Universitat Autònoma de Barcelona

Physics and Information
What is the role of Information in Physics?

Paul Erker

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons: http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons: http://es.creativecommons.org/blog/licencias/

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license: https://creativecommons.org/licenses/?lang=en

Universitat Autònoma de Barcelona

Physics and Information
What is the role of Information in Physics?

Doctoral Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics to the Unitat de Fisica Teorica: Informacio i fenomens quantics, Universitat Autonoma de Barcelona and in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Informatics to the Faculty of Informatics, Universitá della Svizzera Italiana
by

Paul Erker
under supervision of
Stefan Wolf, Andreas Winter & Marcus Huber

2018

Dissertation Committee

UAB = Universitat Autonoma de Barcelona
USI = Universitá della Svizzera Italiana
IQOQI = Institut für Quantenoptik und Quanteninformation

Prof. Dr. Andreas Winter, UAB

Prof. Dr. Stefan Wolf, USI

Dr. Marcus Huber, UAB & IQOQI

Prof. Dr. Antonio Carzaniga, USI

Prof. Dr. Olaf Schenk, USI

Prof. Dr. John Calsamiglia, UAB

Prof. Dr. Časlav Brukner, IQOQI & University of Vienna

PhD Program Director

I certify that except where due acknowledgement has been given, the work presented in this thesis is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; and the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program.

Paul Erker

Acknowledgements

I sincerely thank
Marcus Huber, Andreas Winter & Stefan Wolf
for their support.

Thanks to my parents for their unconditional love.
Thanks my sister for keeping the flow. Thanks to my brother for his propaganda.
Thanks also to the rest of the big family.

Thanks to 369, LIQUID, FiGa and Pratersauna
for all the labouring, working and acting that we did together.
Thanks to Peršmanhof for providing a partizan hideout while I was writing this thesis.

Thanks to all the people and entities that accompanied me on my path and
that were not mentioned here explicitly.
I extend my full gratitude for every little piece of it.
You can be sure that I carry you in my heart.

Never forget!

Tem boi na linha!!!

Abstract

More than a century ago, physicists around the world were collectively developing a theory to describe the newly discovered strange behaviours of some physical systems. This marks the birth of quantum theory. Few decades later, the groundbreaking idea to separate information from its physical carrier led to the establishment of information theory. These, initially independent theories, merged together in the last decades of the former century, leaving us with quantum information theory. This thesis will explore topics at the intersection of mathematics, physics and computer science, trying to elucidate the interwovenness of these three disciplines. While doing so, the results that were established during the years of studies leading up to this work are introduced. The question posed in the title will not be answered fully, as it may be too early still to give a definite answer to this multifaceted question.

Resumen

Hace más de un siglo, los físicos de todo el mundo estaban desarrollando una teoría para describir comportamientos extraños recientemente descubiertos de algunos sistemas físicos, lo que marca el nacimiento de la teoría cuántica. Algunas décadas más tarde, la idea innovadora de separar la información de su portador físico llevó al establecimiento de la teoría de la información. Estas teorías independientes al principio se fusionaron en las últimas décadas del siglo anterior, dando lugar a la teoría de la información cuántica. Esta tesis explorará temas en la intersección entre matemáticas, física y ciencias de la computación, tratando de dilucidar la interconexión entre las tres. Dicha exploración se llevará a cabo junto con la presentación de los resultados obtenidos durante estos años de estudio. La pregunta planteada en el título no será respondida completamente, ya que puede ser demasiado pronto para dar una respuesta definitiva a esta amplia pregunta.

Resum

Fa més d'un segle els físics de tot el món desenvolupaven una teoria per descriure comportaments estranys recentment descobertes d'alguns sistemes físics, això marca el naixement de la Teoria Quàntica. Algunes dècades més tard, la idea revolucionària de separar la informació de la seva companyia física va conduir a l'establiment de la Teoria de la Informació. Aquests al principi es van unir les teories independents en les últimes dècades del segle anterior deixant-nos de Teoria de la Informació Quàntica. Aquesta tesi explorarà una intersecció entre matemàtiques, física i informàtica, tractant d'aclarir l'entrellaçat de l'anterior. Tot seguit, es van establir els resultats que es van establir durant els anys d'estudis que van dur a aquest treball. La pregunta plantejada en el títol no es respondrà plenament ja que pot ser massa aviat per donar una resposta definitiva a aquesta àmplia pregunta.

Contents

Chapter 1

Introduction

This dissertation is written under a cotutela agreement mutually established between the Faculty of Informatics of the Universitá della Svizzera Italiana and the Unitat de Fisica Teorica: Informacio i fenomens quantics of the Universitat Autonoma de Barcelona (see Appendix A). As the concept and understanding of what a doctoral education is changing and has changed throughout history this work starts with a short excursion on the most recent official definitions of the framework underlying this thesis.

Both the Swiss Confederation and the Kingdom of Spain have signed the Bologna Declaration in 1999 [1]. Therefore the concepts and definitions established in the process that followed declaration shall serve as a basis for finding out what a doctorate is within this framework, naturally respecting the national regulations of each side. The European Qualifications Framework [2], the latest step towards a unified higher education in europe, defines the learning outcomes of level 8 as follows

- knowledge at the most advanced frontier of a field of work or study and at the interface between fields
- the most advanced and specialised skills and techniques, including synthesis and evaluation, required to solve critical problems in research and/or innovation and to extend and redefine existing knowledge or professional practice
- demonstrate substantial authority, innovation, autonomy, scholarly and professional integrity and sustained commitment to the development of new ideas or processes at the forefront of work or study contexts including research.

and adds that

> Level 8 is compatible with the Framework for Qualifications of the European Higher Education Area.

So one can go and take a look at the Framework for Qualifications of the European Higher Education Area [3]. It turns out that level 8 of the European Qualifications Framework corresponds to completing the third cycle of the Framework for Qualifications of the European Higher Education Area [3]. In the report one can find that the title corresponding to completing a doctorate shall be awarded to students who:

- have demonstrated a systematic understanding of a field of study and mastery of the skills and methods of research associated with that field;
- have demonstrated the ability to conceive, design, implement and adapt a substantial process of research with scholarly integrity;
- have made a contribution through original research that extends the frontier of knowledge by developing a substantial body of work, some of which merits national or international refereed publication;
- are capable of critical analysis, evaluation and synthesis of new and complex ideas;
- can communicate with their peers, the larger scholarly community and with society in general about their areas of expertise;
- can be expected to be able to promote, within academic and professional contexts, technological, social or cultural advancement in a knowledge based society.

with a footnote defining the term 'research' as follows

> The word 'research' is used to cover a wide variety of activities, with the context often related to a field of study; the term is used here to represent a careful study or investigation based on a systematic understanding and critical awareness of knowledge. The word is used in an inclusive way to accommodate the range of activities that support original and innovative work in the whole range of academic, professional and technological fields, including the humanities, and traditional, performing, and other creative arts. It is not used in any limited or restricted sense, or relating solely to a traditional 'scientific method'.

One may ask the question why all above definitions are cited here. The answer lies in the question about what a doctoral thesis actually is. Unfortunately the living traditions of the academic institutions that weave together a meaning for the theses that written there, are in general not citable. Reading literature on the history and development of academia and specially PhD theses [4, 5, 6], gives a lot of insight but is only a partially satisfying answer to above question. Bureaucrats, however, always seem to find words to categorise the world.

Theses generically do not answer the question what they are supposed to be, but rather start with an introduction to an already highly specialised field. In both the physics and the information theory community self-containedness and consistency are highly regarded properties. That is why I want this thesis to be self-contained and consistent not only in technical terms of the branch of science this thesis is written in but also in the context of this work being a doctoral dissertation, i.e. this thesis starts with what it is.

1.1 Outline

After having introduced what this dissertation is in the introductory chapter, here shall be the outline of the content of this thesis. Chapter 2 starts the overview of quantum information theory, also quantum theory and information theory will be discussed. The technical fundamentals are established and all basic notions necessary are given. Chapter 3 discusses the importance of bases for the description of the systems used in quantum information theory and also includes the introduction of a basis that was developed within this doctoral studies. Chapter 4 then continues with the description and discussion of one of the most essential phenomenon of the field, namely entanglement. This also includes applications and theoretical tools for the detection and quantification of entanglement. Chapter 5 will discuss the interplay of thermodynamics and information. In this chapter quantum machines will be introduced and there is a short excourse on how to go a step further. Chapter 6 will discuss time and specifically how to measure it. For this purpose the notion of the autonomous clock is described and its properties analysed. This thesis ends with Chapter 7 which contains the conclusions.

Chapter 2 is partially based on [7]. Chapter 3 and 4 are partially based on [8, 9, 10]. Chapter 5 and 6 are partially based on [11].

Chapter 2

Quantum Information Theory

This chapter will start with a brief historic overview of the origins of quantum theory and information theory. We will then see how these theories merged together to quantum information theory. After this, all necessary definitions and notions that we will need to understand this intersection between physics, mathematics and informatics will be introduced. This chapter ends with a no-go result on the generalisation of an information measure to situations where identical and independently distributed (i.i.d.) resources are not available.

2.1 Quantum Theory

Unlike other physical theories, quantum theory was not delivered in one piece as a 'quantum' of knowledge. It was developed over several decades by various people until it became the theory that will be presented in this chapter. Here we want to follow the history of how these bits and pieces came together.

The first step towards quantum mechanics was actually an attempt to save classical physics. This happened at a time when young students interested in the physics of nature were advised not to do so, as everything has been done already and there was nothing new to discover [12]. It seemed at that point, that studying physics had been reduced to mere engineering, in the sense that one would study physics solely to learn how to apply the given laws. A small, but groundbreaking step was introduction of the 'Wirkungsquantum' by Max Planck in 1900. The 'Wirkungsquantum', a quantum of action, is a constant that relates the frequency of a electromagnetic wave to its energy. First, it was just a working hypothesis in order to fix the inconsistencies that arose in the description of the black body radiation by the Rayleigh-Jeans and Wien's law, known as the 'ultraviolet catastrophe' and the 'infrared divergence', pointing at the spectra of light where the theory was failing to reproduce the experimental data. It would take another ten years for Planck to make sense of the implications of his newly introduced concept.

It was Albert Einstein who convinced Planck at the first Solvey conference [13], that the theory of quanta of lights, photons, he had used to describe the photoelectric effect made sense and that the 'Wirkungsquantum' was a constant of nature, not just an number to fine-tune an effective theory. Still, it was up to other scientists to figure out how all the puzzle pieces would fit together, as

both Einstein and Planck were not totally satisfied with the genie that they let out of the bottle.

Here the story gets complicated to tell, as a lot developments happened at the same time. Let us start north, where Niels Bohr was one of the first physicist to use the new theory of quantas to come up with a more consistent description of the atom. Although Bohr's model of an atom was later replaced by one of orbitals, it was one of the first milestones. He would continue until his death to work on quantum theory. His principles of correspondence and complementarity form an essential part of the efforts to make sense of this newly arising theory.

Meanwhile, in Hamburg, Wolfgang Pauli a physicist from Vienna, got to know Otto Stern, whose experiments together with Walther Gerlach, known as the Stern-Gerlach experiment, marked another stepping stone for the young theory. With the help of some coincidences they managed to show the quantisation of the spatial direction of the angular momentum in silver atoms. This quantised form of angular momentum was later found by Pauli to be a new degree of freedom, that is nowadays known as spin. After Werner Heisenberg introduced his matrix theory of quantum mechanics, Pauli used it to reproduce the observed spectrum of the hydrogen atom, which in turn gave credibility to Heisenberg's new formalism.

Another physicist from Vienna, named Erwin Schrödinger, living in Switzerland at that time found a different way to model the observed quantisation of the world. In his seminal paper 'Quantisierung als Eigenwertproblem' he presents a differential equation, now known as Schrödinger equation, that also gives the right energy eigenvalues for the hydrogen atom. It quickly became clear that Heisenberg's matrix theory and Schrödinger's wave mechanics are actually equivalent [14]. This discovery is generally attributed to Paul Dirac, whose very convenient notation, called bra-ket notation, is used up to today.

A decade later, the research on quantum theory in central Europe had ceased due to the rise of the fascist Nazi regime. Many of the founders named above had to flee due to their political views or ancestry. One scientist that was fighting the Fascists since the beginning of the thirties was Grete Hermann (principiis obsta sero medicina paratur cum mala per longas convaluere moras). Aside from actively fighting the nazis, she pursued research on the foundations of physics with special focus on quantum theory. Although she was actively involved in the discussion on quantum theory, exchanging ideas with many of the above mentioned, her work was overlooked for decades. Not only had she already conducted research on the distinction between predictability and causality, she had also already given a proof for the impossibility of hidden variable theory for quantum mechanics. This was published in 'Welche Konsequenzen haben die Quantentheorie und die Feldtheorie der modernen Physik für die Theorie der Erkenntnis?' were she criticises John von Neumann's version the proof. Her proof was independently discovered by John Stewart Bell decades later. Despite it being counterfactual, one can imagine how the world could have had developed if she would have been taken seriously.

2.2 Information Theory

World War II marks the end of the completion of quantum theory, as well as it marks the beginning of research in informatics. Already at the beginning of the forties, the Third Reich had gained new territory quickly allowing the Germans to expand their genocide operation. The Nazi troops marching all over Europe were ciphering their important communication with an electro-mechanical rotor machine, called Enigma. Cracking the daily changing encoding of messages through Enigma was not only a substantial aid to the Allies during the war, it also plays a big role in the birth of information science. This was the first time that major resources were dedicated to developing and engineering electromechanical devices that were able to compute in a powerful way. One of the people responsible for this major effort was Alan Turing.

Before continuing to tell the history of computers and how quickly information sciences developed after World War II, let us go a little further back in time. Nearly exactly a hundred years earlier, the first complex computer program was written, when machines that were able to execute it were still futuristic and the amount of resources necessary to build such an device were immense. Augusta Ada King, Countess of Lovelace (Ada Lovelace for short) wrote an algorithm that calculates a series of Bernoulli numbers for a machine, that at this point only existed on paper. This machine was called the Analytical Engine and it was proposed by Charles Babbage in 1837. His designs are the first known plans of an device that fulfills a property that we nowadays call Turing-completeness.

Back in the midst of a horrible war, the people in the United Kingdom Government Code and Cypher School at Bletchley Park did not know about Babbage's design nor Lovelace's algorithm. They constructed their own devices based on the first mathematical description of a computer introduced by Alan Turing [15]. This model, called a Turing machine, is the mathematical description that we use in information sciences to describe computing devices. His article did not only feature the introduction of a mathematical model of computation, he also uses it to follow Gödel's footsteps. Years prior to Turing's article Gödel incompleteness theorems [16] had shattered David Hilbert's dreams of finding mathematics to be complete and consistent with. Turing, fascinated by these results, added on top of this the proof that also the *Entscheidungsproblem* was uncomputable. This proof was given in a different form by Alonzo Church at the same time. Both of the proofs rely on the notion of "effectively calculable" functions. Today we know this assumption as the Church-Turing hypothesis, in plain english it states that that a function on the natural numbers is computable if and only if it is computable by a Turing machine.

Now that we have looked into the beginnings of our understanding of computing devices let us turn to roots of our mathematical understanding what a signal is. We do not have to go far, neither in space nor time. In 1943 Turing was sent to Washington to exchange information with the United States Navy on the topic of the cryptanalytic methods they were using at Bletchley Park. There he met a young and bright scientist who was working at Bell Labs at that time. His name was Claude Shannon and the work he did in the years to come basically mark the genesis of information theory as it will be presented later on in this thesis.

2.3 Quantum Information Theory

After having had a brief excursion into the history of both quantum theory and information theory, let us look at how they started to interact with each other. This happened surprisingly late. Although they share common founders, such as for example John von Neumann, who substantially contributed to both fields, it took decades for scientists to connect them. Maybe because most of the brain power and physical resources were devoted to exploit the two new theories to the maximum. The development of new small transistors accelerated the construction ever faster and powerful computers while physicists convinced formerly hostile world powers to collaborate in the Conseil Européen pour la Recherche Nucléaire (CERN). This huge collaboration was meant to level the playing field under the premise that the possibility to construct another disastrous weapon like the nuclear bomb dormant in the depths of physics. One of the scientists working there in the late fifties was John Stewart Bell. As a hobby, as he put it [17], he was working on the foundations of quantum mechanics in his free time. He was especially interested in a 'paradox' that was brought up by Einstein, Podolski and Rosen (EPR) in 1935 [18]. The basis of what EPR found to be paradoxical is a phenomena that known as entanglement (a term coined by Schrödinger). Bell was able to follow EPR's line of thought in order to derive his famous theorem and the inequalities that it implied. At that time his work did not have an immediate impact but we will see that it plays a central role for what will become quantum information theory.

The first known person to actually interweave quantum and information was Stephen Wiesner, using the polarisation of light for his conjugate coding in 1969 [19]. With these ideas he was far ahead of his time which meant that his research was neither acknowledged nor published for over a decade. He eventually quit academic life and moved back to Israel, it is said that he visits Aharanov's group meetings from time to time [20]. The sixties did not only bring what can be regarded as the genesis article of quantum information theory, the world was also witnessing increasing tensions between the United States and the Soviet Union. This caused a stockpiling of physics graduate students as they were seen as war commodities [21].

The bubble burst in the seventies, leaving behind unemployed physicists that had earned their PhDs at the most prestigious physics departments in the country. At the this time when the New Age movement and counterculture had reached their heights, a group of young physicists founded a spirited, informal discussion group, they called the Fundamental Fysiks Group. These were wild times and the group was in midst of it. Imagine a time in which United States military intelligence was trying to use psychic phenomena, such as remote viewing or mind-reading, for military purposes. In this setting, nothing but an ordinary conservative life seemed possible, several members of the Fundamental Fysiks Group quickly became obsessed with quantum entanglement and Bell's theorem. Bell's theorem seemed to them as finally offer explanations for the paranormal phenomena and even open new ways to understand human consciousness. For a long time, nearly all publications regarding Bell's theorem and quantum entanglement came from participants of the Fundamental Fysiks Group. Furthermore, they had created an extensive, underground network that they used to circulate preprints amongst other things, not much unlike arXiv just much slower. This is how Nick Herbert a Stanford physics graduate disseminated his article on FLASH, a scheme for superluminal communication [22, 23].

Through a series of coincidences the FLASH paper ended up in the hands of Bill Wootters and

Wojciech Zurek. They published a short paper refuting Herbert's with help of the no-cloning theorem [24]. The no-cloning theorem, which states that in general an unknown quantum state cannot be copied, is the foundation for many results in quantum information theory, as for example the security of quantum cryptography just to name one. Wootters and Zurek were not the only ones to see that there was something wrong in the FLASH scheme. Dennis Dieks a Dutch physicist that had met Herbert, while he was visiting Amsterdam, independently discovered the theorem in the same year [25]. Last but not least GianCarlo Ghirardi who had been a referee of Herbert's initial submission, rejected the paper by the same argument [23].

Around the same time, the concept of a quantum computer started pop up. First, Paul Benioff introduces a quantum mechanical model for the description of computers [26] and shortly after Richard Feynmann shows that we do not know any efficient methods to simulate the dynamics of quantum systems on Turing machines [27]. This was independently done a little earlier by Yuri Manin in the Soviet Union [28]. These results meant that there was the need for a more general computational model which also allows to efficiently simulate quantum behaviour. This was done in less then five years by David Deutsch, who introduced the first model of an universal quantum computer [29].

In the beginning of the nineties Artur Ekert caught the attention of the physics and informatics community with the introduction of an entanglement-based secure communication scheme [30]. This afterwards resulted in attention for Charles Bennett's and Gilles Brassard's work on quantum key distribution [31] seven years earlier, that had gone mostly unnoticed. The nineties were marked by the development of a lot of the seminal results in quantum information theory, like the first quantum algorithms [32, 33, 34], the first quantum error correction schemes [35, 36], quantum annealing [37] and quantum coding results [38]. They also mark the time of the first experimental implementations, e.g. quantum logic gates [39, 40], Grover's algorithm [41] and quantum teleportation protocol [42].

With the new millennium, the development of the field of quantum information increased in tempo, and still has a lot of momentum today. So let us go into the fundamentals that allow us to discover this young field in a rigorous manner.

2.4 Preliminaries

Let us define the necessary mathematical notions in order to introduce the postulates which we will use to describe quantum theory.

Definition 2.4.1 (Hilbert space)**.** A **Hilbert space** $\mathcal{H}$ is a complex vector space with an positive definite Hermitian inner product $\langle\cdot|\cdot\rangle$, s.t. it is complete in the metric defined by the norm.

In this work we will use both finite- and infinite-dimensional Hilbert spaces. For some finite-dimensional Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$ their dimensionality will be denoted by $d_A = |A|, d_B = |B|$, respectively. Sometimes $\mathcal{H}_d$ will be written to denote a d-dimensional Hilbert space. Furthermore the space of homomorphisms $\mathcal{M} : \mathcal{H}_A \to \mathcal{H}_B$ is written as $\mathrm{Hom}(\mathcal{H}_A, \mathcal{H}_B)$.

We can associate any vector, element of some Hilbert space $\mathcal{H}_A$, with a *ket* $|\phi\rangle_A \in \text{Hom}(\mathbb{C}, \mathcal{H}_A)$, exploiting the isomorphism $\mathcal{H}_A \cong \text{Hom}(\mathbb{C}, \mathcal{H}_A)$. The *bra* $\langle\phi|_A \in \text{Hom}(\mathcal{H}_A, \mathbb{C})$ is the adjoint of the ket and is an element of the *dual space* of $\mathcal{H}_A$. This is the so-called bra-ket or Dirac notation. Moreover, $\text{End}(\mathcal{H})$ is the set of endomorphisms on $\mathcal{H}$, i.e. the homomorphisms from a Hilbert space $\mathcal{H}$ to itself.

The eigenvalues of some operator $O_A \in \text{End}(\mathcal{H})$ are denoted as $\lambda_i(O_A)$ and its singular values as $s_i(O_A)$, where the subscript specifies the system the operator acts on. For the hermitian adjoint of some operator O we use the expression $O^\dagger$, additionally the set of Hermitian operators on $\mathcal{H}$ is defined as $\text{Herm}(\mathcal{H}) := \{H \in \text{End}(\mathcal{H}) : H = H^\dagger\}$.

An important operation we will make use of a lot is the *trace* of an operator.

Definition 2.4.2 (Trace)**.** The trace of an operator is defined as

$$\text{Tr}(O_A) := \sum_i \langle e_i|O_A|e_i\rangle \tag{2.1}$$

where $\{e_i\}_i$ is any orthonormal basis of $\mathcal{H}_A$.

We will also make use of the trace norm.

Definition 2.4.3 (Trace norm, 1-norm)**.** Let $O \in \text{End}(\mathcal{H})$ then

$$||O||_1 := \text{Tr}\sqrt{O^\dagger O}. \tag{2.2}$$

The identity operator denoted by $\mathbb{1} \in \text{End}(\mathcal{H})$ maps any vector $|\phi\rangle \in \mathcal{H}$ to itself. It can be written as

$$\mathbb{1} = \sum_i |e_i\rangle\langle e_i| \tag{2.3}$$

for any orthonormal basis $\{e_i\}_i$ of $\mathcal{H}$.
The set of positive semi-definite operators that act on $\mathcal{H}$ is denoted as $\mathcal{P}(\mathcal{H})$ and is defined as

$$\mathcal{P}(\mathcal{H}) := \{O \in \text{Herm}(\mathcal{H}) : \lambda_i(O) \geq 0 \quad \forall\ i\}. \tag{2.4}$$

2.5 Postulates of Quantum Theory

Now we have gathered all the necessary tools to introduce the postulates which we will use to describe quantum theory.

Definition 2.5.1 (Postulate 1: States)**.** The state of a quantum system is described by a positive semi-definite Hermitian operator $\rho \in \text{End}(\mathcal{H})$, called a **density matrix**.

Sometimes a shorthand notation $\rho \in \mathcal{H}$ will be used. We will differentiate between *pure* and *mixed* quantum states

Definition 2.5.2 (Pure states)**.** A quantum state is called **pure** iff

$$\rho^2 = \rho. \tag{2.5}$$

If we deal with pure states we will sometimes write just the associated ket $|\phi\rangle$ due to the fact that in the case of pure states

$$\rho_{pure} = |\phi\rangle\langle\phi|. \tag{2.6}$$

Definition 2.5.3 (Mixed states)**.** All quantum states that are not pure, are called **mixed**.

Mixed states can be written as a mixture of pure states

$$\rho_{mixed} = \sum_i p_i |\phi_i\rangle\langle\phi_i| \tag{2.7}$$

where $\{p_i\}$ is some probability distribution with $\sum_i p_i = 1$ and $0 < p_i \in \mathbb{R}$. Furthermore we will introduce a quantity called the *purity* of a quantum state

Definition 2.5.4 (Purity)**.** The **purity** of a quantum state is given by

$$\mathrm{Tr}(\rho)^2 \in \left[\frac{1}{d}, 1\right]. \tag{2.8}$$

The lower bound for the purity is attained for the so called *maximally mixed* state.

Definition 2.5.5 (Maximally mixed state)**.** The state

$$\rho_{max} := \frac{1}{d}\mathbb{1} \tag{2.9}$$

is called **maximally mixed** state.

Definition 2.5.6 (Postulate 2: Observables)**. Observables** corresponding to physical properties of a quantum systems are described by linear operators $O \in \mathrm{Herm}(\mathcal{H})$, where each eigenvalue $\lambda_i(O)$ represents a possible value.

Due to the Hermiticity of the observables the eigenvalue decomposition can be employed, i.e.

$$O = \sum_i \lambda_i P_{\lambda_i} \tag{2.10}$$

where λ_i is an eigenvalue of O and P_{λ_i} is the projector onto the subspace of the respective eigenvalue.

Definition 2.5.7 (Postulate 3: Measurements)**.** Measuring a quantum system in state ρ with an observable O will give the outcome k with probability

$$p_k = \mathrm{Tr}(P_{\lambda_k}\rho). \tag{2.11}$$

Conditioned on the observation of outcome k the post-measurement state is given by

$$\rho_k := \frac{1}{p_k} P_{\lambda_k} \rho P_{\lambda_k}^\dagger. \tag{2.12}$$

This postulate is also known as **Born rule**. Next we need to define the notion of an *expectation value* in the context of quantum theory

Definition 2.5.8 (Expectation value)**.** The **expectation value** of measuring a quantum system in state ρ with an observable O is given by

$$\langle O \rangle_\rho := \mathrm{Tr} O\rho. \tag{2.13}$$

Sometimes the shorthand notation $\langle O \rangle$ will be used instead of $\langle O \rangle_\rho$. To continue with our endeavour to describe quantum theory with a small set of postulates we need to define a special kind of observable, specifically the one corresponding to the energy of a system, called *Hamiltonian*.

Definition 2.5.9 (Hamiltonian)**.** The observable corresponding to the energy of a quantum system is called **Hamiltonian**.

Before we define the postulate on the evolution of quantum states let us take a look at its origins, namely the Schrödinger equation

Definition 2.5.10 (Time-dependent Schrödinger equation)**.** The **Schrödinger equation** describes the time evolution of the state of a closed quantum system,

$$i\hbar \frac{d\,|\psi\rangle}{dt} = H(t)\,|\psi\rangle\,. \tag{2.14}$$

Its stationary version makes the concept of the Hamiltonian as the observable corresponding to energy clearer.

Definition 2.5.11 (Stationary Schrödinger equation)**.** The **time-independent Schrödinger equation** states that

$$H\,|\psi\rangle = E\,|\psi\rangle \tag{2.15}$$

where E is a constant equal to the total energy of the system.

As we have stated all our axioms in terms of density matrices so far with have to do the same for their evolution.

Definition 2.5.12 (Postulate 4: Evolution)**.** The evolution of a quantum state ρ describing a closed system with a Hamiltonian H is given by

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}\,[H, \rho]\,. \tag{2.16}$$

This postulate is also known as the **Liouville-von-Neumann equation** and given a time-dependent state $\rho(t)$[1] the solution is

$$\rho(t) = e^{\frac{i}{\hbar}Ht}\rho(0)e^{-\frac{i}{\hbar}Ht}. \tag{2.17}$$

To go into greater detail regarding the evolution of quantum states let us have a look at a class of operators called *unitaries*.

[1]The picture we have chosen here, namely to have time-dependent states and time-independent observables, is called the **Schrödinger picture**. One may also choose it the other way around which would lead to the **Heisenberg picture**. A description where both states and observables evolve in time is called **Interaction picture** or **Dirac picture**.

Definition 2.5.13 (Unitarity)**.** An operator U is called **unitary** iff

$$UU^\dagger = \mathbb{1} \tag{2.18}$$

or equivalently

$$U = U^{-1}. \tag{2.19}$$

As the Hamiltonian H is Hermitian and t can be seen to be just a scalar[2] we have that the following holds

Lemma 2.5.14. *Let H be Hermitian and t be scalar, then*

$$U = e^{iHt} \tag{2.20}$$

is unitary.

Proof. Due to the Hermiticity of H we have that

$$\left(e^{iHt}\right)^\dagger = e^{-iHt}. \tag{2.21}$$

We continue with

$$e^{iHt}\left(e^{iHt}\right)^\dagger = e^{iHt-iHt} = \mathbb{1} \tag{2.22}$$

where the first equality holds due to the Baker-Campbell-Hausdorff formula for exponential operators (for proof see [43]) and the fact that H commutes with itself. By definition 2.5.13, this concludes the proof. □

Now that we have seen that evolution of quantum states generated by a Hamiltonian H is actually always unitary we could also restate our postulate, in a slightly weaker version, without involving the notion of time explicitly.

Definition 2.5.15 (Postulate 4': Evolution revisited)**.** The evolution of a quantum state ρ to a quantum state $\tilde{\rho}$ is given by

$$\tilde{\rho} = U\rho U^\dagger \tag{2.23}$$

where U is an unitary operator.

Since we have so far only talked about single quantum systems the last postulate will tell us how to compose systems, such that we will be able to describe large multipartite states.

Definition 2.5.16 (Postulate 5: Composition)**.** Let A and B be two quantum systems with associated Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, then **composite system** is described by a density operator $\rho \in \mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_B)$.

Sometimes the shorthand notation $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ will be used.

[2]The problem of the notion of time in quantum theory is a very complex one that will be discussed in great detail in chapter 7.

2.6 Further Definitions

Now that we have seen the postulates of quantum theory that allow us to describe quantum system let make some further definitions that will prove to be useful as we go along our journey into the more technical depths of the field.
Let $\mathcal{S}_=(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : \mathrm{Tr}(\rho) = 1\}$ be the set of all normalised states and $\mathcal{S}_\leq(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : 0 < \mathrm{Tr}(\rho) \leq 1\}$ be the set quantum states including subnormalised states acting on $\mathcal{H}$. If either $\mathcal{S}_=$ or $\mathcal{S}_\leq$ is meant $\mathcal{S}$ will be written.

In order to quantify the distance between quantum states we will use a quantity called *fidelity*

Definition 2.6.1 (Fidelity)**.** Let ρ and σ be two quantum states then the **fidelity** is defined as

$$F(\rho, \sigma) := \left|\left|\sqrt{\rho}\sqrt{\sigma}\right|\right|_1 . \tag{2.24}$$

If however the states we dealing with are subnormalised we need to employ the **generalised fidelity** as a measure of distance

Definition 2.6.2 (Generalised fidelity)**.** Let $\rho, \sigma \in \mathcal{S}_\leq$ then the **generalised fidelity** is defined as

$$F(\rho, \sigma) := \left|\left|\sqrt{\rho}\sqrt{\sigma}\right|\right|_1 + \sqrt{(1 - \mathrm{Tr}\rho)(1 - \mathrm{Tr}\sigma)} \tag{2.25}$$

In the case that either ρ or σ is normalised the above expression reduces to 2.6.1. Let us continue with defining properties of linear maps.

Definition 2.6.3 (Positivity)**.** A linear map $\mathcal{E} \in \mathrm{Hom}(\mathrm{End}(\mathcal{H}_A), \mathrm{End}(\mathcal{H}_B))$ is said to be **positive** iff

$$\mathcal{E}(\rho) \geq 0 \quad \forall \quad \rho \in \mathcal{S}. \tag{2.26}$$

Definition 2.6.4 (Complete positivity)**.** A linear map $\mathcal{E} \in \mathrm{Hom}(\mathrm{End}(\mathcal{H}_A), \mathrm{End}(\mathcal{H}_B))$ is said to be **completely positive** iff

$$\mathcal{E} \otimes \mathbb{1}_C \geq 0 \quad \forall \quad \rho \in \mathcal{S} \tag{2.27}$$

for any Hilbert space $\mathcal{H}_C$.

Definition 2.6.5 (Trace preserving)**.** A linear map $\mathcal{E} \in \mathrm{Hom}(\mathrm{End}(\mathcal{H}_A), \mathrm{End}(\mathcal{H}_B))$ is said to be **trace preserving** iff

$$\mathrm{Tr}\left(\mathcal{E}(\rho)\right) = \mathrm{Tr}\left(\rho\right) \quad \forall \quad \rho \in \mathcal{S} \tag{2.28}$$

Definition 2.6.6 (Trace non-increasing)**.** A linear map $\mathcal{E} \in \mathrm{Hom}(\mathrm{End}(\mathcal{H}_A), \mathrm{End}(\mathcal{H}_B))$ is said to be **trace non-increasing** iff

$$\mathrm{Tr}\left(\mathcal{E}(\rho)\right) \geq \mathrm{Tr}\left(\rho\right) \quad \forall \quad \rho \in \mathcal{S}. \tag{2.29}$$

Also we will make use of the partial trace

Definition 2.6.7 (Partial trace)**.** The partial trace of an operator is defined as

$$\mathrm{Tr}_B(O_{AB}) := \sum_i \langle i_B | O_{AB} | i_B \rangle \tag{2.30}$$

where $\{i\}_B$ is any orthonormal basis of $\mathcal{H}_B$.

Applied here, this means that $\rho_A \equiv \mathrm{Tr}_B\, \rho_{AB}$ is the reduced state on system A.

In the following, we will work in units such that $c = \hbar = k_B = 1$, if not stated otherwise.

2.7 Entropies

In this section quantities of information will be discussed in the classical and in the quantum case. Also results on the possible extension of the quantum conditional mutual information to a non-asymptotic framework will be presented.

2.7.1 Shannon entropies

Shannon's groundbreaking work of establishing information theory in a single stroke is intimately linked to introduction of what today we call the *Shannon entropy*. Given a random variable X the Shannon entropy quantifies the information gain achieved by learning the value of X. Formally it is mostly defined as a function of a probability distribution p_X.

Definition 2.7.1 (Shannon entropy)**.** Given a probability distribution p_X the Shannon entropy is defined as

$$H(X)_p = H(p_X) := -\sum_{x\in\mathcal{X}} p(x)\log(p(x)). \tag{2.31}$$

If not otherwise stated log will refer to the logarithm of base two. [3] In the case of dichotomic random variables and the corresponding binary probability distributions the Shannon entropy solely depends on a single parameter $q := p(0)$. This motivates the definition of the *binary entropy*.

Definition 2.7.2 (Binary entropy)**.** Given a binary probability distribution $p : \{0, 1\} \to [\,0, 1]$ the binary entropy is defined as

$$H(q) := -q\log(q) - (1-q)\log(1-q). \tag{2.32}$$

Continuing the discussion of entropy let us turn to cases where we have more than just one random variable.

Definition 2.7.3 (Joint entropy)**.** Given a joint probability distribution p_{XY} the joint entropy is defined as

$$H(XY)_p := -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log(p(x,y)). \tag{2.33}$$

The joint entropy quantifies the total uncertainty of the random variables X and Y. Asking the question on how much uncertainty remains about a random X conditioned on knowing the value of Y leads to the definition of the *conditional entropy*.

Definition 2.7.4 (Conditional entropy I)**.** Given a joint probability distribution p_{XY} the conditional entropy is defined as

$$H(X|Y)_p := H(XY)_p - H(Y)_p. \tag{2.34}$$

[3] Note that here $0\log(0) = 0$.

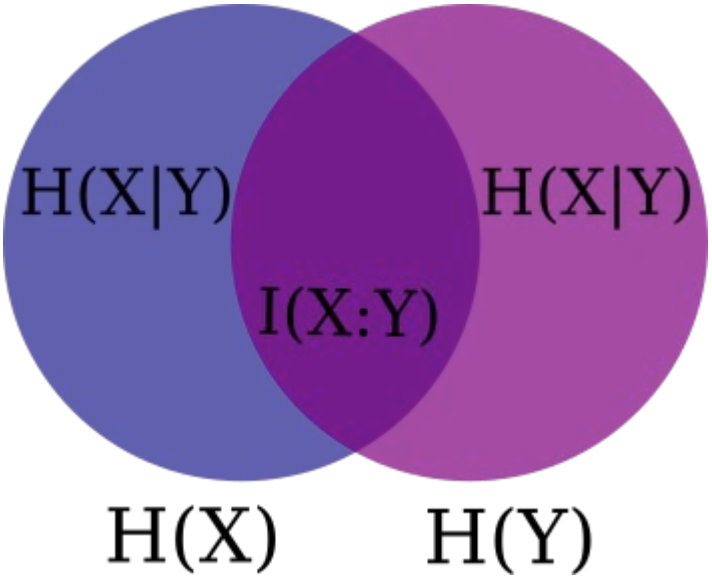

Figure 2.1: Venn diagram of the Shannon entropies of a bipartite system [44].

One can also define the conditional entropy by defining what a conditional probability distribution is. We denote $p_{X|Y}(x|y) := \frac{p_{XY}(x,y)}{p_Y(y)}$ the conditional distribution of X conditioned on knowing Y, where p_Y is the Y-marginal of p_{XY}.

Definition 2.7.5 (Conditional entropy II)**.** Given a joint probability distribution p_{XY} the conditional entropy is defined as

$$H(X|Y)_p := \sum_{y \in \mathcal{Y}} p_Y(y) H(p_{X|Y}) \tag{2.35}$$

The *mutual information* is a measure that quantifies how much X and Y have in common.

Definition 2.7.6 (Mutual Information)**.** Given a joint probability distribution p_{XY} the mutual information is defined as

$$I(X:Y)_p := H(X)_p + H(Y)_p - H(XY)_p. \tag{2.36}$$

Using definition 2.7.4 one can see that $I(X:Y)_p = H(X)_p - H(X|Y)_p$, elucidating the connection between the mutual information and the conditional entropy. The mutual information can also be defined for the case where one conditions on knowing a third variable Z.

Definition 2.7.7 (Conditional Mutual Information)**.** Given a joint probability distribution p_{XYZ} the conditional mutual information is defined as

$$I(X:Y|Z)_p := H(X|Z)_p + H(Y|Z)_p - H(XY|Z)_p. \tag{2.37}$$

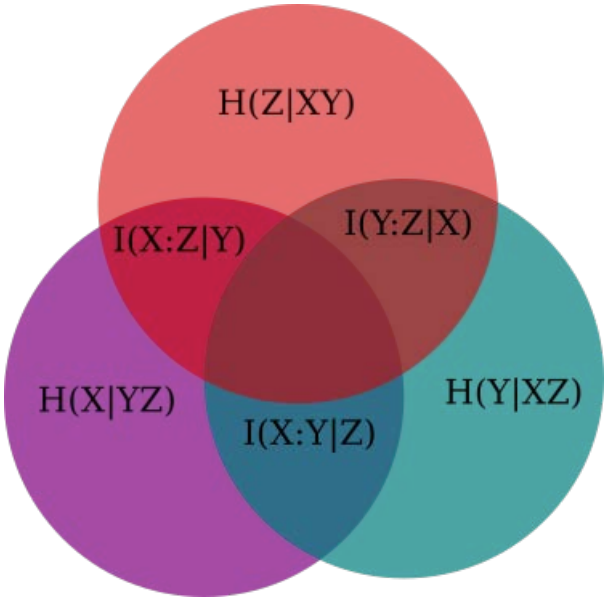

Figure 2.2: Venn diagram of the conditional Shannon entropies of a tripartite system [44].

2.7.2 Von Neumann entropies

One can generalise the definitions given above for the classcial case to the quantum case. While so far we have used probability distributions we will now start to define this quantities as functions of the density operator (see definition 2.5.1). Although the Shannon entropy as introduced above can be seen as a special case of Von Neumann's entropy that we are going to introduce now, it came nearly two decades later [45, 14].

Definition 2.7.8 (von Neumann entropy)**.** The Von Neumann entropy is defined as

$$H(\rho) := -\mathrm{Tr}\rho \log \rho. \tag{2.38}$$

The Von Neumann entropy bears the same symbol as the Shannon entropy, namely H. This can be a little confusing, is justified though. The Von Neumann entropy can be expressed as Shannon entropy of the spectrum of the density operator ρ, i.e.

$$H(\rho) = -\sum_i \lambda_i \log \lambda_i \tag{2.39}$$

where the λ_i are the eigenvalues of ρ. This means that we can encode any probability distributions into the diagonal elements of the density operator. The Von Neumann entropy therefore is a genuine generalisation of the Shannon entropy. For the sake of clarity and comprehensibility in the following we will use the first letters of the alphabet $A, B, C, \ldots$ in order to label quantum systems and the last letters $\ldots, X, Y, Z$ if we speak of strictly classical systems. Moreover, $H(A)_\rho := H(\rho_A)$ and $H(AB)_\rho := H(\rho_{AB})$ will be the convention we will stick to in the following.

We can also define a quantum version of the conditional entropy in complete analogy to the classical case.

Definition 2.7.9 (Conditional von Neumann entropy)**.** Given a bipartite quantum state ρ_{AB} the conditional Von Neumann entropy is defined as

$$H(A|B)_\rho := H(AB)_\rho - H(B)_\rho. \tag{2.40}$$

While the classical conditional entropy $H(X|Y)_p$ is a strictly non-negative quantity its quantum generalisation is not. We will see later on in this thesis that this property can be used in order to detect entanglement.

Mutual information measures are widely used to characterise the correlations in quantum many-body systems as well as in classical systems.

Definition 2.7.10 (von Neumann Mutual Information)**.** Given a bipartite quantum state ρ_{AB} the von Neumann mutual information is defined as

$$I(A:B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho. \tag{2.41}$$

Posing the question on how much can be deduced from a system about a second system conditioned on knowing the state of a third system leads to the definition of the conditional mutual information. In the setting where one is provided with i.i.d. resources this is a well studied quantity and it has been shown that for the case of quantum state redistribution it characterises the optimal amount of qubits needed in order to accomplish this task [46]. Furthermore it is the basis for the squashed entanglement, an important quantity in entanglement theory [47]

Definition 2.7.11 (Quantum Conditional Mutual Information)**.** Given a tripartite quantum state ρ_{ABC} the quantum conditional mutual information is defined as

$$I(A:B|C)_\rho := H(AC)_\rho + H(BC)_\rho - H(C)_\rho - H(ABC)_\rho. \tag{2.42}$$

Markov states in the context of quantum information theory are defined via the quantum conditional mutual information.

Definition 2.7.12 (Markov states)**.** The set of all Markov states is defined as

$$\mathcal{M} := \{\sigma \in \mathcal{D}(\mathcal{H}) : I(A:B|C)_\sigma = 0\}. \tag{2.43}$$

The condition that σ fulfills strong subadditivity with equality (i.e. $I(A:B|C)_\sigma = 0$) is equivalent to the statement that there exists a decomposition of system C as

$$\mathcal{H}_C = \bigoplus_i \mathcal{H}_{C_i^L} \otimes \mathcal{H}_{C_i^R} \tag{2.44}$$

into a direct sum of tensor products such that

$$\sigma_{ABC} = \bigoplus_i p_i \sigma_{AC_i^L} \otimes \sigma_{C_i^L B}\,, \tag{2.45}$$

where $\{p_i\}$ is some probability distribution (see Ref. [48]).

Many of the well-known quantities of information in the asymptotic framework of quantum information can be written in terms of the relative entropies.

Definition 2.7.13 (Quantum relative entropy)**.** Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\sigma \in \mathcal{S}(\mathcal{H})$, then the quantum relative entropy is defined as

$$D_1(\rho\|\sigma) := \mathrm{Tr}\left[\rho\left(\log\rho - \log\sigma\right)\right]. \tag{2.46}$$

Recalling the definitions for the von Neumann entropy $H(A)_\rho$, the conditional von Neumann entropy $H(A|B)_\rho$ and the von Neumann mutual information $I(A:B)_\rho$, it is straightforward to check that

$$\begin{aligned} H(A)_\rho &= -\mathrm{Tr}\,\rho\log\rho && (2.47)\\ &= -D_1(\rho_A\|\mathbb{1}_A), && (2.48) \end{aligned}$$

$$\begin{aligned} H(A|B)_\rho &= H(AB)_\rho - H(B)_\rho && (2.49)\\ &= -\min_{\sigma_B} D_1(\rho_{AB}\|\mathbb{1}_A\otimes\sigma_B) && (2.50) \end{aligned}$$

and

$$\begin{aligned} I(A:B)_\rho &= H(A)_\rho + H(B)_\rho - H(AB)_\rho && (2.51)\\ &= \min_{\sigma_B} D_1(\rho_{AB}\|\rho_A\otimes\sigma_B). && (2.52) \end{aligned}$$

2.7.3 Rényi entropies

A quantity that is used to generalise quantities of information to a situation where i.i.d. (identically and independently distributed) resources are not available is the α-Rényi relative entropy [49, 50, 51]. Information quantities in terms quantum Rényi relative entropies have found applications in various operational tasks [53, 54, 55]. Note that the α-Rényi relative entropy in form it will be defined below actually goes back to Petz [52], we will still drop his name after the definition for the sake of simplicity.

Definition 2.7.14 ((Petz') Quantum α-Rényi relative entropy)**.** Let $\alpha \in (0,1)\cup(1,\infty)$ and let $\rho,\sigma \in \mathcal{S}_{\leq}(\mathcal{H})$ with $supp\,\rho \subseteq supp\,\sigma$. Then the (Petz') quantum Rényi relative entropy of order α is defined as

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\mathrm{Tr}\left(\rho^\alpha\sigma^{1-\alpha}\right). \tag{2.53}$$

Note that in the limit $\alpha\to 1$ we recover the relative entropy (2.46) and that D_α is a monotonically increasing function in α. Moreover we define the 0-relative entropy ($D_{\min}$ in Ref. [49]), which naturally appears in binary hypothesis testing when the probability for type I errors is set to zero.

Definition 2.7.15 (Quantum 0-relative entropy)**.** Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\sigma \in \mathcal{P}(\mathcal{H})$, then the quantum 0-relative entropy is defined as

$$D_0(\rho\|\sigma) := \lim_{\alpha\to 0^+} D_\alpha(\rho\|\sigma) = -\log\mathrm{Tr}\,P_\rho\,\sigma, \tag{2.54}$$

where P_ρ denotes the projector onto the support of ρ.

On the other end of the spectrum we have the quantum relative max-entropy, it corresponds to the case $\alpha \to \infty$ when defining the α-Rényi relative entropy in a slightly different way [56].

Definition 2.7.16 (Quantum relative max-entropy)**.** Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\sigma \in \mathcal{P}(\mathcal{H})$, then the max relative entropy is defined as

$$D_{\max}(\rho\|\sigma) := \inf\{\lambda \in \mathbb{R} : \rho \leq e^{\lambda}\sigma\}\,. \tag{2.55}$$

One can now use the Rényi relative entropies for the definition of entropies for $\alpha \neq 1$, analogously to the Von Neumann case.

Definition 2.7.17 (Quantum α-Rényi entropies)**.** Let $\alpha \in (0,1) \cup (1,\infty)$ and let $\rho_A \in \mathcal{S}_{\leq}(\mathcal{H}_A)$. Then the quantum Rényi entropy of order α is defined as

$$H_\alpha(A)_\rho := H_\alpha(\rho_A) = D_\alpha(\rho_A\|\mathbb{1}_A)\,. \tag{2.56}$$

Considering bipartite system we can define the conditional versions of the quantum α-Rényi entropy.

Definition 2.7.18 (Quantum conditional α-Rényi entropies)**.** Let $\alpha \in (0,1) \cup (1,\infty)$ and let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$. Then the quantum conditional Rényi entropy of order α is defined as

$$H_\alpha(A|B)_\rho := \sup_{\sigma_B} -D_\alpha(\rho_A B\|\mathbb{1}_A \otimes \sigma_B)\,. \tag{2.57}$$

As for the case of the quantum conditional 0/max-entropies we have that one is defined by the other's relative entropy.

Definition 2.7.19 (Quantum conditional 0/max-entropies)**.** Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$, then the quantum min/max-entropy is defined as

$$H_{0/\max}(A|B)_\rho := \sup_{\sigma_B} -D_{\max/0}(\rho_A B\|\mathbb{1}_A \otimes \sigma_B)\,. \tag{2.58}$$

There exist a lot more types of entropies and divergencies that have been developed in the non-i.i.d. context and these are connected in manifold ways. This can be quite confusing and stating all different definitions and relations to each other would go way beyond the scope of this thesis. The whole zoo of entropies has been visualised in an appealing way in [57].
The quantum mutual information measures in the non-asymptotic setting for the different values of α are given in the following way.

Definition 2.7.20 (Quantum α-Rényi mutual information)**.** Let $\alpha \in (0,1) \cup (1,\infty)$ and let $\rho_{AB} \in \mathcal{S}(\mathcal{H})$. Then the quantum Rényi mutual information of order α is defined as

$$I_\alpha(A:B) := \inf_{\sigma_B} D_\alpha(\rho_{AB}\|\rho_A \otimes \sigma_B)\,. \tag{2.59}$$

Here, the analogies stop, as it is becomes a non-trivial questions how to generalise the quantum conditional mutual information to the non-i.i.d. setting. Here the recent years have brought a lot of insights to these questions [58, 59].

2.7.4 How not to

At the time of writing [7] it was still an open question how to generalise the quantum conditional mutual information to the non-i.i.d. setting. In this subsection we will follow the proof of above cited article in order to show how not to do it. To follow the proof we need to introduce some concepts and definitions. The proof will be done by considering the set of all Markov states i.e. states for which the quantum conditional mutual information is zero. An approach where one minimises the quantum α-Rényi relative entropies over this set may seem reasonable as it can be shown that for the case of D_{min} it leads to a non-negative quantity that fulfills a data-processing inequality and a duality relation, as one would expect from a generalisation of the quantum conditional mutual information. However this approach fails as will be shown now.

This failure arrises from the fact that in the case of the totally antisymmetric state this approach leads to a constant lower bound for the whole range of the parameter α while the quantum conditional mutual information goes to zero for large dimensions (with $\mathcal{O}(\frac{1}{d})$). Thus this approach cannot produce a quantity that in general lower bounds the quantum conditional mutual information.
In order to present the promised results we first need the definition of the totally antisymmetric state.

Definition 2.7.21 (Totally antisymmetric state)**.** The totally antisymmetric state as is defined as

$$\gamma_d := \frac{2}{d(d-1)} P_\Lambda \tag{2.60}$$

where P_Λ denotes the projector onto the antisymmetric subspace $\Lambda^2(\mathbb{C}^d)$ in $(\mathbb{C}^d)^{\otimes 2}$.

In Ref. [60] the authors show that the quantity

$$\Delta(\rho_{ABC}) := \inf_{\sigma_{ABC}\in\mathcal{M}} D_1(\rho_{ABC}\|\sigma_{ABC}) \tag{2.61}$$

poses an upper bound to the quantum conditional mutual information, i.e.

$$I(A:B|C)_\rho \leq \Delta(\rho_{ABC}) \tag{2.62}$$

and there exist certain states for which the inequality becomes an equality. Along these lines we will show that it is not possible to generalise the quantum conditional mutual information to the non-i.i.d. setting by optimising quantum Rényi relative entropies over the set of all Markov states. This will be done in the following by the construction of a specific example, where this approach leads to an upper bound of the quantum conditional mutual information for any value of α. Note that this approach may seem reasonable, as it can be shown that

$$\Delta_{\min}(\rho_{ABC}) := \inf_{\sigma_{ABC}\in\mathcal{M}} -\log F^2(\rho_{ABC}, \sigma_{ABC}) \tag{2.63}$$

shares three main properties with the aforementioned. Namely these are non-negativity, non-increasing under data processing and duality for quadripartite pure states [61].

Generalising above approach with the quantum Rényi relative entropies, we have that for $\alpha \geq 0$

$$\Delta_\alpha(\rho_{ABC}) := \inf_{\sigma_{ABC}\in\mathcal{M}} D_\alpha(\rho_{ABC}\|\sigma_{ABC}). \tag{2.64}$$

Our result now reads as follows

Theorem 2.7.22. *Let P_k be the projector onto the antisymmetric subspace $\Lambda^k(\mathbb{C}^d)$ in $(\mathbb{C}^d)^{\otimes k}$ with $\mathcal{H}_A \cong \mathcal{H}_B \cong \mathbb{C}^d$ being the first two tensor factors and let $\rho_{ABC} := \frac{P_k}{d_k}$, where $d_k := \dim \Lambda^k(\mathbb{C}^d) = \binom{d}{k}$. Then,*

$$\Delta_0(\rho_{ABC}) \geq \log \sqrt{\frac{4}{3}} \tag{2.65}$$

Proof. We start out from the definition of the left-hand side making use of the monotonicity under CPTP maps of the quantum Rényi relative relative entropies by tracing out subsystem C.

$$\begin{aligned}
\Delta_0(\rho_{ABC}) &= \inf_{\sigma_{ABC}\in\mathcal{M}} \lim_{\alpha\to 0^+} D_\alpha(\rho_{ABC}\|\sigma_{ABC}) && (2.66)\\
&\geq \inf_{\sigma_{AB}\in\mathcal{S}} \lim_{\alpha\to 0^+} D_\alpha(\gamma_d\|\sigma_{AB}) && (2.67)\\
&= \inf_{\sigma_{AB}\in\mathcal{S}} -\log \operatorname{Tr} P_{\gamma_d}\sigma_{AB} && (2.68)
\end{aligned}$$

Due to the special form of the Markov states (see eq. 2.45) we have that the optimisation in the last two lines runs over the set of all separable states $\mathcal{S} = \{\sigma_{AB} \in \mathcal{D}(\mathcal{H}_{AB}) : \sigma_{AB} = \sum_j p_j \sigma_A^{(j)} \otimes \sigma_B^{(j)}\}$ for some probability distribution $\{p_j\}$. Moreover it is clear that $\rho_{AB} = \mathrm{Tr}_C \rho_{ABC}$ equals the totally antisymmetric state γ_d in $(\mathbb{C}^d)^{\otimes 2}$.
Since the totally antisymmetric state γ_d is invariant under the action $g \otimes g$ where g is unitary, we can restrict the optimisation problem to states obeying the same symmetry. It has be shown that the expression in the last line has a constant lower bound equal to $\log \sqrt{\frac{4}{3}}$ (see proof of Corollary 3 in [62]).

□

We can now directly relate the quantum conditional mutual information to Δ_0 for the state ρ_{ABC} defined above.

Corollary 2.7.23. *Let $\rho_{ABC} = \frac{P_k}{d_k}$ as above then*

$$\Delta_0(\rho_{ABC}) > I(A:B|C)_{\rho_{ABC}}. \tag{2.69}$$

for $k = \lceil \frac{d+1}{2} \rceil$ and $d \geq 27$.

Proof. An explicit evalution of the right-hand side (see Ref. [62]) gives us

$$I(A:B|C)_{\rho_{ABC}} = \begin{cases} 2\log \frac{d+2}{d} & \text{if } d \text{ is even} \\ \log \frac{d+3}{d-1} & \text{if } d \text{ is odd.} \end{cases} \tag{2.70}$$

Finally a simple numerical calculation shows that

$$\log \sqrt{\frac{4}{3}} > \begin{cases} 2\log \frac{d+2}{d} & \text{if } d \text{ is even} \\ \log \frac{d+3}{d-1} & \text{if } d \text{ is odd} \end{cases} \tag{2.71}$$

is the case for $d \geq 27$. □

This implies that the relative gap between Δ_0 and the quantum conditional mutual information can be made arbitrarily big by scaling up the dimension, as

$$\Delta_0(\rho_{ABC}) \geq const. \tag{2.72}$$

while

$$I(A:B|C)_{\rho_{ABC}} \leq \frac{4}{d-1} = \mathcal{O}\left(\frac{1}{d}\right) \tag{2.73}$$

when $k = \lceil \frac{d+1}{2} \rceil$. Observing that the quantum Rényi relative entropies monotonically increase in α it follows that

Corollary 2.7.24. *Let $\rho_{ABC} = \frac{P_k}{d_k}$ as above and $\alpha > 0$, then*

$$\Delta_\alpha(\rho_{ABC}) > I(A:B|C)_{\rho_{ABC}}. \tag{2.74}$$

for $k = \lceil \frac{d+1}{2} \rceil$ and $d \geq 27$.

As in the case of Rényi generalisations of some quantities of information the property of convergence to their von Neumann equivalents in the i.i.d. -limit emerges from an additional limit taken in the smoothing parameter, we show here that our results above also hold in a regularised and smoothed version.
First we need to define a metric on $\mathcal{S}_\leq$, which we choose to be the purified distance (introduced in Ref. [53]).

Definition 2.7.25 (Purified distance)**.** The purified distance is defined as

$$P(\rho,\sigma) := \sqrt{1 - F^2(\rho,\sigma)}. \tag{2.75}$$

Furthermore two states ρ and σ will be called ϵ-close if and only if $P(\rho,\sigma) \leq \epsilon$. This leads to the definition of the ball of ϵ-close states for given state ρ.

Definition 2.7.26 (ϵ-ball)**.** The ball of ϵ-close states around $\rho \in \mathcal{S}_\leq$ is defined as

$$\mathcal{B}^\epsilon(\rho) := \{\rho' \in \mathcal{S}_\leq : P(\rho,\rho') \leq \epsilon\}. \tag{2.76}$$

Now we can define the smoothed version of the Δ_0-quantity

$$\Delta_0^\epsilon(\rho) := \max_{\tilde{\rho} \in \mathcal{B}^\epsilon(\rho)} \Delta_0(\tilde{\rho}) \tag{2.77}$$

One can immediately observe that

$$\Delta_0^\epsilon(\rho) \geq \Delta_0(\rho) \quad \forall\, 0 \leq \epsilon < 1 \tag{2.78}$$

holds. As for the case of regularisation we need to generalise our statements to the case where number of available copies of the given state goes to infinity. Therefore we define the regularised version of Δ_0, i.e.

$$\Delta_0^\infty(\rho) := \lim_{n\to\infty} \frac{1}{n} \Delta_0(\rho^{\otimes n}) = \lim_{n\to\infty} \frac{1}{n} \inf_{\sigma\in\mathcal{M}} D_0(\rho^{\otimes n} \| \sigma). \tag{2.79}$$

Thus we need a statement more general than Theorem 1, which reads as follows;

Theorem 2.7.27. *Let* $\rho_{ABC} = \frac{P_k}{d_k}$ *as above then*

$$\Delta_0(\rho_{ABC}^{\otimes n}) \geq n \log \sqrt{\frac{4}{3}}. \tag{2.80}$$

Proof. Tracing out system C we have that

$$\begin{aligned}
\Delta_0(\rho_{ABC}^{\otimes n}) &= \inf_{\sigma_{ABC} \in \mathcal{M}} \lim_{\alpha \to 0^+} D_\alpha(\rho_{ABC}^{\otimes n} \| \sigma_{ABC}) && (2.81)\\
&\geq \inf_{\sigma_{AB} \in \mathcal{S}} \lim_{\alpha \to 0^+} D_\alpha(\gamma_d^{\otimes n} \| \sigma_{AB}) && (2.82)\\
&= \inf_{\sigma_{AB} \in \mathcal{S}} -\log \text{ Tr } P_{\gamma_d}^{\otimes n} \sigma_{AB} && (2.83)
\end{aligned}$$

The observation that the last line has a lower bound given by $n \log \sqrt{\frac{4}{3}}$ (see Lemma 9 and 12 in Ref. [62]) concludes the proof. □

Putting together the observations Theorem 2.7.27 and Equation (2.78) we can conclude that the results achieved for Δ_0 also hold for its smoothed and regularised version, hence establishing

$$\Delta_0^{\infty,\epsilon}(\rho_{ABC}) := \lim_{n \to \infty} \frac{1}{n} \Delta_0^\epsilon(\rho^{\otimes n}) \geq const. \quad \forall\, 0 \leq \epsilon < 1. \tag{2.84}$$

Chapter 3

Bases

As far as we know quantum theory provides the most accurate description of the world as we perceive it with the help of the most advanced technological measurement apparati. In the last chapter, we have provided the necessary notation and definitions in order to understand what we will discuss in the following. We have seen that already for stating the first postulate of quantum theory 2.5.1 we need the concept of a basis for our underlying Hilbert space. The basis that we use to describe a state (e.g. computational,...) is the connection to the physical world, in the sense that as physicists we have to find the correspondence between the mathematical object (i.e. a set of vectors or operators) and the observable phenomenon (e.g. position/momentum, direction of spin, polarisation,...). Apart from this, the right choice of basis can also save us a lot of time or effort when trying to solve a problem. There are bases that have properties that are especially useful and some come in very handy when parametrising states. These topics will be discussed in this chapter. Furthermore we will introduce a new basis with advantageous properties [8].

3.1 Bloch representations

The Bloch representation is an essential tool for the analysis of characteristic features of quantum systems. Initially introduced by Bloch [63] for qubit systems, it has since found its way in the heart of quantum information theory and quantum theory lectures. It gives (at least for the qubit case) a very comprehensive way to understand the phase space of qubit systems in a geometrical fashion, as we will see in the following. Aside the pedagogical upsides it has been adopted as a technical tool in a wide range of settings (see e.g. [64, 65, 66]). Also it has been generalised for multipartite and higher dimensional systems.

As the Bloch representation uses the expectation values of a complete set of measurements in order to define the density operator, it can help in finding solutions for Hamiltonian evolutions [67] as well as in entanglement theory [68, 69, 70, 71, 72, 73, 74], which will be discussed in the next chapter.

Simply put, the Bloch representation is a decomposition of the density matrix with help of a complete operator basis. For the case of two-level systems the Pauli matrices are the generic choice.

Definition 3.1.1 (Pauli matrices)**.** The Pauli Matrices, including the Identity σ_0, are defined as

$$\sigma_0 := \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.1}$$

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.2}$$

$$\sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{3.3}$$

$$\sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.4}$$

$$\tag{3.5}$$

In Dirac notation we can express them as

$$\sigma_0 = |0\rangle\langle 0| + |1\rangle\langle 1| \tag{3.6}$$

$$\sigma_1 = |0\rangle\langle 1| + |1\rangle\langle 0| \tag{3.7}$$

$$\sigma_2 = -i\,|0\rangle\langle 1| + i\,|1\rangle\langle 0| \tag{3.8}$$

$$\sigma_3 = |0\rangle\langle 0| - |1\rangle\langle 1| \tag{3.9}$$

$$\tag{3.10}$$

Sometimes the Pauli matrices are labelled with x, y, z instead of numbers. Together with the identity σ_0 they form a complete basis for all complex self-adjoint 2×2 matrices. Therefore we can write down any qubit state in terms of the Pauli matrices, i.e. its Bloch decomposition. For qubit states this leads the following very beautiful single-letter formula.

Definition 3.1.2 (Bloch decomposition)**.** Let ρ be a quantum state of dimension $d = 2$, i.e. a qubit, then we call

$$\rho = \frac{1}{2}\left(\mathbb{1} + \vec{n}\,\vec{\sigma}\right), \tag{3.11}$$

the Bloch decomposition of ρ, where the Bloch vector $\vec{n} \in \mathbb{R}^3$ fullfills $|\vec{n}| \leq 1$. Moreover, $\vec{\sigma}$ is the vector with the Pauli matrices as its elements

$$\vec{\sigma} := \begin{pmatrix} \sigma_1 \\ \sigma_3 \\ \sigma_3 \end{pmatrix}. \tag{3.12}$$

From this beautiful parametrisation of the qubit state space one directly gets a intuitive geometrical picture. Namely, all possible Bloch vectors $\vec{n}$ form a ball with the maximally mixed state $\rho_{\vec{0}} = \frac{\mathbb{1}}{2}$ in its origin. The Bloch sphere, formed by $\vec{n}$ with unit length, contains all pure qubit states. This is represented in figure 5.1.

One can take this approach further and generalise it to qudits (quantum systems of dimension d).

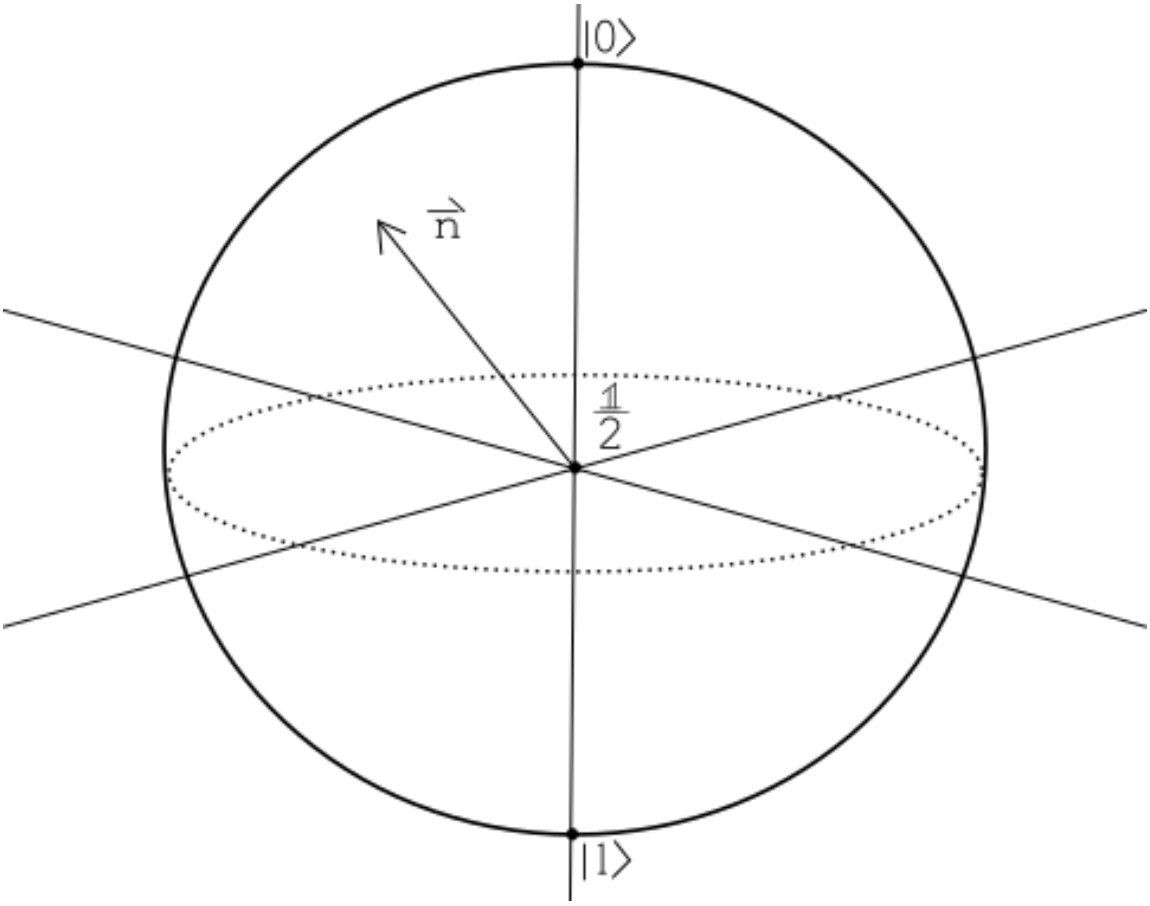

Figure 3.1: Bloch picture: The state space of all possible qubit states can be represented by the Bloch ball, which is formed by all possible Bloch vectors $\vec{n}$. [44]

For this purpose, one needs a complete basis for all density matrices acting on a d-dimensional Hilbert space $\mathcal{H}_d$ which furthermore has to contain elements that are orthogonal and traceless.

Definition 3.1.3 (Generalised Bloch decomposition)**.** Let ρ be a quantum state of dimension d, i.e. a qudit, then we call

$$\rho = \frac{1}{d}\left(\mathbb{1} + \sum_{i=1}^{d^2-1} n_i \mathcal{B}_i\right), \tag{3.13}$$

the generalised Bloch decomposition of ρ, where we have that the orthogonal and traceless operators $\mathcal{B}_i$ together with $\mathcal{B}_0 = \mathbb{1}$ form a complete basis for $\mathcal{S}_=(\mathcal{H}_d)$. The elements of the generalised

Bloch vector n_i are given by

$$n_i := \mathrm{Tr}\left(\rho\, \mathcal{B}_i\right). \tag{3.14}$$

Unfortunately a geometrical understanding of the state space in terms of the (d^2-1)-dimensional generalised Bloch vector (formed by the n_i) is very hard for $d \geq 2$. Already in the qutrit case, i.e. $d = 3$, we are left with very intricate structures [75, 76, 77], whose complexity grows with increasing dimension.

As can be readily checked, the Pauli matrices are unitary and Hermitian. For higher dimensional systems, complete operator bases that come with both features do not exist. Furthermore, the Bloch vector corresponding to non-Hermitian operator bases will in general not be real but also contain complex entries. Such choices, as e.g. Heisenberg-Weyl operators, have been explored in the past [65, 78, 79, 80, 81, 82]. So far, the generalised Gell-Mann matrices, generators of the special unitary group $\mathrm{SU}(d)$, are the most commonly used in such a case. While they are very convenient for some applications (e.g. representing high dimensional spin systems [66, 83]), they perform poorly at others (e.g. where basis elements with full rank are needed). This chapter will include the introduction of a recently proposed Hermitian generalisation of the Pauli matrices namely the Heisenberg-Weyl observables (HWO) [8].

3.2 Mutually unbiased bases

In the Bloch picture (see figure 5.1), taking states corresponding to antipodal points on the sphere results in a basis for the Hilbert space $\mathcal{H}_2$. Constructing another basis by taking Bloch vectors that are orthogonal to the initial ones, results in two bases that are connected by an interesting relation. Namely they are mutual unbiased, the overlaps of their elements are the minimal possible for such a case. This means that if one measures in one basis, there is minimal information about outcomes in the second basis or conversely measuring in two such basis gives maximal information about the state. Such relation between bases is not limited to only two-dimensional spaces, but can be found in any dimension. So let us now look at this very prominent and important property that bases can have, namely mutual unbiasedness. Let us directly start off with the definition of the property.

Definition 3.2.1 (Mutual Unbiased)**.** Sets of basis vectors $\{|v_i^k\rangle\}$ are called mutually unbiased, iff they are both orthonormal $\langle v_i^k|v_j^k\rangle = \delta_{ij}$ and their overlaps are unbiased $|\langle v_i^k|v_j^{k'}\rangle|^2 = \frac{1}{d}$.

Mutual unbiased bases (MUBs) can be used for an efficient tomography [84] and cryptography protocols [85]. They also have a strong connection to Quantum Random Access Codes [86]. Due to their wide range of applicability of this property it constitutes an interesting field for future research since also there are still a lot of open questions regarding MUBs in higher dimensions.

In the next chapter we will use the properties of MUBs in order to detect and quantify entanglement. It shall be remarked that it is still an unsolved problem how many MUBs exist in general (the smallest example being $d = 6$). We however know that in all prime power dimensions exactly $d + 1$ such bases exist. For a more in depth review and a discussion of further applications please see [87].

3.3 Gell-Mann matrices

The commonly known operator basis for describing qudit density matrices are the generalised Gell-Mann matrices (GMM). They constitute a Hermitian generalisation of the Pauli matrices introduced in definition 3.1.1.

Definition 3.3.1 (Generalised Gell-Mann matrices)**.** The generalised Gell-Mann matrices for dimension d are defined by three sets: The symmetric set is given by

$$\mathcal{G}_{jk}^{S} := \sqrt{\frac{d}{2}} \left(|k\rangle \langle j| + |j\rangle \langle k| \right), \tag{3.15}$$

the antisymmetric set is given by

$$\mathcal{G}_{jk}^{A} := -i\sqrt{\frac{d}{2}} \left(|k\rangle \langle j| - |j\rangle \langle k| \right) \tag{3.16}$$

and the diagonal set is given by

$$\mathcal{G}_{l}^{D} := \sqrt{\frac{d}{l(l+1)}} \left(\sum_{j=1}^{l} |j\rangle \langle j| - l \, |l+1\rangle \langle l+1| \right) \tag{3.17}$$

where $1 \leq l \leq d-1$ and $1 \leq j < k \leq d$ holds. To achieve a full orthogonal basis we add the identity operator as $\mathcal{G}_{0}^{D} := \mathbb{1}_d$ to the diagonal set.

For $d = 2$ the GMM reduce to the Pauli matrices.

3.3.1 Spectrum

We want to have a quick look at the spectrum of GMM, specifically the largest eigenvalue for the squares of the GGM as such expressions will come in handy later on. The largest eigenvalue for the squares of matrices of the three sets can be straightfoward computed as

$$\max \lambda \left((\mathcal{G}_{jk}^{S})^2 \right) = \max \lambda \left((\mathcal{G}_{jk}^{A})^2 \right) = \frac{d}{2} \tag{3.18}$$

$$\max \lambda \left((\mathcal{G}_{l}^{D})^2 \right) = \frac{dl}{l+1}. \tag{3.19}$$

Next we want to consider a matrix M that is decomposable into a tensor product whose elements are GGM. For such a matrix M we can calculate the largest eigenvalue of its square by using above equations

$$\max \lambda \left(M^2 \right) = \left(\prod_{l=1}^{d-1} \left(\frac{dl}{l+1} \right)^{\alpha_l} \right) \times \left(\frac{d}{2} \right)^{\beta+\gamma} \tag{3.20}$$

where we have that α_l, β and γ are the numbers specifying how often elements of the diagonal set, the antisymmetric set and the symmetric set appear in the tensor product that constitutes M, respectively.

3.4 Heisenberg-Weyl observables

In this section we will define the discrete phase-space displacement operators or Heisenberg-Weyl operators and then derive a new complete Hermitian operator basis from them [9].

Definition 3.4.1 (Phase-space operators)**.** We define the operator $X = e^{-i2\pi P/d}$ by the action $X|j\rangle = |j+1 \bmod d\rangle$ and the operator $Z = e^{i2\pi Q/d}$ by $Z|j\rangle = e^{i2\pi j/d}|j\rangle$, where Q and P are the discrete position and momentum operators, respectively, describing a $d \times d$ grid.

X and Z are in general non-commuting operators obeying the following relation

$$Z^l X^m = X^m Z^l e^{i2\pi lm/d}. \tag{3.21}$$

We can use these operators to define unitaries that act as discrete displacements on the grid.

Definition 3.4.2. (Displacement operators) Let $\mathcal{D}(l,m) = Z^l X^m e^{-i\pi lm/d}$.

One of the most convenient properties of the displacement operators defined above is their completeness, i.e. they form a complete non-Hermitian basis. As one can easily check via relation (3.21) the displacement operators fulfill the orthogonality condition

$$\mathrm{Tr}\{\mathcal{D}(l,m)\mathcal{D}^\dagger(l',m')\} = d\delta_{l,l'}\delta_{m,m'}. \tag{3.22}$$

As we have seen in the last section every density matrix can be represented via a decomposition in a complete operator basis. Such that the displacement operator also give us the opportunity to do so,

$$\rho = \frac{1}{d}\sum_{l,m=0}^{d^2-1} \mathrm{Tr}\{\rho\mathcal{D}(l,m)\}\mathcal{D}^\dagger(l,m) := \frac{1}{d}(\mathbb{1} + \vec{\xi}\cdot\vec{\mathcal{D}}^\dagger). \tag{3.23}$$

Here the Bloch representation via the vector $\vec{\xi}$ becomes visible. As the displacement operators are not Hermitian the d^2-1 entries of $\vec{\xi}$ are in general complex and therefore do not correspond to the outcomes of physical measurements . So the next step is to a find d^2-1 Hermitian operators whose expectation values we can use to fully describe an arbitrary quantum state of dimension d.

This goal can be achieved with the following definition

Definition 3.4.3. (Heisenberg-Weyl observables) Let $\mathcal{Q}(l,m) = \chi\mathcal{D}(l,m) + \chi^*\mathcal{D}^\dagger(l,m)$, where $\chi = \dfrac{(1 \pm i)}{2}$.

This operators fulfill the following orthogonality condition

$$\mathrm{Tr}\{\mathcal{Q}(l,m)\mathcal{Q}(l',m')\} = d\delta_{l,l'}\delta_{m,m'}. \tag{3.24}$$

Together with the identity identity matrix, $\mathcal{Q}(0,0) = \mathbb{1}_d$, the Heisenberg-Weyl obervables now form a complete orthogonal traceless Hermitian operator basis.

We can therefore decompose an arbitrary quantum state ρ of dimension d in the following form

$$\rho = \frac{1}{d}\sum_{l,m=0}^{d^2-1} \langle\mathcal{Q}(l,m)\rangle\mathcal{Q}(l,m), \tag{3.25}$$

where the expression $\langle \mathcal{Q}(l,m) \rangle$ is strictly real due to the Herminicity of the Heisenberg-Weyl observables.

Note that the Heisenberg-Weyl obervables reduce to the Pauli matrices for $d = 2$. Let us now have a look what happens if we extend the definition of the Heisenberg-Weyl observables to the infinite dimensional case.

3.4.1 Infinite dimensional limit

We can introduce a compact notation for the Heisenberg-Weyl observables to show that they can be systematically extended to the continuous limit of infinite dimensional systems. This is important as operational discretisations for continuous variable systems are of great value for quantum information processing tasks [88, 89].

We will do so by introducing the so called phase-space displacement amplitude,

$$\alpha := \sqrt{\frac{\pi}{d}}(m + il). \tag{3.26}$$

We can use α to define a real vector in a two dimensional space $\alpha := (\alpha_R, \alpha_I)$, such that we can rewrite $\mathcal{Q}(\alpha)$ as

$$\mathcal{Q}(\alpha) := \mathcal{Q}(\sqrt{\frac{d}{\pi}}\alpha_I, \sqrt{\frac{d}{\pi}}\alpha_R), \tag{3.27}$$

where $\mathcal{S} := \{\alpha : \alpha_I = \sqrt{\frac{\pi}{d}}l, \alpha_R = \sqrt{\frac{\pi}{d}}m\}$.
Now the decomposition of a quantum state ρ is given by

$$\rho = \frac{1}{d}\Big(\mathbb{1} + \sum_{\alpha \in \mathcal{S}} \langle \mathcal{Q}(\alpha) \rangle \mathcal{Q}(\alpha)\Big). \tag{3.28}$$

To investigate the limit $d \to \infty$ we consider $\hat{x} = Q\sqrt{2\pi/d}$ and $\hat{p} = P\sqrt{2\pi/d}$ as the position and momentum operators, such that $X^m \equiv e^{-ix\hat{p}}$ indicates position displacement by $x = m\sqrt{2\pi/d}$ and $Z^l \equiv e^{ip\hat{x}}$ displaces the momentum by $p = l\sqrt{2\pi/d}$. As a sanity check we observe that the Heisenberg commutation relation for position and momentum of a continuous variable system $[\hat{x}, \hat{p}] = i$ holds in the limit $d \to \infty$ Now we can reformulate the definition of the displacement operators in the following fashion

$$\mathcal{D}(p,x) := e^{ip\hat{x}}e^{-ix\hat{p}}e^{-ixp/2}. \tag{3.29}$$

To continue our exploration of the continuos limit we will need the so called Baker-Campbell-Hausdorff formula for exponential operators (for proof see [43]), i.e. given that $[A,[A,B]] = 0 = [B,[A,B]]$ we have that

$$fe^{A+B} = e^A e^B e^{-[A,B]/2}. \tag{3.30}$$

As this is the case here we can rewrite equation (3.29) as

$$\mathcal{D}(p,x) = e^{ip\hat{x} - ix\hat{p}}, \tag{3.31}$$

which is only valid in the limit of infinite dimension, i.e. $d \to \infty$.

With help of the creation (annihilation) operators of a bosonic mode $a^\dagger(a)$ we can write

$$\mathcal{D}(\alpha) = e^{\alpha a^\dagger - \alpha^* a}, \tag{3.32}$$

This leads to the following orthogonality condition for the infinite dimensional displacement operators

$$\mathrm{Tr}\{\mathcal{D}^\dagger(\alpha)\mathcal{D}(\alpha')\} = \pi\delta^2(\alpha - \alpha'), \tag{3.33}$$

such that we get

$$\mathrm{Tr}\{\mathcal{Q}(\alpha)\mathcal{Q}(\alpha')\} = \pi\delta^2(\alpha - \alpha'), \tag{3.34}$$

as the continuous analog of (3.24).

We observe that the replacement $\frac{1}{d}\sum_\alpha \to \frac{1}{\pi}\int d^2\alpha$ characterises the discrete-continuous transition as we have shown that the Heisenberg-Weyl observables presented in the last section for discrete systems can also be extended to continuous variable systems.

3.4.2 Anti-commutativity

A distinctive feature of the Pauli operators are their commutation relations, namely they all anti-commute mutually. As we will see in the next chapter this makes them useful in the task of entanglement detection. Here we will investigate the anti-commutation relation of the above introduced Heisenberg-Weyl observables. From the equation (3.21) and the discussion on the infinite limit of these observables we get that

$$\mathcal{D}(\alpha)\mathcal{D}(\alpha') = e^{i2\alpha\times\alpha'}\mathcal{D}(\alpha')\mathcal{D}(\alpha)\,, \tag{3.35}$$

holds for the discrete as well as the continouos displacement operators. We also have that $\mathrm{Im}(\alpha\alpha'^*) = \alpha \times \alpha'$ which allows us to conveniently characterise the commutativity and anti-commutativity among all basis elements. We can observe that for any α and α' such that

$$|\alpha \times \alpha'| = \frac{\pi}{2}(2n+1) \leq \frac{\pi(d-1)^2}{d}, \tag{3.36}$$

we get anti-commutation and therefore also anti-commutation for the corresponding pair of Heisenberg-Weyl observables, i.e.

$$\{\mathcal{Q}(\alpha), \mathcal{Q}(\alpha')\} = 0. \tag{3.37}$$

On the other hand for any α and α' such that

$$|\alpha \times \alpha'| = \pi n \tag{3.38}$$

we get commutativity. As we are more interested in the anti-commutativity for reasons that will become clear to reader in the next chapter, we end the discussion on commutativity of the Heisenberg-Weyl observables here.

Continuing our investigation of the anti-commutativity in the discrete case, we get the following relation

$$|m'l - ml'| = d(2n+1)/2 \leq (d-1)^2. \tag{3.39}$$

This leads to an interesting observation, namely that strict anti-commutativity between Heisenberg-Weyl observables can only hold in odd-dimensional systems. The reason behind this is simple. The right hand side of equation (3.39) is an integer only in the case that d is even while the left hand side is always integer.

3.4.3 The maximal set

Here we will prove that the cardinality of the maximal set of pairwise anti-commuting operators in the HW basis is at most three. The proof was given by Prof. Otfried Gühne during discussions on the topic of [9] and is presented in the Appendix therein. The outline is as follows: consider a set of four observables in the HW basis that are pairwise anti-commuting and then show that such a set cannot exist.
The first step of the proof is to parametrise the four observables in the set by two-dimensional real vectors $\vec{A}, \vec{B}, \vec{C}$ and $\vec{D}$, describing the displacement in phase space. The assumption of anti-commutativity then gives us six conditions on the vectors, the first one being

$$|\vec{A} \times \vec{B}| = k_1 \frac{\pi}{2} \tag{3.40}$$

where k_1 is an odd integer. The other five conditions are analogous, with odd numbers $k_2, \ldots, k_6$. Taking all six conditions into account implies that the following equation has to be satisfied

$$k_1 k_6 + k_2 k_5 - k_3 k_4 = 0. \tag{3.41}$$

This however is impossible as all k_i are odd.

3.4.4 Spectrum

Let us have a short look at the spectrum and the maximum eigenvalue of the Heisenberg-Weyl observables. We can easily compute the expectation value of a Heisenberg-Weyl observable squared with

$$\sqrt{\langle \mathcal{Q}^2(l,m) \rangle} = \sqrt{1 + \mathrm{Im}\langle \mathcal{D}(2l, 2m) \rangle}. \tag{3.42}$$

From the Unitarity of the displacement operators we have that their eigenvalues are bounded by 1. Therefore it is enough to look at the imaginary part, such that we get

$$\sqrt{\langle \mathcal{Q}^2(l,m) \rangle} \leq |q_{\max}| \leq \sqrt{2}. \tag{3.43}$$

if we take $q_{\max}$ to be the maximum eigenvalue.
As the Heisenberg-Weyl observables reduce to the Pauli operators for the case $d = 2$ we have that $\mathrm{Im}\,\mathcal{D}(2l, 2m) = 0$. Going to higher dimensions we see that $\mathrm{Im}\,\mathcal{D}(2l, 2m)$ is non-zero. We can still bound the maximum eigenvalue by $\sqrt{2}$. This means that the absolute values of the eigenvalues of $\sqrt{\mathcal{Q}^2(l,m)}$ for all Heisenberg-Weyl observables are given by

$$|q_n| = \sqrt{1 + \sin \frac{4\pi n}{d}} \tag{3.44}$$

with $n = 0, \cdots, d-1$.

Chapter 4

Entanglement

As already mentioned in the beginning of the last chapter the story of entanglement starts with Albert Einstein, Boris Podolsky and Nathan Rosen (EPR) and their article that was published in 1935 [18]. They were so baffled by the phenomenon that they concluded from its mere existence that quantum theory must be incomplete. This line of reasoning was later dubbed the 'EPR paradox', fortunately for us a paradox is only seemingly contradictory. It was however not those three who gave the phenomenon the name by which we know it today. It was Schrödinger who wrote a letter to Einstein using the german word 'Verschränkung' to describe the, at this point very strange, correlations between two particles. He then continued to publish a seminal paper on the topic, using entanglement as an english translation. In this work he already recognises the importance of the concept for the foundation of quantum theory [90], describing the phenomenon as follows:

> Another way of expressing the peculiar situation is: the best possible knowledge of a whole does not necessarily include the best possible knowledge of all its parts, even though they may be entirely separate and therefore virtually capable of being 'best possibly known', i.e., of possessing, each of them, a representative of its own. The lack of knowledge is by no means due to the interaction being insufficiently known - at least not in the way that it could possibly be known more completely - it is due to the interaction itself.

This describes well what was so discomforting to many physicists at that time. Namely, they were not willing to accept that knowledge of the parts did not automatically imply knowledge of the whole. They saw a need for 'hidden variables' that would explain this peculiar feature of quantum theory. While Grete Hermann's work went unnoticed, many soothed themselves with Niels Bohr's response to the EPR article [91, 92]. In this work he attributes the paradox to the inappropriate use of the 'detached observer'. Which in general is a valid and interesting point [93, 94], but in this case just led to a three decade long silence on the topic.

As discussed already it was up to John Bell to ignite the interest on this matter of quantum correlations. Although entanglement does not allow us to communicate superluminaly, we know today that entanglement is the key resource that takes quantum communication beyond the classically possible. Ranging from secure key distribution [30] to super-dense coding [95] it has found

application, it even improves communication capacities in a general sense [96]. Not only communication, entanglement pops up in nearly all quantum information processing tasks when we want to surpass the limits of classical physics.

So let us dive into the matter with the mathematical tools that we have acquired so far. Entanglement is actually commonly not defined directly, but through the absence of separability. So let us define what we mean by that.

Definition 4.0.1 (Separable states)**.** Given a bipartite quantum state ρ_{AB} it is called separable, iff it can be written as

$$\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i \,. \tag{4.1}$$

That definition in plain english means that we can decompose a separable state into a probabilistic mixture of individually valid marginal quantum states $\rho_{A/B}^i$, where the index i labels the different states. If the state can directly written as product of quantum states on each subsystem, we call it a product state.

Definition 4.0.2 (Product states)**.** Given a n-partite quantum state $\rho_{A_1 \ldots A_n}$ it is called product state, iff it can be written as

$$\rho_{A_1 \ldots A_n} = \bigotimes_{i=1}^{n} \rho_{A_i} \,. \tag{4.2}$$

Coming back to the topic of interest, we can now defined what an entangled state is.

Definition 4.0.3 (Entangled states)**.** Given a bipartite quantum state ρ_{AB} it is called entangled, iff it is not separable.

This means that we cannot write an entangled state as an mixture of product states. The overall state does not permit a description solely by its parts. While for bipartite qubit systems there exists an exhaustive arsenal of tools for a complete analysis, as soon as one increases the dimension the situation gets a lot more difficult. This is what we want to do here as we will especially focus on high dimensional systems in contrast to systems with two-dimensional degrees of freedom. These have the advantage that the capacity of each exchanged quantum system is not limited to one bit of information.

Recent years have seen extensive research devoted to the topic of high-dimensional entanglement, not only proving entanglement for such high dimensional systems [97] but also revealing its underlying dimensionality [98, 99, 100]. Also the general potential for accommodating many dimensions has been studied [101, 102, 103, 104, 105]. The technological developments of the last decades have made high dimensional degrees of freedom available in experiment and therefore allowed for a deeper understanding. For this thesis photonic systems will be of relevance, as they provide entangled photon pairs that are produced in down-conversion processes. Various different

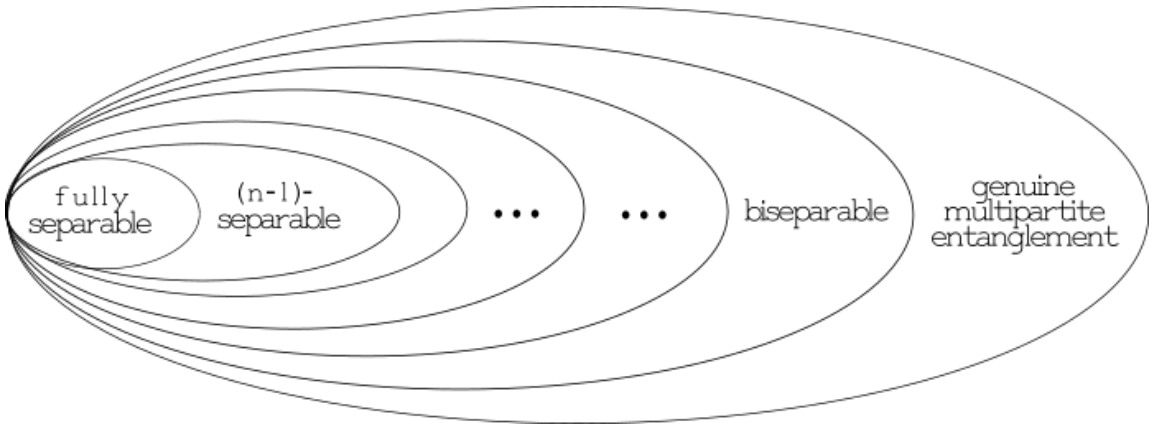

Figure 4.1: Schematic illustration of the structure of the different sets of states, from fully separable to genuine multipartite entangled. [44]

properties of these photon pairs can be exploited, ranging from entangling their paths in waveguides [98, 106, 107, 108] to their orbital angular momentum [99, 100, 109, 110, 111, 112, 113, 114] or energy-time bins [115, 116, 117]. The high-dimensionality of the entanglement and the underlying systems does not only increase the amount of shared information, which increases the efficiency of known protocols, but also allows for quantum communication at noise ratios that would not be possible for qubit systems [118, 119, 120, 121, 122].

Before we come to stating what will be presented in this chapter, let us have a quick look at what happens if we not only increase the dimension but also the number of systems that we consider. This leads to the study of multipartite entanglement. As we are using a negative definition for entanglement, we will have to introduce a generalised notion of separability.

Definition 4.0.4 (k-separable states)**.** Given a n-partite quantum state $\rho_{A_1\dots A_n}$ it is called k-separable, iff it can be written as

$$\rho_{A_1\dots A_n} = \sum_i p_i \bigotimes_{j=1}^{k} \rho^i_{\tilde{A}_j} \,. \tag{4.3}$$

for some partition $\mathcal{H} = \mathcal{H}_{\tilde{A}_1} \otimes \cdots \otimes \mathcal{H}_{\tilde{A}_k}$ of the Hilbert space, where $k \leq n$.

This means that for a k-separable state we can decompose our state into a probabilistic mixture of k-partite product states. This leads to a more complex notion of separability and therefore also entanglement. Only fully separable states do not contain any entanglement.

Definition 4.0.5 (Fully separable states)**.** Given a n-partite quantum state $\rho_{A_1\dots A_n}$ it is called fully separable, iff it can be written as

$$\rho_{A_1\ldots A_n} = \sum_i p_i \bigotimes_{j=1}^{n} \rho^i_{A_j} \,. \tag{4.4}$$

While in principle every state that is not fully separable contains at least some amount of entanglement, the structure of these different forms of entanglement gets very complex. Here we will focus on states that are not even 2-separable, this states are called genuine multipartite entangled (GME).

Definition 4.0.6 (Genuine multipartite entangled states)**.** Given a n-partite quantum state $\rho_{A_1\ldots A_n}$ it is called genuine multipartite entangled, iff it is not 2-separable.

Now that we have introduced all the necessary notions and concepts, let us proceed with decribing what will happen in this chapter. In the first part we will discuss how to detect entanglement, presenting a method that makes use of the advantageous properties of the HWO [8] and constitutes a generalisation of an already known result [123], later also proven differently in [124, 125]. The second part will focus on the possibility of quantifying the amount of entanglement shared between systems. Here a recently developed framework will be presented [9, 10] which makes it possible to do so only using two different global measurement settings. The framework will be analysed in an experimental proposal that will also be introduced in this chapter.

4.1 Detection

In this section we will discuss how to detect entanglement.

4.1.1 Anti-commutativity bound

Here, a theorem bounding sums of squared expectation values in the case of small mutual anti-commutators of the observables is presented and it is shown how this theorem can be used to detect entanglement. This theorem is a generalisation of the theorem presented in [123], later also proven differently in [124, 125], that was valid only in the case of dichotomic anti-commuting observables. Here we will show that this is also possible for observables with an arbitrary spectrum and non-vanishing anti-commutators.

Theorem 4.1.1 (Anticommutativity Bound)**.** *Let $\{\lambda_i\}_{i\in\mathcal{I}}$ with the index set $\mathcal{I} = \{1, 2, \ldots, d^2\}$ denote an orthonormal self-adjoint basis $\mathcal{B}$ of a d-dimensional Hilbert space $\mathcal{H}$ and $\mathcal{A} \subseteq \mathcal{I}$ refer to a subset of $\mathcal{B}$ such that $\frac{1}{2}\sqrt{\sum_{i\neq j\in\mathcal{A}}\langle\{\lambda_i, \lambda_j\}\rangle^2} \leq \mathcal{K}$. The complement is denoted $\bar{\mathcal{A}}$. Then the corresponding Bloch vector components c_i of any density matrix $\rho \in \mathcal{H}$ expressed in $\mathcal{B}$ as $\rho = \sum_{i\in\mathcal{A}} c_i\lambda_i + \sum_{l\in\bar{\mathcal{A}}} c_l\lambda_l$ can be bounded by*

$$\sum_{i\in\mathcal{A}} c_i^2 \leq \frac{\max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle + \mathcal{K}}{[\min_{i\in\mathcal{A}} \mathrm{Tr}(\lambda_i^2)]^2}. \tag{4.5}$$

Proof. Any $\rho \in \mathcal{H}$ can be expressed in an orthonormal basis $\mathcal{B}$. The chosen basis $\mathcal{B}$ can always be divided into $\mathcal{A} \cup \overline{\mathcal{A}}$, where the set $\mathcal{A}$ is at worst trivial. Now consider the observable $\mathcal{O}$

$$\mathcal{O} := \sum_{i \in \mathcal{A}} c_i \lambda_i. \tag{4.6}$$

As we know that the variance of any observable is positive we have that

$$(\Delta \mathcal{O})^2 := \langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2 \geq 0 \tag{4.7}$$

holds. Inserting the explicit representation of $\mathcal{O}$ above we obtain

$$\langle \mathcal{O}^2 \rangle = \langle \sum_{i \in \mathcal{A}} c_i^2 \lambda_i^2 + \sum_{i \neq j \in \mathcal{A}} c_i c_j \lambda_i \lambda_j \rangle \tag{4.8}$$

$$= \langle \sum_{i \in \mathcal{A}} c_i^2 \lambda_i^2 + \frac{1}{2} \sum_{i \neq j \in \mathcal{A}} c_i c_j \underbrace{(\lambda_i \lambda_j + \lambda_j \lambda_i)}_{\{\lambda_i, \lambda_j\}} \rangle \tag{4.9}$$

$$\leq \sum_{i \in \mathcal{A}} c_i^2 \langle \lambda_i^2 \rangle + \underbrace{\sqrt{\sum_{i \neq j \in \mathcal{A}} c_i^2 c_j^2}}_{\leq \sum_i c_i^2} \underbrace{\frac{1}{2} \sqrt{\sum_{i \neq j \in \mathcal{A}} \langle \{\lambda_i, \lambda_j\} \rangle^2}}_{\leq \mathcal{K}}, \tag{4.10}$$

applying the boundedness of the anti commutator of $\mathcal{A}$. We also know that

$$\langle \mathcal{O} \rangle^2 = \left(\sum_{i \in \mathcal{A}} c_i \operatorname{Tr}(\rho \lambda_i) \right)^2 \tag{4.11}$$

$$= \left(\sum_{j \in \mathcal{A}} \operatorname{Tr}(\sum_{i \in \mathcal{A}} c_i \lambda_i) c_j \lambda_j) + \sum_{l \in \overline{\mathcal{A}}} \operatorname{Tr}(\sum_{i \in \mathcal{A}} c_i \lambda_i) c_l \lambda_l) \right)^2 \tag{4.12}$$

$$= \left(\sum_{i \in \mathcal{A}} c_i^2 \operatorname{Tr}(\lambda_i^2) \right)^2 \tag{4.13}$$

by use of orthonormality of $\mathcal{B}$. Summing up above equations yields to the following statements

$$0 \leq \langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2 \tag{4.14}$$

$$\leq \sum_{i \in \mathcal{A}} (c_i^2 \langle \lambda_i^2 \rangle + c_i^2 \mathcal{K}) - (\sum_{i \in \mathcal{A}} c_i^2 \operatorname{Tr}(\lambda_i^2))^2 \tag{4.15}$$

$$\leq \sum_{i \in \mathcal{A}} c_i^2 \Big(\max_{i \in \mathcal{A}} \langle \lambda_i^2 \rangle - (\sum_{i \in \mathcal{A}} c_i^2) \min_{i \in \mathcal{A}} \operatorname{Tr}(\lambda_i^2)^2 \Big) + \sum_{i \in \mathcal{A}} c_i^2 \mathcal{K} \tag{4.16}$$

and thus

$$\sum_{i \in \mathcal{A}} c_i^2 \left(\max_{i \in \mathcal{A}} \langle \lambda_i^2 \rangle - \left(\sum_{i \in \mathcal{A}} c_i^2 \right) \min_{i \in \mathcal{A}} \operatorname{Tr}(\lambda_i^2)^2 \right) + \sum_{i \in \mathcal{A}} c_i^2 \mathcal{K} \tag{4.17}$$

$$\leq \sum_{i \in \mathcal{A}} c_i^2 \left(\max_{i \in \mathcal{A}} \langle \lambda_i^2 \rangle - \left(\sum_{i \in \mathcal{A}} c_i^2 \right) \min_{i \in \mathcal{A}} \operatorname{Tr}(\lambda_i^2)^2 \right)$$

The fact that every element of the sum on the right hand side in 4.17 is positive enforces that

$$0 \leq \max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle - \left(\sum_{i\in\mathcal{A}} c_i^2\right)\min_{i\in\mathcal{A}}\mathrm{Tr}(\lambda_i^2)^2 + \mathcal{K} \Longleftrightarrow \sum_{i\in\mathcal{A}} c_i^2 \leq \frac{\max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle + \mathcal{K}}{\min_{i\in\mathcal{A}}\mathrm{Tr}(\lambda_i^2)^2} \tag{4.18}$$

which concludes the proof. □

The computation of $\mathcal{K}$ can still be quite lengthy such that we want to establish a simple upper bound of the form

$$\mathcal{K} = \frac{1}{2}\sqrt{\sum_{i\neq j\in\mathcal{A}}\langle\{\lambda_i,\lambda_j\}\rangle^2} \leq \frac{1}{2}\sqrt{\sum_{i\neq j\in\mathcal{A}}\|\{\lambda_i,\lambda_j\}\|_\infty^2}. \tag{4.19}$$

Here the Heisenberg-Weyl observables come in handy as

$$\mathrm{Tr}(\lambda_i^2)^2 = d^2\,\forall\,\lambda_i \tag{4.20}$$

and we can relate the coefficients c_i to the Bloch vector representation, i.e.

$$c_i = \frac{1}{d}\langle\mathcal{Q}(\alpha)\rangle\,. \tag{4.21}$$

This gives us the desired result in the form of

$$\sum_{i\in\mathcal{A}}\langle\mathcal{Q}(\alpha_i)\rangle^2 \leq 1 + \max_n\left(\sin\frac{4\pi n}{d}\right) + \frac{1}{2}\sqrt{\sum_{i\neq j\in\mathcal{A}}\|\{\mathcal{Q}(\alpha_i),\mathcal{Q}(\alpha_j)\}\|_\infty^2}\,. \tag{4.22}$$

Now we can proceed on our path to use anti-commutativity to detect entanglement. The first step to do so is to pick a set of nonzero Bloch vector entries of a multipartite quantum state with anti-commuting reductions across the partition one is interested in. This means that the partition $A|\bar{A}$ and the set $\beta = \beta_A \cup \beta_{\bar{A}}$ with elements $\tau_\beta = \mathrm{Tr}(\rho\lambda_i \otimes \lambda_j)$ have that $i \in \beta_A$, $j \in \beta_{\bar{A}}$ and $\lambda_i(\lambda_j)$ are arbitrary observables acting on subsystem $A(\bar{A})$. We can then bound the sum of moduli of these entries for all states which are product states with respect to the selected partitions in the form of

$$\mathrm{Tr}(\rho_A \otimes \rho_{\bar{A}}\lambda_i \otimes \lambda_j) = \mathrm{Tr}(\rho_A\lambda_i)\mathrm{Tr}(\rho_{\bar{A}}\lambda_j)\,, \tag{4.23}$$

and $|\langle u|v\rangle| \leq \| u \|_2 \| v \|_2$.

Let us first look into the case where all observables λ_i are anti-commuting (i.e. $\mathcal{K} = 0$) and normalized [i.e. $\mathrm{Tr}(\lambda_i\lambda_{i'}) = d\delta_{i,i'}$]. This enables us to directly use the anti-commutativity bound to state that

$$\sum_{i\in\beta_A}|\mathrm{Tr}(\rho_A\lambda_i)|^2 \leq \max_{i\in\beta_A}\langle\lambda_i^2\rangle\,, \tag{4.24}$$

and analogously for $\bar{A}$. As a sum of moduli of expectation values is a convex function in the space of density operators we have that the inequality 4.24 is not only valid for all product states but also for all separable states. This additionally means that we can use the same procedure for

the cases of non-normalized or only partially anti-commuting observables.

Here again the HW observables come in very handy as we can use them to simplify the bound to

$$\sum_{i\in\mathcal{A}}\langle\mathcal{Q}(\alpha_i)\rangle^2 \leq q_{\max}^2 + \mathcal{K} \tag{4.25}$$

with $q_{\max}^2 = 1 + \max_{n\in\mathbb{N}} \sin(4\pi n/d)$ being the maximum eigenvalue of a HW observable which is the same for any $\mathcal{Q}(\alpha_j)$ for a given dimension and $\mathcal{K} \leq \frac{1}{2}\sqrt{\sum_{i\neq j\in\mathcal{A}} \|\{\mathcal{Q}(\alpha_i), \mathcal{Q}(\alpha_j)\}\|_\infty^2}$.

It shall be remarked that the case of non-vanishing anti-commutators has also been studied for dichotomic observables in the context of uncertainty relations [126]. Our central quantity here $\mathcal{K}$ is proportional to the 2-norm of the "anti-commutator matrix" introduced therein.

4.1.2 Bipartite case

As an illustration of above mentioned method let us start out with the bipartite case. Here we will use the maximally entangled state with qudits which is defined as follows

Definition 4.1.2 (Maximally entangled state)**.** Given two quantum systems A and B of dimension d the maximally entangled state is defined as

$$|\psi_{max,d}\rangle := \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1} |j\rangle_A |j\rangle_B . \tag{4.26}$$

In the bipartite case this state constitutes the maximal resource for many quantum information processing and communication tasks. We can write the state in its Bloch decomposition using the HW observables.

$$|\phi_d\rangle\langle\phi_d| = \frac{1}{d^2}\Big(\mathbb{1}\otimes\mathbb{1} + \sum_{\alpha\in\mathcal{S}} \mathcal{Q}(\alpha)\otimes\mathcal{Q}(\alpha)^*\Big) \tag{4.27}$$

where $\mathcal{Q}(\alpha)^* = \mathcal{Q}(-\alpha^*)$ denotes the complex conjugate.

As one can see from above decomposition the expectation values for the diagonal elements are all equal to 1. Therefore we can violate the upper bound by measuring only three anti-commuting observables for each party. We can achieve even higher violations by choosing three pairwise anti-commuting observables whose respective amplitudes fulfill the constraint $|\alpha_1 \times \alpha_2| = |\alpha_2 \times \alpha_3| = |\alpha_3 \times \alpha_1| = \pi/2(2n+1)$. This in turn leads to a general procedure for identifying sets of three pairwise anti-commuting observables (the largest cardinality achievable for HW observables as we have seen in Section 3.4.3).

We have that (4.24) for our case here (i.e. $\mathcal{K} = 0$) reads as

$$\sum_{i=1}^{3}\langle\mathcal{Q}(\alpha_i)\otimes\mathcal{Q}(\alpha_i)^*\rangle \overset{\text{DV}}{\leq} q_{\max}^2 \overset{\text{CV}}{\leq} 2, \tag{4.28}$$

where the upper bounds hold for separable states in the discrete (DV) and continuous variable (CV) case. If we let the dimension go to infinity, i.e. in the continous limit $d \to \infty$, the maximally entangled state (definition 4.1.2) becomes the perfectly correlated Einstein-Podolski-Rosen(EPR) state [18] which can also be written as an infinitely squeezed two-mode squeezed state, i.e. $\frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} dx \, |x\rangle \, |x\rangle = \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} dx \, |p\rangle \, |-p\rangle$ with continuous Bloch decomposition $\frac{1}{\pi^2} \int d\alpha^2 \mathcal{Q}(\alpha) \otimes \mathcal{Q}(\alpha)^*$. The set of the three pairwise anti-commuting observables contains the HW observables defined by three equiangular amplitudes with equal lengths

$$|\alpha_j| = \sqrt{\pi/\sqrt{3}} \simeq 1.34 \tag{4.29}$$

mutually separated by an angle $2\pi/3$.

To show the advantages that the HW basis has in contrast to the so far most used generalised Gell-Mann basis (see Section 3.3) let us see what this procedure would look like using the latter. Using the generalised Gell-Mann basis all the correlations used in equation 4.28 are all equal to $2/d$ [65] and thus we have that the necessary number of measurements that are required for detecting entanglement scales with the dimension d, thus making it highly impractical for large dimension. This clearly shows the advantageous properties of the HW observables in the detection of high-dimensional entanglement.

Within one local systems the anti-commuting elements (see section 3.4.3) are limited one can use tensor product bases containing commuting and anti-commuting sets to construct large sets. This can be done by choosing an odd number of anti-commuting factors, as this guarantees the anti-commutativity of the product. Through straightforward combinatorial calculations one can extend this to arbitrary high number systems, given that the commuting and anti-commuting local basis elements are found.

It should also be remarked that the bound presented here can be used to detect entanglement even if the observables are nondichotomic and thus constitutes a genuine generalisation of the results obtained in [123, 124, 125].

4.1.3 Explicit bipartite example in $d = 9$

To give the reader a more detailed insight on this method works let us consider the bipartite example in dimension $d = 9$. The explicit form of the density operator corresponding to the maximally entangled state in dimension $d = 9$ (see definition 4.1.2) reads as

$$\rho_{max,9} = \frac{1}{9} \sum_{i,j=0}^{8} |ii\rangle\langle jj| \, . \tag{4.30}$$

Considering the following five amplitudes $\alpha_1 = \sqrt{\pi}i/3$, $\alpha_2 = \sqrt{\pi}8i/3$, $\alpha_3 = \sqrt{\pi}4/3$, $\alpha_4 = \alpha_1 + \alpha_3$ and $\alpha_5 = \alpha_2 + \alpha_3$ we can define three observables,

$$\mathcal{O}_1 := \mathcal{Q}_A(\alpha_1) \otimes \mathcal{Q}_B(\alpha_2) \tag{4.31}$$
$$\mathcal{O}_2 := \mathcal{Q}_A(\alpha_3) \otimes \mathcal{Q}_B(\alpha_3) \tag{4.32}$$
$$\mathcal{O}_3 := \mathcal{Q}_A(\alpha_4) \otimes \mathcal{Q}_B(\alpha_5), \tag{4.33}$$

each of which has an expectation values of 1 for the maximally entangled state. We have seen in the last chapter there cannot exist HW observables that anti-commute for odd dimensions and therefore $\mathcal{K} \neq 0$. We can use equation 4.22 to establish an upper bound for the sum of moduli of operators that holds for all separable states

$$|\langle\mathcal{O}_1\rangle_{\rho_{max,9}}| + |\langle\mathcal{O}_2\rangle_{\rho_{max,9}}| + |\langle\mathcal{O}_3\rangle_{\rho_{max,9}}| = 3 \underbrace{\leq}_{SEP} 2.41987 \tag{4.34}$$

which is clearly violated by the maximally entangled states

4.1.4 Explicit tripartite example in $d = 4$

For the case of three systems we will use the Greenberger-Horne-Zeilinger, short GHZ, state which is defined as

Definition 4.1.3 (GHZ state)**.** The n-partite GHZ state of dimension d is defined as

$$|\psi\rangle_{GHZ(n,d)} = \frac{1}{\sqrt{d}} \sum_{i,j=0}^{d-1} |\underbrace{i \dots i}_{n}\rangle \ . \tag{4.35}$$

For $n = 2$ it reduced to the maximally entangled state (see definition 4.1.2). Written down as a density matrix for in the tripartite case for ququarts (i.e. $d = 4$) it reads as

$$\rho_{GHZ(3,4)} = \frac{1}{4} \sum_{i,j=0}^{3} |iii\rangle\langle jjj| \, . \tag{4.36}$$

Again we begin by choosing a set of indices, such that we ensure anti-commutativity along a partition. Now we also have to choose a bipartition along which we want to detect entanglement. Here we will consider the partition into the subsystems A and BC. The three measurement settings with amplitudes $\alpha_1 = \sqrt{\pi/4}$, $\alpha_2 = \sqrt{\pi}i$ and $\alpha_3 = \alpha_1 + \alpha_2$ allow us to define the following three observables

$$\mathcal{O}_1 := \mathcal{Q}_A(\alpha_1) \otimes \mathbb{1} \otimes \mathcal{Q}_C(\alpha_1) \tag{4.37}$$
$$\mathcal{O}_2 := \mathcal{Q}_A(\alpha_3) \otimes \mathcal{Q}_B(\alpha_2) \otimes \mathcal{Q}_C(\alpha_3) \tag{4.38}$$
$$\mathcal{O}_3 := \mathcal{Q}_A(\alpha_2) \otimes \mathcal{Q}_B(\alpha_2) \otimes \mathcal{Q}_C(\alpha_2). \tag{4.39}$$

We have constructed these observables such that the first cut $\mathcal{Q}_A$ is anti-commuting as well as $\mathcal{Q}_B$ is anti-commuting, while in contrary $\mathcal{Q}_C$ is commuting. This results in the overall cut $\mathcal{Q}_B \otimes \mathcal{Q}_C$ being anti-commuting as required. If we look at the expectation values of these observables for the GHZ state we get that$\langle\mathcal{O}_1\rangle_{\rho_{GHZ(3,4)}} = 1$, $\langle\mathcal{O}_2\rangle_{\rho_{GHZ(3,4)}} = -1$ and $\langle\mathcal{O}_3\rangle_{\rho_{GHZ(3,4)}} = 1$. We can again

establish an upper bound that holds for all separable states which is clearly violated inserting above values, i.e.

$$|\langle\mathcal{O}_1\rangle_{\rho_{GHZ(3,4)}}| + |\langle\mathcal{O}_2\rangle_{\rho_{GHZ(3,4)}}| + |\langle\mathcal{O}_3\rangle_{\rho_{GHZ(3,4)}}| = 3 \underbrace{\leq}_{SEP} 1. \tag{4.40}$$

Here we in this particular case we have that the maximal eigenvalue of all observables, used in (3.44), is equal to one. To recap how this method works, assume that the state would have been separable into A and BC which means that it would be possible to split the commuting tensor products of ABC into the anti commuting parts A and BC, such that theorem 4.1.1 would hold. This is not the case, as we see a violation of the bound and can therefore conclude that the state $\rho_{GHZ(3,4)}$ is not separable across the split $A|BC$. Note that we could also use this method for any other split (i.e. $B|AC$ and $C|AB$) analogously and see a violation. This means that $\rho_{GHZ(3,4)}$ is not separable across any bipartite split.

In order to see how resistant our method is against white noise (modeled by the totally mixed state of given dimension) we look at the noisy version of the GHZ state

$$\rho_{noisyGHZ(3,4)}(p) = p\rho_{GHZ(3,4)} + \frac{(1-p)}{64}\mathbb{1}\,. \tag{4.41}$$

We then see that with this method we can achieve entanglement detection up to an threshold of $p > \frac{1}{3}$. It should be remarked that the choice of observables above is not unique, although other choices of observables might lead to an diminished noise resistance.

4.2 Quantification

We have discussed in the beginning of the chapter the clear potential of high dimensional systems and entanglement of such. Though for some tasks it might not be sufficient to solely detect entanglement in a system but we rather need to quantify how much entanglement is contained within the system. Especially here, as we go to high dimensions not only the presence of entanglement counts but also its dimensionality. The dimensionality of entanglement (i.e. the Schmidt rank) is given by the rank of the reduced density matrix and denotes the minimal dimension that is needed to reproduce the correlations of the state. Such that by proving a certain dimensionality of entanglement we also prove that the system must have it least this dimension in order to accommodate these correlations. We see that proving the dimensionality of entanglement puts a number on the entanglement in a system, but it can be quite a deceiving one. This is best explained in an example. Consider the following state

$$|\psi_\epsilon\rangle = \frac{1}{\sqrt{2+k\epsilon^2}}\left(|0,0\rangle + |1,1\rangle + \epsilon\sum_{i=1}^{k}|i,i\rangle\right) \tag{4.42}$$

where $\epsilon \ll 1$. Although this state has a $(k+2)$-dimensional entanglement it only takes very little entanglement resources in order to create it, namely, a little more than a single Bell state is sufficient. We see that although it might prove quite useful to prove the dimensionality of entanglement what we really want in order to properly quantify the amount of entanglement contained in a state, is a number that describes the necessary entangled resources in order to create the state under investigation. This quantity we will call entangled bits or e-bits, it is simply the number of number of two-dimensional Bell states that are necessary to produce the target state. It is very hard task to calculate the number of e-bits contained in a state [127]. We have not yet found any efficient way to calculate this number, even if we have full access to a reconstructed density matrix and there might not be any, as the best known algorithm to decide whether this number is nonzero already scales exponential in the system's dimension [128].

Instead of actually calculating this number for a given state, what we want to do here is to give a lower bound in a very efficient way. One way to do so are entanglement witnesses. Entanglement witnesses are a class of operators that has been well researched for a long time now [129]. On the one hand they can be used to detect entanglement and on the other we can utilise the expectation values of these operators for quantifying the amount of entanglement contained in the state [130, 131, 132, 133, 134, 135]. Entanglement witnesses come with the drawback that generically they require a number of local measurement bases that scales with the system size in order to do so. This means that complexity of an implementation rapidly increase with growing dimension.

A way that allows to reveal entanglement with just two local measurements uses mutually unbiased bases (see definition 3.2.1) [136]. This has already been exploited for high dimensional experiments [137, 138, 139]. In the following we will introduce a method that combines the best of both approaches. This will allow us to give a lower bound on the e-bits contained in a state using just two local measurement settings [9]. This method can even be generalised further in order to tailor it to the target state [10].

4.2.1 Entanglement of Formation

As discussed above we want to quantify the amount of entangled bits contained in a given state. A measure that does so is the Entanglement of Formation (EOF) [140], for pure states quantifies the asymptotic conversion rate between maximally entangled states and the quantum state under investigation. It can be formulated in a regularised asymptotic way such that it precisely gives this entanglement cost as a rate of 'target states per Bell state' also for mixed states. In plain english this means, given N copies of qubit Bell states one can create k copies of the target state deterministically using only local operations and classical communication (LOCC). In the limit $N \to \infty$ one can then find this asymptotic conversion rate $\frac{N}{k}$. Let us continue by defining the Entanglement of Formation in technical terms first in the case of pure states.

Definition 4.2.1 (Entanglement of Formation (for pure states))**.** Given a pure state $|\psi_{AB}\rangle$ the Entanglement of Formation is defined as

$$E_{oF}(|\psi_{AB}\rangle) := S(\mathrm{Tr}_{A/B}(|\psi_{AB}\rangle\langle\psi_{AB}|)). \tag{4.43}$$

This expression for pure states corresponds exactly to the Entropy of Entanglement. We can generalise this definition to mixed states by introducing a optimisation over all decompositions into projectors onto pure states, called convex roof construction.

Definition 4.2.2 (Entanglement of Formation)**.** Given a quantum state ρ the Entanglement of Formation is defined as

$$E_{oF}(\rho) = \inf_{\mathcal{D}(\rho)} \sum_i p_i E_{oF}(|\psi_i\rangle). \tag{4.44}$$

This in words means that we take the minimal average entanglement across all possible decompositions as a measure when it comes to general states. As mentioned, even if the whole state ρ is known exactly, it is a hard problem even to decide whether the measure is nonzero [141].
Before we start looking into methods to get around the hardness of this problem let us have a peek on how it is possible to extend the definition of the Entanglement of Formation to also quantify multipartite entanglement. This is a notoriously hard task, contrary to the bipartite case there is no unique 'currency', from which every state can be created via LOCC. Despite the progress made using maximally entangled sets [142, 143], the question about how to evaluate these measures including many high dimensional parties is still an open one. Leaving aside the sufficiency one can come up with a more rudimentary classification, focusing solely on the necessary resources. Along this lines a multipartite generalisation of Entanglement of Formation was proposed in [133, 144, 145].

Definition 4.2.3 (Multipartite Entanglement of Formation)**.** Given a quantum state ρ the multipartite Entanglement of Formation is defined as

$$E_{GME} := \inf_{\mathcal{D}(\rho)} \sum_i p_i \min_{A_i} S(\mathrm{Tr}_{A_i}(|\psi_i\rangle\langle\psi_i|)), \tag{4.45}$$

where A_i denotes a bipartition.

Operationally spoken it covers the minimal necessary average entanglement across every cut for creating the target state via LOCC. As we have solely focused on the necessary resources this measure does not reveal any deeper structure of the entanglement, it is non-zero for any multipartite entangled state and trivial if there exists a decomposition into at least biseparable states.

4.2.2 MUB bound

As it is a hard problem to decide whether the Entanglement of Formation is nonzero even given full knowledge about the state in question ρ [141] we want to find a lower bound to this measure. The framework presented here relies on only two measurement outcomes. It may be surprising that an entanglement of $\log(d)$ can be certified using only two measurement settings but it is possible. This astonishing power comes with a trade-off, the bound is sensitive to noise. After the presenting the bound we will continue with an analysis of the performance of the bound. Let us begin with the bipartite case. The advantage that we have in bipartite systems is that we can use the Schmidt decomposition , which in general does not exist for multipartite states.

Our framework will make use of correlations in two mutually unbiased bases. Natural candidates for two such bases in high dimensions are for example discretised position and momentum correlation. The central figure of merit that we will use is similar to the one developed in [136]. This means that our method is easily applied to already existing experimental data (e.g. from [138, 137]). For bipartite systems the existing method makes use of the sum over all diagonal correlations in m different MUBs, that is

$$C_m(\rho) := \sum_{k=1}^{m} \sum_{i=0}^{d-1} \langle v_i^k (v_i^k)^* | \rho | v_i^k (v_i^k)^* \rangle \,. \tag{4.46}$$

Let us get a deeper look in simplest case, that of the qubit. For $d = 2$ exactly $m = 3$ MUBs exist. Physically that would correspond to the polarisation of photons, path information of an particle in an interferometer or whether an atom is excited or not. Basically any two-dimensional quantum degree of freedom. For the sake of simplicity let us take the polarisation of photons as the example to see what this quantity is about. Corresponding to the three MUBs there are six different polarisation states: Horizontal, Vertical, Diagonal, Anti-diagonal, Left-handed-circular, Right-handed-circular. We will denote theses by their initial letter, H, V, D, A, L, R. Furthermore we want that $\langle X, Y \rangle$ denotes the coincidence counts in X and Y for the first and second photon. This leads to

$$C_3 = \frac{\langle H, H \rangle + \langle V, V \rangle + \langle D, D \rangle + \langle A, A \rangle + \langle L, L \rangle + \langle R, R \rangle}{\langle H, H \rangle + \langle H, V \rangle + \langle V, H \rangle + \langle V, V \rangle}. \tag{4.47}$$

The actual values will depend on the state that we measured, still we can make some general statements. In [136] the authors prove that this quantity is upper bounded for separable states by $1 + \frac{m-1}{d}$. The maximally entangled (see 4.1.2) on the other hand reaches a values of m. Surpassing the value that is allowed for separable states means that we have detected entanglement. Here we are interested in the case of $m = 2$ which is the minimal number bases that one needs in order to detect entanglement. Not only is C_2 the easiest to realise experimentally but also it suffices to detect entanglement in any pure state. Now we have to relate it to the EOF in order to be able to not only detect but also quantify entanglement in a given state.

To continue with our task at hand we look at a motivating example. The three MUBs for the qubit case that we treated above can be chosen to be represented by the Pauli matrices σ_i. Then the six different polarisation states are described by the six eigenstates of the Pauli matrices. If we would only measure in two of the three polarisation directions on both sides, we would for example get information on the following expectation values $\langle \sigma_x \otimes \sigma_x \rangle$ and $\langle \sigma_y \otimes \sigma_y \rangle$. In the case that we would find both of these values to be close to -1, we would be able to make statements about

the possible outcome in the third measurement direction. Specifically the positivity of the density operator implies that it has to be close to -1 as well, i.e. $\langle \sigma_z \otimes \sigma_z \rangle \approx -1$. Such that by just making two measurements we can infer that the state at hand is close to a Bell state and therefore has an EOF close to 1. In the [146] the authors give such an example and furthermore the provide a semi-definite programming characterisation (SDP) to o evaluate the convex roof extended linear entropy, even if the values are not close to -1.

Let us put these statements into a more rigorous and analytical form. The following quantity allows us to quantify entanglement in a noise robust manner for any dimension d with just two measurements is defined as

$$B(\rho) := N\Big[d\left(\sum_{i=0}^{d-1}\langle v_i^2(v_i^2)^*|\rho|v_i^2(v_i^2)^*\rangle\right) - 1 - \sum_{\substack{m\neq n, m\neq l \\ l\neq o, n\neq o}} \sqrt{\langle v_m^1 v_n^1|\rho|v_m^1 v_n^1\rangle\langle v_l^1 v_o^1|\rho|v_l^1 v_o^1\rangle} - \sum_{i\neq j}\sqrt{\langle v_i^1 v_j^1|\rho|v_i^1 v_j^1\rangle\langle v_j^1 v_i^1|\rho|v_j^1 v_i^1\rangle}\Big] \tag{4.48}$$

where $N = \sqrt{\frac{2}{d(d-1)}}$. Whenever this quantity $B(\rho) \geq 0$ we can indeed lower bound the Entanglement of Formation in the following way

$$E_{oF}(\rho) \geq -\log(1 - \frac{B(\rho)^2}{2})\,. \tag{4.49}$$

As the next section 4.2.3 is devoted to the derivation of $B(\rho)$ and equation 4.49, let us continue here with a discussion on the implications and a brief look into noise resistance here.

One interesting observation is that if we take the maximally entangled state $|\psi_{max,d}\rangle$ we have that

$$B(|\psi_{max,d}\rangle\langle\psi_{max,d}|) = \sqrt{2(1-\frac{1}{d})}. \tag{4.50}$$

This means that we can quantify all the entanglement of the maximally entangled state as equal to $E_{oF} = \log(d)$ with just two global measurement setting. It is important to stress that these results hold for any pair of MUBs. Moreover, the additional terms that are needed for computing $B(\rho)$ can be recorded alongside with the original measurements which means that there is no need for adjusting the local measurement settings.

So in order to gauge the practical usefulness it will be essential to study the performance of the bound in experimentally feasible settings with all the specific sources of noise taken into account that correspond to that very experimental setting taken into account. This we will do in section 4.2.6. Here we conduct a general theoretical analysis of the noise resistance for two paradigmatic noise models, dephasing and white noise. This analysis clarifies that the feature that our bound holds for any arbitrary pair of MUBs comes at the cost of having higher sensitivity to noise.

We start dephasing noise, where we mix the intended (for the bipartite case always maximally entangled state as it constitutes the maximal resource) target state is mixed with its dephased version, i.e.

$$\rho_{dp}(p) = p|\psi_{max,d}\rangle\langle\psi_{max,d}| + \frac{(1-p)}{d}\left(\sum_{i=0}^{d-1}|ii\rangle\langle ii|\right). \tag{4.51}$$

Computing our quantity for $\rho_{dp}(p)$ leads to

$$B(\rho_{dp}) = N\left(d\left(p + \frac{(1-p)}{d^2} - 1\right)\right). \tag{4.52}$$

This implies that we can see entanglement up to a noise threshold of $p_{crit} = \frac{1}{d+1}$. We can immediately see that the noise resistance increases with the dimension d. On the other if we look at the white noise model we mix our target state with the maximally mixed state, as we have already done in the sections 4.1.3 and 4.1.4.

$$\rho_{wn}(p) = p|\psi_{max,d}\rangle\langle\psi_{max,d}| + \frac{(1-p)}{d^2}\mathbb{1}_d\,. \tag{4.53}$$

Evaluating our quantity for the white noise model we have that

$$B(\rho_{wn}) = N\left(d\left(p + \frac{(1-p)}{d^2} - 1 - \frac{d-2}{d}(d-1)^2(1-p)\right)\right). \tag{4.54}$$

This actually leads to a decreasing noise resistance of $p_{crit} = \frac{d^2-3d+2}{d^2-2d+2}$ which approaches 1 for $d \to \infty$. In this short analysis we have used very idealised models of noise that are both somewhat unrealistic. As already mentioned we will there conduct a more detailed study of the noise robustness in section 4.2.6 for the specific experimental proposal that will be introduced in section 4.2.4.

4.2.3 Derivation of the bound

We now want to present a detailed derivation of the fact that two MUB measurements can quantify bipartite entanglement for any dimension d, even without any knowledge about the specific structure of the second MUB. We start the derivation by pointing out the lower bounds for the concurrence (i.e. the square root of the linear entropy) developed in [144, 145] as

$$I(\rho) := \sqrt{\frac{2}{d(d-1)}}\left(\sum_{m\neq n}\underbrace{|\langle mm|\rho|nn\rangle|}_{I_1} - \underbrace{\sqrt{\langle mn|\rho|mn\rangle\langle nm|\rho|nm\rangle}}_{I_2}\right) \tag{4.55}$$

$$\leq \inf_{\mathcal{D}(\rho)}\sum_i p_i\sqrt{\left(2\left(1-\mathrm{Tr}({\rho_i}_A{}^2)\right)\right)}\,, \tag{4.56}$$

where $\{p_i\}$ is a probability distribution.

To get a bound on the Entanglement of Formation we can use the relation between the linear entropy $H_L := 1 - \mathrm{Tr}(\rho^2)$ and the Renyi 2-entropy $H_2(A)_\rho = -\log(\mathrm{Tr}(\rho_A^2))$. Moreover we know

that that the family of Renyi entropies is monotonically decreasing in α, i.e. it holds that $H_\alpha \geq H_\beta \, \forall \alpha \leq \beta$. As a reminder we repeat the definition of the Entanglement of Formation here,

$$E_{oF}(\rho) = \inf_{\mathcal{D}(\rho)} \sum_i p_i H(A)_{\rho_i} \, . \tag{4.57}$$

with $\{p_i\}$ again denoting a probability distribution. If we take together above observations we directly find the following lower bound

$$E_{oF}(\rho) \geq -\log(1 - \frac{I^2}{2}) \tag{4.58}$$

So from now on our task will be to lower bound $I(\rho)$ using only two mutually unbiased measurements. The first thing to notice is that I_2 in equation 4.55 is directly accessible from correlations in the first basis $\langle mn|\rho|mn\rangle$, i.e. choosing $\{|v_i^1\rangle\} = \{|i\rangle\}$. Since the term I_2 is strictly positive and it decreases $I(\rho)$. Therefore it is desirable to have these 'wrong' correlations as suppressed as possible. This means, that when analysing experimental data one should choose the basis states such that the most correlated elements between Alice and Bob are labeled by the same numbers m.

Now comes the more tricky part: Estimating the total number of coherences in the term I_2. First we remind ourselves that to be mutually unbiased the bases have to have an overlap that fullfills $|\langle v_i^2|j\rangle|^2 = \frac{1}{d} \, \forall i, j$. That means that we can write the second basis as

$$|v_k^2\rangle = \frac{1}{\sqrt{d}} \sum_{m=0}^{d-1} e^{-i\phi_m^k} |m\rangle \, . \tag{4.59}$$

If we now evaluate the following linear combination of C_m (see equation 4.46) for the two mutual unbiased bases

$$\Sigma(\rho) := C_2 - C_1 = \sum_{k=0}^{d-1} \langle v_k^2 v_k^{2*} | \rho | v_k^2 v_k^{2*} \rangle \tag{4.60}$$

we first notice that $|v_k^2 v_k^{2*}\rangle = \frac{1}{d} \sum_{m,n} e^{-i(\phi_m^k - \phi_n^k)} |mn\rangle$ such that we can write

$$\Sigma(\rho) = \frac{1}{d^2} \sum_k \sum_{m,n,l,o} e^{i(\phi_m^k - \phi_n^k + \phi_o^k - \phi_l^k)} \langle mn|\rho|lo\rangle \, . \tag{4.61}$$

To properly evaluate $\Sigma(\rho)$ we conduct a case by case study for the different values of the indices. We start by observing that

$$\sum_{m=0}^{d-1} e^{i(\phi_m^k - \phi_m^{k'})} = 0 \quad \forall \, k \neq k' \tag{4.62}$$

due to the fact that $\langle v_k^2 | v_{k'}^2 \rangle = 0 \, \forall \, k \neq k'$. From this follows that all the terms such that $m = l$ and $n \neq o$ vanish as we find the pre-factor to be $\sum_{k=0}^{d-1} e^{i(\phi_o^k - \phi_n^k)} = 0$. The same holds true for terms where the indices fullfill $m = n$ and $l \neq o$, $n = o$ and $m \neq l$ or $l = o$ and $m \neq n$.

This means that there are three cases left which we have to evaluate, namely:

1) $m = l$ and $n = o$

2) $m = n$ and $l = o$

3) $m \neq n$, $m \neq l$, $n \neq o$ and $l \neq o$

We will now split the remaining non-vanishing parts of the sum in $\Sigma(\rho)$ into three parts $\Sigma(\rho) = \Sigma_1(\rho) + \Sigma_2(\rho) + \Sigma_3(\rho)$ corresponding to the three different cases. For the first part this simply yields

$$\Sigma_1 = \frac{1}{d} \underbrace{\sum_{m,n} \langle mn|\rho|mn\rangle}_{=\mathrm{Tr}(\rho)=1} = \frac{1}{d}\,. \tag{4.63}$$

The second case is especially interesting as

$$\Sigma_2(\rho) = \frac{1}{d} \sum_{m \neq l} \langle mm|\rho|ll\rangle\,, \tag{4.64}$$

which means that it contains exactly the same expectation values as I_2 in the quantity that we want to bound. Last but not least we have

$$\Sigma_3(\rho) = \frac{1}{d^2} \sum_{\substack{m\neq n, m\neq l \\ l\neq o, n\neq o}} c_{m,n,l,o}\, \mathrm{Re}[\langle mn|\rho|lo\rangle] \tag{4.65}$$

where we have introduced the following object purely for convenience

$$c_{m,n,l,o} := \sum_k e^{i(\phi_m^k - \phi_n^k + \phi_o^k - \phi_l^k)}\,. \tag{4.66}$$

The Cauchy-Schwarz inequality implies that

$$0 \geq d\Sigma_3(\rho) - \sum_{\substack{m\neq n, m\neq l \\ l\neq o, n\neq o}} \sqrt{\langle mn|\rho|mn\rangle \langle lo|\rho|lo\rangle} \tag{4.67}$$

and as we have that $\mathrm{Re}[z] \leq |z|$ we can lower the first part of $I(\rho)$ in the following way

$$I_1 \geq d \left(\Sigma - \frac{1}{d} - \frac{1}{d} \sum_{\substack{m\neq n, m\neq l \\ l\neq o, n\neq o}} \sqrt{\langle mn|\rho|mn\rangle \langle lo|\rho|lo\rangle} \right)\,. \tag{4.68}$$

Thus

$$\begin{aligned} I_1 - I_2 \geq d \left(\Sigma(\rho) - \frac{1}{d} - \frac{1}{d} \sum_{\substack{m\neq n, m\neq l \\ l\neq o, n\neq o}} \sqrt{\langle mn|\rho|mn\rangle \langle lo|\rho|lo\rangle} \right) \\ - \sum_{m\neq n} \sqrt{\langle mn|\rho|mn\rangle \langle nm|\rho|nm\rangle} = \frac{1}{N} B(\rho) \end{aligned} \tag{4.69}$$

where $N = \sqrt{\frac{2}{d(d-1)}}$. It is now straightforward to see that we have finally arrived at our goal, i.e. we shown that

$$E_{oF} \geq -\log\left(1 - \frac{B(\rho)^2}{2}\right). \quad \square \tag{4.70}$$

4.2.4 Experimental proposal

A natural candidate for two mutual unbiased bases are discretised position and momentum. They correlations in theses bases can be produced at a high quality and therefore have already been used to verify entanglement experimentally [147, 148, 149]. The experimental proposal that we will present here to show the usefulness of the method presented above will be based on modern cameras and lenses (which perform a Fourier transformation in the far field). We also utilise down-conversion photons whose spatial correlations we want to quantify. In the scenario that we consider a pair of photons illuminate single-photon sensitive cameras. Simple optics suffice to access two mutual unbiased bases (the position-basis and the momentum-basis) which we will use to evaluate the strength and noise-dependence of our method. Moreover this presents an ideal case to exploit the fact that we do not need to know he exact phase relation between the mutually unbiased bases, saving a lot of effort on side of the experimentalists.

Spontaneous parametric down-conversion (SPDC) in a nonlinear crystal is process that creates the photon pairs which we will use [150]. The creation of correlated photons in this manner is a well established technique in quantum optical experiments. In a type-II SPDC the two photons can be deterministically separated due to their opposite spin. This would also be the case if they have different wavelengths. The two beams are sent onto two different regions of an ICCD (intensified charge-coupled device) camera or two separated cameras, depending on the setup at hand.

The advantage of ICCD cameras is their high detection efficiency, which amounts to about 20% for 800nm and up to 50% for green light. Furthermore they feature are very low dark count rate therefore qualifying as single-photon sensitive devices. One could also use electron multiplying charge-coupled device (EMCCD) cameras which come a long with even higher efficiencies but have the disadvantage of very high dark count rates. In the case one can use averaging techniques in order to infer quantum correlation [105, 151]. We do not use this approach here, as ideally we tap directly into raw count rates of the measured data.

In the first step of the proposal we directly image the crystal onto the camera. This way we should be able to see the strong position correlations produced in the SPDC. For the second step a lens is introduced into the setup (see A in figure 4.2) that allows us to go to the Fourier-plane of the crystal. The second measurement should therefore reveal the strong momentum correlations, which should result in strong position anti-correlations in the image. The ratio between the marginal and conditional probability, called Federov-Ratio F is widely used measure in order to quantify the strength of spatial correlations. Here all the examples will be calculated with a value of 25, which is a realistic value that can be reached in contemporary experiments [150].

As already said, it is essential for gauging the power of our method to study the performance of the bound taking all sources of noise into account. In the proposed experimental setup the three major sources of noise that reduce the detectable number of e-bits through equation (4.49). The

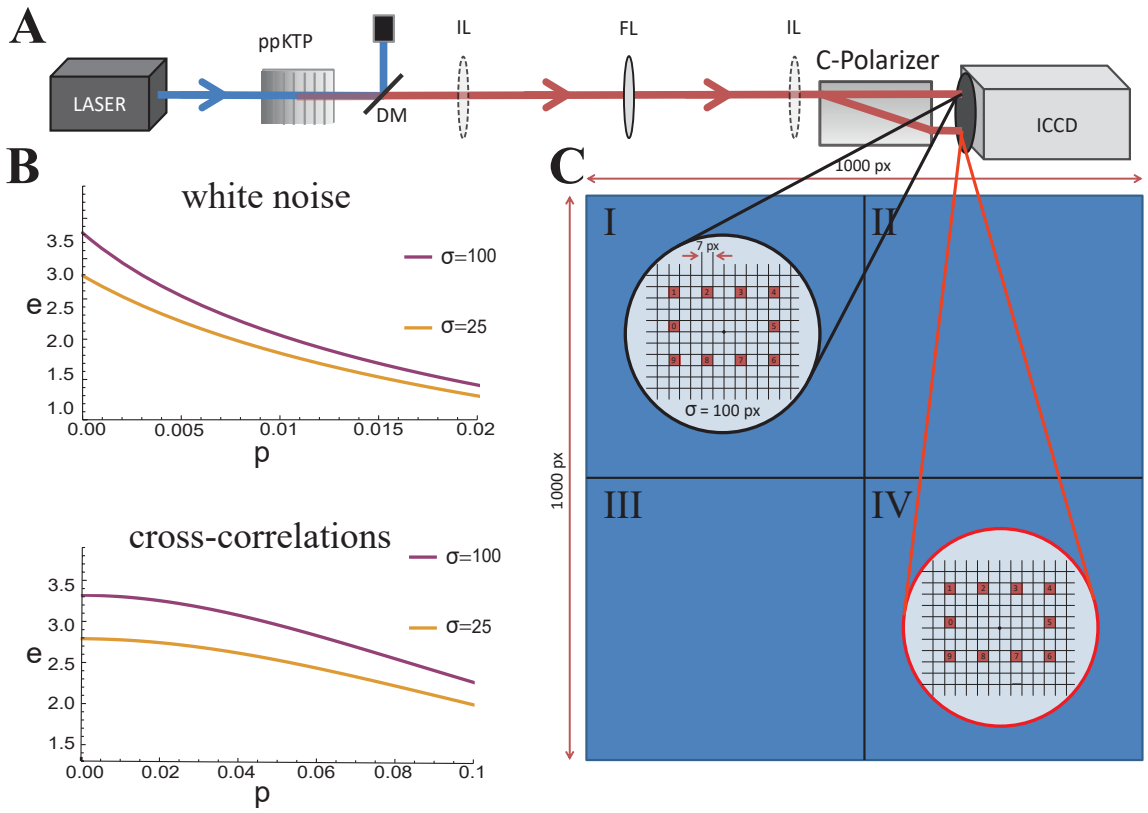

Figure 4.2: **A**: The proposed experimental setup. A laser produces pairs of spatially entangled photons with wavelength of 800nm in a SPDC process in the nonlinear ppKTP crystal. After the pump beam is removed by a dichroic mirror, the photon-pair is deterministically separated with a calcite polariser. The two parts of the beam then go to different positions at a camera. Depending on the lens configuration, the photons are detected in the position-basis (if the two imaging lenses (IL) are in the beam path) or in the momentum-basis (if only the Fourier lens (FL) is used). **B**: Here we showcase two different effects that predominantly reduce the certifiable e-bits, white noise and cross-correlations between different pixels. The purple and orange graphs show different weighting of the individual terms. Purple depicts a beam with $\sigma = 100$ pixels, and orange $\sigma = 25$ pixels. p stands for the probability of white noise or cross-correlations, respectively. **C**: The considered areas of pixels at the screen of the ICCD camera as considered in the calculations. Due to the significant impact of white noise, it is optimal to use only small parts of the camera. Figure and caption from [9].

first effect are dark counts, they should be roughly distributed uniformly over all camera pixels, such that they average out and therefore result in white noise. The second effect are cross-correlations between different states that also impact the the number of certifiable e-bits negatively. Last an non-equal weighting of the different modes also has an effect on the quality of the produced data. In section 4.2.6 these effects are analysed in detail and the graphs of B in figure 4.2 show the interplay of these effects.

As for different configurations of the camera and labelling of the pixels in the analysis of the data, the mentioned sources of noise impact the results differently, we will look at some example cases that are within the range of feasible experiments. In the next section we will perform a careful analysis of these effects. For the moment we will look broadly at what is possible. Intuitively, we observe that if we group large regions together, we can reduce cross-talk but in turn increase the number of dark counts. In figure 4.2C, a carefully chosen configuration is presented that allows for the certification of more than 3 e-bits. In this configuration we group together 7×7 pixels to make up a region whose counts we use to compute the expectation values that include the chosen label. We leave enough space between the regions in order to minimise the cross-talk between them. The concentric arrangement of the considered regions lowers the effect of the unequal weighting of the beam. In this setup we find that the optimal number areas to consider is 10 (i.e. we consider qudits with dimension $d = 10$) for which we can tolerate a noise-level of 0.6% in order to detect 2.4 e-bits. If we reduce the size of the considered areas to 3×3 pixels, we find 3.05 e-bits if the rest of the setup remains the same. It might at first seem a little counterintuitive that a reduced size of the areas considered actually leads to a higher value of certified e-bits. Although the smaller regions decrease the number of photons detected per unit of time (proportional to the area not counted) we have that in this example the chosen wide single-photon beam features a probability distribution for finding a photon in any of the regions that is nearly uniform. Therefore the reduced size leads to a significant decrease of white noise resulting in a higher number of e-bits. The optimisation of the design for the optimal areas can be done after the data is collected, as it is a mere question of labelling. In section 4.2.7 a step-by-step calculation for the case of $d = 3$ is conducted.

Last we have to also look at the stabilisation of the setup for longer times as this poses an experimental challenge. As for ICCDs we that the maximum imaging rate is about 10 per second. That means that we have to choose a very low photon production rate of roughly 1 Hz (=Hertz is the unit of frequency in the International System of Units and is defined as once per second), to prevent multi-photon events which would increase the level of white noise. Also only a small fraction of the pairs (around 1 in 100) actually arrives in the considered regions, taking into account the detection efficiency this leads to a rate for the measured photon pairs in the order of 10^{-4} Hz. At this rate one needs around $100 - 200$ hours to collect the necessary data. There exist various workarounds to this caveat. Switching from the standard 800nm frequency where the efficiency is around 20% to green entangled photons [152] would pump up the detection efficiency to about 50%. If available one could also use novel sCMOS (scientific complementary metal-oxide-semiconductor) cameras [153, 154] which are faster and feature lower noise compared to the ICCDs. Also the use compressed sensing approaches [149] could be an option.

4.2.5 Evaluation

In principle, the method presented here gives one the ability to certify arbitrary amounts of entanglement, for pure states at least. Even with a realistic noise assessment we show that it should be possible with current technology to certify three times more entanglement than then maximally entangled (qubit) state. This can be achieved with only two global measurement settings and without knowledge of the specific phase relations of the two MUBs. Development of more advanced camera technologies or clever detection design should push the number of experimentally certifiable e-bits even higher. One of the reasons we developed this method was to show the potential of down conversion sources for high dimensional quantum communication. Still, this method is not limited to optical systems. In principle it can be applied to any other quasi-continuous variable entangled systems, given the necessary experimental control.

One way to further reduce the sensitivity to noise of our method would be the inclusion of further measurements. This however would require precise knowledge or control over the phase relation between them and is not possible for all systems. Take wavelength-entangled systems for example where one has access to exactly two MUBs, namely wavelength and time. This might prove the great potential of wavelength-entangled quantum systems[155, 156, 157]. Note that the method was recently generalised to be applicable for relation between the two bases and also including a protocol how to tailor the second basis specifically to the target state. This allowed us to show 9-dimensional entanglement in an 11-dimensional system using orbital angular momentum (OAM) states.

Furthermore, this method could prove to be very useful for quantum key distribution protocols (QKD). Up today this protocols show a heavy trade-off between security and implementability. Fully device independent quantum key distribution for example relies on (almost) loophole free Bell inequality violation, which have been achieved recently but with an immense experimental effort [158, 159, 160]. Prepare and measure schemes are easier to implement but suffer the feat that be attacked by hacking the source or the measurement devices. Randomised application of our method could help to certify that entanglement is present in source, not affecting the key rates too much. This would enable a certification of high levels of security for high dimensional QKD through the interpolation between prepare and measure and fully device independent schemes.

4.2.6 Analysing the noise robustness

As promised this section is devoted to conducting a detailed analysis of noise sources in the experimental proposal presented above. As mentioned the three most impactful effects on the amount of certifiable e-bits that we can discuss here are:

i) Cross correlations

ii) Unequal weighting

iii) White noise

Where one can attribute i) to non-perfect correlations, ii) to the gaussian character and iii) to accidental dark counts of the camera.

Furthermore we need to specify how to get from the measured data to the to expectation values $\langle mn|\rho|mn\rangle$ which we need to evaluate the bound. This can be done by taking to be $N_{m,n}$ as the total number of correlated clicks recorded between area m on Alice's side and area n on Bob's side and then computing

$$\langle mn|\rho|mn\rangle = \frac{N_{m,n}}{\sum_{i,j} N_{i,j}}\,, \tag{4.71}$$

analogously for the second basis.

Cross correlations

There are physical constraints in the experimental production of the spatial correlations that we exploit in our proposed setup. Such constraints are for example the size of the pump laser or the length of the crystal that lead to a finite strength of the correlations. This leads to a cross-talk between pixel areas which lead to a smaller number of certified e-bits.

For the sake of simplicity we restrict ourselves to cross-correlations only between directly neighbouring areas. The description of the cross-correlations depends on the geometry of the considered pixel areas and the restriction to direct neighbours that we only take into account first-order contributions. In the configuration that we chose for our proposal (figure 4.2C) all areas have two direct neighbors.

If one takes $NB(x)$ to be the set of neighbours of x, then

$$\overline{|x,x\rangle} = N_C \left(c_x \,|x,x\rangle + \sum_{y \in NB(x)} c_{x,y} \left(|x,y\rangle + |y,x\rangle\right) \right) \tag{4.72}$$

where c_x and $c_{x,y}$ are calculated numerically. Note that we did significantly reduce the impact of cross-correlation to the outcome of the bound by introducing blank areas between the ones that are considered as can be seen in figure 4.2C.

Unequal weighting

The unequal weighting of the different pixel-areas comes mostly from the Gaussian character and is influenced by the discretisation of the grid. One can write the state as

$$|\psi_U\rangle = N_U \left(\sum_{x=0}^{x_{max}} w_x \overline{|x,x\rangle} \right) \tag{4.73}$$

where w_x is calculated numerically and $\overline{|x,x\rangle}$ was discussed above. Note that we could have flattened the distribution by an adjustment of the discretisation, this however results in the adverse effect of increasing the white noise.

White Noise

The majority of white noise can be attributed to dark counts on the ICCD camera but also to photon loss. Both of these effects we expect to be uniformly distributed over the whole camera so that we can use the same model that we have used so far, namely

$$\rho_W = p_W |\psi_U\rangle\langle\psi_U| + \frac{1-p_W}{d^2}\mathbb{1}_{d^2} \tag{4.74}$$

with $|\psi_U\rangle$ being the same as in equation (4.73).

As we have discussed already another source of white noise would be a multi-photon events. We that for every specific setting of dark-count rate, detector-efficiency and used area of the ICCD we to find the optimal value for the photon-pair rate P. We will therefore only take into account events with exactly one photon pair.

Considering only events with exactly two photons and rejecting all others makes our analysis considerably more simple. If we consider two regions, region A and region B, the following events are possible

- Two dark counts
- One dark count, one real count
- Two real counts.

What we are interested in, is the probability of good counts. First we calculate all counts,

$$\begin{aligned} N_{all} = {} & \bar{P}\,(D_1 D_2) + \\ & + P^1\left(D_1 D_2 \bar{\epsilon_1}\bar{\epsilon_2} + D_1\epsilon_2\bar{D}_2\bar{\epsilon_1} + \epsilon_1 D_2\bar{D}_1\bar{\epsilon_2} + \epsilon_1\epsilon_2\bar{D}_1\bar{D}_2\right) + \\ & + P^2\left(D_1 D_2 \bar{\epsilon_1}^2\bar{\epsilon_2}^2 + 2D_1\epsilon_2\bar{D}_2\bar{\epsilon_1}^2\bar{\epsilon_2} + 2\epsilon_1 D_2\bar{D}_1\bar{\epsilon_2}^2\bar{\epsilon_1} + 4\epsilon_1\epsilon_2\bar{D}_1\bar{D}_2\bar{\epsilon_1}\bar{\epsilon_2}\right) + \\ & + \ldots \\ & = \sum_{n=0}^{\infty} P^n\left(D_1 D_2\bar{\epsilon_1}^n\bar{\epsilon_2}^n + nD_1\epsilon_2\bar{D}_2\bar{\epsilon_1}^n\bar{\epsilon_2}^{n-1} + n\epsilon_1 D_2\bar{D}_1\bar{\epsilon_2}^n\bar{\epsilon_1}^{n-1} + n^2\epsilon_1\epsilon_2\bar{D}_1\bar{D}_2\bar{\epsilon_1}^{n-1}\bar{\epsilon_2}^{n-1}\right) \end{aligned} \tag{4.75}$$

where D is a dark count, $\bar{D}$ is a good count, ϵ means the photon is detected and $\bar{\epsilon}$ that it is not. Furthermore we denoted $\bar{P}$ the probability that no photon-pair is created. We are only interested in the good counts for the case that one photon pair is created, i.e.

$$N_{good} = P^1\left(\epsilon_1\epsilon_2\bar{D}_1\bar{D}_2\right) \tag{4.76}$$

and in terms of probability, simply

$$p_{good} = \frac{N_{good}}{N_{all}}. \tag{4.77}$$

As we assume that both the dark counts and the detector efficiency is equal all over the ICCD camera, we furthermore have that $D_1 = D_2$ and $\epsilon_1 = \epsilon_2$. We arrive at the white-noise probability through

$$p_W = 1 - p_{good}\,. \tag{4.78}$$

4.2.7 Step-by-step calculation in three dimensions

To explicitly show how the method works, let us have a look at data that has been simulated for the case $d = 3$ which includes a conservative estimation of noise. We will use the following labels

$$\begin{aligned} &\text{Basis 1}: \quad |0\rangle_1 \quad |1\rangle_1 \quad |2\rangle_1 \\ &\text{Basis 2}: \quad |0\rangle_2 \quad |1\rangle_2 \quad |2\rangle_2 \, . \end{aligned}$$

Given the data we can compute a 3×3 correlation matrices for both bases

$$Corr_i = \begin{pmatrix} \langle 0,0|_i \, \rho \, |0,0\rangle_i & \langle 0,1|_i \, \rho \, |0,1\rangle_i & \langle 0,2|_i \, \rho \, |0,2\rangle_i \\ \langle 1,0|_i \, \rho \, |1,0\rangle_i & \langle 1,1|_i \, \rho \, |1,1\rangle_i & \langle 1,2|_i \, \rho \, |1,2\rangle_i \\ \langle 2,0|_i \, \rho \, |2,0\rangle_i & \langle 2,1|_i \, \rho \, |2,1\rangle_i & \langle 2,2|_i \, \rho \, |2,2\rangle_i \end{pmatrix} . \tag{4.79}$$

We can use a single index i to indicate the basis because we assume that both photons are measured in the same basis. We compute the respective expectation values in the bases through the number of coincidence clicks. So for example $\langle 0,0|_i \, \rho \, |0,0\rangle_1$ can be computed by taking the number of coincidence clicks N_{11} that we have gotten in basis 1 for the region 0 on both sides and the total number of clicks. Using the formula given in the beginning of the section we have that

$$\langle 0,0|_1 \, \rho \, |0,0\rangle_1 = \frac{N_{11}}{\sum_{i,j=1}^{3} N_{ij}} \, . \tag{4.80}$$

Note that this method works the same no matter which type of MUBs one uses. Here we will use data that is simulated for measurements in position (Basis 1) and momentum (Basis 2). Using the noise model we discussed in section 4.2.6 we have the following correlation matrices

$$\mathrm{Corr}_1 = \begin{pmatrix} 1015 & 23 & 9 \\ 17 & 947 & 8 \\ 9 & 28 & 1008 \end{pmatrix} \tag{4.81}$$

$$\mathrm{Corr}_2 = \begin{pmatrix} 1053 & 21 & 7 \\ 29 & 1017 & 25 \\ 5 & 15 & 1023 \end{pmatrix} . \tag{4.82}$$

Now to calculate the main quantity of our bound $B(\rho)$ in equation (4.48) which enables us to lower bound the Entanglement of Formation, we compute the first term. To compute the sum of correlations on the diagonal which are featured in the first term, we simply take the sum of all the diagonal entries and the divide them by the sum of all entries, e.g. for basis 1, $C_1 = \frac{\mathrm{Diag}(\mathrm{Corr}_1)}{\mathrm{Total}(\mathrm{Corr}_1)}$ gives $C_1 = \frac{2970}{3064} = 0.9693$. For the term that appears in $B(\rho)$ we actually need the correlations in basis 2, namely the ones on the diagonal, this is $C_2 = \frac{\mathrm{Diag}(\mathrm{Corr}_2)}{\mathrm{Total}(\mathrm{Corr}_2)} = 0.9681$.

The second term of the bound is formed by the off-diagonal elements of the correlation matrix for basis 1. We have the four indices (m, n, l, o) each of dimension $d = 3$ together with the restrictions $m \neq n$,$m \neq l$,$l \neq o$,$n \neq o$, which results in a total of 18 terms. For the first term $m = 0, n = 1, l = 1, o = 0$ we get $\sqrt{\langle 0,1| \, \rho \, |0,1\rangle \, \langle 1,0| \, \rho \, |1,0\rangle} = \frac{\sqrt{23 \cdot 17}}{Corr_1} = 0.00645$. Analogously, the second

term $m = 0, n = 1, l = 1, o = 2$ reads $\sqrt{\langle 0,1|\,\rho\,|0,1\rangle\,\langle 1,2|\,\rho\,|1,2\rangle} = \frac{\sqrt{23\cdot 8}}{N} = 0.0044$, the third term $m = 0, n = 1, l = 2, o = 0$ evaluates to $\sqrt{\langle 0,1|\,\rho\,|0,1\rangle\,\langle 2,0|\,\rho\,|2,0\rangle} = \frac{\sqrt{13\cdot 9}}{N} = 0.00470$ and so on and so forth. Thus summing up all the terms we get that

$$M_1 = \sum_{\substack{m\neq n, m\neq l \\ l\neq o, n\neq o}} \sqrt{\langle v_m^1 v_n^1|\rho|v_m^1 v_n^1\rangle\langle v_l^1 v_o^1|\rho|v_l^1 v_o^1\rangle} = 0.0852\,. \tag{4.83}$$

The last missing piece is the sum of square roots of the off-diagonal elements defined by the two indices i and j such that restriction $i \neq j$ holds. Again we use the correlation matrix of basis 1. For the first case $i = 1$ and $j = 1$ we have $\sqrt{\langle 0,1|\,\rho\,|0,1\rangle\,\langle 1,0|\,\rho\,|1,0\rangle} = \frac{\sqrt{55\cdot 56}}{N} = 0.00645$, the second case $i = 0$, $j = 2$ we have $\sqrt{\langle 0,2|\,\rho\,|0,2\rangle\,\langle 2,0|\,\rho\,|2,0\rangle} = \frac{\sqrt{9\cdot 9}}{N} = 0.00294$ and so on. In the end

$$M_2 = \sum_{i\neq j} \sqrt{\langle v_i^1 v_j^1|\rho|v_i^1 v_j^1\rangle\langle v_j^1 v_i^1|\rho|v_j^1 v_i^1\rangle} = 0.02856\,. \tag{4.84}$$

Putting together all the pieces we get that

$$B(\rho) = \sqrt{\frac{2}{d(d-1)}}\,(d\cdot C_2 - 1 - M_1 - M_2) \tag{4.85}$$

$$= \sqrt{\frac{1}{3}}\,(3\cdot 0.9681 - 1 - 0.0852 - 0.02856) = 1.0338\,. \tag{4.86}$$

Now we can use the value that we got for $B(\rho)$ to finally lower bound the Entanglement of Formation through (4.49). We find that

$$E_{oF}(\rho) \geq -\log(1 - \frac{B(\rho)^2}{2}) = 1.1\,. \tag{4.87}$$

We have now proven that a state, whose measurement results lead to statistics like the ones given in 4.81 and 4.82, has to contain at least 1.1 bits of nonlocal information. This is not the only the only implication we can derive from this. As the maximum amount of e-bits that a state of dimension $d = 2$ can contain is equal to 1, we have that a value of 1.1 is only possible if the underlying state has at least dimension $d = 3$. In general we can lower bound the entanglement dimensionality, which is equal to the minimal possible Schmidt number as

$$D \geq \lceil 2^{E_{oF}(\rho)} \rceil\,. \tag{4.88}$$

For our example here this means that the Schmidt number of the state is at least

$$D \geq \lceil 2.4 \rceil = 3\,. \tag{4.89}$$

This method of certifying the dimensionality of entanglement has been generalised to any two bases in [10].

4.2.8 The multipartite case

In the bipartite case the criteria that we derived did not require any knowledge of the phase relation between the bases. The lower bound that we derived was genuine, it had no assumption on the underlying state. In the multipartite case this proves to be very hard. Although one can certify all entanglement contained for specific pure multipartite state using just two MUBs, it cannot be done in general. With growing number of systems considered there is exists a multitude of possible correlation structure, giving rise to genuine multipartite entanglement (GME), making the construction of a single universal criterion virtually impossible. One of the big disadvantages that we have in the multipartite case, is that in general there exists nothing similar to the Schmidt decomposition. In principle method can be tailored to any GME state and especially to known states of a specific dimensionality structure [161], for the sake of simplicity we will focus on n-partite d-dimensional GHZ states though. As mentioned they can be readily produced in optical setups [162].

Another difference to the bipartite case is that we cannot ignore the phase relation between the MUBs anymore. Therefore we assume the following relations for the rest of this section

$$|\tilde{i}_k\rangle := \sum_{m=0}^{d-1} \omega^{\tilde{i}m} |m\rangle \,, \tag{4.90}$$

where $\omega := e^{\frac{2\pi i}{d}}$. Furthermore we have to define the functions

$$s_\alpha := \sum_{j=1}^{n} i_j \mod d. \tag{4.91}$$

and

$$f_\alpha := \begin{cases} 1 & \text{if } 0 \leq s_\alpha \leq \left\lfloor \frac{d}{4} \right\rfloor \\ -1 & \text{if } \left\lceil \frac{d}{4} \right\rceil \leq s_\alpha \leq \left\lfloor \frac{3d}{4} \right\rfloor \\ 1 & \text{if } \left\lceil \frac{3d}{4} \right\rceil \leq s_\alpha \leq d-1 \end{cases} \tag{4.92}$$

in order to get to the first quantity of interest. Analogously to the bipartite case where we had C_m, which was a linear combination of diagonal density matrix elements, here we define

$$C_{n,d} := \sum_\alpha f_\alpha \langle \tilde{k}_\alpha | \rho | \tilde{k}_\alpha \rangle \tag{4.93}$$

where $\alpha = i_1, \ldots, i_n$ denotes a multi-index with $i_1, \ldots, i_n \in \{0, \ldots, d-1\}$. We now take p_l to be the number of combinations for which $s_l = l$,

$$\frac{1}{\xi} := \frac{1}{2d^n} \sum_{l=0}^{d-1} p_l |\operatorname{Re} e(\omega^l)| \,, \tag{4.94}$$

and let

$$g = 1 - \frac{2p_0}{d^n} \,. \tag{4.95}$$

This allows us to input the phase relations of equation (4.90) into $C_{n,d}$, such that we arrive at

$$C_{n,d} = -g + \frac{1}{\xi} \sum_{\gamma} \operatorname{Re}\langle k_\alpha|\rho|k_\beta\rangle \tag{4.96}$$

where $\gamma := \{\alpha, \beta | i \in \alpha, k \in \beta : \langle i_j|k_j\rangle = 0 \; \forall \; 1 \leq j \leq n\}$.

The strategy will be the same as in the bipartite case. We will try to relate $C_{n,d}$ to lower bounds on concurrences for multipartite systems [133, 134, 135, 144, 145], in particular the following

$$C_{GME} \geq \sqrt{\frac{2}{d(d-1)}} \left(\sum_{j=0}^{d-1} \sum_{i \neq j} |\langle i|^{\otimes n} \rho |j\rangle^{\otimes n}| - \sum_{\kappa} P_{ij}^{\kappa} \right) \tag{4.97}$$

where

$$P_{ij}^{\kappa} = \sqrt{\langle i|^{\otimes n} \langle j|^{\otimes n} \Pi_\kappa \rho^{\otimes 2} \Pi_\kappa |i\rangle^{\otimes n} |j\rangle^{\otimes n}} \,. \tag{4.98}$$

For this purpose let us have a look at a the following quantity

$$\overline{C}_{n,d} := \xi C_{n,d} - \sum_{j=0}^{d-1} \sum_{i<j} |\langle i|^{\otimes n} \rho |j\rangle^{\otimes n}| - g. \tag{4.99}$$

which allows us to establish a lower bound for C_{GME} of the form

$$C_{GME} \geq \sqrt{\frac{2}{d(d-1)}} \left(\xi C_{n,d} - \overline{C}_{n,d} - \sum_{\kappa} P_{ij}^{\kappa} \right) \,. \tag{4.100}$$

Now we can apply the Cauchy-Schwarz inequality to arrive at our final result

$$C_{GME} \geq \sqrt{\frac{2}{d(d-1)}} \left[\xi \left(C_{n,d} - \sum_{\gamma'} \sqrt{\langle k_\alpha|\rho|k_\alpha\rangle \langle k_\beta|\rho|k_\beta\rangle} \right) + g - \sum_{\kappa} P_{ij}^{\kappa} \right] \,. \tag{4.101}$$

which holds for $\gamma' := \gamma / \{\alpha, \beta | i \in \alpha, k \in \beta : \langle i_j|k_j\rangle = 0 \;,\; i_j = i_h \;,\; k_j = k_h \; \forall \; 1 \leq j, h \leq n\}$.

As mentioned before these bounds need to be tailored to specific GME states in order to properly work. In the next section we will look at explicit examples of this bound in the tripartite case for the GHZ state as mentioned in beginning of the section.

4.2.9 Explicit tripartite examples

This section is devoted to presenting the inequalities that result from the multipartite bound explicitly. The tripartite scenarios, i.e. $n = 3$, will only be presented for qubits and qutrits, $d = 2$ and $d = 3$ respectively.

Qubits

The explicit version of the bound for the genuine multipartite concurrence [144, 145] reads as

$$\begin{aligned} C_{GME} \geq B_{GME}(\rho) =& 2\big(\mathrm{Re}\langle 111|\rho|000\rangle - \sqrt{\langle 001|\rho|001\rangle\langle 110|\rho|110\rangle} \\ & - \sqrt{\langle 010|\rho|010\rangle\langle 101|\rho|101\rangle} - \sqrt{\langle 011|\rho|011\rangle\langle 100|\rho|100\rangle}\big)\,. \end{aligned} \tag{4.102}$$

The quantity $C_{n,d}$ is given as the linear combination of the following expectation values

$$\begin{aligned} C_{3,2} =& \langle +++|\rho|+++\rangle + \langle +--|\rho|+--\rangle + \langle -+-|\rho|-+-\rangle \\ & + \langle --+|\rho|--+\rangle - \langle ++-|\rho|++-\rangle - \langle +-+|\rho|+-+\rangle \\ & - \langle -++|\rho|-++\rangle - \langle ---|\rho|---\rangle \end{aligned} \tag{4.103}$$

$$= 2(\mathrm{Re}\langle 111|\rho|000\rangle + \mathrm{Re}\langle 001|\rho|110\rangle + \mathrm{Re}\langle 010|\rho|101\rangle + \mathrm{Re}\langle 100|\rho|011\rangle) \tag{4.104}$$

which leads to the following bound

$$\begin{aligned} C_{GME} \geq & C_{3,2} - 4\big(\sqrt{\langle 001|\rho|001\rangle\langle 110|\rho|110\rangle} \\ & + \sqrt{\langle 010|\rho|010\rangle\langle 101|\rho|101\rangle} + \sqrt{\langle 011|\rho|011\rangle\langle 100|\rho|100\rangle}\big)\,. \end{aligned} \tag{4.105}$$

This bound is tight for the GHZ state, which it was tailored to. Explicitly this means that

$$B_{GME}(|\psi_{GHZ(3,2)}\rangle\langle\psi_{GHZ(3,2)}|) = C_{GME}(|\psi_{GHZ(3,2)}\rangle\langle\psi_{GHZ(3,2)}|)\,. \tag{4.106}$$

Again we use our white noise model, i.e.

$$\rho_{noise} = p|\psi_{GHZ(3,2)}\rangle\langle\psi_{GHZ(3,2)}| + \frac{1-p}{d^3}\mathbb{1}, \tag{4.107}$$

to find the noise resistance of for this example. We find that the state is still detected to be genuinely multipartite entangled up to a value of $p_{crit} = \frac{3}{5}$.

Qutrits

For the case of qutrits we have the following bound that was shown in [144, 145], written down explicitly it reads

$$
\begin{aligned}
C_{GME} \geq \frac{2}{\sqrt{3}}\big(& \operatorname{Re} e\langle 000|\rho|111\rangle + \operatorname{Re} e\langle 111|\rho|222\rangle + \operatorname{Re} e\langle 222|\rho|000\rangle \\
& - \sqrt{\langle 001|\rho|001\rangle\langle 110|\rho|110\rangle} - \sqrt{\langle 010|\rho|010\rangle\langle 101|\rho|101\rangle} - \sqrt{\langle 100|\rho|100\rangle\langle 011|\rho|011\rangle} \\
& - \sqrt{\langle 112|\rho|112\rangle\langle 221|\rho|221\rangle} - \sqrt{\langle 121|\rho|121\rangle\langle 212|\rho|212\rangle} - \sqrt{\langle 122|\rho|122\rangle\langle 211|\rho|211\rangle} \\
& - \sqrt{\langle 002|\rho|002\rangle\langle 220|\rho|220\rangle} - \sqrt{\langle 020|\rho|020\rangle\langle 202|\rho|202\rangle} - \sqrt{\langle 022|\rho|022\rangle\langle 200|\rho|200\rangle}\big) .
\end{aligned} \tag{4.108}
$$

Already here with qutrits the bound amounts to a quite cumbersome expression, namely

$$
\begin{aligned}
C_{3,3} = & \langle\tilde{0}\tilde{0}\tilde{0}|\rho|\tilde{0}\tilde{0}\tilde{0}\rangle + \langle\tilde{1}\tilde{1}\tilde{1}|\rho|\tilde{1}\tilde{1}\tilde{1}\rangle + \langle\tilde{2}\tilde{2}\tilde{2}|\rho|\tilde{2}\tilde{2}\tilde{2}\rangle + \langle\tilde{0}\tilde{1}\tilde{2}|\rho|\tilde{0}\tilde{1}\tilde{2}\rangle + \langle\tilde{1}\tilde{2}\tilde{0}|\rho|\tilde{1}\tilde{2}\tilde{0}\rangle + \langle\tilde{2}\tilde{0}\tilde{1}|\rho|\tilde{2}\tilde{0}\tilde{1}\rangle \\
& + \langle\tilde{1}\tilde{0}\tilde{2}|\rho|\tilde{1}\tilde{0}\tilde{2}\rangle + \langle\tilde{0}\tilde{2}\tilde{1}|\rho|\tilde{0}\tilde{2}\tilde{1}\rangle + \langle\tilde{2}\tilde{1}\tilde{0}|\rho|\tilde{2}\tilde{1}\tilde{0}\rangle - \langle\tilde{0}\tilde{0}\tilde{1}|\rho|\tilde{0}\tilde{0}\tilde{1}\rangle - \langle\tilde{0}\tilde{1}\tilde{0}|\rho|\tilde{0}\tilde{1}\tilde{0}\rangle - \langle\tilde{1}\tilde{0}\tilde{0}|\rho|\tilde{1}\tilde{0}\tilde{0}\rangle \\
& - \langle\tilde{2}\tilde{2}\tilde{0}|\rho|\tilde{2}\tilde{2}\tilde{0}\rangle - \langle\tilde{2}\tilde{0}\tilde{2}|\rho|\tilde{2}\tilde{0}\tilde{2}\rangle - \langle\tilde{0}\tilde{2}\tilde{2}|\rho|\tilde{0}\tilde{2}\tilde{2}\rangle - \langle\tilde{1}\tilde{1}\tilde{2}|\rho|\tilde{1}\tilde{1}\tilde{2}\rangle - \langle\tilde{1}\tilde{2}\tilde{1}|\rho|\tilde{1}\tilde{2}\tilde{1}\rangle - \langle\tilde{2}\tilde{1}\tilde{1}|\rho|\tilde{2}\tilde{1}\tilde{1}\rangle \\
& - \langle\tilde{0}\tilde{0}\tilde{2}|\rho|\tilde{0}\tilde{0}\tilde{2}\rangle - \langle\tilde{0}\tilde{2}\tilde{0}|\rho|\tilde{0}\tilde{2}\tilde{0}\rangle - \langle\tilde{2}\tilde{0}\tilde{0}|\rho|\tilde{2}\tilde{0}\tilde{0}\rangle - \langle\tilde{1}\tilde{1}\tilde{0}|\rho|\tilde{1}\tilde{1}\tilde{0}\rangle - \langle\tilde{1}\tilde{0}\tilde{1}|\rho|\tilde{1}\tilde{0}\tilde{1}\rangle - \langle\tilde{0}\tilde{1}\tilde{1}|\rho|\tilde{0}\tilde{1}\tilde{1}\rangle \\
& - \langle\tilde{2}\tilde{2}\tilde{1}|\rho|\tilde{2}\tilde{2}\tilde{1}\rangle - \langle\tilde{2}\tilde{1}\tilde{2}|\rho|\tilde{2}\tilde{1}\tilde{2}\rangle - \langle\tilde{1}\tilde{2}\tilde{2}|\rho|\tilde{1}\tilde{2}\tilde{2}\rangle
\end{aligned} \tag{4.109}
$$

$$
\begin{aligned}
= \frac{2}{3}\big(& \operatorname{Re}\langle 000|\rho|111\rangle + \operatorname{Re}\langle 000|\rho|222\rangle + \operatorname{Re}\langle 001|\rho|220\rangle + \operatorname{Re}\langle 002|\rho|110\rangle \\
& + \operatorname{Re}\langle 002|\rho|221\rangle + \operatorname{Re}\langle 010|\rho|121\rangle + \operatorname{Re}\langle 010|\rho|202\rangle + \operatorname{Re}\langle 011|\rho|122\rangle \\
& + \operatorname{Re}\langle 011|\rho|200\rangle + \operatorname{Re}\langle 012|\rho|120\rangle + \operatorname{Re}\langle 012|\rho|201\rangle + \operatorname{Re}\langle 020|\rho|101\rangle \\
& + \operatorname{Re}\langle 020|\rho|212\rangle + \operatorname{Re}\langle 021|\rho|102\rangle + \operatorname{Re}\langle 021|\rho|210\rangle + \operatorname{Re}\langle 022|\rho|100\rangle \\
& + \operatorname{Re}\langle 022|\rho|211\rangle + \operatorname{Re}\langle 100|\rho|211\rangle + \operatorname{Re}\langle 101|\rho|212\rangle + \operatorname{Re}\langle 102|\rho|210\rangle \\
& + \operatorname{Re}\langle 110|\rho|221\rangle + \operatorname{Re}\langle 111|\rho|222\rangle + \operatorname{Re}\langle 112|\rho|220\rangle + \operatorname{Re}\langle 120|\rho|201\rangle \\
& + \operatorname{Re}\langle 121|\rho|202\rangle + \operatorname{Re}\langle 122|\rho|200\rangle - \underbrace{\sum_i \langle i|\rho|i\rangle}_{=1}\big) .
\end{aligned} \tag{4.110}
$$

Here we have the resulting bound

$$
\begin{aligned}
C_{GME} \geq \frac{2}{\sqrt{3}}\Big[\Big(&\frac{3}{2}C_{3,3} + 1 - \sqrt{\langle 001|\rho|001\rangle\langle 220|\rho|220\rangle} - \sqrt{\langle 002|\rho|002\rangle\langle 110|\rho|110\rangle} \\
&- \sqrt{\langle 002|\rho|002\rangle\langle 221|\rho|221\rangle} - \sqrt{\langle 010|\rho|010\rangle\langle 121|\rho|121\rangle} - \sqrt{\langle 010|\rho|010\rangle\langle 202|\rho|202\rangle} \\
&- \sqrt{\langle 011|\rho|011\rangle\langle 122|\rho|122\rangle} - \sqrt{\langle 011|\rho|011\rangle\langle 200|\rho|200\rangle} - \sqrt{\langle 012|\rho|012\rangle\langle 120|\rho|120\rangle} \\
&- \sqrt{\langle 012|\rho|012\rangle\langle 201|\rho|201\rangle} - \sqrt{\langle 020|\rho|020\rangle\langle 101|\rho|101\rangle} - \sqrt{\langle 020|\rho|020\rangle\langle 212|\rho|212\rangle} \\
&- \sqrt{\langle 021|\rho|021\rangle\langle 102|\rho|102\rangle} - \sqrt{\langle 021|\rho|021\rangle\langle 210|\rho|210\rangle} - \sqrt{\langle 022|\rho|022\rangle\langle 100|\rho|100\rangle} \\
&- \sqrt{\langle 022|\rho|022\rangle\langle 211|\rho|211\rangle} - \sqrt{\langle 100|\rho|100\rangle\langle 211|\rho|211\rangle} - \sqrt{\langle 101|\rho|101\rangle\langle 212|\rho|212\rangle} \\
&- \sqrt{\langle 102|\rho|102\rangle\langle 210|\rho|210\rangle} - \sqrt{\langle 110|\rho|110\rangle\langle 221|\rho|221\rangle} - \sqrt{\langle 112|\rho|112\rangle\langle 220|\rho|220\rangle} \\
&- \sqrt{\langle 120|\rho|120\rangle\langle 201|\rho|201\rangle} - \sqrt{\langle 121|\rho|121\rangle\langle 202|\rho|202\rangle} - \sqrt{\langle 122|\rho|122\rangle\langle 200|\rho|200\rangle}\Big) \\
&- \sqrt{\langle 001|\rho|001\rangle\langle 110|\rho|110\rangle} - \sqrt{\langle 010|\rho|010\rangle\langle 101|\rho|101\rangle} - \sqrt{\langle 100|\rho|100\rangle\langle 011|\rho|011\rangle} \\
&- \sqrt{\langle 112|\rho|112\rangle\langle 221|\rho|221\rangle} - \sqrt{\langle 121|\rho|121\rangle\langle 212|\rho|212\rangle} - \sqrt{\langle 122|\rho|122\rangle\langle 211|\rho|211\rangle} \\
&- \sqrt{\langle 002|\rho|002\rangle\langle 220|\rho|220\rangle} - \sqrt{\langle 020|\rho|020\rangle\langle 202|\rho|202\rangle} - \sqrt{\langle 022|\rho|022\rangle\langle 200|\rho|200\rangle}\Big] .
\end{aligned}
\tag{4.111}
$$

This bound is as the first one corresponding to the GHZ state, in this case $|\psi_{GHZ(3,3)}\rangle$. For the white noise model

$$
\rho_{noise} = p|\psi_{GHZ(3,3)}\rangle\langle\psi_{GHZ(3,3)}| + \frac{1-p}{d^3}\mathbb{1} \tag{4.112}
$$

we find a resistance up to $p_{crit} = \frac{32}{59}$. Such a value is attainable in modern quantum optics experiments [162].

Chapter 5

Thermodynamics & Information

Thermodynamics stands out among all the theories that we consider valid within the physics. First, it survived throughout the three centuries since being established without major changes. Second, it was created hands-on, not in order to demystify the world but rather to exploit it more efficiently. Third, it is somehow scale invariant, as far as we know there is no region of validity (except that it has a probabilistic character). Fourth, it is independent of the physical carriers, it doesn't care about what happens on the particular systems on the microscale, but rather gives us a picture on an abstract macroscopic level.

Information theory shares all these properties with Thermodynamics, except the first one. As it was not created as physical theory and later became generalised by quantum information theory. This similarity may be the reason why the interaction between these two theories has been a fruitful one during the last decades. Information theoretic tools where used to strengthen and elucidate the very basis of thermodynamics, as for example the maximum entropy principle. Also the other way around thermodynamics has contributed to understand information theoretic tasks, e.g. through Landauer's principle. Together, they managed to resolve longstanding paradoxes in physics like Maxwell's demon.

Before we continue the discussion, let us state the laws of thermodynamics. As there exists various versions of the different laws and also discussion on whether their number can be reduced, we will introduce them in an informal way.

Zeroth law If a system is in thermal equilibrium with each of two other systems, then these two other systems are in thermal equilibrium with each other.

First law The change of internal energy of a system is equal to the difference between the accumulated heat and the work done.

Second law Heat cannot flow from a colder to a warmer system without any other associated change.

Third law The entropy of a system approaches a constant when approaching absolute zero temperature.

These laws have the following direct implications. The zeroth law implies that we can define the concept of temperature. The first law is equivalent to prohibiting the existence of perpetual

motion machines of the first kind, i.e. machines that produce work without the input of energy. The second law is equivalent to the statement that perpetual motion machines of the second kind, i.e. machines whose sole purpose is to transform heat into work. The third law implies that it is impossible to reach a temperature of absolute zero in a finite amount of operations.

We see that thermodynamics and information theory not only share important figures of merit as for example entropy. They both work very similarly for us, in the sense that they describe the limits of nature on a very abstract level. One of the questions this chapter deals with, is what happens when also quantum theory is taken into account.

Our understanding of this interaction is far from complete and research on this topics has kept scientists busy for the last years. In the next sections we will review the interaction between quantum information theory and thermodynamics. We will end this chapter with thoughts on how to go a step further.

5.1 Demons & Principles

The genesis of thermodynamics is intimately connected with the concept of work extraction. As mentioned, humanity was already using the heat flow between systems of different temperature to gain mechanical work before a complete and consistent theory of such was developed. While so far all who have tried to build a perpetual motion machines failed, back in the 19th century James Clerk Maxwell pointed out a fictional entity that could break the second law [163]. This entity nowadays known as Maxwell's demon, he introduced in the following way:

> ...if we conceive a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are still as essentially finite as our own, would be able to do what is at present impossible to us. For we have seen that the molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics. [164]

This means that the demon has a container filled with a gas at a certain temperature. The container is separated into two parts and the separating wall has a little frictionless trap door that can be operated by the demon. Furthermore, the demon has the ability to see and calculate the movement of all the individual molecules in the container. Now such an entity could act in such a way, that whenever a molecule with a velocity higher than the average comes from left and whenever a molecule with a velocity smaller than the average comes from right, it opens the trap door. Otherwise it keeps the door shut. In principle that would lead to the peculiar situation, that right side of the container will increase its temperature (as the average velocity of the molecules within

increases) and the left side will decrease its temperature (as the average velocity of the molecules within decreases), thus establishing a heat flow from cold to warm. Such an entity could therefore violate the second law of thermodynamics and would be able to construct a perpetual motion machine.

The resolution of this paradox needed various steps. The first of these steps goes in importance far beyond resolving the apparent violation of the second law, it is the introduction of Landauer's principle. In 1961 Rolf Landauer, at that time working for IBM, wrote an article with the title 'Irreversibility and Heat Generation in the Computing Process' [165]. In this article, he argued that there is a fundamental thermodynamic cost associated to the erasure of information.

Theorem 5.1.1 (Landauer's principle)**.** *The minimum amount of heat dissipation required to erase one bit of information is*

$$k_B\, T\, ln2\,. \tag{5.1}$$

For proofs see [165, 221]. This remarkable observation is nowadays widely accepted as a natural law. Among other things, it led to the introduction of reversible computing in order to circumvent it. As the traditional architecture for computers involved logical operation such as the AND operation, which in its action erases some information, scientists developed several schemes that allow for a completely reversible way of computing, therefore showing that in principle computation does not have any intrinsic cost [167]. One of these scientists was Charles Bennett, who was also the one resolved the paradox brought about by Maxwell's demon.

Bennett pointed out that the demon must have some sort of memory where all the steps he does are recorded and that it will eventually run out of space, therefore having to erase some of this information. Due to this process of erasure the demon has to expend at least as much work as it has gained through sorting the particles [168].

5.2 Quantum Thermodynamics

Thermodynamics began with the advent of steam engines and the need of an accurate description of such. Now we are entering the age of quantum machines, a result of the fruitful interaction of quantum information theory and thermodynamics. This interaction resulted in a field of research we can call quantum thermodynamics. At first such an approach might seem unreasonable but it turns out that there exists an intricate connection between the main figures of merit. Correlation in the case of quantum information theory and Work in the case of thermodynamics. In this section we will discuss this relationship and also quantum machines.

Let us begin by defining the most important quantities and concepts that will be used in the following.

Definition 5.2.1 (Average Energy)**.** The average (internal) energy of a system in state $\rho \in \mathcal{H}$ with Hamiltonian H is given by

$$E(\rho) := \langle H \rangle = \mathrm{Tr}\,(H\rho)\,. \tag{5.2}$$

Operations on our system that preserve the (internal) energy of the state are called energy-preserving.

Definition 5.2.2 (Energy-preserving unitaries)**.** Given a state $\rho \in \mathcal{H}$ and a Hamiltonian H we call a unitary U (on average) energy-preserving, iff

$$E(\rho) = \mathrm{Tr}\left(H\rho\right) = \mathrm{Tr}\left(H\left(U\rho U^\dagger\right)\right) = E(U\rho U^\dagger) \tag{5.3}$$

For the case that the unitary operation U commutes with the Hamiltonian H, i.e. $[U, H] = 0$, we directly get that the (internal) energy of a system cannot be changed through the cyclicity of the trace.

To avoid ambiguity between entropies that we denoted with the symbol H and the Hamiltonians that we assigned the same symbol, let us redefine entropy in the context of quantum thermodynamics. We will do so by choosing a definition for the von Neumann quantities in more physical fashion, using the natural logarithm $\ln := \log_e$.

Definition 5.2.3 (Entropy)**.** The entropy of a system in state $\rho \in \mathcal{H}$ is given by

$$S(\rho) := -\mathrm{Tr}\left(\rho \ln(\rho)\right) . \tag{5.4}$$

Analogously we will define the mutual information in this context.

Definition 5.2.4 (Mutual Information)**.** The mutual information between subsystem A and B of a system in state $\rho_{AB} \in \mathcal{H}_{AB}$ is given by

$$I(A:B)_\rho := I(\rho_{AB}) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{5.5}$$

As before we will use the mutual information in order to quantify the correlation between two subsystems A and B. Now that we know how to calculate the (internal) energy and the entropy of a state in this context, we can also compute the free energy of a system.

Definition 5.2.5 (Free Energy)**.** The free energy of a system in state $\rho \in \mathcal{H}$ with Hamiltonian H is given by

$$F(\rho) := E(\rho) - T\,S(\rho), \tag{5.6}$$

where is T is the temperature.

There exists a state that is special in many regards, as it fulfills the maximum entropy principle, it is called the Gibbs state [213].

Definition 5.2.6 (Gibbs state)**.** Given a Hamiltonian H the Gibbs state is defined as

$$\tau(\beta) := \mathcal{Z}^{-1} e^{\beta H} \tag{5.7}$$

where $\beta = \frac{1}{T}$ is the inverse temperature and $\mathcal{Z} = \mathrm{Tr}(e^{\beta H})$ is the partition function.

As discussed, entanglement is an important resource when it quantum information processing. Specifically Entanglement theory can be formulated as a resource theory. The core of resource theories are what is called free resources and free operations. In the case of Entanglement theory these are separable states and LOCC (local operations and classical communication). In quantum thermodynamics the free resources are Gibbs states and the free operations depend on what model of thermodynamics is chosen [169]. This makes sense also from a everyday standpoint, as our environment is filled with systems that have a well-defined temperature. Bringing systems into thermal contact will in general lead to a thermalisation of both. In the previous chapter we had the Entanglement of Formation (EoF), a quantity specifying a rate of conversion between the resource and the target state. It can be shown that EoF, in case one considers Entanglement theory as a resource theory, constitutes a monotone with respect to the free operations, i.e. LOCC. In the case of quantum thermodynamics the free energy also constitutes a monotone, taking thermal operations to be the free operations [214].

Definition 5.2.7 (Thermal operations)**.** We call thermal operations, all operations that can be written as

$$\mathcal{T}(\rho_A) = \mathrm{Tr}_{\tilde{A}} \left(U_{AB} \left(\rho_A \otimes \tau(\beta) \right) U_{AB}^{\dagger} \right) \tag{5.8}$$

where $\tilde{A}$ is any partition of system AB and U_{AB} is energy-preserving.

As it is still an ongoing discussion in the community and an active field of research to find a well-suited definition for the notion of work in quantum thermodynamics [169]. We will use a pragmatic approach to this matter and talk about the work cost of unitaries, which is the difference in energy between the final and the initial state.

Definition 5.2.8 (Work cost)**.** Given a state $\rho \in \mathcal{H}$, a Hamiltonian H and a unitary transformation U, we have that work cost of the unitary U is given by

$$W(U) := \mathrm{Tr}\left(H \left(U \rho U^{\dagger} \right) - H\rho \right) = E(\rho_{final}) - E(\rho_{initial}) \,. \tag{5.9}$$

5.3 Work and correlations

quantum thermodynamics can give us interesting insights about the connection of information and energy. Above we have seen that the Gibbs state can assumed to be available freely. If we consider multiple systems, that means that we have a product state of Gibbs states as the free resource. In order to create correlations between the systems we will need to invest energy and perform global transformations on the target systems. In this section we want to show that there is a fundamental work cost associated to every bit of correlation [170] and further that every bit of correlations allows us to gain work [171].

5.3.1 Work for correlations

The setup is the following [170]. We have two uncorrelated systems S_1 and S_2 that are thermalised with a heat bath of temperature T. This means that we are given the state of the system $\tau_S = \tau(\beta)_{S_1} \otimes \tau(\beta)_{S_2}$ and of the bath $\tau_B = \tau(\beta)_B$, which are free resources. The overall state is therefore

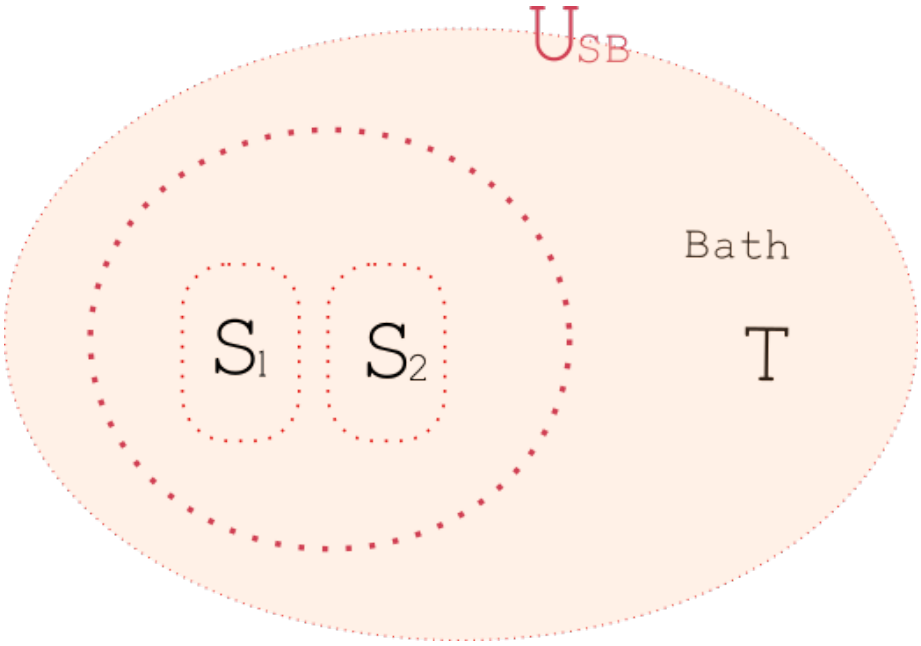

Figure 5.1: Schematic illustration of the setup. [44]

$\tau_{SB} = \tau_S \otimes \tau_B$. Also we assume that the total Hamiltonian has the form $H = H_{S_1} + H_{S_2} + H_B$. In order to correlate the systems S_1 and S_2 we will apply a global unitary U_{SB} and look at its minimum work cost. First, we observe that for any product state, i.e. states that can be written as $\rho_{product} = \rho_A \otimes \rho_B$ we have that

$$I(A:B)_{\rho_{product}} = S(A) + S(B) - \underbrace{S(AB)}_{=S(A)+S(B)} = 0\,. \tag{5.10}$$

We will then proceed to compute the work cost of the unitary U_{SB}, see definition 5.2.8, depending on the increase of correlation between S_1 and S_2, quantified in terms of the mutual information. Let us start by direct computation,

$$W_{corr} = \mathrm{Tr}\left(H\left(U_{SB}\tau_{SB}U_{SB}^\dagger\right) - H\tau_{SB}\right) = \Delta E_S + \Delta E_B\,. \tag{5.11}$$

Note that the overall entropy is left invariant, i.e. $S(\tau_{SB}) = S(U_{SB}\tau_{SB}U_{SB}^\dagger) = S(\rho_{SB})$, such that we can rewrite above expression in terms of the change in free energy

$$
\begin{aligned}
W_{corr} &= \Delta E_S + \Delta E_B && (5.12)\\
&= \Delta F_S + \Delta F_B + T\left(S(\rho_S) + S(\rho_B) - S(\tau_S) - S(\tau_B)\right) && (5.13)\\
&= \Delta F_S + \Delta F_B + T\left(S(\rho_S) + S(\rho_B) - S(\tau_{SB})\right) && (5.14)\\
&= \Delta F_S + \Delta F_B + T\left(S(\rho_S) + S(\rho_B) - S(\rho_{SB})\right) && (5.15)\\
&= \Delta F_S + \Delta F_B + TI(S:B)_\rho\,. && (5.16)
\end{aligned}
$$

We can now split the ΔF_S into the free energy change of the systems S_1 and S_2 and their correlation,

$$\Delta F_S = \Delta F_{S_1} + \Delta F_{S_2} + TI(S_1:S_2)_\rho\,, \tag{5.17}$$

and combine it with the expression (5.16), s.t.

$$W_{corr} = \Delta F_{S_1} + \Delta F_{S_2} + \Delta F_B + TI(S:B)_\rho + TI(S_1:S_2)_\rho\,. \tag{5.18}$$

As all quantities on the right hand side are strictly positive we can assert that,

$$W_{corr} \geq TI(S_1:S_2)_\rho\,, \tag{5.19}$$

This means that ever bit of correlation costs at least $T\,I(S_1:S_2)$ in work. In [170] the authors also show that it is possible to saturate this bound just using a simple set of operations.

5.3.2 Work from correlations

Now we want to look at the inverse question. Now that we have gained the correlations $I(S_1:S_2)$ we want to see what is the maximum amount of work W_{extra} that we can gain from these correlations. So the setup remains the same only that our initial state now is ρ_{corr} with mutual information $I(S_1:S_2)$ and we want to apply a unitary U_{SB} such that we get τ_{SB} as the final state. In [171] the authors show that in this scenario the maximum extractable work is given by

$$W_{extra} \leq TI(S_1:S_2)_\rho\,. \tag{5.20}$$

We observe that we can assign a work value, and therefore an energy value, to every bit of correlation, may it be in terms of creating it or in terms of extracting work from it.

5.4 Quantum machines

As in the macroscopic we want to be able to build and describe machines in the microscopic. These thermal machines are the quantum versions of heat engines or refrigerators. As we are still limited

by the laws of thermodynamics, we have to use at least to heat baths of different temperature in order to make machines that are of interest to us.

The maser invented in the nineteen sixties can actually be seen as the first quantum thermal machine, as it basically works as a heat engine [172]. Overviews on the topic of quantum machines can be found in [173] and in [169].

In general, there are two regimes one can focus on. One is an analysis in terms of the cyclic behaviour of systems in contact with thermal baths and a characterisation in terms of the steady-state that is given by the flows of heat and work in the limit for long times. The other is called the transient regime, is dealing with how the system reaches its stationary state.

Having two heat baths, allows one to consider the first bath as the environment with its ambient temperature, such that the second bath then constitutes an out of equilibrium state with respect to the first bath. The aim is to produce the target quantity (e.g. work, heat or refrigeration) at an optimal rate. In this perspective the quantum engine acts as a catalyst enabling the workings of the machine.

Note that in the near future, quantum machines may become important regarding the facilitation of rapidly emerging technologies such as nano-technology and quantum information processing devices. Thus a complete understanding of the specific behaviour and also the correlations within is of high importance [169].

The smallest quantum engines that are possible go as little as being just a single qutrit or being comprised of 2 qubits [169]. Despite their small size they can still approach the maximal possible efficiency, Carnot efficiency. In the next section the two-qubit heat engine introduced in [211] will be discussed in detail.

5.4.1 Two-qubit heat engine

In this section the two-qubit engine introduced in [211] is reviewed. In the next chapter we will use this engine to power the autonomous quantum clock that will be presented therein. This engine is not limited to this use case, others include refrigeration, heating or providing mechanical work [174, 175].

The cornerstone of the engine are two qubits that are coupled to heat baths of different temperature. One of the qubits, that from now on we will call 'hot', has an energy gap of E_h and is connected to the heat bath with temperature T_h. The other qubit, the 'cold' one, has an energy gap of $E_c < E_h$ and is connected to a heat bath of temperature T_c. Both qubits are coupled to a d-dimensional ladder with equally spaced energy levels, that are spaced with E_w. The ladder is assumed to not interact only with the two qubits it is coupled to and not any heat baths. Thus we have that the following free Hamiltonian for the overall system.

$$H_{free} = \sum_{j=h,c} E_j \left|1\right\rangle_j \left\langle 1\right|_j + \sum_{k=0}^{d-1} kE_w \left|k\right\rangle_w \left\langle k\right|_w , \tag{5.21}$$

where we have introduced the notation that the index $j = h, c$ specifies whether it is the 'hot' or the 'cold' system, the states $\left|k\right\rangle_w$ correspond to the k-th energy level in the ladder system and $\left|1\right\rangle_j$ stands for the excited states of the qubits, respectively.

Our goal is to run the engine as a heat engine, we will therefore introduce the following constraint,

$$E_h = E_c + E_w . \tag{5.22}$$

In this way our machine is enabled to convert the energy of an excited 'hot' qubit to precisely one excitation of the 'cold' qubit and one of the ladder. This constraint means that our engine has to run in resonance. Otherwise it would not be possible to exchange energy between the qubits and the ladder. This also implies that the energy levels, $\left|0\right\rangle_c \left|1\right\rangle_h \left|k\right\rangle_w$ and $\left|1\right\rangle_c \left|0\right\rangle_h \left|k+1\right\rangle_w$, are degenerate.

The following interaction Hamiltonian leads exactly to the desired interplay between the systems

$$H_{\text{int}} = g \sum_{k=0}^{d-1} \left(\left|1\right\rangle_c \left|0\right\rangle_h \left|k+1\right\rangle_w \left\langle 0\right|_c \left\langle 1\right|_h \left\langle k\right|_w + \text{h.c.}\right) , \tag{5.23}$$

where g is the coupling constant and h.c. stands for Hermitian conjugate. As a consequence of the demanded constraint on the energies (5.22), we observe that the engine can even be operated in the weak-coupling regime, which is characterised by the relation $g \ll E_c, E_w$.

This illustrates what the interaction Hamiltonian (5.23) does, but let us also describe it in plain english. Due to the difference in temperatures between the two heat baths we get a heat flow from 'hot' qubit at T_h to the 'cold' qubit at T_c. The constraint (5.22) on the energies that we demanded enables this heat flow. Specifically this implies that we employ one quantum of energy E_h from the 'hot' qubit in order to excite the 'cold' qubit, using up E_c. The leftover energy is transferred to the ladder, where it equals exactly one excitation $E_w = E_h - E_c$. This describes the first term of the the interaction Hamiltonian (5.23), we call it 'forward' process. The Hermitian conjugate part in the interaction Hamiltonian (5.23) does exactly the reverse, it takes one quantum of energy from both the ladder and the 'cold' system in order to excite the 'hot' system. We will call this 'backwards' process.

As we want the engine to deliver work, i.e. raise the energy of the ladder, we want the 'forward' process is more likely than the 'backwards' process. The authors of [211] show that this can be done by choosing the parameters in the following way. If we denote p_1 the probability of occupying the state $\left|0\right\rangle_c \left|1\right\rangle_h$ and p_0 the probability of occupying the state $\left|1\right\rangle_c \left|0\right\rangle_h$, we can add a bias favouring the 'forward' process

$$|0\rangle_c |1\rangle_h |k\rangle_w \to |1\rangle_c |0\rangle_h |k+1\rangle_w \ , \tag{5.24}$$

by simply choosing $p_1 > p_0$.

If we operate the engine in the weak-coupling regime and we observe that both the 'cold' and the 'hot' qubit are very close to the Gibbs states corresponding to the respective temperatures of the heat baths, we will come to the conclusion that the probabilities p_1 and p_0 only depend on the temperatures of the respective heat bath and the energies gaps of the two qubits.

This way we can introduce the bias towards the transition (5.24) by assuming that

$$\frac{E_h}{T_h} < \frac{E_c}{T_c} \, . \tag{5.25}$$

The fact that the interplay of the two coupled qubits with the ladder is ruled by the two states $|0\rangle_c |1\rangle_h$ and $|1\rangle_c |0\rangle_h$ has led the authors of [211] introduce what they call the machine's virtual qubit, with the former two states as the two levels. We can now characterise this virtual qubit by itself. Specifically this means associating a virtual temperature

$$T_v := \frac{E_h - E_c}{\beta_h E_h - \beta_c E_c}, \tag{5.26}$$

with $\beta_{c,h} = 1/T_{c,h}$ and furthermore, a quantity that will be called the population bias, defined as

$$Z_v := \frac{p_0 - p_1}{p_0 + p_1} = \tanh(\beta_v E_w/2) \, . \tag{5.27}$$

Equipped with this newly acquired concept of the virtual qubit and its associated quantities, we can describe the functioning of the engine as follows. If the constraint (5.22) is fulfilled, i.e. we have that the difference between the energy gaps of the 'cold' and the 'hot' qubit is resonant with ladder's energy spacing, we can see the action of the engine as placing a load in thermal contact with the virtual qubit. The load then effectively 'thermalises' with the virtual qubit, whose virtual temperature is given by the population ratio $p_1/p_0 = \mathrm{e}^{-\beta_v E_w}$. If we have a negative virtual temperature T_v, or equivalently a negative population bias Z_v, this 'thermalisation' process makes the load to climb up the ladder. One can show that whenever (5.25) is satisfied, (5.26) and (5.27) will be negative. Note the details thermalisation model ruling the coupling between the different parts of the engine do not affect us here, as we have based our analysis only on the statics of the model.

5.5 Complexity and Emergence

Up to now, we have discussed different entropic measures and information quantifiers. They all, however, make use of events that did not take place [176], often referred to as counterfactuals. There is a measure of information that does not do so, the Kolmogorov complexity. Given an input string, its value is equal to the length of the shortest program that computes the string as an output on a universal Turing machine [178, 179]. This is a very powerful tool, knowing its value would solve a lot of problems for us, e.g. a proper definition for randomness or perfect lossless compression, just to name two. It seems that this quantity might be too powerful for us, as it is provably uncomputable [177] (for large enough strings). Let us look at its definition.

Definition 5.5.1 (Kolmogorov complexity)**.** Given a infinite binary string $\underline{x}$ and a universal Turing machine $\mathcal{U}$, the Kolmogorov complexity is defined as

$$K(\underline{x})_{\mathcal{U}} := \min_{p:\,\mathcal{U}(p)=\underline{x}} |p| \,. \tag{5.28}$$

One can show that this definition is independent of the Turing machine used, as their difference can always be upper bounded by an universal constant c [179], i.e. given universal Turing machines $\mathcal{U}_1$ and $\mathcal{U}_2$ it holds that,

$$|K(\underline{x})_{\mathcal{U}_1} - K(\underline{x})_{\mathcal{U}_2}| \leq c \,. \tag{5.29}$$

This can be seen as follows, given the shortest program to compute a string in one programming language we can always translate it into the shortest program in another programming language without loosing too much. This means that we can drop the subscript by using a fixed Turing machine and all constant overheads. Furthermore, it is possible to define a Kolmogorov complexity for finite strings by encoding them into infinite strings. Take the finite binary string $\underline{a}_{[n]} := (a_1, \ldots, a_n)$ for example, we can turn it into an infinite string $\underline{a} := (a_1, \ldots, a_n, 0, \ldots, 0, \ldots)$, by just adding infinitely many 0 to it. Now, we look at the scaling of K given such a string $\underline{a}$ [179, 180]. We call $\underline{a}$ incompressible if

$$K(\underline{a}) \approx n :\Leftrightarrow \lim_{n\to\infty} \frac{K(\underline{a}_{[n]})}{n} = 1 \tag{5.30}$$

and if

$$K(\underline{a}) \approx 0 :\Leftrightarrow \lim_{n\to\infty} \frac{K(\underline{a}_{[n]})}{n} = 0 \Leftrightarrow K(\underline{a}_{[n]}) = o(n) \tag{5.31}$$

we call it computable, in this sense incompressibility is the extreme form of uncomputability [180]. Now, that we have an idea of what the Kolmogorov complexity looks like, we can try to use it as a descriptive tool for our observations of the world. For this argument we quantise our perception of the world by dichotomic questions, all observations that we can make can therefore be described by a binary string. The Kolmogorov complexity would then give us the length of the

shortest description for any possible observation. In a world comprised of binary strings one can not only describe correlations that are stronger than classical possible, but also do thermodynamics [181].

Furthermore, imagine a steaming cup of coffee and milk being poured into it. At first the liquid is uniformly black, then their is a period in which milk and coffee have not merged yet, sprinkles of black, white and all kind mixtures thereof, in a chaotic dance. Waiting long enough the situation will have settled down and the liquid will have a uniform colour again. While in the world of strings, the situations in the beginning and in the end may seem to be less complex, they are not. Let us use Kolmogorov sufficient statistics in order to describe this example [182]. This allows use to quantify the exploitable 'structure', for this we want that M_k is the smallest set such that $\underline{o} \in M_k$ and $K(M_k) \leq k$ holds for $k \in \mathbb{N}$. Moreover, we define k_0 as the value where the function $\log |M_k|$ starts to have a slope with value -1. In plain english, k_0 would be the point where the string $\underline{o}$ becomes a typical element of the set M_{k_0}, i.e. there is no more structure that we could exploit. Now we have that

$$K(\underline{o}) = k_0 + \log |M_{k_0}| = K(M(\underline{o})) + log|M(\underline{o}))|\,, \tag{5.32}$$

where we have defined $M(\underline{o}) := M_{k_0}$ [199]. $M(\underline{o})$ can be seen as the macrostate that emerged from the underlying microstate $\underline{o}$. We have finally arrived at the point where complexity and emergence meet. Before we continue the discussion, note that $K(\underline{o})$ overall is at least a non-decreasing function according to the second law of thermodynamics that can be derived in this context [181, 199]. Whether it can increase at all, depends on the existence of randomness.

One can think of all our natural laws as short programs that together with the initial conditions serve as an input for a Turing machine, which then outputs predictions. This simplistic view at least holds true for the classical laws of physics, as for quantum systems we still do not know whether there exists an efficient way to simulate them with a Turing machine. In such a case it may come in handy to find sound definitions of quantum models for the Kolmogorov complexity and the Turing machine [183]. Maybe it is even possible to show that even the laws of nature themselves naturally emerge from an underlying structure [184, 185]. Independently of these questions, it would be very interesting to see whether one can find criteria that characterise the phenomenon of emergence. For such questions thermodynamics provides a good testbed. Quantities like temperature emerge from a coarse-graining of the microscopical dynamics. While it is very hard to assign a temperature to single microscopic systems, except the thermal states and qubits (with two values it uniquely fix an exponential distribution), macroscopically it can be easily calculated by an averaging of their velocities. Time, much alike temperature, could itself be an emergent quantity. The problem is, that in general there are infinitely many ways to coarse-grain over microscopic quantities, but not all will lead to physically relevant quantities. Furthermore, ideally we would like to have programs of constant length (natural laws) that accurately describe their behaviour. In general their will most likely be a trade-off between the predictive and the explanatory power for such theories, the question is, if we can find ways to determine under which circumstances we are allowed to trade.

Chapter 6

Time

Time is a mysterious concept. We all perceive its passage, but for science, since its inception, it has been hard to capture it. So this chapter begins with the same quote that stands at the beginning of the chapter of time in [186]. The quote is a comment that Heisenberg wrote in a letter to Pauli in November 1925:

> Your problem of the time sequence plays, of course, a fundamental role, and I had figured out for my own private use (Hausgebrauch) something about it. First I believe that one can distinguish between a coarse and a finer time sequence. If a point in space does not assume in the new theory a definite role or can be formulated only symbolically, then the same is true for an instant of time of an event. But there always will exist a coarse time sequence, like a coarse position in space - that is, within our geometric visualisation one will be able to carry out a coarse description. I think it might be possible that this coarse description is perhaps the only thing that can be demanded from the formalism [of quantum mechanics].[187]

This comment points at the special role that time plays in quantum theory. It is the only parameter that has survived the transition to quantum theory unaltered (speaking in an non-relativistic context). While quantum theory has its roots in the quantisation of nature, the description of time with in it, is not. It is in information theory though, that time is merely a sequence of steps. In quantum theory there is no observable for time. The impossibility of such an observable was famously proven by Pauli [192]. Namely, the introduction of such a time observable would result in an energy operator (Hamiltonian) which has an unbounded spectrum from above and below.

This chapter will take a pragmatic approach to the understanding of time within quantum information theory, namely it will try to describe and analyse our abilities to measure time, as this is also so far the only way we are able to measure time, as a mere sequence of clicks.

So in the following we will discuss time measuring devices, i.e. clocks, and try to find the minimal resources one needs in order to be able to measure time.

6.1 Clocks

Despite the fact that we use quantum systems to construct the most advanced clocks [189, 190, 191], the concept of time itself remains elusive in quantum theory. Since the inception of quantum theory this issue has been studied in various different forms. Specially the connection to energy, which is the time-invariant quantity for closed systems, has brought about fruitful results in terms of limitations on the time needed for quantum system to evolve. These results are now known under name of quantum speed limits [192, 193, 194, 195, 196, 197]. A different direction of studying the notion of time in quantum theory is to promote it to a parameter that is genuinely quantum, rather then continue treating it as classical [198, 199, 200, 201, 202]. This way the evolution of quantum systems is encapsulated in correlations.

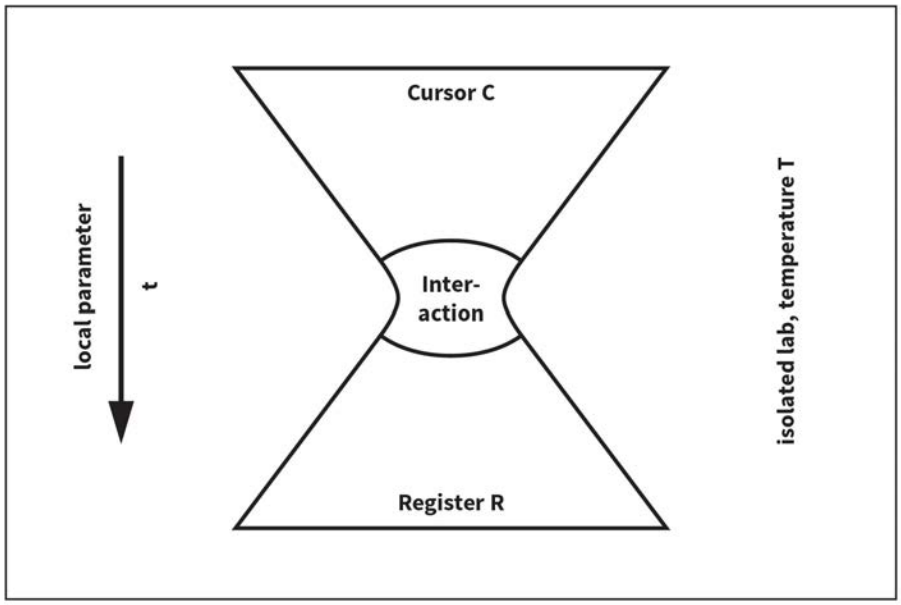

Figure 6.1: Schematic illustration of the Quantum Hourglass [186].

Coming back to the topic of how to measure time, various models for how to use quantum systems for this task have been proposed in the past [203, 204, 205, 206]. All of these models for quantum clocks share the caveat that they need a duly chosen initial state and unitary evolution to function properly. These models in general can be considered as measuring a time interval, therefore having a functionality more comparable to a stopwatch than a clock. The precision of such stopwatches can be related to the properties of the clock, for example its dimension [206]. They, however, rarely specify neither the procedure of preparation nor the measurement. In principle, they can be used as timing devices for the implementation of a given interaction leading to a unitary operation. As they need very accurate timing information themselves for the application of the required unitary transformation, this leads to a sort of circularity in the argument.

So let us have a look what features we want to demand for a device to call it a clock. In contrast to a stopwatch, a clock should be able to continuously give a time reference. It is therefore essential for any complete model of a clock to incorporate the process of information read-out. This leads to the requirement that a clock has to be at least bipartite [186]. See figure 6.1 for the first illustration of such models [186, 207].

In the following we will describe a clock as follows. We will call the first part of the clock, pointer. It is the subsystem to whose dynamics are effectively determined by passage of time and therefore will serve as the reference for the clock. The second part, the register, has the task to store the information obtained on the evolution of the pointer system. This ideally should be done in terms of classical information, allowing an external observer to access it. Summing up, we have the pointer system which is designed such that it produces a sequence of signals which are then recorded by the register as ticks, as it can been seen in figure 6.2.

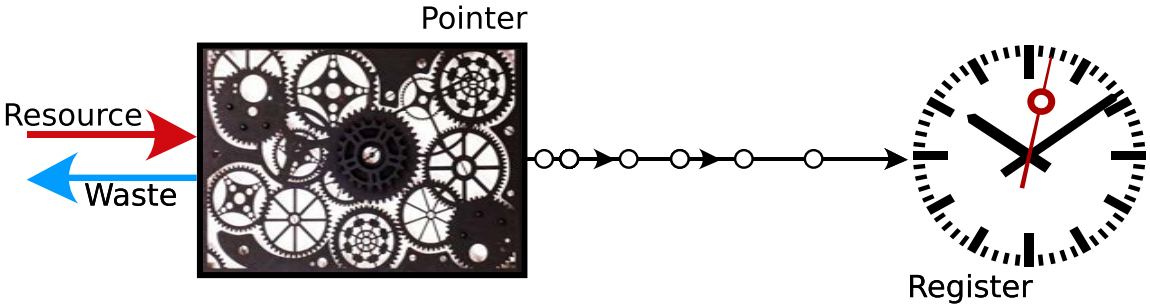

Figure 6.2: A pointer system generates a sequence of events that are recorded by the register [44].

One can observe that the workings of a clock lead to an asymmetric flow of information. This is turn means that this process is irreversible and connects it to the second law of thermodynamics [208], due to the production of entropy associated with irreversible processes. We can thus expect that a system that is suitable to function as a clock will be connected to its tendency to produce entropy. The exact link between these two properties has however not been shown yet.

The question that we are interested in first, is even more general than the asking for the relationship between the ability of a system to act as a clock and its propensity to produce entropy. It is the question about the minimal resources that one needs in order to maintain a clock running. We will answer this question in the next sections by presenting a model for a self-contained time measuring device that works without the need for external control or timing information. Such clocks we will call autonomous clocks. These clocks needs to contain an isolated system whose evolution is determined by a time-independent Hamiltonian and the resources used to power the clock should themselves not require any type of timing information to be prepared.

In particular we will analyse autonomous clocks that a driven by the heat flow between two thermal reservoirs of different temperature. This constitutes the minimal non-equilibrium resource. Specifically, we will discuss the physics of the mechanisms the make the clock work, this includes its

initialisation and power supply. To achieve this we will exploit the results on autonomous thermal machine [209, 210, 211] and make use of thesee machines for the purpose of producing a series of regular ticks. This enables us to show that the production of entropy of the clock ultimately bounds its performance.

This leads to the question about the figures of merit we use to quantify the performance of a clock. Here two quantities have turned out to be useful [186, 207, 10]. Namely the following two:

Resolution — quantifies how frequently a clock ticks

Accuracy — quantifies how many ticks a clock provides before its uncertainty becomes greater than the average time between ticks

We will show in the coming sections that only large enough rate of entropy production allows one to reach given levels of resolution and accuracy simultaneously. If the rate is not high enough a trade-off between both of them appears. This means that one has to sacrifice some resolution in order to attain the desired level of accuracy, or vice versa. Moreover, we will see that even in a regime where the resolution of a clock is arbitrarily low, the entropy production still upper bounds its accuracy. This hints at a fundamental connection between entropy production and the clock's arrow of time, as the relevant entropy production is solely that of the pointer system itself, ignoring other processes as for example measurement or erasure of the register. This point will be demonstrated by considering an explicit model for an autonomous clock and fully computing its dynamics. It is then conjectured on the basis of general thermodynamic arguments, that such trade-offs can be seen in any implementation of an autonomous clock.

6.2 Autonomous Quantum Clocks

As mentioned above, our goal is to find fundamental limitations of quantum clocks. Therefore, we consider clocks that are complete and self-contained, i.e. autonomous clocks. In order to guarantee fair bookkeeping, our devices should not be in need of any resource that itself requires timing information. This assumption allows us to carefully account for all resources that needed for the task of measuring time. In this section we will discuss autonomous clocks in general, while the next section will introduce a particular model of such.

A general feature of autonomous clocks is that their evolution is given by a time-independent Hamiltonian such that in the course of this evolution a sequence of ticks is produced that can be recorded by the register. This is illustrated in figure 6.2. The process of transferring the information of the pointer system to register should ideally be irreversible. Moreover, this process should ideally happen spontaneously, ensuring that no external intervention or time-dependent coupling is present. As we want that the flow of information is uni-directional, we need to have that the probability of the spontaneous process happening is larger compared to the one of its time-reverse. This implies that the free energy of the pointer system will decrease and thus an autonomous clock will need some source of free energy in order continue producing ticks. This source of free energy is necessary to keep the clock out of equilibrium.

In theory any non-equilibrium quantum system has the ability to make the free energy available that is necessary to run a clock. Unfortunately there is a large class of non-equilibrium states that we have to exclude, as their preparation already requires the presence of a clock. The timing information that is needed to prepare such resource states could for example be the period that a resonant driving field has to be applied. As already mentioned such resources are excluded from our model, in order to precent circularity in the arguments and to be able to carefully account for all resources that are consumed. Although it is of no importance to the model, in practice it would desirable for the used resources to be available abundantly in nature or to have an easy way to generate them.

One can argue that the minimal non-equilibrium resource that fulfils all above requirements is given by two heat baths of different temperatures. In practice the presence of a single heat bath is unavoidable, as this is given by the ambient temperature of the environment, which we will denote T_c. Moreover, is it possible to deterministically prepare a second heat bath at a temperature $T_h > T_c$ not possessing any knowledge about the internal structure of the reservoir nor using any operations that would need timing information. The reason behind this, is that any generic quantum system (excluding not-integrable and many-body localised systems) will equilibrate towards the thermal state [212], as it is given through the condition of maximal entropy [213].

Moreover, we have that any other potential resource for running our clock would have lower entropy at the same energies and therefore require additional knowledge or control to in order to be prepared. Thus, an equilibrated resource that features an average energy content that is higher than that of the environment, is the minimal resource that can drive the clock out of equilibrium. Our explicit model for an autonomous clock will hence be a quantum clock that is powered by two thermal baths. We will conduct a quantitative analysis of such a clock in the next sections. Nevertheless it needs to emphasised the concept of the autonomous clock is more general and the explicit model with which we will deal in the following. There exist various different scenarios and also resource states that could be considered as well [215].

6.2.1 Minimal thermal clock model

In this section we will consider a explicit model of a autonomous quantum clock wherein the pointer system will be powered by the flow of heat between to thermal baths of different temperatures. In particular we will use the quantum heat engine introduced in [211], that we have already discussed in detail in section 5.4.1, in order to drive the pointer.

To recap, the machine is constructed by taking two qubits and independently couple them to two different thermal baths. We have that the first qubit with the energy gap E_h is in thermal contact to the hot heat bath at temperature T_h, while the second qubit with the energery gap $E_c < E_h$ is in thermal contact to the cold heat bath with temperature $T_c < T_h$. We will assume that T_c is the ambient temperature of the environment. To end this recap of the machine's design, it is pointed out that the output of engine in terms of work, drives a load up a ladder, that consists of d equally spaced energy levels, where the spacing between the levels is given by $E_w = E_h - E_c$. Figure 6.4 illustrates this setup.

The gradient in temperature between the two thermal reservoirs induces a heat flow. This current runs from the hot to the cold qubit and provides work to load, thereby generating an upwards

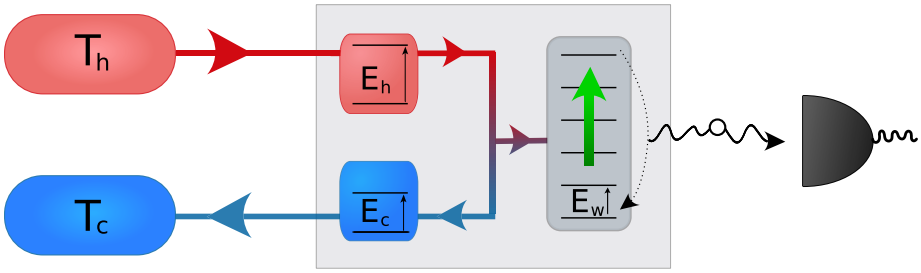

Figure 6.3: We consider a pointer comprising a two-qubit heat engine that drives a thermally isolated load up a ladder, whose highest-energy state undergoes radiative decay back to the ground state. Photons are thus repeatedly emitted and registered by a photodetector as ticks of the clock.[44]

movement on the ladder. As mentioned in section 5.4.1, the machine's working can be understood as a resonant interaction between the load and a virtual qubit [211]. This virtual qubit is comprised of two special states of the engine, namely the ones that couple to the ladder.

Working in a regime where the coupling between engine and ladder is small, we get a thermal distribution of the populations in the two different states of the virtual qubit according to the virtual temperature

$$T_v = \frac{E_h - E_c}{\beta_h E_h - \beta_c E_c}\,, \tag{6.1}$$

where again we have that $\beta_{c,h} = 1/T_{c,h}$. The ratio between the population of the virtual qubit's states is given by

$$\frac{p_1}{p_0} = \mathrm{e}^{-\beta_v E_w}\,, \tag{6.2}$$

with p_0 denoting the population of the higher energy state, p_0 denoting the population of the lower energy state and $\beta_v = 1/T_v$. Whenever the virtual qubit features a negative temperature, called a population inversion, it causes the load to 'climb' up the ladder. In this perspective the load 'thermalises' with the virtual qubit at the given virtual temperature. One can express the virtual temperature in terms of the population of its energy levels,

$$Z_v = \frac{p_0 - p_1}{p_0 + p_1} = \tanh(\beta_v E_w/2). \tag{6.3}$$

We will call this quantity the population bias of the virtual qubit. It will later play an essential role in the characterisation of the performance of our clock.

As promised we will specify the interaction between the pointer system and the register. We start to complete the picture by assuming that the top level of the ladder is unstable and can decay to the ground state, thereby emitting a photon with the energy $E_\gamma = (d-1)E_w$. The clock's ticks

are the detection events in the register that are caused by the arriving photons. Due to the presence of the decay channel we also have the possibility of the inverse process happening. To circumvent that, we assume that the background temperature is very small compared to the photons energy, i.e. $T_c \ll E_\gamma$. This renders the probability of such an inverse process happening negligible.

Figure 6.4: Creative illustration of the minimal thermal clock [44].

To sum up, we have that the heat current in the engine steers the load up the ladder. When it finally reaches the top level, a photon is emitted through the decay to the ground state. The process then starts anew, through its repetition it creates a series of photons which are witnessed by the register, hence making the clock tick. Note that the decay leading to the emission of the photon depends on how the ladder's energy evolves, as this evolution is probabilistic it in turn results in a stochastic sequence of ticks. The resulting distribution of ticks will depend on the 'height' of the ladder, i.e. its dimension d, and the population bias of the virtual qubit Z_v, proportional to the driving force provided by the engine.

Depending on the actual values, the load can become 'smeared out' over the ladder, which means that there is a broad probability distribution over the energy levels, hence resulting in a slowly and irregularly ticking clock. On the contrary, if we a nearly complete population inversion, i.e. $Z_v \to -1$, the probability for the load moving downwards on the ladder becomes very small and we get a clock that ticks faster in more regular time intervals. Lastly, figure 6.4 shows the workings of the minimal thermal clock from a original perspective.

6.2.2 Performance of the clock

Now that we have established the model of the minimal thermal clock, we want to evaluate its performance. Part of this evaluation is the analysis of the power consumption, which is fundamentally

related to the entropy production of the clock. For the clock to tick, the engine must drive the load up the ladder and in the process dissipates energy into the cold bath. Our aim is to relate the dissipated energy to the performance of the clock. We therefore look at the amount of heat that is dissipated into the cold reservoir for each tick, namely

$$Q_c = (d-1)E_c\,. \tag{6.4}$$

This quantity is not equal to the total amount of heat supplied to engine, as this amounts to $Q_h = (d-1)E_h$. A large share of the total energy used for the generation of one tick is carried away by the photon. In theory one can think of a way such that the energy of the photon E_γ could be recycled after its detection and provided to the hot reservoir again. Hence, leaving us with the fundamental cost for one tick in form of the dissipated heat Q_c, in (6.4). This leads to an irreversible entropy production of at least $\beta_c Q_c$ per tick.

As mentioned above we will us the figures of merit, resolution and accuracy, to quantify the performance of the clock. While the resolution is given by the average amount of ticks provided by the clock per unit time, the accuracy is characterised by number of ticks the clock can provide before a tick has an uncertainty equal to the average time interval between the ticks. In theory one could also choose an operational definition for the accuracy, independent of background time [207]. It is however conjectured that in certain limits both of this definitions coincide [216].

So let us now define our figures of merit in a more rigorous manner. Given the distribution of waiting times of the ticks of a clock, we get the following definitions for resolution and accuracy.

Definition 6.2.1 (Resolution)**.** The resolution of a clock is defined as

$$\nu_{\text{tick}} := \frac{1}{t_{\text{tick}}}\,, \tag{6.5}$$

where t_{tick} corresponds to the average waiting time between two consecutive ticks.

Thus the resolution is equal to the average amount of ticks that the clock provides per second. We will now define the accuracy of a clock in the following way.

Definition 6.2.2 (Accuracy)**.** The accuracy of a clock is defined as

$$N := \left(\frac{t_{\text{tick}}}{\Delta t_{\text{tick}}}\right)^2\,, \tag{6.6}$$

where Δt_{tick} corresponds to the standard deviation of the distribution of ticks and t_{tick} to the average waiting time between two consecutive ones.

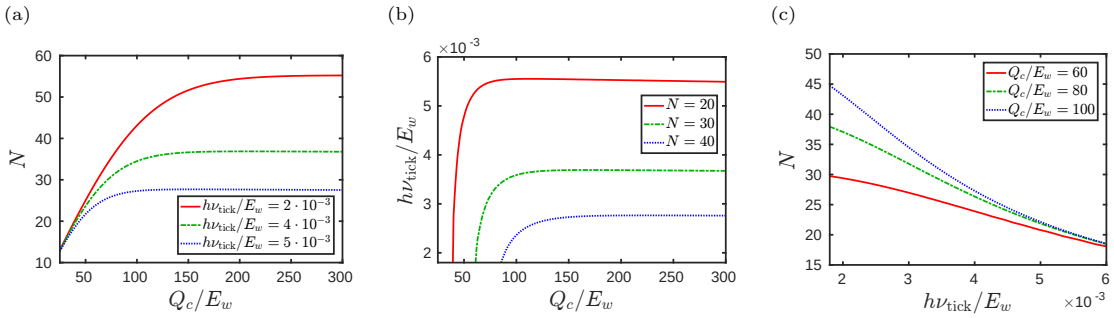

Figure 6.5: Illustration of the fundamental trade-off between the dissipated heat and the achievable accuracy and resolution. (a) Accuracy N as a function of dissipated heat per tick Q_c, for various values of the resolution $\nu_{\rm tick}$. At low energy, the accuracy increases linearly with the dissipated energy, independently of the resolution. However, for higher energies, the accuracy saturates. (b) Resolution $\nu_{\rm tick}$ as a function of dissipated heat per tick Q_c, for various values of the accuracy N. The resolution first increases with dissipated energy, but then quickly saturates to a maximal value. (c) Trade-off between accuracy and resolution when the energy dissipation rate is fixed. The data are computed for fixed values of $T_c = E_w$, $T_h = 1000E_w$ and $g = \hbar\gamma = \hbar\Gamma = 0.05E_w$, while the ladder dimension d and cold qubit energy E_c are varied independently. Note that $d \geq 10$ for all of the plotted points, thus $T_c = E_w \ll E_\gamma = (d-1)E_w$ and we can safely ignore the absorption of a photon (i.e. the reverse of the decay process). Figure and caption from [11].

Regarding the definition of the accuracy it is important to note that in the model of the autonomous clock, we have the assumed that after the emission of the photon, which happens spontaneously, the entire pointer system is reset to the initial state. This means in particular that we have that the engine qubits are in equilibrium respective thermal reservoirs and in a product state with the ground state of the ladder. As in the weak-coupling limit the back action on the engine's qubits through the interplay with the ladder is minimal, we can describe the ticks of the clock as a renewal process. Hence, we get that the waiting time between any pair of consecutive ticks is statistically independent from and also identically distributed (i.i.d) to the waiting time in between any other pair of consecutive ticks.

This assumption is valid in the weak-coupling limit and the property of the independently distributed waiting times leads to the observation that the uncertainty in time of the $n^{\rm th}$ tick is given by $\sqrt{n}\Delta t_{\rm tick}$. Thus we get that the accuracy of the clock here is exactly the amount of ticks N such the uncertainty in time of the$N^{\rm th}$ tick, i.e. $\Delta t_{\rm tick}$, is equal to the mean waiting time between to consecutive ticks $t_{\rm tick}$.

One can do numerical calculations by using the equations of motion discussed in section 6.2.4. The results are shown in figure 6.5 (a-b), where we can observe the behaviour of the accuracy N and the resolution $\nu_{\rm tick}$ in terms of dissipated heat Q_c while fixing the other quantity. Figure 6.5 (c) shows the trade-off relation between the two figure of merit given a fixed amount of dissipated energy. From this results we can conclude that building a good clock, that has both high accuracy and resolution, implicates a large amount of dissipated energy per tick. We can see that in particular in figure 6.5 (c), where the curves are clearly ordered, yielding a better performance of the clock

with increased dissipation of entropy. The amount of dissipated energy hence becomes a resource that allows us to construct good clocks. Furthermore, we see that the trade-off relationship between the two figures of merit is neither trivial nor linear. In the next section we will focus on the low energy dissipation regime, where as we can observe in figure 6.5 (a), the accuracy becomes directly proportional to the entropy production for a fixed resolution.

6.2.3 Accuracy in the weak-coupling limit

We will now analyse a regime of low energy dissipation and the behaviour of the accuracy therein. In contrast to figure 6.5(a), where we had a fixed resolution and the dimension was allowed to vary, we will now try compute the accuracy solely in terms of the dissipated power and the dimension of the ladder. To be able to do so, we have to use approximations that are only valid in the case that there is a vanishing interaction between the engine and the ladder. In section 6.2.5 it will be shown how to derive this description of the clock through a perturbative approximation to the two-qubit engine.

Notably, we will show that the accuracy of the clock becomes independent of the details of its dynamics, it being solely dependent on the population bias of the virtual qubit Z_v and the ladder dimension d. To do so we will approximate the evolution of the ladder system by a biased random walk, determined by the interplay with the virtual qubit. This can be done as the virtual qubit cannot create any coherences in the ladder system. This renders the transition probability of the load going one 'step' up or one 'step' down independent of its position on the ladder.

We will denote the probability of the ladder's population to move upwards $p_\uparrow$ and the probability to move downwards $p_\downarrow$, such that

$$\frac{p_\uparrow}{p_\downarrow} = \mathrm{e}^{-\beta_v E_w} \tag{6.7}$$

is satisfied. Moreover, we assume that d is large enough for any reflections of the load on the boundaries of the ladder are negligible and the clock ticks immediately when the top level is reached. Doing do we can calculate the resolution by the simple one-letter formula,

$$\nu_{\text{tick}} = \frac{p_\uparrow - p_\downarrow}{d}\,. \tag{6.8}$$

An intuitive interpretation of this formula would be to see the dimension d has the 'height' of the ladder and $p_\uparrow - p_\downarrow$ as the 'speed' of the load. Then equation 6.8 can be seen as computing the number of times the load is able to 'climb' up the ladder by unit time.

Given the foregoing assumptions we will show in section 6.2.5, that the accuracy can be computed by,

$$N = d|Z_v|\,. \tag{6.9}$$

As said before, the accuracy is now independent of clocks's dynamics and therefore also its dynamical time scale which is determined by the rates $p_{\uparrow/\downarrow}$. Instead we have that the accuracy is

now dependent solely on the dimensionless quantities Z_v and d.

We have seen in the previous chapter in the detailed description of the two-qubit engine, the population bias of the virtual qubit depends on the dissipated heat. Bringing together the equations (5.26) and (5.27), we can compute the accuracy as,

$$N = d \tanh\left(\frac{(\beta_c - \beta_h)Q_c - \beta_h E_\gamma}{2d}\right). \tag{6.10}$$

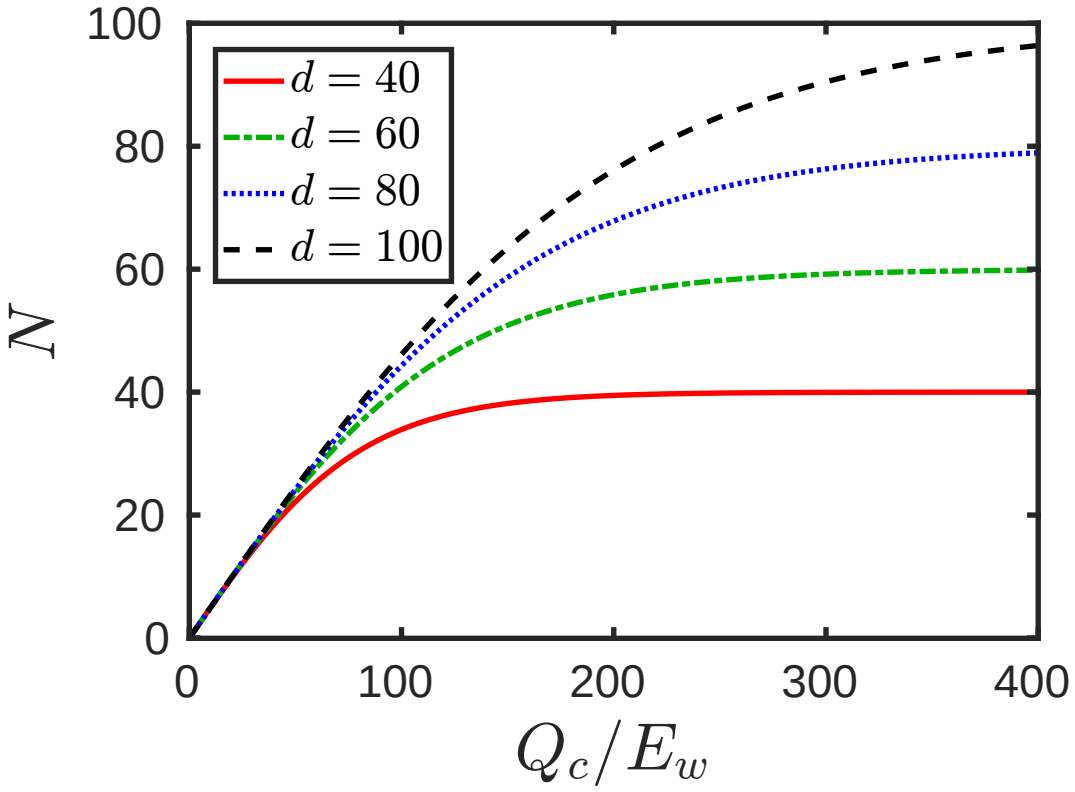

Figure 6.6: Accuracy N versus dissipated energy Q_c for various values of the dimension d of the ladder, according to the approximation (6.10) with the same bath temperatures as in figure 6.5. Figure and caption from [11].

The relationship of the population bias Z_v and the heat flow between the two thermal baths is however more general than what is considered here [217]. Section 6.2.7 includes a detailed discussion on this topic. It should be mentioned here, that we can observe that the accuracy solely

depends on the amount of heat that is dissipated but not its rate.

Figure 6.6 illustrates the behaviour of the accuracy described in equation (6.10) for different dimensions d. As one can see the accuracy starts out by increasing linearly and then saturates to its maximal value $N = d$. We know that with growing Q_c the virtual qubit's bias also grows with the limit $|Z_v| \to 1$ as $Q_c \to \infty$. This means that the accuracy is limited not only by the dissipated heat Q_c, but also by the dimension of the ladder. Therefore we can conclude that for attaining a certain accuracy, the clock needs to have at least a certain minimal dimension as well as it needs to dissipate minimum amount of entropy per tick.

If we look at the case where we let the dimension grow infinitely, i.e. taking the limit $d \to \infty$, we see that

$$N \to \frac{(\beta_c - \beta_h)Q_c - \beta_h E_\gamma}{2}\,. \tag{6.11}$$

This means that the accuracy also in the case of an unbounded dimension is linearly dependent on the dissipated energy, therefore imposing a fundamental limitation on the performance of the clock. It can be seen even more clearly if we use $Q_h = Q_c + E_\gamma$, such that we get

$$N \to \frac{\beta_c Q_c - \beta_h Q_h}{2} = \frac{\Delta S_{tick}}{2}\,, \tag{6.12}$$

where ΔS_{tick} denotes the increase of entropy in the clock. Observing equation (6.12) we can conclude that the regularity of the ticking of the clock is connected to the strength of its the arrow of time. Thus we have achieved a quantitative connection between the irreversibility of a clock and its arrow of time.

6.2.4 Dynamics

Here we have a detailed look at how we model the dynamics of the pointer system that we used to compute the distribution of ticks in the numerical simulations. Here the following formulation of the master equation is used. To model the effect that each thermal bath has on the corresponding qubit, we consider the superoperator

$$\mathcal{L}_j = \gamma_j \mathcal{D}[\sigma_j] + \gamma_j \mathrm{e}^{-\beta_j E_j} \mathcal{D}[\sigma_j^\dagger]\,, \tag{6.13}$$

with the index $j = h, c$, indicating the hot and the cold system, respectively. Where we have used the qubit lowering operators $\sigma_j = |0\rangle_j \langle 1|$ and the dissipator in Lindblad form

$$\mathcal{D}[L]\rho = L\rho L^\dagger - \frac{1}{2}\left\{L^\dagger L, \rho\right\}\,. \tag{6.14}$$

The dissipation rates $\gamma_{h,c}$ govern the time scale of the dissipative processes that act on the engine.

To model the decay channel we couple the ladder system to the electromagnetic field.. Specifically, we design the ladder such that only the highest energy transition $|d-1\rangle_w \to |0\rangle_w$ interacts

significantly with the reservoir of electromagnetic field modes at temperature T_c. We denote Γ the spontaneous emission rate, that is associated with the emission of a photon with energy $(d-1)E_w$. Moreover, we assume that the photon is then measured by a photo-detector that has perfect efficiency and only a negligible time delay. The measurement of the emitted photons produces the macroscopically observable ticks of the clock.

As mentioned we want to be able to ignore the transition $|0\rangle_w \to |d-1\rangle_w$, i.e. an incoming photon exciting the ladder system. We therefore require that $T_c \ll (d-1)E_w$, which means that the background temperature T_c is low enough, to make the probability for the unwanted transition to happen negligible.

In order to get the tick distribution of the clock, in theory one has to compute the density operator of the pointer $\rho(t)$ for all times t. As stated above we assume that the qubits in the engine do not change significantly, remaining in the thermal states and therefore in equilibrium with their respective heat baths. Thus the all the ticks are independent from each other and we are able to compute resolution and accuracy from the distribution in time of a single tick, which makes the computation feasible.

To start the description of the dynamics we denote $\rho_0(t)$, the state that is conditioned on no spontaneous emission having occurred up to time t. Working in this 'no-tick' subspace we assume that the pointer start in the initial state

$$\rho_0(0) = \frac{\mathrm{e}^{-\beta_h E_h \sigma_h^\dagger \sigma_h}}{\mathcal{Z}_h} \otimes \frac{\mathrm{e}^{-\beta_c E_c \sigma_c^\dagger \sigma_c}}{\mathcal{Z}_c} \otimes |0\rangle_w\langle 0| \,, \tag{6.15}$$

where $\mathcal{Z}_{c,h}$ are the respective partition functions. This models the situation where the clock has just ticked, the ladder has decayed into the ground state and the engine's qubit are in perfect equilibrium with their respective heat baths. Subsequently the state of the conditional density operator $\rho_0(t)$ evolves according to the master equation

$$\frac{\mathrm{d}\rho_0}{\mathrm{d}t} = i\left(\rho_0 H_{\text{eff}}^\dagger - H_{\text{eff}}\rho_0\right) + \mathcal{L}_h\rho_0 + \mathcal{L}_c\rho_0 \,, \tag{6.16}$$

where $H_{\text{eff}} = H_0 + H_{\text{int}} + H_{\text{se}}$ is the effective non-Hermitian Hamiltonian and the spontaneous emission contribution therein is

$$H_{\text{se}} = -\frac{i\Gamma}{2} |d-1\rangle_w\langle d-1| \,. \tag{6.17}$$

Due to the non-hermiticity of the evolution the state $\rho_0(t)$ does not remain normalised. Moreover, we have that the trace of the conditional density operator $P_0(t) = \mathrm{Tr}[\rho_0(t)]$ gives us the probability that the clock has not ticked yet. We can now compute the probability density $W(t)$ for the waiting times

$$W(t) = -\frac{\mathrm{d}P_0}{\mathrm{d}t}. \tag{6.18}$$

For the evaluation of the performance of our clock we only need the mean and the variance of the waiting time between two consecutive ticks. These are given by,

$$t_{\text{tick}} = \int_0^\infty \mathrm{d}\tau\ \tau W(\tau)\,, \tag{6.19}$$

$$(\Delta t_{\text{tick}})^2 = \int_0^\infty \mathrm{d}\tau\ (\tau - t_{\text{tick}})^2 W(\tau)\,. \tag{6.20}$$

6.2.5 Biased random walk approximation

In this section we show how we can one can derive the accuracy of the simple thermal clock by assuming that the evolution of the pointer system can be modelled by biased random walk controlled by the virtual qubit. This assumption leads to a diagonal density operator of the ladder system, which can be described as a vector containing the populations of the different energy levels. Furthermore we have to assume that the ladder is big enough, such that there are no reflection effects at the top, i.e. the population distribution does not 'feel' the boundedness of the ladder.

The foregoing assumptions allow us to model the state of the ladder with time-dependent probability distribution $q(n,t)$ on a grid of integers, that correspond to the different energy levels. Thus, we have $n \in \mathbb{Z}$, $q(n,t) > 0$, and $\sum_n q(n,t) = 1$. The evolution of this probability distribution is fully determined by the probability to move upwards $p_\uparrow$ (jumping to next integer), together with the probability to move downwards $p_\downarrow$ (jumping to the previous integer).

This leads to an equation of motion for the probability distribution,

$$\begin{aligned}\frac{\mathrm{d}q(n,t)}{\mathrm{d}t} &= p_\uparrow\, q(n-1,t) + p_\downarrow\, q(n+1,t) \\ &\quad -\ (p_\uparrow + p_\downarrow)\, q(n,t)\,.\end{aligned} \tag{6.21}$$

As we want to use this equation of motion for the computation of the accuracy and the resolution of the clock, we need to know how fast the distribution moves up the ladder and how much it spreads in doing so. For this purpose we consider the mean μ and the variance σ^2 of the distribution $q(n,t)$.

$$\mu(t) := \sum_n n\, q(n,t), \tag{6.22}$$

$$\sigma^2(t) := \sum_n \left(n - \mu(t)\right)^2 q(n,t) \tag{6.23}$$

The speed of the upwards movement is simply given by the time derivative of the mean,

$$\begin{aligned}\frac{\mathrm{d}\mu(t)}{\mathrm{d}t} &= \sum_n n\, \frac{\mathrm{d}q(n,t)}{\mathrm{d}t} \\ &= p_\uparrow - p_\downarrow\,.\end{aligned} \tag{6.24}$$

Analogously we can compute the time derivative of the variance,

$$\frac{\mathrm{d}\sigma^2(t)}{\mathrm{d}t} = \sum_n \left((n-\mu(t))^2 \frac{\mathrm{d}q(n,t)}{\mathrm{d}t} - 2\,(n-\mu(t)) \frac{\mathrm{d}\mu(t)}{\mathrm{d}t} q(n,t) \right) . \tag{6.25}$$

If we now use the definitions of the two quantities, given in equations (6.22) and (6.23), we can see that the whole expression simplifies to,

$$\frac{\mathrm{d}\sigma^2(t)}{\mathrm{d}t} = p_\uparrow + p_\downarrow \,. \tag{6.26}$$

Equipped with the time derivatives of the mean and the variance, we can now relate them to the figures of merit of the clock. Given that the average waiting time between two ticks of the clock is equal to the time that the load needs to 'climb' the d energy levels of the ladder from the bottom to the top, we get that

$$t_{\mathrm{tick}} = \frac{d}{\mathrm{d}\mu(t)/\mathrm{d}t} = \frac{d}{p_\uparrow - p_\downarrow} \,, \tag{6.27}$$

where $d-1$ was replaced by d for the sake of simplicity as it was anyways assumed that the dimension is large. The resolution of the clock ν_{tick} is now simply given by the inverse,

$$\nu_{\mathrm{tick}} = \frac{p_\uparrow - p_\downarrow}{d} \,, \tag{6.28}$$

recovering equation (6.8) as promised. Regarding the accuracy, we notice that within the average time it takes the clock to tick the variance will have increased by

$$\Delta\sigma^2 = t_{\mathrm{tick}} \frac{\mathrm{d}\sigma^2(t)}{\mathrm{d}t} = d \left(\frac{p_\uparrow + p_\downarrow}{p_\uparrow - p_\downarrow} \right) . \tag{6.29}$$

Given a good enough decay mechanism, the uncertainty of the load when it reaches the top of the ladder is equal to the uncertainty of a single tick. This way we can calculate the uncertainty in the time interval between consecutive ticks,

$$\Delta t_{\mathrm{tick}} = \frac{\sigma\,(t = t_{\mathrm{tick}})}{\mathrm{d}\mu(t)/\mathrm{d}t} = \frac{\sqrt{d}}{p_\uparrow - p_\downarrow} \sqrt{\frac{p_\uparrow + p_\downarrow}{p_\uparrow - p_\downarrow}} \,. \tag{6.30}$$

Hence, when we have that the variance grows to be as large as the whole ladder, i.e. $\sigma^2 = d^2$, the uncertainty is equal to the average time it takes for a single tick. We can therefore calculate the number N, such that this is the case,

$$N = d \left(\frac{p_\uparrow - p_\downarrow}{p_\uparrow + p_\downarrow} \right) , \tag{6.31}$$

which equal to the accuracy in equation (6.9), as it holds that $p_\uparrow / p_\downarrow = p_1 / p_0$.

6.2.6 Derivation of the biased random walk model

In this section we derive the model of the biased random walk that we have used abundantly in the previous sections. The essentially classical description of the pointer system stems from understanding the virtual qubit's action through the bias it gives the ladder's energy towards increasing. We now want to put the stochastic treatment of the pointer system on a firmer footing. We will work in a regime where the engine-ladder coupling g as well as the spontaneous emission rate Γ are both small in comparison to the thermal dissipation rates $\gamma_{c,h}$.

For this purpose we use the Nakajima-Zwanzig projection operator technique [218, 219], in the limit of $\gamma_j \gg g, \Gamma$. This allows us to compute the conditional reduced density operator of the ladder $\rho_w(t) = \mathrm{Tr}_{h,c}[\rho_0(t)]$ and the equation governing its evolution. We consider the projector,

$$\mathcal{P}\rho_0(t) = \rho_w(t) \otimes \tau_h \otimes \tau_c \,, \tag{6.32}$$

with $\tau_{h,c}$ denoting a local thermal state of the hot and the cold qubit, respectively. Choosing the index $j = h, c$ the local thermal states are given by

$$\tau_j = \frac{1}{\mathcal{Z}_j} \mathrm{e}^{-\beta_j E_j \sigma_j^\dagger \sigma_j}, \tag{6.33}$$

with $\mathcal{Z}_j = 1 + \mathrm{e}^{-\beta_j E_j}$ being the partition function for corresponding qubits. We can continue by writing equation (6.16) as

$$\frac{\mathrm{d}\rho_0}{\mathrm{d}t} = \mathcal{L}\rho_0 = (\mathcal{L}_0 + \mathcal{H}_{\mathrm{se}} + \mathcal{H}_{\mathrm{int}})\,\rho_0 \,, \tag{6.34}$$

where we have decomposed the Liouvillian as $\mathcal{L} = \mathcal{L}_0 + \mathcal{H}_{\mathrm{se}} + \mathcal{H}_{\mathrm{int}}$, as well as we have defined the Hamiltonian superoperator

$$\mathcal{H}_{\mathrm{se}}\rho = \mathrm{i}\left(\rho H_{\mathrm{se}}^\dagger - H_{\mathrm{se}}\rho\right), \tag{6.35}$$

and $\mathcal{H}_{\mathrm{int}}$ analogously. Now we transform the density operator to a dissipative interaction picture. We can do so by defining

$$\tilde{\rho}_0(t) = \mathrm{e}^{-\mathcal{L}_0 t}\rho_0(t) \,, \tag{6.36}$$

$$\tilde{\mathcal{H}}_{\mathrm{int}}(t) = \mathrm{e}^{-\mathcal{L}_0 t}\mathcal{H}_{\mathrm{int}}\mathrm{e}^{\mathcal{L}_0 t} \tag{6.37}$$

and

$$\tilde{\mathcal{H}}_{\mathrm{se}}(t) = \mathcal{H}_{\mathrm{se}} \,. \tag{6.38}$$

The standard perturbative argument [220] allows us to obtain

$$\frac{\mathrm{d}\mathcal{P}\tilde{\rho}_0}{\mathrm{d}t} = \mathcal{H}_{\mathrm{se}}\mathcal{P}\tilde{\rho}_0(t) + \int_0^t \mathrm{d}t'\; \mathcal{P}\tilde{\mathcal{H}}_{\mathrm{int}}(t)\tilde{\mathcal{H}}_{\mathrm{int}}(t')\mathcal{P}\tilde{\rho}_0(t'), \tag{6.39}$$

which is valid up to second order in the small quantities g and Γ. To continue we can apply the Born-Markov approximation to the t' integral in above expression [220] . This allows us to extend

the lower integration limit to negative infinity and transform $\tilde{\rho}_0(t') \to \tilde{\rho}_0(t)$. We can justify this by the assumption that $\gamma_j \gg g, \Gamma$, guaranteeing that the integrand vanishes quickly in comparison to the time scale of notable changes of $\mathcal{P}\tilde{\rho}_0(t)$.

This simplifies equation (6.39) as we can first expand the commutators, then trace out the engine's qubits and finally transform back to the Schrödinger picture. The result is a master equation in which the evolution of the populations is decoupled from that of the coherences as long as it is expressed in the eigenbasis of B_w. As we have that the initial state, given in equation (6.15), has no coherences, we can focus solely on the contributions for the populations

$$\frac{\mathrm{d}\rho_w}{\mathrm{d}t} = p_\downarrow \mathcal{D}[B_w]\rho_w + p_\uparrow \mathcal{D}[B_w^\dagger]\rho_w - \frac{\Gamma}{2}\left(|d-1\rangle_w \langle d-1|_w \rho_w + \rho_w |d-1\rangle_w \langle d-1|_w\right) . \tag{6.40}$$

This allows us to define the probability vector $\vec{q}$ whose elements are given by

$$q_n(t) = \mathrm{Tr}\left(\rho_w(t) |n\rangle_w \langle n|_w\right) , \tag{6.41}$$

which gives

$$\frac{\mathrm{d}\vec{q}}{\mathrm{d}t} = A\vec{q}, \tag{6.42}$$

where

$$A = \begin{pmatrix} -p_\uparrow & p_\downarrow & & & \\ p_\uparrow & -(p_\uparrow + p_\downarrow) & & & \\ & & \ddots & & \\ & & & -(p_\uparrow + p_\downarrow) & p_\downarrow \\ & & & p_\uparrow & -(p_\downarrow + \Gamma) \end{pmatrix} . \tag{6.43}$$

Except the additional term proportional to Γ, corresponding to the spontaneous decay from the top level, this expression is equal to equation (6.21), where $\vec{q}_n(t) = q(n,t)$. Furthermore, for the forward and backward rates we get Laplace-transformed correlation functions of the engine's qubits

$$p_\downarrow = 2g^2 \int_0^\infty \mathrm{d}t\ \mathrm{e}^{\mathrm{i}E_w t} \left\langle \sigma_h(t)\sigma_h^\dagger(0)\sigma_c^\dagger(t)\sigma_c(0) \right\rangle , \tag{6.44}$$

$$p_\uparrow = 2g^2 \int_0^\infty \mathrm{d}t\ \mathrm{e}^{-\mathrm{i}E_w t} \left\langle \sigma_h^\dagger(t)\sigma_h(0)\sigma_c(t)\sigma_c^\dagger(0) \right\rangle , \tag{6.45}$$

with the angle brackets denoting an average with respect to $\rho_h \otimes \rho_c$. The time dependence of the operators is given through $\sigma_{h,c}(t) = \mathrm{e}^{\mathcal{L}_0^\dagger t}\sigma_{h,c}$, where the adjoint Liouvillian $\mathcal{L}_0^\dagger$ is defined by $\mathrm{Tr}(Q\mathcal{L}_0(P)) = \mathrm{Tr}(\mathcal{L}_0^\dagger(Q)P)$, for arbitrary operators P and Q.

Explicitly, this leads to

$$\sigma_j(t) = \mathrm{e}^{-iE_j t - \frac{\gamma_j \mathcal{Z}_j t}{2}} \sigma_j \tag{6.46}$$

for $j = h, c$, corresponding to the hot and the cold qubit, respectively. This further implies

$$p_\downarrow = \frac{4g^2 \mathrm{e}^{-\beta_c E_c}}{\mathcal{Z}_h \mathcal{Z}_c (\gamma_h \mathcal{Z}_h + \gamma_c \mathcal{Z}_c)}, \tag{6.47}$$

$$p_\uparrow = \frac{4g^2 \mathrm{e}^{-\beta_h E_h}}{\mathcal{Z}_h \mathcal{Z}_c (\gamma_h \mathcal{Z}_h + \gamma_c \mathcal{Z}_c)}, \tag{6.48}$$

from which we can verify that

$$\frac{p_\uparrow}{p_\downarrow} = \mathrm{e}^{-(\beta_h E_h - \beta_c E_c)} = \mathrm{e}^{-\beta_v E_w}\,, \tag{6.49}$$

with the self-consistency condition of the Born-Markov approximation requiring that $p_{\downarrow/\uparrow} \ll \gamma_j$.

6.2.7 Limits of thermally run clocks

Above we have brought forward the argument that the accuracy of an autonomous clock is bounded by the entropy it dissipates per tick. We have done so by using the relation between the bias of the ladder to move upwards and the population bias of the virtual qubit. For the two-qubit engine this satisfies [211],

$$\beta_w E_w = \beta_h E_h - \beta_c E_c\,. \tag{6.50}$$

If we now multiply with the 'height' of the ladder, $d-1$, we get

$$\beta_w E_\gamma = \beta_h \left(Q_c + E_\gamma\right) - \beta_c Q_c\,, \tag{6.51}$$

since $E_c = E_h - E_w$. There is an intuitive way to understand this expression. Each time a virtual qubit is prepared by the thermal machine in an appropriate state to exchange E_w with the ladder system, it necessarily has to absorb E_h from the hot bath and also dissipate E_c to the cold bath.

There exists a large class of autonomous quantum thermal machines [217], for which it has been shown that this statement holds true as well, in the weak-coupling regime at least. While one can find more intricate designs and other tweaks for the engine, making it arbitrarily complex, the result will still be constrained by equation (6.50) [215]. This leads us to conclude that the trade-off between accuracy and power consumption that we have derived above for autonomous clocks, is not limited to the explicit model considered here.

This opens up an interesting direction for future research, namely, investigating clocks that are not in the weak-coupling regime. This is specially intriguing it should be possible to outperform the stochastic models, hence essentially classical, that we have presented above by allowing or willingly

incorporating coherences [216]. Even here where we have considered the simplest case, i.e. the two-qubit engine, there is some coherence present in the subspace in which the interaction between the engine and the ladder takes place. This coherence is maintained during the upwards movement of the ladder and prevents the energy distribution from spreading too much, as the authors of [211] show. Still this is not not enough for the presented model to take full advantage of quantum properties. This shows the great potential of achieving even more accurate clocks by the adoption of stronger couplings and more coherence.

6.3 Fundamental limits

We have seen that in the simple thermal clock model that was discussed extensively the performance of the clock, i.e. its ability to accurately and precisely measure time, comes with an intrinsic cost, the generation of entropy through heat dissipation. Now the question is whether this intrinsic work cost for measuring time is specific aspect of the explicit model that we have presented above or whether this universal feature of all time measurement devices. Here we will argue in favour of it being an intrinsic property of any autonomous clock.

As discussed in section 6.2, we have that due to the bipartite nature of autonomous time measurement devices the ticking of the device requires a transition in the pointer system that leads to a corresponding change in the register [186]. Ideally this is a spontaneous and effectively irreversible process, the results of which are accessible to an external observer. The bias of a transition to occur with a greater likelihood than its time-reverse induces a reduction of the free energy in the pointer system. Thus, for clock to run continuously, the pointer system needs to have access to resources that allow it to renew its free energy. Such can be achieved by providing out of equilibrium resources. The question now is whether a time measurement device can exist, that perfectly turns free energy into ticks, that means without increasing entropy.

Let us first look at the question with focus on clocks which are powered by thermal baths, reiterating some of the argument that were given in the previous section. There exists a wide range of possibilities constructing such clocks, that go beyond the explicit model that we have considered above [215]. Still, even this more general designs will work according to the same principle, using the pointer system to drive a population until an unstable level induces a tick. In order for the time-reverse not to happen equally likely, the energy of the thermal background must be small compared to that of the unstable level. This could as well work with complex ladder structures or multiple unstable levels and machines comprised of much more then two qubits.

Notwithstanding, any of these possible generalisations will still be subject to the laws of thermodynamics. This means that specifically the maximum efficiency of converting free energy into ticks will be limited by the Carnot efficiency $\eta_C = 1 - T_c/T_h$. Note that in the limit of going to Carnot efficiency, that is the limit where the power consumption vanishes, machines become reversible. As a finite power consumption is necessary for a finite resolution [186], this conversely implies that a clock that works with Carnot efficiency will tick infinitely slowly. Thus, even in unfeasible regimes, such as $T_c \to 0$ or $T_h \to \infty$, demanding a clock with a finite resolution leads to a minimum in dissipated heat, which in turn implies a minimum production of entropy.

One can also think of more general non-thermal resources that could be used to power a clock. Even in such scenarios the condition of functioning autonomously, prohibits us using resources that themselves need well-timed operations for their production. This in principle should enable us to build clocks that have a better performance than thermally driven ones. Nevertheless, an autonomous clock that has a finite resolution without dissipating entropy, would embody an autonomous machine operates with unit efficiency at finite power. We can hence conclude that in the case that running a autonomous quantum clock is not intrinsically connected to the generation of entropy, the implications would go far beyond the task of measuring time.

At last, it shall be remarked that we have only considered the minimal entropy dissipation possible by running a clock, ignoring all other sources of heat dissipation. Examples of such are for example the reset of the register, which would come with its own associated energy cost [165, 221], or the inability to recycle the photon's energy. Although there is probably a square root missing [216], let us imagine the bound derived in equation (6.12) to be universal. It, much alike Landauer's principle, vastly underestimates the real costs of running a clock. For example considering a typical atomic clock [222], that has a resolution of around 10^{10} Hz and an accuracy of about 10^{16}, then the bound would imply a minimal power consumption of around 50μW. This still surpasses the power consumption of the most efficient atomic clocks by magnitudes [223].

To sum up, we have that in the model of a autonomous clock presented above each unit of dissipated heat can be either turned into an increase of the resolution or the accuracy. The model has allowed us to find a connection between the second law of thermodynamics and the arrow of time [224, 225]. Moreover, we found that independent of the entropy production, the dimension of the pointer system (specifically the ladder) constrains the performance of the clock. This has also shown to hold in general by the authors of [216]. Our considerations were mainly focused on the explicit example of an autonomous clock, given by the minimal thermal clock, while this already allowed us to gain insights on the limitations of our ability to measure time, it would be interesting to see whether one can extend this results by for example considering multiple conserved quantities [226, 227, 228, 229]. Looking into such question could clarify how the choice of resources, another example being passive states [230], impacts the performance of clocks.

The presented results could have great impact also in tasks like controlling other quantum systems [206, 231, 232], helping in the clarification of the underlying costs. It could also be possible that the model of autonomous clocks becomes an important building block on the road of constructing quantum information processing devices on a microscopical scale. As self-contained modules they could be used to time the interaction for the implementation of time-dependent Hamiltonians.

Lastly, the question remains whether one can increase the performance of autonomous clocks by exploiting genuine quantum effects, as for example coherence or entanglement [233, 234]. While letting classical clocks work in parallel can only increase their performance linearly, the use of quantum phenomena as resources could change that behaviour completely.

Chapter 7

Conclusion

The first chapter looked at definitions of a doctoral thesis in the given context and outlined the content of this one here. The second chapter started with insights into the history of the fields of quantum theory, information theory and quantum information theory. It then continued introducing all the prerequisites for what was to follow. After the introduction of the necessary measure of information, a no-go result on the generalisation of quantum conditional mutual information to the non-asymptotic framework was presented.

The third chapter describes the importance of bases for the modelling of systems in quantum information theory. The convenient Bloch decomposition is presented there as well as a discussion of the property of mutual unbiasedness. The chapter then goes on with an analysis of possible bases that can be used for high-dimensional systems. The fourth chapter deals with entanglement, as well as its detection and quantification. For the detection of entanglement a theorem is presented that makes use of sets of anti-commuting basis elements. For the quantification a framework is introduced, that allows to do so with the use of two measurement settings only.

The fifth chapter is about thermodynamics and information. It introduces quantum thermodynamics and the quantum machines one can construct therein. In the end a little discourse on a algorithmic approach for thermodynamics and emergence is given. The sixth chapter is on time and discuses how to measure it. For this purpose autonomous quantum clocks are introduced and the performance of such is analysed.

Last, regarding the question posed in the title: What is the role of information in physics? This ample question remains without a definite answer. Although we have been going through a lot of notions and different contexts in which information plays a vital role, it not yet clear what information exactly is. From knowledge what a measurement basis is to the decision which measurement basis to use to the correlations that the system under observation might show, all is information. Despite the very rigorous definitions we have in some contexts for information, we cannot pin it down to one in general. We have discussed quantities, such as the Kolmogorov complexity, that seem to be the optimal measure for information, taking out the counterfactuals that we got so used to. Still it remains uncomputable and this may just be the point. Since Gödel we know that mathematics is not the perfect natural language it was believed to be and still it is the most universal

we have. It could be, that this finiteness, limitedness, all languages that we have discovered so far bring along is inevitable [235]. Information transcends physics by far, appearing in places where physics cannot go [236] and only the belief in an absolutely predetermined world can render them equals [199]. In the end, it may be up each and everyone one of us to decide where we are going [237].

Science is dead, long live Science!

Bibliography

[1] European Ministers in charge of Higher Education. *The Bologna Declaration of 19 June 1999: Joint declaration of the European Ministers of Education.* Bologna, (1999)

[2] European Commission. *The European qualifications framework for lifelong learning (EQF).* Office for Official Publications of the European Communities, Luxembourg, (2008)

[3] Bologna Working Group. *A Framework for Qualifications of the European Higher Education Area.* Bologna Working Group Report on Qualifications Frameworks (Copenhagen, Danish Ministry of Science, Technology and Innovation), (2005)

[4] W. von Humboldt. *Zur Theorie der Bildung des Menschen.* Schriften zur Bildung, ed. Gerhard Lauer, Reclam Verlag, (2017)

[5] M. Barton *Dissertations: Past, present, and future.* PhD thesis, University of South Florida, (2005)

[6] D. Boud, A. Lee. *Changing Practices of Doctoral Education.* Routledge, (2009).

[7] P. Erker. *How not to Rényi generalize the Quantum Conditional Mutual Information.* J. Phys. A: Math. Theor. 48 275303, (2015)

[8] P. Erker, M. Krenn, M. Huber. *Quantifying high dimensional entanglement with two mutually unbiased bases.* Quantum 1, 22, (2017)

[9] A. Asadian, P. Erker, M. Huber, C. Klöckl. *Heisenberg-Weyl Observables: Bloch vectors in phase space.* Phys. Rev. A 94, 010301(R), (2016)

[10] J. Bavaresco, N. Herrera Valencia, C. Klöckl, M. Pivoluska, P. Erker, N. Friis, M. Malik, M. Huber. *Measurements in two bases are sufficient for certifying high-dimensional entanglement.* Nature Physics, volume 14, pages 1032-1037, (2018)

[11] P. Erker, M. Mitchison, R. Silva, M. Woods, N. Brunner, M. Huber. *Autonomous Quantum Clocks: Does Thermodynamics Limit Our Ability to Measure Time?.* Phys. Rev. X 7, 031022, (2017)

[12] J. Stein. *Cosmic Numbers: The Numbers that Define Our Universe* . Basic Books (2011)

[13] A. Einstein. Autobiographical Notes. *Albert Einstein: Philosopher-Scientist.* Eds.: P. A. Schilpp, Open Court , p. 2-94, (1949)

[14] J. Von Neumann. *Mathematische Grundlagen der Quantenmechanik.* Springer, (1932)

[15] A. Turing. *On Computable Numbers, with an Application to the Entscheidungsproblem* . Proceedings of the London Mathematical Society, 2 , 42 (1), pp. 230?65, (1937)

[16] K. Gödel. *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I.* Monatshefte für Mathematik und Physik, Volume 38, Issue 1, pp 173?198 (1931)

[17] R. Bertlmann, A. Zeilinger (Eds.). *Quantum (Un)speakables - From Bell to Quantum Information.* R. Bertlmann, A. Zeilinger. Springer, (2002)

[18] A. Einstein, B. Podolsky, N. Rosen. *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?.* Phys. Rev. 47, 777, (1935)

[19] S. Wiesner. *Conjugate coding.* ACM SIGACT, Volume 15 Issue 1, pages 78-88, (1983)

[20] Z. Gavorova. Private communication, (2015)

[21] D. Kaiser. *American Physics and the Cold War Bubble.* University of Chicago Press, in preparation

[22] N. Herbert. *FLASH - A superluminal communicator based upon a new kind of quantum measurement.* Foundations of Physics, 12 (12): 1171-1179, (1982)

[23] D. Kaiser. *How the Hippies Saved Physics.* Norton, (2011)

[24] W. Wootters, W. Zurek. *A single quantum cannot be cloned.* Nature, volume 299, pages 802?803 (1982)

[25] D. Dieks. *Communication by EPR devices.* Physics Letters A 92, 271-272, (1982)

[26] P. Benioff. *Quantum mechanical hamiltonian models of turing machines* Journal of Statistical Physics, Band 29, 515?546, (1982)

[27] R. Feynmann. *Simulating Physics with Computers.* International Journal of Theoretical Physics, vol 21, Nos. 6/7, (1982)

[28] Y. Manin. *Vychislimoe i nevychislimoe (Computable and Noncomputable).* (in Russian). Sov. Radio. pp. 13-15, (1980)

[29] D. Deutsch. *Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer.* Proc. R. Soc., vol. 400 no. 1818, 97?117, (1985)

[30] A. Ekert. *Quantum cryptography based on Bell's theorem.* Phys. Rev. Lett. 67 (6): 661?663, (1991)

[31] C. Bennett, G. Brassard. *Quantum cryptography: Public key distribution and coin tossing.* Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, volume 175, page 8, (1984)

[32] D. Simon. *On the power of quantum computation.* Foundations of Computer Science, 35th Annual Symposium, 116?123, (1995)

[33] P. Shor. *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer.* SIAM Journal on Computing, 26/1997, 1484?1509, (1997)

[34] L. Grover. *A fast quantum mechanical algorithm for database search.* Proceedings, 28th Annual ACM Symposium on the Theory of Computing, p. 212, (1996)

[35] A. Steane. *Multiple-Particle Interference and Quantum Error Correction.* Proc. Roy. Soc. Lond. A. 452 (1954): 2551?2577, (1996).

[36] P. Shor. *Scheme for reducing decoherence in quantum computer memory.* AT&T Bell Laboratories, (1995)

[37] T. Kadowaki, H. Nishimori.*Quantum annealing in the transverse Ising model.* Phys. Rev. E, vol 58, number 5, (1998)

[38] B. Schumacher. *Quantum coding.* Phys. Rev. A 51, 2738, (1995)

[39] J. Cirac and P. Zoller. *Quantum Computations with Cold Trapped Ions.* Phys. Rev. Lett. 74, 4091, (1995)

[40] C. Monroe, D. Meekhof, B. King, W. Itano, D. Wineland. *Demonstration of a Fundamental Quantum Logic Gate.* Physical Review Letters. 75 (25): 4714?4717, (1995)

[41] I. Chuang, N. Gershenfeld, M. Kubinec. *Experimental Implementation of Fast Quantum Searching.* Physical Review Letters. 80 (15): 3408?3411, (1998)

[42] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, A. Zeilinger. *Experimental Quantum Teleportation.* Nature 390, 6660, 575-579, (1997)

[43] W. Greiner, J. Reinhardt. *Field Quantization.* Springer, (1996)

[44] Design by the author, created by Flavia Mudesto, (2018)

[45] C. Shannon. *A Mathematical Theory of Communication.* Bell System Technical Journal. 27 (3): 379-423 & (4): 623-666, (1948)

[46] I. Devetak, J. Yard. *The operational meaning of quantum conditional information.* Phys. Rev. Lett. 100, 230501, (2008)

[47] M. Christandl, A. Winter.*Squashed entanglement: An additive entanglement measure.* J. Math. Phys. Vol 45, No 3, pp. 829-840, (2004)

[48] P. Hayden, R. Jozsa, D. Petz, A. Winter. *Structure of states which satisfy strong subadditivity of quantum entropy with equality.* Communications in Mathematical Physics, 246(2):359374, (2004)

[49] N. Datta. *Min- and Max-Relative Entropies and a New Entanglement Monotone.* IEEE Trans. Inf. Th., vol. 55, no. 6, (2009)

[50] R. Renner. *Security of Quantum Key Distribution.* PhD thesis, ETH Zurich, (2005)

[51] M. Tomamichel. *A Framework for Non-Asymptotic Quantum information Theory.* PhD thesis, ETH Zurich, (2012)

[52] D. Petz. *Quasi-Entropies for finite quantum Systems.* Rep. Math. Phys., 23:57?65, (1984)

[53] M. Tomamichel, R. Colbeck, R. Renner. *Duality Between Smooth Min- and Max-Entropies.*IEEE Trans. Inf. Th., vol. 56, no. 9, (2010)

[54] R. Koenig, R. Renner, C. Schaffner. *The operational meaning of min- and max-entropy.* IEEE Trans. Inf. Th., vol. 55, no. 9, (2009)

[55] M. Mosonyi, F. Hiai. *On the quantum Rényi relative entropies and related capacity formulas.* IEEE Transactions on Information Theory, vol. 57, 2474–2487, (2011)

[56] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, M. Tomamichel. *On quantum Renyi entropies: a new generalization and some properties.* J. Math. Phys. 54, 122203, (2013)

[57] P. Faist. *Entropy zoo.* https://phfaist.com/entropyzoo

[58] M. Berta, K. Seshadreesan. M. Wilde. *Rényi generalizations of the conditional quantum mutual information.* Journal of Mathematical Physics vol. 56, no. 2, article no. 022205, (2015)

[59] O. Fawzi, R. Renner. *Quantum conditional mutual information and approximate Markov chains.* Communications in Mathematical Physics: Volume 340, Issue 2, Page 575-611, (2015)

[60] B. Ibinson, N. Linden, A. Winter. *Robustness of quantum Markov chains.* Communications in Mathematical Physics 2; 289, (2008)

[61] P. Erker. *Duality for the conditional min-information.* Semester thesis, ETH Zurich, (2013)

[62] M. Christandl, N. Schuch, A. Winter. *Entanglement of the Antisymmetric States.* Communications in Mathematical Physics 10; 311(2), (2009)

[63] F. Bloch. *Nuclear Induction.* Phys. Rev. 70, 460, (1946)

[64] U. Fano. *Pairs of two-level systems.* Rev. Mod. Phys. 55, 855 (1983)

[65] R. A. Bertlmann and P. Krammer. *Bloch vectors for qudits.* Journal of Physics A: Mathematical and Theoretical 41, 235303 (2008)

[66] E. Bruning, H. Mäkelä, A. Messina, and F. Petruccione. *Parametrizations of density matrices.* Journal of Modern Optics 59, 1 (2012)

[67] F. T. Hioe and J. H. Eberly. *N-Level Coherence Vector and Higher Conservation Laws in Quantum Optics and Quantum Mechanics.* Phys. Rev. Lett. 47, 838, (1981)

[68] J. I. de Vicente. *Further results on entanglement detection and quantification from the correlation matrix criterion.* Journal of Physics A: Mathematical and Theoretical 41, 065309, (2008)

[69] J. I. de Vicente and M. Huber.*Multipartite entanglement detection from correlation tensors.* Phys. Rev. A 84, 062306, (2011)

[70] P. Badziag, C. Brukner, W. Laskowski, T. Paterek, and M. Zukowski. *Experimentally Friendly Geometrical Criteria for Entanglement.* Phys. Rev. Lett. 100, 140403, (2008)

[71] W. Laskowski, M. Markiewicz, T. Paterek, and M. Zukowski. *Correlation tensor criteria for genuine multiqubit entanglement.* Phys. Rev. A 84, 062305, (2011)

[72] C. Schwemmer, L. Knips, M. C. Tran, A. de Rosier, W. Laskowski, T. Paterek, and H. Weinfurter. *Genuine Multipartite Entanglement without Multipartite Correlations.* Phys. Rev. Lett. 114, 180501, (2015)

[73] C. Klöckl and M. Huber. *Characterizing multipartite entanglement without shared reference frames.* Phys. Rev. A 91, 042339, (2015)

[74] O. Gamel. *Entangled Bloch spheres: Bloch matrix and two-qubit state space.* Phys. Rev. A 93, 062320, (2016)

[75] S. Goyal, B. Simon, R. Singh, S. Simon. *Geometry of the generalized Bloch sphere for qutrits.* J. Phys. A: Math. Theor. 49 165203, (2016)

[76] P. Kurzynski. *Multi-Bloch vector representation of the qutrit.* Quantum Inf. Comp. 11, 361, (2011)

[77] P. Kurzynski, A. Kolodziejski, W. Laskowski, M. Markiewiczi. *Three-dimensional visualisation of a qutrit.* Phys. Rev. A 93, 062126, (2016)

[78] S. Massar and P. Spindel. *Uncertainty Relation for the Discrete Fourier Transform.* Phys. Rev. Lett. 100, 190401, (2008)

[79] R. Namiki and Y. Tokunaga. *Discrete fourier-based correlations for entanglement detection.* Phys. Rev. Lett. 108, 230503, (2012)

[80] A. Asadian, C. Budroni, F. E. S. Steinhoff, P. Rabl, and O. Gühne. *Contextuality in Phase Space.* Phys. Rev. Lett. 114, 250403, (2015)

[81] N. Cotfas and D. Dragoman. *Properties of finite Gaussians and the discrete-continuous transition.* Journal of Physics A: Mathematical and Theoretical 45, 425305, (2012)

[82] A. Vourdas. *Quantum systems with finite Hilbert space.* Reports on Progress in Physics 67, 267, (2004)

[83] C. Spengler, M. Huber, and B. C. Hiesmayr. *A composite parameterization of unitary groups, density matrices and subspaces.* Journal of Physics A: Mathematical and Theoretical 43, 385306, (2010)

[84] W. Wootters, B. Fields. *Optimal state-determination by mutually unbiased measurements.* Ann. Phys. 191, 363, (1989)

[85] N. Cerf, M. Bourennane, A. Karlsson, N. Gisin. *.Security of Quantum Key Distribution Using d-Level Systems.* Phys. Rev. Lett. 88, 127902, (2002)

[86] E. Aguilar, J. Borkala, P. Mironowicz, M. Pawlowski. *Connections Between Mutually Unbiased Bases and Quantum Random Access Codes.* Phys. Rev. Lett. 121, 050501, (2018)

[87] T. Durt, B.-G. Englert, I. Bengtsson, K. Zyczkowski. *On mutually unbiased bases.* Int. J. Quant. Inf. 8, 535, (2010)

[88] P. Vernaz-Gris, A. Ketterer, A. Keller, S. P. Walborn, T. Coudreau, P. Milman. *Continuous discretization of infinite-dimensional Hilbert spaces.* Phys. Rev. A 89, 052311, (2014)

[89] M. Krenn, R. Fickler, M. Huber, R. Lapkiewicz, W. Plick, S. Ramelow, A. Zeilinger. *Entangled singularity patterns of photons in Ince-Gauss modes.* Phys. Rev. A 87, 012326, (2013)

[90] E. Schrödinger. *Discussion of probability relations between separated systems.* Mathematical Proceedings of the Cambridge Philosophical Society. 31 (4): 555-563, (1935)

[91] G. Hermann. *Die naturphilosophischen Grundlagen der Quantenmechanik.* Naturwissenschaften, Volume 23, Number 42, 718-721, (1935)

[92] N. Bohr. *Can Quantum-Mechanical Description of Physical Reality be Considered Complete?* Physical Review, 38: 696-702, (1935)

[93] V. Baumann, A. Hansen, S. Wolf. *The measurement problem is the measurement problem is the measurement problem.* arXiv:1611.01111

[94] V. Baumann, S. Wolf. *On Formalisms and Interpretations.* Quantum 2, 99, (2018)

[95] C. Bennett, S. Wiesner, *Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states.* Phys. Rev. Lett. 69 (20): 288, (1992)

[96] C. H. Bennett, P. W. Shor, J. A. Smolin, A. V. Thapliyal. *Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem.* IEEE Transactions on Information Theory, Vol. 48 (10), 2637 - 2655, (2002)

[97] R. Fickler, R. Lapkiewicz, W. Plick, M. Krenn, C. Schäff, S. Ramelow, A. Zeilinger. *Quantum Entanglement of High Angular Momenta.* Science, 338, 640-643, (2012)

[98] R. Fickler, R. Lapkiewicz, M. Huber, M. Lavery, M. Padgett, A. Zeilinger. *Interface between path and orbital angular momentum entanglement for high-dimensional photonic quantum information.* Nature Communications 5, 4502, (2014)

[99] A. C. Dada, J. Leach, G. S. Buller, M. J. Padgett, E. Andersson. *Experimental high-dimensional two-photon entanglement and violations of generalized Bell inequalities.* Nature Physics 7, 677680, (2011)

[100] M. Krenn, M. Huber, R. Fickler, R. Lapkiewicz, S. Ramelow, A. Zeilinger. *Generation and confirmation of a (100 × 100)-dimensional entangled quantum system.* PNAS 111(17), 6243-6247, (2014)

[101] C. K. Law and J. H. Eberly. *Analysis and Interpretation of High Transverse Entanglement in Optical Parametric Down Conversion.* Phys. Rev. Lett. 92, 127903, (2004)

[102] M. O'Sullivan-Hale, I. Khan, R. Boyd, J. Howell. *Pixel Entanglement: Experimental Realization of Optically Entangled d=3 and d=6 Qudits.* Phys. Rev. Lett. 94, 220501, (2005)

[103] J. Pors, S. Oemrawsingh, A. Aiello, M. van Exter, E. Eliel, G. 't Hooft, J. Woerdman. *Shannon Dimensionality of Quantum Channels and Its Application to Photon Entanglement.* Phys. Rev. Lett. 101, 120502, (2008)

[104] P. Dixon, G. A. Howland, J. Schneeloch, J. Howell. *Quantum Mutual Information Capacity for High-Dimensional Entangled States.* Phys. Rev. Lett. 108, 143603, (2012)

[105] P. Moreau, F. Devaux, E. Luntz. *Einstein-Podolsky-Rosen Paradox in Twin Images.* Phys. Rev. Lett. 113, 160401, (2014)

[106] M. Zukowski, A. Zeilinger, M. Horne. *Realizable higher-dimensional two-particle entanglementsvia multiport beam splitters.* Phys. Rev. A, 55, 2564-2579, (1997)

[107] M. Suda, C. Pacher, M. Peev, M. Dusek, F. Hipp. *Experimental access to higher-dimensional entangled quantum systems using integrated optics.* Quantum Inf. Process. 12: 1915-1945, (2013)

[108] Ch. Schäff, R. Polster, M. Huber, S. Ramelow, A. Zeilinger. *Experimental access to higher-dimensional entangled quantum systems using integrated optics* .Optica 2(6), 523-529, (2015)

[109] A. Vaziri, G. Weihs, A. Zeilinger. *Experimental Two-Photon, Three-Dimensional Entanglement for Quantum Communication.* Phys. Rev. Lett. 89, 240401, (2002)

[110] G. Molina-Terriza, J. Torres, L. Torner. *Orbital angular momentum of photons in noncollinear parametric downconversion.* Opt. Comm. 228 (1–3), pp. 155–160 (2003)

[111] G. Molina-Terriza, A. Vaziri, J. Rehacek, Z. Hradil, A. Zeilinger. *Triggered Qutrits for Quantum Communication Protocols.* Phys. Rev. Lett. 92, 167903 (2004)

[112] M. Agnew, J. Leach, M. McLaren, F. Roux, R. Boyd. *Tomography of the quantum state of photons entangled in high dimensions.* Phys. Rev. A, 84(6), 062101, (2011)

[113] M. McLaren, M. Agnew, J. Leach, F. Roux, M. Padgett, R. Boyd, A. Forbes. *Entangled bessel-gaussian beams.* Optics express, 20(21), 23589-23597, (2012)

[114] D. Giovannini, D. Romero, J. Leach, J. Dudley, A. Forbes, M. Padgett. *Characterization of High-Dimensional Entangled Systems via Mutually Unbiased Measurements.* Phys. Rev. Lett. 110, 143601, (2013)

[115] C. Bernhard, B. Bessire, A. Montina, M. Pfaffhauser, A. Stefanov, S. Wolf. *Non-locality of experimental qutrit pairs.* J. of Phys. A: Math. and Theor., 47(42):424013 (2014)

[116] A. Tiranov, J. Lavoie, A. Ferrier, P. Goldner, V. Verma, S. Woo Nam, R. Mirin, A. Lita, F. Marsili, H. Herrmann, C. Silberhorn, N. Gisin, M. Afzelius, F. Bussieres. *Storage of hyperentanglement in a solid-state quantum memory.* Optica, 2, 279–287, (2015)

[117] A. Tiranov, S. Designolle, E. Zambrini Cruzeiro, J. Lavoie, N. Brunner, M. Afzelius, M. Huber, N. Gisin. *Quantification of multidimensional entanglement stored in a crystal.* Phys. Rev. A 96, 040303, (2017)

[118] C. Wang, F. Deng, Y. Li, X. Liu, G. Long. *Quantum secure direct communication with high-dimension quantum superdense coding.* Phys. Rev. A 71, 044305, (2005)

[119] S. Gröblacher, T. Jennewein, A. Vaziri, G. Weihs, A. Zeilinger. *Experimental quantum cryptography with qutrits.* New J. Phys., 8 , 75-75, (2006)

[120] B. P. Lanyon, *et.al. Simplifying quantum logic using higher-dimensional Hilbert spaces.* Nature Physics 5, 134 - 140, (2009)

[121] M. Huber, M. Pawlowski. *Weak randomness in device independent quantum key distribution and the advantage of using high dimensional entanglement.* Phys. Rev. A 88, 032309, (2013)

[122] M. Mirhosseini, O. Magaña-Loaiza, M. O'Sullivan, B. Rodenburg, M. Malik, M. Lavery, M. Padgett, D. Gauthier, R. Boyd. *High-dimensional quantum cryptography with twisted light.* New J. Phys. 17, 033033, (2015)

[123] G. Toth, O. Gühne. *Entanglement detection in the stabilizer formalism.* Phys. Rev. A 72, 022340, (2005)

[124] P. Kurzynski, T. Paterek, R. Ramanathan, W. Laskowski, D. Kaszlikowski. *Correlation Complementarity Yields Bell Monogamy Relations.* Phys. Rev. Lett. 106, 180402, (2011)

[125] M. Wiesniak, K. Maruyama. *Package of facts and theorems for efficiently generating entanglement criteria for many qubits.* Phys. Rev. A 85, 062315, (2012)

[126] J. Kaniewski, M. Tomamichel, S. Wehner. *Entropic uncertainty from effective anticommutators.* Phys. Rev. A 90, 012332, (2014)

[127] Ch. Eltschka, J. Siewert. *Quantifying entanglement resources.* J. Phys. A: Math. Theor. 47 424005, (2014)

[128] A. W. Harrow, A. Natarajan, X. Wu. *An Improved Semidefinite Programming Hierarchy for Testing Entanglement.* Commun. Math. Phys. 352: 881, (2017)

[129] D. Bruß. *Characterizing Entanglement.* J. Math. Phys. 43, 4237, (2002)

[130] H.F. Hofmann, S. Takeuchi. *Violation of local uncertainty relations as a signature of entanglement.* Phys. Rev. A 68, 032103, (2003)

[131] O. Gühne, M. Reimpell, R.F. Werner. *Estimating Entanglement Measures in Experiments.* Phys. Rev. Lett. 98, 110502, (2007)

[132] J. Eisert, F. Brandao, K. Audenaert. *Quantitative entanglement witnesses.* New J. Phys. 9, 46, (2007)

[133] Z. Ma, Z. Chen, J.-L. Chen, C. Spengler, A. Gabriel, M. Huber. *Measure of genuine multipartite entanglement with computable lower bounds.* Phys. Rev. A 83, 062325 (2011)

[134] J. Wu, H. Kampermann, D. Bruß, C. Klöckl, M. Huber. *Determining lower bounds on a measure of multipartite entanglement from few local observables.* Phys. Rev. A 86, 022319, (2012)

[135] S. Hashemi Rafsanjani, M. Huber, C. Broadbent, J. Eberly. *Genuinely multipartite concurrence of N-qubit X matrices.* Phys. Rev. A 86, 062303, (2012)

[136] C. Spengler, M. Huber, S. Brierley, T. Adaktylos, B. Hiesmayr. *Entanglement detection via mutually unbiased bases.* Phys. Rev. A 86, 022311, (2012)

[137] B. Hiesmayr, W. Löffler. *Complementarity reveals bound entanglement of two twisted photons.* New J. Phys. 15, 083036, (2013)

[138] D. Tasca, L. Rudnicki, R. Aspden, M. Padgett, P. Souto Ribeiro, S. Walborn. *Testing for entanglement with periodic coarse-graining.* Phys. Rev. A 97, 042312,(2018)

[139] E. Paul, D. Tasca, L. Rudnicki, S. Walborn. *Detecting entanglement of continuous variables with three mutually unbiased bases.* Phys. Rev. A 94, 012303, (2016)

[140] W. Wootters. *Entanglement of Formation of an Arbitrary State of Two Qubits.* Phys. Rev. Lett. 80, 2245, (1998)

[141] L. Gurvits. *Classical deterministic complexity of Edmonds' problem and quantum entanglement.* Proceedings of the thirty-fifth annual ACM symposium on Theory of computing, 10, (2003)

[142] J. de Vicente, C. Spee, B. Kraus. *Maximally Entangled Set of Multipartite Quantum States.* Phys. Rev. Lett. 111, 110502, (2013)

[143] K. Schwaiger, D. Sauerwein, M. Cuquet, J. de Vicente, B. Kraus. *Operational Multipartite Entanglement Measures.* Phys. Rev. Lett. 115, 150502, (2015)

[144] M. Huber, J. de Vicente. *Structure of Multidimensional Entanglement in Multipartite Systems.* Phys. Rev. Lett. 110, 030501, (2013)

[145] M. Huber, M. Perarnau-Llobet, J. de Vicente. *Entropy vector formalism and the structure of multidimensional entanglement in multipartite systems.* Phys. Rev. A 88, 042328, (2013)

[146] G. Toth, T. Moroder, O. Gühne. *Evaluating Convex Roof Entanglement Measures.* Phys. Rev. Lett. 114, 160501, (2015)

[147] D. Tasca, L. Rudnicki, R. Gomes, F. Toscano, S. Walborn.*Reliable Entanglement Detection Under Coarse–Grained Measurements.* Phys. Rev. Lett. 110, 210502, (2013)

[148] J. Schneeloch, P. Dixon, G. Howland, C. Broadbent, J. Howell. *Violation of Continuous Variable EPR Steering with Discrete Measurements.* Phys. Rev. Lett. 110, 130407, (2013)

[149] G. Howland, J. Howell. *Efficient high-dimensional entanglement imaging with a compressive sensing, double-pixel camera.* Phys. Rev. X 3, 011013, (2013)

[150] S. Walborn, C. Monken, S. Padua, P. Souto Ribeiro. *Spatial correlations in parametric down-conversion.* Physics Reports, vol. 495, (2010)

[151] M. Edgar, *et.al. Imaging high-dimensional spatial entanglement with a camera.* Nat. commun., 3, 984, (2012)

[152] N. Phan, M. Cheng, D. Bessarab, L. Krivitsky. *Interaction of Fixed Number of Photons with Retinal Rod Cells.* Physical Review Letters, 112(21), 213601, (2014)

[153] M. Jachura, R. Chrapkiewicz. *Shot-by-shot imaging of Hong?Ou?Mandel interference with an intensified sCMOS camera.* Optics letters, 40(7), 1540–1543, (2015)

[154] R. Chrapkiewicz, M. Jachura, K. Banaszek, W. Wasilewski. *Hologram of a single photon.* Nature Photonics 10, 576-579, (2016)

[155] Z. Xie, et al. *Harnessing high-dimensional hyperentanglement through a biphoton frequency comb.* Nature Photonics 9, 536–542, (2015)

[156] J. Roslund, R. De Araujo, S. Jiang, C. Fabre, N. Treps. *Wavelength-multiplexed quantum networks with ultrafast frequency combs.* Nature Photonics, 8(2), 109-112, (2014)

[157] S. Yokoyama, et al. *Ultra-Large-Scale Continuous-Variable Cluster States Multiplexed in the Time Domain.* Nature Photonics, 7(12), 982–986, (2013)

[158] M. Giustina, et al. *Significant-Loophole-Free Test of Bell?s Theorem with Entangled Photons.* Phys. Rev. Lett. 115, 250401, (2015)

[159] B. Hensen, et al. *Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres.* Nature volume 526, pages 682?686, (2015)

[160] L. Shalm, et al. *Strong Loophole-Free Test of Local Realism.* Phys. Rev. Lett. 115, 250402, (2015)

[161] J. Cadney, M. Huber, N. Linden, A. Winter. *Inequalities for the Ranks of Quantum States.* Linear Algebra and Applications, vol. 452, pp. 153-171, (2014)

[162] M. Malik, M. Erhard, M. Huber, M. Krenn, R. Fickler, A. Zeilinger. *Multi-photon entanglement in high dimensions.* Nature Photonics 10, 248-252, (2016)

[163] J. Maxwell. *Theory of Heat.* Longmans, Green, and Co., (1872)

[164] H. Leff, A. Rex (eds). *Maxwell's Demon: Entropy, Information, Computing.* Adam-Hilger, (1990)

[165] R. Landauer. *Irreversibility and heat generation in the computing process.* IBM Journal of Research and Development, vol. 5, pp. 183-191, (1961)

[166] J. Vaccaro, S. Barnett. *Information Erasure Without an Energy Cost.* Proc. R. Soc. A, 467 (2130): 1770-1778, (2011)

[167] C. Bennett. *Logical reversibility of computation.* IBM Journal of Research and Development, vol. 17, no. 6, pp. 525-532, (1973)

[168] C. Bennett. *The Thermodynamics of Computation – A Review.* International Journal of Theoretical Physics, vol. 21, no. 12, pp. 905-940, (1982)

[169] J. Goold, M. Huber, A. Riera, L. del Rio, P. Skrzypczyk. *The role of quantum information in thermodynamics — a topical review.* J. Phys. A: Math. Theor. 49, 143001, (2016)

[170] D. Bruschi, M. Perarnau-Llobet, N. Friis, K. Hovhannisyan, M. Huber. *The thermodynamics of creating correlations: Limitations and optimal protocols.* Phys. Rev. E 91, 032118, (2015)

[171] M. Perarnau-Llobet, K. Hovhannisyan, M. Huber, P. Skrzypczyk, N. Brunner, A. Acín. *Extractable Work from Correlations.* Phys. Rev. X 5, 041011, (2015)

[172] J. Geusic, E. Schulz-DuBois, H.Scovil. *Quantum Equivalent of the Carnot Cycle.* Phys. Rev.156, 343, (1967)

[173] R. Kosloff, A. Levy. Annu. *Quantum heat engines and refrigerators: continuous devices.* Rev. Phys. Chem.65, 365, (2014)

[174] F. Clivaz, R. Silva, G. Haack, J. Bohr Brask, N. Brunner, M. Huber. *Unifying paradigms of quantum refrigeration: how resource-control determines fundamental limits.* arXiv:1710.11624

[175] A. Ghosh, V. Mukherjee, W. Niedenzu, G. Kurizki. *Are quantum thermodynamic machines better than classical?* arXiv:1804.05659

[176] A. Peres. *Unperformed experiments have no results.* American Journal of Physics 46, 745 (1978)

[177] G. Chaitin, A. Arslanov, C. Calude. *Program-size Complexity Computes the Halting Problem.* Bulletin of the EATCS 57, (1995)

[178] A. Kolmogorov. *Three approaches to the quantitative definition of information.* Problemy Peredachi Informatsii, Vol. 1, No. 1, pp. 3?11, (1965)

[179] M. Li, P. Vitanyi. *An Introduction to Kolmogorov Complexity and Its Applica-tions.* Springer-Verlag, (2008)

[180] S. Wolf. *An all-or-nothing flavor to the Church-Turing hypothesis.* roceedings of 14th Annual Conference on Theory and Applications of Models of Computation (TAMC 2017), (2017)

[181] S. Wolf. *Second thoughts on the second law.* Adventures Between Lower Bounds and Higher Altitudes, Lecture Notes in Computer Science, LNCS 11011, (2018)

[182] P. Gacs, J. Tromp, P. Vitanyi. *Algorithmic statistics.* IEEE Transactions on Information Theory, Vol. 47, No. 6, pp. 2443?2463, (2001)

[183] M. Müller. *Quantum Kolmogorov Complexity and the Quantum Turing Machine.* PhD thesis, Technical University of Berlin, (2007)

[184] M. Müller. *Law without law: from observer states to physics via algorithmic information theory.* arXiv:1712.01826

[185] M. Müller. *Could the physical world be emergent instead of fundamental, and why should we ask? (short version).* arXiv:1712.01816

[186] P. Erker. *The Quantum Hourglass - approaching time measurement with quantum information theory.* Master Thesis, ETH Zurich, (2014)

[187] A. Hermann, K. Meyen, V. Weisskopf. *Wolfgang Pauli, Scientific Correspondence with Bohr, Einstein, Heisenberg.* Springer, Berlin, (1979)

[188] W. Pauli. *Handbuch der Physik.* Volume 23, p.1-278, Springer, Berlin, (1926)

[189] T. Nicholson, S. Campbell, R. Hutson, G. Marti, B. Bloom, R. McNally, W. Zhang, M. Barrett, M. Safronova, G. Strouse, W. Tew, J. Ye. *Systematic evaluation of an atomic clock at* 2×10^{-18} *total uncertainty.* Nature Comm. 6, 6896, (2015)

[190] N. Hinkley, J. Sherman, N. Phillips, M. Schioppo, N. Lemke, K. Beloy, M. Pizzocaro, C. Oates, A. Ludlow. *An atomic clock with* 10^{-18} *instability.* Science v. 341: 1215-1218, (2013)

[191] C. Chou, D. Hume, J. Koelemeij, D. Wineland, T. Rosenband. *Frequency Comparison of Two High-Accuracy* Al^+ *Optical Clocks.* Phys. Rev. Lett. 104, 070802, (2010)

[192] W. Pauli. *Handbuch der Physik.* Volume 23, p.1-278, Springer, Berlin, (1926)

[193] L. Mandelstam, I. Tamm. *The Uncertainty Relation Between Energy and Time in Non-relativistic Quantum Mechanics.* J. Phys USSR 9, 249-254, (1945)

[194] N. Margolus, L. Levitin. *The maximum speed of dynamical evolution.* Physica D 120, 188-195, (1998)

[195] P. Kosinski, M. Zych. *Elementary proof of the bound on the speed of quantum evolution.* Phys. Rev. A 73, 024303, (2006)

[196] I. Marvian, R. W. Spekkens, P. Zanardi. *Quantum speed limits, coherence, and asymmetry.* Phys. Rev. A 93, 052331, (2016)

[197] D. Piers, M. Cianciaruso, L. Céleri, G. Adesso, D. Soares-Pinto. *Generalized Geometric Quantum Speed Limits.* Phys. Rev. X 6 021031, (2016)

[198] A. Miyake. *Entropic time endowed in quantum correlation.* arXiv:1111.2855

[199] Ä. Baumeler, S. Wolf. *Causality - Complexity - Consistency: Can Space-Time Be Based on Logic and Computation?* Birkhäuser, (2017)

[200] D. Page, W. Wootters. *Evolution without evolution: Dynamics described by stationary observables.* Phys. Rev. D 27, 2885, (1983)

[201] W. Wootters. *Time replaced by quantum correlation.* Int. J. Phys 23, 701, (1984)

[202] V. Giovannetti, S. Lloyd, L. Maccone. *Quantum time.* Phys. Rev. D 92, 045033, (2015)

[203] A. Peres, *Measurement of time by quantum clocks.* American Journal of Physics 48, 552, (1980)

[204] J. Lindkvist, C. Sabin, G. Johansson, I. Fuentes. *Motion and gravity effects in the precision of quantum clocks.* Sci. Rep. 5, 10070, (2015)

[205] V. Buzek, R. Derka, S. Massar, *Optimal quantum clocks.* Phys. Rev. Lett. 82, 2207, (1999)

[206] M. Woods, R. Silva, J. Oppenheim. *Autonomous quantum machines and finite sized clocks.* arXiv:1607.04591

[207] S. Rankovic, Y.-C. Liang, R. Renner. *Quantum clocks and their synchronisation - the Alternate Ticks Game.* arXiv:1506.01373

[208] L. Boltzmann. *On certain questions of the theory of gases.* Nature 51, 413-415, (1895)

[209] N. Linden, S. Popescu, P. Skrzypczyk, *How small can thermal machines be? The smallest possible refrigerator.* Phys. Rev. Lett. 105, 130401, (2010)

[210] A. Levy, R. Kosloff. *Quantum Absorption Refrigerator.* Phys. Rev. Lett. 108, 070604, (2012)

[211] N. Brunner, N. Linden, S. Popescu, P. Skrzypczyk. *Virtual qubits, virtual temperatures, and the foundations of thermodynamics.* Phys. Rev. E 85, 051117, (2012)

[212] C. Gogolin, J. Eisert. *Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems.* Rep. Prog. Phys. 79, 056001, (2016)

[213] E. T. Jaynes. *Information theory and statistical mechanics I.* Phys. Rev. 106, 620, (1957)

[214] F. Brandao, M. Horodecki, N. Ng, J. Oppenheim, S. Wehner. *The second laws of quantum thermodynamics.* PNAS 112, 3275, (2015)

[215] M. Schwarzhans. *The fundamental energy cost of letting a clock tick.* Master Thesis, University of Vienna, (2018)

[216] M. Woods, R. Silva, G. Pütz, S. Stupar, R. Renner. *Quantum clocks are more accurate than classical ones.* arXiv:1806.00491

[217] R. Silva, G. Manzano, P. Skrzypczyk and N. Brunner, *Performance of autonomous quantum thermal machines: Hilbert space dimension as a thermodynamic resource.* Phys. Rev. E **94**, 032120, (2016)

[218] S. Nakajima. *On Quantum Theory of Transport Phenomena* (in German), Progress of Theoretical Physics 20 (6): pp. 948-959, (1958)

[219] R. Zwanzig. *Ensemble Method in the Theory of Irreversibility*(in German), Journal of Chemical Physics 33 (5): pp. 1338-1341, (1960)

[220] H. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*, Oxford University Press, Oxford, (2007)

[221] D. Reeb and M. Wolf. *An improved Landauer Principle with finite-size corrections.* New J. Phys. 16, 103011, (2014)

[222] R. Li, K. Gibble, K. Szymaniec. *Improved accuracy of the NPL-CsF2 primary frequency standard: evaluation of distributed cavity phase and microwave lensing frequency shifts.* Metrologia 48, 5, (2011)

[223] R. Lutwak. *The Chip-Scale Atomic Clock - Prototype Evaluation.* 36th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, (2007)

[224] L. Mlodinow , T. Brun. *Relation between the psychological and thermodynamic arrows of time.* Phys. Rev. E 89, 052102, (2014)

[225] L. Maccone. *A quantum solution to the arrow-of-time dilemma.* Phys. Rev. Lett. 103, 080401, (2009)

[226] Y. Guryanova, S. Popescu, A. Short, R. Silva, P. Skrzypczyk. *Thermodynamics of quantum systems with multiple conserved quantities.* Nature Commun. 7, 12049, (2016)

[227] N. Halpern, P. Faist, J. Oppenheim, A. Winter. *Microcanonical and resource-theoretic derivations of the thermal state of a quantum system with noncommuting charges.* Nature Communications 7, 12051, (2016)

[228] M. Lostaglio, D. Jennings, T. Rudolph. *Thermodynamic resource theories, non-commutativity and maximum entropy principles.* New J. Phys. 19 043008, (2017)

[229] T. Langen, S. Erne, R. Geiger, B. Rauer, T. Schweigler, M. Kuhnert, W. Rohringer, I. Mazets, T. Gasenzer, J. Schmiedmayer. *Experimental observation of a generalized Gibbs ensemble.* Science 348, 207-211, (2015)

[230] W. Pusz, S. Woronowicz. *Passive states and KMS states for general quantum systems.* Comm. Math. Phys. 58, no. 3, 273–290, (1978)

[231] S. Campbell, S. Deffner. *Trade-off between speed and cost in shortcuts to adiabaticity.* Phys. Rev. Lett. 118, 100601, (2017)

[232] A. Malabarba, A. Short, P. Kammerlander. *Clock-driven quantum thermal engines.* New J. Phys. 17 045027, (2015)

[233] R. Jozsa, D. Abrams, J. Dowling, C. Williams. *Quantum clock synchronization based on shared prior entanglement.* Phys. Rev. Lett. 85, 2010, (2000)

[234] P. Komar, E. Kessler, M. Bishof, L. Jiang, A. Sorensen, J. Ye and M. Lukin. *A quantum network of clocks.* Nature Physics 10, 582-587, (2014)

[235] A. Hansen, S. Wolf. *The Measurement Problem Is the "Measurement" Problem.* arXiv:1810.04573

[236] C. Hidalgo. *Why Information Grows: The Evolution of Order, from Atoms to Economies.* Basic Books, (2015)

[237] D. Hofstadter. *Gödel, Escher, Bach. An Eternal Golden Braid.* Basic Books, (1979)

Appendix A

Cotutela

Here a scan of the original contract established between the Universitá della Svizzera Italiana and the Universitat Autonoma de Barcelona is included.

Università della Svizzera italiana

Faculty of Informatics

UAB
Universitat Autònoma de Barcelona
Escola de Doctorat

Agreement for international joint supervision of a thesis

between

UNIVERSITÀ DELLA SVIZZERA ITALIANA (USI), represented by the President, Piero Martinoli

and

The UNIVERSITAT AUTONOMA DE BARCELONA (UAB), represented by Mr. Ferran Sancho Pifarré, Rector of the Universitat Autònoma de Barcelona (UAB), by virtue of the powers derived from article 75, paragraph m) of the UAB Statutes, and in her name Mrs. Xavier Gabarrell Durany, vice-rector of Research, 2011, and with the functions conferred upon him,

relating to the joint supervision of the thesis of Mr Paul ERKER.

The setting up of the doctoral thesis co-tutorship is based on the following legal prerequisites at USI:

· General rules and recommendations of the CRUS for co-tutelle programs.
· University Statutes
· PhD Program Regulations of the Faculty of Informatics at USI

And with respect to UAB, each candidate must comply with the academic and admission requirements established for the PhD courses or programme to which the student has been admitted, along with the general regulations established by the

following Royal Decrees:

· 99/2011, of 28 of January, which regulates the tribunal, the presentation and the evaluation of the PhD theses

Title 1 – Administrative procedure

1-1 The UAB and USI agree, in accordance with the applicable laws, rules and regulations in force in each of their respective countries and Institutions, to jointly organise a doctoral thesis for the benefit of the following doctoral candidate:

name and surname: Paul Erker

date of birth: 05.05.1987

place of birth: Klagenfurt, Carinthia, Austria

theme of doctoral thesis (preliminary title): The role of Information in Physics

The principles governing the joint doctoral thesis as well as the relevant administrative procedures and education related matters are established in this agreement.

1-2 The enrolment in the jointly supervised doctoral thesis of Mr Paul ERKER in the doctoral school of the USI and in the Universitat Autonoma de Barcelona will be valid from the moment of its formalisation. It will be renewed each year during the jointly supervised research work in both institutions.

1-3 Students must formalise their enrolment on PhD courses by paying enrolment fees to contracting institutions according to the chosen regulations. The student must pay his/her annual enrolment fees by satisfying the administrative rates of the contracting institutions. The academic protection-tuition shall be paid alternating every year between the two institutions. Starting in relation to the Universitat Autonoma de Barcelona for the 1st academic year of PhD, where his admission was accepted on the 15.12.2014.

1-4 For his registration, the doctoral student will have to produce all the written proof of address, income, social security and civil liability coverage in each of these countries.

1-5 Time spent preparing the thesis will be divided between the UNIVERSITÀ DELLA SVIZZERA ITALIANA and the Universitat Autonoma de Barcelona with alternate periods depending on the research progress.
The different periods during which the candidate is resident at the participating

universities as part of this co-tutorship will be distributed in accordance with the following schedule:

-In relation to the UAB, the 1st academic year of PhD.
-In relation to the USI , the remaining academic years of PhD.

Title 2 - Teaching

2-1 Thesis work will be jointly supervised by:

-Mr Stefan WOLF of the USI
Title and function: Professor - Thesis supervisor
Laboratory: Faculty of Informatics

-Mr Andreas WINTER of the UAB
Title and function: Professor - Thesis supervisor
Laboratory: Física Teòrica: Informació i Fenòmens Quàntics, Grup d'Informació Quàntica.

-Mr Marcus HUBER of the UAB
Title and function: Postdoctoral Researcher - Thesis supervisor
Laboratory: Física Teòrica: Informació i Fenòmens Quàntics, Grup d'Informació Quàntica.

All three commit to exercise fully their function of director towards the doctoral student. Thesis work will be focused on the role of information in physics.

2-2 When the doctoral student has to validate complementary trainings (scientific or professional), the institutions will specify the mutual recognition procedures, in agreement with the thesis supervisors and the doctorate students.

Title 3 – Oral examination

3-1 The jury will be composed from 3 members minimum and to 8 members maximum. It will consist of an equal number of scientists from both institutions and include besides a majority of external members to both institutions. The supervisors of the thesis only will be able to participate in the 6 or more members tribunal. 2 members will also be included in the tribunal as substitutes.
3-2 The thesis will be written in English and the oral summary will be in English.

3-3 A single oral examination of the thesis will be required in English and will take place at the UNIVERSITÀ DELLA SVIZZERA ITALIANA.
Both institutions will recognize the validity of the examination.

The jury chairperson will write a defence report, signed by all the jury members. The tribunal will make its evaluation in accordance with the regulations of the university at which the thesis is presented. In any case, when a thesis is not presented at the UAB, it will still be necessary for the presentation report of the UAB to be completed and signed.

The subject of the thesis must be protected, as well as its submission, identification and reproduction in compliance with the specific procedures applicable in each country involved in joint supervision of the doctorate.

3-4 Each of the two Institutions undertakes to award a doctoral degree for the same thesis following a favourable report issued by the examination commission.

The UAB will award a "doctor per la UAB en el programa de doctorat en física".

The USI will award a doctoral degree in Computational Sciences.

The degree is issued mentioning the agreement within USI and UAB for a joint doctoral thesis.

3-5 On the basis of a single presentation of a PhD thesis, both universities agree to award the corresponding title of PhD following payment of the issuance fees. The PhD certificates issued by either of the universities will mention the co-tutorship.
The diploma(s) will bear the information as established in the regulation applicable in each university.

Title 4 – Conclusion

4–1 Doctoral candidates shall observe the rules and customs of the host Institution.

4-2 The contracting Institutions, through the offices of their respective thesis supervisors, undertake to notify each other of all the information and documentation useful for the purposes of organising the joint thesis that is the subject matter of this present agreement.

4-3 The presentation, deposit and reproduction of the thesis shall be done in each and every country in accordance with the applicable regulations in force. The protection of

the subject matter of the thesis and likewise the publication, exploitation and protection of the results obtained by the candidate's research in the contracting Institutions shall be subject to the applicable law in force and guaranteed in compliance with the specific procedures in this regard of each of the countries involved in the joint thesis.

If requested, the provisions in connection with intellectual property rights may be agreed in specific protocols or documents.

4-4 This agreement shall be effective as and from the date of its execution by the authorised representative of each contracting Institution and shall be valid until the end of the academic year during which the thesis or study will be orally defended.

In the event that the candidate does not register in one or other of the contracting Institutions, renounces in writing or is not authorised to continue researching and writing the thesis by virtue of a decision made by one of the two thesis supervisors, the contracting Institutions shall jointly and without delay terminate this agreement.

4-5 This agreement is drawn up in three originals in English, which have binding legal force.

For the UNIVERSITÀ DELLA SVIZZERA ITALIANA

Prof. Dr. Piero Martinoli
President
(date, signature)

Prof. Dr. Stefan WOLF
Joint thesis supervisor
(date, signature)

Prof. Dr. Michele ~~LANZA~~ Kai HORMANN
Dean Faculty of Informatics
(date, signature)

For the Universitat Autonoma de Barcelona

15/10/15

Prof. Xavier Gabarrell Durany
Academic Secretary of the Doctoral School
(date, signature)

Prof. Andreas WINTER
Joint thesis supervisor
(date, signature)

Dr. Marcus HUBER
Joint thesis supervisor
(date, signature)

Prof. Santiago Suriñach Cornet
Director of the Department of Physics
(date, signature)

Prof. Jordi Mompart
Coordinator of the PhD of Physics
(date, signature)

Paul ERKER
The doctoral student
(date, signature)

Appendix B

Curriculum Vitae

Current Affiliations

Università della Svizzera italiana (USI)
Via Giuseppe Buffi 13
6900 Lugano
Ticino
Switzerland
&
Universitat Autònoma de Barcelona (UAB)
Plaça Cívica
08193 Cerdanyola
Barcelona
Spain

Education

2007-2011 University of Vienna
BSc in Physics
Thesis: *Genuine Vielteilchenverschränkung und Dicke-Zustände* (in german)
Advisor: Beatrix Hiesmayr

2011-2012 University of Vienna
MSc in Physics

2012-2014 ETH Zürich
MSc in Physics
Thesis: *The Quantum Hourglass - Approaching time measurement with Quantum Information Theory*
Advisor: Renato Renner

2015- ongoing Universitat Autònoma de Barcelona
PhD in Physics
Advisors: Andreas Winter, Marcus Huber

2016- ongoing Università della Svizzera italiana
PhD in Informatics
Advisor: Stefan Wolf

Conferences and academic visits

June 2010 7th Central European Quantum Information Processing Workshop, Valtice, Czech Republic

January 2011 Testing Foundations of Quantum Mechanics at Different Energy Scales, Schneeberg, Austria

January 2011 14th Workshop on Quantum Information Processing, Singapore, Singapore

June 2011 8th Central European Quantum Information Processing Workshop), Znojmo, Czech Republic

April 2012 Quantum Malta 2012: Fundamental Problems in Quantum Physics, Valletta, Malta

June 2012 9th Central European Quantum Information Processing Workshop, Smolenice, Slovakia

December 2012 Workshop on "Theory Meets Experiment: Finding Proposals Towards Testing Foundational Issues in Particle Physics", Vienna, Austria

January 2013 IONS 2013, Zurich/Lausanne , Switzerland

February 2013 Universitat Autònoma de Barcelona/ Institute of Photonic Sciences, Barcelona, Spain

June 2013 10th Central European Quantum Information Processing Workshop, Valtice, Czech Republic

August 2013 Centro de Ciencias de Benasque Pedro Pascual, Benasque, Spain

February 2014 17th Workshop on Quantum Information Processing, Barcelona, Spain

February 2014 Universitat Autònoma de Barcelona, Barcelona, Spain

May 2014 Beyond I.I.d in information theory, Singapore, Singapore

May 2014 Theory of Quantum Computation, Communication and Cryptography, Singapore, Singapore

June 2014 NCCR QSIT Junior Meeting, Passugg, Switzerland

June 2014 Quantum [Un]Speakables II: 50 Years of Bell's Theorem, Vienna, Austria

June 2014 11th Central European Quantum Information Processing Workshop, Znojmo, Czech Republic

June 2014 University of Vienna, Vienna, Austria

November 2014 Universitat Autònoma de Barcelona, Barcelona, Spain

February 2015 Jornada d'Investigadors Predoctorals Interdisciplinària 2015, Barcelona, Spain

April 2015 COST conference on Quantum Thermodynamics, Mallorca, Spain

May 2015 New Trends in Complex Quantum System Dynamics, Cartagena, Spain

June 2015 Quantum Information Workshop, Benasque, Spain

August 2015 V Quantum Information School and Workshop, Paraty, Brazil

September 2015 Time in Physics, Zurich, Switzerland

November 2015 University of Siegen, Siegen, Siegen

February 2016 University of Vienna, Vienna, Austria

February 2016 Universitat Autònoma de Barcelona, Barcelona, Spain

April 2016 Institute of Science and Technology Austria, Klosterneuburg, Austria

April 2016 University of Geneva, Geneva, Switzerland

May 2016 University of Geneva, Geneva, Switzerland

June 2016 Institut für Quantenoptik und Quanteninformation, Vienna, Austria

June 2016 3rd Seefeld Quantum Information Workshop, Seefeld, Austria

September 2016 University of Geneva, Geneva, Switzerland

October 2016 Young Quantum Information Scientists, Barcelona, Spain

November 2016 University of Bern, Bern, Switzerland

December 2016 Mathematical Foundations of Quantum Thermodynamics, Smolenice, Slovakia

February 2017 QSIT Meeting, Arosa, Switzerland

February 2017 Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

March 2017 Institut für Quantenoptik und Quanteninformation, Vienna, Austria

July 2017 Institut für Quantenoptik und Quanteninformation, Vienna, Austria

July 2017 Quantum Information Workshop, Benasque, Spain

August 2017 VI Quantum Information Workshop, Paraty, Brazil

August 2017 University of Cordoba, Cordoba, Agentina

August 2017 University of Buenos Aires, Buenos Aires, Argentina

August 2017 Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

September 2017 International symposium "Worlds of Entanglement", Brussels, Belgium

October 2017 Young Quantum Information Scientists, Erlangen, Germany

February 2018 Institut für Quantenoptik und Quanteninformation, Vienna, Austria

February 2018 QSIT meeting, Arosa, Switzerland

September 2018 Young Quantum Information Scientists, Vienna, Austria

Posters and Talks

December 2010 Talk at the Vienna Theory Lunch Seminar

January 2011 Co-Author of a Poster at 14th Workshop on Quantum Information Processing (QIP)

June 2011 Poster at the 8th Central European Quantum Information Processing Workshop (CEQIP)

March 2012 Talk at the Quantum Information Group seminar at the University of Bristol

April 2012 Poster at Quantum Malta 2012: Fundamental Problems in Quantum Physics

June 2012 Poster at the 9th Central European Quantum Information Processing Workshop (CEQIP)

January 2013 Poster and talk at IONS 2013, Zurich/Lausanne , Switzerland

June 2013 Poster at the 10th Central European Quantum Information Processing Workshop (CEQIP)

May 2014 Poster at Beyond I.I.d in information theory

May 2014 Poster at Theory of Quantum Computation, Communication and Cryptography

June 2014 Talk at NCCR QSIT Junior Meeting

June 2014 Poster at Quantum [Un]Speakables II: 50 Years of Bell's Theorem

November 2014 Talk at the Quantum Information Group seminar at the Universitat Autònoma de Barcelona

February 2015 Talk at ornada d'Investigadors Predoctorals Interdisciplinària (Best Talk award)

April 2015 Poster at COST conference on Quantum Thermodynamics

May 2015 Poster at New Trends in Complex Quantum System Dynamics

August 2015 Poster at V Quantum Information Workshop

September 2015 Poster at Time in Physics

November 2015 Talk at University of Siegen

June 2016 Poster at 3rd Seefeld Quantum Information Workshop

October 2016 Talk at Young Quantum Information Scientists

December 2016 Talk at Mathematical Foundations of Quantum Thermodynamics

February 2017 Talk at QSIT Meeting in Arosa

February 2017 Talk at Universidade Federal de Minas Gerais

August 2017 Talk at VI Quantum Information Workshop in Paraty

August 2017 Talk at University of Cordoba

August 2017 Talk at University of Buenos Aires

September 2017 Talk at international symposium "Worlds of Entanglement"

October 2017 Talk at Young Quantum Information Scientists

Scholarships and Grants

January 2011 Student support for the 14th Workshop on Quantum Information Processing (QIP)

April 2012 Student support for Quantum Malta 2012: Fundamental Problems in Quantum Physics

June 2012 Young scientist funding from University of Vienna for the 9th Central European Quantum Information Processing Workshop (CEQIP)

December 2012 Student support for workshop on "Theory Meets Experiment: Finding Proposals Towards Testing Foundational Issues in Particle Physics", Vienna, Austria

September 2013 Excellence Grant for studying abroad by the Federation of Carinthian Industries, Klagenfurt, Austria

May 2017 Cotutelles de thèse Grant by swissuniversities, Lugano, Switzerland

List of Publications

P1 Marcus Huber, Paul Erker, Hans Schimpf, Andreas Gabriel, Beatrix C. Hiesmayr
Experimentally feasible set of criteria detecting genuine multipartite entanglement in n-qubit Dicke states and in higher dimensional systems
Phys. Rev. A **83**, 040301(**R**) (2011)

P2 Paul Erker
How not to Renyi generalize the Quantum Conditional Mutual Information
J. Phys. A: Math. Theor. 48 (2015) 275303

P3 Paul Erker
The Quantum Hourglass
https://e-collection.library.ethz.ch/view/eth:48076

P4 Paul Erker, Mario Krenn, Marcus Huber
Quantifying high dimensional entanglement with two mutually unbiased bases
Quantum 1, 22 (2017)

P5 Ali Asadian, Paul Erker, Marcus Huber, Claude Klöckl
Heisenberg-Weyl Observables: Bloch vectors in phase space
Phys. Rev. A 94, 010301(R)(2016)

P6 Paul Erker, Mark T. Mitchison, Ralph Silva, Mischa P. Woods, Nicolas Brunner, Marcus Huber
Autonomous quantum clocks: how thermodynamics limits our ability to measure time
Phys. Rev. X 7, 031022 (2017)

P7 Jessica Bavaresco, Natalia Herrera Valencia, Claude Klöckl , Matej Pivoluska, Paul Erker, Nicolai Friis, Mehul Malik, Marcus Huber
Measurements in two bases are sufficient for certifying high-dimensional entanglement
Nature Physics, vol. 14 (2018)

Supervision

Master thesis Co-Supervision
Emanuel Schwarzhans
The fundamental energy cost of letting a clock tick
University of Vienna
August 2018

Organisation of conferences and workshops

Several small workshops and talks Facoltà indipendente di Gandria, Switzerland
since 2016

Young Quantum Information Scientists University of Vienna, Austria
September 2018

Teaching activities

2011-2012 University of Vienna
Teaching assistant for physics and mathematics course

2011-2012 University of Vienna
Sophomore for physics courses

2013-2014 ETH Zurich
Teaching assistant for Physics Laboratory

2015-2016 USI
Teaching assistant for Information & Physics

2015-2016 USI
Teaching assistant for Probability & Statistics

2016-2017 USI
Teaching assistant for Discrete Structures

2017 USI
Teaching assistant for Quantum Computing

2017 ETH Zurich
Teaching assistant for Quantum Information and Cryptography

2018 USI
Teaching assistant for Quantum Computing

2018 ETH Zurich
Teaching assistant for Information and Physics