QUANTUM INFORMATION THEORY OF MEASUREMENT

By

Jennifer Ranae Glick

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics—Doctor of Philosophy

2017

ABSTRACT

QUANTUM INFORMATION THEORY OF MEASUREMENT

By

Jennifer Ranae Glick

Quantum measurement lies at the heart of quantum information processing and is one of the criteria for quantum computation. Despite its central role, there remains a need for a robust quantum information-theoretical description of measurement. In this work, I will quantify how information is processed in a quantum measurement by framing it in quantum information-theoretic terms. I will consider a diverse set of measurement scenarios, including weak and strong measurements, and parallel and consecutive measurements. In each case, I will perform a comprehensive analysis of the role of entanglement and entropy in the measurement process and track the flow of information through all subsystems. In particular, I will discuss how weak and strong measurements are fundamentally of the same nature and show that weak values can be computed exactly for certain measurements with an arbitrary interaction strength. In the context of the Bell-state quantum eraser, I will derive a trade-off between the coherence and "which-path" information of an entangled pair of photons and show that a quantum information-theoretic approach yields additional insights into the origins of complementarity. I will consider two types of quantum measurements: those that are made within a closed system where every part of the measurement device, the ancilla, remains under control (what I will call unamplified measurements), and those performed within an open system where some degrees of freedom are traced over (amplified measurements). For sequences of measurements of the same quantum system, I will show that information about the quantum state is encoded in the measurement chain and that some of this information is "lost" when the measurements are amplified—the ancillae become

equivalent to a quantum Markov chain. Finally, using the coherent structure of unamplified measurements, I will outline a protocol for generating remote entanglement, an essential resource for quantum teleportation and quantum cryptographic tasks.

Dedicated in memory of my beloved father, Kevin White Patterson.

ACKNOWLEDGMENTS

I will always feel extraordinarily lucky and grateful to my advisor, Chris Adami, who graciously welcomed me into his group. He gave me great freedom and support to pursue what interested me most in physics, despite his research commitments being mainly in the biological sciences. I am truly inspired by his multidisciplinary yet holistic approach to science, using his training as a theoretical nuclear physicist to blaze exciting new trails in biology, artificial intelligence, and, yes, even black hole physics (information theory being the unifying tool he artfully threads through these seemingly disparate fields). As anyone who knows him can readily attest, his exceptional zest for science and discovery is infectious, and made my time at MSU immeasurably more rewarding. I dare say I even managed to learn some evolutionary biology from the Adami Lab along the way.

I would like to thank Drs. Norman O. Birge, Phil Duxbury, Brian O'Shea, and Scott Pratt, for serving on my committee and for their guidance throughout my dissertation research. I am especially grateful to Scott for the kindness and support he extended to me from the first day I arrived at MSU. I learned much from many outstanding professors, in and out of the classroom, in particular from Norman, Phil, Carl Schmidt, and Vladimir Zelevinsky. I gratefully acknowledge the financial support from Michigan State University and the Physics Department through my University Enrichment Fellowship, Dissertation Completion Fellowship, and several travel grants, which minimized teaching burdens and enabled me to focus on research. I owe much of my initial scientific development to Lincoln D. Carr and the incredible physics faculty at the Colorado School of Mines, who inspired and equipped me for this journey over six years ago.

I will miss the rollicking pick-up basketball games with fellow MSU physics students, especially Victor Aguilar, Joe Williams, and Bill Martinez. Thanks also to Bill and Aimee Shore for organizing (and how!) and hosting the Thursday game nights and for introducing me to the inner world of board games, which, as it turns out, extends far beyond Monopoly and Battleship. Finally, I would like to thank Rachel Graef, my dearest friend of over 15 years, for her love and encouragement and for always making me feel that no time at all had passed whenever I visited home.

I would like to acknowledge and thank my whole family, in particular, my parents, Kim and Ed, my siblings, Alex, Andrew, and Katie, and their spouses, Caroline and Susie, as well as the entire Glick family. I am especially grateful to my Mom who has encouraged and supported me in my every ambition. My family has been an endless source of strength for me during the last six years. I suspect they would not have let me leave Colorado if they (and, honestly, I) had realized just how long a Ph.D. would really be. The last years were, thankfully, much impurroved by the companionship of *F. c.* Fermi and Nambu.

Beyond all of the guidance and support I have been so fortunate to receive from those listed thus far, the person who deserves the foremost acknowledgement is my husband, Joseph Glick. I would not have achieved the things I have without his unwavering belief in me. He has been my strongest advocate and, despite contending with his own set of challenges as a Ph.D. student, has always had limitless reserves of encouragement and insight for me to draw from. Who could have foreseen our trajectory, whose humble origin traces back to that eternally Golden Spring?

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO SYMBOLS

$\rho(A)$	Density matrix of system A
$S(A)$	The von Neumann entropy of system A
$H[p]$	Shannon entropy of a probability distribution p_i
$S(A\|B)$	Conditional entropy of A, given B
$S(A{:}B)$	Mutual entropy of A and B
$S(A{:}B\|C)$	Conditional mutual entropy of A and B, given C
$S(A{:}B{:}C)$	Ternary mutual entropy of A, B and C
Q	The quantum system
R	Reference system
A_i	Ancilla used to measure a quantum system
D_i	Device used to measure (amplify) ancilla A_i
$U^{(i)}$	Transformation between the eigenbases of observables i and $i-1$.
S_i	Shorthand for the marginal entropy of A_i or D_i
S_{ij}	Shorthand for the joint entropy of $A_i\, A_j$

Chapter 1

Introduction

1.1 Introduction and Historical Overview

John Archibald Wheeler famously suggested that information is fundamental to the physical world with his 1990 "it from bit" doctrine [1]. In our "participatory universe", we understand the physical world by interacting with it, and we obtain information about its state through measurement. When the systems we probe are quantum in nature, so too is the information they contain. The study of quantum information has grown during the last several decades into an extensive and versatile field known generally as quantum information science. It now encompasses diverse areas including quantum communication and cryptography, quantum error correction, and quantum computation. Its far-reaching technological implications become more evident each year, while its most profound impacts are yet to be felt. The remarkable progress in this field has promised dramatic impacts to computing and, therefore, to the global economy. For instance, the development of large-scale quantum processors will allow for the efficient simulation of complex systems in quantum chemistry, the design of new materials and medicine, and has created the need for future quantum-proof cybersecurity systems.

An appreciation for quantum information is perhaps best formed by first recalling its classical counterpart. Formulated by Claude Shannon in his set of 1948 papers [2, 3], classical information theory quantifies information in terms of zeros and ones, bits, and describes its transmission over communication channels. Quantum information theory (QIT), on the other hand, unifies the quantum behavior of *physical* systems with *information* theory, a prime example being entanglement. While much of quantum information processing relies on consuming this non-classical phenomenon as a physical resource (e.g., in quantum teleportation), entanglement can be understood and quantified entirely with the tools provided by quantum information theory. Instead of the bit, the basic unit of information is now the quantum bit, or "qubit", and its state can be a superposition of the classical bit values. As a result, an n-qubit system can store all 2^n possible bit values, instead of just one at a time. This "parallelism" is what would allow a quantum device with roughly 50 logical qubits to execute certain tasks that no existing classical supercomputer could perform in a realistic period of time [4,5]. In fact, a 300-qubit device would have a state space of dimension larger than the number of atoms in the universe.

Quantum information theory began to develop rapidly in the 1990s. Two complementary quantum protocols foundational to quantum communication were discovered at this time: superdense coding [6] in 1992, in which two bits of classical information can be communication using one qubit, followed by teleportation [7] in 1993, where one qubit is transferred by communicating two classical bits. Quantum error-correcting codes, which are a necessary component of fault-tolerant quantum devices, were developed by Calderbank, Shor, and Steane starting in 1995. Their work was followed quickly by important papers from several other groups (see, e.g., Refs. [8–14]). These codes showed that it is possible to have reliable communication over noisy quantum channels and effective quantum computation in the presence of decoherence. Today, research groups around the world are continually redefining the state of the art in quantum information processing. There are now several viable technologies for building quantum devices (see, e.g., [15] for a short review), including trapped ions, quantum dots, superconducting qubits, and exotic topological qubits, along with successful demonstrations of entangling multiple qubits and implementations of single- and two-qubit logic gates.

Quantum measurement is ubiquitous to quantum information processing. As part of the DiVincenzo criteria [16], it is one of the key requirements for building a quantum computer. For instance, after an algorithm is run on a quantum device, a measurement is performed and the result of the computation is read out. In fault-tolerant devices employing error correction, syndrome measurements are performed to correct errors acquired by qubits during the computations. And, the standard quantum teleportation protocol relies on quantum measurement to transfer a quantum state to a remote receiver. Part of the challenge in realizing large-scale quantum processors is improving the measurement process to reduce unwanted coupling to the device. Such interactions interfere with the ability to manipulate and coherently control the device's qubits. From this perspective, it is essential to have a robust quantum information-theoretical understanding of how information and entanglement are processed in quantum measurement.

Over the last several decades many models of quantum measurement have been proposed. Some of the most well-known include the standard Copenhagen model, Everett's relative state formalism [17], the two-state vector formalism [18], Griffiths' consistent histories [19], and Zurek's environment-induced decoherence model [20, 21]. To varying degrees, some of these models have been put forward, in part, to address the so-called measurement problem. In order to be compatible with empirical observations, it is traditionally stated that a measurement of a quantum system probabilistically collapses its state to an eigenstate of the measurement operator corresponding to the measured observable. This collapse is irreversible and not represented by unitary dynamics, and so stands in contradiction to the unitary evolution prescribed by the Schrödinger equation. Those seeking solutions to the measurement problem aim to reconcile this apparent and irreversible collapse with unitary dynamics.

The formulation of quantum measurement I present in this work is similar to Everett's relative state formalism inasmuch as a strong emphasis is placed on the interaction of a quantum system with a measuring device that is treated quantum mechanically, and their subsequent unitary evolution, as first formulated by John von Neumann [22]. However, I greatly expand upon these ideas and model quantum measurement in the framework of quantum information theory, with a focus on understanding the flow of information and entanglement throughout the quantum subsystems. A solution is not provided to the measurement problem as I argue that there is no problem and that accounting for a collapse mechanism is not necessary for a complete description of the measurement process. In other words, there is nothing inherently mysterious or inconsistent about the measurement process and the probabilistic nature of measurement outcomes is simply a consequence of the fact that we can only access part of a large entangled system. Lastly, I prefer to avoid excursions into interpretations and metaphysics, and instead to direct all efforts towards developing sound and concrete mathematical formalisms with relevant experimental applications and whose predictions can be realized and tested in the lab.

This view of quantum measurement is at the heart of the well-established weak measurement formalism [23] where a small coupling is induced between a quantum system and measurement device. Indeed, in the first experimental realization of weak measurements Ritchie, Story, and Hulet wrote that [24] "a measurement is performed by the interaction of a system with a device which provides some information about the population of the eigenstates of an observable of the system." In essence, quantum measurement is nothing more than a unitary interaction between a quantum system and a measuring apparatus that yields information about the state of the quantum system.

With this perspective in mind, the main focus of this work has been to frame quantum measurement in quantum information-theoretic terms and argue that quantum measurements should be treated on the same footing as other unitary quantum operations. In short, I aim to quantify how information is processed in quantum measurement. Considering diverse measurement scenarios such as weak and strong measurements, and parallel and consecutive measurements, I perform a comprehensive analysis of the role of entanglement and entropy in the measurement process and track the flow of information through all subsystems. Along the way, I will show how weak and strong measurements are of the same nature, derive a new complementarity relation for the Bell-state quantum eraser, and harness the coherence of the measurement process to develop quantum protocols for generating remote entanglement.

There are four areas of focus in this dissertation: (1) weak and strong quantum measurements, (2) parallel measurements made on entangled quantum systems, (3) consecutive measurements of a single quantum system, and (4) quantum disentangling protocols. Chapters 2–4 contain the necessary background for the subsequent results presented in Chs. 5–8. Chapter 2 will begin with a review of the physics of entangled quantum systems, while the main concepts of classical and quantum information theory will be presented in Ch. 3. Here, some of the useful tools for modeling the information content and entanglement of quantum systems will be outlined. The standard description of quantum measurement is reviewed in Ch. 4 along with examples of how the theoretical models for controlling and measuring qubits are realized in typical experimental settings.

With this foundation, I will introduce the first area of focus in Ch. 5. I will provide the model of quantum measurement that emphasizes the unitary entanglement between a quantum system and a quantum measuring device, often referred to as an ancilla, which is at the core of this work. I will show how such a model is fully compatible with the weak measurements formalism and that measurements of any interaction strength should be treated on the same footing. Finally, I will analyze in detail strong quantum measurements from the perspective of quantum information and elucidate the role entanglement and information play in the measurement process.

The second area of focus will be parallel quantum measurements in Ch. 6. Using the Bell-state quantum eraser, I will derive a new complementarity relation for a photon traveling through a double-slit apparatus. Specifically, I will compute the trade-off between the information about the photon's path obtained by measurement and the coherence of its density matrix. This demonstrates that an information-theoretic approach yields additional insights into the origins of complementarity.

An analysis of consecutive measurements of a single quantum system in Ch. 7 forms the third topic. I will describe how sequences of measurements that remain unamplified are coherent and that, due to the unique structure of the resulting ancilla density matrix, the joint entropy of the ancilla chain is determined only by the boundary. I will show that amplifying a measurement changes the amount of information gained about a quantum system and that amplified measurements are equivalent to a quantum Markov chain.

The final topic will be quantum protocols in Ch. 8. The coherence and entanglement properties of the measurement chains derived in Ch. 7 will be employed to develop quantum protocols based on disentangling operations. I will use such operations to outline a new protocol for generating remote entanglement, a key component of quantum communication protocols including teleportation and quantum key distribution.

Chapter 2

Preliminaries

Before delving into quantum information theory and quantum measurement, I will give a short review of the mathematical description of quantum systems. I will start with qubits, which are simple two-level quantum systems. I will discuss how the state of any qubit can be illustrated by a point on or within the Bloch sphere, with unitary transformations simply effecting a rotation of the qubit's state vector. The density matrix, which will be discussed next, describes both pure (zero entropy) and mixed quantum states (non-zero entropy) and is a useful tool for modeling multi-qudit (d-dimensional) systems. I will outline the types of correlations and entanglement that occur in these composite systems, which are key to the measurement process (see Chs. 4–7). In later chapters, I will show how these correlations can be more deeply understood using quantum information theory (see Ch. 3) and entropy Venn diagrams. I will then outline common unitary quantum operations on single- and two-qubit systems. Finally, I will discuss the teleportation protocol and superdense coding. These ideas form the foundation of many quantum communication protocols and will be useful for understanding the disentangling operations of Ch. 8.

2.1 Qubits

The state of a two-level system (a qubit) can be written generally as

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle, \tag{2.1}$$

with complex amplitudes that satisfy $|\alpha|^2 + |\beta|^2$. The states $|0\rangle$ and $|1\rangle$ form the computational basis for the qubit's vector space. There are many physical realizations of a qubit, for example, the spin of an electron or the polarization of a photon. I will describe later in Ch. 4.2 some typical experimental implementations of qubits and how they are manipulated and measured.

A qubit can be represented geometrically by considering Eq. (2.1) to be a point on the surface of a sphere [25], called the Bloch sphere. The north pole ($+\hat{z}$) of the Bloch sphere, shown in Fig. 2.1, represents the state $|0\rangle$, while the south pole ($-\hat{z}$) represents $|1\rangle$. Each pair of points on opposite sides of the sphere corresponds to orthogonal state vectors. For example, the $\pm\hat{x}$ axes are written $(|0\rangle \pm |1\rangle)/\sqrt{2}$, while the $\pm\hat{y}$ axes are $(|0\rangle \pm i|1\rangle)/\sqrt{2}$. As we will see in the next section, points on the surface of the sphere represent pure quantum states, while interior points are mixed states. In this picture, the general representation of a qubit's state is given by

$$|\psi\rangle = \cos(\theta/2)\,|0\rangle + e^{i\phi}\,\sin(\theta/2)\,|1\rangle. \tag{2.2}$$

This state is characterized by two real numbers θ and ϕ, where $0 \le \theta \le \pi$ is the polar angle and $0 \le \phi \le 2\pi$ is the azimuthal angle, and an unobservable overall phase has been dropped. For example, the symmetric state $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ is obtained with $\theta = \pi/2$ and $\phi = 0$, which is equivalent to the $\hat{x}$ axis of the Bloch sphere.

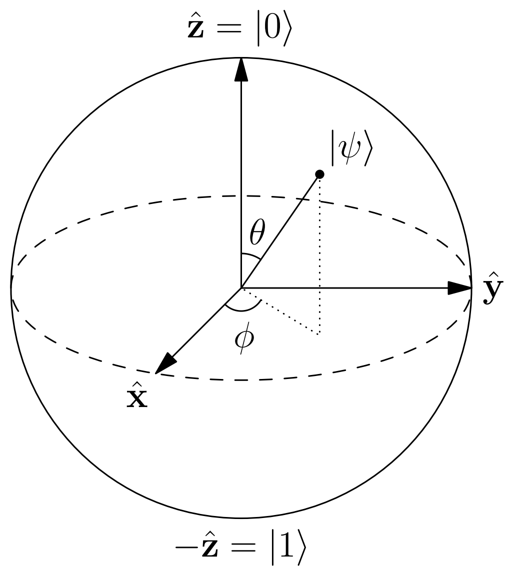

Figure 2.1: The Bloch sphere. The state $|\psi\rangle$ of a qubit can be represented as a unit-length vector in the sphere with coefficients given by the angles θ and ϕ. Image by Gosser.ca, distributed under a CC BY-SA 3.0 license.

Applying unitary operations to (2.2) transforms the state of a qubit. An arbitrary rotation by an angle α about the $\hat{n}$ axis is implemented with the operator

$$R_{\hat{n}}(\alpha) = e^{-i\alpha\,\hat{n}\cdot\vec{\sigma}/2} = \cos(\alpha/2)\,\mathbb{1} - i\,(\hat{n}\cdot\vec{\sigma})\sin(\alpha/2), \tag{2.3}$$

where $\vec{n} = n_x\,\hat{x} + n_y\,\hat{y} + n_z\,\hat{z}$ is a unit vector with normalization $\hat{n}^2 = n_x^2 + n_y^2 + n_z^2 = 1$, and $\vec{\sigma} = \sigma_x\,\hat{x} + \sigma_y\,\hat{y} + \sigma_z\,\hat{z}$ is a vector of Pauli matrices. The unit matrix is denoted $\mathbb{1}$. In the standard basis, $\{|0\rangle, |1\rangle\}$, these matrices are written

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.4}$$

and satisfy the following relations,

$$(\vec{\sigma}\cdot\hat{n})^2 = \mathbb{1}, \tag{2.5}$$

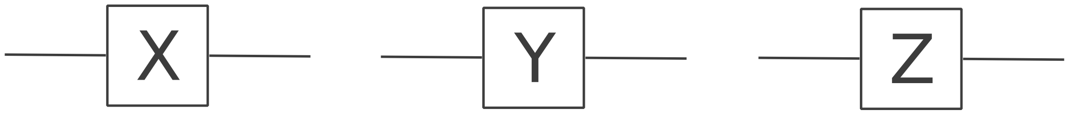

Figure 2.2: The three Pauli gates. The lines indicate the incoming and outgoing state of the qubit. The NOT gate X flips the state of the qubit, Z changes the phase, while Y performs both a bit and phase flip.

and

$$(\vec{\sigma} \cdot \hat{n}) \, (\vec{\sigma} \cdot \hat{m}) = (\hat{n} \cdot \hat{m}) \, \mathbb{1} + i \, (\hat{n} \times \hat{m}) \cdot \vec{\sigma}. \tag{2.6}$$

An arbitrary operation on a qubit is constructed from a rotation and a phase provided by γ:

$$U = e^{i\gamma} \, R_{\hat{n}}(\alpha). \tag{2.7}$$

Three basic transformations of a qubit are implemented with the Pauli gates, σ_x, σ_y, and σ_z (sometimes the notation X, Y, and Z is used) shown in Fig. 2.2. In the standard basis the bit-flip, or quantum NOT gate, is the Pauli matrix σ_x and reverses the state of a qubit. Up to an overall phase, σ_x is equivalent to a rotation by π around the $\hat{x}$ axis,

$$\sigma_x|0\rangle = |1\rangle \ , \quad \sigma_x|1\rangle = |0\rangle. \tag{2.8}$$

The phase-flip gate is the Pauli matrix σ_z and changes the phase using a rotation by π around the $\hat{z}$ axis,

$$\sigma_z|0\rangle = |0\rangle \ , \quad \sigma_z|1\rangle = -|1\rangle. \tag{2.9}$$

A bit and phase flip together is implemented with the Pauli matrix σ_y via a rotation by π

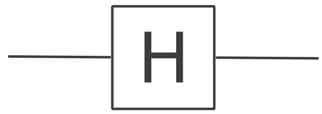

Figure 2.3: The Hadamard gate performs a rotation by an angle π around the axis $(\hat{x}+\hat{z})/\sqrt{2}$.

around the $\hat{y}$ axis,

$$\sigma_y|0\rangle = i|1\rangle \ , \quad \sigma_y|1\rangle = -i|0\rangle. \tag{2.10}$$

Another important single-qubit transformation is the Hadamard gate, H, shown in Fig. 2.3. It is equivalent, up to an overall phase, to a rotation about the axis $(\hat{x}+\hat{z})/\sqrt{2}$ by an angle $\alpha = \pi$,

$$H = \frac{\sigma_x + \sigma_z}{\sqrt{2}}. \tag{2.11}$$

For instance, this operator maps the state $|0\rangle$ into $(|0\rangle + |1\rangle)/\sqrt{2}$ and the state $|1\rangle$ into $(|0\rangle - |1\rangle)/\sqrt{2}$. In terms of the Bloch sphere, the Hadamard gate is also equivalent to two rotations: first a rotation by $\pi/2$ around the $\hat{y}$ axis followed by a rotation by π around the $\hat{x}$ axis.

For n qubits initialized in their ground states, the Hadamard gate produces a superposition of basis states with equal weight,

$$H \otimes \ldots \otimes H \, |0 \ldots 0\rangle = \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle, \tag{2.12}$$

where $|x\rangle$ is the set of n-qubit states, $|0\ldots 00\rangle$, $|0\ldots 01\rangle$, $|0\ldots 10\rangle$, $|0\ldots 11\rangle$, etc. The tensor product, denoted by $\otimes$, of two operators $\mathbf{A}$ (with dimension d_A) and $\mathbf{B}$ (with dimension d_B) acts in an extended vector space $\mathcal{A} \otimes \mathcal{B}$ of dimension $d_A \times d_B$. Therefore, each Hadamard gate on the left side of (2.12) acts on their respective qubit in the joint state $|0\dots0\rangle$. The expression (2.12) is known as the Hadamard transform and is a key element of many quantum algorithms, such as the period-finding subroutine in Shor's algorithm [26, 27] or Grover's search algorithm [28, 29].

2.2 The Density Matrix

The general state of a two-level system, which was written previously in Eq. (2.1), can be generalized to d-dimensional systems, called qudits,

$$|\psi\rangle = \sum_{i=0}^{d-1} \alpha_i \, |i\rangle. \tag{2.13}$$

The complex amplitudes satisfy $\sum_i |\alpha_i|^2 = 1$ and the states $|i\rangle$ form an orthonormal basis. For example, systems with $d = 3$ are called qutrits.

The quantum state (2.13) is referred to as a *pure* state since it has zero entropy. A pure state can be written down as a linear superposition of kets. In contrast, suppose a quantum system is one of many (not necessarily orthogonal) pure states, $|\psi_i\rangle$, with some probability p_i. Such a system is described with a density matrix and is written as a mixture,

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i|. \tag{2.14}$$

In this case, the quantum system is in a mixed state with non-zero entropy. Later, in Ch. 3.4 we will see how to compute the entropy of general mixed states.

The expression (2.14) for the density matrix is not unique. That is, there may exist different sets $\{p_i, \, |\psi_i\rangle\}$ and $\{p_i', \, |\psi_i'\rangle\}$ that give rise to the same density matrix. Furthermore, the number of terms in the sum (2.14) is not necessarily the same for these different decompositions. For example, for a qubit consider the orthogonal states $|\psi_0\rangle = |0\rangle$ and $|\psi_1\rangle = |1\rangle$ with weights p_0 and p_1. One could equivalently write the density matrix with the non-orthogonal states $|\psi_0\rangle = \sqrt{p_0}\,|0\rangle + \sqrt{p_1}\,|1\rangle$ and $|\psi_1\rangle = \sqrt{p_0}\,|0\rangle - \sqrt{p_1}\,|1\rangle$ with weights 1/2, 1/2.

The density matrix, ρ, is a Hermitian operator with unit trace (meaning its diagonal elements sum to one) and eigenvalues that are both real and non-negative (ρ is a "positive operator"). The trace of a matrix is defined as

$$\begin{aligned} \mathrm{Tr}(\rho) &= \sum_n \langle n|\rho|n\rangle \\ &= \sum_{in} p_i \,\langle\psi_i|n\rangle\langle n|\psi_i\rangle = \sum_i p_i = 1, \end{aligned} \tag{2.15}$$

where $|n\rangle$ is a complete set of orthonormal states of ρ.

In a chosen basis, the diagonal elements of the density matrix correspond to the probability to observe the quantum system in a given state. For example, expressing (2.14) in the computational basis, $|j\rangle$,

$$\rho = \sum_{ijj'} p_i \; c_{ij} \, c^*_{ij'} \, |j\rangle\langle j'|, \tag{2.16}$$

shows that the state $|j\rangle$ would be observed with probability $\sum_i p_i \, |c_{ij}|^2$. The off-diagonal elements of ρ, which are equal to $\sum_i p_i \, c_{ij} \, c^*_{ij'}$, characterize how coherent, or how pure, the state is.

The purity of a quantum state is calculated from the trace of the square of the density matrix, $\mathrm{Tr}(\rho^2)$, and is equal to one when ρ is a pure state, and is otherwise less than one. For example, a completely mixed state of dimension d has the density matrix $\rho = \frac{1}{d}\,\mathbb{1}$ with purity equal to $1/d$. If $\rho = |\psi\rangle\langle\psi|$, then $\rho^2 = \rho$ and the purity is equal to one. But, if ρ is mixed the purity is

$$\mathrm{Tr}(\rho^2) = \sum_{ii'} p_i\, p_{i'}\, |\langle\psi_i|\psi_{i'}\rangle|^2, \tag{2.17}$$

which reduces to $\mathrm{Tr}(\rho^2) = \sum_i p_i^2$ if $|\psi_i\rangle$ are orthonormal.

The state of a general (meaning, not necessarily pure) qubit can be expanded in terms of the identity and Pauli matrices as

$$\rho = \frac{1}{2}\big(\mathbb{1} + \vec{a}\cdot\hat{\sigma}\big) = \frac{1}{2}\big(\mathbb{1} + a_x\,\sigma_x + a_y\,\sigma_y + a_z\,\sigma_z\big), \tag{2.18}$$

since these matrices form a basis for all two-dimensional Hermitian matrices. Here, $\vec{a}$ is known as the Bloch vector and indicates the point within the sphere corresponding to the mixed state ρ [25]. The eigenvalues of (2.18) are $(1 \pm \|\vec{a}\|)/2$, where the norm of the Bloch vector satisfies $\|\vec{a}\| \leq 1$. For pure states, $\|\vec{a}\| = 1$ so that $\mathrm{Tr}(\rho^2) = 1$, implying that pure states are restricted to the surface of the Bloch sphere, while mixed states lie in the interior.

The density matrix formulation is a powerful tool in quantum information theory. It is used for describing general quantum systems and operations, and, in particular, for understanding subsystems that are part of a larger entangled pure state. In later chapters, the density matrix will be used, among other things, to characterize the coherence and entanglement in quantum measurement.

2.3 Composite Systems and Entanglement

Our discussion of single-qubit (or qudit) states in Sec. 2.1 can be extended to describe composite, or multi-qudit states. For instance, we can write a general pure state of two qudits, A and B, as

$$|AB\rangle = \sum_{i,j=0}^{d-1} c_{ij}\,|i\rangle \otimes |j\rangle, \tag{2.19}$$

where $|c_{ij}|^2 = 1$ and $\otimes$ indicates the tensor product (which we will often drop for notational convenience). If A and B are each d-dimensional systems, then their joint state, $|AB\rangle$, is $d \times d$–dimensional. The simplest example of a composite system occurs when A and B are uncorrelated. In this case, the coefficients in the wave function factor, $c_{ij} = a_i\, b_j$, yielding a product of two pure states,

$$|AB\rangle = \sum_i a_i\,|i\rangle \otimes \sum_j b_j\,|j\rangle = |\psi_A\rangle \otimes |\psi_B\rangle. \tag{2.20}$$

For general, non-factorizable coefficients c_{ij} the two systems are said to be *entangled*. By performing a partial trace over just the states of B in (2.19), we find that A is a mixed state,

$$\rho(A) = \mathrm{Tr}_B\big(|AB\rangle\langle AB|\big) = \sum_{ii'j} c_{ij}\, c^*_{i'j}\,|i\rangle\langle i'|. \tag{2.21}$$

Although the joint state $|AB\rangle$ is pure and fully known, the entanglement of the state yields subsystems that are mixed and uncertain.

A special case of the state (2.19) is known as a Bell state (or EPR pair after the famous 1935 paper by Einstein, Podolsky, and Rosen [30]),

$$|\Phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \tag{2.22}$$

and plays an essential role in quantum computing and many quantum protocols including quantum teleportation and superdense coding. The state (2.22) is entangled as it cannot be written as a tensor product of two pure states. Furthermore, the entanglement is maximum since each qubit is in a maximally mixed state,

$$\rho(A) = \mathrm{Tr}_B\big(|\Phi^+\rangle\langle\Phi^+|\big) = \frac{1}{2}\,\mathbb{1} = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}. \tag{2.23}$$

The Bell state (2.22) can be trivially extended to n qubits where it is referred to as a GHZ state, named for Greenberger, Horne and Zeilinger [31]. An important instance is $n = 3$,

$$|\mathrm{GHZ}\rangle = \frac{|000\rangle + |111\rangle}{\sqrt{2}}, \tag{2.24}$$

and is used in many quantum communication protocols. For example, in quantum cryptography the original proposal for quantum secret sharing [32] used a GHZ state to distribute quantum information across multiple parts such that two of them are required to reconstruct the original state. The GHZ state is also important in that it represents one of two distinct classes of tripartite entanglement [33], the other being the W state,

$$|W\rangle = \frac{|001\rangle + |010\rangle + |100\rangle}{\sqrt{3}}. \tag{2.25}$$

The states (2.24) and (2.25) are inequivalent since they cannot be transformed into each other via local operations.

Along with three other similar maximally entangled states, the state (2.22) forms a complete basis for any two-qubit state that is known as the Bell basis:

$$\begin{aligned} |\Phi^\pm\rangle &= \frac{|00\rangle \pm |11\rangle}{\sqrt{2}}, \\ |\Psi^\pm\rangle &= \frac{|01\rangle \pm |10\rangle}{\sqrt{2}}. \end{aligned} \tag{2.26}$$

It is easy to show that the reduced density matrix of all four states is $\frac{1}{2}\,\mathbb{1}$, revealing that they are indeed maximally entangled. Another notation that we will sometimes use for the Bell states is $|\beta_{00}\rangle = |\Phi^+\rangle$, $|\beta_{10}\rangle = |\Phi^-\rangle$, $|\beta_{01}\rangle = |\Psi^+\rangle$, and $|\beta_{11}\rangle = |\Psi^+\rangle$.

The Bell states (2.26) can be transformed into one another by operations on just one of the qubits (that is, via local operations). For example, application of the phase-flip operator to the second qubit transforms $\mathbb{1} \otimes \sigma_z \ |\beta_{00}\rangle = |\beta_{10}\rangle$. In general, the Bell states can be written as

$$|\beta_{zx}\rangle = \left(\mathbb{1} \otimes X^x Z^z\right) |\beta_{00}\rangle, \tag{2.27}$$

where X^x and Z^z are powers of the Pauli matrices. In other words, local operations cannot break the entanglement of the pair. For instance, writing each single-qubit state in the diagonal basis $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ yields equivalently entangled states. Later, when discussing the von Neumann entropy and the Schmidt decomposition in Ch. 3, it will be clear that the Schmidt coefficients, and therefore the entanglement, are basis independent in a bipartite pure state.

Experimentally, the set of two-qubit states (2.26) can be generated in different ways. One way this can be achieved is with spontaneous parametric down-conversion [34]. In this method, a laser is focused on a beta-barium borate (BBO) crystal, a nonlinear optical material, such that one incident photon is converted into two photons that are emitted with correlated polarizations. Alternatively, Eq. (2.26) can be produced using a sequence of Hadamard and controlled-NOT gates, discussed in more detail below, on two initially uncorrelated qubits.

Perhaps the most common two-qubit unitary transformation is the controlled-NOT gate, U_{CNOT}. This simple gate, shown in Fig. 2.4, can be used to entangle (or disentangle) qubits

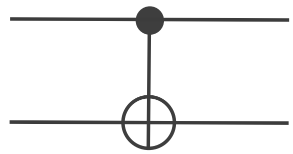

Figure 2.4: The controlled-NOT gate flips the state of the target qubit (open circle) only if the control qubit (filled circle) is in the state $|1\rangle$.

by performing a conditional bit-flip. With the first qubit as the control, this gate will flip the state of the second qubit, the target, if and only if the first qubit is in the state $|1\rangle$, and otherwise will do nothing. The operator can be written in the following way,

$$U_{\text{CNOT}} = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes \sigma_x, \tag{2.28}$$

so that its action on states in the computational basis is

$$\begin{aligned} U_{\text{CNOT}}: \quad & |00\rangle \mapsto |00\rangle \\ & |01\rangle \mapsto |01\rangle \\ & |10\rangle \mapsto |11\rangle \\ & |11\rangle \mapsto |10\rangle. \end{aligned} \tag{2.29}$$

The set of single-qubit operations and the controlled-not gate is known as a universal set. Any quantum gate can be decomposed into these gates so that, together, they are sufficient to perform any quantum computation [35].

Using one Hadamard and one controlled-NOT gate, we can easily construct an EPR pair, the Bell state (2.22). Suppose two qubits are each initialized in the state $|0\rangle$. Applying the sequence of gates $U_{\text{CNOT}}\,(H \otimes \mathbb{1})$, where the control is on the first system, yields

$$U_{\text{CNOT}}\,(H \otimes \mathbb{1})\;|0\rangle \otimes |0\rangle = U_{\text{CNOT}}\,\frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |0\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}\,. \tag{2.30}$$

The controlled-NOT gate can be generalized to a controlled-U gate,

$$C(U) = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes U, \tag{2.31}$$

where the bit-flip operation, σ_x, is replaced by an arbitrary unitary operation, U. It is easy to check that the controlled-U gate, as with any other quantum gate, is unitary. This general operation will be useful when discussing quantum disentangling operations in Ch. 8.

In addition to the pure state in Eq. (2.19), we can also consider mixed states for two qudits. Two systems, A and B, are *separable* if their joint state can be written as combination of product states,

$$\rho(AB) = \sum_i p_i\, \rho_i^A \otimes \rho_j^B, \tag{2.32}$$

where p_i are non-negative probabilities and $\sum_i p_i = 1$. The reduced density matrices for each of the subsystems, ρ_i^A and ρ_i^B, could be pure or mixed. Such states exhibit classical correlations between the subsystems. Of course, if there is only one non-zero p_i then the separable state (2.32) reduces to an *uncorrelated* state, $\rho(AB) = \rho(A) \otimes \rho(B)$. If the joint state $\rho(AB)$ cannot be written in the separable form (2.32), then it is said to be entangled.

2.4 The No-Cloning Theorem

A striking feature of quantum systems is that it is *impossible* to perfectly copy, or clone, an arbitrary quantum state. This was proved in 1982 by Wootters and Zurek [36] as well as by Dieks who showed that this property precludes the possibility of superluminal communication via EPR pairs [37].

To demonstrate the essence of the no-cloning theorem, suppose that there exists a unitary transformation, U, that can indeed clone an arbitrary state. Then, for two such orthogonal states $|\psi\rangle$ and $|\phi\rangle$, the transformations $U(|\psi\rangle|0\rangle) = |\psi\rangle|\psi\rangle$ and $U(|\phi\rangle|0\rangle) = |\phi\rangle|\phi\rangle$ are possible. Since U must be a linear operator, its application to an *arbitrary* superposition, $|\eta\rangle = \alpha\,|\psi\rangle + \beta\,|\phi\rangle$, results in

$$U\big(|\eta\rangle|0\rangle\big) = \alpha\,|\psi\rangle|\psi\rangle + \beta\,|\phi\rangle|\phi\rangle, \tag{2.33}$$

where $|\alpha|^2 + |\beta|^2 = 1$. However, if U is a cloning operation, then it must be true that

$$U\big(|\eta\rangle|0\rangle\big) = |\eta\rangle|\eta\rangle = \alpha^2\,|\psi\rangle|\psi\rangle + \beta^2\,|\phi\rangle|\phi\rangle + \alpha\beta\,|\psi\rangle|\phi\rangle + \alpha\beta\,|\phi\rangle|\psi\rangle. \tag{2.34}$$

Clearly, Eqs. (2.33) and (2.34) are not the same. The presence of the cross terms in (2.34) implies that this cloning operation is impossible.

The no-cloning theorem is usually discussed for pure quantum states but was generalized in 1996 by Barnum et al. to include mixed states via the no-broadcast theorem [38]. Although it is impossible to perfectly clone an arbitrary quantum state, *approximate* cloning is permitted. Indeed, the highest achievable fidelity for qubit cloning has been calculated to be 5/6 [39, 40].

The no-cloning theorem has many interesting consequences in addition to the impossibility of superluminal communication mentioned above. For instance, classical error correction, which relies on the redundant encoding of bits in a repetition code, cannot be employed in quantum algorithms (see, e.g., [41]). The theorem is also responsible for the distinct nature of quantum measurements as compared to classical measurements. As we will see in later chapters, the inability to perfectly clone arbitrary quantum states suggests that our measurement devices do not and cannot reflect the true underlying state of the quantum system.

2.5 Quantum Teleportation and Superdense Coding

Two fundamental and complementary protocols in quantum information processing are quantum teleportation and superdense coding. These protocols rely heavily on the use of the maximally entangled Bell states. In this section, I will outline both protocols as the main ideas will be helpful for understanding the quantum disentangling operations in Ch. 8.

Quantum teleportation, discovered in 1993 by Bennett et al. [7], is a method for transferring a qubit to a receiver using two bits of classical information. Following the original proposal a few years later, teleportation was implemented using the polarization states of photons in 1997 and 1998 [42–44]. Also in 1998, the first "complete" teleportation (where the final corrections are applied to Bob's qubit) was achieved across inter-atomic distances in the nuclear spins of trichloroethylene using nuclear magnetic resonance (NMR) [45]. In 2004, teleportation across micrometer distances was realized using trapped ions [46, 47]. More recently, teleportation of the state of superconducting transmon qubits over 6 millimeters was performed in 2013 [48] and of the spin states of nitrogen vacancy centers in diamond over 3 meters in 2014 [49]. As of July 4, 2017 the world record for long-distance quantum

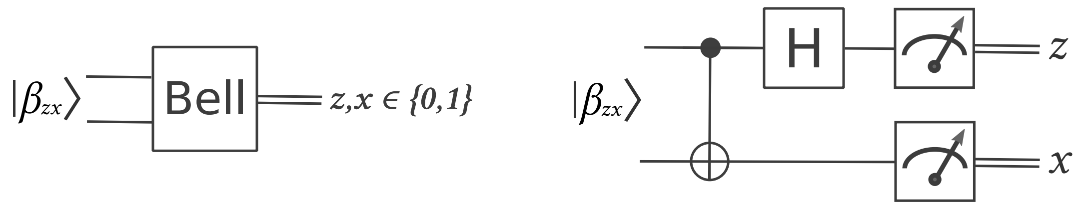

Figure 2.5: Shorthand notation for the Bell measurement circuit is shown on the left together with the complete circuit on the right. A Bell state $|\beta_{zx}\rangle$ is converted into the two classical bits z and x (drawn with double solid lines) after a controlled-NOT, Hadamard, and two single-qubit measurements in the standard basis.

teleportation with photons is nearly 1400 kilometers, the distance between an observatory in Tibet and a satellite in low Earth orbit [50] (see also [51] for the team's demonstration of first distributing the entanglement over such length scales).

In the quantum teleportation protocol, two parties (usually referred to as Alice and Bob) initially share a maximally entangled state of two qubits A and B, say, $|\Phi^+\rangle_{AB} = |\beta_{00}\rangle_{AB}$. The sender, Alice, wishes to transfer an arbitrary quantum state $|\psi\rangle_Q = \alpha|0\rangle + \beta|1\rangle$ to a receiver, Bob. To do so, Alice first performs a Bell measurement (see Fig. 2.5 for a circuit diagram of a Bell measurement) on her half of the entangled pair (qubit A) and the unknown qubit Q in the state $|\psi\rangle_Q$. She sends the result of the measurement, two classical bits, to Bob, who then performs the necessary corrections to his half of the Bell state (qubit B). Afterwards, his qubit B will be in the state $|\psi\rangle$. Interestingly, as we will see, neither Alice nor Bob need to know the actual state $|\psi\rangle$ that they are sending.

Before going through each step of the protocol in detail, we first make an observation. Quantum teleportation can be summarized by the following simple relation for the initial state of all three qubits Q, A, and B before the protocol is implemented,

$$|\psi\rangle_Q \otimes |\beta_{00}\rangle_{AB} = \frac{1}{2} \sum_{z,x\in\{0,1\}} |\beta_{zx}\rangle_{QA} \otimes X^x Z^z \,|\psi\rangle_B. \tag{2.35}$$

Here, $|\beta_{zx}\rangle$ are the four Bell states, and X and Z are Pauli operators (see Sec. 2.3). From this expression, we can see that the state $|\psi\rangle$ of Q will be transferred to B, up to a unitary correction. To be clear, only the *state* of Q is transferred to Bob's qubit, not the *physical* particle Q. Finally, at the end of the protocol, the state (2.35) evolves to

$$\frac{1}{4} \sum_{z,x\in\{0,1\}} |\beta_{zx}\rangle_{QA}\langle\beta_{zx}| \otimes |\psi\rangle_B\langle\psi| = \frac{1}{2}\mathbb{1}_Q \otimes \frac{1}{2}\mathbb{1}_A \otimes |\psi\rangle_B\langle\psi|, \tag{2.36}$$

where qubits Q and A are completely mixed and the state of the Bob's qubit B is $|\psi\rangle$, which is disentangled from the rest of the system.

We can understand the protocol better by considering each step in more detail. The initial state (2.35) of the three qubits is the product state

$$|\phi_0\rangle = |\psi\rangle_Q \otimes |\beta_{00}\rangle_{AB} = \Big(\alpha|0\rangle_Q + \beta|1\rangle_Q\Big) \otimes \frac{1}{\sqrt{2}}\Big(|00\rangle_{AB} + |11\rangle_{AB}\Big), \tag{2.37}$$

where Q labels the arbitrary state to be teleported from Alice to Bob, and A and B label Alice's and Bob's qubits, respectively. The labels will be dropped from now on since the ordering of states is always maintained.

In the first step, Alice performs a Bell measurement by applying a CNOT gate to Q and A, with the control on Q, followed by a Hadamard gate on Q. This transforms Eq. (2.37) to

$$|\psi_1\rangle = \frac{1}{2}\Big[|00\rangle\Big(\alpha|0\rangle + \beta|1\rangle\Big) + |01\rangle\Big(\alpha|1\rangle + \beta|0\rangle\Big) + |10\rangle\Big(\alpha|0\rangle - \beta|1\rangle\Big) + |11\rangle\Big(\alpha|1\rangle - \beta|0\rangle\Big)\Big]. \tag{2.38}$$

Expression (2.38) already suggests that the conditional state of B is nearly equivalent to the desired state $|\psi\rangle$.

In the second step, Alice measures the qubits Q and A in the computational basis. For example, if she observes the outcome "00" then, according to (2.38), the corresponding state

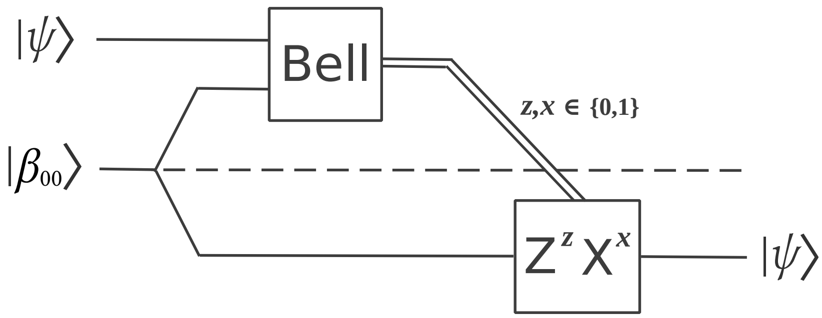

Figure 2.6: The quantum teleportation circuit. The dashed line separates the systems under Alice's (upper) and Bob's (lower) control. Alice transfers the state $|\psi\rangle$ to Bob by performing a Bell measurement and sending the two resulting classical bits of information z and x (denoted by the double solid lines) to Bob. Conditionally on these two bits, Bob performs a correction $Z^z X^x$ to his qubit to obtain the state $|\psi\rangle$.

of B is evidently $\alpha|0\rangle_B + \beta|1\rangle_B$, which is the original state $|\psi\rangle$.

For each of the four possible outcomes in Alice's measurement, Bob's qubit will be in one of four states. Each of these states is related to $|\psi\rangle$ by a simple single-qubit operation. That is, Eq. (2.38) can be rewritten as

$$|\psi_1\rangle = \frac{1}{2}\Big[|00\rangle\,|\psi\rangle + |01\rangle\,\sigma_x|\psi\rangle + |10\rangle\,\sigma_z|\psi\rangle + |11\rangle\,(-i\sigma_y)|\psi\rangle\Big]. \tag{2.39}$$

In this form, it is clear which operation Bob must perform on his qubit to recover $|\psi\rangle$. For example, suppose Alice found the outcome "01". In this case, upon receiving Alice's two-bit message Bob corrects his qubit $\sigma_x|\psi\rangle$ with the bit-flip gate σ_x, to obtain $|\psi\rangle$. Similarly, for the other measurement outcome "10", Bob applies the phase-flip gate σ_z, while for "11" he applies $i\sigma_y$. Thus, with two bits of classical communication, a sender can successfully transfer an arbitrary qubit to a distant receiver. See Fig. 2.6 for the complete teleportation circuit.

A modified version of quantum teleportation has also been proposed by Brassard, Braunstein and Cleve [52, 53] and is implemented without performing a Bell measurement. In this case, a sequence of controlled-NOT and Hadamard gates replaces the Bell measurement, which is no longer performed on the first two qubits, Q and A. Nevertheless, the state $|\psi\rangle$ is still transferred to Bob.

The second fundamental quantum protocol is superdense coding, which was published in 1992 by Bennett and Wiesner [6] one year before the teleportation protocol. In superdense coding, two bits of classical information can be communicated using a single qubit. It provides a way to transmit classical information over a quantum channel, and is, in a sense, the inverse of the teleportation scheme. Superdense coding was first experimentally realized in 1996 in an optical setting by Mattle et al. [54]. In February 2017, the highest bit rate to date of 1.67 bits per qubit (the maximum, of course, being 2) was achieved using linear optics over a fiber optic link [55].

In superdense coding, Alice and Bob each have one half of an entangled pair of qubits, for instance, the Bell state $|\Phi^+\rangle = |\beta_{00}\rangle$ in (2.22). Depending on the value of the two bits that Alice wishes to communicate to Bob, she encodes her message by performing one of four operations on her qubit, A. She then sends A to Bob, who decodes the message. Superdense coding is based on a simple identity for the Bell state before the protocol is implemented,

$$(H \otimes \mathbb{1})\, U_{\text{CNOT}}\, (Z^z X^x \otimes \mathbb{1})\, |\beta_{00}\rangle = |z\rangle |x\rangle. \tag{2.40}$$

That is, a set of unitary operations on the initially entangled qubits yields a product state.

To see how this works, suppose that Alice wants to send the bit string "00". In this case, she performs the unitary operation, $\mathbb{1}$, and sends qubit A to Bob. The decoding scheme that Bob follows is always the same regardless of which message Alice sends, and is the inverse of the entangling scheme described in (2.30) for creating Bell states. Namely, Bob applies the

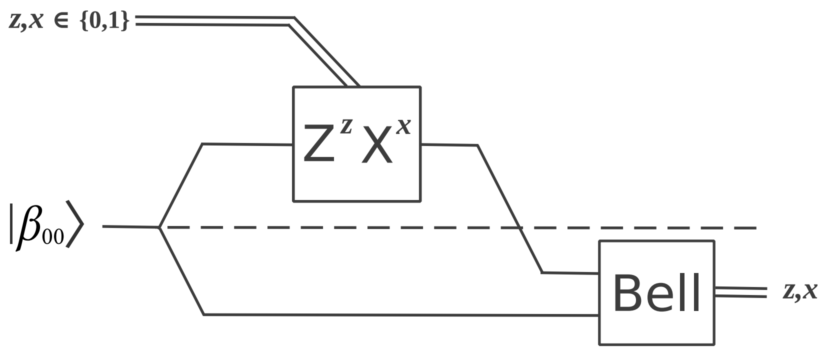

Figure 2.7: The superdense coding circuit. The dashed line separates the systems under Alice's (upper) and Bob's (lower) control. The two bits z, x (indicated by the double solid line) are transferred to Bob when Alice sends her encoded qubit.

controlled-NOT operation (2.28), where the control is on qubit A, followed by the Hadamard gate (2.11) on A. For this example, decoding yields the state

$$(H \otimes \mathbb{1})\, U_{\text{CNOT}} \left[(\mathbb{1} \otimes \mathbb{1})\ |\beta_{00}\rangle \right] = |00\rangle. \tag{2.41}$$

If Bob measures his qubits in the computational basis, he will find with certainty the outcomes "00", which is precisely the message that Alice intended to send. Or, suppose that Alice wants to send the message "10". In this case, she encodes the message by applying $X = \sigma_x$ to her qubit. The four possible messages that Alice can send to Bob are summarized in Table 2.1 and the circuit is shown in Fig. 2.7.

The set of unitary operations that Alice performs on her qubit A can be achieved using two additional qubits, a and a', which are initially in one of the four basis states $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$, corresponding to the message that Alice wishes to send. This way, Alice does not have to directly implement the controlled operations on her qubit A and instead can send in two qubits prepared in the state $|ij\rangle$, corresponding to the message ij, to Bob. In this case, the circuit diagram in Fig. 2.7 is modified to include qubits a and a' and a controlled-NOT gate between A and a followed by a controlled-Z gate between A and a'.

Table 2.1: Encoding and decoding schemes in the superdense coding protocol.

Alice's encoding ($Z^z X^x$)	State after encoding ($\|\beta_{zx}\rangle$)	State after decoding (z, x)
$\mathbb{1}$	$\|\Phi^+\rangle = \|\beta_{00}\rangle$	$\|00\rangle$
X	$\|\Psi^+\rangle = \|\beta_{01}\rangle$	$\|01\rangle$
Z	$\|\Phi^-\rangle = \|\beta_{10}\rangle$	$\|10\rangle$
$ZX = iY$	$\|\Psi^-\rangle = \|\beta_{11}\rangle$	$\|11\rangle$

If an eavesdropper, Eve, intercepts qubit A when it is in transit to Bob, she cannot ascertain the message. This is because the marginal state of A is mixed,

$$\rho(A) = \mathrm{Tr}_B\Big(Z^z X^x \otimes \mathbb{1}\ |\beta_{00}\rangle\langle\beta_{00}| X^x Z^z \otimes \mathbb{1}\Big) = \mathrm{Tr}_B\Big(|\beta_{zx}\rangle\langle\beta_{zx}|\Big) = \frac{1}{2}\,\mathbb{1}, \tag{2.42}$$

and doesn't reveal anything about the value of z and x. It is the entanglement between the qubits that allows two bits of information to be sent instead of just one.

2.6 Summary

In this chapter, I outlined the mathematical description of quantum systems, with particular focus on the two-level system known as the qubit. I reviewed the density matrix formalism and the properties of general quantum states. I discussed pure and mixed states, as well as the types of correlations, classical and quantum, between multiple quantum systems. I then showed how the no-cloning theorem places severe restrictions on the ability to copy quantum states and, as we will see in Ch. 5, is responsible for the non-classical nature of quantum measurements. Finally, I reviewed the quantum teleportation and superdense coding protocols, which will be useful for the protocols presented in Ch. 8.

Chapter 3

Quantum Information Theory

3.1 Introduction

The departure of quantum states from their classical counterparts can be more deeply understood using quantum information theory (QIT). The tools provided by QIT allow us to characterize the structure of quantum states and, in particular, the amount of information that can be obtained about a quantum system, or how entanglement and entropy are distributed in composite systems. QIT can also be used to place constraints on quantum systems. For example, Holevo's theorem places an upper limit to the amount of information that can be extracted about the state of a quantum system via measurement, while the entropic relationship between several quantum systems is bounded by the strong subadditivity of quantum entropy. In this chapter, I first review the basic ideas of information theory according to Shannon in order to set the stage for quantum information theory. Then I will summarize several useful quantum information-theoretic concepts including the Schmidt decomposition and Holevo's theorem.

3.2 Classical Information Theory

The foundations of information theory were originally laid by Claude Shannon in 1948 when he published the set of transformational papers [2,3]. His theory provided a way to quantify information and to understand the communication of information over a (noisy) channel. A measure of information in a random variable is related to the variable's entropy, which quantifies the uncertainty of the value of the variable. Equivalently, the entropy reveals how much information one gains upon learning the value of the variable. The entropy that is shared between two random variables (the mutual entropy) is a measure of the amount of information that one variable has about the other. Shannon's noisy channel coding theorem established a maximum rate for the information that can be reliably transmitted through a channel with a particular noise level in terms of a channel capacity, which is a function of this mutual entropy.

In this section, I will review the main ideas in classical information theory, and provide the mathematical definitions of entropy and information. By constructing joint and mutual entropies, it will become clear how information is distributed among many correlated subsystems. This will facilitate the transition to the next section on quantum information theory, which is able to describe the properties of complex quantum entangled systems.

3.2.1 Shannon Entropy

The entropy of a random variable is a function of the set of probabilities corresponding to the values that the variable can take. The Shannon entropy [2, 56] of a system A is defined as

$$H(A) = -\sum_{i=0}^{d-1} p_i \log p_i, \tag{3.1}$$

where the probability of the ith state is p_i and the dimension of the system A is d. When the logarithm is taken to base 2, the entropy is in units of bits. States that occur with zero probability ($p_i = 0$ for some i) do not contribute to the entropy according to the limit $\lim_{x\to 0} x \log x = 0$. Equation (3.1) quantifies what is required to store information about the variable A so that one is able to reconstruct it later [25]. For example, if A takes on one of four states with equal probability, then two bits of information are required to specify its state.

The marginal entropy in Eq. (3.1) is extended to a joint entropy for two (or more) systems A and B via their joint probability distribution p_{ij},

$$H(AB) = -\sum_{ij} p_{ij} \log p_{ij}, \tag{3.2}$$

which is a measure of the uncertainty in both systems. The marginal and joint probability distributions give rise to a conditional probability, $p_{i|j} = p_{ij}/p_j$, which is used to construct the conditional entropy

$$H(A|B) = -\sum_{ij} p_{ij} \log p_{i|j}. \tag{3.3}$$

This characterizes the average reduction in entropy of A, given the state of system B.

On the other hand, the mutual probability distribution $p_{i:j} = p_i\, p_j / p_{ij}$ yields the mutual entropy

$$H(A:B) = -\sum_{ij} p_{ij} \log p_{i:j}, \tag{3.4}$$

or the information shared between systems A and B. In terms of marginal and joint entropies, the conditional and mutual entropies are defined as [56]

$$H(A|B) = H(AB) - H(B), \tag{3.5}$$

$$H(A:B) = H(A) + H(B) - H(AB), \tag{3.6}$$

or,

$$H(A) = H(A|B) + H(A:B). \tag{3.7}$$

In the next section, we will see how Shannon entropies are bounded according to the possible correlations that exist between subsystems.

3.2.2 Inequalities for Shannon Entropy

The Shannon entropy of two systems obeys two inequalities, the first of which is known as subadditivity. Suppose that two systems, A and B, are uncorrelated (independent). In this case, their joint probability distribution factors into a product of two marginal probabilities, $p_{ij} = p_i\, p_j$, and the joint entropy reduces to a sum of marginal entropies, $H(AB) = H(A) + H(B)$. In general, when A and B are correlated, the joint entropy is reduced by the mutual entropy as in (3.6). Specifically,

$$H(AB) \leq H(A) + H(B). \tag{3.8}$$

In other words, $H(AB)$ is subadditive and some of the information in the joint system AB can be found in the correlations between subsystems A and B. From (3.8), one finds that the mutual Shannon entropy (3.6) is non-negative, $H(A:B) \geq 0$. The proof of statement (3.8) can be shown with the inequality, $\log_2 x \leq (x-1)/\ln 2$, for all positive x [25]. Note that this comes from using $\ln x = (\log_2 x)\,(\ln 2)$ and the fact that $\ln x \leq x - 1$. Using this in expression (3.4), one arrives at (3.8).

The second inequality for the entropy of two systems establishes a lower bound to the joint Shannon entropy. That is, the entropy of a composite system cannot be less than the entropy of any of its parts,

$$H(AB) \geq H(A),\, H(B). \tag{3.9}$$

It follows that the conditional entropy (3.5) is always non-negative, $H(A|B) \geq 0$. This can also be shown by considering the expression (3.3) for the conditional entropy in terms of its probability distributions. As the conditional probability is bounded by $0 \leq p_{i|j} \leq 1$, the expression $-\log p_{i|j}$ must be non-negative. Therefore (3.3) is also non-negative.

3.3 Quantum Information Theory

In Shannon's theory of information, the basic unit of information is the bit, which can take one of two values. In quantum information theory, information can be contained in the states of quantum systems, and so the basic unit becomes the qubit. As we saw in Ch. 2.1, the state of a qubit can correspond to any point on the surface of the Bloch sphere and so can be in a linear superposition of the classical bit values. In Ch. 2.3, we studied the types of correlations that arise in multi-qudit systems and know from Ch. 2.5 that some of the most important applications of quantum information science rely on entanglement as a resource. Entanglement is a property that is unique to quantum systems and cannot be accounted for in the classical Shannon theory. More generally, since quantum systems can exist in superpositions, it is necessary to use density matrices in place of probability distributions for a useful definition of quantum entropy, as the density matrix encodes the phase information of the quantum system. Classical entropies are functions only of probability distributions, which correspond to the diagonal elements of a density matrix. Therefore, it is important to have a theory of quantum information that can correctly describe the behavior of quantum systems. In the next sections, I review the definition of entropy as proposed by von Neumann, and show how superposition and entanglement manifest themselves in information theory.

3.3.1 von Neumann Entropy

The entropy of a quantum state, first proposed by von Neumann in 1927 [22,57] (remarkably, twenty-one years before Shannon's classical entropy), is defined as a trace over the density matrix ρ of a system A,

$$S(A) = -\mathrm{Tr}\big[\rho(A)\log\rho(A)\big]. \tag{3.10}$$

Note that we will use the symbol S for the von Neumann entropy and H for the Shannon entropy. Diagonalizing the density matrix, this expression can be written as a summation of the eigenvalues, λ_i, of $\rho(A)$,

$$S(A) = -\sum_i \lambda_i \log \lambda_i. \tag{3.11}$$

If the state of system A corresponds to a completely mixed state, its density matrix is proportional to the unit matrix, $\rho(A) = \frac{1}{2}\mathbb{1}$, with eigenvalues $1/2, 1/2$. The entropy of A is then $S(A) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1$ bit (since the logarithm is base 2) and coincides with the Shannon entropy. On the other hand, if A is a pure state, then its entropy vanishes, $S(A) = 0$, since its density matrix has a single nonzero eigenvalue equal to 1.

A similar expression can be written for the joint entropy of a density matrix $\rho(AB)$ of two systems A and B. The conditional probability (3.3) from Shannon's theory is extended to quantum systems with a conditional "amplitude matrix" $\rho(A|B)$ so that the conditional von Neumann entropy is

$$S(A|B) = -\mathrm{Tr}\big[\rho(AB)\log\rho(A|B)\big]. \tag{3.12}$$

The conditional amplitude matrix is not a density matrix since $\mathrm{Tr}[\rho(A|B)] \neq 1$, similarly to $\sum_{ij} p_{i|j} \neq 1$ for classical probabilities. Nonetheless, it is a positive semi-definite Hermitian matrix with non-negative and real eigenvalues. Cerf and Adami defined in [58] the conditional amplitude matrix as

$$\rho(A|B) = \exp\Big(\log\rho(AB) - \log\big[\mathbb{1}_A \otimes \rho(B)\big]\Big) = \lim_{n\to\infty}\Big(\rho(AB)^{1/n}\big[\mathbb{1}_A \otimes \rho(B)\big]^{-1/n}\Big)^n, \tag{3.13}$$

where the eigenvalues of this operator are related to the separability of the underlying bipartite state $\rho(AB)$ [59, 60]. There have been other definitions proposed in addition to the one described here (see, e.g., Ref. [61]).

With the definition (3.13), the conditional entropy (3.12) is equivalently

$$S(A|B) = S(AB) - S(B), \tag{3.14}$$

in direct analogy with (3.5). In the classical limit the matrix (3.13) is diagonal, with elements $p_{i|j}$, and the conditional entropy (3.12) coincides with the classical result (3.3). However, it is possible for the eigenvalues of $\rho(A|B)$ to exceed unity. This leads to a *negative* conditional von Neumann entropy, something that is impossible in classical physics. That is, under certain conditions $S(AB) < S(B)$, which occurs when A and B are entangled [58, 59]. For the Bell state $|AB\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, the joint entropy is zero, $S(AB) = 0$, while the marginal entropies are $S(A) = S(B) = 1$, so that the conditional entropies are negative, $S(A|B) = S(B|A) = -1$. This negativity of conditional entropies is associated with a violation of entropic Bell inequalities [62] and has an operational meaning in the context of partial quantum information and quantum state merging protocols [63].

Similarly to the conditional entropy (3.12), the mutual von Neumann entropy,

$$S(A:B) = -\mathrm{Tr}\big[\rho(AB)\log\rho(A:B)\big], \tag{3.15}$$

can be defined with a mutual amplitude matrix [58],

$$\rho(A{:}B) = \exp\Big(\log\big[\rho(A)\otimes\rho(B)\big] - \log\rho(AB)\Big) = \lim_{n\to\infty}\Big(\big[\rho(A)\otimes\rho(B)\big]^{1/n}\rho(AB)^{-1/n}\Big)^n. \tag{3.16}$$

This yields the quantum version of (3.6),

$$S(A:B) = S(A) + S(B) - S(AB) = S(A) - S(A|B). \tag{3.17}$$

The mutual entropy is a fundamental quantity used in the description of the capacity of quantum channels and for quantum error-correcting codes [64].

3.3.2 Inequalities for von Neumann Entropy

The von Neumann entropy of two systems A and B satisfies the subadditivity condition,

$$S(AB) \leq S(A) + S(B), \tag{3.18}$$

in direct analogy with the classical result (3.8). However, in contrast to (3.9), the joint von Neumann entropy has the lower bound

$$S(AB) \geq |S(A) - S(B)|, \tag{3.19}$$

which is known as the triangle inequality or the Araki-Lieb inequality [65]. For example, if A and B are in an entangled pure state, then $S(AB) = 0$ and $S(A) = S(B) \geq 0$. This implies that the joint state of the system is fully known, but it cannot be determined from the subsystems alone since they are entropic.

Both of these inequalities can be extended to an inequality for three systems, A, B and C. This property is known as strong subadditivity [66, 67],

$$S(ABC) + S(B) \leq S(AB) + S(BC), \tag{3.20}$$

and there is a corresponding version for Shannon entropy. This expression can be recast in two equivalent forms. First, in terms of conditional entropies,

$$S(C|B) - S(C|BA) \geq 0, \tag{3.21}$$

and second in terms of a conditional mutual entropy,

$$S(A : C|B) \geq 0. \tag{3.22}$$

When discussing consecutive measurements in Chapter 7, we will see that this quantity vanishes for a particular set of systems A, B, and C, and is called a quantum Markov chain [68–70].

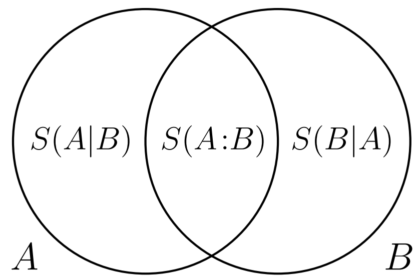

Figure 3.1: Entropy Venn diagram for two systems A and B.

3.3.3 Entropy Venn Diagrams

A useful tool for visualizing how entropy is distributed between correlated systems is the entropy Venn diagram [59]. A general diagram for two systems, A and B, is shown in Fig. 3.1. In general, the regions that overlap are mutual entropies, $S(A:B)$, while the non-overlapping regions are conditional entropies, $S(A|B)$ and $S(B|A)$. The entropy of the entire system is found by adding together all entries in the diagram. In this case, $S(A|B) + S(A:B) + S(B|A) = S(AB)$. The marginal entropy of one system is the sum of all entropies in its circle. For instance, the entropy of A is obtained by adding together $S(A|B)$ and $S(A:B)$, which yields $S(A)$. To visualize the mathematical operation of tracing over a system in a density matrix, simply ignore the entries that are only in its circle. For example, tracing over B implies that one is left with $S(A|B)$ and $S(A:B)$, which is just $S(A)$, the entropy of A.

If the state $\rho(AB)$ underlying the diagrams in Fig. 3.1 is a pure state, then the joint entropy necessarily vanishes, $S(AB) = 0$. From the Schmidt decomposition (discussed later in Sec. 3.4.1), it follows that $S(A) = S(B)$, and hence

$$S(A|B) = -S(A), \tag{3.23}$$

$$S(A:B) = 2\,S(A). \tag{3.24}$$

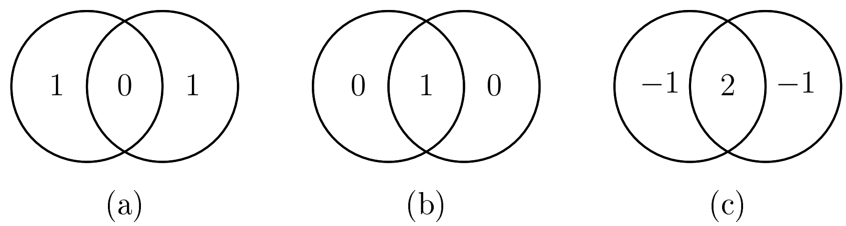

Figure 3.2: Examples of entropy Venn diagrams for two systems. (a) Both systems are independent (uncorrelated); (b) perfectly correlated; and (c) quantum entangled.

There are three important cases that capture the types of correlations that can exist between two systems, A and B (recall the correlations in the context of density matrices that were discussed in Ch. 2.3). We show the entropy Venn diagram for each case in Fig. 3.2. First, the two systems can be completely uncorrelated, $\rho(AB) = \rho(A) \otimes \rho(B)$. In this case, the total entropy is the sum of the marginal entropies so that $S(A|B) = S(A)$, $S(B|A) = S(B)$, and $S(A:B) = 0$. Second, they can be correlated such that their conditional entropies vanish. That is, the state of one system determines the state of the other system. For example, one possible correlated state is $\rho(AB) = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$. Here, $S(A|B) = S(B|A) = 0$ and $S(A:B) = 1$. These two situations can occur in both classical and quantum systems. The third case occurs when A and B are quantum entangled and is characterized by a negative conditional entropy. One example of an entangled state is $|AB\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, the Bell state. Here, $S(A|B) = S(B|A) = -1$ and $S(A:B) = 2$.

Quantum entanglement constitutes a major departure from the possible correlations in the Shannon theory. Using (3.9) and (3.19) one can show [71] that the mutual entropy $S(A:B)$ is bounded from above by $S(A:B) \leq 2\min[S(A), S(B)]$, which can be twice the classical upper bound $H(A:B) \leq \min[H(A), H(B)]$. This feature is exploited in superdense coding protocols [6]. Quantum entanglement is ubiquitous in quantum computing and communication protocols, one of the most famous being quantum teleportation.

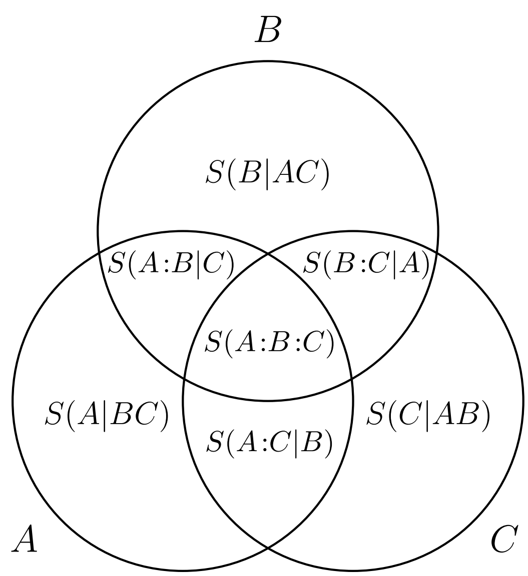

Figure 3.3: Entropy Venn diagram for three systems A, B, and C. The entropy shared by all three systems is the ternary mutual entropy $S(A:B:C)$ and vanishes if $\rho(ABC)$ is a pure state.

The entropy relationships that have been discussed thus far can be easily extended to characterize multipartite systems. Frequently, one is interested in the entropy relationships (see Fig. 3.3) between three systems A, B, and C [59, 60, 72]. For example, the entropy of A conditional on the joint system BC is

$$S(A|BC) = S(ABC) - S(BC). \tag{3.25}$$

The conditional mutual entropy is defined as the mutual entropy between A and B, given the system C:

$$\begin{aligned} S(A:B|C) &= S(A|C) - S(A|BC) \\ &= S(AC) + S(BC) - S(C) - S(ABC). \end{aligned} \tag{3.26}$$

The entropy shared by all three systems is the ternary mutual entropy and is defined as

$$\begin{aligned} S(A:B:C) &= S(A:B) - S(A:B|C) \\ &= S(A) + S(B) + S(C) - S(AB) - S(AC) - S(BC) + S(ABC). \end{aligned} \tag{3.27}$$

If the state $\rho(ABC)$ underlying the system ABC is a pure state, then the ternary mutual entropy always vanishes, $S(A:B:C) = 0$. That is, using the Schmidt decomposition (see Sec. 3.4.1) a bipartite "cut" of the joint system ABC can be made such that $S(A) = S(BC)$, $S(B) = S(AC)$, and $S(C) = S(AB)$. In this case, the entries of the Venn diagram 3.3 are

$$S(A\,|\,BC) = -S(A), \tag{3.28}$$

$$S(A:B\,|\,C) = S(B) + S(A) - S(C), \tag{3.29}$$

$$S(A:B:C) = 0. \tag{3.30}$$

As an illustrative example of the tripartite Venn diagram, consider the following entangled state of three qubits (also known as a GHZ state [31]),

$$|ABC\rangle = \frac{1}{\sqrt{2}}\big(|000\rangle + |111\rangle\big). \tag{3.31}$$

Since the state is evidently pure and entangled, the ternary mutual entropy vanishes and the conditional entropies are negative. However, tracing over any part of the state reveals that each resulting subsystem is a mixed state with an entropy of 1 bit. See Fig. 3.4 (a). This is an example of the principle known as the monogamy of entanglement [73]. That is, if two systems A and B are maximally entangled (e.g., in a Bell state), they cannot be correlated with a third system, C. Partitioning (3.31) into any two subsystems, say, AB and C, reveals that the two halves are maximally entangled. However, tracing over A, B or C yields a

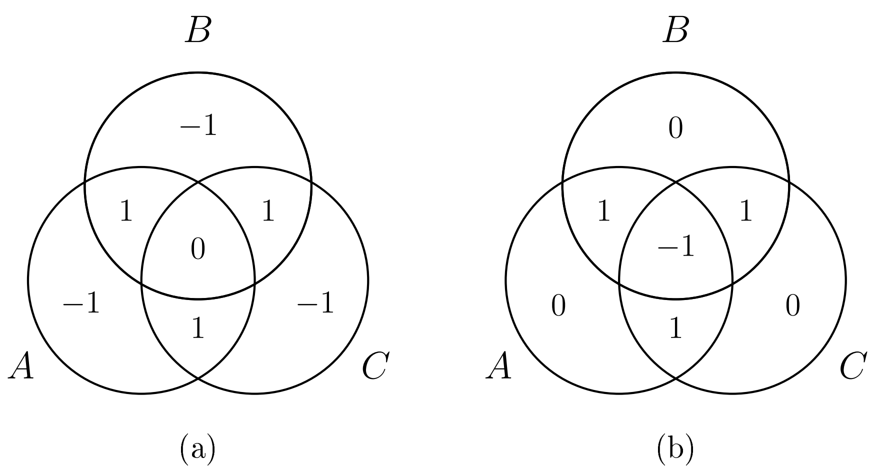

Figure 3.4: Entropy Venn diagram for (a) the GHZ state (3.31) and for (b) XOR encryption.

mixed state for the remaining system, which cannot be a maximally entangled state.

A second example of tripartite Venn diagrams demonstrates a simple type of encryption. Suppose there is a joint system that satisfies the following: $S(ABC) = S(AB) = S(AC) = S(BC) = 2$ bits and $S(A) = S(B) = S(C) = 1$ bit. In this case, the Venn diagram is as shown in Fig. 3.4(b). If the joint state of any two systems is specified, then the state of the third is fully known. That is, $S(A|BC) = 0$. Furthermore, tracing over any subsystem leaves the other two with no information, $S(A : B) = 1 - 1 = 0$. However, conditioning on the state of the third system yields full information, $S(A : B|C) = 1$. For example, consider a message in plain text (A) and an encryption key (B) that produces an encrypted cipher text (C) according to the XOR gate. Without the key, one cannot gain any information about the plain text from the cipher text since $S(A : C) = 0$. With the key, $S(A : C|B) = 1$, and one can extract the original message.

For a general n-partite system, the entropy relationships can be written down in the form of "chain rules" for quantum entropies. Namely, the joint entropy of n systems is

$$S(A_1 \dots A_n) = S(A_1) + S(A_2|A_1) + S(A_3|A_1A_2) + \dots, \tag{3.32}$$

while

$$S(A_1 \ldots A_n : A_{n+1}) = S(A_1 : A_{n+1}) + S(A_2 : A_{n+1}|A_1) + S(A_3 : A_{n+1}|A_1 A_2) + \ldots, \quad (3.33)$$

quantifies the mutual entropy between the n systems and an additional system A_{n+1}. Remarkably, we will see in Ch. 6 that the complementarity relation in the famous Bell-state quantum eraser is due to the chain rule (3.33) for $n + 1 = 3$ systems.

3.4 Tools in Quantum Information

In this section, some of the most prevalent tools and techniques used in quantum information theory will be discussed. I will start with the Schmidt decomposition, which is a way of characterizing the degree of entanglement in bipartite entangled quantum systems. This is closely related to the idea of purification, where a mixed quantum state can be viewed as a pure state in an enlarged Hilbert space. I will also describe how to compute the entropy of mixed quantum states, and in particular derive the joint entropy theorem. Finally, I will derive Holevo's theorem, which establishes an upper bound to the amount of information one can obtain about a quantum system through measurement.

3.4.1 The Schmidt Decomposition and Purification

The Schmidt decomposition is a very useful method for describing bipartite quantum systems and leads to the idea of Schmidt numbers, which are important for characterizing the degree of entanglement between systems. If a composite system, AB, is represented by a pure state,

$|\psi\rangle$, then there exist orthonormal states $|i_A\rangle$ for A and $|i_B\rangle$ for B such that,

$$|\psi\rangle = \sum_i \lambda_i \, |i_A\rangle \, |i_B\rangle. \tag{3.34}$$

Here, λ_i are real, non-negative numbers (one can always absorb any phases into the definitions of the states $|i\rangle$) called Schmidt coefficients that satisfy $\sum_i \lambda_i^2 = 1$.

To show this [25], suppose A and B have Hilbert spaces of the same dimension and that $|j\rangle$ and $|k\rangle$ are orthogonal bases for A and B, respectively. Then, in general the full wave function can be expressed in terms of these bases as

$$|\psi\rangle = \sum_{jk} a_{jk} \, |j\rangle \, |k\rangle, \tag{3.35}$$

where a_{jk} are the complex elements of a matrix a. The singular value decomposition states that this matrix can be expressed as a product of three matrices $a = u\,\sigma\,v$. The diagonal matrix σ has non-negative elements and u, v are unitary matrices. Writing this matrix product in index form,

$$|\psi\rangle = \sum_{ii'jk} u_{ji}\,\sigma_{ii'}\,v_{i'k}\,|j\rangle\,|k\rangle = \sum_{ijk} u_{ji}\,\sigma_{ii}\,v_{ik}\,|j\rangle\,|k\rangle. \tag{3.36}$$

The second equality follows since σ is a diagonal matrix.

Next, redefine the states $|j\rangle$ and $|k\rangle$ as $|i_A\rangle = \sum_j u_{ji}\,|j\rangle$ and $|i_B\rangle = \sum_k v_{ik}\,|k\rangle$. These new states are still orthonormal since $\langle i'_A|i_A\rangle = \sum_j u_{ji}\,u^*_{ji'} = \delta_{ii'}$ and $\langle i'_B|i_B\rangle = \sum_k v_{ik}\,v^*_{i'k} = \delta_{ii'}$. Finally, defining $\lambda_i = \sigma_{ii}$ yields the Schmidt decomposition,

$$|\psi\rangle = \sum_i \lambda_i \, |i_A\rangle \, |i_B\rangle. \tag{3.37}$$

It follows directly from this decomposition, that the eigenvalues of the density matrices for A and B are identical since

$$\rho_A = \sum_i \lambda_i^2 \, |i_A\rangle\langle i_A|,$$
$$\rho_B = \sum_i \lambda_i^2 \, |i_B\rangle\langle i_B|. \tag{3.38}$$

Therefore, the two subsystems have the same entropy,

$$S(A) = S(B) = H[\lambda^2]. \tag{3.39}$$

Note that $|i_A\rangle$ and $|i_B\rangle$ are called Schmidt bases for A and B, and that the Schmidt number corresponds to the number of non-zero values λ_i. In particular, a state $|\psi\rangle$ is a product state of A and B if and only if it has a Schmidt number of one. Therefore, the Schmidt number is a measure of the entanglement between A and B. A maximally entangled state (of dimension d) has Schmidt coefficients $\lambda_i = 1/d$.

The Schmidt decomposition (3.34) is not unique since one can always perform unitary transformations on the two systems individually. That is, if $|\psi\rangle = \sum_i \lambda_i \, |i_A\rangle \, |i_B\rangle$ is a Schmidt decomposition for A and B, then so too is $U \otimes V|\psi\rangle = \sum_i \lambda_i \left(U|i_A\rangle\right) \left(V|i_B\rangle\right)$. This implies that the entanglement of a bipartite pure state is invariant under unitary transformations.

Interestingly, a Schmidt decomposition does not exist in general for a tripartite system [74,75]. A tripartite pure state $|ABC\rangle = \sum_i \lambda_i \, |i_A\rangle \, |i_B\rangle \, |i_C\rangle$ has a Schmidt decomposition if the "bi-Schmidt basis" $\langle i_A|ABC\rangle$ (and similarly for $\langle i_B|ABC\rangle$ and$\langle i_C|ABC\rangle$) has a Schmidt number of one [74].

Using the Schmidt decomposition one can use a technique known as purification. Suppose one has a quantum state $\rho(A)$ for some system A. By introducing a second "reference" system R, one finds that the total state of A and R can be written as a pure state $|AR\rangle$. This procedure is called purification [25] and is commonly used in quantum information theory. To show that this can indeed be done, first write A in its eigenbasis as $\rho(A) = \sum_i p_i \left|i_A\right\rangle\left\langle i_A\right|$. Then, with the reference system, which is defined to have the same state space as A and an orthonormal basis $\left|i_R\right\rangle$, one can write the state of A and R together as

$$\left|AR\right\rangle = \sum_i \sqrt{p_i} \left|i_A\right\rangle \left|i_R\right\rangle. \tag{3.40}$$

Tracing out R recovers the state $\rho(A)$ so that $\left|AR\right\rangle$ is indeed a purification of $\rho(A)$.

3.4.2 Joint Entropy Theorem

Very often in quantum information theory one encounters density matrices that are block-diagonal in structure, sometimes called classical-quantum states. These states will be used frequently in the following chapters when discussing parallel and consecutive measurements. The joint entropy theorem [25], which I will show a simple proof of, provides a way to compute the entropy of such states.

Suppose one has a density matrix of the form

$$\rho = \sum_i p_i \left|i\right\rangle\left\langle i\right| \otimes \rho_i. \tag{3.41}$$

Here, p_i are probabilities, $\left|i\right\rangle$ are orthogonal states for system A, and ρ_i is any set of density matrices for system B. This is called a *classical-quantum* state because A is diagonal in the basis $\left|i\right\rangle$ (i.e., classical), while B is a general (pure or mixed) quantum state ρ_i (see, e.g., [76]).

The states ρ_i of B can be diagonalized such that their spectral decomposition is

$$\rho_i = \sum_j \lambda_{i,j} \, |e_{i,j}\rangle\langle e_{i,j}|, \tag{3.42}$$

where $\lambda_{i,j}$ are eigenvalues that sum to one for each i: $\sum_j \lambda_{i,j} = 1$. The orthogonality of the eigenstates guarantees that

$$\langle e_{i,j'} | e_{i,j} \rangle = \delta_{jj'}. \tag{3.43}$$

However, in general,

$$\langle e_{i',j} | e_{i,j} \rangle \neq \delta_{ii'}, \tag{3.44}$$

unless ρ_i have orthogonal support.

Using this decomposition, the state (3.41) becomes

$$\rho = \sum_{ij} p_i \, \lambda_{i,j} \, |i \; e_{i,j}\rangle\langle i \; e_{i,j}|. \tag{3.45}$$

The joint states $|i \; e_{i,j}\rangle$ are orthogonal because $\langle i' | i \rangle = \delta_{ii'}$,

$$\langle i' \; e_{i',j} | i \; e_{i,j} \rangle = \delta_{ii'}. \tag{3.46}$$

It follows that the entropy of (3.41) is

$$S(\rho) = -\sum_{ij} p_i \, \lambda_{i,j} \, \log \, p_i \, \lambda_{i,j}. \tag{3.47}$$

Expanding the logarithm and performing the partial sums, this simplifies to

$$S(\rho) = H[p] + \sum_i p_i \, S(\rho_i), \tag{3.48}$$

where $H[p]$ is the Shannon entropy of the probability distribution p_i and

$$S(\rho_i) = -\sum_j \lambda_{i,j} \, \log \, \lambda_{i,j}. \tag{3.49}$$

Although the basis states of ρ_i are not orthogonal to those of $\rho_{i'}$, the block-diagonal structure of ρ yields an exact result for $S(\rho)$.

3.4.3 Entropy of Mixed States

Recall from Ch. 2.2 that a mixed quantum state is written as a summation of states ρ_i that each occur with some probability p_i. That is, the density matrix for a system A is

$$\rho(A) = \sum_i p_i \, \rho_i. \tag{3.50}$$

In general, the ρ_i may not be orthogonal so that the entropy of $\rho(A)$ falls somewhere between a lower and upper bound.

A lower bound is established on $S(A)$ by using the concavity of entropy. Consider an additional system, B, with an orthonormal basis $|i\rangle$ such that the joint state of A and B is

$$\rho(AB) = \sum_i p_i \, \rho_i \otimes |i\rangle_B \langle i|. \tag{3.51}$$

One can easily check that tracing this state over B recovers the state of A written above. From the joint entropy theorem in the previous Sec. 3.4.2, the entropy of this block-diagonal state is

$$S(AB) = H[p] + \sum_i p_i \, S(\rho_i). \tag{3.52}$$

The subadditivity of entropy (see Sec. 3.3.2) states that this quantity is less than or equal to the sum of the entropies of its subsystems. That is,

$$S(AB) \leq S(A) + S(B). \tag{3.53}$$

The entropy of system B is simply $S(B) = H[p]$, as can be seen by tracing $\rho(AB)$ over A. Together with Eqs. (3.52) and (3.53), this yields the lower bound on the entropy of the mixed state $\rho(A)$,

$$\begin{aligned} S(A) &\geq S(AB) - S(B) \\ &\geq \sum_i p_i \, S(\rho_i), \end{aligned} \tag{3.54}$$

with equality if and only if the ρ_i are identical.

If the states ρ_i in Eq. (3.50) are orthogonal then the entropy of A can be computed exactly. The spectral decomposition of the states ρ_i yields

$$\rho_i = \sum_j \lambda_{i,j} \, |e_{i,j}\rangle\langle e_{i,j}|, \tag{3.55}$$

where $\lambda_{i,j}$ are the eigenvalues of ρ_i and sum to one for a given i: $\sum_j \lambda_{i,j} = 1$. The orthogonality of the eigenstates guarantees that

$$\langle e_{i,j'} | e_{i,j}\rangle = \delta_{jj'}. \tag{3.56}$$

Since the ρ_i were required to be orthogonal, they have support on orthogonal subspaces so that every basis state in Hilbert space $\mathcal{H}_i$ is orthogonal to every other basis state in Hilbert space $\mathcal{H}_{i'}$:

$$\langle e_{i',j} | e_{i,j} \rangle = \delta_{ii'}. \tag{3.57}$$

Inserting the spectral decomposition (3.55) into (3.50),

$$\rho(A) = \sum_{ij} p_i \, \lambda_{i,j} \, |e_{i,j}\rangle\langle e_{i,j}|, \tag{3.58}$$

it is clear that $p_i \, \lambda_{i,j}$ are the eigenvalues and $|e_{i,j}\rangle$ are the eigenvectors of $\rho(A)$. The entropy of $\rho(A)$ is then simply

$$S(A) = -\sum_{ij} p_i \, \lambda_{i,j} \, \log \, p_i \, \lambda_{i,j}. \tag{3.59}$$

Expanding the logarithm and performing the partial sums, this simplifies to

$$S(A) = H[p] + \sum_i p_i \, S(\rho_i) \tag{3.60}$$

where $H[p]$ is the Shannon entropy of the probability distribution p_i and

$$S(\rho_i) = -\sum_j \lambda_{i,j} \, \log \, \lambda_{i,j}. \tag{3.61}$$

Note that if the states ρ_i did not have support on orthogonal subspaces, the following inequality would hold

$$S(A) < H[p] + \sum_i p_i S(\rho_i). \tag{3.62}$$

Therefore, a general state (3.50) satisfies the two inequalities,

$$\sum_i p_i S(\rho_i) \leq S(A) \leq H[p] + \sum_i p_i S(\rho_i). \tag{3.63}$$

For instance, if the states ρ_i in Eq. (3.50) are all pure, then the entropy of the system A is bounded by $0 \leq S(A) \leq H[p]$. If, in addition, the states are orthogonal, then $S(A) = H[p]$.

3.4.4 Holevo's Theorem

Holevo's theorem establishes an upper bound to the amount of accessible information that can be extracted about a quantum state through measurement [25]. To derive the limit, first suppose that there is a preparer P that has a classical random variable X with probabilities p_x. In addition, the preparer has access to a set of density matrices, ρ_x, of a quantum system, Q. The total state is described by the classical-quantum density matrix,

$$\rho(PQ) = \sum_x p_x \, |x\rangle\langle x| \otimes \rho_x. \tag{3.64}$$

Tracing over P, the state of Q is given by the mixed density matrix,

$$\rho(Q) = \sum_x p_x \, \rho_x = \rho. \tag{3.65}$$

Using the results of the joint entropy theorem in Sec. 3.4.2, the mutual entropy between the preparer P and the quantum preparation Q is

$$S(P:Q) = S(\rho) - \sum_x p_x \, S(\rho_x) = \chi, \tag{3.66}$$

where χ is usually called the Holevo information.

A measurer, M, can attempt to determine the state x of the quantum system through a measurement of Q. The measurement is implemented with the unitary entangling operation $U = \sum_x P_x \otimes U_x$ between Q and M, where $P_x = |x\rangle\langle x|$ are projectors on the state of Q and the unitary matrices U_x move the ancilla from the initial state $|0\rangle$ to the final state $|x\rangle = U_x|0\rangle$. This operator will be discussed in greater detail later in Ch. 4. If the measurer starts in the initial state $|0\rangle$, then after the measurement the full state becomes (using primes to clearly distinguish the post-measurement systems)

$$\rho(P'Q'M') = \sum_{xyz} p_x\, |x\rangle\langle x| \otimes P_y\, \rho_x\, P_z^\dagger \otimes |y\rangle\langle z|. \tag{3.67}$$

Tracing out the quantum system Q, the density matrix of the preparer and the measurer is

$$\rho(P'M') = \sum_{xy} p_x \operatorname{Tr}\big(P_y\, \rho_x\big)\, |x\rangle\langle x| \otimes |y\rangle\langle y|. \tag{3.68}$$

where $\operatorname{Tr}(P_y\, \rho_x\, P_z^\dagger) = \operatorname{Tr}(P_y\, \rho_x)\, \delta_{yz}$. The entropy of this state is just the Shannon entropy of the probability distribution $q_{xy} = p_x \operatorname{Tr}\big(P_y\, \rho_x\big)$,

$$S(P'M') = H(P'M') = H\,[q] = -\sum_{xy} \big[p_x \operatorname{Tr}\big(P_y\, \rho_x\big)\big] \log \big[p_x \operatorname{Tr}\big(P_y\, \rho_x\big)\big]\,. \tag{3.69}$$

The mutual entropy between the classical preparer P and the measurer M after the measurement $S(P' : M')$ is now shown to be bounded from above by the mutual entropy of P and Q before the measurement $S(P : Q)$. For the remainder of the proof, I will follow the derivation laid out by Cerf and Adami in Ref. [77]. The measurement (3.67) does not change the entanglement between P and QM so that

$$S(P' : Q'M') = S(P : QM) = S(P : Q). \tag{3.70}$$

The second equality follows from the fact that before the measurement M is uncorrelated with P and Q. According to the chain rule, $S(P' : Q'M') = S(P' : M') + S(P' : Q'|M')$, which, together with the previous expression, leads to

$$S(P' : M') = S(P : Q) - S(P' : Q'|M'). \tag{3.71}$$

By the strong subadditivity of quantum entropy [66,67] in Eq. (3.20), the conditional mutual entropy is always non-negative $S(P' : Q'|M') \geq 0$, and it follows that

$$S(P' : M') \leq S(P : Q) = \chi. \tag{3.72}$$

The quantity χ is the maximum amount of information that the measurer can extract about the state of the quantum system.

For example, suppose the states ρ_x in (3.64) are pure. The Holevo information (3.66) in this case is simply $\chi = S(\rho)$. This quantity is bounded between zero and $H[p]$ according to Sec. 3.4.3. In particular, if the states ρ_x are pure *and* orthogonal, then the Holevo information is the Shannon entropy of the probability distribution p_x, $\chi = H[p]$. The Holevo information is further reduced from $S(\rho)$ if the states ρ_x are mixed.

3.5 Summary

In this chapter, I reviewed the Shannon theory of information in order to set the stage for quantum information theory. I showed how quantum superposition and entanglement lead to striking differences between the quantum and classical theory of information, which can be illustrated using entropy Venn diagrams. These diagrams will prove useful in the following chapters on quantum measurement. I discussed in Sec. 3.4 some of the important tools that are commonly used to describe quantum systems. For example, I showed how the Schmidt decomposition can be used to characterize entanglement as well as methods for computing entropies of entangled systems.

Chapter 4

Introduction to Quantum Measurement

4.1 Basics of Quantum Measurement Theory

As we have seen from the no-cloning theorem, classical intuitions fail us when carried over to the measurement of quantum systems. That these systems can be in superpositions or entangled necessarily makes measuring them a non-trivial task. Classically, we can always, at least in principle, perfectly "copy" the state of an object onto a measuring device and succeed in getting the same result every time. The no-cloning theorem of Sec. 2.4 generally removes this feature in quantum systems. In fact, the measuring device becomes entangled with the quantum system and, as a result, the measurement outcomes are probabilistic.

The measurement postulate of quantum mechanics states that for a measurement of a quantum system, Q, there exists a set of measurement operators, $\{M_m\}$, that act on the Hilbert space of Q, yielding an outcome m with some probability,

$$p_m = \langle\psi|M_m^\dagger M_m|\psi\rangle, \tag{4.1}$$

where $|\psi\rangle$ is the state of Q before the measurement [25]. For an outcome m, the state of Q after the measurement becomes

$$|\psi\rangle \mapsto \frac{M_m\,|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}. \tag{4.2}$$

The probabilities, p_m, sum to one, which implies the completeness relation for the measurement operators,

$$\sum_m M_m^\dagger M_m = \mathbb{1}. \tag{4.3}$$

Within quantum measurement theory, there are two important cases known as Positive Operator-Valued Measure (POVM) measurements and Projection-Valued Measure (PVM) measurements. The latter, which are also known as projective or von Neumann measurements, were discussed in detail by von Neumann [22]. In this work, I will focus primarily on projective measurements (both weak and strong measurements, which will be discussed later), although it is useful to have an understanding of the features of both types of measurement scenarios.

Measurements in quantum mechanics can be generally described using two mathematical tools known as a Positive Operator-Valued Measure (POVM) and a Projection-Valued Measure (PVM). The latter are used to implement projective, or von Neumann measurements. Although projective measurements are seen as a special case of POVM measurements, by Naimark's theorem (sometimes also called Neumark's theorem), a POVM can always be realized by a projective measurement on an extended Hilbert space [78]. That is, the combination of unitary interactions between the quantum system and an ancillary system, followed by a projective measurement is equivalent to a POVM measurement. In this section, I provide an outline of POVM and PVM measurements, as well as a description of measurement when accounting for an ancillary system.

Recall that a measurement corresponds to a set of measurement operators, M_m. In a

POVM measurement, one defines the operators

$$E_m = M_m^\dagger M_m, \tag{4.4}$$

and calls them the elements of the POVM. The elements of a POVM do not necessarily commute with each other, but are Hermitian ($E_m = E_m^\dagger$) and positive ($\langle\psi|E_m|\psi\rangle \geq 0$) operators that satisfy the completeness relation (4.3). The set of all elements, $\{E_m\}$, is called the POVM. The apparatus that one uses to measure the quantum system is represented by the POVM and the elements, E_m, are chosen to correspond to each possible measurement outcome such that they yield the measurement probabilities

$$p_m = \langle\psi|E_m|\psi\rangle. \tag{4.5}$$

If the quantum system is described by a density matrix, ρ, then the probabilities are given instead by the trace,

$$p_m = \mathrm{Tr}\big(E_m\,\rho\big), \tag{4.6}$$

and the post-measurement state is

$$\rho \mapsto \rho_m = \frac{M_m\,\rho\,M_m^\dagger}{\mathrm{Tr}\big(M_m^\dagger M_m\,\rho\big)}. \tag{4.7}$$

For projective measurements, one imposes the requirement that the measurement operators, M_m, are orthogonal projectors, $P_m = |m\rangle\langle m|$, so that

$$P_m^\dagger\,P_{m'} = \delta_{mm'}\,P_m. \tag{4.8}$$

That is, the POVM elements are the same as the measurement operators, $E_m = P_m^\dagger P_m = P_m$. In this case, it is straightforward to show that repeated projective measurements yield the same measurement outcome: if we measure a system and find the result m, then subsequent measurements will also yield m, assuming that no decoherence has taken place between consecutive measurements. This is not true for POVM measurements since the measurement operators, M_m, are not necessarily all orthogonal. According to the spectral decomposition, a Hermitian operator $\hat{A}$ corresponding to some observable can be diagonalized and written in the form

$$\hat{A} = \sum_m \lambda_m \, P_m, \tag{4.9}$$

with eigenvalues λ_m and projectors P_m. Since the operators P_m are orthogonal, the number of operators in the decomposition (4.9) is equal to the dimension of A's Hilbert space. In contrast, a POVM can have an unlimited number of elements.

According to Naimark's theorem [78], the measurement process can be described as a unitary interaction between a quantum system, Q, and an ancillary system, M, sometimes called the meter. The coupling induced by the measurement is due to the interaction Hamiltonian,

$$\mathcal{H} = g \, \hat{A} \otimes \hat{M}. \tag{4.10}$$

The parameter g controls the strength of the coupling between Q and M, and the operator corresponding to the observable being measured is $\hat{A}$. The operator $\hat{M}$ acts on the pointer variable of the meter and moves it from its initial state depending on the state of the quantum system. Here, it is assumed that either the measurement occurs on a short enough time scale so that the free evolution of Q can be neglected, or that the observable commutes with the free Hamiltonian of Q. Note that since $\hat{A}$ and $\hat{M}$ act in different Hilbert spaces, they commute. If the measurement has a duration of t, the Hamiltonian generates a translation in time via the unitary operator

$$U = e^{-it\mathcal{H}/\hbar}. \tag{4.11}$$

In a more realistic scenario, the coupling parameter in the measurement Hamiltonian (4.10) is time dependent. If $g(t)$ is nonzero in the interval $[0, T]$ and normalized according to $\int_0^T g(t)dt = g_0$, then we replace expression (4.11) by

$$U = \mathcal{T}\, e^{-i\int_o^T \mathcal{H}(t)dt/\hbar}, \tag{4.12}$$

where $\mathcal{T}$ indicates the time-ordering operator (which can be ignored if $\mathcal{H}$ commutes with itself at different times). In protective quantum measurements [79–81], for example, the interaction is taken to be very slow and weak so that the free evolution of the quantum system and meter can no longer be neglected. In such a measurement, the interaction is adiabatic if $g(t)$ is sufficiently smooth and thus the quantum system remains in its initial eigenstate throughout the measurement. In this work, we assume the interaction is fast (the impulse approximation) and use (4.11) for the time evolution.

If the operator $\hat{A}$ in the Hamiltonian of Eq. (4.10) is written in its spectral decomposition with eigenvalues a_n and d eigenvectors $|a_n\rangle$,

$$\hat{A} = \sum_n a_n\, P_n = \sum_n a_n\, |a_n\rangle\langle a_n|, \tag{4.13}$$

then the unitary operator (4.11) can also be expressed as

$$U = \sum_n P_n \otimes U_n = \sum_n P_n \otimes e^{-i\,g\,t\,a_n\,\hat{M}/\hbar}, \tag{4.14}$$

where $U_n = \exp(-i\,g\,t\,a_n\,\hat{M}/\hbar)$ are unitary operators on the pointer variable. In this form, it is clear that if the quantum system is in the state $|a_n\rangle$, the pointer is shifted by an amount controlled by the eigenvalue a_n. That is, writing the initial quantum system in the eigenbasis of the observable as $|Q_i\rangle = \sum_m \alpha_m\,|a_m\rangle$ with complex amplitudes α_m, the measurement unitary (4.14) entangles Q and M according to

$$U\,|Q_i\rangle|M_i\rangle = \sum_m \alpha_m\,|a_m\rangle \otimes e^{-i\,g\,t\,a_m\,\hat{M}/\hbar}\,|M_i\rangle = \sum_m \alpha_m\,|a_m\rangle \otimes |a_m\rangle_M. \tag{4.15}$$

Here, $|M_i\rangle$ is the initial state of the meter, and $|a_m\rangle_M = \exp(-i\,g\,t\,a_m\,\hat{M}/\hbar)\,|M_i\rangle$ are its d final states.

The final state of the quantum system can be computed from the total wave function (4.15). Tracing over the states of the meter, the density matrix of the quantum system is

$$\rho_Q = \mathrm{Tr}_M\Big(U|Q_i\rangle\langle Q_i| \otimes |M_i\rangle\langle M_i|U^\dagger\Big), \tag{4.16}$$

which has been left general since the final pointer states may not be orthogonal. In contrast, the states $|a_m\rangle$ of the quantum system are orthogonal, so that the density matrix of the meter is

$$\rho_M = \sum_m |\alpha_m|^2\,|a_m\rangle_M\langle a_m|. \tag{4.17}$$

One of the most well-known examples of quantum measurement is the Stern-Gerlach experiment [82]. Here, the spin component of a spin-1/2 particle is measured by sending it through a Stern-Gerlach apparatus. The interaction with a inhomogeneous magnetic field couples the spin of the atom with its spatial degree of freedom (the meter) resulting in two deflected beams with different spin. If these beams are then observed, one has obtained information about the state of Q. For example, a measurement of the spin along $\hat{z}$ corresponds to the Hamiltonian $\mathcal{H} = -g\,\hat{\sigma}_z \otimes \hat{z}$. The eigenvalues of the observable are ± 1 so that the final state of the quantum system and pointer is

$$\begin{aligned} U\,|Q_i\rangle|M_i\rangle &= \alpha_0\,|+1\rangle \otimes |+1\rangle_M + \alpha_1\,|-1\rangle \otimes |-1\rangle_M \\ &= \alpha_0\,|+1\rangle \otimes e^{-i\,g\,t\hat{z}/\hbar}\,|M_i\rangle + \alpha_1\,|-1\rangle \otimes e^{+i\,g\,t\hat{z}/\hbar}\,|M_i\rangle. \end{aligned} \tag{4.18}$$

Thus, the pointer is shifted up or down according to $\exp(\pm i\,g\,t\,\hat{z}/\hbar)$.

4.2 Qubits and Quantum Measurements in Experimental Settings

There are many different physical realizations of qubit systems, ranging from photons [83–85] and the spin states of trapped ions [86–88] or quantum dots [89], to artificial atoms fabricated from superconducting circuits. These latter qubits are fabricated with Josephson junctions [90] and are designed to operate optimally under different conditions (see Refs. [91, 92] for a review). The three basic types of superconducting qubits are the charge qubit [93–96] (logical qubit states correspond to the number of excess Cooper pairs on a superconducting island), also known as the Cooper pair box, the phase qubit [97–99] (logical states are associated with the phase across the junction), and the flux qubit [100, 101] (logical states correspond to supercurrent flowing clockwise or anticlockwise around a loop). In this section, the theoretical description of quantum systems and measurements discussed thus far will be connected to typical experimental settings. In particular, I will summarize single-qubit operations and measurements in optical systems where the qubit states are constructed from the degrees of freedom of a photon.

4.2.1 Photonic Qubits

Linear optical quantum computing (LOQC), which uses photons as the basic unit of information, is one promising pathway to universal quantum computation [83, 84, 102, 103]. In the dual-rail representation [25], the qubit is a single photon in one of two optical modes. These modes can be the spatial, polarization, or time degrees of freedom of the photon. The state $|n\rangle_a \otimes |n'\rangle_b$ is one where there are n (n') photons in mode a (b). The logical states of the qubit are written in terms of the occupation numbers of the two modes as

$$\begin{aligned} |0\rangle_L &= |1\rangle \otimes |0\rangle, \\ |1\rangle_L &= |0\rangle \otimes |1\rangle. \end{aligned} \tag{4.19}$$

An alternative encoding is the single-rail representation [104], where one considers only one optical mode but with two distinct Fock states. For example, the logical states could be $|0\rangle_L = |0\rangle$ and $|1\rangle_L = |1\rangle$, which are the vacuum and single-photon states. The discussion here will be in the context of the dual-rail representation.

Single-qubit operations on the logical states are performed with standard optical elements. Furthermore, arbitrary single-photon unitary gates can be formed from a combination of beam splitters and phase shifters [25, 105]. Measurements of photonic qubits can be made in a destructive manner using photodetectors to convert incident photons into a current, or, the in the case of weak measurements, with optical elements like wave plates [106] and birefringent crystals [24]. First, a description of photonic qubits encoded with spatial states is given, followed by a discussion of polarization states.

The logical states of a spatial qubit can be written in terms of the occupation numbers for the two ports of a beam splitter, for example. That is, $|0\rangle_L = |10\rangle$ and $|1\rangle_L = |01\rangle$ could be used to indicate a single photon entering the first or the second port, respectively. Operations on these logical states can be performed with beam splitters and phase shifters. A beam splitter is an optical element containing a partially reflective surface with some degree of reflection and transmission. For a 50-50 (symmetric) beam splitter, the reflection and transmission coefficients are both 1/2. The unitary matrix for a general beam splitter is [83, 84]

$$U_{BS} = \begin{pmatrix} \cos\theta & -e^{i\Delta}\sin\theta \\ e^{-i\Delta}\sin\theta & \cos\theta \end{pmatrix}, \tag{4.20}$$

where the transmission and reflection is controlled by θ and Δ accounts for possible phase shifts due to the material used in the beam splitter. For instance, a beam splitter with no phase shift is

$$U_{BS}(\Delta = 0) = R_{\hat{y}}(2\theta), \tag{4.21}$$

which is a rotation by 2θ around the $\hat{y}$ axis (recall Eq. (2.3) for a general rotation in the Bloch sphere representation). If the beam splitter is symmetric, it acts on the logical states to produce

$$\begin{aligned} U_{SBS}\,|0\rangle_L &= \frac{|0\rangle_L + |1\rangle_L}{\sqrt{2}}, \\ U_{SBS}\,|1\rangle_L &= \frac{|0\rangle_L - |1\rangle_L}{\sqrt{2}}. \end{aligned} \tag{4.22}$$

On the other hand, a beam splitter with phase shift $\Delta = \pi/2$ is a rotation around the $\hat{x}$ axis,

$$U_{BS}(\Delta = \pi/2) = R_{\hat{x}}(2\theta). \tag{4.23}$$

Another basic optical element is the phase shifter with the corresponding gate [103]

$$U_P = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}. \tag{4.24}$$

This shifts just the mode $|1\rangle_L$ by $e^{i\phi}$. Up to an overall phase, this gate is equivalent to a rotation by ϕ about the $\hat{z}$ axis,

$$U_P = e^{i\phi/2}\, R_{\hat{z}}(\phi). \tag{4.25}$$

It can be shown [25] that an arbitrary single-qubit operation can be decomposed into rotations about $\hat{z}$ and about $\hat{y}$. Thus, beam splitters and phase gates can implement any single-qubit gate. Furthermore, they can be combined to construct other types of gates, for example, the Hadamard gate (2.11). First, note that the symmetric beam splitter with $\Delta = -\pi/2$,

$$U_{SBS}(\Delta = -\pi/2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = R_{\hat{x}}(-\pi/2), \tag{4.26}$$

is equivalent to a $\sqrt{\text{NOT}}$ gate (which has the property that $\sqrt{\text{NOT}}\,\sqrt{\text{NOT}} = \text{NOT}$) up to an overall phase [102, 103, 107]. By placing phase shifters at the input and output ports of the beam splitter, this turns into the Hadamard gate:

$$H = U_P(-\pi/2)\; U_{SBS}(\Delta = -\pi/2)\; U_P(-\pi/2). \tag{4.27}$$

Similarly, a single π phase gate at the input produces the Hadamard gate,

$$H = U_{SBS}(\Delta = 0)\; U_P(\pi). \tag{4.28}$$

In order to construct two-qubit gates (which are used to entangle qubits) two photons must be made to interact. In principle, this can be done using nonlinear media via the optical Kerr effect [108]. In this way, gates that are necessary for universal quantum computation can be realized (e.g., the controlled-phase and controlled-NOT gates) [102]. However, the nonlinearities required of a material must be very large to induce a π phase shift (according to O'Brien [85], no such material yet exists) and so the proposal for LOQC without nonlinear media [83, 109] may be more promising [110].

In contrast to the spatial qubit, the logical states of a polarization qubit are, for example, the horizontal and vertical linear polarization of the photon, $|0\rangle_L = |H\rangle$ and $|1\rangle_L = |V\rangle$. A general state of a polarization qubit is written in terms of these logical states as (dropping the subscript L)

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle. \tag{4.29}$$

For a photon propagating along the $\hat{z}$ direction, one can define $|0\rangle = |H\rangle$ as the (linear) horizontal polarization along the $\hat{x}$ direction and $|1\rangle = |V\rangle$ as the (linear) vertical polarization along $\hat{y}$. A general linearly polarized photon is then a superposition of these states

$$|\psi\rangle = a\,|H\rangle + b\,|V\rangle, \tag{4.30}$$

where a and b are real.

In the Bloch sphere representation, linearly polarized photons are described by vectors in the xz plane: $|H\rangle$ and $|V\rangle$ are the eigenstates of σ_z, while the eigenstates of σ_x are the diagonal and anti-diagonal polarization states,

$$\begin{aligned} |D\rangle &= \frac{|H\rangle + |V\rangle}{\sqrt{2}}, \\ |A\rangle &= \frac{|H\rangle - |V\rangle}{\sqrt{2}}. \end{aligned} \tag{4.31}$$

On the other hand, circularly polarized photons are written in terms of the left- and right-circular polarization states

$$\begin{aligned} |L\rangle &= \frac{|H\rangle + i\,|V\rangle}{\sqrt{2}}, \\ |R\rangle &= \frac{|H\rangle - i\,|V\rangle}{\sqrt{2}}, \end{aligned} \tag{4.32}$$

which are the eigenstates of σ_y.

The polarization states of a photon are manipulated using standard optical elements like wave plates. Wave plates are made from birefringent material and are used to rotate the polarization of single photons (see Ch. 6 for wave plates in the context of the quantum eraser experiment) by imparting a polarization-dependent phase shift to the incident photon. The wave plate has different refractive indices along the orthogonal principle axes. They can be constructed such that the extraordinary axis coincides with the fast axis, while the ordinary axis is the slow axis. Photons travel faster along the fast axis than the slow axis since the index of refraction is smaller in that direction and so pick up different phases. Two types of wave plates, the half-wave (HWP) and the quarter-wave plate (QWP), are the most commonly used.

To describe the effect of the wave plate, we can use the Jones matrix, which performs a rotation of the photon's polarization. For a general wave plate (WP) oriented with its fast axis at an angle β to the coordinate system of the incident photon (say, $|H\rangle, |V\rangle$), the Jones matrix is [111, 112]

$$U = \begin{pmatrix} \cos(\frac{\alpha}{2}) + i\sin(\frac{\alpha}{2})\cos(2\beta) & i\sin(\frac{\alpha}{2})\sin(2\beta) \\ i\sin(\frac{\alpha}{2})\sin(2\beta) & \cos(\frac{\alpha}{2}) - i\sin(\frac{\alpha}{2})\cos(2\beta) \end{pmatrix}, \tag{4.33}$$

If the wave plate is oriented such that $\beta = 0$, then this reduces to

$$U = \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix}, \tag{4.34}$$

where the relative phase shift has $\alpha = \pi/2$ for a QWP and $\alpha = \pi$ for a HWP. Thus, the QWP introduces a complex relative phase and can convert linear to circular polarization and vice versa. On the other hand, the HWP adds a real relative phase (dropping the overall phase of i) and performs rotations of linear polarization or changes the handedness of circular polarization.

The Jones matrix for a HWP oriented at arbitrary angle β is (dropping the overall phase)

$$U_{\text{HWP}} = \begin{pmatrix} \cos(2\beta) & \sin(2\beta) \\ \sin(2\beta) & -\cos(2\beta). \end{pmatrix}. \tag{4.35}$$

Right away, we see that choosing $\beta = \pi/8$ yields the Hadamard gate. On the other hand, the Jones matrix for a QWP with $\beta = \pi/4$ to the $|H\rangle$ direction is

$$U_{\text{QWP}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = R_{\hat{x}}(-\pi/2), \tag{4.36}$$

which is a rotation by $-\pi/2$ around the $\hat{x}$ axis. This QWP transforms the linearly polarized states $|H\rangle$ and $|V\rangle$ into circularly polarized states,

$$\begin{aligned} U_{\text{QWP}}\,|H\rangle &= \frac{|H\rangle + i|V\rangle}{\sqrt{2}} = |L\rangle, \\ U_{\text{QWP}}\,|V\rangle &= \frac{|V\rangle + i|H\rangle}{\sqrt{2}} = i|R\rangle. \end{aligned}$$

Note that choosing $\beta = 0$ leaves the polarization linear (H, V), while $0 < \beta < \pi/4$ yields elliptical polarization.

To transform between the spatial and polarization encodings of an optical qubit, we can use a polarizing beam splitter (PBS). This device, which is similar to a standard beam splitter, splits the path of the photon depending on its polarization. For example, if the PBS is oriented in the $|H\rangle$, $|V\rangle$ basis then photons that are horizontally polarized are transmitted, while vertically polarized are reflected. The PBS can be oriented to distinguish between any two orthogonal polarization states, $|H\rangle$ and $|V\rangle$, or $|D\rangle$ and $|A\rangle$, or $|L\rangle$ and $|R\rangle$. In this way, we can measure the polarization of a photon in any basis. The PBS is also equivalent to the controlled-NOT gate, so that the state of the location qubit is flipped if the polarization qubit is in the "1" state [103].

A standard method for measuring photonic qubits is with a photodetector, which yields information about the number of photons in a given mode incident on the device. There are two types of detectors: photon-number detectors, which can detect the number of incident photons, and bucket detectors, which cannot discriminate between different photon numbers and can only detect the presence or absence of photons [113]. For a photon-number detector, the detection probability using Fock states was first derived in [114]. Given an input state ρ_i, the conditional probability to detect n photons is $p(n|i) = \mathrm{Tr}(|n\rangle\langle n|\,\rho_i) = \langle n|\rho_i|n\rangle$, where $|n\rangle\langle n|$ is a projector formed from the number states. However, for bucket detectors, we use the POVM description with elements $P_{\text{no click}} = |0\rangle\langle 0|$ and $P_{\text{click}} = \sum_{n=1}^{\infty} |n\rangle\langle n|$. The probability to record a click given input state ρ_i is $p(\text{click}|i) = \sum_{n=1}^{\infty}\langle n|\rho_i|n\rangle$ [113, 115].

To account for possible inefficiencies in a photodetector, we can introduce a model for an inefficient number-resolving device. Such a device can be characterized (neglecting dark counts, which occur when the detector responds even when there is no incident photon [113]) by $p_D(t)$, the probability to detect t photons [114, 115],

$$p_D(t) = \sum_{i=t}^{\infty} \binom{i}{t} \eta^t (1-\eta)^{i-t} p_S(i), \tag{4.37}$$

where η is the efficiency of the detector, and $p_S(i)$ is the probability the photon source produced i photons. The binomial coefficient accounts for the number of ways t atoms in the photodetector can be excited by incident photons, from i atoms. For a general source, the probability to detect a single photon $t = 1$ is

$$p_D(1) = \sum_{i=1}^{\infty} i\,\eta\,(1-\eta)^{i-1}\,p_S(i). \tag{4.38}$$

For instance, if the source produces only one photon, the probability to detect one photon is $p_D(1) = \eta$, and zero photons is $p_D(0) = 1 - \eta$. The conditional probability of detecting t photons given that i photons were actually present is

$$p_D(t|i) = \binom{i}{t} \eta^t (1-\eta)^{i-t}. \tag{4.39}$$

Unlike the destructive nature of photodetectors, it is possible to alternatively measure photonic qubits using optical devices. This is commonly encountered in the context of weak measurements. Such measurements do not disturb the quantum system as much as standard strong measurements since they involve small couplings between the quantum system and the ancilla. A few specific experiments using weak measurements are discussed here, while the details of the theoretical model are provided later in Ch. 5.

The first experimental implementation of weak measurements was performed by Ritchie, Story, and Hulet in 1991 using a birefringent crystal to measure the polarization of a photon [24]. Depending on the incident photon's polarization, the crystal separated the beam into two components with orthogonal polarizations. If the distance between the component beams is small (smaller than the waist of the Gaussian beam), then a weak measurement was performed. Here, the position of the photon was used as the pointer variable and was coupled to the polarization. The measurement Hamiltonian [116] can be written as

$$\hat{\mathcal{H}} = g\,\hat{Z} \otimes \hat{p}, \tag{4.40}$$

where $\hat{Z} = |H\rangle\langle H| - |V\rangle\langle V|$ is the Pauli matrix in the $|H\rangle, |V\rangle$ basis, and $\hat{p}$ is the momentum operator for the pointer. The interaction, with coupling strength parameterized by g, shifts the position x of the pointer "up" or "down" according to the eigenvalues (± 1) of the observable. In this way, a (weak) measurement of the polarization is made.

In another experiment by Lundeen et al., a weak measurement of the transverse position of a photon was performed by coupling it to its polarization (the pointer variable) using a half-wave plate [106]. If the rotation induced by the plate is small then the measurement is weak. Here, the measurement Hamiltonian can be written as [106, 117]

$$\mathcal{H} = \theta\,\hat{\pi}_x \otimes \hat{\sigma}_y, \tag{4.41}$$

where $\hat{\pi}_x = |x\rangle\langle x|$ is a projector formed from the position states. The Pauli matrix $\hat{\sigma}_y$ acts on the pointer variable and rotates it according to the angle θ. Polarization measurements have been made in other ways. For example, Hosten and Kwiat used the Spin Hall Effect of Light (SHEL) and a variable angle prism to measure different spin projections of a photon [118].

4.3 Summary

In this final introductory chapter, I gave an overview of the theory of quantum measurement, including the basic formalism of projective and POVM measurements. I ended by connecting the theory to physical realizations of qubits in an optical setting and how they are controlled and measured in typical experiments. I showed how the description of measurement that includes the entangling interaction between a quantum system and a meter is realized in weak measurements via the coupling of a photon's polarization to its spatial states (the meter) or vice versa. The next chapter will explore in more detail these types of measurements, where the focus will be on coupling strengths that are weak and strong, as well as on how information is processed in a quantum measurement.

Chapter 5

Weak and Strong Quantum Measurements

5.1 Introduction

In addition to the description of measurement outlined in the previous chapter, quantum measurements can also be weak or strong. Weak measurements are usually named for the weak coupling induced between the quantum system and the measuring device. As a result, the measurement reveals little information about the quantum system on average. In contrast, strong measurements (e.g., projective measurements) yield more information but greatly disturb the system.

When weak measurements and their associated weak values were first introduced by Aharonov, Albert, and Vaidman in 1988 [23], they were greeted with a degree of skepticism and doubts as to their practical importance. That one can obtain large expectation values outside the range of the measured observable's eigenvalues was viewed by some as merely an interesting theoretical puzzle. Nearly 30 years later, weak measurements are now understood to be a useful measurement technique with a wide range of applicability (see Ref. [116] for a review). They can be used, for example, to amplify small signals [24, 118], to directly measure quantum states [106, 119, 120], or to continuously monitor a quantum system with a sequence of many weak measurements, yielding a so-called quantum trajectory [121, 122].

We should keep in mind that there are different definitions of weak measurements in the literature (see, e.g., [123] for an introduction). In one definition, a POVM is constructed such that one measurement outcome is very likely compared to all of the others. In this way, the measurement almost always leaves the system in nearly the same state, thus yielding little information. There is a small probability that one of the unlikely outcomes will occur and, in this case, the measurement will greatly change the state of the system and reveal much more information.

In another definition, which will be the one used in the rest of this chapter, weak measurements are those where *all* possible outcomes lead to small changes in the system and only a little information is ever obtained. In this case, the measurement operators $\hat{M}_m$ [recall Eq. (4.2)] are equal to the identity matrix plus a small correction.

A concept that is closely associated with weak measurements is the weak value. It arises when a system that is prepared in a certain state (it is preselected) undergoes first a weak and then a strong measurement, with those measurement outcomes postselected. For a quantum system prepared in an initial state $|Q_i\rangle$ and postselected in a final state $|Q_f\rangle$, the weak value for the observable $\hat{A}$ is defined as [23]

$$\langle A\rangle_W = \frac{\langle Q_f|\hat{A}|Q_i\rangle}{\langle Q_f|Q_i\rangle}. \tag{5.1}$$

To understand the origin of this quantity [116], consider a quantum system that is prepared in the state $|Q_i\rangle$. A strong (projective) measurement yields the standard probability to detect an outcome corresponding to a final state $|Q_f\rangle$,

$$p = |\langle Q_f|Q_i\rangle|^2. \tag{5.2}$$

If, before this strong measurement, a weak measurement is made, then the probability (5.2) is changed to

$$p' = |\langle Q_f|\hat{U}_\epsilon|Q_i\rangle|^2, \tag{5.3}$$

where the unitary operator $\hat{U}_\epsilon = \exp(-i\epsilon\hat{A})$ is associated with the observable $\hat{A}$ of the weak measurement.

Since the interaction is weak, $\epsilon \ll 1$, the operator can be approximated using its Taylor expansion, $\hat{U}_\epsilon = \mathbb{1} - i\epsilon\hat{A} + \ldots$, and Eq. (5.3) becomes (to first order in ϵ)

$$p' = |\langle Q_f|Q_i\rangle|^2 + 2\epsilon\,\mathrm{Im}\,\langle Q_i|Q_f\rangle\langle Q_f|\hat{A}|Q_i\rangle. \tag{5.4}$$

Renormalizing this quantity (for $p \neq 0$), yields

$$\frac{p'}{p} = 1 + 2\epsilon\,\mathrm{Im}\,\langle A\rangle_W, \tag{5.5}$$

where $\langle A\rangle_W$ is the weak value in (5.1). Thus, a weak measurement of the preselected state alters the original measurement outcome probability, $p = |\langle Q_f|Q_i\rangle|^2$. The result (5.5) is linear in the weak value and is valid in the weak interaction regime. Higher order terms can be dropped when $p'/p - 1$ is sufficiently small and when the linear correction $\epsilon\,\mathrm{Im}\,\langle A\rangle_W$ is sufficiently larger than the contribution from all higher order terms [116, 124].

If the preselected and postselected states are the same, then (5.1) reduces to $\langle Q_i|\hat{A}|Q_i\rangle$, which is the usual expectation value. However, if the postselected state is chosen to be nearly orthogonal to the preselected state then the weak value can be very large, greater than the largest eigenvalue of the operator $\hat{A}$, since we divide by $\langle Q_f|Q_i\rangle \ll 1$. If the postselected state is *exactly* orthogonal to the preselected state, then the divergence of the weak value is countered by the vanishing probability of a successful postselection (that is, p in (5.2) is zero). Furthermore, the weak value can be complex since the numerator and denominator of the expression (5.1) are in general complex.

The meaning of the weak value becomes more clear if we consider the influence of the measurement on the pointer variable of the meter that is coupled to the quantum system. In the original weak measurement formalism [23], the meter is described by a Gaussian wave function for its position x, centered at $x = 0$ with some width σ. When the quantum system and meter become entangled in the measurement, both the position *and* momentum of the pointer variable are modified. The shift in the position, Δx, is related to the real part of the weak value, while the change in momentum corresponds to the imaginary part [125–127]. The measurement is considered weak when the width of the pointer's distribution is much larger than the subsequent shift in position, $\sigma \gg \Delta x$. As a result, there is uncertainty in the measurement outcomes. However, if the measurement is repeated on an ensemble of such systems, the average shift of the pointer (the weak value) can be accurately determined [126]. From another perspective, the weak value has been shown to be related to the amplitudes of the quantum wave function [106]. In this sense, the complex amplitudes of a quantum state can be directly read off of the weak value result.

Since weak values can be large and complex, there have been many interesting theoretical predictions [23] and paradoxical results in experimental settings. Some examples include the quantum three-box paradox, which was first described by [128,129] and later experimentally confirmed in [130], and Hardy's paradox [131]. Large weak values can be used to amplify signals that would otherwise be very small [23,24,118,132,133] and the weak measurement formalism was shown to be related to Vaidman's two-state formalism [125,128]. The Lundeen group used weak measurements to directly measure the quantum wave function of a photon [106] instead of the conventional yet indirect technique of quantum state tomography [134]. Although the initial implementations of weak measurements were in optical experiments, more recently, weak measurements have been made of solid state qubits (quantum dots) [135, 136] as well as superconducting (transmon) qubits [137, 138]. See Ref. [116] for an introduction to weak values and their experimental significance.

5.2 Extended Model of Weak and Strong Measurements

In the standard description of weak measurements and weak values, the interaction between the quantum system and ancilla is assumed to be weak. In this way, the expansion of the measurement interaction can be approximated by the unit operation plus a small correction that weakly couples the quantum system and ancilla. It was shown, however, that in some cases weak values can be computed exactly for interactions that are not weak (see, e.g., Ref. [139]) and that even strong measurements can directly probe the wave function in a unitary manner [117].

Here, I outline a general description of quantum measurement that accounts for interactions of arbitrary strength, and different choices of measured observables. I argue that weak and strong measurements should be treated on the same footing as it allows us to understand measurements in a more general sense. In particular, I derive new results for the shift in the expectation values of a set of noncommuting observables for the pointer variable, which yields the real and imaginary components of the weak value. I show that for two-dimensional pointers (qubits), the imaginary part of the weak value is proportional to the expectation value of an observable along an axis perpendicular to the plane of initialization, while the real part depends on the expectation values in the plane. The results hold for any postselection of the quantum system as well as for arbitrary initializations of the pointer. I focus on qubit pointers, but the generalization to higher dimensional qudits is straightforward.

Although I discuss measurements in the following sections in the context of spin-1/2 systems, the formalism is general and can be used for any effective two-level system. In particular, I first consider the case of spin measurements using a qubit pointer, and later consider position measurements. Such measurements have been implemented experimentally in many different settings. Recently, for example, weak measurements of a transmon qubit have been made using another transmon as the ancilla [137]. For position measurements, recall the discussion of Ref. [106] in Ch. 4.2.1, in which a weak measurement of a photon's position is made by coupling its polarization (the pointer) to its position using a wave plate.

5.2.1 Spin Measurements with a Qubit Pointer

5.2.1.1 The Measurement Process

As we have seen many times already, to measure an observable of a quantum system, we couple the system to an ancillary system (the pointer variable of a meter) according to the unitary von Neumann interaction. The time-evolution operator for the measurement Hamiltonian, $\mathcal{H}$, is given by $U = \exp(-i\mathcal{H}t/\hbar)$, where t indicates the length of the measurement interaction.

The interaction Hamiltonian, $\mathcal{H}$, for a qubit quantum system Q and pointer M corresponding to a measurement of the spin along an arbitrary axis, $\hat{n}$, is

$$\mathcal{H} = \theta\,(\vec{s}\cdot\hat{n}) \otimes (\vec{s}\cdot\hat{m}), \tag{5.6}$$

with interaction strength given by θ. The observable we will measure is $\vec{s}\cdot\hat{n} = s_x n_x + s_y n_y +$

$s_z n_z$ where $s_j = \sigma_j/2$ is the spin operator. The operator acting on the pointer, $(\vec{\sigma}\cdot\hat{m})/2$, will lead to a rotation of its initial state about the axis $\hat{m}$ by an angle $\theta/2$. To see this, we write the unitary operator $U = e^{-i\mathcal{H}t/\hbar}$ (setting $t = 1$ and $\hbar = 1$) using the series expansion for the exponential,

$$U = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{-i\,\theta\,(\vec{\sigma}\cdot\hat{n})\otimes(\vec{\sigma}\cdot\hat{m})}{4}\right)^{k}. \tag{5.7}$$

From the property of Pauli matrices, $(\vec{\sigma}\cdot\hat{n})^2 = \mathbb{1}$, we can evaluate the series exactly to find

$$U = \cos(\theta/4)\,\mathbb{1}\otimes\mathbb{1} - i\,\sin(\theta/4)\,(\vec{\sigma}\cdot\hat{n})\otimes(\vec{\sigma}\cdot\hat{m}). \tag{5.8}$$

When the interaction strength vanishes, $\theta = 0$, the operator (5.8) reduces to $\mathbb{1}\otimes\mathbb{1}$, so that no entanglement is created between Q and M. To first order in θ, the interaction for a weak measurement is

$$U_W = \mathbb{1}\otimes\mathbb{1} - i\,(\theta/4)\,(\vec{\sigma}\cdot\hat{n})\otimes(\vec{\sigma}\cdot\hat{m}). \tag{5.9}$$

The interaction (5.8) can be rewritten in a slightly more useful form. First, note that the Pauli matrices can be expressed in terms of projection operators along the axis $\hat{n}$ as

$$\begin{aligned} P_0^{(n)} &= \frac{1}{2}\big(\mathbb{1} + \vec{\sigma}\cdot\hat{n}\big) = |0_n\rangle\langle 0_n|, \\ P_1^{(n)} &= \frac{1}{2}\big(\mathbb{1} - \vec{\sigma}\cdot\hat{n}\big) = |1_n\rangle\langle 1_n|, \end{aligned} \tag{5.10}$$

where $|0_n\rangle$ and $|1_n\rangle$ are the eigenstates of $\vec{\sigma}\cdot\hat{n}$. Equation (5.8) can then be expressed in terms of projections on the quantum system as

$$U = P_0^{(n)}\otimes R_{\hat{m}}(\theta/2) + P_1^{(n)}\otimes R_{\hat{m}}(-\theta/2), \tag{5.11}$$

where the operators acting on the pointer are simply rotation operators around the axis $\hat{m}$ by an angle $\theta/2$,

$$R_{\hat{m}}(\pm\theta/2) = \cos(\theta/4)\,\mathbb{1} \mp i\sin(\theta/4)\,(\vec{\sigma}\cdot\hat{m}). \tag{5.12}$$

In this form, it is clear that the measurement rotates the pointer by an angle $\pm\theta/2$ depending on the state of Q.

In a strong measurement, the final pointer states are orthogonal so that they can be reliably distinguished and the quantum state is strongly disturbed from its initial state. To ensure this orthogonality requirement, the pointer variable must be initialized in a state $|M_i\rangle$ that is an eigenstate of an operator $\vec{\sigma}\cdot\hat{m}'$ with a basis that is orthogonal to the pointer's rotation axis $\hat{m}$: $\hat{m}\cdot\hat{m}' = 0$. In other words, the pointer is initialized in the plane perpendicular to its rotation axis. For instance, for a pointer that is prepared in the state $|M_i\rangle = |0_z\rangle$ with respect to the $\hat{z}$ axis, the rotation must be around an axis in the xy plane. In general, the overlap between the two final pointer states is

$$\langle M_i|R_{\hat{m}}(-\theta/2)R_{\hat{m}}(\theta/2)|M_i\rangle = \cos(\theta/2) + i\sin(\theta/2)\ \langle M_i|\vec{\sigma}\cdot\hat{m}|M_i\rangle, \tag{5.13}$$

which does not necessarily vanish when $\theta = \pi$. Writing the initial pointer state in the eigenbasis $|0_m\rangle$, $|1_m\rangle$ of the operator $\vec{\sigma}\cdot\hat{m}$ as $|M_i\rangle = a|0_m\rangle + b|1_m\rangle$ with complex amplitudes a and b, the matrix element in (5.13) evaluates to $\langle M_i|\vec{\sigma}\cdot\hat{m}|M_i\rangle = 2|a|^2 - 1$. Thus, the final pointer states are orthogonal at $\theta = \pi$ only if $|a| = 1/\sqrt{2}$ since at this angle (5.13) is $\langle M_i|R_{\hat{m}}(-\pi/2)R_{\hat{m}}(\pi/2)|M_i\rangle = i(2|a|^2 - 1)$.

To see how strongly the quantum system is disturbed by the measurement, we compute its density matrix after the coupling to the pointer. The initial state of the quantum system, written in the eigenbasis of the measured observable, is

$$|Q_i\rangle = \alpha|0_n\rangle + \beta|1_n\rangle. \tag{5.14}$$

Applying (5.11) to the initial system $|Q_i\rangle|M_i\rangle$ and tracing out the pointer variable yields the final density matrix for Q,

$$\rho'_Q = \cos^2(\theta/4)\,|Q_i\rangle\langle Q_i| + \sin^2(\theta/4)\,(\vec{\sigma}\cdot\hat{n})|Q_i\rangle\langle Q_i|(\vec{\sigma}\cdot\hat{n}). \tag{5.15}$$

Here, we used the fact that $\mathrm{Tr}[|M_i\rangle\langle M_i|\vec{\sigma}\cdot\hat{m}] = 0$. Writing this in the eigenbasis $|0_n\rangle$, $|1_n\rangle$ of the operator $\vec{\sigma}\cdot\hat{n}$ yields

$$\rho'_Q = \begin{pmatrix} |\alpha|^2 & \alpha^*\beta\,\cos(\theta/2) \\ \alpha\beta^*\cos(\theta/2) & |\beta|^2 \end{pmatrix}. \tag{5.16}$$

From this expression, it is clear that a strong measurement corresponds to a state ρ'_Q that is maximally disturbed from its initial state (it is entirely diagonal) and the final pointer states are orthogonal. In a weak measurement, ρ'_Q is nearly unchanged from its initial pure state, $\rho'_Q \approx |Q_i\rangle\langle Q_i|$, while the final pointer states of the meter are almost parallel according to (5.13) and cannot be perfectly distinguished. Therefore, there is a trade-off between the amount of information the meter has about Q and the resulting loss of coherence to the quantum state.

For example, suppose the pointer is initialized in the state $|M_i\rangle = |0_z\rangle$ and the rotation axis is $\hat{m} = \hat{y}$. Writing Q in the eigenbasis of the observable $\vec{\sigma}\cdot\hat{n}$ as in (5.14), the effect of (5.11) when the interaction is strong is

$$U\,|Q_i\rangle\,|0_z\rangle = \alpha\,|0_n\rangle\,|0_x\rangle + \beta\,|1_n\rangle\,|1_x\rangle, \tag{5.17}$$

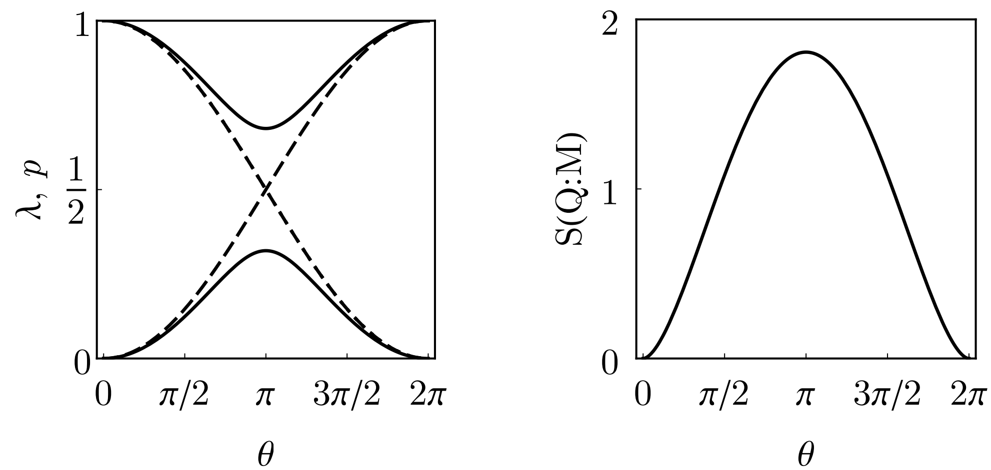

Figure 5.1: Left: The eigenvalues $\lambda_\pm$ (solid) and diagonal elements p_0, p_1 (dashed) of ρ_M, the pointer density matrix. For a strong measurement, $\theta = \pi$, the eigenvalues exactly correspond to the square of the amplitudes $|\alpha|^2$ and $|\beta|^2$ of the initial quantum state $|Q_i\rangle$. Right: The shared entropy between Q and M. The pointer has the most information about the state of Q at $\theta = \pi$ where the mutual entropy is maximum.

where $|0_x\rangle$ and $|1_x\rangle$ are the eigenstates of σ_x. For a general angle θ we find instead

$$\begin{aligned} U\,|Q_i\rangle\,|0_z\rangle = \alpha\,|0_n\rangle \Big(\cos(\theta/4)\,|0_z\rangle + \sin(\theta/4)\,|1_z\rangle\Big) \\ + \beta\,|1_n\rangle \Big(\cos(\theta/4)\,|0_z\rangle - \sin(\theta/4)\,|1_z\rangle\Big). \end{aligned} \tag{5.18}$$

Tracing out the quantum system, the diagonal elements and eigenvalues of the pointer density matrix ρ_M are

$$p_0 = \cos^2(\theta/4)\,, \quad p_1 = \sin^2(\theta/4), \tag{5.19}$$

and

$$\lambda_\pm = \frac{1}{2}\left(1 \pm \sqrt{1 - 4|\alpha|^2\,|\beta|^2 \sin^2(\theta/2)}\right), \tag{5.20}$$

which are shown in Fig. 5.1. At $\theta = \pi$, the eigenvalues reduce to $|\alpha|^2$, $|\beta|^2$, which are the squares of the amplitudes in the initial quantum state (5.14). When $\theta \ll 1$, the eigenvalues are approximately $|\alpha|^2|\beta|^2 \sin^2(\theta/2)$ and $1-|\alpha|^2|\beta|^2 \sin^2(\theta/2)$. Also in Fig. 5.1 is the shared entropy between Q and M, $S(Q:M) = S(Q)+S(M)-S(QM) = 2S(M)$, which is maximum at $\theta = \pi$ and goes to zero as $\theta \to 0$, demonstrating that weak measurements extract less information about the quantum state.

5.2.1.2 Weak Values

In the original weak measurement formalism [23], the quantum system is projected onto a final state (postselected) after the sequence of preparation and weak measurement. In this way, the weak value can be determined. If the initial wave function of Q is not known, then the weak value can be used to compute the coefficients in $|Q_i\rangle$ [106]. In this section we compute the weak value for a general initialization of the pointer and arbitrary interaction strength.

Supposing Q is prepared (preselected) in the initial state $|Q_i\rangle$ and the pointer is initialized in the state $|M_i\rangle$, after the measurement interaction (5.8) the joint state is

$$U\,|Q_i\rangle|M_i\rangle = \cos(\theta/4)\,|Q_i\rangle|M_i\rangle - i\,\sin(\theta/4)\,(\vec{\sigma}\cdot\hat{n})|Q_i\rangle\,(\vec{\sigma}\cdot\hat{m})|M_i\rangle. \tag{5.21}$$

If the quantum state is postselected to be in the state $|Q_f\rangle$, then the final (conditional) state of the meter is

$$\langle Q_f|U|Q_i\rangle\,|M_i\rangle = \cos(\theta/4)\,\langle Q_f|Q_i\rangle\,|M_i\rangle - i\,\sin(\theta/4)\,\langle Q_f|\vec{\sigma}\cdot\hat{n}|Q_i\rangle\,(\vec{\sigma}\cdot\hat{m})|M_i\rangle. \tag{5.22}$$

Renormalizing this expression yields the final state $|M_f\rangle$ of the meter,

$$|M_f\rangle = |M_i\rangle - i\,\tan(\theta/4)\,\frac{\langle Q_f|\vec{\sigma}\cdot\hat{n}|Q_i\rangle}{\langle Q_f|Q_i\rangle}\,(\vec{\sigma}\cdot\hat{m})|M_i\rangle. \tag{5.23}$$

The weak value of the observable $\hat{A} = \vec{\sigma}/2 \cdot \hat{n}$ is defined as the matrix element

$$\langle A \rangle_W = \frac{\langle Q_f | \hat{A} | Q_i \rangle}{\langle Q_f | Q_i \rangle} = \frac{1}{2} \frac{\langle Q_f | \vec{\sigma} \cdot \hat{n} | Q_i \rangle}{\langle Q_f | Q_i \rangle}. \tag{5.24}$$

Choosing $|Q_f\rangle$ nearly orthogonal to $|Q_i\rangle$ can lead to weak values that are much larger than the largest eigenvalue of $\hat{A}$.

To determine the weak value, the shift in expectation values for a set of complementary observables are computed in the final state of the meter (5.23). Together, these yield the real and imaginary parts of the weak value. By straightforward calculation, the expectation value of the observable $\vec{\sigma} \cdot \hat{m}'$ in the final state (5.23) is

$$\begin{aligned} \langle \vec{\sigma} \cdot \hat{m}' \rangle_f = &\left(1 - 4\tan^2(\theta/4)\, |\langle A \rangle_W|^2\right) \langle \vec{\sigma} \cdot \hat{m}' \rangle_i \\ &- 4\tan(\theta/4)\ \mathrm{Re}\big[\langle A \rangle_W\big]\, \langle (\hat{m} \times \hat{m}') \cdot \vec{\sigma} \rangle_i \\ &+ 4\tan(\theta/4)\ \mathrm{Im}\big[\langle A \rangle_W\big]\, (\hat{m} \cdot \hat{m}'), \end{aligned} \tag{5.25}$$

where $\langle \hat{O} \rangle_f \equiv \langle M_f | \hat{O} | M_f \rangle$ and $\langle \hat{O} \rangle_i \equiv \langle M_i | \hat{O} | M_i \rangle$ are the final and initial expectation values of the operator $\hat{O}$ in the pointer states. The first two terms in (5.25) are contributions from the expectation value in the plane of initialization, while the last term shifts the pointer out of the plane. Setting $\hat{m}' = \hat{x}$, $\hat{y}$, $\hat{z}$, for example, in the above formula will yield a set of three equations that can be solved for the weak value.

The choice of the rotation axis $\hat{m}$ is arbitrary and simply sets the reference frame for the pointer. For example, if we choose $\hat{m} = \hat{y}$, then the pointer must be initialized in the xz plane. In this case, we consider the set of expectation values for $\hat{m}' = \hat{x}$, $\hat{y}$, $\hat{z}$. The first two terms of (5.25) contribute to $\langle \sigma_x \rangle_f$ and $\langle \sigma_z \rangle_f$ via $\langle \sigma_x \rangle_i$ and $\langle \sigma_z \rangle_i$, while the third term shifts the pointer out of plane by producing a nonzero $\langle \sigma_y \rangle_f$. Specifically, the expectation values from (5.25) for σ_x, σ_y, and σ_z are:

$$\langle \sigma_z \rangle_f = \Big(1 - 4\tan^2(\theta/4)\,|\langle A \rangle_W|^2\Big)\,\langle \sigma_z \rangle_i - 4\tan(\theta/4)\ \mathrm{Re}\big[\langle A \rangle_W\big]\,\langle \sigma_x \rangle_i, \tag{5.26}$$

$$\langle \sigma_x \rangle_f = \Big(1 - 4\tan^2(\theta/4)\,|\langle A \rangle_W|^2\Big)\ \langle \sigma_x \rangle_i + 4\tan(\theta/4)\,\mathrm{Re}\big[\langle A \rangle_W\big]\,\langle \sigma_z \rangle_i, \tag{5.27}$$

and

$$\langle \sigma_y \rangle_f = 4\tan(\theta/4)\,\mathrm{Im}\big[\langle A \rangle_W\big]. \tag{5.28}$$

Here, we used the orthogonality condition for the final pointer states, which requires that $\langle \sigma_y \rangle_i = 0$ since $|M_i\rangle$ must be in the xz plane. Thus, $\langle \sigma_y \rangle_f$ is proportional to just the imaginary part of the weak value.

We can rewrite the expectation values (5.26), (5.27), and (5.28) in terms of the probabilities $\mathcal{P}$ to observe polarization measurement outcomes. If

$$\mathcal{P}_0^{(m')} = |\langle 0_{m'}|M_f\rangle|^2 = \langle M_f|P_0^{(m')}|M_f\rangle, \tag{5.29}$$

$$\mathcal{P}_1^{(m')} = |\langle 1_{m'}|M_f\rangle|^2 = \langle M_f|P_1^{(m')}|M_f\rangle, \tag{5.30}$$

are the probabilities to observe the outcomes $|0_{m'}\rangle$ and $|1_{m'}\rangle$ in a measurement of polarization along the $\hat{m}'$ axis, then

$$\langle \vec{\sigma} \cdot \hat{m}' \rangle_f = \mathcal{P}_0^{(m')} - \mathcal{P}_1^{(m')}. \tag{5.31}$$

With the pointer's initialization left arbitrary in the xz plane (the rotation axis $\hat{m}$ is still $\hat{y}$), the real and imaginary components of the weak value can be written in terms of measurement probabilities using (5.26), (5.27), and (5.28) as

$$\begin{aligned}\mathrm{Re}[\langle A\rangle_W] &= \frac{\langle\sigma_z\rangle_i\,\langle\sigma_x\rangle_f - \langle\sigma_x\rangle_i\,\langle\sigma_z\rangle_f}{4\,\tan(\theta/4)} \\ &= \frac{\langle\sigma_z\rangle_i\Big(\mathcal{P}_0^{(x)} - \mathcal{P}_1^{(x)}\Big) - \langle\sigma_x\rangle_i\Big(\mathcal{P}_0^{(z)} - \mathcal{P}_1^{(z)}\Big)}{4\,\tan(\theta/4)}\,,\end{aligned} \tag{5.32}$$

and

$$\mathrm{Im}[\langle A\rangle_W] = \frac{\langle\sigma_y\rangle_f}{4\,\tan(\theta/4)} = \frac{\mathcal{P}_0^{(y)} - \mathcal{P}_1^{(y)}}{4\,\tan(\theta/4)}\,. \tag{5.33}$$

To arrive at these results, we used the fact that

$$\langle\sigma_z\rangle_i^2 + \langle\sigma_x\rangle_i^2 = 1, \tag{5.34}$$

since the pointer is initialized in the $\hat{x}\hat{z}$ plane. That is, $|M_i\rangle = \sin\alpha|0_z\rangle + \cos\alpha|1_z\rangle$ yields $\langle\sigma_z\rangle_i = -\cos(2\alpha)$ and $\langle\sigma_x\rangle_i = \sin(2\alpha)$. Therefore, the real part of the weak value has contributions from both x and z polarization measurement probabilities, while the imaginary part depends only on the y measurement probabilities. In addition, in this form, the weak value is directly obtained from the statistics in the measurement (counts from a photodetector, for example).

As a specific example, let's choose the pointer to be initialized in the state $|M_i\rangle = |0_z\rangle$ with respect to the z axis. In this case, $\langle\sigma_x\rangle_i = 0$ and $\langle\sigma_z\rangle_i = 1$. Then, Eqs. (5.26), (5.27), and (5.28) evaluate to

$$\langle\sigma_z\rangle_f = 1 - 4\tan^2(\theta/4)\,|\langle A\rangle_W|^2, \tag{5.35}$$

$$\langle\sigma_x\rangle_f = 4\tan(\theta/4)\,\mathrm{Re}\big[\langle A\rangle_W\big], \tag{5.36}$$

$$\langle\sigma_y\rangle_f = 4\tan(\theta/4)\,\mathrm{Im}\big[\langle A\rangle_W\big]. \tag{5.37}$$

We see that the expectation value of σ_x is proportional to the real part of the weak value, while the expectation value of σ_y is proportional to the imaginary part,

$$\begin{aligned} \langle A \rangle_W &= \frac{\langle \sigma_x \rangle_f + i \langle \sigma_y \rangle_f}{4 \tan(\theta/4)} \\ &= \frac{\left(\mathcal{P}_0^{(x)} - \mathcal{P}_1^{(x)}\right) + i\left(\mathcal{P}_0^{(y)} - \mathcal{P}_1^{(y)}\right)}{4 \tan(\theta/4)}. \end{aligned} \tag{5.38}$$

Therefore, the probabilities $\mathcal{P}$ from the experiment can be used to obtain the weak value.

5.2.1.3 Computing the Wave Function

Here we review how weak measurements can be used to directly compute the quantum wave function rather than employing standard state tomography techniques. This was first performed experimentally for pure states by Lundeen et al. in Ref. [106] and quickly was followed by a generalization to mixed states [119].

If the probabilities $\mathcal{P}$ from the previous sections are determined in an experiment, then (5.38) can be used to compute the weak value. For a general initialization of the pointer, the weak value is instead given by Eqs. (5.32) and (5.33), from which the coefficients in the wave function $|Q_i\rangle$ can be directly determined. From (5.24),

$$\langle A \rangle_W = \frac{\langle Q_f | P_0^{(n)} | Q_i \rangle}{\langle Q_f | Q_i \rangle} - \frac{1}{2} = \frac{\langle Q_f | 0_n \rangle \langle 0_n | Q_i \rangle}{\langle Q_f | Q_i \rangle} - \frac{1}{2}. \tag{5.39}$$

Since $|Q_f\rangle$ can be selected in the experiment, the terms $\langle Q_f | 0_n \rangle$ and $\langle Q_f | 1_n \rangle$ are known.

The initial quantum system, written in the eigenbasis of the observable, is

$$|Q_i\rangle = \alpha |0_n\rangle + \beta |1_n\rangle. \tag{5.40}$$

Thus, the coefficient α is computed from

$$\langle A\rangle_W = \alpha\,\frac{\langle Q_f|0_n\rangle}{\langle Q_f|Q_i\rangle} - \frac{1}{2}. \tag{5.41}$$

Alternatively, with the second coefficient β,

$$\langle A\rangle_W = \frac{1}{2} - \beta\,\frac{\langle Q_f|1_n\rangle}{\langle Q_f|Q_i\rangle}. \tag{5.42}$$

The overlap $\langle Q_f|Q_i\rangle$ can be determined by normalizing the wave function,

$$|Q_i\rangle = \frac{1}{2}\langle Q_f|Q_i\rangle\left(\frac{1+2\langle A\rangle_W}{\langle Q_f|0_n\rangle}\;|0_n\rangle + \frac{1-2\langle A\rangle_W}{\langle Q_f|1_n\rangle}\;|1_n\rangle\right). \tag{5.43}$$

From this expression, it is clear that the complex amplitudes of the quantum wave function can be directly seen from the measurement outcomes via (5.38).

5.2.1.4 Weak Measurements without Postselection

We can also consider the situation where the quantum state is not postselected after the weak measurement. In this case, we compute the standard expectation value of an observable in the initial quantum state from the pointer. The density matrix for the pointer is found by tracing out Q from the final joint state (5.21),

$$\begin{aligned}\rho_M = {}&\cos^2(\theta/4)\,|M_i\rangle\langle M_i| + \sin^2(\theta/4)\,\vec{\sigma}\cdot\hat{m}|M_i\rangle\langle M_i|\vec{\sigma}\cdot\hat{m}\\ &+\frac{i}{2}\sin(\theta/2)\,\langle\vec{\sigma}\cdot\hat{n}\rangle_Q\Big[|M_i\rangle\langle M_i|\vec{\sigma}\cdot\hat{m} - \vec{\sigma}\cdot\hat{m}|M_i\rangle\langle M_i|\Big],\end{aligned} \tag{5.44}$$

where

$$\langle\vec{\sigma}\cdot\hat{n}\rangle_Q = \mathrm{Tr}_Q\Big[\vec{\sigma}\cdot\hat{n}\,|Q_i\rangle\langle Q_i|\Big] \tag{5.45}$$

is the expectation value of $\vec{\sigma} \cdot \hat{n}$ in the initial quantum state. Similarly to the previous sections, we compute the expectation value of $\vec{\sigma} \cdot \hat{m}'$ in the final state of M. In contrast to Eq. (5.25), we now find

$$\begin{aligned} \langle \vec{\sigma} \cdot \hat{m}' \rangle_f &= \mathrm{Tr}_M \Big[(\vec{\sigma} \cdot \hat{m}') \, \rho_M \Big] \\ &= \cos(\theta/2) \, \langle \vec{\sigma} \cdot \hat{m}' \rangle_i + \sin(\theta/2) \, \langle \vec{\sigma} \cdot \hat{n} \rangle_Q \, \langle \hat{m}' \times \hat{m} \cdot \vec{\sigma} \rangle_i, \end{aligned} \tag{5.46}$$

which does not have a term proportional to $\hat{m}' \cdot \hat{m}$. That is, there are no imaginary contributions to the expectation value (as expected) and $\langle \sigma_y \rangle_f = 0$. In addition,

$$\langle \sigma_x \rangle_f = \cos(\theta/2) \, \langle \sigma_x \rangle_i + \sin(\theta/2) \, \langle \vec{\sigma} \cdot \hat{n} \rangle_Q \, \langle \sigma_z \rangle_i, \tag{5.47}$$

$$\langle \sigma_z \rangle_f = \cos(\theta/2) \, \langle \sigma_z \rangle_i - \sin(\theta/2) \, \langle \vec{\sigma} \cdot \hat{n} \rangle_Q \, \langle \sigma_x \rangle_i. \tag{5.48}$$

When the interaction is weak, these expectation values are, to first order in θ, equivalent to those in Eqs. (5.26) and (5.27) obtained by postselecting Q:

$$\langle \sigma_x \rangle_f \sim \langle \sigma_x \rangle_i + \theta \, \langle \vec{s} \cdot \hat{n} \rangle_Q \, \langle \sigma_z \rangle_i, \tag{5.49}$$

$$\langle \sigma_z \rangle_f \sim \langle \sigma_z \rangle_i - \theta \, \langle \vec{s} \cdot \hat{n} \rangle_Q \, \langle \sigma_x \rangle_i. \tag{5.50}$$

When $|Q_f\rangle = |Q_i\rangle$, the weak value is equal to the expectation value $\langle \vec{s} \cdot \hat{n} \rangle_Q$, which has no imaginary contributions since $\vec{s} \cdot \hat{n}$ is Hermitian.

5.2.2 Position Measurements with a Qubit Pointer

We now derive the weak value for measurements of an observable that takes the form of a projector. For instance, we could measure the position with $|x\rangle\langle x|$, or momentum with $|p\rangle\langle p|$. Here, we suppose that we are measuring the position x of a quantum system using a qubit pointer. Since the dimension of the ancillary system is two, this scheme yields only information about whether the quantum system is at location x or is not. A higher-dimensional ancilla is needed to distinguish between more than two spatial states.

The interaction Hamiltonian for a position measurement is written as

$$\mathcal{H} = \theta\, \hat{\pi}_n \otimes (\vec{s} \cdot \hat{m}), \tag{5.51}$$

where $\hat{\pi}_n = |x_n\rangle\langle x_n|$ is a projector composed of one of n position eigenstates. The series expansion for the time-evolution operator, U, can be exactly evaluated to

$$U = \mathbb{1} \otimes \mathbb{1} + \hat{\pi}_n \otimes \Big(R_{\hat{m}}(\theta) - \mathbb{1}\Big). \tag{5.52}$$

Since the projectors add to unity, $\sum_n \hat{\pi}_n = \mathbb{1}$, this can be rewritten as

$$U = \sum_{n' \neq n} \hat{\pi}_{n'} \otimes \mathbb{1} + \hat{\pi}_n \otimes R_{\hat{m}}(\theta), \tag{5.53}$$

where it is clear that the pointer is only rotated by an angle θ about the m axis when the quantum system is found at location x_n. When $\theta = 0$, the interaction reduces to the identity. When $\theta = \pi$, the final pointer states $\mathbb{1}|M_i\rangle$ and $R_{\hat{m}}(\pi)|M_i\rangle$ are orthogonal, as long as the pointer variable is initialized in the plane perpendicular to its rotation axis, $\hat{m}$.

We now proceed as before. With the quantum system preselected in the state $|Q_i\rangle$ and the pointer initialized as $|M_i\rangle$, the measurement interaction (5.52) leads to

$$\begin{aligned} U|Q_i\rangle|M_i\rangle = |Q_i\rangle|M_i\rangle &+ \hat{\pi}_n|Q_i\rangle\big(\cos(\theta/2) - 1\big)|M_i\rangle \\ &- \hat{\pi}_n|Q_i\rangle\, i\sin(\theta/2)\, \vec{\sigma} \cdot \hat{m}|M_i\rangle \end{aligned} \tag{5.54}$$

Defining the weak value,

$$\langle A \rangle_W = \frac{\langle Q_f | \hat{\pi}_n | Q_i \rangle}{\langle Q_f | Q_i \rangle}, \tag{5.55}$$

the final pointer state after postselecting the quantum system in the state $\langle Q_f|$ and renormalizing is

$$|M_f\rangle = \Big[1 - 2\langle A \rangle_W \sin^2(\theta/4)\Big] \, |M_i\rangle - i\langle A \rangle_W \sin(\theta/2) \, \vec{\sigma} \cdot \hat{m} \, |M_i\rangle \tag{5.56}$$

The expectation value of the operator $\vec{\sigma} \cdot \hat{m}'$ in the final pointer state (5.56) is

$$\begin{aligned} \langle \vec{\sigma} \cdot \hat{m}' \rangle_f = & \Big(1 - \mathrm{Re}[\langle A \rangle_W] \, \epsilon_\theta - |\langle A \rangle_W|^2 \, \epsilon_\theta \, \cos(\theta/2)\Big) \, \langle \vec{\sigma} \cdot \hat{m}' \rangle_i \\ & - \sin(\theta/2) \left(2 \, \mathrm{Re}[\langle A \rangle_W] - |\langle A \rangle_W|^2 \, \epsilon_\theta\right) \langle (\hat{m} \times \hat{m}') \cdot \vec{\sigma} \rangle_i \\ & + \sin(\theta/2) \left(2 \, \mathrm{Im}[\langle A \rangle_W]\right) \, (\hat{m} \cdot \hat{m}'), \end{aligned} \tag{5.57}$$

where we defined

$$\epsilon_\theta = 4 \sin^2(\theta/4). \tag{5.58}$$

If we set the rotation axis of the pointer variable to be $\hat{m} = \hat{y}$, then it is initialized in the xz plane. The three equations obtained from (5.57) for $\hat{m}' = \hat{x}$, $\hat{y}$, $\hat{z}$ yield the real and imaginary parts of the weak value. The imaginary part is

$$\mathrm{Im}[\langle A \rangle_W] = \frac{\langle \sigma_y \rangle_f}{2 \, \sin(\theta/2)} = \frac{\mathcal{P}_0^{(y)} - \mathcal{P}_1^{(y)}}{2 \, \sin(\theta/2)}, \tag{5.59}$$

while real part is

$$
\begin{aligned}
\mathrm{Re}[\langle A\rangle_W] &= \frac{1}{2}\Big(1 - g_z(\theta)\,\langle\sigma_z\rangle_f - g_x(\theta)\,\langle\sigma_z\rangle_f\Big) \\
&= \frac{1}{2}\Big(1 - g_z(\theta)\,\Big(\mathcal{P}_0^z - \mathcal{P}_1^z\Big) - g_x(\theta)\Big(\mathcal{P}_0^x - \mathcal{P}_1^x\Big)\Big).
\end{aligned}
\tag{5.60}
$$

The θ dependence of $\mathrm{Re}[\langle A\rangle_W]$ is contained in the functions

$$
g_z(\theta) = \langle\sigma_z\rangle_i + \cot(\theta/2)\,\langle\sigma_x\rangle_i, \tag{5.61}
$$

$$
g_x(\theta) = \langle\sigma_x\rangle_i - \cot(\theta/2)\,\langle\sigma_z\rangle_i. \tag{5.62}
$$

Here, we used the fact that

$$
\langle\sigma_z\rangle_i^2 + \langle\sigma_x\rangle_i^2 = 1, \tag{5.63}
$$

since the pointer variable is initialized in the xz plane.

For a strong measurement, $\theta = \pi$, these functions reduce to $g_z(\theta) = \langle\sigma_z\rangle_i$ and $g_x(\theta) = \langle\sigma_x\rangle_i$. For a weak measurement, $g_z(\theta) \sim \frac{2}{\theta}\langle\sigma_x\rangle_i$ and $g_x(\theta) \sim -\frac{2}{\theta}\langle\sigma_z\rangle_i$. It is interesting to note that Eqs. (5.32) and (5.33) are identical to (5.60) and (5.59) when θ is small. In other words, the weak value is the same for spin measurements and position measurements when the interaction is weak.

If we initialize the pointer variable along $\hat{z}$ such that $\langle\sigma_z\rangle_i = 1$ and $\langle\sigma_x\rangle_i = 0$, the more general results derived in this section reduce to those of Vallone et al. in Ref. [117], and of Lundeen et al. in Ref. [106] for small θ. In [117], it was shown that strong measurements, as opposed to the weak approximation made in [106], yield a more accurate result in a direct measurement of the wave function. Our result, Eq. (5.57) with $\hat{m}' = \hat{x}$, $\hat{y}$, $\hat{z}$, yields

$$\begin{aligned}
\langle \sigma_z \rangle_f &= 1 - \mathrm{Re}[\langle A \rangle_W]\, \epsilon_\theta - |\langle A \rangle_W|^2\, \epsilon_\theta \cos(\theta/2), && (5.64)\\
\langle \sigma_x \rangle_f &= \sin(\theta/2) \left(2\, \mathrm{Re}[\langle A \rangle_W] - |\langle A \rangle_W|^2\, \epsilon_\theta \right), && (5.65)\\
\langle \sigma_y \rangle_f &= \sin(\theta/2) \left(2\, \mathrm{Im}[\langle A \rangle_W] \right). && (5.66)
\end{aligned}$$

Solving the first two equations yields the real part of the weak value

$$\begin{aligned}
\mathrm{Re}[\langle A \rangle_W] &= \frac{\mathcal{P}_0^{(x)} - \mathcal{P}_1^{(x)} + \tan(\theta/2) \left(1 - \mathcal{P}_0^{(z)} + \mathcal{P}_1^{(z)} \right)}{2\, \tan(\theta/2)} \\
&= \frac{\mathcal{P}_0^{(x)} - \mathcal{P}_1^{(x)} + 2 \tan(\theta/4)\, \mathcal{P}_1^{(z)}}{2\, \sin(\theta/2)}.
\end{aligned} \tag{5.67}$$

In the second equality we eliminated $\mathcal{P}_0^{(z)}$ by using

$$\langle \sigma_z \rangle_f = \langle M_f | M_f \rangle - 2\mathcal{P}_1^{(z)} = 1 - \epsilon_\theta \mathrm{Re}[\langle A \rangle_W] + \epsilon_\theta |\langle A \rangle_W|^2 - 2\mathcal{P}_1^{(z)}. \tag{5.68}$$

This normalization of the state $|M_f\rangle$ is valid only for the initialization along $\hat{z}$. For small θ the weak value is

$$\langle A \rangle_W = \frac{1}{\theta} \Big(\mathcal{P}_0^{(x)} - \mathcal{P}_1^{(x)} \Big) + \frac{i}{\theta} \Big(\mathcal{P}_0^{(y)} - \mathcal{P}_1^{(y)} \Big), \tag{5.69}$$

while for $\theta = \pi$,

$$\begin{aligned}
\langle A \rangle_W &= \frac{1}{2} \Big(1 - \mathcal{P}_0^{(z)} + \mathcal{P}_1^{(z)} \Big) + \frac{i}{2}\, \Big(\mathcal{P}_0^{(y)} - \mathcal{P}_1^{(y)} \Big) \\
&= \frac{1}{2} \Big(\mathcal{P}_0^{(x)} - \mathcal{P}_1^{(x)} + 2\mathcal{P}_1^{(z)} \Big) + \frac{i}{2}\, \Big(\mathcal{P}_0^{(y)} - \mathcal{P}_1^{(y)} \Big).
\end{aligned} \tag{5.70}$$

Therefore, the results derived here, Eqs. (5.59) and (5.60), are more general and account for arbitrary pointer initializations and interaction strengths.

5.2.3 Relationship Between Measurement Interactions

In this section we will see how different measurement interactions for two-level systems can be related through global rotations of the quantum system and/or pointer. If the observable $\hat{A}$ is a two-dimensional projector $\hat{P}$, then the corresponding measurement operator with $\hat{A} = \hat{P}$ is related to the measurement operator with $\hat{A} = \vec{\sigma} \cdot \hat{n}$. Specifically, writing the projector as $\hat{P} = \frac{1}{2}(\mathbb{1} + \vec{\sigma} \cdot \hat{n})$ for the Hamiltonian $\mathcal{H} = \theta \, \hat{P} \otimes \vec{\sigma} \cdot \hat{m}$, yields the measurement operator

$$U = \left(\mathbb{1} \otimes R_{\hat{m}}(\theta/2)\right) e^{-i\theta/4 \, \vec{\sigma}\cdot\hat{n} \otimes \vec{\sigma}\cdot\hat{m}}. \tag{5.71}$$

Thus, the operators with $\hat{A} = \hat{P}$ and $\hat{A} = \vec{\sigma} \cdot \hat{n}$ are related by an overall rotation of the pointer.

Similarly, if both the measured observable and pointer variable are written as two-dimensional projectors, then $\mathcal{H} = \theta \hat{P} \otimes \hat{Q}$. Writing the projectors as $\hat{P} = \frac{1}{2}(\mathbb{1} + \vec{\sigma} \cdot \hat{n})$ and $\hat{Q} = \frac{1}{2}(\mathbb{1} + \vec{\sigma} \cdot \hat{m})$,

$$U = e^{-i\theta/4} \left(R_{\hat{n}}(\theta/2) \otimes R_{\hat{m}}(\theta/2)\right) e^{-i\theta/4 \, \vec{\sigma}\cdot\hat{n} \otimes \vec{\sigma}\cdot\hat{m}}. \tag{5.72}$$

In this case, the operators with $\hat{A} = \hat{P}$, $\hat{M} = \hat{Q}$ and $\hat{A} = \vec{\sigma} \cdot \hat{n}$, $\hat{M} = \vec{\sigma} \cdot \hat{m}$ are related by an overall rotation of the quantum system and the pointer.

5.3 Quantum Information Theory of Strong Quantum Measurements

In the rest of this chapter, I will focus on projective (strong) measurements. I will expand upon the ideas introduced previously and will analyze how information is processed in a strong measurement from the perspective of quantum information theory. Using entropy Venn diagrams, I will track the distribution of entanglement and entropy in the measurement of a single quantum system and in parallel measurements made on an entangled system [59, 140, 141]. I will also distinguish between measurements made on prepared (pure) and unprepared (completely mixed) quantum states.

5.3.1 The Measurement Process

Suppose a given quantum system is in the initial state

$$|Q\rangle = \sum_{x_1=0}^{d-1} \alpha_{x_1}^{(1)} |\widetilde{x}_1\rangle, \tag{5.73}$$

where $\alpha_{x_1}^{(1)}$ are complex amplitudes. Here, Q is expressed in terms of the d orthonormal basis states $|\widetilde{x}_1\rangle$ associated with the observable that we will measure. The von Neumann measurement is implemented with a unitary operator that entangles the quantum system Q with an ancillary system (the pointer) A_1,

$$U_{QA_1} = \sum_{x_1=0}^{d-1} P_{x_1} \otimes U_{x_1}, \tag{5.74}$$

where $P_{x_1} = |\widetilde{x}_1\rangle\langle\widetilde{x}_1|$ are projectors on the state of Q. As we've seen many times already, the operators U_{x_1} transform the initial state $|M_i\rangle_{A_1}$ of the ancilla to the final state $U_{x_1}|M_i\rangle_{A_1} =$

$|x_1\rangle_{A_1}$. Since we are considering strong measurements, the states of the ancilla, $|x_1\rangle_{A_1}$, are orthogonal. From now on, we drop the subscripts A_1 on the final states of the ancilla. The unitary interaction (5.74) between the quantum system and the ancilla leads to the entangled state [59]

$$|QA_1\rangle = U_{QA_1}\,|Q\rangle\,|M_i\rangle_{A_1} = \sum_{x_1} \alpha^{(1)}_{x_1}\,|\widetilde{x}_1\rangle\,|x_1\rangle. \tag{5.75}$$

The coefficients $\alpha^{(1)}_{x_1}$ reflect the degree of entanglement between Q and A_1: the number of non-zero coefficients is the Schmidt number [25] of the Schmidt decomposition.

The measurement process thus described contains the essence of the no-cloning theorem [36], which states that it is impossible to perfectly copy a quantum state unless it is in an eigenstate of the measurement operator (see Sec. 2.4). In other words, if the ancilla A_1 had faithfully copied Q, then the final state would be the *separable* wave function $|QA_1\rangle = \sum_{x_1} \alpha^{(1)}_{x_1}\,|\widetilde{x}_1\rangle \otimes \sum_{x_1} \alpha^{(1)}_{x_1}\,|x_1\rangle$ instead of Eq. (5.75). Therefore, in general, the entanglement produced in quantum measurement prohibits the measurement device from making a perfect copy of the quantum system. This is very different from classical measurements, where the measurement device can, in principle, always perfectly reflect the state of the system.

Information about the measurement outcomes is obtained from the state of the ancilla. Tracing over (5.75), the marginal density matrix of A_1 (and similarly for Q) is

$$\rho(A_1) = \mathrm{Tr}_Q\left(|QA_1\rangle\langle QA_1|\right) = \sum_{x_1} |\alpha^{(1)}_{x_1}|^2\,|x_1\rangle\langle x_1|. \tag{5.76}$$

From the symmetry of the state (5.75), the marginal von Neumann entropy of A_1 is the same as Q, which, in turn, is equal to the Shannon entropy of the probability distribution

$q^{(1)}_{x_1} = |\alpha^{(1)}_{x_1}|^2$:

$$S(Q) = S(A_1) = H[q^{(1)}] = -\sum_{x_1} q^{(1)}_{x_1} \log_d q^{(1)}_{x_1}. \tag{5.77}$$

Recall that we denote the Shannon entropy of a d-dimensional probability distribution p_{x_i} by $H[p] = -\sum_{x_i=0}^{d-1} p_{x_i} \log_d p_{x_i}$. The von Neumann entropy of a density matrix $\rho(X)$ is defined as $S(X) = S(\rho(X)) = -\mathrm{Tr}\,[\rho(X)\log_d \rho(X)]$, which on account of the logarithm to the base d, gives entropies the units "dits".

The ancilla and quantum system are not classically correlated in (5.75) (as is required for decoherence models, e.g., [142]), but in fact are entangled. This entanglement is characterized by a negative conditional entropy [59, 143], $S(A_1|Q) = S(QA_1) - S(Q) = -S(A_1)$, where the joint entropy vanishes since (5.75) is pure. We illustrate the entanglement between A_1 and Q with an entropy Venn diagram [59] in Fig. 5.2(a). The mutual entropy at the center of the diagram, $S(Q : A_1) = S(Q) + S(A_1) - S(QA_1)$, reflects the entropy that is shared between both systems and is twice as large as the classical upper bound [59, 71, 143].

5.3.2 Unprepared Quantum States

In the previous section, we considered measurements of a quantum system that is prepared in a pure state (5.73) with coefficients $\alpha^{(1)}_{x_1}$. Suppose instead that we are given a quantum system about which we have no information, that is, where no previous measurement results could inform us of the state of Q. In this case, we write the quantum system's initial state as a maximum entropy mixed state

$$\rho(Q) = \frac{1}{d}\sum_{x_0=0}^{d-1} |\widetilde{x}_0\rangle\langle\widetilde{x}_0|\,, \tag{5.78}$$

with coefficients that now correspond to a uniform probability distribution. We call this an *unprepared* quantum system. We can "purify" $\rho(Q)$ by defining a higher-dimensional pure state where Q is entangled with a reference system R [25],

$$|QR\rangle = \frac{1}{\sqrt{d}} \sum_{x_0=0}^{d-1} |\widetilde{x}_0\rangle |x_0\rangle \,, \tag{5.79}$$

such that $\rho(Q)$ is recovered by tracing (5.79) over R. In this way, we can see explicitly how the total system evolves unitarily as a pure state. Here and earlier, the states of Q are written with a tilde, $|\widetilde{x}_0\rangle$, to distinguish them from the states of R, which are denoted by $|x_0\rangle$. In this section, we assume that Q is an unprepared (or "unknown") state with maximum entropy so that it is maximally entangled with R, as in (5.79). With such an assumption, we do not bias any subsequent measurements [144].

To measure Q with an ancilla A_1, we express the quantum system in the eigenbasis $|\widetilde{x}_1\rangle$ of the observable that ancilla A_1 will measure using the unitary matrix $U^{(1)}_{x_0 x_1} = \langle \widetilde{x}_1 | \widetilde{x}_0 \rangle$. The orthonormal basis states of the ancilla, $|x_1\rangle$, with $x_1 = 0, \ldots, d-1$, automatically serve as the "interpretation basis" [145]. We then entangle [59] Q with A_1, which, as before, is in the initial state $|M_i\rangle_{A_1}$, using the unitary entangling operation U_{QA_1} in Eq. (5.74),

$$\begin{aligned} |QRA_1\rangle &= \mathbb{1}_R \otimes U_{QA_1} \, |QR\rangle \, |M_i\rangle_{A_1} \\ &= \frac{1}{\sqrt{d}} \sum_{x_0 x_1} U^{(1)}_{x_0 x_1} \, |\widetilde{x}_1\rangle \, |x_0\rangle \, |x_1\rangle, \end{aligned} \tag{5.80}$$

where $\mathbb{1}_R$ is the identity operation on R. We always write the states on the right hand side in the same order as they appear in the ket on the left hand side. We express the reference in a new basis by defining $|x_1\rangle_R = \sum_{x_0} U^{(1)\intercal}_{x_1 x_0} \, |x_0\rangle_R$ with the transpose of $U^{(1)}$, so that the

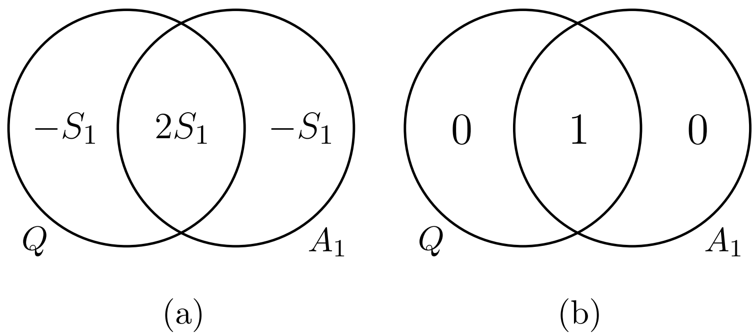

Figure 5.2: Entropy Venn diagrams [59] for the quantum system and ancilla. (a) For prepared quantum states, Q and A_1 are entangled according to Eq. (5.75). (b) For unprepared quantum states, Q and A_1 are correlated according to Eq. (5.82) when the reference R has been traced out. In this figure, we use the notation $S_1 = S(A_1)$ for the marginal entropy of ancilla A_1.

joint system QRA_1 appears as

$$|QRA_1\rangle = \frac{1}{\sqrt{d}} \sum_{x_1} |\widetilde{x}_1\rangle|x_1\rangle|x_1\rangle \ . \tag{5.81}$$

Note that (5.81) is a tripartite Schmidt decomposition of the joint density matrix $\rho(QRA_1) = |QRA_1\rangle\langle QRA_1|$, which is possible here because the entanglement operator U_{QA_1} ensures the bi-Schmidt basis ${}_R\langle x_1|QRA_1\rangle$ has Schmidt number one [74].

Tracing out the reference system from the full density matrix $\rho(QRA_1)$, we note that the ancilla is perfectly correlated with the quantum system,

$$\rho(QA_1) = \frac{1}{d} \sum_{x_1} |\widetilde{x}_1\, x_1\rangle\langle\widetilde{x}_1\, x_1|, \tag{5.82}$$

in contrast to Eq. (5.75) where A_1 and Q are entangled. Such correlations are indicated by a vanishing conditional entropy [59], $S(A_1|Q) = S(QA_1) - S(Q) = 0$. Tracing over (5.82), we find that each system has maximum entropy $S(Q) = S(A_1) = 1$. In Fig. 5.2, we compare the entropy Venn diagrams that are constructed from the states (5.75) and (5.82).

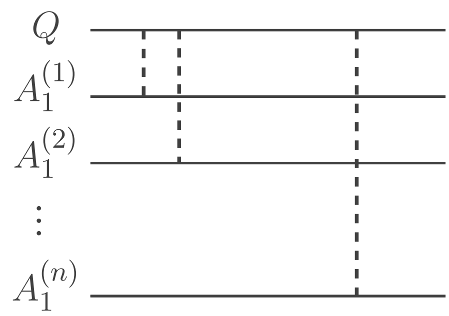

Figure 5.3: Composition of the quantum ancilla. Dashed lines indicate entanglement between the quantum system, Q, and each of the n qudits, $A_1^{(1)}, \ldots, A_1^{(n)}$, in the ancilla A_1. Time proceeds from left to right.

We note in passing that R can be thought of as representing all previous measurements of the quantum system that have occurred before A_1. We contrast measurements of unprepared quantum states (5.79) as described in this section, with measurements of *prepared* quantum states (see Sec. 5.3.1), which are initially pure states (5.73) defined without a reference system R.

5.3.3 Composition of the Quantum Ancilla

The ancilla A_1 may, in practice, be composed of many qudits $A_1^{(1)} \ldots A_1^{(n)}$ that all participate in the measurement of Q according to the sequence of entangling operations $U_{QA_1^{(n)}} \cdots U_{QA_1^{(1)}}$ between Q and $A_1^{(i)}$ (see Fig. 5.3). In this case, Eq. (5.75), for example, is extended to

$$|QA_1\rangle = \sum_{x_1} \alpha_{x_1}^{(1)} \, |\widetilde{x}_1\rangle \, |x_1\rangle_{A_1^{(1)}} \cdots |x_1\rangle_{A_1^{(n)}} \, . \tag{5.83}$$

Tracing out the quantum system from Eq. (5.83), the joint state of the entire ancilla is $\rho(A_1) = \rho(A_1^{(1)} \ldots A_1^{(n)}) = \sum_{x_1} |\alpha_{x_1}^{(1)}|^2 |x_1 \ldots x_1\rangle\langle x_1 \ldots x_1|$. That is, each component of A_1 is perfectly correlated with every other component, so that A_1 is internally self-consistent ("all parts of A_1 tell the same story"). However, while A_1 appears classical, and could conceivably consist of a macroscopic number of components, it is potentially *fragile*, in the sense that its entanglement with other devices may become hidden when any part $A_1^{(i)}$ of A_1 is lost (traced over). In the following chapters, we will distinguish "amplifiable" from non-amplifiable devices. That is, a state is amplifiable if tracing over any of its components does not modify the correlations between its subsystems.

To do this, we will consider in our discussion of Markovian quantum measurements in Ch. 7, an additional step to the measurement process by introducing a macroscopic detector D_1 that is used to measure the quantum ancilla A_1. In other words, D_1 observes the quantum observer A_1. This second system, which is also composed of many qudits, amplifies the measurement with A_1, by recording the outcome on a macroscopic device. While A_1 may be fragile depending on the situation, D_1 is robust: any part of D_1 could be traced over without altering its correlations with other macroscopic measurement devices. While such a procedure (a quantum system observed by a quantum ancilla, which is observed by a classical device) may appear arbitrary, it represents a convenient way of splitting up the second stage of von Neumann's measurement [22] to better keep track of the fate of entanglement. This gives rise to two classes of quantum measurement: those performed within a closed system where every part of the measurement device remains under control (unamplified measurements), and those performed within an open system, where some degrees of freedom are traced over (amplified measurements). In fact, a similar construction is used in the weak measurement formalism discussed earlier in this chapter (a quantum system is weakly coupled to a pointer variable, followed by a strong measurement of the pointer) as well as in other measurement settings (e.g., superconducting qubits can be measured by first coupling to another superconducting qubit, the ancilla, whose state is later measured and read out).

In Ch. 7, I will formally define the concept of a quantum Markov chain that is used

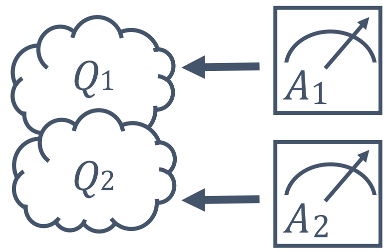

Figure 5.4: Illustration of parallel quantum measurements. Two different measurement devices, A_1 and A_2, are used to separately measure the components, Q_1 and Q_2, of an entangled quantum system.

in this work, in the context of consecutive measurements of a quantum system. I also further develop the formalism to describe unamplified measurements with quantum ancillae, A_i, which I will show are non-Markovian, and amplified measurements with macroscopic detectors, D_i, which are Markovian.

5.3.4 Parallel Quantum Measurements

In the previous section, we saw how to perform a single measurement of a quantum system. We can extend this description by considering measurements made in parallel on composite systems (see Fig. 5.4). A common realization of this idea is the set of (in)compatible measurements made on Bell states to test for violations of Bell inequalities [146, 147]. I will describe these measurements in some detail below since the results will be helpful for understanding the calculations in Ch. 6 on the Bell-state quantum eraser.

Consider the Bell state for two qubits Q_1 and Q_2, written in the eigenbasis $|0\rangle$ and $|1\rangle$ of the operator σ_z,

$$|Q_1Q_2\rangle = \frac{1}{\sqrt{2}}\sum_i |ii\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big). \tag{5.84}$$

To perform a parallel measurement of this composite system, consider a measurement of Q_1

using device A_1 and of Q_2 using A_2. Here, the first measurement is of the spin component along the $\hat{n}_1$ direction where $\hat{n}_1 \cdot \hat{z} = \cos\theta_1$, while the other measurement is of the spin component along the $\hat{n}_2$ direction where $\hat{n}_2 \cdot \hat{z} = \cos\theta_2$. Since these measurements are made on separate components of the entangled system, the time ordering has no effect on the subsequent calculations.

Entangling the ancillae with the quantum system, the state (5.84) becomes

$$|Q_1 Q_2 A_1 A_2\rangle = \frac{1}{\sqrt{2}} \sum_{ijk} U_{ij}\, V_{ik}\, |jk\, jk\rangle. \tag{5.85}$$

If we restrict the qubit states to the xz plane of the Bloch sphere, then we can write the transformation between the basis of σ_z and $\vec{\sigma} \cdot \hat{n}_1$ as the rotation matrix U,

$$U = \begin{pmatrix} \cos(\theta_1) & -\sin(\theta_1) \\ \sin(\theta_1) & \cos(\theta_1) \end{pmatrix}, \tag{5.86}$$

and similarly for $\vec{\sigma} \cdot \hat{n}_2$. Tracing over the composite quantum system, the density matrix of the ancillae is

$$\rho(A_1 A_2) = \frac{1}{2} \sum_{ii'jk} U_{ij} U^*_{i'j}\, V_{ik} V^*_{i'k}\, |jk\rangle\langle jk| = \frac{1}{2} \sum_{jk} \Big| \sum_i U_{ij}\, V_{ik} \Big|^2 |jk\rangle\langle jk|. \tag{5.87}$$

The probability to observe outcome j in device A_1 together with outcome k in device A_2 is the coherent distribution,

$$p_{jk} = \frac{1}{2} \Big| \sum_i U_{ij}\, V_{ik} \Big|^2, \tag{5.88}$$

while the marginal probabilities of A_1 and A_2 are perfectly random (both equal to 1/2). Thus, the conditional probability to observe outcome j with A_1, given that outcome k was obtained with A_2 (and vice versa), is

$$p_{j|k} = p_{k|j} = \Big| \sum_i U_{ij}\, V_{ik} \Big|^2. \tag{5.89}$$

The effects of the entanglement between systems Q_1 and Q_2 can be understood through the correlation between measurement outcomes. First, consider compatible measurements in which both measurements are of σ_z so that $\theta_1 = \theta_2 = 0$. In this case, the full wave function (5.85) becomes a four-particle GHZ state [recall Eq. (2.24)],

$$A_1[\sigma_z],\ A_2[\sigma_z]: \qquad |Q_1 Q_2 A_1 A_2\rangle = \frac{1}{\sqrt{2}} \sum_i |iiii\rangle = \frac{1}{\sqrt{2}} \Big(|0000\rangle + |1111\rangle \Big). \tag{5.90}$$

The measurement outcomes are described by the joint probability

$$p_{jk} = \frac{1}{2}\, \delta_{jk}, \tag{5.91}$$

which yields the conditional probabilities,

$$p_{j|k} = p_{k|j} = \delta_{jk}. \tag{5.92}$$

We see that the measurement outcomes are perfectly correlated, so that whichever of the two states we observe, we know that the corresponding outcome for the other system must be the same.

This perfect correlation of measurement outcomes does not imply faster-than-light signaling. Indeed, the choice of basis and the measurement outcome for A_1 would have to be sent along a causal classical communication channel to A_2, where both results can be compared [37, 148]. It is straightforward to show that for any set of parallel measurements of the same observable, the outcomes will be perfectly correlated. That is, due to the Schmidt decomposition, the degree of entanglement is invariant under a change of basis so that the measurement outcomes will always be correlated in the same way.

For the second example, we consider incompatible measurements where we measure σ_z using A_1 and σ_x using A_2. Such measurements are incompatible in the sense that the observables do not commute, $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\,\sigma_k$. In this case, the wave function (5.85) is

$$A_1[\sigma_z],\ A_2[\sigma_x]:\quad |Q_1Q_2A_1A_2\rangle = \frac{1}{\sqrt{2}}\sum_{ik} U'_{ik}\,|ikik\rangle = \frac{1}{2}\Big(|0000\rangle - |0101\rangle + |1010\rangle + |1111\rangle\Big). \tag{5.93}$$

The measurement outcomes are now described by the joint distribution

$$p_{jk} = \frac{1}{4}, \tag{5.94}$$

and conditional probabilities

$$p_{j|k} = p_{k|j} = \frac{1}{2}. \tag{5.95}$$

In other words, the outcomes are completely uncorrelated and random. If the state j is observed with A_1, then this has no correlation to the state k observed with A_2.

Using entropy Venn diagrams, we can examine how the entanglement is distributed for different sets of measurements. The diagrams for three different measurements are shown in Fig. 5.5. In all three scenarios, the center of the diagram, the ternary mutual entropy, vanishes because the underlying state (5.85) is pure. The negative conditional entropies are indicative of the entanglement between the subsystems. As we already saw, tracing over Q_1 and Q_2 always leaves the ancilla correlated. Interestingly, when making incompatible measurements, the ancillae are uncorrelated with each other, but are even more strongly entangled with the quantum systems as compared to the case of compatible measurements [59,72].

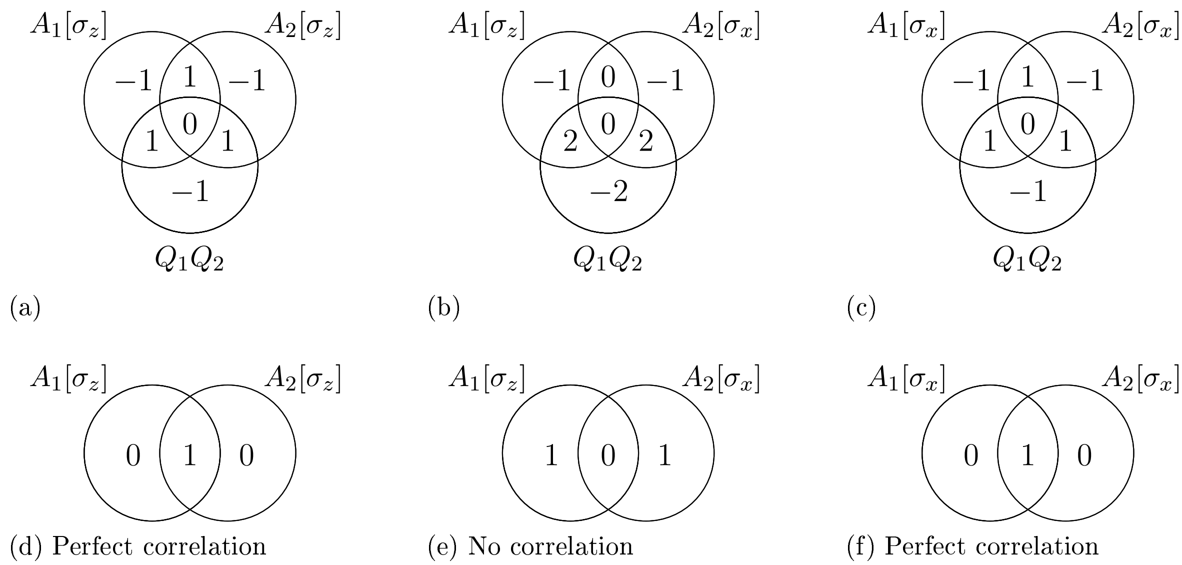

Figure 5.5: Entropy diagrams for three different parallel measurements of a Bell state. When incompatible spins (middle) are measured, the ancillae are uncorrelated. The ancillae are fully correlated for compatible measurements (left and right).

5.4 Conclusions

As part of Chs. 2-4, I introduced the main ideas underlying quantum correlations, quantum information theory, and quantum measurement. Here, in this chapter, I described the basic model of measurement that is used in this work. Starting with an expanded model of weak measurements, I showed how the formalism can equally well describe measurements of arbitrary interaction strengths and derived generalized results for the weak values corresponding to spin and position measurements. I discussed strong measurements from the perspective of quantum information theory, using entropy Venn diagrams to track how information is processed in a measurement. I focused on measurements of a single quantum state and parallel measurements of an entangled system. I will expand upon this foundation in the following chapters in the context of parallel and consecutive measurements. In the final chapter, these results will be used to construct a protocol for generating remote entanglement.

Chapter 6

Parallel Quantum Measurements

6.1 Introduction

Parallel measurements of a composite quantum system appear in many measurement schemes such as, for example, in Bell inequality violations [146, 147] and the quantum eraser [149]. In this chapter, I will apply the parallel measurement formalism to the Bell-state quantum eraser experiment, where one part of an entangled pair of photons is sent through a double-slit apparatus. I will show that a full quantum information-theoretic analysis can be used to derive the complementarity relation that is fundamental to this system. That is, I will construct the trade-off between the information extracted about the photon's path through the slits (conventionally termed the "which-path" information) and the coherence of its density matrix.

Quantum systems can display particle- or wave-like properties, depending on the type of measurement that is performed on them. The Bell-state quantum eraser is an experiment that brings the duality to the forefront, as a single measurement can retroactively be made to measure particle-like or wave-like properties (or anything in between). Here we develop a unitary information-theoretic description of this quantum measurement situations that sheds light on the trade-off between the quantum and classical features of the measurement. In particular, we show that both the coherence of the quantum state and the classical information obtained from it can be described using quantum-information-theoretic tools only, and that those two measures satisfy an equality on account of the chain rule for entropies. The coherence information and the which-path information have simple interpretations in terms of state preparation and state determination, and suggest ways to account for the relationship between the classical and the quantum world.

Wave-particle duality is an iconic feature of quantum mechanics, one not shared by classical systems in which an object cannot simultaneously have a wave and particle nature. Unraveling the mysteries behind wave-particle duality has occupied the better part of the last century, while significant advances in our understanding have come both from clever experimental approaches as well as theoretical developments. Pivotal experiments were the quantum optical implementations of Wheeler's [150] delayed-choice experiment (see, e.g., [151], as well as [152, 153], which were based on the theoretical proposal in [154]), which are equivalent in principle to delayed-choice quantum eraser experiments [34, 155]. For a thorough review of delayed-choice experiments, see [156] and references therein. The theoretical advances have framed the discussion of the wave-particle duality in terms of a quantum-information-theoretic trade-off between the coherence of the quantum system and the information that one may attempt to obtain about the particle path in the interferometer or double-slit experiment (the "which-path" information) [157–161].

The delayed-choice experiments highlight an important feature of this new understanding of wave-particle duality: while as per Bohr's complementarity principle [162] it is the nature of the experiment that determines whether we shall observe wave- or particle-like behavior, it is clear that the nature of the experiment can be changed after it has already taken place. In other words, the same experiment can retroactively be made to measure wave or particle properties, or anything in between [152, 153, 163]. Such a state of affairs is often greeted with skepticism, as the experiments seem (to some) to imply that the delayed choice of the measurement changes the quantum state retroactively, thus violating causality (see, e.g., the discussion in [163]). The natural interpretation of these results is that a quantum system has both particle-like and wave-like properties at the same time, and that the results of measurements can reveal one or a mixture of those characteristics. Due to the classical nature of the measuring devices, however, they do not—in fact, cannot—reflect the true nature of the quantum state. In the following, we will make these arguments in a strictly information-theoretic setting.

We develop the framework of (possibly delayed) complementary measurements (which-path or which-phase) in terms of a quantum information-theoretic description of the Bell-state quantum eraser, but the formalism is general and applies equally to any situation where measurements are made in parallel on two (and in an obvious extension to several) entangled quantum systems, such as the Garisto-Hardy entanglement eraser [164].

We first describe the ordinary double-slit experiment performed on one half of a Bell state, then the polarization-tagged version where which-path information can be extracted, followed by the quantum erasure procedure. In the next section we describe these steps in terms of classical and quantum information theory that results in an information-theoretic equality that mirrors (and is completely analogous to) the trade-off between distinguishability and visibility of Greenberger and Yasin [157] as well as Englert [158]. The equality involves the coherence of the quantum system and the information obtained about its path just as Bagan et al. have recently shown [161], but we do not assume a specific form for the measure of coherence as it emerges naturally from the information-theoretic analysis. The "conservation law" between coherence and information appears simply as a consequence of the chain rule for entropies. We offer conclusions in which we suggest what information is actually encoded in the measurement devices, given that it cannot reflect information about the quantum state.

6.2 The Bell-State Quantum Eraser

The Bell-state quantum eraser experiment [165], as illustrated in Fig. 6.1, proceeds as follows. An entangled pair of photons, A and B, is prepared by spontaneous parametric down-conversion (SPDC) [34] in the Bell state,

$$|\Psi\rangle_{AB} = \frac{|h\rangle|v\rangle + |v\rangle|h\rangle}{\sqrt{2}} \ , \tag{6.1}$$

where the first and second states refer to photons A and B, respectively, and $|h\rangle, |v\rangle$ are the horizontal and vertical linear polarization states. Note that an obvious extension of the present construction is to allow for arbitrary entangled photon pairs, such as for example the q-Bell states [71] in which a parameter q interpolates between product states ($q = 0$ or 1) and fully entangled states ($q = 1/2$). We mention briefly the consequence of $q \neq 1/2$ at the end of Sec. 6.4.

Photon B, called the "idler", travels along the upper path where a polarizing beam splitter (PBS) with its optical axis oriented at an angle θ to the $|h\rangle, |v\rangle$ basis allows for polarization measurements in specific bases. Meanwhile, photon A, called the "signal", travels along the lower path towards a double-slit apparatus. Photon A will pass through the double slit to subsequently be detected by a CCD camera (denoted D_X) from which it is possible to construct an interference pattern. The pattern is erased by placing two quarter-wave plates (QWPs) in front of each slit. This tags the path of photon A and provides path information. See Fig. 6.1 for a schematic of the experiment.

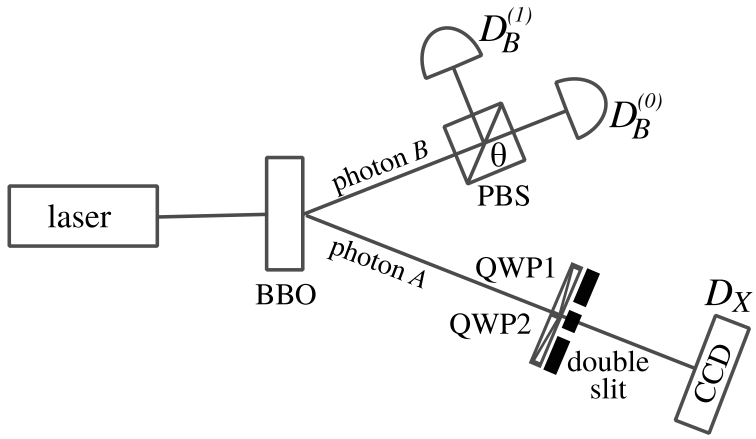

Figure 6.1: Schematic [165] of the double-slit Bell-state quantum eraser experiment. After production of the Bell state (6.1) by type-II spontaneous parametric down-conversion (SPDC) in the BBO (β-barium borate) crystal, photon B travels along the upper branch to a polarizing beam splitter (PBS), with optical axis oriented at angle θ relative to the $|h\rangle$, $|v\rangle$ basis. Its polarization is subsequently measured in the rotated basis by the photodetectors $D_B^{(0)}$, $D_B^{(1)}$. Photon A travels down to the quarter-wave plates (QWPs) and double slit, and then to a CCD camera, denoted D_X, which plays the role of an interference screen.

6.2.1 Splitting the Photon Path

The full wave function describing the entangled pair of photons A and B is

$$|\Psi\rangle_{AB} = \frac{|h\rangle_P|v\rangle_B + |v\rangle_P|h\rangle_B}{\sqrt{2}} \otimes |\psi\rangle_Q, \tag{6.2}$$

where the Hilbert space $\mathcal{H}_A = \mathcal{H}_P \otimes \mathcal{H}_Q$ of photon A is composed of polarization P and spatial Q degrees of freedom. The polarizations of photons A and B, entangled in a Bell state, are decoupled from the spatial state $|\psi\rangle$ of photon A. We drop the spatial states of photon B as they remain decoupled throughout. Sending photon A through the double slit transforms only the spatial states of A so that Eq. (6.2) evolves to

$$|\Psi\rangle_{AB} = \frac{|h\rangle_P|v\rangle_B + |v\rangle_P|h\rangle_B}{\sqrt{2}} \otimes \frac{|\psi_1\rangle_Q + |\psi_2\rangle_Q}{\sqrt{2}}. \tag{6.3}$$

The states $|\psi_j\rangle$ denote the path of photon A corresponding to slit j. Note that the extension of this framework to allow for N-path devices is straightforward. The spatial degree of freedom of A, denoted by Q, is still independent from its polarization.

Tracing over the polarization states of photons A and B, the density matrix describing the spatial modes of photon A is the pure state

$$\rho_Q = \frac{1}{2}\Big(|\psi_1\rangle_Q\langle\psi_1| + |\psi_2\rangle_Q\langle\psi_2| + |\psi_1\rangle_Q\langle\psi_2| + |\psi_2\rangle_Q\langle\psi_1|\Big). \tag{6.4}$$

The expectation value in the position basis $|x\rangle$ of the screen D_X yields the probability to observe photon A at the spatial location x

$$\langle x|\rho_Q|x\rangle = p(x) = \frac{1}{2}\big|\psi_1(x) + \psi_2(x)\big|^2, \tag{6.5}$$

where we define the coefficients $\psi_j(x) = \langle x|\psi_j\rangle$. This probability distribution is a coherent superposition and the usual double-slit interference fringes will be observed on the screen. In the appendix we show how the characteristic fringes can be derived from a von Neumann measurement of Q by the detector D_X.

6.2.2 Tagging the Photon Path

To extend this discussion to a quantum eraser experiment, a tagging operation is performed on the two branches of the double-slit apparatus in order to provide information about the path of photon A. In practice, this is implemented by placing a quarter-wave plate (QWP) in front of each slit. Recall from Eq. (4.33) in Ch 4.2.1 that the Jones matrix for a general wave plate oriented at an angle β (the fast axis) to our coordinate system (in this case, $|h\rangle$ and $|v\rangle$) is [111, 112]

$$U = \begin{pmatrix} \cos(\frac{\alpha}{2}) + i\sin(\frac{\alpha}{2})\cos(2\beta) & i\sin(\frac{\alpha}{2})\sin(2\beta) \\ i\sin(\frac{\alpha}{2})\sin(2\beta) & \cos(\frac{\alpha}{2}) - i\sin(\frac{\alpha}{2})\cos(2\beta) \end{pmatrix}, \tag{6.6}$$

where $\alpha = \pi/2$ for a QWP. More specifically, the QWP in front of slit 1 (slit 2) has its fast axis at $\beta = 45°$ ($\beta = -45°$), which leads to

$$U_{\text{QWP}}^{(\pm)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \pm i \\ \pm i & 1 \end{pmatrix}, \tag{6.7}$$

where $U_{\text{QWP}}^{(+)} = U_{\text{QWP}}^{(1)}$ and $U_{\text{QWP}}^{(-)} = U_{\text{QWP}}^{(2)}$ are the matrices associated with slit 1 and 2, respectively. These transform linearly polarized states $|h\rangle$ and $|v\rangle$ into circularly polarized states according to

$$\begin{aligned} U_{\text{QWP}}^{(1)}\,|h\rangle &= \frac{|h\rangle + i|v\rangle}{\sqrt{2}} = |L\rangle, \\ U_{\text{QWP}}^{(1)}\,|v\rangle &= \frac{|v\rangle + i|h\rangle}{\sqrt{2}} = i|R\rangle, \\ U_{\text{QWP}}^{(2)}\,|h\rangle &= \frac{|h\rangle - i|v\rangle}{\sqrt{2}} = |R\rangle, \\ U_{\text{QWP}}^{(2)}\,|v\rangle &= \frac{|v\rangle - i|h\rangle}{\sqrt{2}} = -i|L\rangle, \end{aligned}$$

where $|R\rangle$ ($|L\rangle$) denotes right-handed (left-handed) circular polarization.

When photon A passes through the QWPs and the double slit, its polarization becomes entangled with its spatial degree of freedom so that the wave function (6.2) evolves to

$$|\widetilde{\Psi}\rangle_{AB} = \frac{1}{\sqrt{2}}\left(\frac{|L\rangle_P|v\rangle_B + i\,|R\rangle_P|h\rangle_B}{\sqrt{2}} \otimes |\psi_1\rangle_Q + \frac{|R\rangle_P|v\rangle_B - i\,|L\rangle_P|h\rangle_B}{\sqrt{2}} \otimes |\psi_2\rangle_Q \right), \tag{6.8}$$

where the tilde indicates that the tagging operation has been performed. Grouping together

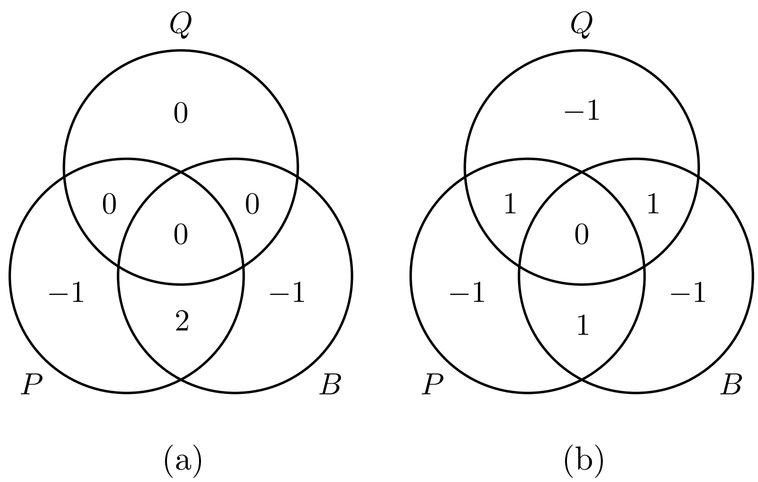

Figure 6.2: Entropy Venn diagrams showing the effect of the tagging operation. (a) Before tagging, the polarization of photons A (denoted by P) and B are entangled in a Bell state. (b) After tagging, the spatial degree of freedom of photon A (denoted by Q) becomes entangled with the polarizations of A and B according to Eq. (6.9).

the polarization states of photon B, we can equivalently express this state as

$$|\widetilde{\Psi}\rangle_{AB} = \frac{1}{\sqrt{2}}\left(\frac{|\psi_1\rangle_Q|L\rangle_P + |\psi_2\rangle_Q|R\rangle_P}{\sqrt{2}} \otimes |v\rangle_B + i\,\frac{|\psi_1\rangle_Q|R\rangle_P - |\psi_2\rangle_Q|L\rangle_P}{\sqrt{2}} \otimes |h\rangle_B\right). \quad (6.9)$$

The entanglement between the two degrees of freedom of photon A causes the spatial modes Q to become completely mixed

$$\rho_Q = \frac{1}{2}\big(|\psi_1\rangle_Q\langle\psi_1| + |\psi_2\rangle_Q\langle\psi_2|\big), \quad (6.10)$$

so that interference is no longer observed on the screen. In Fig. 6.2 we show the entropy Venn diagrams [59] before (a) and after (b) the tagging operation with the QWPs. In these diagrams, the sum of all the entries in a circle add up to the entropy of the subsystem, and the entropy shared between subsystems is indicated in the overlap between circles. Conditional entropies appear in unshared areas of the circle, and can be negative in quantum mechanics [58] (they must be positive if they are classical Shannon entropies). Entropies shared between three systems (the center of the diagrams in Fig. 6.2) can be negative both in classical and quantum physics [59]. All of the von Neumann entropies $S(\rho) = -\mathrm{Tr}\rho \log \rho$ can be calculated in a straightforward manner from the density matrix $\rho_{QPB} = |\widetilde{\Psi}\rangle_{AB}\langle\widetilde{\Psi}|$ and the marginalized density matrices $\rho_Q = \mathrm{Tr}_{PB}(\rho_{QPB})$, $\rho_B = \mathrm{Tr}_{QP}(\rho_{QPB})$, etc.

Before tagging [see Eq. (6.2)], Q is completely decoupled from the polarization P of photon A and of photon B, which together are entangled in a Bell state. After tagging [see Eq. (6.9)], all three variables Q, P, and B are in a tripartite entangled state. Note that the ternary mutual entropy $S(Q:P:B)$ vanishes in both diagrams since the underlying density matrix is a pure state [59]. The expression for the ternary shared entropy in terms of subsystem entropies can be read off the Venn diagram in general as $S(Q:P:B) = S(Q) + S(P) + S(B) - S(QB) - S(QP) - S(PB) + S(QPB)$, and similarly for any pairwise shared entropies.

6.2.3 Erasing the Photon Path

As is by now well-known [166], it is still possible to extract an interference pattern from the screen data, even when the system Q has been tagged, *if* we have additional information about the state of photon B. Suppose we perform a polarization measurement of B in a basis that is described by an angle θ relative to the $|h\rangle$, $|v\rangle$ basis. For a general change of basis, the polarization states of B are written as

$$|v\rangle_B = U_{00}|0\rangle_B + U_{01}|1\rangle_B, \tag{6.11}$$

$$|h\rangle_B = U_{10}|0\rangle_B + U_{11}|1\rangle_B. \tag{6.12}$$

For simplicity, we use the following parametrization for the matrix elements U_{ij} of the rotation operator that transforms $|h\rangle, |v\rangle$ to the new basis spanned by $|0\rangle, |1\rangle$ in terms of the single angle θ:

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} . \tag{6.13}$$

Rewriting the states of B in the basis $|0\rangle, |1\rangle$ and entangling B with a polarization detector D_B transforms Eq. (6.9) to [59]

$$|\widetilde{\Psi}\rangle_{ABD_B} = \frac{1}{2}\sum_{mk} i^m \; |\psi_m^k\rangle_Q \otimes |m\rangle_P \otimes |kk\rangle_{BD_B}, \tag{6.14}$$

where $k = 0, 1$ labels the polarization of B and D_B while $m = 0\,(\text{“}L\text{”}), 1\,(\text{“}R\text{”})$ denotes the polarization of A, and where we defined the spatial states $|\psi_m^k\rangle_Q$ of photon A

$$|\psi_L^k\rangle_Q = U_{0k}|\psi_1\rangle_Q - iU_{1k}|\psi_2\rangle_Q, \tag{6.15}$$

$$|\psi_R^k\rangle_Q = U_{1k}|\psi_1\rangle_Q - iU_{0k}|\psi_2\rangle_Q. \tag{6.16}$$

These states describe the spatial state of photon A (the system Q) when it has a circular polarization m and is correlated with photon B that has polarization k. Only the states for a given polarization m are orthonormal

$$\langle\psi_m^{k'}|\psi_m^k\rangle \;\; = \;\; \delta_{kk'}, \tag{6.17}$$

$$\langle\psi_L^{k'}|\psi_R^k\rangle \;\; = \;\; U_{0k'}^*\, U_{1k} + U_{0k}\, U_{1k'}^*. \tag{6.18}$$

Of course, the state ρ_Q derived from (6.14) is still completely mixed as in (6.10) so that no interference can be observed on the screen. However, as long as the erasure angle θ is nonzero it is now possible to extract an interference pattern given the outcome of the polarization measurement of photon B (even if that measurement occurs much later than the measurement of photon A).

From the wave function (6.14), we can compute the joint density matrix for photon $A \equiv QP$ (spatial Q and polarization P) and detector D_B. Tracing over B yields

$$\rho_{AD_B} = \frac{1}{2} \sum_k \rho_A^k \otimes |k\rangle_{D_B} \langle k|, \tag{6.19}$$

where the (orthonormal) states of photon A, conditional on the state k of detector D_B, are $\rho_A^k = |\phi_k\rangle\langle\phi_k|$ and $|\phi_k\rangle = \frac{1}{\sqrt{2}} \sum_m i^m \, |\psi_m^k\rangle_Q \otimes |m\rangle_P$. The effect of the erasure is contained in the behavior of these states as the measurement angle θ is varied. For a measurement in the original basis ($\theta = 0$), detector D_B prepares photon A in one of the fully entangled states: $|\phi_0\rangle \propto |\psi_1\rangle_Q |L\rangle_P + |\psi_2\rangle_Q |R\rangle_P$ and $|\phi_1\rangle \propto |\psi_1\rangle_Q |R\rangle_P - |\psi_2\rangle_Q |L\rangle_P$. From these expressions we can infer, with a polarization measurement of photon A, its path. For instance, outcome $k = 0$, $m = L$ is associated with the spatial state $|\psi_1\rangle$ for slit 1. Therefore, polarization measurements of photon B at $\theta = 0$ yield full path information and no interference fringes.

On the other hand, for a measurement in the diagonal basis at $\theta = \pi/4$, detector D_B prepares photon A in one of the completely decoupled states: $|\phi_0\rangle \propto (|\psi_1\rangle_Q - i|\psi_2\rangle_Q) \otimes (|L\rangle_P + i|R\rangle_P)$ and $|\phi_1\rangle \propto (|\psi_1\rangle_Q + i|\psi_2\rangle_Q) \otimes (-|L\rangle_P + i|R\rangle_P)$. Now, a polarization measurement of A cannot reveal path information, and the coherently summed spatial modes lead to interference fringes. In the appendix, we compute these interference patterns and show their dependence on the erasure angle θ. Regardless of the temporal order of the two polarization measurements, the measurement of B can be seen as state *preparation*, while the measurement of A is state *determination*, that is, extraction of which-path information.

6.3 Information Theory

6.3.1 State Preparation

We illustrate the quantum erasure mechanism by building on the entropy Venn diagrams in Fig. 6.2 and constructing the entropic relationships between the random variables Q, P, and D_B from the joint and marginal entropies associated with Eq. (6.19). In the basis $|\phi_k\rangle \otimes |k\rangle$ it is clear that the entropy of (6.19) is $S(QPD_B) = 1$, while the marginal entropies are all $S(Q) = S(P) = S(D_B) = 1$. Tracing over detector D_B, the total entropy of $A \equiv QP$ is also $S(QP) = 1$. The pairwise entropy $S(PD_B)$ is equal to $S(QB)$ by the Schmidt decomposition of (6.14), which in turn is equal to $S(QD_B)$ by the symmetry between the B and D_B states. The remaining pairwise entropy is computed from the density matrix

$$\rho_{QD_B} = \frac{1}{2}\sum_k \rho_Q^k \otimes |k\rangle_{D_B}\langle k|, \tag{6.20}$$

where the state of Q conditional on the outcome k of detector D_B is

$$\begin{aligned} \rho_Q^k &= \frac{1}{2}\sum_m |\psi_m^k\rangle_Q\langle\psi_m^k| \\ &= \frac{1}{2}\Big(|\psi_1\rangle_Q\langle\psi_1| + |\psi_2\rangle_Q\langle\psi_2| + i\,(-1)^k \sin 2\theta\Big[|\psi_1\rangle_Q\langle\psi_2| - |\psi_2\rangle_Q\langle\psi_1|\Big]\Big). \end{aligned} \tag{6.21}$$

From this expression, we see that for a given outcome k, the spatial degree of freedom of photon A is generally no longer mixed [as in Eq. (6.10)] and has a coherence that is controlled by the sine of the measurement angle. In turn, it is possible to extract interference fringes from the screen. With the Bloch vector $\vec{a} = -(-1)^k \sin(2\theta)\,\hat{y}$, Eq. (6.21) can be expressed as $\rho_Q^k = \frac{1}{2}(\mathbb{1} - (-1)^k \sin(2\theta)\,\sigma_y)$, where $\hat{y}$ is a unit vector, $\mathbb{1}$ is the identity matrix of dimension two, and σ_y is a Pauli matrix. This density matrix varies from a fully mixed state $|\vec{a}| = 0$ at

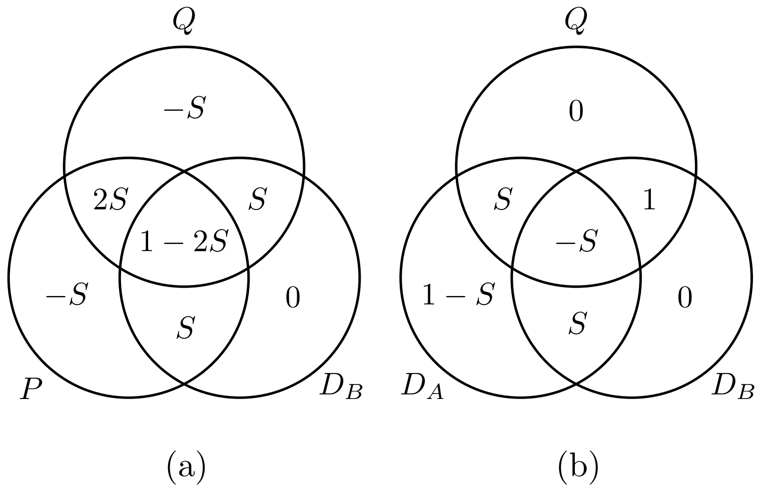

Figure 6.3: Entropy Venn diagrams for (a) state preparation with detector D_B [see Eq. (6.14)], and (b) state determination with detector D_A [see Eq. (6.23)]. Here, S is the entropy of Eq. (6.21).

$\theta = 0$ to a pure state $|\vec{a}| = 1$ at $\theta = \pi/4$.

To compute the entropy of the block-diagonal matrix in Eq. (6.20), we first find the entropy of ρ_Q^k. The eigenvalues of ρ_Q^k are $\lambda_\pm = \frac{1}{2}(1 \pm |\vec{a}|) = \frac{1}{2}[1 \pm \sin(2\theta)]$ and are independent of the index k, leading to equal entropies $S(\rho_Q^0) = S(\rho_Q^1)$. Therefore, the entropy of (6.20), which is also equal to the entropy $S(PD_B)$, is [25]

$$S(QD_B) = 1 + \frac{1}{2}\sum_k S(\rho_Q^k) = 1 + S, \tag{6.22}$$

where $S = -\lambda_+ \log \lambda_+ - \lambda_- \log \lambda_-$ is the entropy of (6.21), and varies from $S = 1$ at $\theta = 0$ to $S = 0$ at $\theta = \pi/4$.

The relationship between the variables Q, P, and D_B is summarized by the entropy Venn diagram in Fig. 6.3(a). As a result of tagging the path of photon A, the spatial and polarization modes of A are entangled, given the state of D_B, for $S > 0$ ($\theta < \pi/4$). The amount of entanglement S varies with the erasure angle θ and specifies the degree to which the polarization P can reveal information about the spatial mode Q. The non-zero ternary mutual entropy $S(Q\!:\!P\!:\!D_B) = 1 - 2S$ indicates that the mutual entropy of Q and P can be shared by detector D_B [71].

6.3.2 State Determination

In order to reveal information about the path of photon A, its polarization is measured after it passes the double-slit apparatus using a detector D_A in the circular $|L\rangle, |R\rangle$ basis. After the measurement with D_A, the wave function (6.14) evolves to

$$|\Psi'\rangle_{AD_ABD_B} = \frac{1}{2}\sum_{mk} i^m |\psi_m^k\rangle_Q \otimes |mm\rangle_{PD_A} \otimes |kk\rangle_{BD_B}. \tag{6.23}$$

The entropic relations between the variables Q, D_A, and D_B are computed from their joint density matrix, which is found by tracing out the polarization states of photons A and B from (6.23),

$$\rho_{QD_AD_B} = \frac{1}{4}\sum_{mk} |\psi_m^k\rangle_Q\langle\psi_m^k| \otimes |m\rangle_{D_A}\langle m| \otimes |k\rangle_{D_B}\langle k|. \tag{6.24}$$

For a set of outcomes k and m, the corresponding spatial state of photon A is $|\psi_m^k\rangle_Q$. It is straightforward to show that the entropy of Eq. (6.24) is $S(QD_AD_B) = 2$, while the marginal entropies are $S(Q) = S(D_A) = S(D_B) = 1$.

Tracing over the spatial states Q in Eq. (6.24), we find that the two polarization detectors are uncorrelated

$$\rho_{D_AD_B} = \frac{1}{2}\mathbb{1}_{D_A} \otimes \frac{1}{2}\mathbb{1}_{D_B}, \tag{6.25}$$

with a joint entropy $S(D_A D_B) = 2$ and

$$S(D_A : D_B) = 0. \tag{6.26}$$

Thus, the measurement with D_A reveals no information about the state of D_B. This is not surprising as the QWPs act as a "controlled-NOT" gate on the polarization. Indeed, conditioning on the spatial states of photon A yields

$$S(D_A : D_B|Q) = S \geq 0. \tag{6.27}$$

In a sense, therefore, the states of Q encrypt the relationship between the state preparation with D_B and its readout with D_A.

Tracing out the states of the polarization detector D_A in Eq. (6.24), the joint density matrix ρ_{QD_B} is unchanged from Eq. (6.20), since the measurement with D_A does not affect the correlations between Q and D_B.

Tracing over the states of D_B in Eq. (6.24), the joint density matrix for Q and D_A in turn is

$$\rho_{QD_A} = \frac{1}{2}\sum_m \rho_Q^m \otimes |m\rangle_{D_A}\langle m| = \frac{1}{2}\mathbb{1}_Q \otimes \frac{1}{2}\mathbb{1}_{D_A}, \tag{6.28}$$

where the density matrix of Q, conditional on the polarization outcome m of detector D_A, is

$$\rho_Q^m = \frac{1}{2}\sum_k |\psi_m^k\rangle_Q\langle\psi_m^k| = \frac{1}{2}\Big(|\psi_1\rangle_Q\langle\psi_1| + |\psi_2\rangle_Q\langle\psi_2|\Big), \tag{6.29}$$

which is independent of the polarization index m, and is equivalent to the full density matrix ρ_Q. That is, $\rho_Q^m = \rho_Q = \frac{1}{2}\mathbb{1}_Q$ is a completely mixed state. Finally, the joint entropy of Q

and D_A is $S(QD_A) = 2$. The entropic relationships between the variables Q, D_A and D_B are summarized by the Venn diagram in Fig. 6.3(b).

6.3.3 Information-Theoretic Origins of Coherence and Path Information

From the marginal and joint entropies computed in the previous section, we can construct information-theoretic relationships between the variables Q, D_A, and D_B.

6.3.3.1 Coherence

The information shared between the preparation with detector D_B and the quantum system Q (the spatial state of photon A) is given by the mutual entropy

$$S(Q : D_B) = 1 - S \leq 1 \; , \tag{6.30}$$

with S defined in Eq. (6.22). We can understand how this entropy depends on the erasure angle θ by considering two cases. First, from the joint density matrix ρ_{QD_B} in Eq. (6.20), a measurement of photon B at an angle $\theta = 0$ decouples Q from detector D_B

$$\theta = 0 : \quad \rho_{QD_B} = \frac{1}{2}\mathbb{1}_Q \otimes \frac{1}{2}\mathbb{1}_{D_B}, \tag{6.31}$$

and the conditional state ρ_Q^k in Eq. (6.21) becomes a statistical mixture, i.e., interference cannot be observed on the screen. In this case, $S(Q : D_B) = 0$ and there is no information shared between the two variables. However, increasing the erasure angle to $\theta = \pi/4$ leads to perfect correlation

$$\theta = \frac{\pi}{4}: \quad \rho_{QD_B} = \frac{1}{2}\Big(|f\rangle_Q\langle f| \otimes |0\rangle_{D_B}\langle 0| + |a\rangle_Q\langle a| \otimes |1\rangle_{D_B}\langle 1|\Big), \tag{6.32}$$

where $|f\rangle_Q = \frac{1}{\sqrt{2}}(|\psi_1\rangle_Q - i|\psi_2\rangle_Q)$ corresponds to a fringe pattern and $|a\rangle_Q = \frac{1}{\sqrt{2}}(|\psi_1\rangle_Q + i|\psi_2\rangle_Q)$ corresponds to an anti-fringe (phased-shifted) pattern. Now, $\rho_Q^0 = |f\rangle_Q\langle f|$ and $\rho_Q^1 = |a\rangle_Q\langle a|$ are coherent superpositions, i.e., it is possible to extract interference on the screen. At this angle, $S(Q : D_B) = 1$. Therefore, the mutual entropy $S(Q : D_B)$ is related to the coherence of the conditional states ρ_Q^k, and in turn, to the visibility of interference fringes as we will see below.

6.3.3.2 Path Information

From the joint density matrix ρ_{QD_A} computed in Eq. (6.28), the polarization measurement with detector D_A reveals nothing about the spatial degree of freedom of A since the joint state is completely decoupled. It follows that the mutual information vanishes

$$S(Q : D_A) = 0. \tag{6.33}$$

In other words, if we do not know the outcome of the polarization measurement of photon B, an attempt to measure the polarization of A after it traverses the double slit and QWPs will not reveal anything about the spatial state of A. On the other hand, if we do know the state of D_B, then the conditional mutual information is

$$S(Q : D_A|D_B) = S \geq 0, \tag{6.34}$$

and varies with the erasure angle θ.

To understand the behavior of this quantity as a function of θ, consider the state of Q and D_A, conditional on the outcome k of D_B. According to Eq. (6.24), this state is $\rho^k_{QD_A} = \frac{1}{2}\sum_m |\psi^k_m\rangle_Q\langle\psi^k_m| \otimes |m\rangle_{D_A}\langle m|$. For an outcome $k = 0$ of a polarization measurement of photon B at angle $\theta = 0$,

$$\theta = 0 : \quad \rho^0_{QD_A} = \frac{1}{2}\Big[|\psi_1\rangle_Q\langle\psi_1| \otimes |L\rangle_{D_A}\langle L| + |\psi_2\rangle_Q\langle\psi_2| \otimes |R\rangle_{D_A}\langle R| \Big]. \tag{6.35}$$

At this angle, the conditional mutual information is maximal, $S(Q : D_A|D_B) = 1$, and it is clear that the state of the polarization detector D_A is associated with one of the states $|\psi_j\rangle_Q$, so that the measurement can reveal information about the path of photon A. For instance, an outcome L corresponds to the state $|\psi_1\rangle_Q$. As the erasure angle θ increases from zero to $\pi/4$, the information we have about the spatial state of A is reduced to zero, since the density matrix becomes decoupled:

$$\theta = \frac{\pi}{4} : \quad \rho^0_{QD_A} = |f\rangle_Q\langle f| \otimes \frac{1}{2}\,\mathbb{1}_{D_A}. \tag{6.36}$$

At this angle, $S(Q : D_A|D_B) = 0$. Therefore, the tagging operation with the QWPs only reveals information about the path of A as long as we have additional information about the state of photon B from its polarization measurement with detector D_B. Thus, Eq. (6.34) is the correct expression for which-path information. Note further that

$$S(Q : D_A D_B) = 1\ , \tag{6.37}$$

which implies that—given the outcomes of both polarization detectors D_A and D_B—it is possible to predict *with certainty* the outcome of a direct measurement of the path of photon A.

6.4 Discussion

The quantities considered in the previous section can be used to generalize the usual concepts of coherence and path information, allowing us to construct a more fundamental relationship that is derived from information-theoretic principles and a unitary model of quantum measurement [59].

Recapitulating the results from the previous section, we know that whether we extract fringes or antifringes from the screen is controlled by the state of D_B. The visibility of the fringes is, in turn, related to the coherence of the conditional state ρ_Q^k of Q in Eq. (6.21) and the mutual information $S(Q : D_B)$. As we have already seen, at angle $\theta = 0$ ($\theta = \pi/4$) the two variables Q and D_B are completely uncorrelated (correlated), the conditional state ρ_Q^k is a statistical mixture (coherent superposition), and we observe no interference (full interference).

On the other hand, information about the path of photon A is determined by the correlation between its polarization (via the state of detector D_A) and its spatial states. This correlation is computed from the conditional mutual information $S(Q\!:\!D_A|D_B)$, and must be conditioned on the state of D_B since Q and D_A are otherwise uncorrelated. When $\theta = 0$ ($\theta = \pi/4$), the variables Q and D_A are completely correlated (uncorrelated) given D_B, so that D_A can (cannot) reveal path information, and we extract no interference (full interference) from the screen.

These two information-theoretic quantities, namely, the coherence $S(Q\!:\!D_B)$ of system Q and its path information $S(Q\!:\!D_A|D_B)$, are fundamentally linked through the chain rule for entropies

$$S(Q\!:\!D_B) + S(Q\!:\!D_A|D_B) = S(Q\!:\!D_A D_B) = 1, \tag{6.38}$$

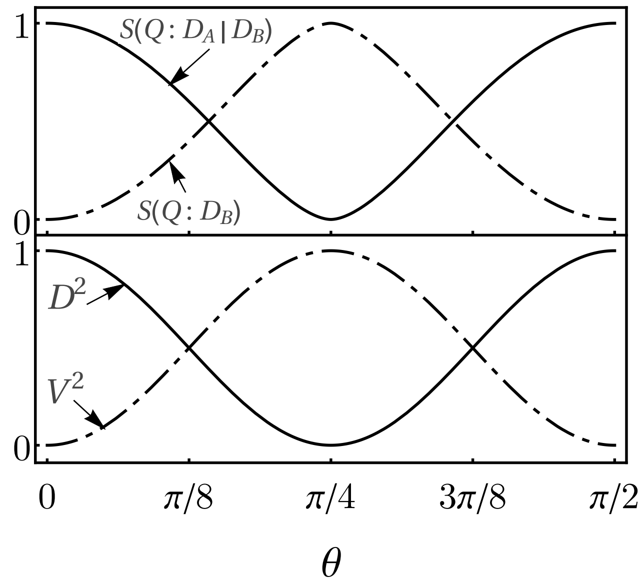

Figure 6.4: Top: Relationship between coherence and path information in (6.38) for the Bell-state quantum eraser as a function of the erasure angle θ. Shown are the information-theoretic quantities for the path information $S(Q:D_A|D_B) = S$ (solid) and coherence $S(Q:D_B) = 1 - S$ (dashed). Bottom: The square of the distinguishability D (solid) and fringe visibility V (dashed), defined in the text, as a function of the erasure angle θ. When $\theta = 0$ ($\theta = \pi/4$) there is full (no) path information and no (full) coherence.

and their sum is conserved throughout the erasure process. This information-theoretic formulation of complementarity generalizes earlier attempts [160,161,167] by explicitly referencing the measurement devices. We note that the absence of correlations between detectors D_A and D_B, $S(D_A : D_B) = 0$, is crucial to enforce complementarity.

We show in the top of Fig. 6.4 the coherence and information in Eq. (6.38) as a function of the erasure angle θ. From the eigenvalues $\lambda_\pm$, we can derive an alternative form for the entropy, S, appearing in the complementarity relation (6.38). First, we rewrite the argument of the sine function as $2\lambda_\pm = 1 \pm \sin(2[\theta + \pi/4 - \pi/4])$. Using the sum-difference formula, this can be expressed as $2\lambda_\pm = 1 \mp \cos(2\theta + \pi/2)$. Finally, by the double-angle formula, we find $2\lambda_\pm = 1 \mp (1 - 2\sin^2(\theta + \pi/4)$. Therefore, the entropy is $S = H[\sin^2(\theta + \pi/4)]$. We can compare these expressions to two other measures that are commonly used to discuss the wave-particle duality, namely the distinguishability, D, and the visibility, V [157, 158]. In general, $D^2 + V^2 \leq 1$, but becomes an equality when the detectors are prepared in pure states. In the particular case we are discussing, $D^2 = \cos^2(2\theta)$, while the fringe visibility is $V^2 = \sin^2(2\theta)$. They are shown in the bottom part of Fig. 6.4 and exhibit a remarkably similar behavior when compared to the information-theoretic complementarity principle.

From Eq. (6.38), we can derive additional information-theoretic relations with conditional and mutual entropies. With the definition of conditional mutual information [59], $S(Q : D_A|D_B) = S(Q|D_B) - S(Q|D_A D_B) = S(Q|D_B)$, Eq. (6.38) becomes a relation between a mutual entropy and conditional entropy:

$$S(Q{:}D_B) + S(Q|D_B) = 1. \tag{6.39}$$

Furthermore, $S(Q{:}D_A) - S(Q{:}D_A|D_B) = S(Q{:}D_A{:}D_B) \leq 0$, so that Eq. (6.38) becomes

$$0 \leq S(Q{:}D_A) + S(Q{:}D_B) \leq 1, \tag{6.40}$$

where the lower bound comes from the non-negativity of mutual entropies. This can be rewritten in terms of conditional entropies as

$$1 \leq S(Q|D_A) + S(Q|D_B) \leq 2, \tag{6.41}$$

where we used $S(Q) = 1$.

Bagan et al. recently constructed an entropic complementarity relation between coherence and path information in an interferometer [161] using an entropic measure for coherence. For path states with equal probability, orthogonal detector states, and orthogonal measurements, their relation can be written as an equality very similar to ours

$$C_{\mathrm{rel\,ent}}(\rho) + H(M:D) = 1, \tag{6.42}$$

where $C_{\mathrm{rel\,ent}}(\rho) = 1 - S(\rho)$ is a measure of the coherence of the particle's state ρ in the interferometer, and $H(M:D) = S(\rho)$ is the path information, which is the mutual entropy of the path detector states D and the results M of probing them with a measurement. The connection to our result (6.38) is immediately obvious, as $S(\rho)$ in Eq. (6.42) is indeed equivalent to our S, the entropy of the conditional state ρ_Q^k of photon A in Eq. (6.21). However, our measures of coherence and path information emerge naturally from a full information-theoretic analysis and yield more insight into the origins of their complementarity, in particular how the entropy of system Q is distributed among the detectors D_A and D_B (as summarized by Fig. 6.3).

We end this discussion by noting that if we prepare photons A and B in an imperfectly entangled state, e.g., the q-Bell state [71] $\sqrt{q}\,|h\rangle|v\rangle + \sqrt{1-q}\,|v\rangle|h\rangle$, the coherence and path information we derived are modified from their original forms. For an erasure angle $\theta = \pi/4$, we can write down a simple replacement for Eqs. (6.30) and (6.34) that appear in the complementarity relation of Eq. (6.38). That is,

$$\theta = \frac{\pi}{4}: \qquad S(Q:D_B) = 1 - S(q), \tag{6.43}$$

$$S(Q:D_A|D_B) = S(q), \tag{6.44}$$

where $S(q)$ replaces S, the entropy of the state (6.21). The quantity $S(q)$ is defined by the eigenvalues $\lambda_\pm = 1/2 \pm \sqrt{q(1-q)}$. The parameter q controls the initial entanglement between photons A and B and allows us to extract nonzero path information, at the cost of reduced coherence, even when $\theta = \pi/4$. Setting $q = 1/2$ recovers the original result of zero path information and full coherence, $S(q = 1/2) = S = 0$, at this erasure angle.

6.5 Interference Patterns

Here we derive the spatial intensity distribution for photon A that is incident on a screen D_X by modeling the interaction as a von Neumann measurement of the spatial location of photon A. Expanding the spatial states of A in terms of the position basis yields

$$|\psi_j\rangle_Q = \sum_{x=1}^{n} \psi_j(x)\, |x\rangle_Q, \tag{6.45}$$

where $j = 0, 1$ labels each slit. The states $|x\rangle$ can be discretized into n distinct locations according to

$$\begin{aligned} |x = 1\rangle &= |100 \cdots 0\rangle, \\ |x = 2\rangle &= |010 \cdots 0\rangle, \\ &\vdots \\ |x = n\rangle &= |0 \cdots 001\rangle, \end{aligned}$$

which denote the location x at which a photon is detected by D_X. Inserting this basis into the expression (6.14) and performing the measurement of Q with D_X (which starts in the initial state $|x = 0\rangle = |0 \cdots 0\rangle$), we come to

$$|\Psi'\rangle_{AD_X BD_B} = \frac{1}{2} \sum_{xmk} i^m \psi_m^k(x) |xx\rangle_{QD_X} \otimes |m\rangle_P \otimes |kk\rangle_{BD_B}, \tag{6.46}$$

where we define the coefficients $\psi_m^k(x) = \langle x|\psi_m^k\rangle_Q$.

Tracing over A and B in Eq. (6.46), we arrive at the (classical) joint density matrix for detector D_B and the screen

$$\begin{aligned}\rho_{D_X D_B} &= \frac{1}{4}\sum_{xmk} |\psi_m^k(x)|^2 \, |x\rangle_{D_X}\langle x| \otimes |k\rangle_{D_B}\langle k| \\ &= \frac{1}{2}\sum_{k} \rho_{D_X}^k \otimes |k\rangle_{D_B}\langle k|,\end{aligned} \tag{6.47}$$

where

$$\rho_{D_X}^k = \sum_x p_k(x) \, |x\rangle_{D_X}\langle x|, \tag{6.48}$$

are the conditional states of the screen D_X with corresponding conditional probability distribution

$$p_k(x) = \frac{1}{2}\sum_m |\psi_m^k(x)|^2. \tag{6.49}$$

Tracing out detector D_B from (6.47) yields the full density matrix for D_X

$$\rho_{D_X} = \frac{1}{2}\sum_k \rho_{D_X}^k = \sum_x p(x) \, |x\rangle_{D_X}\langle x|, \tag{6.50}$$

where the total probability distribution of the screen is

$$p(x) = \frac{1}{2}\sum_k p_k(x). \tag{6.51}$$

It is straightforward to show that the total probability distribution $p(x)$ for the screen is completely incoherent due to the cancellation of the cross terms of the two conditional probabilities. That is,

$$p(x) = \frac{1}{2}\Big(|\psi_1(x)|^2 + |\psi_2(x)|^2\Big). \tag{6.52}$$

This distribution describes two overlapping intensity peaks on the screen corresponding to the pattern obtained from each slit individually. From the data as a whole (i.e., when we do not know the outcome of detector D_B) no interference is observed on the screen.

However, suppose we do know the outcome of the polarization measurement of photon B. For an outcome k, the conditional state of the screen D_X is given by Eq. (6.48). To discern the type of interference pattern the probability distribution (6.49) of this density matrix describes, we rewrite the conditional probability in terms of the original coefficients $\psi_j(x) = \langle x|\psi_j\rangle_Q$, which leads to

$$p_k(x) = \frac{1}{2}\Big[|\psi_1(x)|^2 + |\psi_2(x)|^2 + i\,(-1)^k \sin 2\theta \Big(\psi_1(x)\,\psi_2(x)^* - \psi_1(x)^*\,\psi_2(x)\Big)\Big]. \quad (6.53)$$

where we used $U_{1k}\,U_{0k}^* + U_{1k}^*\,U_{0k} = (-1)^k \sin 2\theta$. In general, this expression will describe interference fringes with a visibility that is controlled by the magnitude of the coefficient $\sin 2\theta$ in front of the cross terms. Let us consider two specific cases of the erasure angle θ.

First, $\theta = 0$ corresponds to a measurement of photon B in the linear $|h\rangle$, $|v\rangle$ basis. In this case, expression (6.53) reduces to an incoherent sum

$$\theta = 0: \quad p_k(x) = \frac{1}{2}\Big(|\psi_1(x)|^2 + |\psi_2(x)|^2\Big), \quad (6.54)$$

and describes two overlapping intensity peaks on the screen with no interference. In turn, we have full information about the path of photon A. Second, $\theta = \pi/4$ describes a measurement of photon B in the diagonal $|\pm\rangle = \frac{1}{\sqrt{2}}(|h\rangle \pm |v\rangle)$ basis. In this case, expression (6.53) becomes a perfectly coherent sum

$$\theta = \frac{\pi}{4}: \quad p_k(x) = \frac{1}{2}\Big|\psi_1(x) - i\,(-1)^k \psi_2(x)\Big|^2. \quad (6.55)$$

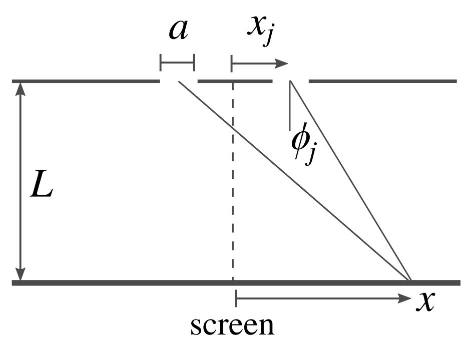

Figure 6.5: Geometry of the double-slit apparatus. The slit width is a, the distance from the origin to the center of slit j is x_j, and the distance between slits is $d = |x_2 - x_1|$. The distance from the slits to the screen is L, while the angle from the center of slit j to point x on the screen is given by $\tan\phi_j = (x - x_j)/L$.

That is, the effect of tagging has been erased since the standard double-slit diffraction pattern can be observed. In general, given an outcome k for detector D_B, the corresponding state of the screen is $\rho^k_{D_X}$ with probability distribution $p_k(x)$. This leads to fringes with a level of visibility that is determined by the erasure angle θ. The distribution for $k = 0$ is phase shifted relative to $k = 1$, so that depending on the state of D_B, one observes either fringes or antifringes. Therefore, measuring photon B in a basis characterized by the angle θ allows one to tune the visibility of the interference fringes from the standard two-slit diffraction to single-slit diffraction [153].

To explicitly compute the interference patterns, we write the coefficients of the jth slit for a photon of wavelength λ as [168]

$$\psi_j(x) = \frac{\sin\alpha}{\alpha}\, e^{-2i\,\alpha\, x_j/a}\,, \tag{6.56}$$

where $\alpha = \pi a \sin\phi_j/\lambda$. The geometry of the double-slit apparatus (see Fig. 6.5) is described by the slit width a, the distance x_j to the center of the jth slit, the angle $\phi_j = \tan^{-1}((x - x_j)/L)$ from the center of slit j to the position x on the screen, and the distance L from the

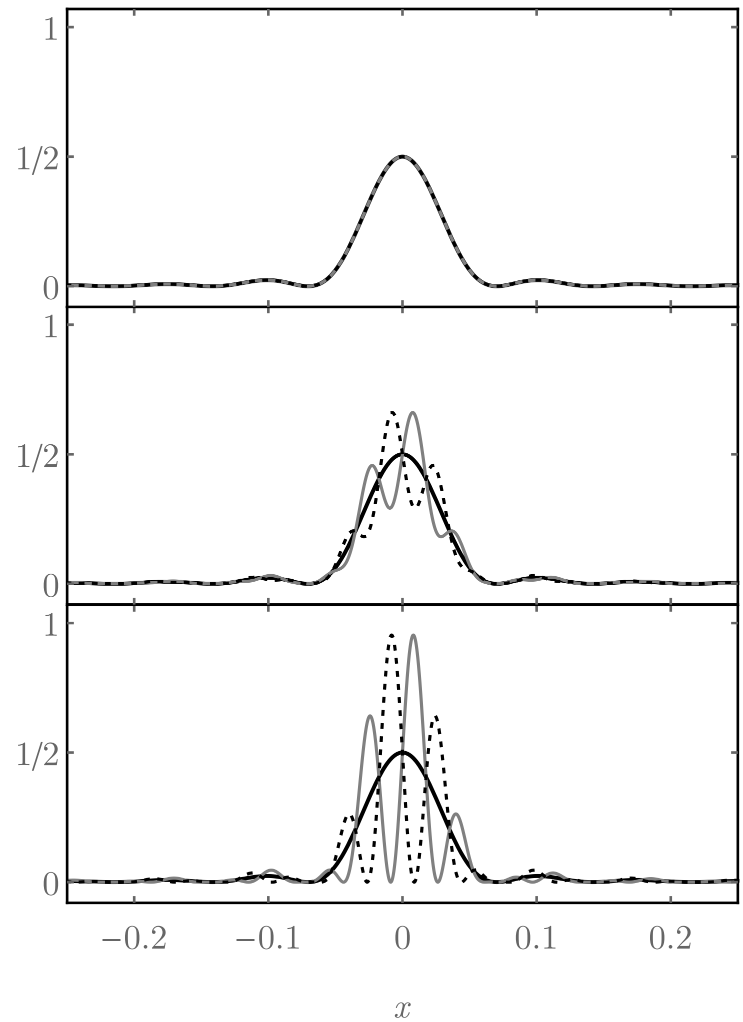

Figure 6.6: Intensity distributions in the quantum eraser. The conditional distributions $p_k(x)$ are plotted as a function of the position on the screen, x, in meters, and normalized so that the maximum is at 1. The solid gray (black dotted) oscillations describe the interference pattern $p_0(x)$ ($p_1(x)$) of photon A conditional on outcome $k = 0$ ($k = 1$) of detector D_B. The solid black line is the total distribution $p(x)$ and is the single-slit diffraction result. Parameters for this specific case are $a = 10\,\mu$m, $d = 20\,\mu$m, $L = 1$ m, and $\lambda = 702$ nm. Top to bottom: probability distributions for three erasure angles $\theta = 0$, $\theta = \pi/16$, $\theta = \pi/4$.

slits to the screen. For a single slit at the origin, the intensity corresponding to detection of a photon at position x on the screen is

$$|\psi_j(x)|^2 = \left|\frac{\sin\alpha}{\alpha}\right|^2 , \tag{6.57}$$

which is the standard result for single-slit Fraunhofer diffraction. For two slits separated by a distance $d = |x_2 - x_1|$, the coefficients for each slit are coherently added. In the far-field limit $L \gg d$, we can use the approximation $\phi_j = \phi = \tan^{-1}(x/L)$. Using the coefficients $\psi_j(x)$ in the expression (6.53) for $p_k(x)$, leads to the interference patterns for the intensity shown in Fig. 6.6. The two patterns $p_0(x)$ and $p_1(x)$ are shifted relative to each other on the screen, and the envelope of each pattern is a single-slit diffraction pattern. We show the distributions for three erasure angles: $\theta = 0$, $\pi/16$, $\pi/4$. For a measurement in the linear $|h\rangle$, $|v\rangle$ basis ($\theta = 0$), there is no interference on the screen, since there is full information about the path of photon A. As the erasure angle increases to $\pi/4$ (a measurement in the diagonal $|\pm\rangle$ basis), the oscillations increase to the level of the usual interference pattern for two-slit diffraction. The solid black line is the total distribution $p(x)$, which is the full data observed in the experiment, and shows no interference. Only by knowing the outcome k of detector D_B can one extract the associated conditional distribution $p_k(x)$ from the full distribution $p(x)$.

6.6 Conclusions

We prefer to tread lightly when using our results to discuss aspects of quantum theory that have been discussed in a controversial manner since the discussions between Bohr and Einstein concerning these matters [169]. Nevertheless, we believe some statements can be made unequivocally. For example, it is now clear (and has been pointed out repeatedly before us), that a quantum system not only carries both particle and wave attributes, but that these quantities are manifested in measurement devices in a fluid manner. In particular, the dynamics of the Bell-state quantum eraser, which allows us to give measurements different "meanings" depending on what state preparation we may choose *after* the state determination has taken place, cannot possibly be consistent with a picture of quantum measurement in which the quantum state is irreversibly projected so as to be consistent with the state of the measurement device. The actuality of not only information erasure, but the production of alternative outcomes via the retroactive manipulation of state preparation, confirms the picture that the wave function after measurement continues to carry amplitudes that are *not* consistent with the state of the measurement device.

That the quantum state can be inconsistent with the state of the measurement device should not come as a surprise to practitioners of quantum information science. After all, the idea of the classical measurement, in which the state of the system to be measured is copied onto the state of the measurement device, cannot carry over to quantum mechanics on account of the no-cloning theorem [36, 37]. Indeed, the central idea of classical measurement—in which the variation of the system is fully correlated with the variation in the device—is impossible for pure quantum states that carry no entropy whatsoever.

Of course, mixed quantum states (pure joint states with a reference state traced out) can carry entropy, and this entropy can be shared with classical measurement devices. This appears to be the case in the construction described here, as the entropy of the system Q is exactly one bit (in the ideal case whose extension was discussed in Sec. 6.4). If the classical device (here the device D_A) cannot carry information about the state of Q, what information does it reflect? In our view, a classical device's state must be consistent with the state of other classical measurement devices, so as to ensure a causally consistent world. Here, the information $S(Q : D_A|D_B)$ predicts the outcome of a measurement of the which-path information that would be obtained if a device was placed squarely in the path of the beam. Of course, such a device would record a random outcome (the photon would be found in state ψ_1 half the time), and D_A would perfectly predict this random outcome. Still, neither of these states predict the state of Q, which after all is neither here nor there. We are thus forced to admit that our classical devices do not (and *cannot*) reveal to us the quantum reality underlying our classical world [170]. However, experimental (and theoretical) ingenuity has allowed us to be aware of our classical device's deceptions, and shown us the path to perhaps design even more clever schemes to lift the veil from the underlying quantum reality.

Chapter 7

Consecutive Quantum Measurements

7.1 Introduction

Consecutive measurements performed on the same quantum system can reveal fundamental insights into quantum theory's causal structure, and probe different aspects of the quantum measurement problem. According to the Copenhagen interpretation, measurements affect the quantum system in such a way that the quantum superposition collapses after the measurement, erasing any knowledge of the prior state. We show here that counter to this view, unamplified measurements (measurements where all variables comprising a pointer are controllable) have coherent ancilla density matrices that encode the memory of the entire set of quantum measurements, and that the quantum chain of a set of consecutive unamplified measurements is non-Markovian. In contrast, sequences of amplified measurements (measurements where at least one pointer variable has been lost) are equivalent to a quantum Markov chain. An analysis of arbitrary non-Markovian quantum chains of measurements reveals that all of the information necessary to reconstruct the chain is encoded on its boundary (the state preparation and the final measurement), reminiscent of the holographic principle.

The physics of consecutive (sequential) measurements on the same quantum system has enjoyed increased attention as of late, as it probes the causal structure of quantum mechanics [171]. It is of interest to researchers concerned about the apparent lack of time-reversal invariance of Born's rule [172,173], as well as to those developing a consistent formulation of covariant quantum mechanics [174,175], which does not allow for a time variable to define the order of (possibly non-commuting) projections [176].

Consecutive measurements can be seen to challenge our understanding of quantum theory in an altogether different manner, however. According to standard theory, a measurement causes the state of a quantum system to "collapse", repreparing it as an eigenstate of the measured operator so that after multiple consecutive measurements on the quantum system any information about the initial preparation is erased. However, recent investigations of sequential measurements on a single quantum system with the purpose of optimal state discrimination have already hinted that quantum information survives the collapse [177,178], and that information about a chain of sequential measurements can be retrieved from the final quantum state [179].

Here we investigate the circumstances that make chains of quantum measurements "Markovian" (meaning that each consecutive measurement "wipes the slate clean" so that retrodiction of quantum states [179] is impossible) and under what conditions the quantum trajectory remains coherent so that the memory of previous measurements is preserved.

In particular, we study the *relative state* of measurement devices (both quantum and classical) in terms of quantum information theory, to ascertain how much information about the quantum state appears in the measurement devices, and how this information is distributed. We find that a crucial distinction refers to the "amplifiability" of a quantum measurement, that is, whether a result is encoded in the states of a closed or an open system, and conclude that a unitary relative-state description makes predictions that are different from a formalism that assumes quantum state reduction.

While the suggestion that the relative state description of quantum measurement [17] (see also [59, 140, 141, 145, 180]) and the Copenhagen interpretation are at odds and may lead to measurable differences has been made before [145, 180], here we frame the problem of consecutive measurements in the language of quantum information theory, which allows us to make these differences manifest.

We begin with the unitary description of quantum measurement [59, 140, 141] discussed previously in Chs. 4 and 5, and apply it in Sec. 7.2 to a sequence of quantum measurements where the pointer—meaning a set of quantum ancilla states—remains under full control of the experimenter. In such a closed system, the pointer can in principle decohere if it is composed of more than one qubit, but this decoherence can be reversed in general. We prove in Theorems 1 and 2 properties of the entropy of a chain of consecutive measurements that imply that the entropy of such chains resides in the last (or first and last) measurements. We then show that for coherence to be preserved in such chains, measurements cannot be arbitrarily amplified—in contrast to the macroscopic measurement devices that are necessarily open systems.

In Sec. 7.3, we analyze sequences of amplifiable—that is, macroscopic—measurements and prove in Theorem 3 that amplified measurement sequences are Markovian. Corollary 3.1 asserts an information-theoretic statement of the general idea that two macroscopic measurements anywhere on a Markov chain must be uncorrelated given the state of all the measurement devices that separate them in the chain. This corollary epitomizes the essence of the Copenhagen idea of quantum state reduction in terms of the conditional independence of measurement devices that are not immediately in each other's past or future. It is consistent with the notion that the measurements collapsed the state of the wave function, erasing any conditional information that a detector could have had about prior measurements. However, no irreversible reduction occurs and all coefficients in the underlying pure-state wave

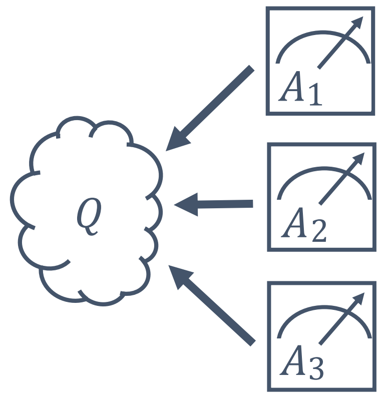

Figure 7.1: Illustration of consecutive measurements of a quantum system, Q, using devices A_1, A_2, and A_3 (numbers indicate the time ordering of the measurement sequence). We will see how the ancillae A_1, A_2, and A_3 become entangled with the quantum system during the measurement process and compute how information is distributed throughout the subsystems. This scheme can be extended to a sequence of n consecutive measurements, which will be considered in the later sections.

function continue to evolve unitarily.

Section 7.4 unifies the two previous sections by proving three statements (Theorems 4, 5, and 6) that relate information-theoretic quantities pertaining to unamplified measurements to the corresponding expressions for amplified measurements. We show that, in general, amplification leads to a loss of information.

After a brief application of the collected concepts and results to standards such as quantum state preparation and the quantum Zeno effect in Sec. 7.5, we close with conclusions.

7.2 Non-Markovian Quantum Measurements

In Ch. 5 we introduced the concept of non-Markovian measurements as those sequences of measurements that are not amplified by macroscopic devices, which we called D. In preparation for Theorem 3 in Sec. 7.3.4 that establishes this correspondence, we first consider consecutive measurements (see Fig. 7.1) with quantum ancillae of prepared and unprepared quantum states (recall Ch. 5.3 for these definitions), and demonstrate the non-Markovian character of the chain of ancillae. Throughout, we assume that the measurements are made on a short enough time scale so that the free evolution of the quantum system during the sequence of measurements can be neglected. In the following, we will use entropy Venn diagrams to study the correlations between subsystems and the distribution of entropies during consecutive measurements.

7.2.1 Consecutive Measurements Prepared Quantum States

Building on the discussion from Ch. 5 where we described a single measurement of a quantum system, we now introduce a second ancilla A_2 that measures Q. This measurement corresponds to a new observable with an eigenbasis $|\widetilde{x}_2\rangle$ that is rotated with respect to the basis of the first observable, $|\widetilde{x}_1\rangle$, via the unitary transformation $U^{(2)}_{x_1x_2} = \langle\widetilde{x}_2|\widetilde{x}_1\rangle$. Unitarity requires that

$$\begin{aligned}\sum_{x_2} U^{(2)}_{x_1x_2}\, U^{(2)*}_{x'_1x_2} &= \delta_{x_1x'_1}\,,\\ \sum_{x_1} U^{(2)}_{x_1x_2}\, U^{(2)*}_{x_1x'_2} &= \delta_{x_2x'_2}\,.\end{aligned} \tag{7.1}$$

After entangling Q and A_2 with an operator analogous to (5.74), the wave function (5.75) evolves to

$$|QA_1A_2\rangle = \sum_{x_1x_2} \alpha^{(1)}_{x_1}\, U^{(2)}_{x_1x_2}\, |\widetilde{x}_2\, x_1x_2\rangle, \tag{7.2}$$

where the orthogonal basis of the second ancilla, A_2, is formed from the states $|x_2\rangle$.

Tracing out Q, the quantum ancillae are correlated according to the joint density matrix,

$$\rho(A_1 A_2) = \sum_{x_1 x'_1 x_2} \alpha^{(1)}_{x_1} \alpha^{(1)*}_{x'_1} U^{(2)}_{x_1 x_2} U^{(2)*}_{x'_1 x_2} |x_1 x_2\rangle\langle x'_1 x_2|, \tag{7.3}$$

while A_1 and A_2 together are entangled with the quantum system. The marginal ancilla density matrices, obtained from (7.3), are

$$\rho(A_i) = \sum_{x_i} q^{(i)}_{x_i} |x_i\rangle\langle x_i|\,, \qquad i = 1, 2 \tag{7.4}$$

where $q^{(1)}_{x_1} = |\alpha^{(1)}_{x_1}|^2$ is the probability distribution of ancilla A_1, while the probability distribution of A_2 is the incoherent sum (a sum of squares), $q^{(2)}_{x_2} = \sum_{x_1} |\alpha^{(1)}_{x_1}|^2\, |U^{(2)}_{x_1 x_2}|^2$. We can compare this expression to the coherent probability distribution (a square of sums), $\big|\sum_{x_1} \alpha^{(1)}_{x_1} U^{(2)}_{x_1 x_2}\big|^2$ for A_2 had the first measurement with A_1 never occurred. The marginal entropy of both A_1 and A_2 is the Shannon entropy $S(A_i) = H[q^{(i)}]$ of the probability distribution $q^{(i)}_{x_i}$. Recall from Chs. 3.2 and 5.3 that we denote the Shannon entropy of a d-dimensional probability distribution p_{x_i} by $H[p] = -\sum_{x_i=0}^{d-1} p_{x_i} \log_d p_{x_i}$, while the von Neumann entropy of a density matrix $\rho(X)$ is defined as $S(X) = S(\rho(X)) = -\mathrm{Tr}\,[\rho(X) \log_d \rho(X)]$. Taking the logarithm to the base d gives entropies the units "dits".

A third measurement of Q with an ancilla A_3 yields

$$|QA_1A_2A_3\rangle = \sum_{x_1 x_2 x_3} \alpha^{(1)}_{x_1} U^{(2)}_{x_1 x_2} U^{(3)}_{x_2 x_3} |\widetilde{x}_3\, x_1 x_2 x_3\rangle, \tag{7.5}$$

where $U^{(3)}_{x_2 x_3} = \langle \widetilde{x}_3 | \widetilde{x}_2 \rangle$ describes the relative transformation between the bases of the third and second observables, and $|x_3\rangle$ are the basis states of ancilla A_3. The quantum system is entangled with all three ancillae in (7.5), as illustrated by the negative conditional entropies

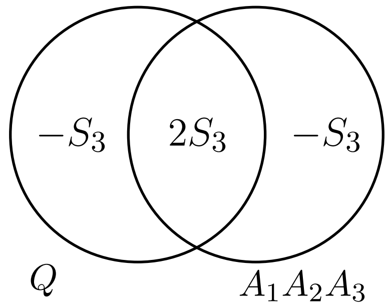

Figure 7.2: Entropy Venn diagram for the state (7.5). The presence of negative conditional entropies reveals that the quantum system, Q, is entangled with all three ancillae, A_1, A_2, and A_3. In this figure, we use the notation $S_3 = S(A_3)$, which is the marginal entropy of the last ancilla A_3 in the measurement sequence. To generalize this diagram from three to n consecutive measurements of a prepared quantum system, S_3 is replaced by S_n, the entropy of the last ancilla, A_n, in the chain.

in Fig. 7.2. The degree of entanglement is controlled by the marginal entropy $S(A_3) = H[q^{(3)}]$ of ancilla A_3, for the probability distribution $q^{(3)}_{x_3} = \sum_{x_1 x_2} |\alpha^{(1)}_{x_1}|^2 \, |U^{(2)}_{x_1 x_2}|^2 \, |U^{(3)}_{x_2 x_3}|^2$. This procedure can be repeated for an arbitrary number of consecutive measurements and can be used to succinctly describe the quantum Zeno and anti-Zeno effects (see Sec. 7.5.1). In the following sections we will quantify how information is distributed in states like (7.5), and how that distribution changes when the measurements are amplified.

7.2.2 Consecutive Measurements of Unprepared Quantum States

Sequential measurements of an unprepared quantum system yield entropy distributions between the quantum system and ancillae that are different from those created by measurements of prepared quantum systems (see Sec. 7.2.1). In this section, we consider a sequence of measurements of an unprepared quantum system that is initially entangled with a reference system as in (5.79). Adding to the calculations in Sec. 5.3.2, we measure Q again in a rotated basis $U^{(2)}_{x_1 x_2} = \langle \widetilde{x}_2 | \widetilde{x}_1 \rangle$, by entangling it with an ancilla A_2. Then, with $|x_2\rangle$ the basis states of ancilla A_2, the wave function (5.81) becomes

$$|QRA_1A_2\rangle = \frac{1}{\sqrt{d}}\sum_{x_1x_2} U^{(2)}_{x_1x_2}\,|\widetilde{x}_2\,x_1x_1x_2\rangle\,. \tag{7.6}$$

It is straightforward to show that the marginal ancilla density matrices are maximally mixed, $\rho(A_1){=}\rho(A_2){=}1/d\,\mathbb{1}$, where $\mathbb{1}$ is the identity matrix of dimension d. It follows that A_1 and A_2 have maximum entropy $S(A_1) = S(A_2) = 1$. Recall that all logarithms are taken to the base d, giving entropies units of *dits*. If $d = 2$, the units are *bits*. The joint state of A_1 and A_2 is diagonal in the ancilla product basis,

$$\rho(A_1A_2) = \frac{1}{d}\sum_{x_1}|x_1\rangle\langle x_1|\otimes\sum_{x_2}|U^{(2)}_{x_1x_2}|^2|x_2\rangle\langle x_2|\,, \tag{7.7}$$

in contrast to Eq. (7.3). Still, the quantum ancillae A_1 and A_2 are correlated. Equations (7.3) and (7.7) immediately imply that if the quantum system is measured repeatedly in the same basis ($U^{(2)}_{x_1x_2} = \delta_{x_1x_2}$) by independent devices, all of those devices will be perfectly correlated and will reflect the same outcome [59, 140].

Let us entangle a third ancilla, A_3, with the quantum system and perform a measurement of an observable with eigenbasis rotated via $U^{(3)}_{x_2x_3} = \langle\widetilde{x}_3|\widetilde{x}_2\rangle$. We find that (7.6) evolves to

$$|QRA_1A_2A_3\rangle = \frac{1}{\sqrt{d}}\sum_{x_1x_2x_3} U^{(2)}_{x_1x_2}\,U^{(3)}_{x_2x_3}\,|\widetilde{x}_3\,x_1x_1x_2x_3\rangle. \tag{7.8}$$

The entropic relationships between the variables Q, R, and $A_1A_2A_3$ are shown in Fig. 7.3. The zero ternary mutual entropy, $S(Q : R : A_1A_2A_3) = 0$, indicates that the entropy $S(A_1A_3) = S_{13}$ that is shared by R and $A_1A_2A_3$ is not shared with the quantum system. Tracing out the reference state, we find that the quantum system is entangled with all three ancillae. However, this entanglement is now shared with the reference system, which yields

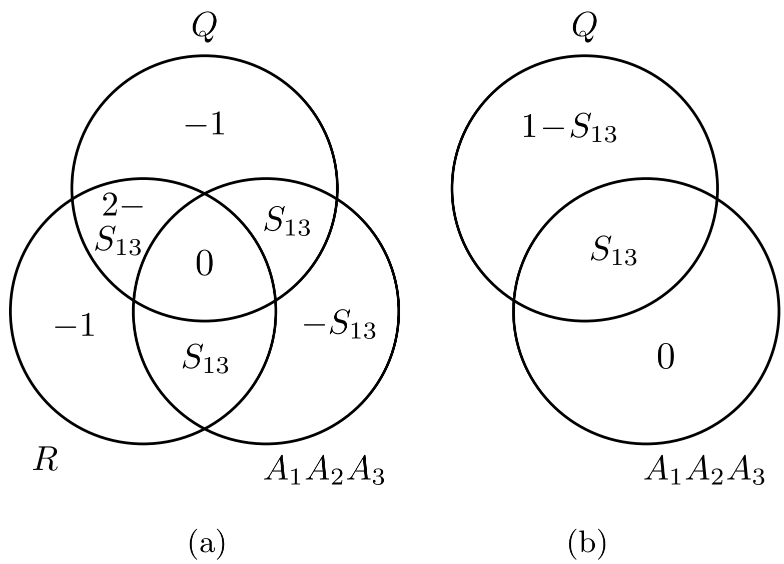

Figure 7.3: Entropy Venn diagrams for the pure state (7.8), where $S_{13} = S(A_1A_3)$ is the joint entropy of Eq. (7.11). (a) The entropy S_{13} that is shared by the reference state, R, and the chain of ancillae, $A_1A_2A_3$, is not shared with the quantum system, Q, since the ternary mutual entropy vanishes: $S(Q : R : A_1A_2A_3) = 0$. (b) Tracing out the reference leaves Q entangled with all ancillae. To generalize these diagrams from three to n consecutive measurements of an unprepared quantum system, S_{13} is replaced by S_{1n}, the joint entropy of the first and last ancillae, A_1 and A_n, in the chain.

a Venn diagram that is different from Fig. 7.2.

Consecutive measurements provide a unique opportunity to extract information about the state of the quantum system from the correlations created between the ancillae, as we do not directly observe either the quantum system or the reference. Tracing out Q and R from the full density matrix associated with Eq. (7.8) yields the joint state of the three ancillae,

$$\rho(A_1A_2A_3) = \frac{1}{d}\sum_{x_1} |x_1\rangle\langle x_1| \otimes \sum_{x_2x_2'} U^{(2)}_{x_1x_2}\, U^{(2)*}_{x_1x_2'}\, |x_2\rangle\langle x_2'| \otimes \sum_{x_3} U^{(3)}_{x_2x_3}\, U^{(3)*}_{x_2'x_3}\, |x_3\rangle\langle x_3|. \quad (7.9)$$

Unlike the pairwise state $\rho(A_1A_2)$ in Eq. (7.7), the state of all three ancillae is not an incoherent mixture. Performing a third measurement has, in a sense, revived the coherence of the A_2 subsystem.

An apparent collapse has taken place after the second consecutive measurement in Eq. (7.7) as the corresponding density matrix has no off-diagonal terms. However, the third measurement seemingly *undoes* this projection, as can be seen from the appearance of off-diagonal terms in Eq. (7.9). This "reversal" is different from protocols that can "un-collapse" weak measurements [181,182], because it is clear that the wave function (7.8) underlying the density matrix (7.9) remains a pure state. The presence of the cross terms in Eq. (7.9) has fundamental consequences for our understanding of the measurement process, and may open up avenues for developing new quantum protocols. In particular, the cross terms in Eq. (7.9) enable the implementation of disentangling protocols [183] (see Ch. 8).

As mentioned in Sec. 5.3.3, the ancilla A_i may be composed of a large number of qudits. To account for a possibly macroscopic ancilla, we suppose that n qudits $A_i^{(1)} \cdots A_i^{(n)}$, which comprise the ith ancilla A_i, measure the quantum system. In this case, the joint density matrix (7.9) is extended to

$$\begin{aligned}\rho(A_1A_2A_3) = \frac{1}{d}\sum_{x_1} |x_1\ldots x_1\rangle\langle x_1\ldots x_1| &\otimes \sum_{x_2x_2'} U^{(2)}_{x_1x_2}\,U^{(2)*}_{x_1x_2'}\,|x_2\ldots x_2\rangle\langle x_2'\ldots x_2'| \\ &\otimes \sum_{x_3} U^{(3)}_{x_2x_3}\,U^{(3)*}_{x_2'x_3}\,|x_3\ldots x_3\rangle\langle x_3\ldots x_3|.\end{aligned} \tag{7.10}$$

In principle, accounting for macroscopic ancillae does not destroy the coherence of the joint state (7.10), which is concentrated in the A_2 subsystem. The coherence is protected as long as no qudits in the intermediate ancilla A_2 are 'lost', implying a trace over their states, which removes all off-diagonal terms. In practical implementations, it may be effectively impossible to prevent decoherence when the number of qudits is sufficiently large. On the other hand, the pairwise density matrices $\rho(A_1A_2)$, $\rho(A_2A_3)$, and $\rho(A_1A_3)$ are unaffected by a loss of qudits as they are already diagonal. In addition, it can be easily shown that the coherence in Eqs. (7.9) and (7.10) is fully destroyed if just the A_2 measurement is amplified by a device

D_2 (as we will see in Sec. 7.3.3). That is, amplification of the first and last ancillae has no effect on the coherence of (7.9) and (7.10).

From the joint ancilla density matrix (7.9), we now derive several properties of the chain of quantum ancillae and summarize them using an entropy Venn diagram between A_1, A_2, and A_3. First, we construct all three pairwise ancilla density matrices and compute their entropies. Tracing out A_3 from the joint density matrix (7.9) recovers $\rho(A_1A_2)$ in Eq. (7.7), as it should because the interaction between Q and A_3 does not influence the past interactions of Q with A_1 and A_2. Tracing over A_2 in Eq. (7.9) gives

$$\rho(A_1A_3) = \frac{1}{d}\sum_{x_1} |x_1\rangle\langle x_1| \otimes \sum_{x_2x_3} |U^{(2)}_{x_2x_3}|^2\, |U^{(3)}_{x_2x_3}|^2\, |x_3\rangle\langle x_3|, \tag{7.11}$$

while tracing over A_1 yields

$$\rho(A_2A_3) = \frac{1}{d}\sum_{x_2} |x_2\rangle\langle x_2| \otimes \sum_{x_3} |U^{(3)}_{x_2x_3}|^2\, |x_3\rangle\langle x_3|\ . \tag{7.12}$$

All three pairwise density matrices are diagonal in the ancilla product basis (see Theorem 2 in Sec. 7.2.3 for a general proof). We take "diagonal in the ancilla product basis" to be synonymous with "classical". From Eqs. (7.7), (7.11), and (7.12), we can calculate the entropy of each pair of ancillae and of the joint state of all three ancillae from Eq. (7.9). The pairwise entropies are

$$S(A_1A_2) = 1 - \frac{1}{d}\sum_{x_1x_2} |U^{(2)}_{x_1x_2}|^2 \log_d |U^{(2)}_{x_1x_2}|^2, \tag{7.13}$$

$$S(A_2A_3) = 1 - \frac{1}{d}\sum_{x_2x_3} |U^{(3)}_{x_2x_3}|^2 \log_d |U^{(3)}_{x_2x_3}|^2, \tag{7.14}$$

$$S(A_1A_3) = 1 - \frac{1}{d}\sum_{x_1x_3} |\beta^{(13)}_{x_1x_3}|^2 \log_d |\beta^{(13)}_{x_1x_3}|^2. \tag{7.15}$$

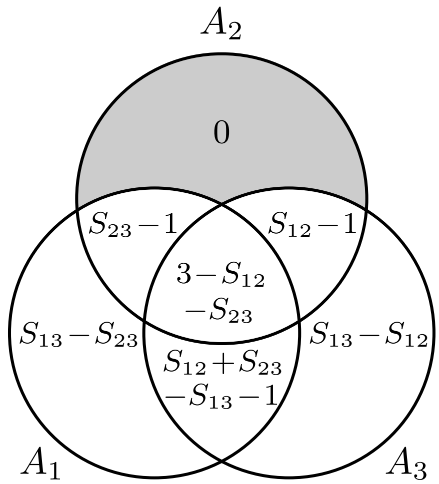

Figure 7.4: Entropy relationships for three quantum ancillae that measured an unprepared quantum system. The vanishing conditional entropy, $S(A_2|A_1A_3) = 0$, is a consequence of the property that the entropy of the entire ancilla chain is equal to the entropy of the borders, $S(A_1A_2A_3) = S(A_1A_3)$. This will be discussed more in Sec. 7.2.3. In this figure, $S(A_iA_j) = S_{ij}$ denotes the pairwise entropy of any two ancillae A_i and A_j.

where $|\beta^{(13)}_{x_1x_3}|^2 = \sum_{x_2} |U^{(2)}_{x_1x_2}|^2 |U^{(3)}_{x_2x_3}|^2$. Furthermore, it is straightforward to show that $S(A_1A_2A_3)$, the entropy of $\rho(A_1A_2A_3)$ in Eq. (7.9), is equal to $S(A_1A_3)$. This equality holds for any set of three consecutive measurements in an arbitrarily-long measurement chain as we will later prove in Theorem 2 of Sec. 7.2.3. With these joint entropies, we construct the entropy Venn diagram for the three ancillae that consecutively measured an unprepared quantum system, as shown in Fig. 7.4.

We apply the formalism presented thus far to the specific case of qubits (Hilbert space dimension $d = 2$). The eigenbasis of the second observable that is measured using ancilla A_2 is rotated relative to the basis of the first observable by an angle θ_2. Similarly, the eigenbasis of the third observable measured using A_3 is at an angle θ_3 relative to the second observable. If we consider observables with eigenbases in the xz plane of the Bloch sphere, we can implement the basis transformations with the rotation matrix,

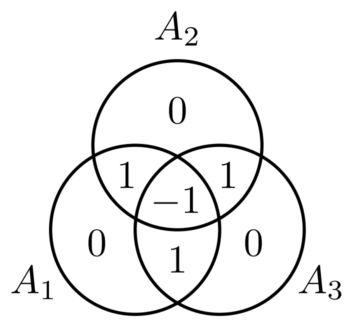

Figure 7.5: Diagram for three qubit ancillae that measured an unprepared quantum system. The eigenbasis of the second observable is at an angle $\theta_2 = \pi/4$ relative to the basis of the first observable and likewise for the third observable.

$$U^{(i)} = \begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix} . \tag{7.16}$$

For measurements at $\theta_2 = \theta_3 = \pi/4$, for example, we have $|U^{(2)}_{x_1 x_2}|^2 = |U^{(3)}_{x_2 x_3}|^2 = 1/2$, and we expect each ancilla to be maximally entropic: $S(A_1) = S(A_2) = S(A_3) = 1$ bit. The joint entropy of each pair of ancillae is two bits, as can be read off of Eqs. (7.13-7.15). Because of the non-diagonal nature of $\rho(A_1A_2A_3)$ in Eq. (7.9), the joint density matrix of the three ancillae (using σ_z, the third Pauli matrix, and $\mathbb{1}$, the 2×2 identity matrix),

$$\rho(A_1A_2A_3) = \frac{1}{8} \begin{pmatrix} \mathbb{1} & -\sigma_z & 0 & 0 \\ -\sigma_z & \mathbb{1} & 0 & 0 \\ 0 & 0 & \mathbb{1} & \sigma_z \\ 0 & 0 & \sigma_z & \mathbb{1} \end{pmatrix} , \tag{7.17}$$

has entropy $S(A_1A_2A_3) = 2$ bits, as can be checked by finding the eigenvalues of (7.17). Figure 7.5 summarizes the entropic relationships for unamplified consecutive qubit measurements at $\theta_2 = \theta_3 = \pi/4$.

It is instructive to note that the Venn diagram in Fig. 7.5 is the same as the one obtained for a one-time binary cryptographic pad where two classical binary variables (the source and the key) are combined to a third (the message) via a controlled-NOT operation [184] (the density matrices underlying the Venn diagrams are very different, however). The Venn diagram implies that the state of any one of the three quantum ancillae can be predicted from knowing the *joint* state of the two others. However, the prediction of A_3, for example, cannot be achieved using expectation values from A_2's and A_1's states separately, as the diagonal of Eqs. (7.9) and (7.17) corresponds to a uniform probability distribution. Thus, quantum coherence can be seen to encrypt classical information about past states.

7.2.3 Coherence of the Chain of Unamplified Measurements

So far we have seen that the joint ancilla density matrices describing unamplified measurements generally contain a non-vanishing degree of coherence. This suggests that coherence is not lost in the measurement sequence, but is actually contained in specific ancilla subsystems. In this section, we extend our unitary description of consecutive unamplified measurements of a quantum system to an arbitrarily-long chain of ancillae, and derive several properties of the measurement chain.

Many of the joint ancilla density matrices that we have encountered in describing consecutive quantum measurements are so-called "classical-quantum states" (recall Ch. 3.4.2, or see Ref. [76]). Such states have a block-diagonal structure of the form $\rho = \sum_i p_i \, \rho_i \otimes |i\rangle\langle i|$, where the density matrix ρ_i appears with probability p_i. In addition, the ancilla states that we derive here have the property that the density matrices ρ_i are always pure quantum superpositions (we defined the purity of a quantum state in Ch. 2.2, which is 1 for a pure state and $1/d$ for a completely mixed state of dimension d).

For measurements of a prepared quantum system, classical-quantum states occur in the joint density matrices of two or more consecutive ancillae. For instance, recall the state $\rho(A_1 A_2)$ from Eq. (7.3) that resulted from two measurements of a prepared quantum system. We can diagonalize this state with the set of non-orthogonal states $\alpha^{(2)}_{x_2}\,|\psi_{x_2}\rangle = \sum_{x_1} \alpha^{(1)}_{x_1}\, U^{(2)}_{x_1 x_2}\,|x_1\rangle$ for subsystem A_1, so that (7.3) appears as

$$\rho(A_1 A_2) = \sum_{x_2} q^{(2)}_{x_2}\,|\psi_{x_2}\rangle\langle\psi_{x_2}| \otimes |x_2\rangle\langle x_2|, \tag{7.18}$$

where the normalization is equal to the probability distribution for the second ancilla A_2,

$$q^{(2)}_{x_2} = |\alpha^{(2)}_{x_2}|^2 = \sum_{x_1} |\alpha^{(1)}_{x_1}|^2\,|U^{(2)}_{x_1 x_2}|^2. \tag{7.19}$$

On the other hand, classical-quantum states occur for measurements of unprepared quantum systems when there are at least *three* consecutive measurements, as the first measurement in that sequence can be viewed as the state preparation. For example, Eq. (7.9) can be diagonalized with the set of non-orthogonal states $\beta^{(13)}_{x_1 x_3}\,|\phi_{x_1 x_3}\rangle = \sum_{x_2} U^{(2)}_{x_1 x_2}\, U^{(3)}_{x_2 x_3}\,|x_2\rangle$ for the A_2 subsystem, so that

$$\rho(A_1 A_2 A_3) = \frac{1}{d}\sum_{x_1 x_3} p^{(13)}_{x_1 x_3}\,|x_1 x_3\rangle\langle x_1 x_3| \otimes |\phi_{x_1 x_3}\rangle\langle\phi_{x_1 x_3}|, \tag{7.20}$$

where the normalization is

$$p^{(13)}_{x_1 x_3} = |\beta^{(13)}_{x_1 x_3}|^2 = \sum_{x_2} |U^{(2)}_{x_1 x_2}|^2\,|U^{(3)}_{x_2 x_3}|^2. \tag{7.21}$$

Evidently, from (7.18) and (7.20), each density matrix ρ_i in the general state $\rho = \sum_i p_i\,\rho_i \otimes |i\rangle\langle i|$ corresponds to a pure state in our ancilla density matrices. This leads to an interesting observation that the entropy of a chain of ancillae is contained in either just the last device or in both the first and last devices together. In the first example above for $\rho(A_1A_2)$, it is straightforward to show using Eq. (7.18) that $S(A_1A_2) = S(A_2)$. That is, the entropy of the sequence A_1A_2 is found at the end of the chain, A_2. From the definition of conditional entropy [58], it follows that the entropy of A_1 vanishes (it is in the pure state $|\psi_{x_2}\rangle$), given the state of A_2:

$$S(A_1|A_2) = S(A_1A_2) - S(A_2) = 0. \tag{7.22}$$

In the second example above for $\rho(A_1A_2A_3)$, we find from Eq. (7.20) that $S(A_1A_3) = S(A_1A_2A_3)$. In other words, the entropy of the chain resides in the boundary, A_1 and A_3. It follows that, given the joint state of A_1 and A_3, A_2's state has zero entropy (see the gray region in Fig. 7.4) and is fully determined (it is in the pure state $|\phi_{x_1x_3}\rangle$):

$$S(A_2|A_1A_3) = S(A_1A_2A_3) - S(A_1A_3) = 0\,. \tag{7.23}$$

In the following Theorems 1 and 2, we extend these results to an arbitrarily-long chain of quantum ancillae. These findings are important as they show that unamplified measurement chains retain a finite amount of coherence. Specifically, for measurements on unprepared quantum states, the coherence is contained in all ancillae up to the last, while for unprepared quantum states it is contained in all ancillae except for the boundary.

To begin, we define (ancilla) random variables A_i that take on states x_i with probabilities $q^{(i)}_{x_i}$. Each ancilla has d orthogonal states and the set of outcomes for the ith ancilla is labeled by the index x_i, where $x_i = 0, \ldots, d-1$.

Theorem 1. *The density matrix describing $j+1$ ancillae that consecutively measured a prepared quantum system is a classical-quantum state such that its joint entropy is contained only in the last device in the measurement chain. That is,*

$$S(A_1 \ldots A_{j+1}) = S(A_{j+1}). \tag{7.24}$$

Proof. Generalizing the result (7.5), the wave function $|\Psi\rangle = |QA_1 \dots A_{j+1}\rangle$ for $j+1$ consecutive measurements of a prepared quantum state is

$$|\Psi\rangle = \sum_{x_1 \dots x_{j+1}} \alpha^{(1)}_{x_1} \, U^{(2)}_{x_1 x_2} \dots U^{(j+1)}_{x_j x_{j+1}} \, |\widetilde{x}_{j+1} \, x_1 x_2 \dots x_j x_{j+1}\rangle. \tag{7.25}$$

The first ket $|\widetilde{x}_{j+1}\rangle$ in the joint state on the right hand side of (7.25) describes the quantum system, which is written in the eigenbasis of the last measured observable. Each A_i measures an observable with an eigenbasis that is rotated relative to the basis of the previous observable such that $U^{(i)}_{x_{i-1},x_i} = \langle \widetilde{x}_i | \widetilde{x}_{i-1}\rangle$. The unitarity of $U^{(i)}$ requires that

$$\begin{aligned} \sum_{x_{i-1}} U^{(i)}_{x_{i-1}x_i} U^{(i)*}_{x_{i-1}x'_i} &= \delta_{x_i x'_i} \, , \\ \sum_{x_i} U^{(i)}_{x_{i-1}x_i} U^{(i)*}_{x'_{i-1}x_i} &= \delta_{x_{i-1} x'_{i-1}} \, . \end{aligned} \tag{7.26}$$

Recasting expression (7.25) in terms of the following set of non-orthogonal states,

$$\alpha^{(j+1)}_{x_{j+1}} \, |\psi_{x_{j+1}}\rangle = \sum_{x_1 \cdots x_j} \alpha^{(1)}_{x_1} \, U^{(2)}_{x_1 x_2} \cdots \, U^{(j+1)}_{x_j x_{j+1}} \, |x_1 \cdots x_j\rangle, \tag{7.27}$$

yields

$$|\Psi\rangle = \sum_{x_{j+1}} \alpha^{(j+1)}_{x_{j+1}} \, |\widetilde{x}_{j+1} \; \psi_{x_{j+1}} \; x_{j+1}\rangle. \tag{7.28}$$

This is not a true tripartite Schmidt decomposition [74] as the states $|\psi_{x_{j+1}}\rangle$ are not orthogonal: the partial inner product $\langle \psi_{x_{j+1}} | \Psi\rangle$ does not give a state with a Schmidt number of one. Although the states $|\psi_{x_{j+1}}\rangle$ are not orthogonal, they are normalized according to

$$q^{(j+1)}_{x_{j+1}} = |\alpha^{(j+1)}_{x_{j+1}}|^2 = \sum_{x_1 \cdots x_j} |\alpha^{(1)}_{x_1}|^2 \; |U^{(2)}_{x_1 x_2}|^2 \cdots \; |U^{(j+1)}_{x_j x_{j+1}}|^2, \tag{7.29}$$

which is the probability distribution of ancilla A_{j+1}.

Tracing out the quantum system from the density matrix $|\Psi\rangle\langle\Psi|$ formed from (7.28), the state of all $j+1$ ancillae can be written as

$$\rho(A_1 \dots A_{j+1}) = \sum_{x_{j+1}} q_{x_{j+1}}^{(j+1)} \; |\psi_{x_{j+1}}\rangle\langle\psi_{x_{j+1}}| \otimes |x_{j+1}\rangle\langle x_{j+1}|. \tag{7.30}$$

This state is non-diagonal in the ancilla product basis $|x_1 \cdots x_{j+1}\rangle$, but is diagonalized by (7.27). The density matrix (7.30) is a classical-quantum state where the first j ancillae are in the pure state $|\psi_{x_{j+1}}\rangle$. In the following, we will compute the entropy of the joint state (7.30) and show that it is equal to the entropy of the last ancilla, A_{j+1}.

The appearance of classical-quantum states in the sequence of measurements leads to the interesting (and perhaps surprising) observation that the joint entropy of all ancillae in Eq. (7.30) resides only in the last device in the measurement chain. Since the joint state $|\psi_{x_{j+1}}\rangle \otimes |x_{j+1}\rangle$ is orthonormal, it is easy to see that the entropy of (7.30) is equal to the Shannon entropy of the probability distribution $q_{x_{j+1}}^{(j+1)}$. This is equivalent to the entropy of the last ancilla, so that

$$S(A_1 \dots A_{j+1}) = S(A_{j+1}). \tag{7.31}$$

$\square$

Note that this implies that there is an upper bound to the joint entropy: $\max[S(A_1 \dots A_{j+1})] = \max[S_{j+1}] = 1$.

From this property, it immediately follows that the entropy of the first j ancillae, condi-

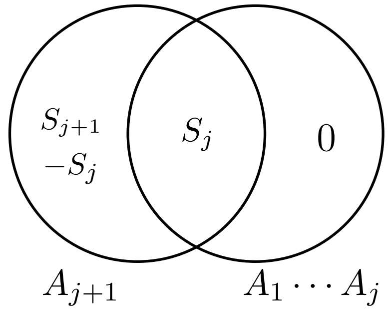

Figure 7.6: Diagram for the unamplified measurement sequence with ancillae A_1, A_2, ... , A_j, A_{j+1}. According to (7.30), the joint entropy of the first j ancillae vanishes when given A_{j+1}, since the entropy resides only at the end of the chain. In this figure, we use the notation $S_k = S(A_k)$ for the marginal entropy of the kth ancilla.

tional on the state of the last ancilla, vanishes,

$$S(A_1 \dots A_j | A_{j+1}) = S(A_1 \dots A_{j+1}) - S(A_{j+1}) = 0. \tag{7.32}$$

Therefore, if the state of the end of the measurement chain is known, then all preceding ancillae exist in a pure quantum superposition: The state of $A_1 \dots A_j$ is fully determined (a zero entropy state), given A_{j+1}. This implies that the entropy of all ancillae in an arbitrarily-long sequence of measurements resides only at the end of the chain. The entropy Venn diagram for these two subsystems is shown in Fig. 7.6.

Theorem 2. *For $j+1$ consecutive measurements of an unprepared quantum system, where the reference is traced out, the density matrix for three or more consecutive ancillae is a classical-quantum state such that its joint entropy is contained only in the first and last device of the measurement chain. That is,*

$$S(A_{i-1} A_i \dots A_j A_{j+1}) = S(A_{i-1} A_{j+1}). \tag{7.33}$$

Proof. Generalizing the result (7.8), the wave function $|\Psi'\rangle = |QRA_1 \dots A_{j+1}\rangle$ of $j+1$

ancillae that consecutively measured an unprepared quantum state is

$$|\Psi'\rangle = \frac{1}{\sqrt{d}} \sum_{x_1 \dots x_{j+1}} U^{(2)}_{x_1 x_2} \dots U^{(j+1)}_{x_j x_{j+1}} |\widetilde{x}_{j+1}\, x_1\, x_1 x_2 \dots x_{j+1}\rangle. \tag{7.34}$$

Of the full set of consecutive measurements, consider the subset $A_{i-1}, A_i, \dots, A_j, A_{j+1}$, where $1 < i < j$. Tracing out Q, the reference, and all other ancilla states from the full density matrix $|\Psi'\rangle\langle\Psi'|$, and using the unitarity of each $U^{(i)}$ as stated in Eq. (7.26), the density matrix for this subset can be written as

$$\rho(A_{i-1} \dots A_{j+1}) = \frac{1}{d} \sum_{\substack{x_{i-1} \\ x_{j+1}}} p^{(i-1,j+1)}_{x_{i-1} x_{j+1}}\, |x_{i-1}\rangle\langle x_{i-1}| \otimes |\phi_{x_{i-1} x_{j+1}}\rangle\langle\phi_{x_{i-1} x_{j+1}}| \otimes |x_{j+1}\rangle\langle x_{j+1}|. \tag{7.35}$$

In the following, we will compute the entropy of this state and show that it is equal to the entropy of $\rho(A_{i-1} A_{j+1})$.

The density matrix (7.35) is a classical-quantum state with the intermediate ancillae $A_i, \dots, A_j$ in the pure state $|\phi_{x_{i-1} x_{j+1}}\rangle$. In the ancilla product basis $|x_{i-1} x_i \dots x_j x_{j+1}\rangle$, this matrix is block-diagonal due to the non-diagonality of the subsystem $A_i, \dots, A_j$. However, it is diagonalized by the non-orthogonal states

$$\beta^{(i-1,j+1)}_{x_{i-1} x_{j+1}}\, |\phi_{x_{i-1} x_{j+1}}\rangle = \sum_{x_i \cdots x_j} U^{(i)}_{x_{i-1} x_i} \dots U^{(j+1)}_{x_j x_{j+1}}\, |x_i \dots x_j\rangle, \tag{7.36}$$

which are normalized according to

$$p^{(i-1,j+1)}_{x_{i-1} x_{j+1}} = |\beta^{(i-1,j+1)}_{x_{i-1} x_{j+1}}|^2 = \sum_{x_i \cdots x_j} |U^{(i)}_{x_{i-1} x_i}|^2 \dots |U^{(j+1)}_{x_j x_{j+1}}|^2. \tag{7.37}$$

These normalization coefficients obey the sum rule

$$\sum_{x_{i-1}} p^{(i-1,j+1)}_{x_{i-1}x_{j+1}} = \sum_{x_{j+1}} p^{(i-1,j+1)}_{x_{i-1}x_{j+1}} = 1\,. \tag{7.38}$$

The density matrix for any two ancillae is already diagonal in the ancilla product basis (it is classical). For example, the joint state of A_{i-1} and A_{j+1} is

$$\rho(A_{i-1}A_{j+1}) = \frac{1}{d}\sum_{\substack{x_{i-1}\\x_{j+1}}} p^{(i-1,j+1)}_{x_{i-1}x_{j+1}}\ |x_{i-1}x_{j+1}\rangle\langle x_{i-1}x_{j+1}|, \tag{7.39}$$

so that its entropy reduces to the Shannon entropy $H[p^{(i-1,j+1)}/d]$ of the distribution $p^{(i-1,j+1)}_{x_{i-1}x_{j+1}}/d$. However, the density matrix for three or more consecutive ancillae corresponds to a classical-quantum state (7.35). This state has non-zero coherence that is contained in the subsystem of the intermediate ancillae, which are in the (non-orthogonal) pure state $|\phi_{x_{i-1}x_{j+1}}\rangle$. Since the joint state $|x_{i-1}\rangle \otimes |\phi_{x_{i-1}x_{j+1}}\rangle \otimes |x_{j+1}\rangle$ is still orthonormal, it is straightforward to show that the entropy of (7.35) is equal to the (Shannon) entropy of (7.39), despite the fact that the underlying state (7.35) is non-classical:

$$S(A_{i-1}A_i \ldots A_j A_{j+1}) = S(A_{i-1}A_{j+1}). \tag{7.40}$$

$\square$

It follows directly that the entropy of the intermediate ancillae $A_i, \ldots, A_j$ vanishes when

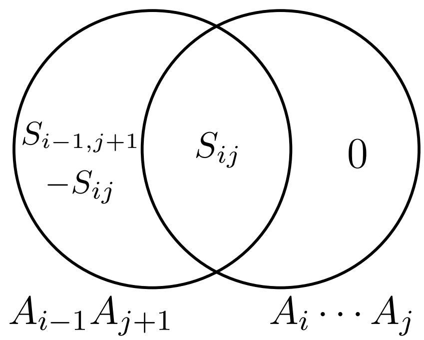

Figure 7.7: Diagram for an unamplified measurement sequence with ancillae A_{i-1}, A_i, ... , A_j, A_{j+1}. According to (7.35), the entropy of all intermediate ancillae, $A_i, \ldots, A_j$, vanishes when given A_{i-1} and A_{j+1}, since the entropy resides at the boundary of the chain. In this figure, we use the notation $S_{k\ell} = S(A_k A_\ell)$ for the pairwise entropy of any two ancillae A_k and A_ℓ.

given the joint state of the ancillae A_{i-1} and A_{j+1},

$$\begin{aligned} S(A_i \ldots A_j | A_{i-1} A_{j+1}) &= S(A_{i-1} A_i \ldots A_j A_{j+1}) \\ &\quad - S(A_{i-1} A_{j+1}) \\ &= 0. \end{aligned} \tag{7.41}$$

Evidently, if the state of the *boundary* of the chain is known, then the intermediate ancillae exist in a pure quantum superposition. The joint state of $A_i, \ldots, A_j$ is fully determined (a zero-entropy state), given the joint state of A_{i-1} that measured Q in the past, together with A_{j+1} that measured Q in the future. Thus, for measurements on unprepared quantum systems, the entropy of an arbitrarily-long ancilla chain is found only in its boundary. The entropy Venn diagram for the boundary and the bulk of the measurement chain is shown in Fig. 7.7.

That the entropy of a chain of measurements is determined entirely by the entropy of the chain's boundary may seem remarkable, but is reminiscent of the holographic principle [185–187]. Indeed, it is conceivable that an extension of the one-dimensional quantum chains we discussed here to tensor networks [188] could make this correspondence more precise [189]. We contrast this result with the previous Theorem 1 for measurements on prepared quantum systems, where the entropy resided only at the end of the chain since the preparation was already known.

7.3 Markovian Quantum Measurements

The non-Markovian measurements we have been discussing up to this point are potentially fragile: while the pointers can consist of many subsystems (even a macroscopic number), the entanglement they potentially display with other quantum systems will be lost even if only a single qudit escapes our control (and therefore, mathematically speaking, must be traced over). In this section we discuss a second step within von Neumann's second stage of quantum measurement, where we observe the fragile quantum ancilla using a secondary observer. While this quantum "observer of the observer" also potentially consists of many different subsystems, it is robust in the sense that tracing over any of the degrees of freedom making up the pointer variable does not modify the relative state of the pointer and the quantum system or other devices.

7.3.1 Amplifying Quantum Measurements

To amplify a measurement, we observe the first quantum observer (denoted by A_1) by measuring A_1 with a device D_1. This additional interaction with the first ancilla in (5.75) leads to the tripartite entangled state

$$|QA_1D_1\rangle = \mathbb{1}_Q \otimes U_{A_1D_1} |QA_1\rangle |M_i\rangle_{D_1} = \sum_{x_1} \alpha^{(1)}_{x_1} |\widetilde{x}_1 x_1 x_1\rangle, \tag{7.42}$$

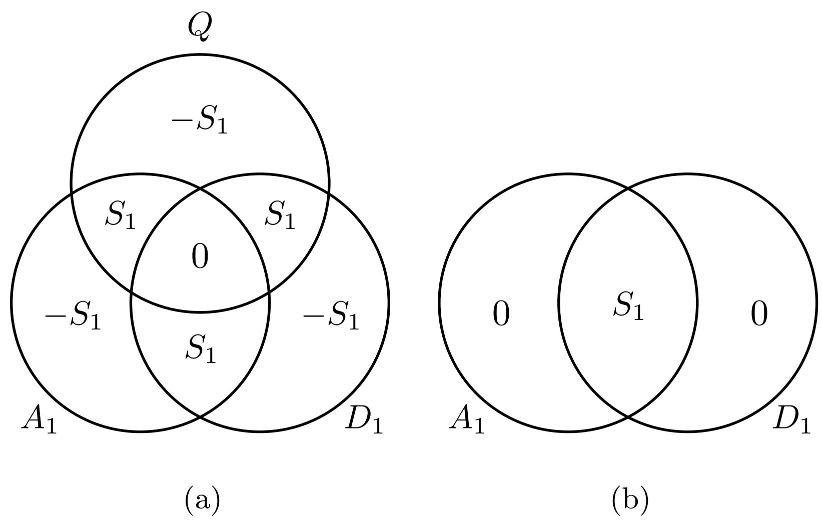

Figure 7.8: Effects of amplification for (a) the tripartite entangled state (7.42). (b) Tracing over the quantum system, A_1 and D_1 are perfectly correlated as in Eq. (7.43). The $S(A_1 : D_1) = S_1$ bits of information gained in the measurement are not shared with the quantum system since the mutual ternary entropy vanishes, $S(Q : A_1 : D_1) = 0$. The quantity $S_1 = H[q^{(1)}]$ is the marginal entropy of each of the three subsystems, Q, A_1 and D_1.

where $|M_i\rangle_{D_1}$ is the initial state of the device D_1 and $|x_1\rangle$ are its final orthogonal states. Tracing over the quantum system, we find that D_1 is perfectly correlated with the quantum ancilla A_1 according to the density matrix

$$\rho(A_1 D_1) = \sum_{x_1} q_{x_1}^{(1)} \, |x_1 x_1\rangle\langle x_1 x_1|, \tag{7.43}$$

where $q_{x_1}^{(1)} = |\alpha_{x_1}^{(1)}|^2$. That is, they consistently reflect the same measurement outcomes. Together, A_1 and D_1 are still entangled with the quantum system. In Fig. 7.8 we show the entropy Venn diagrams for the entangled state (7.42) and the correlated state (7.43). Since the underlying state (7.42) is pure, the ternary mutual entropy vanishes, $S(Q : A_1 : D_1) = 0$. In other words, the correlations that are created between the devices (the $S(A_1 : D_1)$ dits of information that are gained in the measurement) are not shared with the quantum system.

The macroscopic device D_1 is composed of many qudits $D_1^{(1)}, \ldots, D_1^{(n)}$ that all measure

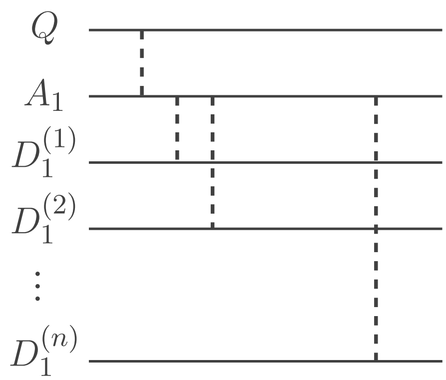

Figure 7.9: Observing the quantum observer A_1 with a device D_1. Dashed lines indicate the entanglement created by the measurement between the ancilla A_1 and each of the n qudits $D_1^{(1)}, \ldots, D_1^{(n)}$ that comprise D_1. Time proceeds from left to right.

the quantum ancilla A_1 according to the sequence of entangling operations $U_{A_1 D_1^{(n)}} \ldots U_{A_1 D_1^{(1)}}$ (see Fig. 7.9). That is, Eq. (7.42) can be expanded to

$$|QA_1D_1\rangle = \sum_{x_1} \alpha_{x_1}^{(1)} |\widetilde{x}_1\rangle |x_1\rangle |x_1\rangle_{D_1^{(1)}} \ldots |x_1\rangle_{D_1^{(n)}} . \tag{7.44}$$

The measurement outcome is read out from the state of the joint system

$$\rho(D_1^{(1)} \ldots D_1^{(n)}) = \sum_{x_1} q_{x_1}^{(1)} |x_1 \ldots x_1\rangle\langle x_1 \ldots x_1|, \tag{7.45}$$

where it is clear that the device D_1 is self-consistent and all of its components reflect the same measurement outcome. This state is robust in the sense that it is not necessary to "keep track" of all qudits in D_1 to observe correlations. Thus, tracing over any of the states in the expression above returns an equivalently self-consistent state.

In the following two sections, we amplify a chain of consecutive measurements of a prepared and an unprepared quantum system. Unlike our previous results for unamplified measurements, we will find that the joint state of devices D_1, D_2, $\ldots$, is now always classical (diagonal in the ancilla product basis), leading to entropy distributions that are significantly different from those of the unamplified ancillae.

7.3.2 Amplifying Consecutive Measurements of Prepared Quantum States

We begin by first considering the amplification of consecutive measurements of a prepared quantum state. Introducing a second pair of devices A_2 and D_2, Eq. (7.42) evolves to

$$|QA_1D_1A_2D_2\rangle = \sum_{x_1x_2} \alpha_{x_1}^{(1)}\, U_{x_1x_2}^{(2)}\, |\widetilde{x}_2\, x_1x_1\, x_2x_2\rangle. \tag{7.46}$$

Again, we find D_2 to be perfectly correlated with the quantum ancilla A_2. The joint state of D_1 and D_2 is the classical density matrix

$$\rho(D_1D_2) = \sum_{x_1x_2} |\alpha_{x_1}^{(1)}|^2\, |U_{x_1x_2}^{(2)}|^2\, |x_1x_2\rangle\langle x_1x_2|. \tag{7.47}$$

This state is diagonal in the ancilla product basis, unlike the state (7.3) before amplification. Thus, the effect of amplifying the ancillae is a removal of all off-diagonal elements in the joint density matrices.

From (7.47), we see that for repeated measurements in the same basis ($U_{x_1x_2}^{(2)} = \delta_{x_1x_2}$) the results are fully correlated. The joint density matrix (7.47) reduces to $\rho(D_1D_2) = \sum_{x_1} |\alpha_{x_1}^{(1)}|^2\, |x_1x_1\rangle\langle x_1x_1|$ so that the entropy of D_2 given D_1 vanishes, $S(D_2|D_1) = S(D_1D_2) - S(D_1) = 0$. The conditional probability to record the outcome x_2, given that the first measurement yielded x_1, is simply $p(x_2|x_1) = \delta_{x_1x_2}$. In other words, both devices agree on the outcome, as expected. It appears as if the quantum system had indeed "collapsed" into an eigenstate of the first observable since the second device D_2 correctly confirms the measurement outcome. This result is consistent with the Copenhagen view of the quantum state during the measurement sequence as $|Q\rangle \to |\widetilde{x}_1\rangle \to |\widetilde{x}_1\rangle$. However, we see that no collapse assumption is needed for a consistent description of the measurement outcomes and their correlations, and in fact, all amplitudes of the quantum system are preserved. That is, (7.46) continues to evolve as a pure state.

In addition, the probability distribution for the second measurement with the pair A_2D_2 is consistent with a collapse postulate as it is given by the incoherent sum (a sum of squares), $q^{(2)}_{x_2} = \sum_{x_1} |\alpha^{(1)}_{x_1}|^2 \, |U^{(2)}_{x_1 x_2}|^2$, instead of the coherent expression (a square of sums), $\big|\sum_{x_1} \alpha^{(1)}_{x_1} \, U^{(2)}_{x_1 x_2}\big|^2$, which is the result if the first measurement with A_1D_1 had never occurred.

7.3.3 Amplifying Consecutive Measurements of Unprepared Quantum States

In this section, we study consecutive measurements of an unprepared quantum state, which will yield an entropy Venn diagram that differs significantly from Fig. 7.4 for the quantum ancillae. To begin, we follow the procedure introduced in Sec. 7.3.1, and amplify the state (7.8) of three consecutive measurements of an unprepared quantum state.

First, we show that amplifying the qubits on the boundary of the chain of measurements does not alter the coherence of the joint state (7.9). Introducing devices D_1 and D_3 that amplify the quantum ancillae A_1 and A_3, respectively, the state (7.8) evolves to

$$|QRA_1D_1A_2A_3D_3\rangle \; = \; \frac{1}{\sqrt{d}} \sum_{x_1 x_2 x_3} U^{(2)}_{x_1 x_2} \; U^{(3)}_{x_2 x_3} \; |\widetilde{x}_3 \, x_1 x_1 x_1 \, x_2 \, x_3 x_3\rangle. \tag{7.48}$$

As before, each pair of systems A_iD_i are perfectly correlated and reflect the same outcome from their measurement of Q. Tracing over the density matrix formed from this wave func-

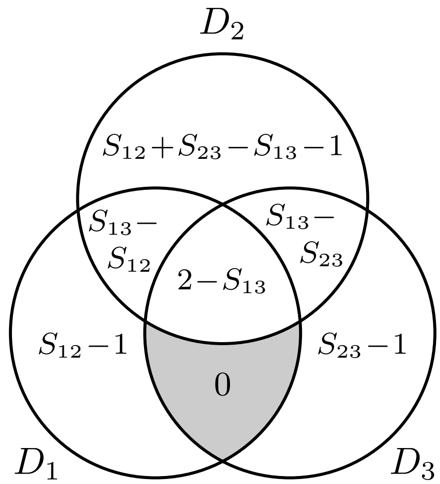

Figure 7.10: A sequence of devices D_1, D_2, and D_3 that observe (amplify) the quantum ancillae A_1, A_2, and A_3, according to (7.49). Only amplification of the intermediate ancilla A_2 is sufficient to destroy the coherence in the original state (7.9). The zero conditional mutual entropy, $S(D_1 : D_3|D_2) = 0$, indicates that the last device D_3 has no information about the first, from the perspective of the second. This property will be connected to quantum Markov chains in Sec. 7.3.4. Note that all the pairwise entropies are unchanged by amplification, $S(A_iA_j) = S(D_iD_j) = S_{ij}$.

tion, we find that the new state of $A_1A_2A_3$ is unchanged from Eq. (7.9).

In contrast, amplifying the *intermediate* ancilla destroys all of the coherence in the original state (7.9). That is, measuring A_2 with D_2 leads to a fully incoherent density matrix for $A_1A_2A_3$ that is now equivalent to the joint state

$$\rho(D_1D_2D_3) = \frac{1}{d}\sum_{x_1x_2x_3} |U^{(2)}_{x_1x_2}|^2|U^{(3)}_{x_2x_3}|^2|x_1x_2x_3\rangle\langle x_1x_2x_3|. \tag{7.49}$$

We can contrast this state to the result we obtained for unamplified measurements in Eq. (7.9) using entropy Venn diagrams. Compare the diagram in Fig. 7.10 for the state $\rho(D_1D_2D_3)$ [Eq. (7.49)] to the diagram in Fig. 7.4 for the unamplified state $\rho(A_1A_2A_3)$ [Eq. (7.9)]. Clearly, amplification of just the intermediate ancilla A_2 (or, equivalently, all three quantum ancillae) has destroyed the coherence of the original state $\rho(A_1A_2A_3)$, which was encoded in the A_2 subsystem. Note that pairwise entropies are the same for both

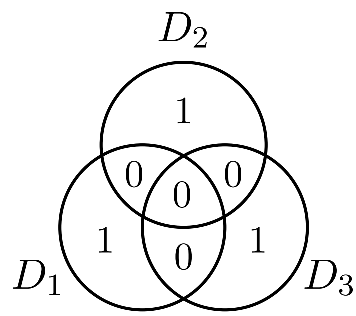

Figure 7.11: Amplification with D_1, D_2, and D_3 in (7.50) of three qubit ancillae that measured an unprepared quantum system. The second measurement of Q is of an observable with eigenbasis rotated by $\theta_2 = \pi/4$ relative to the eigenbasis of the first observable. And, the third observable is at $\theta_3 = \pi/4$ relative to the second. In this case, all three devices are uncorrelated. Amplification of A_2 with D_2 alone is sufficient to destroy the coherence in (7.17).

amplified and unamplified measurements of unprepared quantum systems, e.g., $S(A_iA_j) = S(D_iD_j)$. We proved previously in Theorem 2 of Sec. 7.2.3 that pairwise density matrices (7.39) are always diagonal, so that amplifying those ancillae does not modify their joint density matrix.

We apply these results to the case of qubit measurements ($d = 2$), which we implemented with the rotation matrix in Eq. (7.16). For three consecutive measurements with $\theta_2 = \theta_3 = \pi/4$, the joint density matrix of $D_1D_2D_3$, which we show for comparison to the unamplified state (7.17), is diagonal:

$$\rho(D_1D_2D_3) = \frac{1}{8}\begin{pmatrix} \mathbb{1} & 0 & 0 & 0 \\ 0 & \mathbb{1} & 0 & 0 \\ 0 & 0 & \mathbb{1} & 0 \\ 0 & 0 & 0 & \mathbb{1} \end{pmatrix}. \tag{7.50}$$

As with the unamplified state (7.17), the pairwise entropies here are also 2 bits. However, the tripartite entropy has increased to $S(D_1D_2D_3) = 3$ bits from the 2 bits we found for $S(A_1A_2A_3)$. Compare the resulting entropy Venn diagram in Fig. 7.11 to the diagram in Fig. 7.5 obtained for unamplified qubit measurements.

The difference between the unamplified density matrix $\rho(A_1A_2A_3)$ in Eq. (7.9) and the amplified state $\rho(D_1D_2D_3)$ in Eq. (7.49) can be ascertained by revealing the off-diagonal terms via quantum state tomography (see, e.g., [190]), by measuring just a single moment [191] of the density matrix, such as $\mathrm{Tr}[\rho(A_1A_2A_3)^2]$, or else by direct measurement of the wave function [106].

The results in the two preceding sections are compatible with the usual formalism for orthogonal measurements [192, 193], where the conditional probability $p(x_2|x_1)$ to observe outcome x_2, given that the previous measurement yielded outcome x_1, is given by

$$p(x_2|x_1) = |U^{(2)}_{x_1x_2}|^2 \,. \tag{7.51}$$

Indeed, our findings thus far are fully consistent with a picture in which a measurement "collapses" the quantum state (or alternatively, where a measurement recalibrates an observer's "catalogue of expectations" [194–196]).

To see this, we write the joint density matrix $\rho(D_1D_2)$, found by tracing (7.49) over D_3, in the collapse picture. For a device D_1 that records outcome x_1 with probability $1/d$, and a device D_2 corresponding to a measurement of Q at an angle determined by the rotation matrix $U^{(2)}$, the resulting density matrix is

$$\widetilde{\rho}\,(D_1D_2) = \frac{1}{d}\sum_{x_1} |x_1\rangle\langle x_1| \otimes \rho^{x_1}_{D_2}\,, \tag{7.52}$$

where the state $\rho^{x_1}_{D_2}$ of D_2 is defined using the projection operators $P_{x_1} = |x_1\rangle\langle x_1|$ on the state of D_1,

$$\rho^{x_1}_{D_2} = \frac{\mathrm{Tr}_{D_1}\Big[P_{x_1}\,\rho(D_1D_2)P^{\dagger}_{x_1}\Big]}{\mathrm{Tr}_{D_1D_2}\Big[P_{x_1}\,\rho(D_1D_2)P^{\dagger}_{x_1}\Big]} = \sum_{x_2} |U^{(2)}_{x_1x_2}|^2\,|x_2\rangle\langle x_2|. \tag{7.53}$$

In other words, the state $\rho(D_1D_2)$ that was obtained in a unitary formalism is equivalent to the collapse version $\widetilde{\rho}(D_1D_2)$. However, despite these consistencies with the collapse picture, we emphasize that the actual measurements induce no irreversible collapse and that all coefficients in the underlying pure-state wave function (7.48) are preserved and evolve unitarily throughout the measurement process.

7.3.4 Quantum Markov Chains

One of the key differences between the entropy Venn diagrams in Figs. 7.4 and 7.10 is the vanishing conditional mutual entropy [59] for amplified measurements, $S(D_1 : D_3|D_2) = 0$. Before amplification, the equivalent quantity for the quantum ancillae is in general non-zero, $S(A_1 : A_3|A_2) \geq 0$. Evidently, the intermediate measurement with D_2 has, from the perspective of D_2 (meaning, given the state of D_2) *erased all correlations* between the first device D_1 and the last device D_3 in the measurement sequence. The vanishing of the conditional mutual entropy is precisely the condition that is fulfilled by quantum Markov chains as we will outline below.

Using the results for unprepared quantum states (this holds equally for prepared quantum states), we demonstrate that the chain of devices, D_1, D_2, D_3, which amplified consecutive measurements of a quantum system, is Markovian as defined in [69] (see also [68] and references therein). We prove later in this section in Theorem 3 that this result can be extended to any number of consecutive measurements, not just three. To show that $S(D_1 : D_3|D_2)$ is indeed zero, we compute the joint entropy $S(D_1D_2D_3)$ of all three devices. From Eq. (7.49), we find

$$S(D_1D_2D_3) = 1 - \frac{1}{d}\sum_{x_1x_2} |U^{(2)}_{x_1x_2}|^2 \log_d |U^{(2)}_{x_1x_2}|^2 - \frac{1}{d}\sum_{x_2x_3} |U^{(3)}_{x_2x_3}|^2 \log_d |U^{(3)}_{x_2x_3}|^2, \quad (7.54)$$

or, $S(D_1D_2D_3) = S(D_1) + S(D_2|D_1) + S(D_3|D_2)$. However, using the chain rule for entropies [59], the tripartite entropy can also be written generally in the form $S(D_1D_2D_3) = S(D_1) + S(D_2|D_1) + S(D_3|D_2D_1)$. From these two expression, we see immediately that

$$S(D_3|D_2D_1) = S(D_3|D_2). \tag{7.55}$$

Thus, the entropy of D_3 is not reduced by conditioning on more than the state of the previous device D_2. This is the Markov property for entropies [68, 69].

The Markov property further implies that D_1 and D_3 are independent from the perspective of D_2, since the conditional mutual entropy [59] vanishes (see the gray region in Fig. 7.10),

$$S(D_1 : D_3|D_2) = S(D_3|D_2) - S(D_3|D_2D_1) = 0\,. \tag{7.56}$$

This result is consistent with the notion that the measurement with D_2 collapsed the state of the wave function, erasing any (conditional) information that D_3 could have had about the prior measurement with D_1. The conditional mutual entropy does not vanish for unamplified measurements, $S(A_1 : A_3|A_2) \geq 0$, reflecting the fundamentally non-Markovian nature of the chain of quantum ancillae. In other words, as long as the measurement chain remains unamplified (for example, the A_2 subsystem in (7.9)), the intermediate measurement does not erase the correlations between A_1 and A_3 (compare the gray region in Fig. 7.10 to the same region in Fig. 7.4).

We now provide a formal proof of the statement that the chain of devices that amplified the quantum ancillae is equivalent to a quantum Markov chain.

Theorem 3. *A set of consecutive quantum measurements is non-Markovian until it is am-*

plified. Specifically, the sequence of devices $D_i, \ldots, D_j$, with $i < j$, that measure (amplify) the quantum ancillae $A_i, \ldots, A_j$ (which themselves measured a quantum system Q) forms a quantum Markov chain:

$$S(D_j|D_{j-1}\ldots D_i) = S(D_j|D_{j-1}). \tag{7.57}$$

Proof. We first show that the Markov property of probabilities implies the Markov property for entropies (see, e.g., Refs. [68,69]). If consecutive measurements on a quantum system can be modeled as a Markov process, the probability to observe outcome x_j in the jth device, conditional on previous measurement outcomes, depends only on the last outcome x_{j-1},

$$p(x_j|x_{j-1}\ldots x_i) = p(x_j|x_{j-1}). \tag{7.58}$$

Inserting Eq. (7.58) into the expression for the conditional entropy [58] gives

$$\begin{aligned} S(D_j|D_{j-1}\ldots D_i) &= -\sum_{x_i\cdots x_j} p(x_i\ldots x_j)\log_d p(x_j|x_{j-1}\ldots x_i) \\ &= -\sum_{x_i\cdots x_j} p(x_i\ldots x_j)\log_d p(x_j|x_{j-1})\,. \end{aligned} \tag{7.59}$$

A partial summation over the joint probability distribution gives

$$p(x_{j-1}x_j) \;= \sum_{x_i\cdots x_{j-2}} p(x_i\ldots x_j), \tag{7.60}$$

so that the entropic condition satisfied by a quantum Markov chain is

$$\begin{aligned} S(D_j|D_{j-1}\ldots D_i) &= -\sum_{x_{j-1}x_j} p(x_{j-1}x_j)\;\log_d p(x_j|x_{j-1}) \\ &= S(D_j|D_{j-1}),. \end{aligned} \tag{7.61}$$

We now show that the chain of amplified measurements satisfies the entropic Markov property (7.61). For n consecutive measurements, the state $|\Psi\rangle = |QA_1 \dots A_n\rangle$ of Q and all ancillae is given by

$$|\Psi\rangle = \sum_{x_1 \cdots x_n} \alpha^{(i)}_{x_i}\, U^{(2)}_{x_1 x_2} \dots U^{(n)}_{x_{n-1} x_n}\, |\widetilde{x}_n\, x_1 \dots x_n\rangle. \tag{7.62}$$

After amplifying this state, we find that the density matrix for the joint set of sequential devices, $D_i, \dots, D_j$, with $i < j$, is diagonal, as expected,

$$\rho(D_i \dots D_j) = \sum_{x_i} q^{(i)}_{x_i} |x_i\rangle\langle x_i| \otimes \sum_{x_{i+1}} |U^{(i+1)}_{x_i x_{i+1}}|^2\, |x_{i+1}\rangle\langle x_{i+1}| \otimes \cdots \otimes \sum_{x_j} |U^{(j)}_{x_{j-1} x_j}|^2\, |x_j\rangle\langle x_j|. \tag{7.63}$$

The probability distribution $q^{(i)}_{x_i}$ of the ith device can be obtained from (7.29). The entropy of (7.63) is

$$\begin{aligned} S(D_i \dots D_j) = &- \sum_{x_i \cdots x_{j-1}} \Big[q^{(i)}_{x_i} |U^{(i+1)}_{x_i x_{i+1}}|^2 \dots |U^{(j-1)}_{x_{j-2} x_{j-1}}|^2\Big] \log_d \Big[q^{(i)}_{x_i} |U^{(i+1)}_{x_i x_{i+1}}|^2 \dots |U^{(j-1)}_{x_{j-2} x_{j-1}}|^2\Big] \\ &- \sum_{x_{j-1} x_j} q^{(j-1)}_{x_{j-1}}\, |U^{(j)}_{x_{j-1} x_j}|^2 \log_d |U^{(j)}_{x_{j-1} x_j}|^2, \end{aligned} \tag{7.64}$$

where $q^{(j-1)}_{x_{j-1}}$ is the probability distribution of D_{j-1}. The first term in Eq. (7.64) is just the joint entropy $S(D_i \dots D_{j-1})$, so that the entropy of the jth device, conditional on the previous devices, is

$$\begin{aligned} S(D_j | D_{j-1} \dots D_i) &= S(D_i \dots D_j) - S(D_i \dots D_{j-1}) \\ &= - \sum_{x_{j-1} x_j} q^{(j-1)}_{x_{j-1}}\, |U^{(j)}_{x_{j-1} x_j}|^2 \, \log_d |U^{(j)}_{x_{j-1} x_j}|^2. \end{aligned} \tag{7.65}$$

All that remains is to show that (7.65) is equal to $S(D_j | D_{j-1})$. A simple calculation using the density matrix for two amplified consecutive measurements with D_{j-1} and D_j,

$$\rho(D_{j-1}D_j) = \sum_{x_{j-1}x_j} q_{x_{j-1}}^{(j-1)}\, |U_{x_{j-1}x_j}^{(j)}|^2\, |x_{j-1}x_j\rangle\langle x_{j-1}x_j|, \tag{7.66}$$

yields the joint entropy,

$$S(D_{j-1}D_j) = -\sum_{x_{j-1}} q_{x_{j-1}}^{(j-1)} \log_d q_{x_{j-1}}^{(j-1)} - \sum_{x_{j-1}x_j} q_{x_{j-1}}^{(j-1)} |U_{x_{j-1}x_j}^{(j)}|^2 \log_d |U_{x_{j-1}x_j}^{(j)}|^2. \tag{7.67}$$

The first term in this expression is the entropy of D_{j-1} (all marginal density matrices and entropies are the same for amplified and unamplified ancillae; this is proved formally later in Lemma 2 of Sec. 7.4.1),

$$S(D_{j-1}) = H[q^{(j-1)}] = -\sum_{x_{j-1}} q_{x_{j-1}}^{(j-1)} \log_d q_{x_{j-1}}^{(j-1)}. \tag{7.68}$$

The conditional entropy $S(D_j|D_{j-1})$ is thus

$$\begin{aligned} S(D_j|D_{j-1}) &= S(D_{j-1}D_j) - S(D_{j-1}) \\ &= -\sum_{x_{j-1}x_j} q_{x_{j-1}}^{(j-1)} |U_{x_{j-1}x_j}^{(j)}|^2 \log_d |U_{x_{j-1}x_j}^{(j)}|^2, \end{aligned} \tag{7.69}$$

which is the same as (7.65). □

We emphasize that the result that amplified measurements are Markovian holds for measurements of unprepared as well as prepared quantum states.

Corollary 3.1. *The Markovian nature of amplified measurements implies that the devices D_i and D_j share no entropy (are independent) from the perspective of the intermediate devices,*

$D_{i+1}, \ldots, D_{j-1}$, since the conditional mutual entropy vanishes:

$$S(D_i : D_j | D_{i+1} \ldots D_{j-1}) = 0. \tag{7.70}$$

Proof. The conditional mutual entropy is defined [59] as a difference between two conditional entropies,

$$S(D_i : D_j | D_{i+1} \ldots D_{j-1}) = S(D_j | D_{j-1} \ldots D_{i+1}) - S(D_j | D_{j-1} \ldots D_i). \tag{7.71}$$

From Theorem 3, the two quantities on the right hand side of this expression are both equal to $S(D_j | D_{j-1})$. Therefore the conditional mutual entropy vanishes [69]. $\square$

For three devices, the Markov property is

$$S(D_{i-1} : D_{i+1} | D_i) = S(D_{i+1} | D_i) - S(D_{i+1} | D_i D_{i-1}) = 0. \tag{7.72}$$

We see that, from the strong subadditivity (SSA) of quantum entropy [66, 67],

$$S(D_{i+1} | D_i D_{i-1}) \leq S(D_{i+1} | D_i), \tag{7.73}$$

amplified measurements satisfy SSA with equality.

The previous theorem established that the sequence of amplified measurements is a quantum Markov chain. Now, we will demonstrate that unamplified measurements are non-Markovian. In the following calculation, we use the state (7.34) for measurements of unprepared quantum states for simplicity. We will find that the Markov property (7.61) is violated in this case, so that in general unamplified measurements are non-Markovian.

First, consider the joint density matrix for the sequence of quantum ancillae $A_i, \ldots, A_j$

(with $i < j$), similarly to (7.34). As in Eq. (7.35), we find

$$\rho(A_i \dots A_j) = \frac{1}{d} \sum_{x_i x_j} p^{(ij)}_{x_i x_j} \, |x_i\rangle\langle x_i| \otimes |\phi_{x_i x_j}\rangle\langle\phi_{x_i x_j}| \otimes |x_j\rangle\langle x_j|, \tag{7.74}$$

where the coefficients $p^{(ij)}_{x_i x_j} = |\beta^{(ij)}_{x_i x_j}|^2$ and the normalized, but non-orthogonal states $|\phi_{x_i x_j}\rangle$ were defined in Eq. (7.36). The joint states $|x_i \, \phi_{x_i x_j} \, x_j\rangle$ are orthonormal, so the entropy of Eq. (7.74) is simply

$$S(A_i \dots A_j) = 1 - \frac{1}{d} \sum_{x_i x_j} p^{(ij)}_{x_i x_j} \log_d p^{(ij)}_{x_i x_j}. \tag{7.75}$$

The coefficients $p^{(ij)}_{x_i x_j}$ can be equivalently expressed in terms of $U^{(j)}$ as

$$p^{(ij)}_{x_i x_j} = |\beta^{(ij)}_{x_i x_j}|^2 = \sum_{x_{j-1}} p^{(i,j-1)}_{x_i x_{j-1}} \, |U^{(j)}_{x_{j-1} x_j}|^2. \tag{7.76}$$

Next, we use the log-sum inequality [56] to rewrite the joint entropy (7.75) as an inequality. The log-sum inequality states that for non-negative numbers $a_1, \dots, a_d$ and $b_1, \dots, b_d$,

$$\sum_{x_i=1}^{d} a_{x_i} \log \frac{a_{x_i}}{b_{x_i}} \geq \left(\sum_{x_i=1}^{d} a_{x_i} \right) \log \frac{\sum_{x_i=1}^{d} a_{x_i}}{\sum_{x_i=1}^{d} b_{x_i}} \,, \tag{7.77}$$

with equality if and only if $a_{x_i}/b_{x_i} = \text{const}$. Inserting (7.76) into (7.75) and using the log-sum inequality with $b_{x_{j-1}} = 1$ and $a_{x_{j-1}} = p^{(i,j-1)}_{x_i x_{j-1}} \, |U^{(j)}_{x_{j-1} x_j}|^2$, we find that the joint entropy is bounded from below by

$$\begin{aligned} S(A_i \dots A_j) \geq & -\frac{1}{d} \sum_{x_i x_{j-1}} p^{(i,j-1)}_{x_i x_{j-1}} \log_d p^{(i,j-1)}_{x_i x_{j-1}} \\ & - \frac{1}{d} \sum_{x_{j-1} x_j} |U^{(j)}_{x_{j-1} x_j}|^2 \log_d |U^{(j)}_{x_{j-1} x_j}|^2. \end{aligned} \tag{7.78}$$

The first term on the right hand side of Eq. (7.78) is simply $S(A_i \dots A_{j-1}) - 1$, while the second term is $S(A_{j-1}A_j) - 1$. Given that $S(A_{j-1}) = 1$, it is straightforward to show that Eq. (7.78) can be rewritten as a difference between two conditional entropies,

$$S(A_j|A_{j-1}) - S(A_j|A_{j-1} \dots A_i) \leq 1, \tag{7.79}$$

with equality only when $p^{(i,j-1)}_{x_i x_{j-1}} |U^{(j)}_{x_{j-1}x_j}|^2$ is a constant. This occurs when $|U^{(j)}_{x_{j-1}x_j}|^2 = 1/d$ and $|U^{(\ell)}_{x_{\ell-1}x_\ell}|^2 = 1/d$ for one or more of the $\ell = i+1, \dots, j-1$ matrices. This shows that conditioning on more than just the state of the last ancilla A_{j-1} will reduce the conditional entropy of ancilla A_j (by at most 1). Since Eq. (7.79) is not equal to zero in general, we conclude that the sequence of unamplified measurements is non-Markovian.

In the next section, we will compute the information a measuring device has about the quantum system or other devices in a sequence of measurements. In particular, we will show that amplifying measurements (which yields a quantum Markov chain) reduces the information that can be obtained about other systems, as compared to measurements that remain unamplified.

7.4 Effects of Amplifying Quantum Measurements

In the previous Secs. 7.2 and 7.3, we focused on consecutive measurements of a quantum system and discussed the concepts of non-Markovian (unamplified) and Markovian (amplifiable) sequences, respectively. It is reasonable to ask whether there are entropic relationships between those two kinds of measurements. Introducing a second step to von Neumann's second stage serves precisely to establish such relationships. In this section, we establish the following three properties: Markovian devices carry less information about the quantum system than non-Markovian devices; the shared entropy between consecutive non-Markovian devices is larger than the respective quantity for amplified measurements; the last Markovian device in a quantum chain is inherently more random than its non-Markovian counterpart, given the combined state of all previous devices.

7.4.1 Information About the Quantum System

We first calculate how much information about the quantum system, Q, is encoded in the last device in a chain of consecutive measurements of Q. To do this, we prove two Lemmas that state that the marginal entropy of the quantum system is always equal to the entropy of the last ancilla in the chain of measurements, and that the marginal entropy of a quantum ancilla is unaffected by amplification.

Lemma 1. *The entropy of the quantum system, Q, is equal to the entropy of the last ancilla, A_n, in the chain of measurements:*

$$S(Q) = S(A_n). \tag{7.80}$$

Proof. Consider a series of consecutive measurements on a quantum system, Q, with n ancillae. In general, following the measurements, the joint state of the quantum system and all ancillae $|\Psi\rangle = |QA_1 \dots A_n\rangle$ is given by the pure state [see also Eq. (7.25)]

$$|\Psi\rangle = \sum_{x_1 \dots x_n} \alpha^{(1)}_{x_1} \, U^{(2)}_{x_1 x_2} \dots U^{(n)}_{x_{n-1} x_n} \, |\widetilde{x}_n \; x_1 \dots x_n\rangle. \tag{7.81}$$

The density matrix for the quantum system is found by tracing out all ancilla states from the full density matrix associated with (7.81),

$$\rho(Q) = \mathrm{Tr}_{A_1 \dots A_n}(|\Psi\rangle\langle\Psi|) = \sum_{x_n} q_{x_n}^{(n)} \; |\widetilde{x}_n\rangle\langle\widetilde{x}_n|, \tag{7.82}$$

where $q_{x_n}^{(n)}$ is the probability distribution for the last ancilla A_n that can be obtained generally from Eq. (7.29). Clearly, (7.82) is equivalent to the density matrix for the last ancilla, and so the corresponding entropies are the same: $S(Q) = S(A_n) = H[q^{(n)}]$. An alternative proof is to note that a Schmidt decomposition of the pure state $|\Psi\rangle\langle\Psi|$ implies that $S(Q) = S(A_1 \dots A_n)$. And, by Theorem 6 (see Sec. 7.4.2), $S(A_n) = S(A_1 \dots A_n)$, so that $S(Q) = S(A_n)$. □

Lemma 2. *The entropy of a quantum ancilla, A_i, is unchanged if it is measured by an amplifying device, D_i, so that for all i in the chain of measurements:*

$$S(A_i) = S(D_i). \tag{7.83}$$

Proof. Amplifying the ith ancilla A_i in (7.81) with D_i yields the joint density matrix for A_i and D_i,

$$\rho(A_i D_i) = \sum_{x_i} q_{x_i}^{(i)} \; |x_i x_i\rangle\langle x_i x_i|, \tag{7.84}$$

where $q_{x_i}^{(i)}$ is the probability distribution for A_i, as defined in (7.29). The two subsystems are perfectly correlated so that the density matrix and marginal entropy of A_i is equivalent to D_i: $S(D_i) = S(A_i) = H[q^{(i)}]$. □

In the remaining sections, we will use the shortened notation $S(A_i) = S(D_i) = S_i$ for the marginal entropies. Using Lemmas 1 and 2, we are now ready to prove the first theorem regarding information about the quantum system.

Theorem 4. *The information that the last device in a series of measurements has about the quantum system is reduced when the measurements are amplified. That is,*

$$S(Q:D_n) \leq S(Q:A_n), \tag{7.85}$$

for n consecutive measurements of a prepared quantum state, Q.

Proof. We start with the state (7.81) for an unamplified chain of consecutive measurements of a prepared quantum state, Q, with n ancillae. Tracing out all previous ancilla states from (7.81), the joint density matrix for the quantum system and the last ancilla is

$$\rho(QA_n) = \sum_{x_{n-1}x_n x_n'} q_{x_{n-1}}^{(n-1)}\, U_{x_{n-1}x_n}^{(n)}\, U_{x_{n-1}x_n'}^{(n)\,*}\, |\widetilde{x}_n\, x_n\rangle\langle\widetilde{x}_n'\, x_n'|, \tag{7.86}$$

where $q_{x_{n-1}}^{(n-1)}$ is A_{n-1}'s probability distribution.

If we amplify the measurement chain (or, equivalently, just the last measurement) the state (7.86) becomes diagonal. That is,

$$\rho(QD_n) = \sum_{x_{n-1}x_n} q_{x_{n-1}}^{(n-1)}\, |U_{x_{n-1}x_n}^{(n)}|^2\, |\widetilde{x}_n\, x_n\rangle\langle\widetilde{x}_n\, x_n|. \tag{7.87}$$

Note that the amplification is equivalent to a completely dephasing channel [197–199] (see also [25]) since we can write

$$\rho(QD_n) = \sum_{x_n} P_{x_n}\, \rho(QA_n)\, P_{x_n}, \tag{7.88}$$

where $P_{x_n} = |x_n\rangle\langle x_n|$ are projectors on the state of A_n. In other words, $\rho(QD_n)$ is formed from the diagonal elements of $\rho(QA_n)$. The dephasing channel is also called the phasedamping channel and serves as a quantum mechanical noise model where phase information about a quantum state is lost. For example, if the initial quantum state of a system A is $|\psi\rangle = \alpha|0\rangle_A + \beta|1\rangle_A$, with density matrix ρ_A, then the output of the dephasing channel is $\rho'_A = (1-p)\,\rho_A + p\,(P_0\,\rho_A\,P_0 + P_1\,\rho_A\,P_1)$. That is, with probability $(1-p)$ nothing happens to the quantum state, and with probability $p/2$ one of the projectors $P_i = |i\rangle_A\langle i|$ is applied. The channel is equivalent to the phase-flip channel since it can be written as $\rho'_A = (1-p/2)\,\rho_A + p/2\,Z\,\rho_A\,Z$, where the Pauli matrix Z performs a phase flip. In the basis $|0\rangle_A$, $|1\rangle_A$, the depolarizing channel yields

$$\rho'_A = \begin{pmatrix} |\alpha|^2 & (1-p)\,\alpha\beta^* \\ (1-p)\,\alpha^*\beta & |\beta|^2 \end{pmatrix}, \tag{7.89}$$

so that for a completely dephasing channel, $p = 1$, the off-diagonal elements vanish and the state of A becomes completely incoherent.

To show that the amplified mutual entropy is reduced as in Eq. (7.85), it is sufficient to show that the joint entropy is increased. The mutual entropy for two subsystems is defined [59] as $S(Q:A_n) = S(Q) + S(A_n) - S(QA_n)$ and similarly for $S(Q:D_n)$. Since, by Lemma 2, the marginal entropies are unchanged by the amplification, $S(A_n) = S(D_n)$, we have

$$S(Q:D_n) = S(Q:A_n) + S(QA_n) - S(QD_n). \tag{7.90}$$

Therefore, we just need to show that $S(QD_n) \geq S(QA_n)$, which is easiest by considering the relative entropy of coherence [200, 201]. This quantity, $C_{\text{rel.ent.}}(\rho) = S(\rho_{\text{diag}}) - S(\rho)$, is the difference between the entropies of a density matrix ρ and a matrix ρ_{diag} that is formed from the diagonal elements of ρ. It is derived by minimizing the relative entropy

$S(\rho \,\|\, \delta) = \mathrm{Tr}(\rho \log \rho - \rho \log \delta)$ (see, e.g., [202, 203]) over the set of incoherent matrices δ. By Klein's inequality [204], the relative entropy is non-negative so that $S(\rho_{\mathrm{diag}}) \geq S(\rho)$, with equality if and only if ρ is an incoherent matrix. In our case, ρ and ρ_{diag} are given by $\rho(Q : A_n)$ and $\rho(Q : D_n)$, respectively. Therefore, it follows that $S(QD_n) \geq S(QA_n)$ and

$$S(Q : D_n) \leq S(Q : A_n), \tag{7.91}$$

with equality if and only if $\rho(QA_n)$ is already diagonal in the ancilla product basis. $\square$

To directly compute the mutual entropies in Theorem 4, we first diagonalize the density matrix (7.86) with the orthonormal states $|\Phi_{x_{n-1}}\rangle = \sum_{x_n} U^{(n)}_{x_{n-1}x_n} |\widetilde{x}_n \, x_n\rangle$, so that

$$\rho(QA_n) = \sum_{x_{n-1}} q^{(n-1)}_{x_{n-1}} \, |\Phi_{x_{n-1}}\rangle\langle\Phi_{x_{n-1}}|. \tag{7.92}$$

The joint entropy of this state is simply the marginal entropy of A_{n-1}. That is, $S(QA_n) = S(A_{n-1}) = S_{n-1}$, which can also be derived using the Schmidt decomposition and the results of Theorem 6 (see Sec. 7.4.2). Thus, using Lemma 1, the information that the last ancilla has about the quantum system is

$$S(Q : A_n) = 2S_n - S_{n-1}. \tag{7.93}$$

If we now amplify the measurement chain (or, equivalently, just the last measurement) the information that D_n has about Q will be reduced from (7.93). From Eq. (7.87), the joint density matrix of Q and D_n can also be written as

$$\rho(QD_n) = \sum_{x_n} q^{(n)}_{x_n} \, |\widetilde{x}_n \, x_n\rangle\langle\widetilde{x}_n \, x_n|, \tag{7.94}$$

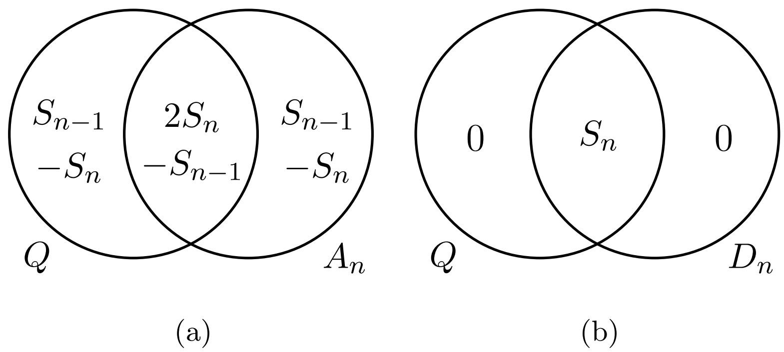

Figure 7.12: Effects of amplification on (a) the quantum system and the unamplified ancilla according to Eq. (7.86), and on (b) the quantum system and the amplifying device according to Eq. (7.94). The information that the last device has about the quantum system is reduced when the measurement is amplified. That is, $S(Q:D_n) \leq S(Q:A_n)$.

which leads to $S(QD_n) = S(D_n) = S_n$. Therefore, amplifying the measurement reduces the quantity (7.93) to

$$S(Q:D_n) = S_n, \tag{7.95}$$

where we used Lemmas 1 and 2 to write $S(Q) = S(A_n) = S(D_n) = S_n$. This quantity depends explicitly on only the last measurement, unlike (7.93), which depends on the last two. The amount of information that the last device has about the quantum system before amplification, (7.93), and after, (7.95), is related by

$$S(Q:D_n) = S(Q:A_n) + S_{n-1} - S_n\ . \tag{7.96}$$

Thus, the marginal entropies in a chain of consecutive measurements never decrease, $S_n \geq S_{n-1}$, since $S(Q:D_n) \leq S(Q:A_n)$. The entropy Venn diagrams for the devices A_n and D_n and the quantum system are shown in Fig. 7.12.

We can illustrate this loss of information about the quantum system by considering consecutive qubit measurements. Suppose that the eigenbasis of observable $(n-1)$ is at an angle $\theta_{n-1} = 0$ relative to observable $(n-2)$, and that observable n is at an angle $\theta_n = \pi/4$ relative to observable $(n-1)$. In this case, the marginal entropies are $S_{n-1} = S_{n-2} = H[q^{(n-2)}]$ and $S_n = 1$ bit. The last device, D_n, has one bit of information about the quantum system, which is less than that of the unamplified ancilla: $S(Q : A_n) = 2 - H[q^{(n-2)}] \geq 1$. Interestingly, how much we know about the state of Q prior to amplification is controlled by the entropy of an ancilla, A_{n-2}, located two steps down the measurement chain.

7.4.2 Information About Past Measurements

We now calculate how much information is encoded in a measurement device about the state of the measurement device that just preceded it in the quantum chain. In particular, we will show that the shared entropy $S(A_n : A_{n-1})$ between the last two devices in the measurement chain is reduced by the amplification process so that $S(D_n : D_{n-1}) \leq S(A_n : A_{n-1})$. These calculations have obvious relevance for the problem of quantum retrodiction [179], but we do not here derive optimal protocols to achieve this.

Theorem 5. *The information that the last device has about the previous device is reduced when that measurement is amplified. That is,*

$$S(D_n : D_{n-1}) \leq S(A_n : A_{n-1})\,. \tag{7.97}$$

Proof. From the wave function (7.81), the density matrix for the last two ancillae in the measurement chain is

$$\rho(A_{n-1}A_n) = \sum_{\substack{x_{n-2}x_{n-1}\\ x'_{n-1}x_n}} q^{(n-2)}_{x_{n-2}}\, U^{(n-1)}_{x_{n-2}x_{n-1}} U^{(n-1)*}_{x_{n-2}x'_{n-1}}\, U^{(n)}_{x_{n-1}x_n} U^{(n)*}_{x'_{n-1}x_n}\, |x_{n-1}x_n\rangle\langle x'_{n-1}x_n|. \tag{7.98}$$

Amplification removes the off-diagonals of $\rho(A_{n-1}A_n)$ so that

$$\rho(D_{n-1}D_n) = \sum_{x_{n-1}} P_{x_{n-1}}\ \rho(A_{n-1}A_n)\, P_{x_{n-1}}, \tag{7.99}$$

where $P_{x_{n-1}} = |x_{n-1}\rangle\langle x_{n-1}|$ are projectors on the state of A_{n-1}. Note that, from (7.98), it is sufficient to amplify just the second-to-last measurement with A_{n-1}. Since the marginal entropies are unchanged by the amplification (Lemma 2), the amount of information before amplification, $S(A_n : A_{n-1})$, and after, $S(D_n : D_{n-1})$, is related by

$$S(D_n\!:\!D_{n-1}) = S(A_n\!:\!A_{n-1}) + S(A_{n-1}A_n) - S(D_{n-1}D_n). \tag{7.100}$$

In a similar fashion to the calculations in Theorem 4, it is evident from (7.99) that the joint entropy is increased, $S(D_{n-1}D_n) \geq S(A_{n-1}A_n)$. It follows that the information that the last device has about the device that preceded it in the measurement sequence is reduced:

$$S(D_n : D_{n-1}) \leq S(A_n : A_{n-1}), \tag{7.101}$$

with equality if and only if $\rho(A_{n-1}A_n)$ is already diagonal in the ancilla product basis. $\square$

Using the case of qubits, we can show how amplification reduces the amount of information about past measurements. In this example, suppose that the last two measurements in the chain are each made at the relative angle $\pi/4$. As expected, the amplified density matrix (7.99) becomes uncorrelated, $\rho(D_{n-1}D_n) = \frac{1}{2}\,\mathbb{1}_{D_{n-1}} \otimes \frac{1}{2}\,\mathbb{1}_{D_n}$, where $\mathbb{1}$ is the 2×2 identity matrix, and the shared entropy vanishes $S(D_n : D_{n-1}) = 0$. In other words, the last device has no information about the one preceding it. In contrast, prior to amplification the density matrix (7.98) is coherent with joint entropy $S(A_{n-1}A_n) = 1 + S_{n-2}$. Therefore, the corresponding shared entropy is nonzero, $S(A_n : A_{n-1}) = 1 - S_{n-2} = 1 - H[q^{(n-2)}]$, revealing that information about the previous measurement survives the sequential $\pi/4$ measurements (as long as A_{n-1} is not amplified).

The calculations described above can be extended to include the information that the last device has about *all* previous devices in the measurement chain. We claim in Theorem 6 that the amplification process reduces this information by a specific minimum (calculable) amount. To prove this statement, we make use of Theorem 1, where we showed that the joint entropy of all quantum ancillae that measured a prepared quantum system is simply equal to the entropy of last ancilla in the unamplified chain.

Theorem 6. *For n consecutive measurements of a quantum system, the information that the last device has about all previous measurements is reduced by amplification by at least an amount Σ_n:*

$$S(D_n : D_{n-1} \ldots D_1) \leq S(A_n : A_{n-1} \ldots A_1) - \Sigma_n \ , \tag{7.102}$$

where $\Sigma_n = S(A_{n-1}|A_n) \geq 0$ is a non-negative conditional entropy that quantifies the uncertainty about the prior measurement given the last.

Proof. We begin by recognizing that the amplified mutual entropy $S(D_n : D_{n-1} \ldots D_1)$ for the full measurement chain is equal to $S(D_n : D_{n-1})$ by the Markov property (see Theorem 3). Then, by Theorem 5, we can place an upper bound on the amplified information

$$S(D_n : D_{n-1} \ldots D_1) = S(D_n : D_{n-1}) \leq S(A_n : A_{n-1}), \tag{7.103}$$

where $S(A_n : A_{n-1})$ is the mutual entropy before amplifying the measurement. Next, we will relate $S(A_n : A_{n-1})$ to $S(A_n : A_{n-1} \ldots A_1)$. From Theorem 1, the latter quantity can be written simply as

$$S(A_n : A_{n-1} \dots A_1) = S_{n-1}, \tag{7.104}$$

so that with the definition of $S(A_n : A_{n-1})$, we come to

$$S(A_n : A_{n-1} \dots A_1) = S(A_n : A_{n-1}) + \Sigma_n, \tag{7.105}$$

where $\Sigma_n = S(A_{n-1}|A_n)$ represents the information gained by conditioning on all previous measurements. Inserting (7.105) into the inequality (7.103), we come to

$$S(D_n : D_{n-1} \dots D_1) \leq S(A_n : A_{n-1} \dots A_1) - \Sigma_n. \tag{7.106}$$

The information is reduced as long as $\Sigma_n \geq 0$. To show this, we recall the joint density matrix (7.98) for A_{n-1} and A_n. This state can be written as a classical-quantum state

$$\rho(A_{n-1}A_n) = \sum_{x_n} q_{x_n}^{(n)} \ \rho_{x_n} \otimes |x_n\rangle\langle x_n|, \tag{7.107}$$

where

$$q_{x_n}^{(n)} \ \rho_{x_n} = \sum_{x_{n-2}} q_{x_{n-2}}^{(n-2)} \ p_{x_{n-2}x_n}^{(n-2,n)} \ |\phi_{x_{n-2}x_n}\rangle\langle\phi_{x_{n-2}x_n}|, \tag{7.108}$$

and the non-orthogonal states $|\phi_{x_{n-2}x_n}\rangle$ were previously defined in Eq. (7.36). In this block-diagonal form, the entropy is

$$S(A_{n-1}A_n) = S_n + \sum_{x_n} q_{x_n}^{(n)} \, S(\rho_{x_n}), \tag{7.109}$$

so that the quantity of interest, Σ_n, can be written as

$$\Sigma_n = S(A_{n-1}|A_n) = \sum_{x_n} q_{x_n}^{(n)}\, S(\rho_{x_n}) \geq 0. \tag{7.110}$$

This quantity is clearly non-negative since both $q_{x_n}^{(n)} \geq 0$ and $S(\rho_{x_n}) \geq 0\ \forall\ x_n$. Therefore, with $\Sigma_n \geq 0$, we find that the information is indeed reduced by the amplification process, and by at least an amount equal to Σ_n. □

Continuing with our qubit example that followed Theorem 5, if the last two measurements were each made at the relative angle $\pi/4$, the ancilla A_n has 1 bit of information about the joint state of all previous ancillae. That is, $S(A_n : A_{n-1} \ldots A_1) = 1$ bit, while the amplifying device D_n has no information at all, $S(D_n : D_{n-1} \ldots D_1) = 0$.

Corollary 6.1. *Amplifying the measurement chain increases the entropy of the last device, when conditioned on all previous devices, by at least an amount Σ_n:*

$$S(D_n|D_{n-1} \ldots D_1) \geq S(A_n|A_{n-1} \ldots A_1) + \Sigma_n. \tag{7.111}$$

Proof. By definition, the mutual entropy and conditional entropy are related by

$$S(D_n : D_{n-1} \ldots D_1) = S_n - S(D_n|D_{n-1} \ldots D_1), \tag{7.112}$$

which, from Theorem 6, is bounded from above by $S(A_n : A_{n-1} \ldots A_1) - \Sigma_n = S_n - S(A_n|A_{n-1} \ldots A_1) - \Sigma_n$. Therefore,

$$S(D_n|D_{n-1} \ldots D_1) \geq S(A_n|A_{n-1} \ldots A_1) + \Sigma_n, \tag{7.113}$$

and the uncertainty in the last measurement is increased by at least an amount Σ_n. □

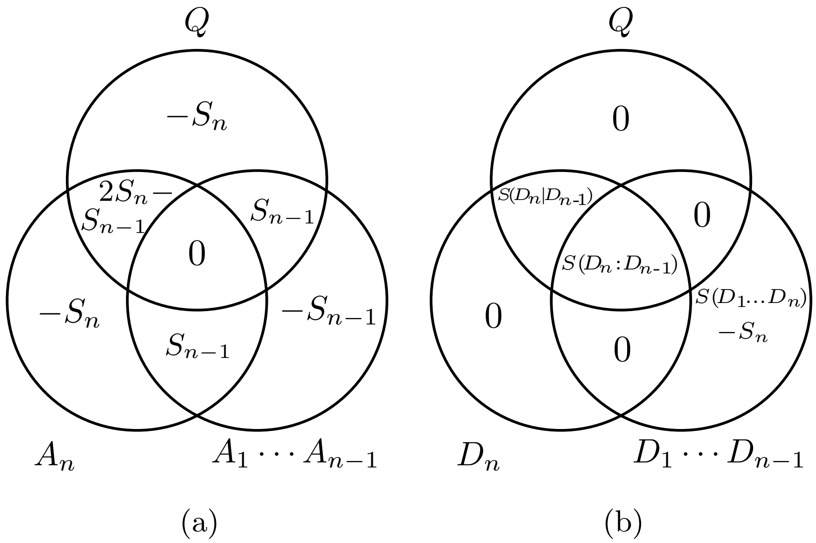

Figure 7.13: Diagram (a) before amplification with n ancillae that consecutively measured a quantum system Q, and (b) after amplification. The information $S(Q : A_n)$ is reduced to $S(Q : D_n)$, and $S(A_n : A_{n-1} \ldots A_1)$ to $S(D_n : D_{n-1} \ldots D_1)$. The zero ternary mutual entropy in (a) indicates that the underlying state of $QA_1 \ldots A_n$ is pure.

This section quantified a number of unsurprising, but nevertheless important results: amplifying measurements reduces information, and increases uncertainty about quantum states. The key quantity that characterizes the difference between unamplified and amplified chains is Σ_n, which quantifies how much we do *not* know about the state preparation, A_{n-1}, given the state determination, A_n. Depending on the relative state between A_{n-1} and A_n, we may know nothing ($\Sigma_n = 1$), or everything ($\Sigma_n = 0$). We summarize the results presented in this section with the entropy Venn diagrams in Fig. 7.13.

7.5 Applications of Consecutive Quantum Measurements

The formalism developed in this chapter can be directly applied to several interesting situations. Here, we focus specifically on the quantum Zeno effect and quantum state preparation.

7.5.1 Quantum Zeno and Anti-Zeno Effects

In the quantum Zeno effect [192,205,206], a quantum system that is observed repeatedly and sufficiently rapidly using projective measurements will be protected against state transitions from its initial state. If a series of n measurements are made in a time T (so that the measurements are spaced in time by T/n), then the probability the state will survive goes to one as $n \to \infty$. In this section, we derive results for the quantum Zeno and anti-Zeno effects in the context of unitary consecutive measurements. Instead of the standard approach where a time-varying quantum state is controlled by quantum measurements of the same observable, we study a static quantum state consecutively measured by quantum detectors where the measured observable changes in time. At the end of the calculation, we show that these two perspectives are equivalent.

We start with a two-level quantum system in the state

$$|Q\rangle = \sqrt{p}\,|0\rangle + \sqrt{1-p}\,|1\rangle, \tag{7.114}$$

with arbitrary p, which is written in the eigenbasis of the first observable to be measured with the device D_1. It is then subsequently measured by $D_2 \ldots D_{n+1}$, with each observable's eigenbasis at an angle $\pi/(4n)$ relative to the previous one, completing a full $\pi/4$ rotation after n observations. The density matrix for the preparation with D_1 (equivalently for Q) is $\rho(D_1) = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$, which has an entropy $S(D_1) = -p\log_2 p - (1-p)\log_2(1-p)$. The density matrix for the second device is

$$\rho(D_2) = \sum_j \left(p\,|U_{0j}|^2 + (1-p)\,|U_{1j}|^2\right)|j\rangle\langle j|\,, \tag{7.115}$$

where the transformation between the two observables' eigenbases can be written

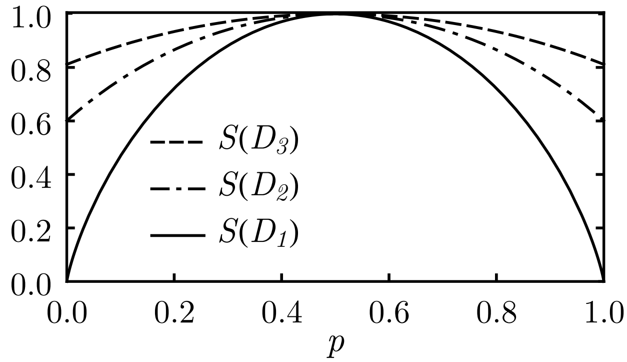

Figure 7.14: Entropies for two consecutive measurements D_2 and D_3, after the preparation with D_1, of the quantum state (7.114). The variable p characterizes the initial state of Q. The eigenbasis of each observable is at an angle of $\pi/8$ relative to the previous one ($n = 2$).

$$U = \begin{pmatrix} \cos(\frac{\pi}{4n}) & -\sin(\frac{\pi}{4n}) \\ \sin(\frac{\pi}{4n}) & \cos(\frac{\pi}{4n}) \end{pmatrix} . \tag{7.116}$$

The entropy of the second device is $S(D_2) = -q^{(2)} \log_2 q^{(2)} - (1 - q^{(2)}) \log_2(1 - q^{(2)})$ with

$$q^{(2)} = 1/2 + (p - 1/2) \cos\left(\frac{\pi}{2n}\right) , \tag{7.117}$$

the probability to observe the state $|0\rangle$ for the second measurement. Figure 7.14 shows the entropies $S(D_1)$, $S(D_2)$, and $S(D_3)$ for measurements using D_2 and D_3 after the preparation with device D_1.

In general, following the preparation, the probability $q^{(n+1)}$ to observe the state $|0\rangle$ after *all* n measurements is

$$q^{(n+1)} = \frac{1}{2} + \left(p - \frac{1}{2}\right) \cos^n\left(\frac{\pi}{2n}\right) \to p \quad \text{as } n \to \infty . \tag{7.118}$$

We see that the measurement probabilities after the sequence of measurements are the same as those of the initial quantum system and the density matrix of the last device is equal to that of the preparation with D_1. For polarization measurements, for example, this results in perfect transmission of an initially polarized beam even though the n polarization rotators would eventually rotate the polarization to an orthogonal state [207].

In the usual description of the quantum Zeno effect, the quantum state evolves unitarily in time between a sequence of measurements made in the same basis. Suppose that after the first measurement using D_1, the quantum state is rotated through an angle $-\pi/(4n)$ (using, e.g., polarization rotators in an optics setting to rotate a photon's polarization [208]) according to (7.116). Applying such a transformation to the quantum state, $U(-\pi/4n)\otimes\mathbb{1}\,|QD_1\rangle$, and then measuring a second time using D_2 in the same basis as the first measurement yields a density matrix for D_2 (and Q) with probabilities identical to (7.117). Therefore, both descriptions are equivalent since the relative angles between the states are the same in each case.

The anti-Zeno effect is often described as the complete destruction of a quantum state due to incoherent consecutive measurements [209–211]. In the present language, this corresponds to the randomization of a given (prepared) quantum state after consecutive measurements at random angles with respect to the initial state. We begin again with the prepared state (7.114), but now observe it consecutively using measurement devices D_k and relative angles θ_k drawn from a uniform distribution on the interval $[0,\pi/4]$. The probability to observe the state $|0\rangle$ after all n measurements with random angles is now

$$q^{(n+1)} = \frac{1}{2} + \left(p - \frac{1}{2}\right)\Pi_{k=1}^{n}\cos(2\theta_k)\ . \tag{7.119}$$

In order to obtain the most likely state probability we calculate the expectation value,

$$E\left[\Pi_{k=1}^{n}\cos(2\theta_k)\right] = \Pi_{k=1}^{n}E\left[\cos(2\theta_k)\right] = \left(\frac{2}{\pi}\right)^{n}, \tag{7.120}$$

so that $E\left[q^{(n+1)}\right] \to 1/2$ as $n \to \infty$. Thus, any quantum state is randomized via consecutive quantum measurements in random bases. A similar result was derived for the dephasing of photon polarization in Ref. [207].

7.5.2 Preparing Quantum States

For our final application, we discuss how to prepare quantum states by considering consecutive measurements on unprepared quantum states. Suppose a quantum system is prepared in the known state

$$\rho(Q) = \sum_{x=0}^{d-1} p_x \, |\widetilde{x}\rangle\langle\widetilde{x}| \,, \tag{7.121}$$

which we already wrote in the eigenbasis of the first observable to be measured after the preparation. We can always prepare a state like (7.121) by measuring an unprepared quantum state (5.79), with the pair A_1D_1 in a given, but arbitrary basis. Then, a second measurement with A_2D_2 of an observable at a relative angle θ_2 gives rise to the state

$$|QRA_1D_1A_2D_2\rangle = \frac{1}{\sqrt{d}} \sum_{x_1x_2} U^{(2)}_{x_1x_2} \, |\widetilde{x}_2 \; x_1 \, x_1x_1 \, x_2x_2\rangle. \tag{7.122}$$

From this we can compute the operator [59] describing the state of the quantum system, conditional on the state of the first device, D_1,

$$\begin{aligned} \rho(Q|D_1) &= \rho(QD_1)\left(\rho(D_1)^{-1} \otimes \mathbb{1}_Q\right) \\ &= \sum_{x_1} \rho_Q^{x_1} \otimes |x_1\rangle\langle x_1| \,, \end{aligned} \tag{7.123}$$

where $\rho(D_1)^{-1}$ is the inverse of the density matrix. The density matrix $\rho_Q^{x_1}$ is the prepared state (7.121) of the quantum system, given that the outcome x_1 was observed in the first measurement,

$$\rho_Q^{x_1} = \frac{\mathrm{Tr}_{D_1}\Big[P_{x_1}\rho(QD_1)P_{x_1}^\dagger\Big]}{\mathrm{Tr}_{QD_1}\Big[P_{x_1}\rho(QD_1)P_{x_1}^\dagger\Big]} = \sum_{x_2} |U_{x_1 x_2}^{(2)}|^2\, |\widetilde{x}_2\rangle\langle\widetilde{x}_2|. \tag{7.124}$$

Here, $P_{x_1} = |x_1\rangle\langle x_1|$ are projectors on the state of D_1. If we choose for the quantum state preparation the outcome $x_1 = 0$, for example, then $p_{x_2} = |U_{0x_2}^{(2)}|^2$ provides the probability distribution for the quantum system, and we arrive at the desired prepared state (7.121) from (7.124).

The purification of (7.121) in terms of the orthogonal states of ancilla A_2 is

$$|QA_2\rangle = \sum_{x_2} \sqrt{p_{x_2}}\; |\widetilde{x}_2\rangle|x_2\rangle, \tag{7.125}$$

which is an entangled state with the marginal entropies $S(Q)=S(A_2)=H[p]$. If we rename A_2 to A_1, then expression (7.125) is equivalent to (5.75). Equipped with this state preparation, we can now make the usual consecutive (amplified or unamplified) measurements of Q with A_2D_2, A_3D_3, etc.

7.6 Conclusions

Conventional wisdom in quantum mechanics dictates that the measurement process "collapses" the state of a quantum system so that the probability that a particular detector fires depends only on the state preparation and the measurement chosen. This assertion can be tested by considering sequences of measurements of the same quantum system. If a "memory" of the first measurement (the state preparation) persists beyond the second measurement, then a reduction of the wave packet can be ruled out. We discussed two classes of quantum measurement: those performed within a closed system where every part of a measurement device is under control (unamplified measurements), and those performed within an open system, where part of the pointer variable is ignored (amplified measurements). We found that sequences of quantum measurements in closed systems are non-Markovian (retaining the memory of past measurements) while sequences of open-system measurements obey the Markov property. In the latter case, the probability distribution of future measurement results only depends on the state preparation and the measurement chosen. It is clear from our construction that the Markovian measurements are a special case of the non-Markovian ones, and that the loss of memory is not a fundamental property of quantum measurements, but is merely a consequence of the loss of quantum information when tracing over degrees of freedom that participated in the measurement. We quantified this loss by calculating the amount of information lost when observing coherent quantum ancillae using incoherent devices.

We have found that the entropy of coherent chains of measurements is entirely determined by the entropy at the boundary of the chain, namely the entropy of the state *preparation* (the first measurement in the chain) and the last measurement. (If the chain is started on a known state, then the entropy of the chain is contained in the last measurement only). This property is a direct consequence of the unitarity of quantum measurements, and signifies that any quantum measurement outcome is constrained by its immediate past and its immediate future. It has not escaped our attention that this property of quantum chains is reminiscent of the holographic principle, which posits that the description of a system can be encoded entirely on its boundary alone. Because the holographic principle is often thought to have its origin in an information-theoretic description of space-time [212], it is perhaps not surprising that an information-theoretic analysis of chains of measurements would yield precisely such an outcome. In particular, it is perhaps not too hard to imagine that the past-future relationship that consecutive quantum measurements entail create precisely the partial order required for the "causal sets" program for quantum gravity [213]. Of course, to recover space-time from sets of measurements we would need to consider not just sequential measurements on the same system, but multiple parallel chains that are entangled with each other, creating a network rather than a chain (we have already shown in Ch. 6 that the unitary formalism deployed here can be extended to parallel measurements when discussing the Bell-state quantum eraser [141]). In that respect, the network of quantum measurements is more akin to van Raamsdonk's [214] tensor networks, created using entangling and disentangling operations (see also [215]). Incidentally, the present formalism implies the existence of a disentangling operation for consecutive quantum measurements that can serve as a primitive for generating remote quantum entanglement [183] (see Ch. 8).

Using a quantum-information-theoretic approach, we have argued that a collapse picture makes predictions that differ from those of the unitary (relative state) approach if multiple consecutive non-Markovian measurements are considered. Should future experiments corroborate the manifestly unitary formulation we have outlined, such results would further support the notion of the reality of the quantum state [216] and that the wave function is not merely a bookkeeping device that summarizes an observer's knowledge about the system [195, 196]. We hope that moving discussions about the nature of quantum reality from philosophy into the empirical realm will ultimately lead to a more complete (and satisfying) understanding of quantum physics.

Chapter 8

Quantum Disentangling Operations

8.1 Introduction

One of the fundamental results of the consecutive measurement scheme discussed in Ch. 7 is that the chain of quantum ancillae is coherent as long as it remains unamplified. That is, the joint density matrices for the unamplified quantum ancillae are non-diagonal in the ancilla product basis and, specifically, are classical-quantum states. In this chapter, I seek to exploit this fact and construct a general disentangling protocol in which a chain of ancillae can be partially disentangled using a set of unitary operations. I show that disentangling operations are in fact quite common and appear in many quantum protocols, including quantum teleportation. I first start by outlining a general method for disentangling ancilla chains using local operations and consider two different cases that depend on whether the quantum system was initially prepared in a pure state or started in a maximum entropy mixed state. Finally, I use the general construction to establish a protocol for generating remote entanglement.

8.2 One-Bit Disentangling Scheme

The first disentangling scheme considered here requires operations that are conditional on the state of a single qubit. We know from Ch. 7.2.3 that if $j+1$ consecutive measurements are made on a prepared quantum system, then the joint entropy of the resulting ancilla chain is determined only by the last ancilla in the sequence. This is a consequence of the specific coherent structure of the ancilla density matrices. We will first review the relevant results of consecutive quantum measurements before introducing a protocol to disentangle the first j ancillae from the rest of the system. In particular, we can understand the essence of the protocol by studying a sequence of just two measurements.

We saw previously in Ch. 7.2.1 that for two measurements of a prepared quantum system, Q, using ancillae A and B, the combined wave function is

$$|QAB\rangle = \sum_{ij} \alpha_i \, U_{ij} \, |j \, ij\rangle. \tag{8.1}$$

Ignoring the readout stage with the devices A' and B', the density matrix for the quantum ancillae is

$$\rho_{AB} = \sum_{ii'j} \alpha_i \, \alpha^*_{i'} \, U_{ij} \, U^*_{i'j} \, |i\rangle\langle i'| \otimes |j\rangle\langle j|. \tag{8.2}$$

This state has a particular structure that will make it possible to disentangle ancilla A from the rest of the system (Q and B). To see this, we recall from Ch. 7.2.3 that we can diagonalize the matrix (8.2) using the set of non-orthogonal states,

$$\alpha'_j \, |\psi_j\rangle = \sum_i \alpha_i \, U_{ij} \, |i\rangle, \tag{8.3}$$

with normalization equal to the probability distribution for ancilla B,

$$|\alpha'_j|^2 = q_j = \sum_i |\alpha_i|^2 \, |U_{ij}|^2. \tag{8.4}$$

Thus, Eq. (8.2) can also be written in the classical-quantum form,

$$\rho_{AB} = \sum_j q_j \; \rho_j \otimes |j\rangle\langle j| \,, \tag{8.5}$$

where $\rho_j = |\psi_j\rangle\langle\psi_j|$ are the (conditional) pure states of A.

As a result of the classical-quantum structure of the state (8.5), the entropy of the sequence of ancillae (in this case, AB) is contained in just the last ancilla, B. That is, $S_{AB} = S_B$, so that the entropy of A, conditional on B, vanishes,

$$S(A|B) = S_{AB} - S_B = 0. \tag{8.6}$$

This zero conditional entropy suggests the possibility of transforming the joint state of A and B in such a way that A becomes disentangled from the rest of the system. In particular, the vanishing conditional entropy suggests that there exists a unitary operation, V, on ρ_{AB} that leads to

$$V \, \rho_{AB} \, V^\dagger = \rho'_{AB} = \rho'_A \otimes \rho'_B, \tag{8.7}$$

where

$$\rho'_A = |\psi\rangle\langle\psi|. \tag{8.8}$$

As a result, ancilla A is forced into the pure state $|\psi\rangle$, and is in a product state with B. In fact, A has been disentangled from both B and Q.

The operator that performs the transformation (8.7) is a controlled unitary of the form

$$V = \sum_j V^{(j)} \otimes P_j, \tag{8.9}$$

where $P_j = |j\rangle_B\langle j|$ are projectors on B and $V^{(j)}$ are d unitary operators on ancilla A that depend on the state j of B.

If we choose the set of operators $V^{(j)}$ such that they describe a transformation between each of the conditional states $|\psi_j\rangle$ to a new state $|\psi\rangle$,

$$|\psi_j\rangle = V^{(j)\dagger}\,|\psi\rangle, \tag{8.10}$$

then (8.5) is rewritten as

$$\rho_{AB} = \sum_j q_j\; V^{(j)\dagger}\,|\psi\rangle\langle\psi|\,V^{(j)} \otimes |j\rangle\langle j|. \tag{8.11}$$

Clearly, applying (8.9) transforms this state to

$$V\,\rho_{AB}V^\dagger = \rho'_{AB} = |\psi\rangle\langle\psi| \otimes \sum_j q_j\,|j\rangle\langle j|, \tag{8.12}$$

where A is fully disentangled from B [and also from Q via the symmetry between Q and B in the wave function (8.1)] and is in the pure state $|\psi\rangle$.

It is straightforward to generalize this scheme to an arbitrarily-long sequence of measurements. It turns out that the first $n-1$ ancillae can be disentangled from the nth ancilla and Q. In general, this requires a set of operations analogous to (8.10) but for the joint state of $n-1$ ancillae, which may be rather difficult to implement in practice. The entropy Venn diagram for the disentangling scheme is shown in Fig. 8.1 for a general chain $A_1, A_2, \ldots, A_j, A_{j+1}$. Notice how the conditional entropy of the first $n-1$ ancillae always vanishes.

To see how this disentangling procedure might work, we consider the case of qubit measurements of two observables with eigenbases at the relative angle $\theta = \pi/4$. The non-

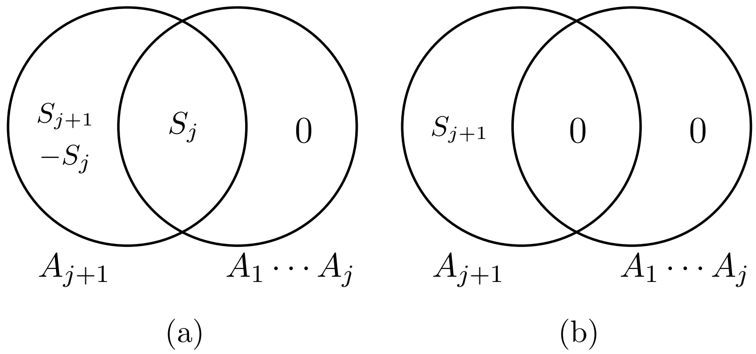

Figure 8.1: Correlations of an arbitrarily-long chain of ancillae A_1, A_2, ... , A_j, A_{j+1} (a) before and (b) after disentangling. The two zeros in (b) indicate that the state of A_1, ... , A_j is pure and disentangled from A_{j+1}.

orthogonal states $|\psi_j\rangle$ of ancilla A are written

$$\begin{aligned}\alpha_0'\,|\psi_0\rangle &= \frac{1}{\sqrt{2}}\big(\alpha_0|0\rangle + \alpha_1|1\rangle\big),\\ \alpha_1'\,|\psi_1\rangle &= -\frac{1}{\sqrt{2}}\big(\alpha_0|0\rangle - \alpha_1|1\rangle\big).\end{aligned} \tag{8.13}$$

Suppose we choose the disentangled state of A to be $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$. Then, the set of conditional operators, $V^{(j)}$, on qubit A is

$$\begin{aligned}V^{(0)\dagger} &= \mathbb{1},\\ V^{(1)\dagger} &= -Z,\end{aligned} \tag{8.14}$$

where Z is the Pauli matrix. The full operator V is then simply a controlled-phase gate, where the control is on the state of B. In terms of these operators, the joint state of AB is

$$\rho_{AB} = q_0\,|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0| + q_1\,Z|\psi\rangle\langle\psi|Z \otimes |1\rangle\langle 1|, \tag{8.15}$$

so that applying V to this state yields (8.12).

If qubits A and B are spatially separated, then one could implement this operation in

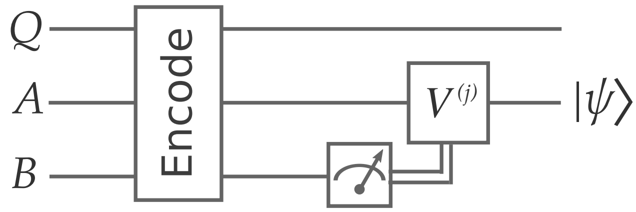

Figure 8.2: Circuit for disentangling the joint state $\rho_{AB} \rightarrow \rho'_{AB} = |\psi\rangle\langle\psi| \otimes \rho'_B$. The encoding scheme consists of a series of two measurements of Q with ancillary qudits A and B. The (conditional) operator shown acting on A is one of the set of d operators $V^{(j)}$.

practice using local operations and classical communication (LOCC). Suppose Alice and Bob have qubits A and B, respectively. To perform the controlled unitary (8.9), Bob could measure (amplify) his qubit B with B' and send the result of the measurement (one bit) to Alice. Based on the value of the bit that Bob sends, Alice would perform one of the two operations $V^{(j)}$ on her qubit. Afterwards, Alice has disentangled A from B (and Q). In this case, the degrees of freedom of system B' should be accounted for, and the final state of QB will no longer be pure, but mixed. See Fig. 8.2 for the corresponding circuit diagram. The "encode" operation refers to the sequence of two measurements with ancillary qubits A and B, while the double solid line indicates the single bit of information that Bob communicates to Alice about the state of qubit B.

To consider more general disentangling scenarios, we must keep in mind that a universal-NOT gate does not exist [217–220]. That is, it is not possible to construct a unitary gate that flips *any* input state to its orthogonal version. For example, σ_x will correctly flip the computational states $|0\rangle$, $|1\rangle$, but does not flip $|0\rangle \pm |1\rangle$ to $|0\rangle \mp |1\rangle$ (instead, σ_z would). Flipping a qubit to an orthogonal state is equivalent to an inversion of the Bloch sphere. Such an inversion preserves angles and corresponds, via Wigner's theorem [221], either to a unitary or antiunitary operation. Proper rotations (determinant equal to 1) are implemented with unitary operations, while orthogonal transformations with determinant -1 correspond to antiunitary operations. Therefore, since it is not possible to construct a unitary operation that flips the state of an arbitrary qubit, there is no universal-NOT gate.

However, if we restrict our states to the equatorial plane ($\theta = \pi/2$) or the xz plane ($\phi = 0$) of the Bloch sphere, then it is possible to unitarily flip the state of an arbitrary qubit. We can show this by considering two states $|\psi_i\rangle = a_i|0\rangle + b_i|1\rangle$, with $i = 1, 2$. A unitary transformation U to an orthogonal state, $U|\psi_i\rangle = b_i^*|0\rangle - a_i^*|1\rangle = |\psi_i^\perp\rangle$, must correspond to $\langle\psi|U^\dagger U|\phi\rangle = \langle\psi|\phi\rangle$ so that either $\theta = \pi/2$ or $\phi = 0$. For general angles, the transformation ($U \to A$) is antiunitary, $A|\psi_i\rangle = b_i^*|0\rangle - a_i^*|1\rangle = |\psi_i^\perp\rangle$, and has $\langle\psi|A^\dagger A|\phi\rangle = \langle\psi|\phi\rangle^*$. In the following protocols, we restrict the qubit states to the xz plane so that if the conditional states $|\psi_j\rangle$ are transformed to an orthogonal state, we can still construct unitary operations to flip them. Of course, if we are not transforming the states $|\psi_j\rangle$ into orthogonal states, then we can always find a unitary operation.

In a general qubit scenario, we can implement the disentangling operation by first transforming the conditional state $|\psi_1\rangle$ into $|\psi_0\rangle$. If we assume real coefficients (so that we are in the xz plane), the two non-orthogonal states of qubit A are

$$\begin{aligned} |\psi_0\rangle &= a'|0\rangle + b'|1\rangle, \\ |\psi_1\rangle &= -c'|0\rangle + d'|1\rangle, \end{aligned} \tag{8.16}$$

where $a' = a/\sqrt{a^2 + b^2}$ (and similarly for the other three coefficients) and where $a = \alpha_0 \cos\theta$, $b = \alpha_1 \sin\theta$, $c = \alpha_0 \sin\theta$ and $d = \alpha_1 \cos\theta$.

We can rotate these states into any other state in the xz plane of the Bloch sphere with a rotation about the y-axis by an angle α. The state $|\psi_1\rangle$ is transformed into $|\psi_0\rangle$ by the rotation

$$R_y(\alpha) = \cos(\alpha/2)\mathbb{1} - i\,\sin(\alpha/2)\,Y, \tag{8.17}$$

at an angle

$$\alpha = 2\cos^{-1}\left(b'd' - a'c'\right). \tag{8.18}$$

In terms of this rotation operator, the state of AB is

$$\rho_{AB} = q_0\,|\psi_0\rangle\langle\psi_0| \otimes |0\rangle\langle 0| + q_1\,R_y^\dagger(\alpha)|\psi_0\rangle\langle\psi_0|R_y(\alpha) \otimes |1\rangle\langle 1|. \tag{8.19}$$

The operator that disentangles qubit A is

$$V = \mathbb{1} \otimes |0\rangle\langle 0| + R_y(\alpha) \otimes |1\rangle\langle 1|, \tag{8.20}$$

and transforms ρ_{AB} into

$$V\rho_{AB}V^\dagger = |\psi_0\rangle\langle\psi_0| \otimes \sum_j q_j\,|j\rangle\langle j|, \tag{8.21}$$

where it is clear that A is pure [the first state in (8.16)] and in a product state with B.

8.3 Two-Bit Disentangling Scheme

In the previous section we saw that the joint ancilla states resulting from measurements of prepared quantum states required only operations conditional on the state of the last ancilla in the chain to disentangle A from the rest of the system. In this section, we extend this procedure to the states resulting from measurements of unprepared quantum states. We will see that, in this case, the disentangling operations are now conditional on the states of the last *and* first ancillae. In particular, we will study three consecutive measurements and find that if the ancillae are spatially separated, this feature corresponds to the communication of two bits of information to place ancilla B in a product state with A and C.

Recall from Ch. 7.2.2 that the joint ancilla state for the three ancilla that measured an unprepared quantum system is non-diagonal in the ancilla product basis. It can, however, be written in a diagonal and classical-quantum form,

$$\rho_{ABC} = \frac{1}{d}\sum_{ik} |\epsilon_{ik}|^2 \, |i\rangle\langle i| \otimes \rho_{ik} \otimes |k\rangle\langle k| \, , \tag{8.22}$$

where $\rho_{ik} = |\phi_{ik}\rangle\langle\phi_{ik}|$ are the (conditional) pure states of B. This set of non-orthogonal states is defined as

$$\epsilon_{ik} \, |\phi_{ik}\rangle = \sum_j U_{ij} \, U'_{jk} \, |j\rangle \, , \tag{8.23}$$

which are normalized according to

$$|\epsilon_{ik}|^2 = \sum_j |U_{ij}|^2 \, |U'_{jk}|^2 \, . \tag{8.24}$$

We know from Ch. 7.2.3 that states such as (8.22) have the property that any intermediate ancillae (in this case, just B) do not contribute to the total entropy. That is, the joint entropy of (8.22) is $S_{ABC} = S_{AC}$, so that the entropy of the ancilla chain resides only in the boundaries. It follows that the conditional entropy of the intermediate ancilla vanishes,

$$S(B|AC) = S_{ABC} - S_{AC} = 0. \tag{8.25}$$

In other words, the state of the intermediate ancilla, B, is fully known (has zero entropy) when given the joint states of the first and last ancillae, A and C.

The presence of a vanishing conditional entropy suggests that there exists a unitary operation, V, that disentangles the joint state (8.22), transforming B into an unconditionally pure state, $|\phi\rangle$. That is,

$$V\ \rho_{ABC}\ V^{\dagger} = \rho'_{ABC} = \rho'_{AC} \otimes \rho'_{B}\,, \tag{8.26}$$

where

$$\rho'_{B} = |\phi\rangle\langle\phi|\,. \tag{8.27}$$

Now the intermediate ancilla, B, is in the pure state $|\phi\rangle$ and is in product with the rest of the system, AC (and also Q due to the symmetry between Q and C in the underlying wave function). After disentangling, the conditional entropy $S'(B|AC)$ still vanishes since $S'_{B} = 0$, so that $S'_{ABC} = S'_{AC}$.

The full unitary operator V is of the form

$$V = \sum_{ik} P_{ik} \otimes V^{(ik)}, \tag{8.28}$$

where $V^{(ik)}$ is a set of d^2 unitary operations on B, and $P_{ik} = |i\rangle_A\langle i| \otimes |k\rangle_C\langle k|$ are projectors on A and C. If we choose the set of operators, $V^{(ik)}$, such that they transform each of the states $|\phi_{ik}\rangle$ of ancilla B into a new state $|\phi\rangle$,

$$|\phi_{ik}\rangle = V^{(ik)\dagger}\ |\phi\rangle, \tag{8.29}$$

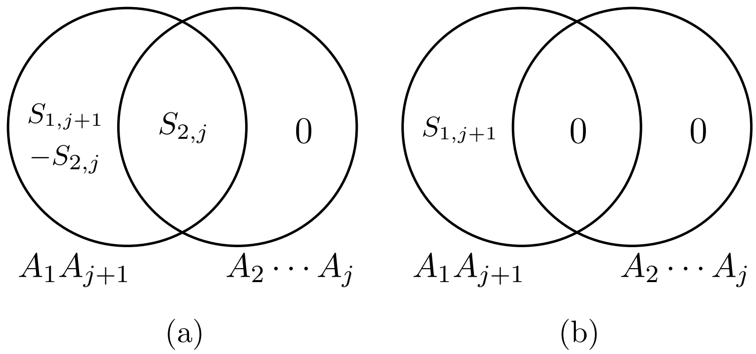

Figure 8.3: Correlations of an arbitrarily-long chain of ancillae A_1, A_2, ... , A_j, A_{j+1} (a) before and (b) after disentangling. The two zeros in (b) indicate that the state of A_2, ... , A_j is pure and disentangled from A_1 and A_{j+1}.

then (8.22) is rewritten as

$$\rho_{ABC} = \frac{1}{d} \sum_{ik} |\epsilon_{ik}|^2 \, |i\rangle\langle i| \otimes V^{(ik)\dagger} |\phi\rangle\langle\phi| \, V^{(ik)} \otimes |k\rangle\langle k|. \tag{8.30}$$

Clearly, applying the full operator (8.28) to this state disentangles ancilla B from AC,

$$V \rho_{ABC} V^{\dagger} = \rho'_{ABC} = \left(\frac{1}{d} \sum_{ik} |\epsilon_{ik}|^2 \, |i\rangle\langle i| \otimes |k\rangle\langle k| \right) \otimes |\phi\rangle\langle\phi|. \tag{8.31}$$

Similarly to the one-bit scheme, this scheme can be generalized to any number of measurements. The entropy Venn diagram for the two-bit version is shown in Fig. 8.3 for a general chain $A_1, A_2, \ldots, A_j, A_{j+1}$.

We now apply this disentangling scheme to the case of qubit measurements. The ability to force ancilla B into a pure state that is in a product state with the rest of the system suggests a simple protocol. Namely, a single-qubit state $|\phi\rangle$ can be extracted from the joint state of ABC by a sequence of unitary gates and measurements as shown in Fig. 8.4. In such a protocol, we assume that we have access to an unprepared quantum state. Three parties

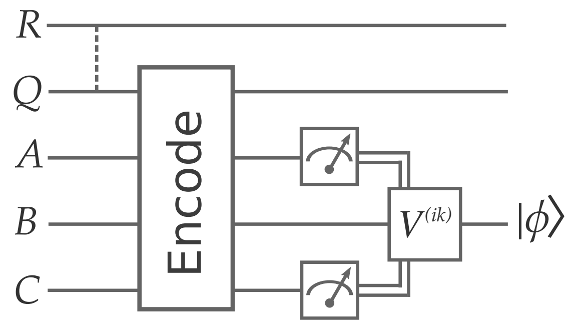

Figure 8.4: Circuit for disentangling the ancilla state $\rho_{ABC} \to \rho'_{ABC} = \rho'_{AC} \otimes |\phi\rangle\langle\phi|$. The encoding scheme consists of three consecutive measurements of Q with ancillary qudits A, B, and C. The (conditional) operator shown acting on B is one of the set of d^2 operators, $V^{(ik)}$. The initial entanglement between Q and the purifying reference R is indicated by the dashed line.

(Alice, Bob and Charlie) that each possess a qubit, measure an observable of Q, creating the entangled state ρ_{ABC} in (8.22). Alice and Charlie each amplify their qubit with a second qubit A' and C', respectively, and then send the results of their measurements (two classical bits i and k) to Bob. Based on the values of the two bits, Bob can perform one of the four unitary transformations, $V^{(ik)}$, on his qubit B. Following these operations, Bob has produced the pure state $|\phi\rangle$ that is in a product state with the rest of the system.

Suppose that the measurements of Q with ancillae B and C were of observables with relative angles $\theta = \theta' = \pi/4$. In this case, the four possible states, $|\phi_{ik}\rangle$, of ancilla B are

$$
\begin{aligned}
|\phi_{00}\rangle &= \frac{1}{\sqrt{2}}\Big(|0\rangle - |1\rangle\Big), \\
|\phi_{01}\rangle &= -\frac{1}{\sqrt{2}}\Big(|0\rangle + |1\rangle\Big), \\
|\phi_{10}\rangle &= \frac{1}{\sqrt{2}}\Big(|0\rangle + |1\rangle\Big), \\
|\phi_{11}\rangle &= -\frac{1}{\sqrt{2}}\Big(|0\rangle - |1\rangle\Big).
\end{aligned}
\tag{8.32}
$$

If we choose the disentangled state of ancilla B to be $|\phi\rangle = |0\rangle$, the corresponding conditional operators are rotations about the $\hat{y}$ axis by an angle $\pm\pi/2$,

$$\begin{aligned}
V^{(00)\dagger} &= \frac{1}{\sqrt{2}}\big(\mathbb{1} + iY\big) = R_{\hat{y}}(-\pi/2), \\
V^{(01)\dagger} &= -\frac{1}{\sqrt{2}}\big(\mathbb{1} - iY\big) = -R_{\hat{y}}(\pi/2), \\
V^{(10)\dagger} &= \frac{1}{\sqrt{2}}\big(\mathbb{1} - iY\big) = R_{\hat{y}}(\pi/2), \\
V^{(11)\dagger} &= -\frac{1}{\sqrt{2}}\big(\mathbb{1} + iY\big) = -R_{\hat{y}}(-\pi/2).
\end{aligned} \tag{8.33}$$

With these transformations, the state (8.22) is written as

$$\rho_{ABC} = \frac{1}{4}\sum_{ik} |i\rangle\langle i| \otimes V^{(ik)\dagger}\,|0\rangle\langle 0|\,V^{(ik)} \otimes |k\rangle\langle k|. \tag{8.34}$$

Applying the full operator (8.28) to the state (8.34) disentangles ancilla B from the rest of the system,

$$V\rho_{ABC}V^{\dagger} = \frac{1}{2}\,\mathbb{1}_A \otimes |0\rangle\langle 0| \otimes \frac{1}{2}\,\mathbb{1}_C\,, \tag{8.35}$$

where $1/2\,\mathbb{1}$ are maximally mixed states.

Note that one could also select the disentangled state to be $|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, instead of $|\phi\rangle = |0\rangle$ in the example above. Then, the corresponding set of conditional operators (for measurements at the relative angles $\theta = \theta' = \pi/4$) can be written as

$$\begin{aligned}
V^{(00)\dagger} &= Z, \\
V^{(01)\dagger} &= -X, \\
V^{(10)\dagger} &= \mathbb{1}, \\
V^{(11)\dagger} &= -ZX = -iY,
\end{aligned} \tag{8.36}$$

instead of (8.33).

In a more general qubit scenario, suppose we implement the disentangling operation by first transforming $|\phi_{01}\rangle$, $|\phi_{10}\rangle$ and $|\phi_{11}\rangle$ into $|\phi_{00}\rangle$. If we assume real coefficients so that we are in the xz plane of the Bloch sphere, the four non-orthogonal states of qubit B are

$$\begin{aligned}
|\phi_{00}\rangle &= a'|0\rangle - b'|1\rangle, \\
|\phi_{11}\rangle &= -b'|0\rangle + a'|1\rangle, \\
|\phi_{01}\rangle &= -c'|0\rangle - d'|1\rangle, \\
|\phi_{10}\rangle &= d'|0\rangle + c'|1\rangle.
\end{aligned} \tag{8.37}$$

Here, we defined $a' = a/\sqrt{a^2+b^2}$ (and similarly for the other three coefficients) and where $a = \cos\theta\,\cos\theta'$, $b = \sin\theta\,\sin\theta'$, $c = \cos\theta\,\sin\theta'$ and $d = \sin\theta\,\cos\theta'$. First, we note that two of the states can be written in terms of the other two as

$$\begin{aligned}
|\phi_{11}\rangle &= X|\phi_{00}\rangle, \\
|\phi_{01}\rangle &= -X|\phi_{10}\rangle,
\end{aligned} \tag{8.38}$$

so that we only have to find a rotation for $|\phi_{10}\rangle$ into $|\phi_{00}\rangle$. Using the results from the one-bit disentangling scheme, we write the transformation as a rotation about the y-axis by an angle

$$\alpha = 2\cos^{-1}\left(a'd' - b'c'\right). \tag{8.39}$$

Thus, with this set of transformations, qubit B ends in the state $|\phi_{00}\rangle$.

8.4 Examples of Disentangling Operations

It turns out that many quantum protocols utilize a disentangling scheme similar to what was described previously. Here, we discuss two well-known examples, the Garisto-Hardy disentanglement eraser and the teleportation protocol.

8.4.1 Quantum Disentanglement Erasers

The disentanglement eraser was formulated by Garisto and Hardy in 1999 [164] as a way to recover entanglement in a quantum system that was destroyed through additional correlations with another system. In the disentanglement eraser, we start with the entangled state

$$|AB\rangle = \frac{1}{\sqrt{2}}\Big(|00\rangle + |11\rangle\Big). \tag{8.40}$$

Additional correlations are introduced by tagging the components of this state with a tagger qubit T using a controlled-NOT gate (A or B can be the control qubit):

$$U_{CNOT}\,|AB\rangle\,|0\rangle_T = |ABT\rangle = \frac{1}{\sqrt{2}}\Big(|000\rangle + |111\rangle\Big). \tag{8.41}$$

The initial entanglement of AB has been destroyed since it has become completely mixed, $\rho_{AB} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$.

The entanglement can be restored by measuring the tagger T with another ancilla T' in some basis. If one chooses the basis $|\pm\rangle_{T'} = (|0\rangle_{T'} \pm |1\rangle_{T'})/\sqrt{2}$ then the joint state becomes

$$\begin{aligned} U\,|ABT\rangle\,|+\rangle_{T'} &= |ABTT'\rangle \\ &= \frac{1}{\sqrt{2}}\left[\frac{|00\rangle_{AB} + |11\rangle_{AB}}{\sqrt{2}}\,|++\rangle_{TT'} + \frac{|00\rangle_{AB} - |11\rangle_{AB}}{\sqrt{2}}\,|--\rangle_{TT'}\right]. \end{aligned} \tag{8.42}$$

If one registers the outcome $|+\rangle_{T'}$, then the entangled state $|00\rangle_{AB}+|11\rangle_{AB}$ is fully recovered.

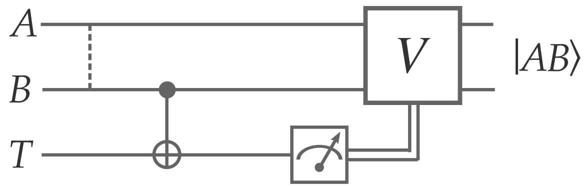

Figure 8.5: Garisto-Hardy disentanglement eraser. After tagging with system T the entanglement of AB (indicated by the dashed line) is lost. Measuring T with T' in a rotated basis and applying a conditional unitary to AB restores the original entangled state $|AB\rangle$.

However, for the outcome $|-\rangle_{T'}$, a phase shift restores the entanglement:

$$|ABTT'\rangle = \frac{1}{\sqrt{2}}\left[\left[\frac{|00\rangle_{AB} + |11\rangle_{AB}}{\sqrt{2}}\right]|{+}{+}\rangle_{TT'} + \left[\mathbb{1}_A \otimes \sigma_z \frac{|00\rangle_{AB} + |11\rangle_{AB}}{\sqrt{2}}\right]|{-}{-}\rangle_{TT'}\right]. \quad (8.43)$$

Conditional on the outcome of the T' measurement, one applies the appropriate unitary to AB to obtain

$$V\,|ABTT'\rangle = \left[\frac{|00\rangle_{AB} + |11\rangle_{AB}}{\sqrt{2}}\right]\left[\frac{|++\rangle_{TT'} + |--\rangle_{TT'}}{\sqrt{2}}\right]. \quad (8.44)$$

Thus, A and B are restored to their initial entangled state. The conditional unitary that implements the disentangling operations can be written as

$$V = \mathbb{1}_T \otimes \Big(\mathbb{1}_A \otimes \mathbb{1}_B \otimes |+\rangle_{T'}\langle+| + \mathbb{1}_A \otimes \sigma_z \otimes |-\rangle_{T'}\langle-|\Big). \quad (8.45)$$

A quantum circuit diagram for the disentanglement eraser is shown in Fig. 8.5.

8.4.2 Quantum Teleportation

The standard teleportation protocol was presented previously in Ch. 2. Recall from (2.35) that the initial state of the unknown qubit and the shared entangled pair is

$$|\psi\rangle \otimes |\beta_{00}\rangle = \frac{1}{2} \sum_{zx} |\beta_z x\rangle \otimes V^{(zx)\dagger} |\psi\rangle, \tag{8.46}$$

where $V^{(zx)\dagger} = X^x Z^z$. In this form, it is clear that the conditional unitary operator that disentangles Bob's qubit is

$$V = \sum_{zx} P_{zx} \otimes V^{(zx)}, \tag{8.47}$$

where $P_{zx} = |\beta_{zx}\rangle\langle\beta_{zx}|$ are projectors in the Bell basis. After applying (8.47) to (8.46), Bob's qubit is disentangled and is in the state $|\psi\rangle$.

8.5 Undoing Quantum Measurements

A natural question that we might ask is what the state of the quantum system is after the disentangling process. Specifically, whether it is equivalent to the situation where certain measurements were never performed so that the disentangling operation is equivalent to undoing a measurement. For example, recall the simple case of two measurements on a prepared quantum system. After disentangling ancilla A, the wave function is

$$V\,|QAB\rangle = \sum_j \alpha'_j \; |b_j\, j\rangle \otimes |\psi\rangle, \tag{8.48}$$

where the amplitudes satisfy

$$|\alpha'_j|^2 = \sum_i |\alpha_i|^2 \, |U_{ij}|^2. \tag{8.49}$$

If the measurement with A had never occurred, and instead we only measured Q using B, the resulting wave function would have been

$$|QB\rangle = \sum_j \widetilde{\alpha}'_j \, |b_j \, j\rangle, \tag{8.50}$$

with amplitudes that satisfy

$$|\widetilde{\alpha}'_j|^2 = \Big| \sum_i \alpha_i \, U_{ij} \Big|^2. \tag{8.51}$$

The difference between (8.49) and (8.51) is that the amplitudes in the disentangled state are an incoherent sum of terms, reflecting the fact that a measurement of Q had been performed before B. Thus, the disentangling operation is not equivalent to undoing past measurements. This means that if the measurement with A represented an "error" in a quantum algorithm, then disentangling A does not reset the state of QB to the pre-error state, and cannot be used as an error-correcting protocol.

8.6 Extracting Entanglement

The schemes in the previous sections can be extended to disentangle multi-qudit entangled states. Here, we consider three measurements of a prepared quantum system, from which an entangled state of AB can be extracted. Recall from Ch. 7.2.1 that the ancilla density matrix after such a set of measurements is

$$\rho_{ABC} = \sum_k q''_k \, |\psi_k\rangle\langle\psi_k| \otimes |k\rangle\langle k|, \tag{8.52}$$

where

$$\alpha''_k \, |\psi_k\rangle = \sum_{ij} \alpha_i \, U_{ij} \, U'_{jk} \, |ij\rangle, \tag{8.53}$$

and

$$|\alpha''_k|^2 = q''_k = \sum_{ij} |\alpha_i|^2 \, |U_{ij}|^2 \, |U'_{jk}|^2. \tag{8.54}$$

For instance, for $\theta = \pi/4$, the two conditional states of AB in Eq. (8.52) are

$$\begin{aligned} |\psi_0\rangle &= a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle, \\ |\psi_1\rangle &= b|00\rangle - a|01\rangle - d|10\rangle + c|11\rangle, \end{aligned} \tag{8.55}$$

where $a = \alpha_0 \cos\theta'$, $b = -\alpha_0 \sin\theta'$, $c = \alpha_1 \cos\theta'$ and $d = \alpha_1 \sin\theta'$. It is straightforward to show that applying $-Z \otimes iY$ to the second state yields the first state:

$$(-Z \otimes iY) \, |\psi_1\rangle = |\psi_0\rangle. \tag{8.56}$$

Thus, the operator

$$\begin{aligned} V &= \sum_k Z^k \otimes (-iY)^k \otimes |k\rangle\langle k| \\ &= \mathbb{1} \otimes \mathbb{1} \otimes |0\rangle\langle 0| - Z \otimes iY \otimes |1\rangle\langle 1|, \end{aligned} \tag{8.57}$$

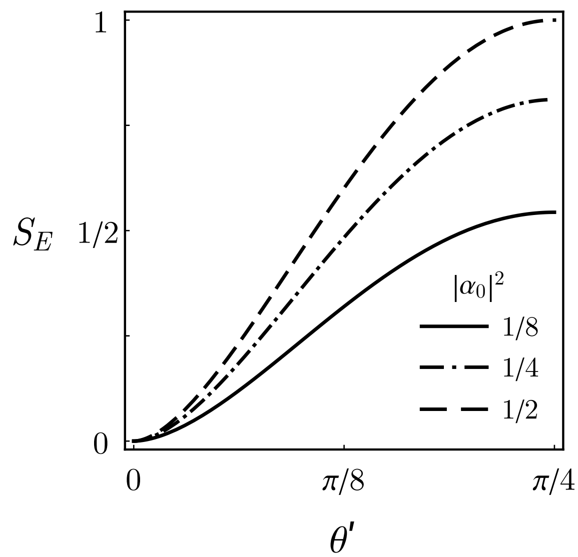

Figure 8.6: Entanglement entropy S_E of the disentangled state $|\psi_0\rangle$ of A and B in (8.55). Three curves plotted as a function of the second measurement angle θ' (with $\theta = \pi/4$) for different amplitudes α_0 of the initial quantum state. Maximum entanglement occurs at $\alpha_0 = 1/\sqrt{2}$ and $\theta' = \pi/4$.

leaves the state ρ_{ABC} disentangled,

$$V\rho_{ABC}V^\dagger = |\psi_0\rangle\langle\psi_0| \otimes \frac{1}{2}\mathbb{1}_C. \tag{8.58}$$

In this way, a two-qubit entangled state is extracted from the ancilla chain. The entanglement of $|\psi_0\rangle$ in (8.55) is characterized by the entanglement entropy S_E, computed from the entropy of the reduced density matrix for A or B. This is shown in Fig. 8.6 as a function of the measurement angle θ'. Evidently, increasing θ' produces a state that is more entangled. In particular, a maximally entangled state is produced when $\alpha_0 = 1/\sqrt{2}$ and $\theta' = \pi/4$.

8.7 Generating Remote Entanglement

Shared entanglement between spatially separated systems is an essential resource for quantum information processing including long-distance quantum cryptography and teleportation. Here, I describe a protocol for generating a maximally entangled state between remote locations that requires only local operations and does not rely on communication between the separated parties.

Much of quantum information processing relies on entanglement as a resource. For example, entanglement that is shared between distant parties is necessary to implement Ekert's quantum key distribution protocol for secure communication [222], to transfer quantum states using teleportation [7], or to establish large-scale quantum networks. Remote entanglement generation has been realized in many systems such as with optical photons [223–225], the nitrogen vacancy centers of solid state qubits [226], and superconducting qubits [227–229].

I describe a method for deterministically generating remote entanglement between two qubits using local operations on pairs of separated qubits. I consider two versions—one with and one without classical communication—and show how the degree of entanglement created can be tuned by the choice of encoding parameters.

8.7.1 Encoding Scheme

To generate remote entanglement, four ancillary qubits A, B, C, and D are first encoded via a sequence of unamplified measurements [70] of a quantum system Q. Such measurements are equivalent to the unitary entangling operations implemented in weak measurements [23, 24, 106, 116, 118, 119] (see Ch. 5), but the interaction considered here is strong (see, e.g., [117, 139]). Following the encoding, local operations are performed on the qubit pairs AB and CD, which are spatially separated from each other, such that a shared entangled state of BC is produced. The entanglement of this state is will depend on the details of the four measurements.

The density matrix of the qubit quantum system, Q, is taken to be proportional to the identity matrix, $\rho_Q = \frac{1}{2}\,\mathbb{1}$, so that it is a maximum entropy state. The four consecutive entangling operations between the quantum system and the ancillary qubits lead to the total wave function [59, 70],

$$|QRABCD\rangle = \frac{1}{\sqrt{2}} \sum_{ijk\ell} U_{ij}\, U'_{jk}\, U''_{k\ell}\, |\ell i\, ijk\ell\rangle, \tag{8.59}$$

where the initial mixed state of Q has been purified with a reference R and each system is of dimension two. The measured observables are characterized by the matrix elements of U, U', and U''. That is, the eigenbasis of the first observable is rotated relative to the eigenbasis of the second via $|i\rangle = \sum_j U_{ij}\,|j\rangle$, and similarly for the third and fourth observables with U' and U''. Since the measurements are strong, the final states of the ancillary qubits, $|i\rangle_A$, $|j\rangle_B$, $|k\rangle_C$, and $|\ell\rangle_D$, are orthogonal. If we consider only measurements of observables corresponding to the xz plane of the Bloch sphere, such a transformation can be implemented with a rotation by an angle θ according to

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{8.60}$$

and similarly for U' and U'' with angles θ' and θ'', respectively.

The encoding operation starts by first entangling qubits A and B with Q, with a relative angle of θ between the first two observables, and sending them to Alice. Qubits C and D are subsequently entangled with Q, with a relative angle θ'' between the last two observables, and are sent to Bob. As we will see, the intermediate angle θ' can be left arbitrary in the protocol so that it is not necessary for Alice or Bob to know the measurement bases the other chose.

The coherence of the chain of ancillary qubits will be used to create the remote entanglement. Their joint state can be written in terms of a new basis for B and C such that tracing (8.59) over Q and R it appears as

$$\rho_{ABCD} = \frac{1}{2} \sum_{i\ell} p_{i\ell} \, |i\rangle\langle i| \otimes |\phi_{i\ell}\rangle\langle\phi_{i\ell}| \otimes |\ell\rangle\langle\ell|. \tag{8.61}$$

The four non-orthogonal states of BC,

$$\epsilon_{i\ell} \, |\phi_{i\ell}\rangle = \sum_{jk} U_{ij} \, U'_{jk} \, U''_{k\ell} \, |jk\rangle, \tag{8.62}$$

are normalized according to

$$p_{i\ell} = |\epsilon_{i\ell}|^2 = \sum_{jk} |U_{ij}|^2 \, |U'_{jk}|^2 \, |U''_{k\ell}|^2. \tag{8.63}$$

Consistent with the choice of encoding parameters (θ, θ', and θ'') Alice and Bob construct a set of conditional unitary operations to perform locally on their pairs of qubits (AB and CD) to generate remote entanglement. At the end they share an entangled state of B and C that is in a product state with the rest of the system.

As local unitary operations alone cannot change the entanglement of a state (recall from Ch. 3.4.1 that local operations do not change the Schmidt coefficients in the Schmidt decomposition of a bipartite pure state and, therefore, do not change the entanglement [25]), remote entanglement can only be generated in this protocol by local operations on B and C when the entanglement entropies of each state $|\phi_{i\ell}\rangle$ in (8.62) are the same.

We consider two versions of the protocol. The first, which is more general, occurs when Alice chooses the relative angle $\theta = \pi/4$ (Bob could, equivalently, have picked $\theta'' = \pi/4$). We will see that, in this case, classical communication is required between Alice and Bob concerning the states of their qubits A and D (but *not* the states of B and C). Furthermore, the final degree of entanglement between B and C will depend only on Bob's measurement angle θ''. The second version is a special case of the previous scheme where both the first and third angles are $\theta = \theta'' = \pi/4$. Here, communication about the states of A and D will no longer be required and a maximally entangled state will be produced with only local operations on the two pairs of qubits.

8.7.2 Protocol with Communication

The first scenario requires Alice's measurement angle to be $\theta = \pi/4$, which necessitates classical information to be communicated between Alice and Bob. Using local operations on B and C in addition to classical communication about the states of A and D, they generate a new state of BC that is entangled according to Bob's measurement angle θ''.

To encode the ancillary qubits, Alice selects the relative angle $\theta = \pi/4$ while Bob's angle θ'' is left arbitrary. When $\theta = \pi/4$, the density matrix (8.61) can be written as

$$\rho_{ABCD} = \frac{1}{4} \sum_{i\ell} |i\rangle\langle i| \otimes V^{(i\ell)\dagger} |\phi_{00}\rangle\langle\phi_{00}| V^{(i\ell)} \otimes |\ell\rangle\langle\ell|, \tag{8.64}$$

where the coefficients $p_{i\ell} = 1/2$, and

$$V^{(i\ell)\dagger} = Z^{i+\ell} X^{\ell} \otimes X^{\ell} \tag{8.65}$$

is a conditional unitary operator on qubits B and C. Here, Z and X are Pauli operators and $(i + \ell)$ is modulo two.

The conditional joint states (8.62) of qubits BC are functions of the angles θ' and θ''. The key observation is that each of them can be written in terms of local operations on the state $|\phi_{00}\rangle$ as

$$|\phi_{i\ell}\rangle = V^{(i\ell)\dagger}|\phi_{00}\rangle = Z^{i+\ell}X^{\ell} \otimes X^{\ell}\,|\phi_{00}\rangle. \tag{8.66}$$

The $i = \ell = 0$ state can be written as

$$|\phi_{00}\rangle = -\sin\theta'\,|\widetilde{\beta}_{01}\rangle + \cos\theta'\,|\widetilde{\beta}_{10}\rangle, \tag{8.67}$$

where the states

$$\begin{aligned}
|\widetilde{\beta}_{00}\rangle &= \sin\theta''|00\rangle + \cos\theta''|11\rangle,\\
|\widetilde{\beta}_{01}\rangle &= \sin\theta''|01\rangle + \cos\theta''|10\rangle,\\
|\widetilde{\beta}_{10}\rangle &= \cos\theta''|00\rangle - \sin\theta''|11\rangle,\\
|\widetilde{\beta}_{11}\rangle &= \cos\theta''|01\rangle - \sin\theta''|10\rangle,
\end{aligned} \tag{8.68}$$

form the generalized Bell basis.

The entanglement entropy [25] S_E, which is a standard measure of the entanglement of a bipartite pure state, is the same for each state (8.66) since local operations do not change the amount of entanglement. The entanglement entropy is computed from the von Neumann entropy of one of the subsystems $\rho_B^{(i\ell)} = \mathrm{Tr}_C(|\phi_{i\ell}\rangle\langle\phi_{i\ell}|)$, and turns out to be independent of the angle θ',

$$\begin{aligned}
S_E &= S\big(\rho_B^{(i\ell)}\big)\\
&= -\cos^2\theta''\log\cos^2\theta'' - \sin^2\theta''\log\sin^2\theta''.
\end{aligned} \tag{8.69}$$

Evidently, the conditional states (8.66) are in a product state (uncorrelated) when $\theta'' = 0$ and are fully entangled when $\theta'' = \pi/4$ (this second case is the one considered in the next

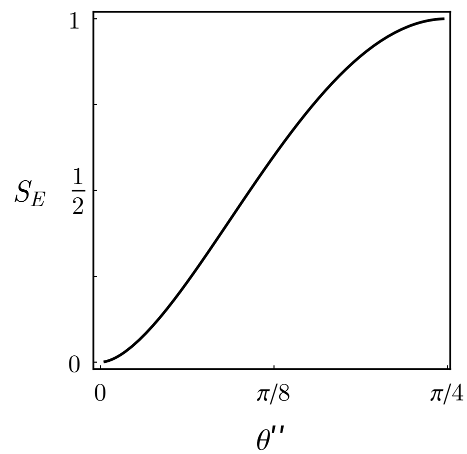

Figure 8.7: Entanglement entropy S_E of the conditional states (8.66) of qubits B and C for $\theta = \pi/4$. In this case, S_E is independent of the intermediate angle θ'. The states are maximally entangled at $\theta'' = \pi/4$.

section). Thus, Bob can control how entangled the shared state of BC is by choosing a particular measurement angle θ'' (see Fig. 8.7).

The operation that disentangles the qubits B and C from A and D takes the form

$$V = \sum_{i\ell} |i\rangle\langle i| \otimes V^{(i\ell)} \otimes |\ell\rangle\langle\ell|, \tag{8.70}$$

where the unitary operators $V^{(i\ell)}$ were defined in (8.65). This does not completely factor into two separate operations on the pairs of qubits AB and CD. In other words, classical communication is necessary between Alice and Bob in order to implement the disentangling operation. From (8.65), it is clear that Alice must know the state ℓ of Bob's qubit D before performing her controlled unitary on A and B, while Bob does not need to know the state i of Alice's qubit A. Despite the communication requirement, the conditional operations (8.65) are independent of the angles θ' and θ'' so that Alice and Bob do not need to know each other's measurement bases.

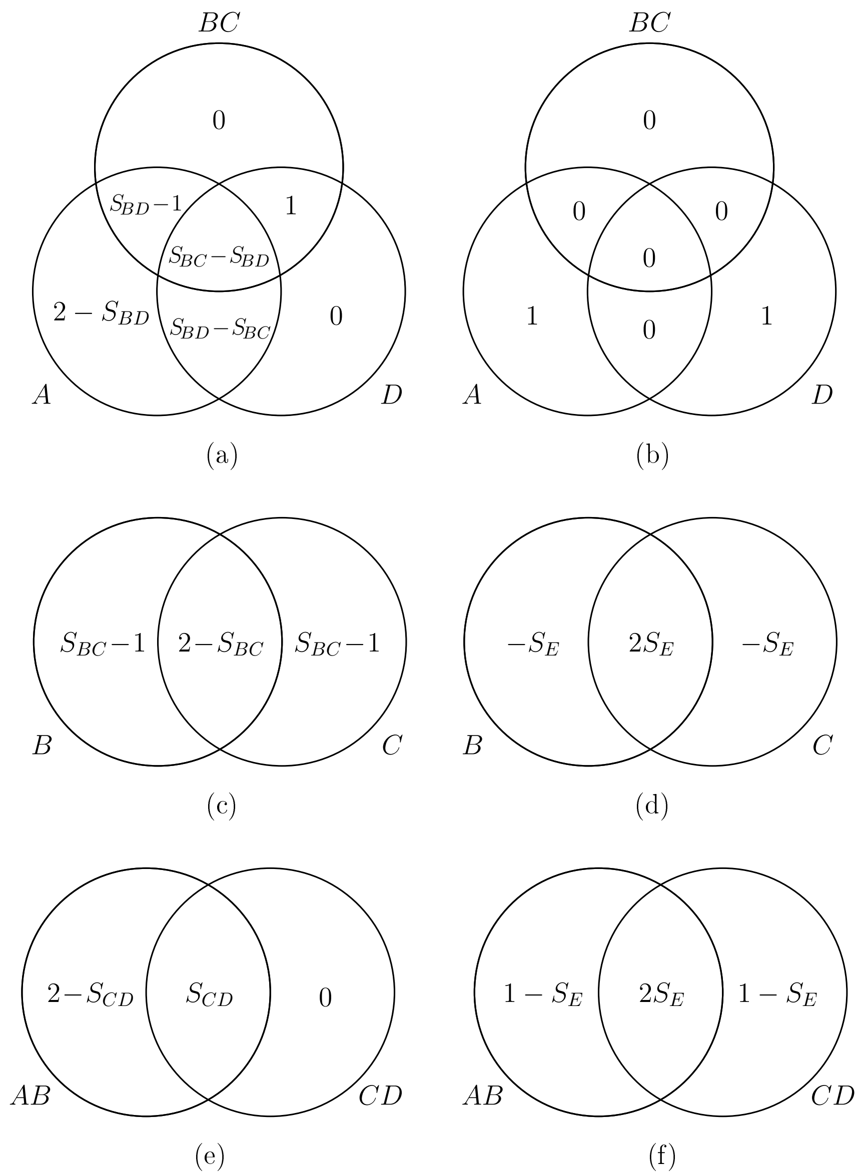

Figure 8.8: Entropy Venn diagrams for the ancillary qubits before (a,c,e) and after (b,d,f) disentangling with (8.65). The first angle is $\theta = \pi/4$, while θ' and θ'' are left arbitrary. The entropy S_{BD} (S_{BC}) [S_{CD}] depends on θ' and θ'' (θ') [θ''] and the entanglement entropy S_E is defined in (8.69).

After applying (8.70) to the state (8.64),

$$V\rho_{ABCD}V^\dagger = \frac{1}{2}\mathbb{1} \otimes |\phi_{00}\rangle\langle\phi_{00}| \otimes \frac{1}{2}\mathbb{1}, \tag{8.71}$$

qubits B and C are successfully disentangled from the rest of the system, leaving qubits A

and D in maximally mixed states.

The entropy relationships between qubits A, B, C, and D are illustrated using entropy Venn diagrams [59] in Fig. 8.8 before and after the disentangling operation (8.70). Note that the shared entropy $S(A:D)$ is always zero, and that Alice and Bob will share a maximally entangled state when Bob chooses the angle $\theta'' = \pi/4$ so that $S_E = 1$.

8.7.3 Discussion

8.7.3.1 Conditions for Entanglement Generation

In the most general scenario, for local operations to successfully disentangle the joint state of qubits B and C from the rest of the system, it is necessary for the entanglement entropy of (8.62) to be a constant for all i, ℓ. In other words, disentanglement is only possible if the Schmidt coefficients in the Schmidt decomposition of each state (8.62) are the same. This, however, does not guarantee that the resulting state of BC will be *entangled*. For instance, at $\theta = 0$ or $\theta'' = 0$, the entanglement entropies are indeed the same, but vanish, so that the resulting state of BC is completely uncorrelated and no shared entanglement is created.

When nonzero entanglement between qubits B and C is successfully generated, the shared entropy of A and D vanishes, $S(A:D) = 0$. In all other situations, $S(A:D) > 0$. Given the correlated structure of the coefficients $\frac{1}{2}\, p_{i\ell}$ [see (8.63)] in the density matrix ρ_{AD}, a vanishing mutual entropy can only occur if at least one of the three angles is $\pi/4$ so that $\frac{1}{2}\, p_{i\ell} = 1/4$. In turn, this corresponds precisely to a constant and *nonzero* entanglement entropy. Thus, the necessary and sufficient condition for the entanglement generation scheme described here is simply

$$S(A:D) = 0. \tag{8.72}$$

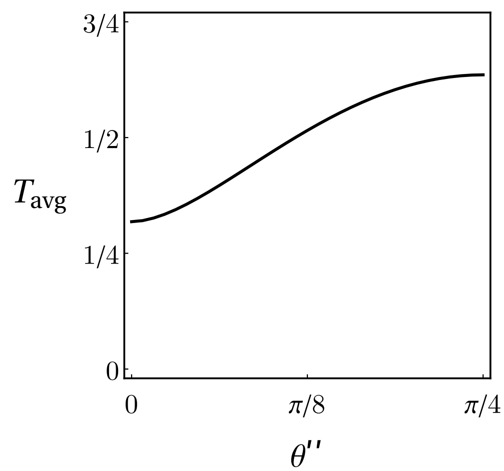

Figure 8.9: The trace distance, T, averaged over all intermediate measurement angles θ', as a function of θ''. Here, $\theta = \pi/4$.

8.7.3.2 Reliability

We can characterize how coherent the joint state of all four ancillary qubits is before the disentangling operation (8.70) is applied. The closer the state is to its classical counterpart, the more robust it is to decoherence during the phase of the protocol when the qubit pairs are sent to Alice and Bob. We quantify how close ρ_{ABCD} is to the classical (completely decoherent) version, σ_{ABCD}, which has only diagonal elements, by computing the trace distance between (8.61) and the decoherent state

$$\sigma_{ABCD} = \frac{1}{d} \sum_{ijk\ell} |U_{ij}|^2\, |U'_{jk}|^2\, |U''_{k\ell}|^2\, |ijk\ell\rangle\langle ijk\ell|. \tag{8.73}$$

The trace distance between two states ρ and σ is defined as [25] $T = \frac{1}{2}\mathrm{Tr}(|\rho - \sigma|) = \frac{1}{2}\mathrm{Tr}(\sqrt{(\rho - \sigma)^2})$, which is bounded between 0 and 1. The trace distance, averaged over all intermediate angles θ', is shown in Fig. 8.9 for $\theta = \pi/4$. This distance measure increases (the state becomes more coherent) as Bob's angle θ'' increases, and the resulting state of BC becomes more entangled.

8.7.4 Protocol without Communication

In the first protocol, the ancillary qubits were encoded using $\theta = \pi/4$, while θ' and θ'' were left arbitrary. Here, I consider at a special case where both the first and third relative angles are set at $\theta = \theta'' = \pi/4$. We will see that this removes entirely the communication requirement between Alice and Bob because the disentangling operation completely factorizes.

For this set of angles, the density matrix (8.64) can be written with

$$\widetilde{V}^{(i\ell)\dagger} = Z^i \otimes (-Z)^\ell, \tag{8.74}$$

which are conditional unitary operators on B and C. There are two important features of (8.74): the operators on B and C are completely factorized, and they do not depend on the intermediate angle θ'. That is, the operator Z^i, with only the index i, acts on B, and $(-Z)^\ell$, with only the index ℓ, acts on C. Thus, communication between Alice and Bob is no longer required to generate remote entanglement since they do not have to know the state of the other's qubit.

The four conditional states of BC can be obtained by applying (8.74) to the $i = \ell = 0$ state,

$$|\phi_{00}\rangle = -\sin\theta' \, |\beta_{01}\rangle + \cos\theta' \, |\beta_{10}\rangle, \tag{8.75}$$

where the standard Bell basis can be written as

$$|\beta_{zx}\rangle = (\mathbb{1} \otimes X^x Z^z) \, |\beta_{00}\rangle, \tag{8.76}$$

with $|\beta_{00}\rangle = |\Phi^+\rangle$ the usual Bell state. It is clear from (8.74) that if Alice and Bob each

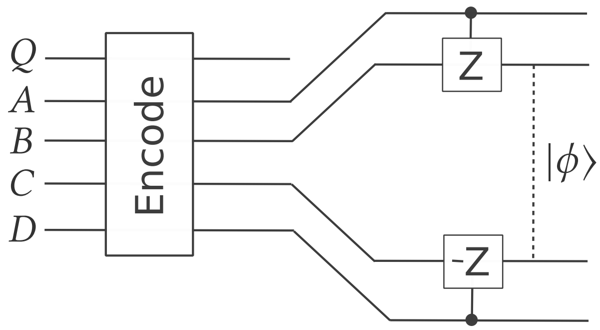

Figure 8.10: Circuit schematic for generating shared entanglement. Alice holds qubits A and B, while Bob holds qubits C and D. The state of qubits $ABCD$ after the encoding scheme is found from (8.59). In this figure, the entangled state $|\phi\rangle$ that is generated corresponds to $|\phi_{00}\rangle$ in (8.75).

perform a controlled unitary on their set of qubits of the form,

$$\begin{aligned}\widetilde{V} &= \widetilde{V}_{AB} \otimes \widetilde{V}_{CD} \\ &= \left[\sum_i |i\rangle\langle i| \otimes Z^i\right] \otimes \left[\sum_\ell (-Z)^\ell \otimes |\ell\rangle\langle\ell|\right],\end{aligned} \tag{8.77}$$

the resulting state will be (8.71), where $|\phi_{00}\rangle$ is now given by (8.75). Note that according to (8.77), Alice and Bob each perform a controlled-phase gate on their pairs of qubits.

Thus, by using only local operations on their pairs of qubits, Alice and Bob disentangle the joint state of BC from the rest of the system. I emphasize that this does not require any classical communication between Alice and Bob, and afterwards they share one half each of the entangled state (8.75). This state (8.75) is maximally entangled regardless of the angle θ' (its entanglement entropy is equal to one). Since the conditional unitary operators (8.77) on B and C are independent of the angle θ', the disentangling can occur when Alice and Bob do not know θ'. In other words, they do not need to know each other's measurement bases.

Figure 8.10 shows a simple schematic of the complete entanglement generation process without communication. The operation denoted "encode" is the sequence of measurements

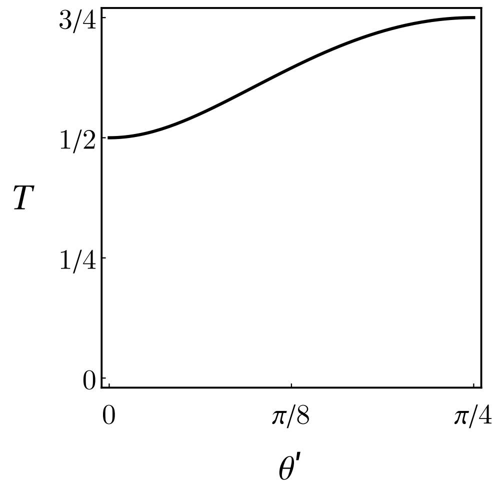

Figure 8.11: The trace distance, T, as a function of the intermediate angle θ'. Here, both Alice's and Bob's relative measurement angle is $\theta = \theta'' = \pi/4$.

that produces the state (8.59). In the figure, the final state $|\phi\rangle$ corresponds to (8.75). The trace distance is plotted as a function of the intermediate angle, θ', in Fig. 8.11. The entanglement between subsystems A, BC, and D is shown in Fig. 8.12 before and after the disentangling operation (8.77).

8.8 Conclusions

Consecutive measurements on a single quantum have a wide range of applications, from extracting joint weak values in weak measurements [230] and modeling the quantum Zeno effect [192, 206, 209–211] to probing the nature of quantum measurement [177–179]. Here, I showed that the consecutive measurement construction can be used to generate remote entanglement. The protocol uses a simple encoding scheme and requires only local operations on the resulting entangled qubits. In the general case, the degree of entanglement can be tuned by the choice of encoding parameters (the set of measured observables), while the second version always produces a maximally entangled state.

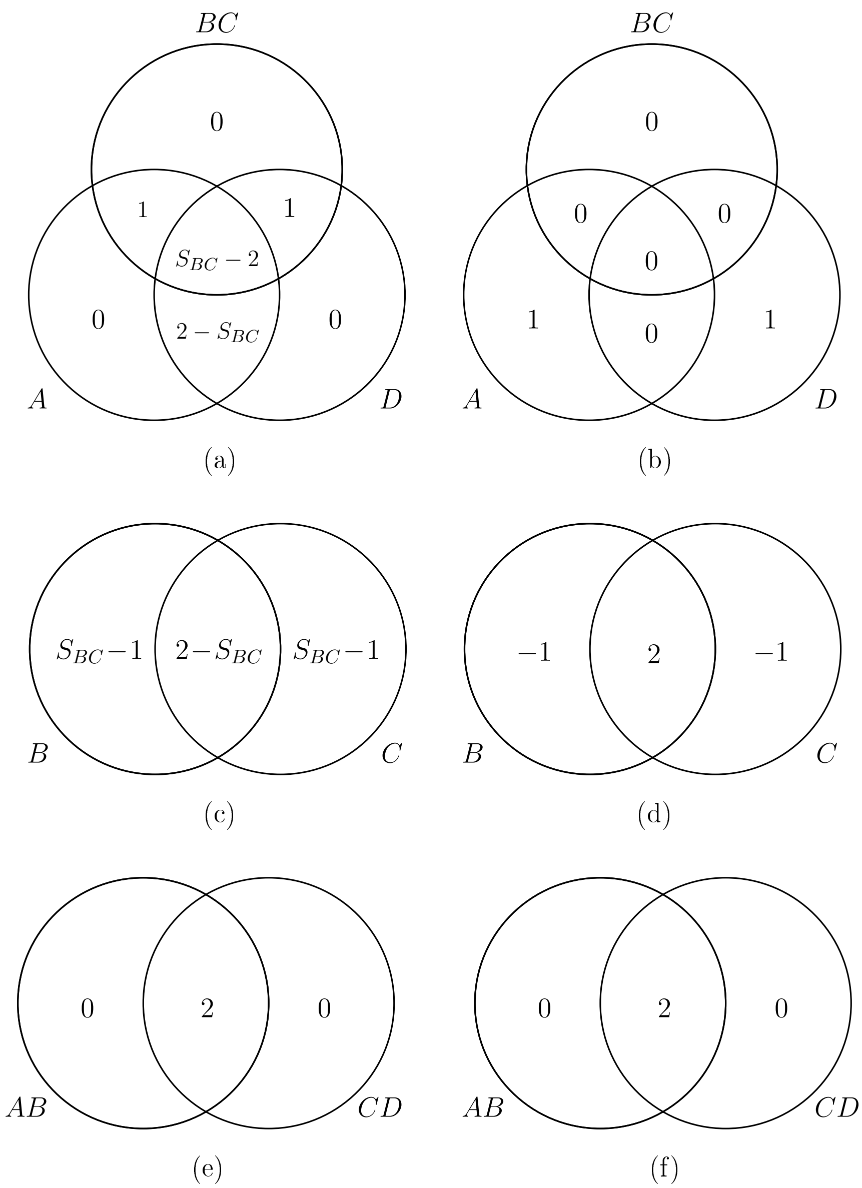

Figure 8.12: Entropy Venn diagrams for $\theta = \theta'' = \pi/4$ before (a,c,e) and after (b,d,f) the disentangling operation (8.77). The intermediate angle θ' is left arbitrary.

Chapter 9

Conclusions

9.1 Summary of Results and Future Work

The primary focus of this dissertation was the formulation of a unitary model of quantum measurement that is treated on the same footing as quantum operations and cast in the framework of quantum information theory. In this manner, I was able to quantify how information is processed in quantum measurements. I began with the early model of measurement that was first introduced in 1932 by von Neumann where a quantum measuring device is entangled with a quantum system. From there, I built on the ideas of Adami and Cerf who, starting in the '90s, described measurement in the context of quantum information theory and introduced the useful quantum entropy Venn diagrams. In this framework, they showed the striking departures of quantum entropy from its classical counterpart, in particular, the emergence of negative conditional quantum entropy from entangled systems, which stands in stark contrast to the non-negativity constraints of classical conditional entropy.

In this work, I expanded upon these ideas and quantified how entanglement and entropy are distributed and evolve in diverse measurement scenarios. I considered parallel and consecutive measurements as well as measurements with variable interaction strengths. Along the way, I discussed how strong measurements are not fundamentally different from weak measurements (Ch. 5), derived a new quantum information-theoretic complementarity relation for the quantum eraser (Ch. 6), explored the Markovianity of sequences of measurements (Ch. 7), and developed protocols based on disentangling operations (Ch. 8).

The main results of this dissertation are organized into four sections. First, in Ch. 5 I described a generalization of weak and strong measurements. Then, in Ch. 6 I described parallel measurements that are made on an a composite quantum system, while in Ch. 7 I analyzed consecutive measurements made on the same quantum system. Finally, I used the findings of Ch. 7 to construct measurement-based quantum protocols, which were the subject of Ch. 8.

In Ch. 5, I argued that within our model of measurement, weak and strong measurements should be treated on the same footing. Furthermore, I showed that certain measurement interactions do not have to be approximated by a weak coupling and can be computed exactly. I focused on the case of spin and position measurements with a qubit pointer and derived the real and imaginary components of the weak value. I only considered qubit systems, but the formalism could readily be extended to include higher-dimensional quantum systems and pointers.

In Ch. 6, I described parallel measurements of an entangled quantum system in the context of the Bell-state quantum eraser experiment. The results of the quantum eraser in all of its various forms have been well established for quite some time so my aim was not to make new experimental predictions, but instead to analyze the (possibly delayed) manipulation of the interference patterns in the eraser mechanism in the light of quantum information theory. I showed how the experiment can be broken down into distinct stages that can be understood using entropy Venn diagrams, which elucidate not only the flow of entropy and entanglement between subsystems, but also how these quantities are ultimately connected to the structure of the photon's interference pattern. To tie all the results together, I derived a complementarity relation between the coherence and "which-path" information of the photon traveling through the eraser apparatus, which is fundamentally a consequence of the chain rule for entropies. Unlike many previous efforts in this area, our approach did not require definitions for coherence and path information to be specified beforehand. In this way, an information-theoretic approach offers additional insights into the origins of complementarity.

In Ch. 7, I turned from parallel measurements to an arbitrarily-long sequence of consecutive measurements of a single quantum system. I utilized the notion of unprepared quantum states, which in my view are more basic than prepared states. Such states can be thought of as arising from a (possibly infinite) sequence of measurements that randomize the quantum system, leaving it in a maximum entropy state, quite similarly to the quantum anti-Zeno effect. I studied the differences that arise from measuring prepared and unprepared quantum states in terms of the structure of the resulting density matrices and entropy Venn diagrams. I described two types of quantum measurements (unamplified and amplified) and quantified how information is "lost" if the measurements are amplified. I distinguished measurements that remain coherent from those that are amplified, and proved that amplified measurement sequences are equivalent to quantum Markov chains. In contrast, I showed that the state of unamplified ancillae remains coherent regardless of the size of the measurement chain. The unique structure of these joint ancilla states leads to the interesting property that their entropy is encoded in the boundary of the measurement chain. Using the tools of quantum information theory, I quantified the amount of information a measurement device has about the quantum system and about previous devices in the measurement chain, and showed that this information is always reduced by amplification.

A natural next step may be to extend and generalize the results in this work to tensor networks. I already alluded to a possible connection to the holographic principle from the fact that the entropy of measurement chains are described by the boundary only. The simple one-dimensional description presented here could be extended to multidimensional measurement networks where the entropy of the entire network may be found on the surface of the volume. In this case, we might ask what sequences of measurements would satisfy this holographic property and what would be the structures of the underlying density matrices.

Another possible avenue of future investigation is to study in more detail quantum Markov chains and the structure of states that satisfy the Markov property. I showed that amplified measurement chains are always Markovian, but we could investigate how non-Markovian are unamplified chains. That is, how the structure of unamplified chains compares to a quantum Markov chain.

Finally, in Ch. 8 I harnessed the coherent structure of the joint ancilla density matrices to construct disentangling schemes. I first considered general one- and two-bit disentangling protocols in order to demonstrate the basic construction. Then, I outlined a protocol for generating remote entanglement, which is a necessary component of quantum information processing. I discussed a method for encoding a set of ancillary qubits via consecutive measurements and applying conditional operators to extract an entangled state shared between two spatially separated locations. Future work in this area could include investigating methods for certifying the entanglement generation. That is, whether Alice and Bob can verify that the entangled state is indeed produced. This is important for quantum key distribution, where we need to determine whether a third party (Eve) has tampered with any of the qubits. We could also probe further the structure of the ancilla chains and the conditions under which disentangling operations are allowed.

REFERENCES

REFERENCES

[1] W. H. Zurek. *Complexity, Entropy, and the Physics of Information: The Proceedings of the 1988 Workshop on Complexity, Entropy, and the Physics of Information Held May–June, 1989, in Santa Fe, New Mexico.* Proceedings volume in the Santa Fe Institute studies in the sciences of complexity. Westview Press, 1990.

[2] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27(3):379–423, 1948.

[3] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27(4):623–656, 1948.

[4] J. Preskill. Quantum computing and the entanglement frontier. *arXiv:1203.5813*, 2012.

[5] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, J. M Martinis, and H. Neven. Characterizing quantum supremacy in near-term devices. *arXiv:1608.00263*, 2016.

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

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

[8] P. W. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52:R2493–R2496, 1995.

[9] A. M. Steane. Error correcting codes in quantum theory. *Phys. Rev. Lett.*, 77:793–797, 1996.

[10] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54:1098–1105, 1996.

[11] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek. Perfect quantum error correcting code. *Phys. Rev. Lett.*, 77:198–201, 1996.

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

[13] A. M. Steane. Simple quantum error-correcting codes. *Phys. Rev. A*, 54:4741–4751, 1996.

[14] D. Gottesman. Class of quantum error-correcting codes saturating the quantum hamming bound. *Phys. Rev. A*, 54:1862–1868, 1996.

[15] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien. Quantum computers. *Nature*, 464(7285):45–53, 2010.

[16] D. P. DiVincenzo. The physical implementation of quantum computation. *Fortschr. Phys.*, 48(9-11):771–783, 2000.

[17] H. Everett. "Relative State" Formulation of Quantum Mechanics. *Rev. Mod. Phys.*, 29:454–462, 1957.

[18] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz. Time Symmetry in the Quantum Process of Measurement. *Phys. Rev.*, 134:B1410–B1416, 1964.

[19] R. B. Griffiths. Consistent histories and the interpretation of quantum mechanics. *J. Stat. Phys.*, 36(1-2):219–272, 1984.

[20] J. P. Paz and W. H. Zurek. Environment-induced decoherence, classicality, and consistency of quantum histories. *Phys. Rev. D*, 48:2728–2738, 1993.

[21] W. H. Zurek. Preferred states, predictability, classicality and the environment-induced decoherence. *Prog. Theor. Phys.*, 89:281–312, 1993.

[22] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Julius Springer, Berlin, 1932.

[23] Y. Aharonov, D. Z. Albert, and L. Vaidman. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.*, 60:1351–1354, 1988.

[24] N. W. M. Ritchie, J. G. Story, and R. G. Hulet. Realization of a measurement of a "weak value". *Phys. Rev. Lett.*, 66:1107–1110, 1991.

[25] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition.* Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[26] P. W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 124–134. IEEE, 1994.

[27] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.

[28] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th annual ACM symposium on Theory of computing*, pages 212–219. ACM, 1996.

[29] L. K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.*, 79:325–328, 1997.

[30] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.*, 47:777–780, 1935.

[31] D. M. Greenberger, M. A. Horne, and A. Zeilinger. *Going Beyond Bell's Theorem*, pages 69–72. Springer Netherlands, Dordrecht, 1989.

[32] M. Hillery, V. Bužek, and A. Berthiaume. Quantum secret sharing. *Phys. Rev. A*, 59:1829–1834, 1999.

[33] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62:062314, 2000.

[34] T. J. Herzog, P. G. Kwiat, H. Weinfurter, and A. Zeilinger. Complementarity and the Quantum Eraser. *Phys. Rev. Lett.*, 75:3034–3037, 1995.

[35] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, 52:3457–3467, 1995.

[36] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, 1982.

[37] D. Dieks. Communication by EPR devices. *Phys. Lett. A*, 92(6):271–272, 1982.

[38] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher. Noncommuting mixed states cannot be broadcast. *Phys. Rev. Lett.*, 76:2818–2821, 1996.

[39] V. Bužek and M. Hillery. Quantum copying: Beyond the no-cloning theorem. *Phys. Rev. A*, 54:1844–1852, 1996.

[40] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin. Optimal universal and state-dependent quantum cloning. *Phys. Rev. A*, 57:2368–2378, 1998.

[41] D. A. Lidar and T. A. Brun. *Quantum Error Correction*. Cambridge University Press, 2013.

[42] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger. Experimental quantum teleportation. *Nature*, 390(6660):575–579, 1997.

[43] D. Boschi, S. Branca, F. De Martini, L. Hardy, and S. Popescu. Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 80:1121–1125, 1998.

[44] A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik. Unconditional quantum teleportation. *Science*, 282(5389):706–709, 1998.

[45] M. A. Nielsen, E. Knill, and R. Laflamme. Complete quantum teleportation using nuclear magnetic resonance. *Nature*, 396(6706):52–55, 1998.

[46] M. Riebe, H. Häffner, C. F. Roos, W. Hänsel, J. Benhelm, G. P. T. Lancaster, T. W. Körber, C. Becher, F. Schmidt-Kaler, D. F. V. James, and R. Blatt. Deterministic quantum teleportation with atoms. *Nature*, 429(6993):734, 2004.

[47] M. D. Barrett, J. Chiaverini, T. Schaetz, J. Britton, W. M. Itano, J. D. Jost, E. Knill, C. Langer, D. Leibfried, R. Ozeri, and D. J. Wineland. Deterministic quantum teleportation of atomic qubits. *Nature*, 429(6993):737, 2004.

[48] L. Steffen, Y. Salathe, M. Oppliger, P. Kurpiers, M. Baur, C. Lang, C. Eichler, G. Puebla-Hellmann, A. Fedorov, and A. Wallraff. Deterministic quantum teleportation with feed-forward in a solid state system. *Nature*, 500(7462):319, 2013.

[49] W. Pfaff, B. J. Hensen, H. Bernien, S. B. van Dam, M. S. Blok, T. H. Taminiau, M. J. Tiggelman, R. N. Schouten, M. Markham, D. J. Twitchen, and R. Hanson. Unconditional quantum teleportation between distant solid-state quantum bits. *Science*, 345(6196):532–535, 2014.

[50] J.-G. Ren et al. Ground-to-satellite quantum teleportation. *arXiv:1707.00934*, 2017.

[51] J. Yin et al. Satellite-based entanglement distribution over 1200 kilometers. *Science*, 356(6343):1140–1144, 2017.

[52] G. Brassard, S. L. Braunstein, and R. Cleve. Teleportation as a quantum computation. *Physica D: Nonlinear Phenomena*, 120(1):43–47, 1998.

[53] S. L. Braunstein. Quantum teleportation without irreversible detection. *Phys. Rev. A*, 53:1900–1902, 1996.

[54] K. Mattle, H. Weinfurter, P. G. Kwiat, and A. Zeilinger. Dense coding in experimental quantum communication. *Phys. Rev. Lett.*, 76:4656–4659, 1996.

[55] B. P. Williams, R. J. Sadlier, and T. S. Humble. Superdense coding over optical fiber links with complete Bell-state measurements. *Phys. Rev. Lett.*, 118:050501, 2017.

[56] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley, 2012.

[57] J. von Neumann. Thermodynamik quantenmechanischer gesamtheiten. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1927:273–291, 1927.

[58] N. J. Cerf and C. Adami. Negative entropy and information in quantum mechanics. *Phys. Rev. Lett.*, 79:5194–5197, 1997.

[59] N. J. Cerf and C. Adami. Information theory of quantum entanglement and measurement. *Physica D: Nonlinear Phenomena*, 120(1–2):62–81, 1998. Proceedings of the Fourth Workshop on Physics and Consumption.

[60] N. J. Cerf and C. Adami. Quantum extension of conditional probability. *Phys. Rev. A*, 60:893–897, 1999.

[61] M. S. Leifer and R. W. Spekkens. Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference. *Phys. Rev. A*, 88:052130, 2013.

[62] N. J. Cerf and C. Adami. Entropic Bell inequalities. *Phys. Rev. A*, 55:3371–3374, 1997.

[63] M. Horodecki, J. Oppenheim, and A. Winter. Partial quantum information. *Nature*, 436(7051):673–676, 2005.

[64] N. J. Cerf and R. Cleve. Information-theoretic interpretation of quantum error-correcting codes. *Phys. Rev. A*, 56:1721–1732, Sep 1997.

[65] H. Araki and E. H. Lieb. Entropy inequalities. *Comm. Math. Phys.*, 18(2):160–170, 1970.

[66] E. H. Lieb and M. B. Ruskai. A Fundamental Property of Quantum-Mechanical Entropy. *Phys. Rev. Lett.*, 30:434–436, 1973.

[67] E. H. Lieb and M. B. Ruskai. Proof of the strong subadditivity of quantum-mechanical entropy. *J. Math. Phys.*, 14(12):1938–1941, 1973.

[68] N. Datta and M. M. Wilde. Quantum Markov chains, sufficiency of quantum channels, and Rényi information measures. *J. Phys. A: Math. Theor.*, 48(50):505301, 2015.

[69] P. Hayden, R. Jozsa, D. Petz, and A. Winter. Structure of States Which Satisfy Strong Subadditivity of Quantum Entropy with Equality. *Comm. Math. Phys.*, 246(2):359–374, 2004.

[70] J. R Glick and C. Adami. Markovian and non-Markovian quantum measurements. *arXiv:1701.05636*, 2017.

[71] C. Adami and N. J. Cerf. von Neumann capacity of noisy quantum channels. *Phys. Rev. A*, 56:3470–3483, 1997.

[72] C. Adami and N. J. Cerf. Prolegomena to a Non-equilibrium Quantum Statistical Mechanics. *Chaos, Solitons & Fractals*, 10(10):1637–1650, 1999.

[73] V. Coffman, J. Kundu, and W. K. Wootters. Distributed entanglement. *Phys. Rev. A*, 61:052306, 2000.

[74] A. K. Pati. Existence of the Schmidt decomposition for tripartite systems. *Phys. Lett. A*, 278(3):118–122, 2000.

[75] A. Peres. Higher order Schmidt decompositions. *Phys. Lett. A*, 202(1):16–17, 1995.

[76] L. Henderson and V. Vedral. Classical, quantum and total correlations. *J. Phys. A: Math. and Gen.*, 34(35):6899, 2001.

[77] N. J. Cerf and C. Adami. Accessible information in quantum measurement. *arxiv:quant-ph/9611032*, 1996.

[78] M. A. Neumark. Spectral functions of a symmetric operator. *Izv. Akad. Nauk SSSR Ser. Mat.*, 4(3):277–318, 1940.

[79] Y. Aharonov, J. Anandan, and L. Vaidman. Meaning of the wave function. *Phys. Rev. A*, 47:4616–4626, 1993.

[80] Y. Aharonov and L. Vaidman. Measurement of the Schrödinger wave of a single particle. *Phys. Lett. A*, 178(1-2):38–42, 1993.

[81] N. D. Hari Dass and T. Qureshi. Critique of protective measurements. *Phys. Rev. A*, 59:2590–2601, 1999.

[82] W. Gerlach and O. Stern. Der experimentelle nachweis der richtungsquantelung im magnetfeld. *Zeitschrift für Physik A Hadrons and Nuclei*, 9(1):349–352, 1922.

[83] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, 2001.

[84] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn. Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.*, 79:135–174, 2007.

[85] J. L. O'Brien. Optical quantum computing. *Science*, 318(5856):1567–1570, 2007.

[86] J. I. Cirac and P. Zoller. Quantum computations with cold trapped ions. *Phys. Rev. Lett.*, 74:4091–4094, 1995.

[87] D. Kielpinski, C. Monroe, and D. J. Wineland. Architecture for a large-scale ion-trap quantum computer. *Nature*, 417(6890):709–711, 2002.

[88] H. Häffner, C. F. Roos, and R. Blatt. Quantum computing with trapped ions. *Phys. Rep.*, 469(4):155–203, 2008.

[89] D. Loss and D. P. DiVincenzo. Quantum computation with quantum dots. *Phys. Rev. A*, 57:120–126, 1998.

[90] B. D. Josephson. Possible new effects in superconductive tunnelling. *Phys. lett.*, 1(7):251–253, 1962.

[91] M. H. Devoret, A. Wallraff, and J. M Martinis. Superconducting qubits: A short review. *arxiv:cond-mat/0411174*, 2004.

[92] J. Clarke and F. K. Wilhelm. Superconducting quantum bits. *Nature*, 453(7198):1031–1042, 2008.

[93] V. Bouchiat, D. Vion, P. Joyez, D. Esteve, and M. H. Devoret. Quantum coherence with a single Cooper pair. *Physica Scripta*, 1998(T76):165, 1998.

[94] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature*, 398(6730):786–788, 1999.

[95] K. W. Lehnert, K. Bladh, L. F. Spietz, D. Gunnarsson, D. I. Schuster, P. Delsing, and R. J. Schoelkopf. Measurement of the excited-state lifetime of a microelectronic circuit. *Phys. Rev. Lett.*, 90:027002, 2003.

[96] D. A. Rodrigues. *Superconducting Charge Qubits*. PhD thesis, University of Bristol, 2003.

[97] J. Q. You and F. Nori. Superconducting circuits and quantum information. *Physics Today*, 58(11):42–47, 2005.

[98] R. W. Simmonds, K. M. Lang, D. A. Hite, S. Nam, D. P. Pappas, and J. M. Martinis. Decoherence in Josephson phase qubits from junction resonators. *Phys. Rev. Lett.*, 93:077003, 2004.

[99] K. B. Cooper, M. Steffen, R. McDermott, R. W. Simmonds, S. Oh, D. A. Hite, D. P. Pappas, and J. M. Martinis. Observation of quantum oscillations between a Josephson phase qubit and a microscopic resonator using fast readout. *Phys. Rev. Lett.*, 93:180401, 2004.

[100] I. Chiorescu, Y. Nakamura, C. J. P. M. Harmans, and J. E. Mooij. Coherent quantum dynamics of a superconducting flux qubit. *Science*, 299(5614):1869–1871, 2003.

[101] I. Chiorescu, P. Bertet, K. Semba, Y. Nakamura, C. J. P. M. Harmans, and J. E. Mooij. Coherent dynamics of a flux qubit coupled to a harmonic oscillator. *Nature*, 431(7005):159–162, 2004.

[102] I. L. Chuang and Y. Yamamoto. Simple quantum computer. *Phys. Rev. A*, 52:3489–3496, 1995.

[103] N. J. Cerf, C. Adami, and P. G. Kwiat. Optical simulation of quantum logic. *Phys. Rev. A*, 57:R1477–R1480, 1998.

[104] A. P. Lund and T. C. Ralph. Nondeterministic gates for photonic single-rail quantum logic. *Phys. Rev. A*, 66:032307, 2002.

[105] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.*, 73:58–61, 1994.

[106] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber. Direct measurement of the quantum wavefunction. *Nature*, 474(7350):188–91, 2011.

[107] C. Adami and N. J. Cerf. *Quantum Computation with Linear Optics*, pages 391–401. Springer Berlin Heidelberg, 1999.

[108] J. Kerr. XL. a new relation between electricity and light: Dielectrified media birefringent. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(332):337–348, 1875.

[109] M. Koashi, T. Yamamoto, and N. Imoto. Probabilistic manipulation of entangled photons. *Phys. Rev. A*, 63:030301, 2001.

[110] J. H. Shapiro. Single-photon Kerr nonlinearities do not help quantum computation. *Phys. Rev. A*, 73:062305, 2006.

[111] R. C. Jones. A New Calculus for the Treatment of Optical Systems, I. Description and Discussion of the Calculus. *J. Opt. Soc. Am.*, 31(7):488–493, Jul 1941.

[112] E. Collett. *Field Guide to Polarization*. SPIE Press, Bellingham, WA, 2005.

[113] P. Kok and B. W. Lovett. *Introduction to optical quantum information processing*. Cambridge University Press, 2010.

[114] M. O. Scully and W. E. Lamb. Quantum Theory of an Optical Maser. III. Theory of Photoelectron Counting Statistics. *Phys. Rev.*, 179:368–374, 1969.

[115] H. Lee, U. Yurtsever, P. Kok, G. M. Hockney, C. Adami, S. L. Braunstein, and J. P. Dowling. Towards photostatistics from photon-number discriminating detectors. *J. Mod. Opt.*, 51(9-10):1517–1528, 2004.

[116] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd. Colloquium: Understanding quantum weak values: Basics and applications. *Rev. Mod. Phys.*, 86:307–316, 2014.

[117] G. Vallone and D. Dequal. Strong measurements give a better direct measurement of the quantum wave function. *Phys. Rev. Lett.*, 116:040502, 2016.

[118] O. Hosten and P. Kwiat. Observation of the spin hall effect of light via weak measurements. *Science*, 319(5864):787–790, 2008.

[119] J. S. Lundeen and C. Bamber. Procedure for direct measurement of general quantum states using weak measurement. *Phys. Rev. Lett.*, 108(7):070402, 2012.

[120] S. Wu. State tomography via weak measurements. *Sci. Rep.*, 3:1193, 2013.

[121] A. N. Korotkov. *Noisy Quantum Measurement of Solid-State Qubits: Bayesian Approach*, pages 205–228. Springer Netherlands, Dordrecht, 2003.

[122] K. Jacobs and D. A. Steck. A straightforward introduction to continuous quantum measurement. *Contemp. Phys.*, 47(5):279–303, 2006.

[123] T. A. Brun. A simple model of quantum trajectories. *Am. J. Phys.*, 70(7):719–737, 2002.

[124] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan. The sense in which a "weak measurement" of a spin-1/2 particle's spin component yields a value 100. *Phys. Rev. D*, 40:2112–2117, 1989.

[125] Y. Aharonov and L. Vaidman. Properties of a quantum system during the time interval between two measurements. *Phys. Rev. A*, 41:11–20, 1990.

[126] J. S. Lundeen and K. J. Resch. Practical measurement of joint weak values and their connection to the annihilation operator. *Phys. Lett. A*, 334(5):337–344, 2005.

[127] R. Jozsa. Complex weak values in quantum measurement. *Phys. Rev. A*, 76:044103, 2007.

[128] Y. Aharonov and L. Vaidman. Complete description of a quantum system at a given time. *J. Phys. A: Math. Gen.*, 24(10):2315, 1991.

[129] R. Ravon and L. Vaidman. The three-box paradox revisited. *J. Phys. A: Math. Theor.*, 40(11):2873, 2007.

[130] K. J. Resch, J. S. Lundeen, and A. M. Steinberg. Experimental realization of the quantum box problem. *Phys. Lett. A*, 324(2–3):125–131, 2004.

[131] J. S. Lundeen and A. M. Steinberg. Experimental Joint Weak Measurement on a Photon Pair as a Probe of Hardy's Paradox. *Phys. Rev. Lett.*, 102:020404, 2009.

[132] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell. Ultrasensitive beam deflection measurement via interferometric weak value amplification. *Phys. Rev. Lett.*, 102:173601, 2009.

[133] A. Feizpour, X. Xing, and A. M. Steinberg. Amplifying single-photon nonlinearity using weak measurements. *Phys. Rev. Lett.*, 107:133603, 2011.

[134] J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat. Photonic state tomography. *Adv. At. Mol. Opt. Phys.*, 52:105–159, 2005.

[135] A. Romito, Y. Gefen, and Y. M. Blanter. Weak values of electron spin in a double quantum dot. *Phys. Rev. Lett.*, 100:056801, 2008.

[136] N. S. Williams and A. N. Jordan. Weak values and the Leggett-Garg inequality in solid-state qubits. *Phys. Rev. Lett.*, 100:026804, 2008.

[137] J. P. Groen, D. Ristè, L. Tornberg, J. Cramer, P. C. de Groot, T. Picot, G. Johansson, and L. DiCarlo. Partial-measurement backaction and nonclassical weak values in a superconducting circuit. *Phys. Rev. Lett.*, 111:090506, 2013.

[138] T. C. White, J. Y. Mutus, J. Dressel, J. Kelly, R. Barends, E. Jeffrey, D. Sank, A. Megrant, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, I. C. Hoi, C. Neill, P. J. J. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, and J. M. Martinis. Preserving entanglement during weak measurement demonstrated with a violation of the Bell–Leggett–Garg inequality. *NPJ Quantum Inf.*, 2:15022, 2016.

[139] K. Nakamura, A. Nishizawa, and M.-K. Fujimoto. Evaluation of weak measurements to all orders. *Phys. Rev. A*, 85:012113, 2012.

[140] N. J. Cerf and C. Adami. Quantum Mechanics of Measurement. *arxiv*, 1997.

[141] J. R. Glick and C. Adami. Quantum information theory of the Bell-state quantum eraser. *Phys. Rev. A*, 95:012105, 2017.

[142] W. H. Zurek. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.*, 75:715–775, 2003.

[143] N. J. Cerf and C. Adami. Negative Entropy and Information in Quantum Mechanics. *Phys. Rev. Lett.*, 79:5194–5197, 1997.

[144] W. K. Wootters. Quantum measurements and finite geometry. *Found. Phys.*, 36:112–126, 2006.

[145] D. Deutsch. Quantum Theory as a Universal Physical Theory. *Internat. Journ. Theor. Phys.*, 24:1–41, 1985.

[146] M. Giustina et al. Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons. *Phys. Rev. Lett.*, 115:250401, 2015.

[147] B. Hensen et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, 2015.

[148] A. Peres and D. R. Terno. Quantum information and relativity theory. *Rev. Mod. Phys.*, 76:93–123, 2004.

[149] M. O. Scully and K. Drühl. Quantum eraser: A proposed photon correlation experiment concerning observation and "delayed choice" in quantum mechanics. *Phys. Rev. A*, 25:2208–2213, 1982.

[150] J. A. Wheeler. The "past" and the "delayed-choice" double-slit experiment. In A. R. Marlow, editor, *Mathematical Foundations of Quantum Theory*, pages 9–48. Academic Press, New York, NY, 1978.

[151] V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, and J.-F. Roch. Experimental realization of Wheeler's delayed-choice gedanken experiment. *Science*, 315:966–8, 2007.

[152] A. Peruzzo, P. Shadbolt, N. Brunner, S. Popescu, and J. L. O'Brien. A quantum delayed-choice experiment. *Science*, 338:634–7, 2012.

[153] F. Kaiser, T. Coudreau, P. Milman, D. B. Ostrowsky, and S. Tanzilli. Entanglement-Enabled Delayed-Choice Experiment. *Science*, 338(6107):637–640, 2012.

[154] R. Ionicioiu and D. R. Terno. Proposal for a Quantum Delayed-Choice Experiment. *Phys. Rev. Lett.*, 107:230406, 2011.

[155] Y. H. Kim, R. Yu, S. P. Kulik, Y. H. Shih, and M. O. Scully. Delayed "Choice" quantum eraser. *Phys. Rev. Lett.*, 84:1–5, 2000.

[156] X.-s. Ma, J. Kofler, and A. Zeilinger. Delayed-choice gedanken experiments and their realizations. *Rev. Mod. Phys.*, 88:015005, 2016.

[157] D. M. Greenberger and A. Yasin. Simultaneous wave and particle knowledge in a neutron interferometer. *Phys. Lett. A*, 128(8):391–394, 1988.

[158] B.-G. Englert. Fringe Visibility and Which-Way Information: An Inequality. *Phys. Rev. Lett.*, 77:2154–2157, 1996.

[159] P. J. Coles, J. Kaniewski, and S. Wehner. Equivalence of wave-particle duality to entropic uncertainty. *Nat. Commun.*, 5:5814, 2014.

[160] R. M. Angelo and A. D. Ribeiro. Wave-particle duality: an information-based approach. *Found. Phys.*, 45:1407–1420, 2015.

[161] E. Bagan, J. A. Bergou, S. S. Cottrell, and M. Hillery. Relations between Coherence and Path Information. *Phys. Rev. Lett.*, 116:160406, 2016.

[162] N. Bohr. Discussions with Einstein on epistemological problems in atomic physics. In J. A. Wheeler and W. H. Zurek, editors, *Quantum Theory and Measurement*, pages 9–49. Princeton University Press, Princeton, NJ, 1984.

[163] Y. Aharonov and M. S. Zubairy. Time and the quantum: erasing the past and impacting the future. *Science*, 307:875–9, 2005.

[164] R. Garisto and L. Hardy. Entanglement of projection and a new class of quantum erasers. *Phys. Rev. A*, 60:827–831, 1999.

[165] S. P. Walborn, M. O. Terra Cunha, S. Pádua, and C. H. Monken. *Phys. Rev. A*, 65:033818, 2002.

[166] M. O. Scully, B.-G. Englert, and H. Walther. Quantum optical tests of complementarity. *Nature*, 351:111–116, 1991.

[167] W. K. Wootters W. H. Zurek. Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr's principle. *Phys. Rev. D*, 19(2):473–484, 1979.

[168] D. Arsenović, M. Bozić, and L. Vusković. Is the classical law of the addition of probabilities violated in quantum interference? *J. Opt. B: Quantum and Semiclass. Opt.*, 4(4):S358, 2002.

[169] J. A. Wheeler and W. H. Zurek, editors. *Quantum Theory and Measurement.* Princeton University Press, Princeton, NJ, 1984.

[170] C. Adami and N. J. Cerf. What information theory can tell us about quantum reality. *Lecture Notes in Computer Science*, 1509:258–268, 1999.

[171] Č. Brukner. Quantum causality. *Nat. Phys.*, 10:259–263, 2014.

[172] C. Rovelli. An Argument Against the Realistic Interpretation of the Wave Function. *Found. Phys.*, 46(10):1229–1237, 2016.

[173] O. Oreshkov and N. J. Cerf. Operational formulation of time reversal in quantum theory. *Nat. Phys.*, 11, 2015.

[174] M. Reisenberger and C. Rovelli. Spacetime states and covariant quantum theory. *Phys. Rev. D*, 65:125016, 2002.

[175] S. J. Olson and J. P. Dowling. Information and measurement in generally covariant quantum theory. 2014.

[176] O. Oreshkov and N. J. Cerf. Operational quantum theory without predefined time. *New J. Phys.*, 18(7):073037, 2016.

[177] E. Nagali, S. Felicetti, P.-L. de Assis, V. D'Ambrosio, R. Filip, and F. Sciarrino. Testing sequential quantum measurements: how can maximal knowledge be extracted? *Sci. Rep.*, 2:443, 2012.

[178] J. Bergou, E. Feldman, and M. Hillery. Extracting Information from a Qubit by Multiple Observers: Toward a Theory of Sequential State Discrimination. *Phys. Rev. Lett.*, 111:100501, 2013.

[179] M. Hillery and D. Koch. Retrodiction of a sequence of measurement results in qubit interferometers. *Phys. Rev. A*, 94:032118, 2016.

[180] H. D. Zeh. Toward a quantum theory of observation. *Found. Phys.*, 3(1):109–116, 1973.

[181] A. N. Korotkov and A. N. Jordan. Undoing a Weak Quantum Measurement of a Solid-State Qubit. *Phys. Rev. Lett.*, 97:166805, 2006.

[182] A. N. Jordan and A. N. Korotkov. Uncollapsing the wavefunction by undoing quantum measurements. *Contemp. Phys.*, 51(2):125–147, 2010.

[183] J. R. Glick and C. Adami. Generating remote entanglement via disentangling operations. In preparation, 2017.

[184] E. Schneidman, S. Still, M. J. Berry, and W. Bialek. Network Information and Connected Correlations. *Phys. Rev. Lett.*, 91:238701, 2003.

[185] G. 'tHooft. Dimensional reduction in quantum gravity. In J. Ellis and S. Randjbar-Daemi, editors, *Salamfestschrift: A collection of talks*, volume 4 of *World Scientific Series in 20th Century Physics*, 1993.

[186] L. Susskind. The world as a hologram. *J. Math. Phys.*, 36(11):6377–6396, 1995.

[187] L. Susskind and E. Witten. The Holographic Bound in Anti-de Sitter Space. 1998.

[188] G. Evenbly and G. Vidal. Tensor network states and geometry. *J. Stat. Phys.*, 145:891–918, 2011.

[189] B. Swingle. Entanglement renormalization and holography. *Phys. Rev. D*, 86:065007, 2012.

[190] A. G. White, D. F. V. James, P. H. Eberhard, and P. G. Kwiat. Nonmaximally Entangled States: Production, Characterization, and Utilization. *Phys. Rev. Lett.*, 83:3103, 1999.

[191] T. Tanaka, Y. Ota, M. Kanazawa, G. Kimura, H. Nakazato, and F. Nori. Determining eigenvalues of a density matrix with minimal information in a single experimental setting. *Phys. Rev. A*, 89:012117, 2014.

[192] A. Peres. *Quantum Theory: Concepts and Methods.* Kluwer Academic, Dordrecht, 1995.

[193] A. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory.* Edizioni della Normale, 2011.

[194] E. Schrödinger. Die gegenwärtige Situation in der Quantenmechanik. *Die Naturwissenschaften*, 48:807–812, 1935.

[195] B.-G. Englert. On quantum theory. *Eur. Phys. J. D*, 67:238, 2013.

[196] C. A. Fuchs, N. D. Mermin, and R. Schack. An introduction to QBism with an application to the locality of quantum mechanics. *Am. J. Phys.*, 82:749–754, 2014.

[197] S. Lloyd. Capacity of the noisy quantum channel. *Phys. Rev. A*, 55:1613–1622, 1997.

[198] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin. Capacities of Quantum Erasure Channels. *Phys. Rev. Lett.*, 78:3217–3220, 1997.

[199] M. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), U. Sen, and B. Synak-Radtke. Local versus nonlocal information in quantum-information theory: Formalism and phenomena. *Phys. Rev. A*, 71:062307, 2005.

[200] T. Baumgratz, M. Cramer, and M. B. Plenio. Quantifying coherence. *Phys. Rev. Lett.*, 113:140401, 2014.

[201] Z. Xi, Y. Li, and H. Fan. Quantum coherence and correlations in quantum system. *Sci. Rep.*, 5:10922, 2015.

[202] F. Hiai and D. Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Commun. Math. Phys.*, 143(1):99–114, 1991.

[203] V. Vedral. The role of relative entropy in quantum information theory. *Rev. Mod. Phys.*, 74:197–234, 2002.

[204] D. Ruelle. *Statistical mechanics: Rigorous results.* World Scientific, 1999.

[205] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland. Quantum Zeno effect. *Phys. Rev. A*, 41:2295–2300, 1990.

[206] D. Home and M. A. B. Whitaker. A conceptual analysis of quantum Zeno; Paradox, measurement, and experiment. *Ann. Phys.*, 258:237–285, 1997.

[207] A. G. Kofman, G. Kurizki, and T. Opatrný. Zeno and anti-Zeno effect for photon polarization dephasing. *Phys. Rev. A*, 63:042108, 2001.

[208] P. G. Kwiat, A. G. White, J. R. Mitchell, O. Nairz, G. Weihs, H. Weinfurter, and A. Zeilinger. High-efficiency quantum interrogation measurements via the quantum Zeno effect. *Phys. Rev. Lett.*, 83:4725–4728, 1999.

[209] B. Kaulakys and V. Gontis. Quantum anti-Zeno effect. *Phys. Rev. A*, 56:1131–1137, 1997.

[210] M. Lewenstein and K. Rzążewski. Quantum anti-Zeno effect. *Phys. Rev. A*, 61:022105, 2000.

[211] A. Luis. Zeno and anti-Zeno effects in two-level systems. *Phys. Rev. A*, 67:062113, 2003.

[212] J. A. Wheeler. Information, physics, quantum: The search for links. In W.H. Zurek, editor, *Complexity, Entropy and the Physics of Information*. Addison-Wesley Publishing Co., Redwood City, CA, 1990.

[213] L. Bombellui, J. Lee, D. Meyer, and R. D. Sorkin. Spacetime as causal sets. *Phys. Rev. Lett.*, 59:521–524, 1987.

[214] M. Van Raamsdonk. Building up spacetime with quantum entanglement. *Gen. Relativ. Gravit.*, 42(10):2323–2329, 2010.

[215] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill. Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence. *J. High Energ. Phys.*, 2015(6):149, 2015.

[216] M. F. Pusey, J. Barrett, and T. Rudolph. On the reality of the quantum state. *Nat. Phys.*, 8(6):475–478, 2012.

[217] V. Bužek, M. Hillery, and R. F. Werner. Optimal manipulations with qubits: Universal-NOT gate. *Phys. Rev. A*, 60:R2626–R2629, 1999.

[218] N. Gisin and S. Popescu. Spin flips and quantum information for antiparallel spins. *Phys. Rev. Lett.*, 83:432–435, 1999.

[219] F. De Martini, V. Bužek, F. Sciarrino, and C. Sias. Experimental realization of the quantum universal NOT gate. *Nature*, 419(6909):815–818, 2002.

[220] F. De Martini, V. Bužek, F. Sciarrino, and C. Sias. Flipping qubits. *Part. Nucl. Lett.*, (1):116, 2003.

[221] E. P. Wigner. *Group theory and its application to the quantum mechanics of atomic spectra*. Pure Appl. Phys. Academic Press, New York, NY, 1959. Trans. from the German.

[222] A. K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67:661–663, 1991.

[223] C.-W. Chou, H. De Riedmatten, D. Felinto, S. V. Polyakov, S. J. Van Enk, and H. J. Kimble. Measurement-induced entanglement for excitation stored in remote atomic ensembles. *Nature*, 438(7069):828–832, 2005.

[224] J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, and H. Weinfurter. Heralded entanglement between widely separated atoms. *Science*, 337(6090):72–75, 2012.

[225] D. L. Moehring, P. Maunz, S. Olmschenk, K. C. Younge, D. N. Matsukevich, L. M. Duan, and C. Monroe. Entanglement of single-atom quantum bits at a distance. *Nature*, 449:68–71, 2007.

[226] H. Bernien, B. Hensen, W. Pfaff, G. Koolstra, M. S. Blok, L. Robledo, T. H. Taminiau, M. Markham, D. J. Twitchen, L. Childress, and R. Hanson. Heralded entanglement between solid-state qubits separated by three metres. *Nature*, 497(7447):86–90, 2013.

[227] N. Roch, M. E. Schwartz, F. Motzoi, C. Macklin, R. Vijay, A. W. Eddins, A. N. Korotkov, K. B. Whaley, M. Sarovar, and I. Siddiqi. Observation of measurement-induced entanglement and quantum trajectories of remote superconducting qubits. *Phys. Rev. Lett.*, 112:170501, 2014.

[228] A. Narla, S. Shankar, M. Hatridge, Z. Leghtas, K. M. Sliwa, E. Zalys-Geller, S. O. Mundhada, W. Pfaff, L. Frunzio, R. J. Schoelkopf, and M. H. Devoret. Robust concurrent remote entanglement between two superconducting qubits. *Phys. Rev. X*, 6:031036, 2016.

[229] K. Koshino, K. Inomata, Z. R. Lin, Y. Tokunaga, T. Yamamoto, and Y. Nakamura. Theory of deterministic entanglement generation between remote superconducting atoms. *Phys. Rev. Applied*, 7:064006, 2017.

[230] G. Mitchison, R. Jozsa, and S. Popescu. Sequential weak measurement. *Phys. Rev. A*, 76:062105, 2007.