures are not directly related to physically measurable quantities and are difficult to calculate, especially for higher-dimensional quantum systems. Here, the problem of measuring the entanglement of multipartite pure states generated in NISQ computers is studied. This class of devices is characterized by noisy quantum gates, which severely limit their usefulness. In this context, a major research subject is the development of algorithms that can work under such disadvantageous conditions.

This work revolves around the geometric measure of entanglement, which characterizes the entanglement of a n-qubit pure state $|\psi\rangle$ as the distance to the nearest separable pure state. This entanglement measure can be calculated by maximizing the fidelity of $|\psi\rangle$ in the set of separable states. The fidelity can be experimentally obtained in a NISQ computer by projecting $|\psi\rangle$ onto a set of n local bases. The complex simultaneous perturbation stochastic approximation algorithm is then used to solve the optimization problem in a classical computer. Thereby, a variational determination of the geometric entanglement measure is presented. In addition to its purity, the method does not require any a priori information about the state $|\psi\rangle$. The GME is well suited for being implemented in NISQ computers, because the measurements necessary to evaluate the GME are local and require a single unitary transformation acting on each qubit. Then,

VDGE has circuit depth 1, which helps to decrease error accumulation.

Numerical simulations indicate that the method reproduces known results on the geometric measure of entanglement of superpositions of 3-qubit GHZ and W states. Simulations with randomly chosen states of 3, 4, 5, and 6 qubits show that the algorithm can provide accurate values of the geometric entanglement, which is controlled by the number of iterations and the size of the ensemble used to measure the fidelity. The simulations are extended up to 25 qubits using matrix product state techniques, where convergence towards the optimum with a reasonable number of fidelity evaluations is also observed, despite the large size of the dimension.

A demonstration of the method is also demonstrated using IBM Quantum `ibmq_lima` and `ibmq_bogota` devices. In particular, the geometric measure of entanglement of a GHZ state for 3, 4, and 5 qubits is calculated, obtaining relative errors in the order of 10^{-2} for 3 and 4 qubits and in the order of 10^{-1} for 5 qubits. The demonstration uses 150 iterations and 10 repetitions. Following the simulations results, an increase in these figures leads to a reduction of relative error. Another important factor is the size of the ensemble used to measure the fidelity, which is currently limited to 2^{13} shots. A consequence of this is that the accuracy in the estimation of the fidelity decreases as the number of qubits increases, which in turn reduces the convergence rate of VDGE.

These results find direct application in the entanglement quantification of high-dimensional pure entangled states. Recently, two quantum-hardware platforms have been used to generate high-dimensional entanglement. The IBM Quantum `ibmq_montreal` device has reportedly generated [90] a 27-qubit GHZ state and detected its genuine multipartite entanglement. On the IBM Quantum `ibmq_rome` device, the generation of a three-qutrit GHZ state has been carried out for the first time [91] and its genuine multipartite entanglement has also been demonstrated. These quantum devices also allow the implementation of our method, which makes it possible to characterize the entanglement of arbitrary states of a large number of qubits and even qudits. Finally, we mention the recent demonstration of a universal qudit quantum processor using trapped ions [92]. Here, local qudits up to dimension 7 are created and accurately controlled. In particular, high accuracy local gates are implemented, which is the basic resource for the variational determination of the geometric measure of entanglement on any dimension.

Appendix A

Monte-Carlo Simulations

Monte Carlo [93] is a form of quadrature or numerical integration, in which finite summations are used to estimate definite integrals. Although Monte Carlo is inherently involved with the concept of probability, it can be applied to problems that have no apparent connection with probabilistic phenomena.

Significantly, it also can be applied to a great variety of problems for which the integral formulation is not posed explicitly. Often, the complex mathematics needed in many analytical applications can be avoided entirely by simulation. Thus, Monte Carlo methods provide extremely powerful ways to address realistic problems that are not amenable to solution by analytic techniques.

One of the fundamental basis for Monte Carlo analysis is the *law of large numbers*, which has strong and weak forms.

Let's consider a sequence of N random samples, denoted by $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$, such that all members of the random sequence are independent and identically distributed, that is, each random variable has the same probability distribution as the others and all are mutually independent. The expectation value of each random variable satisfies that

$$\mu = \mathbb{E}(X_1) = \mathbb{E}(X_2) = \ldots = \mathbb{E}(X_N), \tag{A.1}$$

while the sample mean is given by

$$\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i. \tag{A.2}$$

The *weak law of large numbers* states that, for all positive values of ϵ, no matter how small,

$$\lim_{N\to\infty} P\left(\left|\bar{X}_N - \mu\right| \geq \epsilon\right) = 0. \tag{A.3}$$

The weak law states that for any nonzero tolerance ϵ, with a sufficiently large sample there is a high probability that the sample average will be close to the expected value, i.e., within the tolerance margin. This type of convergence for $\bar{X}_N$ is technically called convergence in probability or weak convergence of a random variable.

The weak law is a statement about the limiting behavior of sums of random variables. It is possible, however, to make a stronger statement and say something about the behavior of the sum $\sum_{i=1}^{N} X_i$ on the way to the limit $N \to \infty$. The *strong law of large numbers* states that for any small $\epsilon > 0$ and $0 < \delta < 1$, a value of n exists such that, for any specified $m > 0$,

$$P\left(\left|\bar{X}_N - \mu\right| \geq \epsilon\right) \leq \delta, \tag{A.4}$$

where $N = n, n+1, \ldots, n+m$. Equivalently, the strong law requires

$$P\left(\lim_{N\to\infty} \bar{X}_N = \mu\right) = 1. \tag{A.5}$$

The strong law asserts that, for any specified $\epsilon > 0$ and $0 < \delta < 1$, one can identify a number of trials n such that (A.4) holds for a sequence of values of $N > n$ of arbitrary finite length m, requiring that the sample mean approaches the expectation value in a more demanding way.

The weak law states that $\left|\bar{X}_N - \mu\right|$ eventually becomes small but that not every value on the way is small. It may be that, for some N, the difference is relatively large. The strong law says that the probability of such a large event, however, is extremely small. Either form of the law is sufficient to form a basis for the Monte Carlo method because both indicate that $\lim_{N\to\infty} \bar{X}_N = \mu$.

Let's consider a collection of n random variables $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$, where each component can be either discrete or continuous. The function $f(\mathbf{x})$ is the joint probability density function of $\mathbf{X}$ if it obeys that

$$f(\mathbf{x}) \geq 0 \quad \text{for all} \quad \mathbf{x} \in V \quad \text{and} \quad \int_V f(\mathbf{x}) d\mathbf{x} = 1, \tag{A.6}$$

where V defines the volume over which $\mathbf{x}$ is defined.

Then, let Z represent a stochastic process that is a function of the random variable $\mathbf{X}$, where $\mathbf{X}$ is governed by the joint probability density function $f(\mathbf{x})$. Then $Z(\mathbf{x})$ is also a random variable and one can define its the expectation value as

$$\mathbb{E}(Z) \equiv \int_V z(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}. \tag{A.7}$$

The heart of a Monte Carlo analysis is to obtain an estimate of the expectation value $\mathbb{E}(Z)$. If one forms the estimate

$$\bar{Z} = \frac{1}{N} \sum_{i=1}^{N} z\left(\mathbf{x}_i\right) \tag{A.8}$$

where the $\mathbf{x}_i$ are suitably sampled from $f(\mathbf{x})$, the law of large numbers states that, as long as the mean exists and the variance is bounded,

$$\lim_{N \to \infty} \bar{Z}_N = \mathbb{E}(Z). \tag{A.9}$$

Appendix B

Haar Measure

To test the VDGE algorithm, the geometrical measure of entanglement of randomly generated states was calculated for multiple dimensions using Monte Carlo simulations. Monte Carlo simulations are made under the assumptions that random samples are independent and identically distributed, so it is necessary to ensure that the random states are taken uniformly from the Hilbert space and not from a privileged region of the space state. This is solved by sampling uniformly at random from the Haar measure [94].

Measure theory [95] is the study of measures. It generalizes to mathematical spaces and even higher dimensions the intuitive notions of length, area, and volume. Let Ω be a set, and let Σ be a family of subsets of Ω. Σ is called a *σ-algebra* if it is an algebra of sets and is stable under the countable union operation, that is, for any sequence $A_n, n \in \mathbb{N}$, of elements of the algebra Σ their union is an element of Σ.

A set function $\mu : \Sigma \to \mathbb{R}$ is called a *measure* if it satisfies the following conditions:

1. (Non-negativity): $\mu(A) \geqslant 0$ for any $A \in \Sigma$.
2. (Countable additivity): If $A_1, A_2, \ldots, A_n \in \Phi$, the sets A_k are pairwise disjoint, and $\bigcup_{k=1}^{n} A_k \in \Phi$, then $\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu\left(A_k\right)$.

A triple (Ω, Σ, μ) is called a *measure space*. If μ is a probability measure, that is, $\mu(\Omega) = 1$, then (Ω, Σ, μ) is called a *probability space*. In probability theory the set Ω is referred to as the space of elementary events, the elements of the σ-algebra Σ as events, and $\mu(A)$ as the probability of the event A taking place.

Operations in quantum computing are described by unitary matrices, which can be expressed in terms of a fixed set of parameters. An $N \times N$ unitary matrix $U = (u_{jk})$ is defined by the relation $U^\dagger U = UU^\dagger = I$, which in terms of the matrix elements reads

$$\sum_{k=1}^{N} u_{jk}^\dagger u_{kl} = \sum_{k=1}^{N} u_{kj}^* u_{kl} = \delta_{jl} \quad \text{and} \quad \sum_{k=1}^{N} u_{jk} u_{kl}^\dagger = \sum_{k=1}^{N} u_{jk}^* u_{lk} = \delta_{jl}. \tag{B.1}$$

For every dimension N, the set of unitary matrices of size $N \times N$ constitute the unitary group $\mathrm{U}(N)$, which corresponds to a compact Lie group with real dimension N^2. Every compact Lie group has a unique (up to an arbitrary constant) left and right invariant measure, known as the *Haar measure*. In other words, if we denote the Haar measure on $\mathrm{U}(N)$ by $d\mu_{\mathrm{H}}(U)$, we have

$$d\mu_{\mathrm{H}}(VU) = d\mu_{\mathrm{H}}(UW) = d\mu_{\mathrm{H}}(U), \quad V, W \in \mathrm{U}(N). \tag{B.2}$$

The $\mathrm{U}(N)$ group can then be made into a probability space by assigning $d\mu_{\mathrm{H}}(U)$ as a distribution. Intuitively, the Haar measure tells us how to weight the elements of $\mathrm{U}(N)$. For example, let f be a function that acts on elements of $\mathrm{U}(N)$. The integral of f over the group is obtained as

$$\int_{V \in \mathrm{U}(N)} f(V) d\mu_H(V). \tag{B.3}$$

The sampling of random states of the Haar-uniform distribution can be made using the QR[1] decomposition of complex-valued matrices with the following steps:

1. Generate an $N \times N$ matrix Z with complex numbers $a + bi$, where both a and b are normally distributed with mean 0 and variance 1.
2. Compute a QR decomposition $Z = QR$.
3. Compute the diagonal matrix $\Lambda = \mathrm{diag}\left(R_{ii} / |R_{ii}|\right)$.
4. Compute $Q' = Q\Lambda$, which will be a Haar-random matrix.

[1] Any complex square matrix Z may be decomposed as

$$Z = QR, \tag{B.4}$$

where Q is a unitary matrix and R is an upper triangular matrix, $R_{ij} = 0$ if $i > j$

Appendix C

Matrix Product States (MPS)

As said before, matrix product states are a well-known representation of quantum states, that is useful for performing numerical simulations in many-body quantum systems.

Let's consider a lattice of n sites with d-dimensional local state spaces $\{|i\rangle\}$ on sites $i = 1, \ldots, n$. The most general pure quantum state on the lattice can be written as

$$|\psi\rangle = \sum_{i_1,\ldots,i_n} c_{i_1\ldots i_n} |i_1\rangle |i_2\rangle \cdots |i_n\rangle , \tag{C.1}$$

with exponentially many coefficients $c_{i_1\ldots i_n}$. The *left-canonical matrix product state* of $|\psi\rangle$ can be constructed as follows [96]. First, the state vector with d^n components is reshaped into a matrix Ψ of dimension $(d \times d^{n-1})$, where the coefficients are related as

$$\Psi_{i_1,(i_2\ldots i_n)} = c_{i_1\ldots i_n}. \tag{C.2}$$

Then, the singular value decomposition of Ψ gives

$$c_{i_1\ldots i_n} = \Psi_{i_1,(i_2\ldots i_n)} = \sum_{a_1}^{r_1} U_{i_1,a_1} D_{a_1,a_1} \left(V^\dagger\right)_{a_1,(i_2\ldots i_n)} \equiv \sum_{a_1}^{r_1} U_{i_1,a_1} c_{a_1 i_2\ldots i_n}, \tag{C.3}$$

where the rank is $r_1 \leq d$. In the last equality, D and $V^\dagger$ have been multiplied and the resulting matrix has been reshaped back into the vector $c_{a_1 i_2\ldots i_n}$.

The matrix U can then be decomposed into a collection of d row vectors A^{i_1} with entries $A^{i_1}_{a_1} = U_{i_1,a_1}$. At the same time, the vector $c_{a_1 i_2\ldots i_n}$ is reshaped into a matrix

$\Psi_{(a_1 i_2),(i_3 \ldots i_n)}$ of dimension $(r_1 d \times d^{n-2})$, which gives

$$c_{i_1 \ldots i_n} = \sum_{a_1}^{r_1} A_{a_1}^{i_1} \Psi_{(a_1 i_2),(i_3 \ldots i_n)}. \tag{C.4}$$

The singular value decomposition is applied again, obtaining

$$c_{i_1 \ldots i_n} = \sum_{a_1}^{r_1} \sum_{a_2}^{r_2} A_{a_1}^{i_1} U_{(a_1 i_2),a_2} D_{a_2,a_2} \left(V^\dagger\right)_{a_2,(i_3 \ldots i_n)} \tag{C.5}$$

$$= \sum_{a_1}^{r_1} \sum_{a_2}^{r_2} A_{a_1}^{i_1} A_{a_1,a_2}^{i_2} \Psi_{(a_2 i_3),(i_4 \ldots i_n)}, \tag{C.6}$$

where U is replaced by a set of d matrices A^{i_2} of dimension $(r_1 \times r_2)$ with entries $A_{a_1,a_2}^{i_2} = U_{(a_1 i_2),a_2}$. The matrices D and $V^\dagger$ are multiplied and reshaped into a matrix Ψ of dimension $(r_2 d \times d^{n-3})$, where $r_2 \leq r_1 d \leq d^2$.

The singular value decomposition and reshaping is performed repeatedly, obtaining

$$c_{i_1 \ldots i_n} = \sum_{a_1,\ldots,a_{n-1}} A_{a_1}^{i_1} A_{a_1,a_2}^{i_2} \ldots A_{a_{n-2},a_{n-1}}^{i_{n_1}} A_{a_{n-1}}^{i_n} \equiv A_{[1]}^{i_1} A_{[2]}^{i_2} \ldots A_{[n-1]}^{i_{n-1}} A_{[n]}^{i_n}. \tag{C.7}$$

where $A_{[m]}^{i}$ are $(r_{m-1} \times r_m)$ matrices with $r_0 = r_n = 1$ and satisfy that

$$\sum_i A_{[m]}^{i} A_{[m]}^{i\dagger} = I_{r_m} \text{ for all } 1 \leq m \leq n. \tag{C.8}$$

Moreover, if $\chi = \max_m r_m$, the MPS is said to have *bond dimension* χ [97].

The arbitrary quantum state is now represented exactly in the form of a matrix product state:

$$|\psi\rangle = \sum_{i_1,\ldots,i_n} A_{[1]}^{i_1} A_{[2]}^{i_2} \ldots A_{[n-1]}^{i_{n-1}} A_{[n]}^{i_n} |i_1\rangle |i_2\rangle \cdots |i_n\rangle . \tag{C.9}$$

To cover all the states in the Hilbert space, χ needs to be exponentially large in the system size. The maximal dimensions of the matrices are reached when, for each singular value decomposition done, the number of non-zero singular values is equal to the upper bound. The dimensions may maximally be

$(1 \times d), (d \times d^2), \ldots, \left(d^{n/2-1} \times d^{n/2}\right), \left(d^{n/2} \times d^{n/2-1}\right), \ldots, (d^2 \times d), (d \times 1)$, going from the first to the last site. This shows that in practical calculations it will usually be impossible to carry out this exact decomposition explicitly, as the matrix dimensions grow up exponentially.

However, it is known that low-energy states of gapped local Hamiltonians in one-diemnsional systems can be efficiently approximated with almost arbitrary accuracy by a MPS with a finite value of D [85].

The $|\mathrm{GHZ}_N\rangle$ and $|\mathrm{W}_N\rangle$ states both can be represented with MPS with bond dimension $\chi = 2$ for any number of qubits N. The $|\mathrm{GHZ}_N\rangle$ state can be represented by a MPS with the following matrices $A^i_{[m]}$:

$$A^0_{[1]} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad A^1_{[1]} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \tag{C.10}$$

$$A^0_{[j]} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A^1_{[j]} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{for all} \quad 1 < j < N \tag{C.11}$$

$$A^0_{[N]} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad A^1_{[N]} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{C.12}$$

Bibliography

[1] E. Schrödinger. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(4): 555–563, 1935. doi: 10.1017/S0305004100013554.

[2] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 5 1935. doi: 10.1103/PhysRev.47.777.

[3] C. H. Bennett. Quantum information and computation. *Physics Today*, 48 (10):24–30, 1995. doi: 10.1063/1.881452.

[4] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, October 1997. doi: 10.1137/S0097539795293172.

[5] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96*. ACM Press, 1996. doi: 10.1145/237814.237866. URL https://doi.org/10.1145/237814.237866.

[6] 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, 3 1993. doi: 10.1103/PhysRevLett.70.1895.

[7] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Physical Review Letters*, 69(20):2881–2884, November 1992. doi: 10.1103/physrevlett.69.2881. URL https://doi.org/10.1103/physrevlett.69.2881.

[8] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, 2014. doi: 10.1016/j.tcs.2014.05.025. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.

[9] C. L. Degen, F. Reinhard, and P. Cappellaro. Quantum sensing. *Rev. Mod. Phys.*, 89:035002, 7 2017. doi: 10.1103/RevModPhys.89.035002.

[10] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum-enhanced measurements:

Beating the standard quantum limit. *Science*, 306(5700):1330–1336, 2004. doi: 10.1126/science.1104149.

[11] V. Giovannetti, S. Lloyd, and L. Maccone. Advances in quantum metrology. *Nature Photonics*, 5(4):222–229, 4 2011. doi: 10.1038/nphoton.2011.35.

[12] W. K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245–2248, 3 1998. doi: 10.1103/PhysRevLett.80.2245.

[13] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight. Quantifying entanglement. *Physical Review Letters*, 78(12):2275–2279, March 1997. doi: 10.1103/physrevlett.78.2275. URL https://doi.org/10.1103/physrevlett.78.2275.

[14] V. Vedral and M. B. Plenio. Entanglement measures and purification procedures. *Physical Review A*, 57(3):1619–1633, March 1998. doi: 10.1103/physreva.57.1619. URL https://doi.org/10.1103/physreva.57.1619.

[15] A. Shimony. Degree of entanglementa. *Annals of the New York Academy of Sciences*, 755(1):675–679, 1995. doi: 10.1111/j.1749-6632.1995.tb39008.x.

[16] H Barnum and N Linden. Monotones and invariants for multi-particle quantum states. *Journal of Physics A: Mathematical and General*, 34(35):6787–6805, August 2001. doi: 10.1088/0305-4470/34/35/305. URL https://doi.org/10.1088/0305-4470/34/35/305.

[17] M. Hayashi, D. Markham, M. Murao, M. Owari, and S. Virmani. Bounds on multipartite entangled orthogonal state discrimination using local operations and classical communication. *Physical Review Letters*, 96(4), February 2006. doi: 10.1103/physrevlett.96.040501. URL https://doi.org/10.1103/physrevlett.96.040501.

[18] Román Orús, Sébastien Dusuel, and Julien Vidal. Equivalence of critical scaling laws for many-body entanglement in the lipkin-meshkov-glick model. *Physical Review Letters*, 101(2), July 2008. doi: 10.1103/physrevlett.101.025701. URL https://doi.org/10.1103/physrevlett.101.025701.

[19] Ofer Biham, Michael A. Nielsen, and Tobias J. Osborne. Entanglement monotone derived from grover's algorithm. *Physical Review A*, 65(6), June 2002. doi: 10.1103/physreva.65.062312. URL https://doi.org/10.1103/physreva.65.062312.

[20] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018. doi: 10.22331/q-2018-08-06-79. URL https://doi.org/10.22331/q-2018-08-06-79.

[21] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 9 2021. doi: 10.1038/s42254-021-00348-9.

[22] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand,

M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik. Noisy intermediate-scale quantum NISQ algorithms, 2021. URL https://arxiv.org/abs/2101.08448v1.

[23] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):4213, 7 2014. doi: 10.1038/ncomms5213.

[24] D. Wecker, M. B. Hastings, and M. Troyer. Progress towards practical quantum variational algorithms. *Phys. Rev. A*, 92:042303, 10 2015. doi: 10.1103/PhysRevA.92.042303.

[25] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, 2 2016. doi: 10.1088/1367-2630/18/2/023023.

[26] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549(7671):242–246, 9 2017. doi: 10.1038/nature23879.

[27] R. Kaubruegger, P. Silvi, C. Kokail, R. van Bijnen, A. M. Rey, J. Ye, A. M. Kaufman, and P. Zoller. Variational spin-squeezing algorithms on programmable quantum sensors. *Phys. Rev. Lett.*, 123:260505, 12 2019. doi: 10.1103/PhysRevLett.123.260505.

[28] B. Koczor, . Endo, T. Jones, Y. Matsuzaki, and S. C. Benjamin. Variational-state quantum metrology. *New Journal of Physics*, 22(8):083038, 8 2020. doi: 10.1088/1367-2630/ab965e.

[29] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 9 2017. doi: 10.1038/nature23474.

[30] IBM Quantum, 2021. URL https://quantum-computing.ibm.com/.

[31] Henk Tijms. *Understanding Probability*. Cambridge University Press, 2009. doi: 10.1017/cbo9781139206990. URL https://doi.org/10.1017/cbo9781139206990.

[32] Venkatarama Krishnan. *Probability and Random Processes*. John Wiley & Sons, Inc., November 2005. doi: 10.1002/0471998303. URL https://doi.org/10.1002/0471998303.

[33] J. J. Sakurai and Jim Napolitano. *Modern Quantum Mechanics*. Cambridge University Press, September 2017. doi: 10.1017/9781108499996. URL https://doi.org/10.1017/9781108499996.

[34] N. Zettili. *Quantum Mechanics: Concepts and Applications*. Wiley, 2009. ISBN 9780470026786.

[35] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2009. doi: 10.1017/cbo9780511976667. URL https://doi.org/10.1017/cbo9780511976667.

[36] Vlatko Vedral. *Introduction to Quantum Information Science*. Oxford University Press, September 2006. doi: 10.1093/acprof:oso/9780199215706.001.0001. URL https://doi.org/10.1093/acprof:oso/9780199215706.001.0001.

[37] P. A. M. Dirac. A new notation for quantum mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 35(3):416–418, July 1939. doi: 10.1017/s0305004100021162. URL https://doi.org/10.1017/s0305004100021162.

[38] Radhika Vathsan. *Introduction to Quantum Physics and Information Processing*. CRC Press, August 2015. doi: 10.1201/b18767. URL https://doi.org/10.1201/b18767.

[39] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2009. doi: 10.1017/cbo9781139525343. URL https://doi.org/10.1017/cbo9781139525343.

[40] Martin B. Plenio and Shashank Virmani. An introduction to entanglement measures. *Quantum Info. Comput.*, 7(1):1–51, jan 2007. ISSN 1533-7146.

[41] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of Modern Physics*, 81(2):865–942, June 2009. doi: 10.1103/revmodphys.81.865. URL https://doi.org/10.1103/revmodphys.81.865.

[42] Otfried Gühne and Géza Tóth. Entanglement detection. *Physics Reports*, 474(1-6):1–75, April 2009. doi: 10.1016/j.physrep.2009.02.004. URL https://doi.org/10.1016/j.physrep.2009.02.004.

[43] Dagmar Bruß and G. Leuchs, editors. *Lectures on Quantum Information*. Wiley, November 2006. doi: 10.1002/9783527618637. URL https://doi.org/10.1002/9783527618637.

[44] Dominique Spehner. Quantum correlations and distinguishability of quantum states. *Journal of Mathematical Physics*, 55(7):075211, July 2014. doi: 10.1063/1.4885832. URL https://doi.org/10.1063/1.4885832.

[45] Gregg Jaeger. *Quantum Information*. Springer New York, 2007. doi: 10.1007/978-0-387-36944-0. URL https://doi.org/10.1007/978-0-387-36944-0.

[46] Reinhard F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Physical Review A*, 40 (8):4277–4281, October 1989. doi: 10.1103/physreva.40.4277. URL https://doi.org/10.1103/physreva.40.4277.

[47] A. V. Arkhangel'skii and L. S. Pontryagin, editors. *General Topology I*.

Springer Berlin Heidelberg, 1990. doi: 10.1007/978-3-642-61265-7. URL https://doi.org/10.1007/978-3-642-61265-7.

[48] Zhihao Ma, Fu-Lin Zhang, and Jing-Ling Chen. Fidelity induced distance measures for quantum states. *Physics Letters A*, 373(38):3407–3409, September 2009. doi: 10.1016/j.physleta.2009.07.042. URL https://doi.org/10.1016/j.physleta.2009.07.042.

[49] Multi partite entanglement. Multi-partite entanglement, 2017.

[50] Ingemar Bengtsson and Karol Zyczkowski. *Geometry of Quantum States*. Cambridge University Press, 2017. doi: 10.1017/9781139207010. URL https://doi.org/10.1017/9781139207010.

[51] Guifré Vidal. Entanglement monotones. *Journal of Modern Optics*, 47(2-3):355–376, February 2000. doi: 10.1080/09500340008244048. URL https://doi.org/10.1080/09500340008244048.

[52] Armin Uhlmann. Roofs and convexity. *Entropy*, 12(7):1799–1832, July 2010. doi: 10.3390/e12071799. URL https://doi.org/10.3390/e12071799.

[53] Ralph Tyrell Rockafellar. *Convex Analysis*. Princeton University Press, December 1970. doi: 10.1515/9781400873173. URL https://doi.org/10.1515/9781400873173.

[54] Edward Witten. A mini-introduction to information theory. *La Rivista del Nuovo Cimento*, 43(4):187–227, March 2020. doi: 10.1007/s40766-020-00004-5. URL https://doi.org/10.1007/s40766-020-00004-5.

[55] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Physical Review A*, 53(4):2046–2052, April 1996. doi: 10.1103/physreva.53.2046. URL https://doi.org/10.1103/physreva.53.2046.

[56] Alexander Streltsov, Hermann Kampermann, and Dagmar Bruß. Linking a distance measure of entanglement to its convex roof. *New Journal of Physics*, 12(12):123004, December 2010. doi: 10.1088/1367-2630/12/12/123004. URL https://doi.org/10.1088/1367-2630/12/12/123004.

[57] T.-C. Wei, M. Ericsson, P. M. Goldbart, and W. J. Munro. Connections between relative entropy of entanglement and geometric measure of entanglement. *Quantum Info. Comput.*, 4(4):252–272, July 2004. doi: 10.26421/QIC4.4-2.

[58] A. Acín, A. Andrianov, L. Costa, E. Jané, J. I. Latorre, and R. Tarrach. Generalized schmidt decomposition and classification of three-quantum-bit states. *Physical Review Letters*, 85(7):1560–1563, August 2000. doi: 10.1103/physrevlett.85.1560. URL https://doi.org/10.1103/physrevlett.85.1560.

[59] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two

inequivalent ways. *Physical Review A*, 62(6), November 2000. doi: 10.1103/physreva.62.062314. URL https://doi.org/10.1103/physreva.62.062314.

[60] Michael Seevinck and Jos Uffink. Partial separability and entanglement criteria for multiqubit quantum states. *Physical Review A*, 78(3), September 2008. doi: 10.1103/physreva.78.032101. URL https://doi.org/10.1103/physreva.78.032101.

[61] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde. Four qubits can be entangled in nine different ways. *Physical Review A*, 65(5), April 2002. doi: 10.1103/physreva.65.052112. URL https://doi.org/10.1103/physreva.65.052112.

[62] L. Lamata, J. León, D. Salgado, and E. Solano. Inductive classification of multipartite entanglement under stochastic local operations and classical communication. *Physical Review A*, 74(5), November 2006. doi: 10.1103/physreva.74.052336. URL https://doi.org/10.1103/physreva.74.052336.

[63] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Physical Review A*, 63(1), December 2000. doi: 10.1103/physreva.63.012307. URL https://doi.org/10.1103/physreva.63.012307.

[64] Jens Eisert and Hans J. Briegel. Schmidt measure as a tool for quantifying multiparticle entanglement. *Physical Review A*, 64(2), July 2001. doi: 10.1103/physreva.64.022306. URL https://doi.org/10.1103/physreva.64.022306.

[65] T.-C. Wei and P. M. Goldbart. Geometric measure of entanglement and applications to bipartite and multipartite quantum states. *Phys. Rev. A*, 68: 042307, 10 2003. doi: 10.1103/PhysRevA.68.042307.

[66] C. Ferrie. Self-guided quantum tomography. *Phys. Rev. Lett.*, 113:190404, 11 2014. doi: 10.1103/PhysRevLett.113.190404.

[67] A. Utreras-Alarcón, M. Rivera-Tapia, S. Niklitschek, and A. Delgado. Stochastic optimization on complex variables and pure-state quantum tomography. *Scientific Reports*, 9(1), November 2019. doi: 10.1038/s41598-019-52289-0. URL https://doi.org/10.1038/s41598-019-52289-0.

[68] J.C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3):332–341, March 1992. doi: 10.1109/9.119632. URL https://doi.org/10.1109/9.119632.

[69] J.C. Spall. An overview of the simultaneous perturbation method for efficient optimization. *Johns Hopkins apl technical digest*, 19(4):482–492, 1998.

[70] Ken Kreutz-Delgado. The complex gradient operator and the cr-calculus, 2009.

[71] Richard P. Feynman. Simulating physics with computers. *International*

Journal of Theoretical Physics, 21(6-7):467–488, June 1982. doi: 10.1007/bf02650179. URL https://doi.org/10.1007/bf02650179.

[72] Masuo Suzuki. Generalized trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Communications in Mathematical Physics*, 51(2):183–190, June 1976. doi: 10.1007/bf01609348. URL https://doi.org/10.1007/bf01609348.

[73] Andrew G. Taube and Rodney J. Bartlett. New perspectives on unitary coupled-cluster theory. *International Journal of Quantum Chemistry*, 106 (15):3393–3401, 2006. doi: 10.1002/qua.21198. URL https://doi.org/10.1002/qua.21198.

[74] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Physical Review A*, 98(3), September 2018. doi: 10.1103/physreva.98.032309. URL https://doi.org/10.1103/physreva.98.032309.

[75] Maria Schuld, Ville Bergholm, Christian Gogolin, Josh Izaac, and Nathan Killoran. Evaluating analytic gradients on quantum hardware. *Physical Review A*, 99(3), March 2019. doi: 10.1103/physreva.99.032331. URL https://doi.org/10.1103/physreva.99.032331.

[76] Matteo M. Wauters, Emanuele Panizon, Glen B. Mbeng, and Giuseppe E. Santoro. Reinforcement-learning-assisted quantum optimization. *Physical Review Research*, 2(3), September 2020. doi: 10.1103/physrevresearch.2.033446. URL https://doi.org/10.1103/physrevresearch.2.033446.

[77] Ken M. Nakanishi, Keisuke Fujii, and Synge Todo. Sequential minimal optimization for quantum-classical hybrid algorithms. *Physical Review Research*, 2(4), October 2020. doi: 10.1103/physrevresearch.2.043158. URL https://doi.org/10.1103/physrevresearch.2.043158.

[78] Xiao Yuan, Suguru Endo, Qi Zhao, Ying Li, and Simon C. Benjamin. Theory of variational quantum simulation. *Quantum*, 3:191, October 2019. doi: 10.22331/q-2019-10-07-191. URL https://doi.org/10.22331/q-2019-10-07-191.

[79] B. Efron and R.J. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall/CRC, 5 1994. doi: 10.1201/9780429246593.

[80] D. J. Wales and J. P. K. Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, 1997. doi: 10.1021/jp970984n.

[81] R. Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117–158, 2014. doi: 10.1016/j.aop.2014.06.013.

[82] A. Klumper, A. Schadschneider, and J. Zittartz. Equivalence and solution of anisotropic spin-1 models and generalized t-j fermion models in one dimension.

Journal of Physics A: Mathematical and General, 24(16):L955–L959, 8 1991. doi: 10.1088/0305-4470/24/16/012.

[83] M. Fannes, B. Nachtergaele, and R. F. Werner. Finitely correlated states on quantum spin chains. *Communications in Mathematical Physics*, 144(3): 443–490, 3 1992. doi: 10.1007/bf02099178.

[84] A. Klümper, A. Schadschneider, and J. Zittartz. Matrix product ground states for one-dimensional spin-1 quantum antiferromagnets. *Europhysics Letters (EPL)*, 24(4):293–297, 11 1993. doi: 10.1209/0295-5075/24/4/010.

[85] Steven R. White. Density matrix formulation for quantum renormalization groups. *Physical Review Letters*, 69(19):2863–2866, November 1992. doi: 10.1103/physrevlett.69.2863. URL https://doi.org/10.1103/physrevlett.69.2863.

[86] F. Verstraete, D. Porras, and J. I. Cirac. Density matrix renormalization group and periodic boundary conditions: A quantum information perspective. *Phys. Rev. Lett.*, 93:227205, 11 2004. doi: 10.1103/PhysRevLett.93.227205.

[87] J. J. García-Ripoll. Quantum-inspired algorithms for multivariate analysis: from interpolation to partial differential equations. *Quantum*, 5:431, April 2021. doi: 10.22331/q-2021-04-15-431.

[88] B.-Q. Hu, X.-J. Liu, J.-H. Liu, and H.-Q. Zhou. Geometric entanglement from matrix product state representations. *New Journal of Physics*, 13(9): 093041, sep 2011. doi: 10.1088/1367-2630/13/9/093041.

[89] P. Teng. Accurate calculation of the geometric measure of entanglement for multipartite quantum states. *Quantum Information Processing*, 16(7), June 2017. doi: 10.1007/s11128-017-1633-8.

[90] G. J. Mooney, G. A. L. White, C. D. Hill, and L. C. L. Hollenberg. Generation and verification of 27-qubit greenberger-horne-zeilinger states in a superconducting quantum computer. *Journal of Physics Communications*, 5(9):095004, 9 2021. doi: 10.1088/2399-6528/ac1df7.

[91] A. Cervera-Lierta, M. Krenn, A. Aspuru-Guzik, and A. Galda. Experimental high-dimensional greenberger-horne-zeilinger entanglement with superconducting transmon qutrits, 2021. URL https://arxiv.org/abs/2104.05627.

[92] M. Ringbauer, M. Meth, L. Postler, R. Stricker, R. Blatt, P. Schindler, and T. Monz. A universal qudit quantum processor with trapped ions, 2021. URL https://arxiv.org/abs/2109.06903.

[93] William L. Dunn and J. Kenneth Shultis. *Exploring Monte Carlo Methods*. Elsevier, Amsterdam, 2012. ISBN 978-0-444-51575-9. doi: https://doi.org/10.1016/B978-0-444-51575-9.00002-6.

[94] Francesco Mezzadri. How to generate random matrices from the classical

compact groups. *Notices of the American Mathematical Society*, 54(5):592 – 604, May 2007. ISSN 0002-9920.

[95] Vladimir Kadets. *A Course in Functional Analysis and Measure Theory*. Springer International Publishing, 2018. doi: 10.1007/978-3-319-92004-7. URL https://doi.org/10.1007/978-3-319-92004-7.

[96] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96–192, January 2011. doi: 10.1016/j.aop.2010.09.012. URL https://doi.org/10.1016/j.aop.2010.09.012.

[97] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac. Matrix product state representations. *Quantum Info. Comput.*, 7(5):401–430, jul 2007. ISSN 1533-7146.