Quantum Error Correction and Entanglement Wedge Reconstruction

Ben Karsberg

August 19, 2022

MSc in Theoretical Physics
The University of Edinburgh
2021

Abstract

This is where you summarise the contents of your dissertation. It should be at least 100 words, but not more than 200 words.

Declaration

I declare that this dissertation was composed entirely by myself.

Chapters 2 and 3 provide an introduction to the subject area and a description of previous work on this topic. They do not contain original research.

Chapter 4 describes work that was done entirely by me. The results of this chapter have been obtained previously by Anne T Matta, but the methods used here are different in some important (or minor) ways.

Chapters 4 through 6 contain my original work. The work described in Chapter 4 was done in collaboration with Professor Carole Ann O'Malley and her PhD student Jake O'Bean. Chapter 5 presents original work done entirely by me.

State whether calculations were done using Mathematica, SymPy, etc, with (or without) gamma matrix code, master integrals, the Super-Duper software package, etc. In other words, you should refer to any software that you used during your project. For example, Monte Carlo simulation packages, hydrodynamics packages, measurement code, fitting code, tensor algebra or calculus packages, Feynman diagram packages, etc.

State whether any software you used was written by you from scratch, by your supervisor (or by whoever), or if it's a standard package.

Personal Statement

You **must** *include a Personal Statement in your dissertation. This should describe what you did during the project, and when you did it. Give an account of problems you faced and how you attempted to overcome them. The examples below are based on personal statements from MSc and MPhys projects in previous years, with (mostly-obvious) changes to make them anonymous.*

Example 1: an analytical project

The project began with an introduction to the spinor-helicity formalism in four dimensions, with my main source material being H. Elvang's "Scattering Amplitudes in Gauge Theory and Gravity" [1]. I read the first chapter, and acquainted myself with the formalism, and how it worked in a practical sense.

Once I felt more comfortable with it, we moved onto the six-dimensional spinor-helicity formalism paper, where I spent some time gaining as strong an understanding of how the formalism worked, and proving identities.

The next stage was to learn about the generalised unitarity procedure, with the end goal being to use it to calculate coefficients for some one loop integral, likely involving massive particles. Learning how this worked took some time, and proved to be some of the most difficult material for me to understand. [1] [2]

It wasn't until later that we began to consider applying what I had learned to a Kaluza-Klein reduction, which ended up being the main focus of the project. It mixed well with the general theme of "extra-dimensional theory" the project began with, and allowed me to apply all that I'd learned and prepared for so far. The vast majority of my remaining time was spent calculating coefficients for the scalar box contribution to the gluon-gluon to two-Kaluza-Klein-particle amplitude, overcoming a number of problems and errors, to finally have human-readable, and presentable results.

During the course of the project, I met with my supervisor every week, in order to discuss my progress and the direction I would head next. Toward the end, the frequency of our meetings increased somewhat, as I began to finish my calculations.

I started writing this dissertation in mid-July, and I spent the first three weeks of August working on it full-time.

Overall, I feel that the project was a success, and I found it to be extremely enjoyable throughout.

Example 2: a computational project

I spent the first 2 weeks of the project reading the material surrounding my project - mainly [1] and [2]. I also began to plan out how I would implement the algorithms in C++, in doing this I gained an understanding of what the main goals of the first half of my project would be and how they could be achieved. I identified which Monte Carlo observables would be useful to measure in these simulations.

For the next 3 weeks I implemented the standard Atlantic City algorithm and debugged my code whilst developing analysis tools in python. I compared the results from my simulations to the results from [3] (for the Random Osculator) and [4] for the EvenMoreRandom Osculator. Having obtained positive results for the Random Osculator I started reading up on Heaviside Articulation. I examined how to integrate a Heaviside Articulator into the simulation in order to produce the most efficient simulation - the solution I decided on was to use a package called HeaviArt[5].

Following this I began to integrate the Heaviside Articulator into my code and test it against the regular algorithm. In addition to this I ran longer simulations to verify my findings without Articulation.

In mid July I finished implementing Heaviside Articulation into my code and began looking into how to quantify any improvement in speed gained by this algorithm. As July progressed I started looking into how to integrate the EvenMoreRandom Osculator into my code - this was the most complicated part of the project, as discussed in the body of this report. Despite much effort on my part, I couldn't get the results produced by the new algorithm to agree with the old ones. Following further study of the literature, and long discussions with Jack O'Bean, it turned out that the original form of Heaviside Articulation didn't applied to the EvenMoreRandom Osculator. With the help of Jack and my supervisor, I then developed the new version described in this report. I also did analytical calculations of the Four-Point Green-and-White- Function to two orders higher than had been published previously in the literature.

For the final parts of the summer I worked mainly on perfecting the algorithm for the Random Osculator and implementing the EvenMoreRandom Osculators algorithm with the improved Heaviside Articulation. The final results were encouraging, but more work is clearly needed. To this end, I have been awarded a studentship by the British University of Lifelong Learning to extend this work during my PhD Studies at the non-existent Scottish Highlands Institute of Technology in Inveroxter.

I started writing this dissertation in mid-July, and I spent the first three weeks of August working on it full-time.

Example 3: a very mathematical project

[In preparation]

Acknowledgements

Put your acknowledgements here. Thanking your supervisor for his/her help is standard practice, but it's not compulsory…

I'd like to thank my supervisor Professor Carole Ann O'Malley for making this project possible, and her PhD student Jack O'Bean for his patience and his detailed functional explanations of how classical symmetries can be broken by quantum effects. Thanks also to Wally Bee and Ken Garoo for sending me their hopping-parameter expansions.

Finally, none of this would have been possible without Catriona Sutherland's witchcraft.

This document has its origins in the dissertation template for the MSc in High Performance Computing, which is apparently descended from a template developed by Professor Charles Duncan for MSc students in Meteorology. His acknowledgement follows:

This template has been produced with help from many former students who have shown different ways of doing things. Please make suggestions for further improvements.

Some parts of this template were lifted unashamedly from the Edinburgh MPhys project report guide, with little or no modification. I have no idea who wrote the first version of that…

You don't have to use LaTeX for your dissertation. You can use Microsoft Word, Apple's Pages, LibreOffice (or similar) if you prefer, but it's *much* easier to typeset equations in LaTeX, and references look after themselves. Whatever you use, your dissertation should have the same general structure as this one, and it should look similar – especially the front page.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The Introduction should contain a description of your project and the problem you are trying to solve. It should start off at a level that should be understandable by anyone with a degree in physics, but it can become more technical later

Where appropriate you should include references to work that has already been done on your topic and anything else which lets you set your work in context.

One of the things you will need to do is to ensure that you have a suitable list of references. To do this you should see [**?**] or some other suitable reference. Note the format of the citation used here is the style favoured in this School. Here is another reference [**?**] for good measure.

Alternatively, you can use BibTeX. See later for some details on this.

You will also want to make sure you have no spelling or grammatical mistakes. To help identify spelling mistukes you caan use the commands *ispell* or *spell* on any Linux/unix machine. See the appropriate manual pages. Remember that spelling mistakes are not the only errors which can occur. Spelling checkers will not find errors which are, in fact, valid words such as *there* for *their*, nor will they find repeated repeated words which sometimes occur if your concentration is broken when typing. **There is no substitute for thorough proof reading!**

Your dissertation should be no longer than 15,000 words. In terms of pages, 30 pages are ok. 50 pages are fine. But it shouldn't be much longer than that.

Chapter 2

Error Correction

In this chapter, I will introduce quantum error correction. I will start with some classical preliminaries to build intuition, before moving to the quantum regime. I will discuss the concept of noise, and introduce quantum operations (not be to be confused with operators) in the context of closed and open systems, before moving onto quantum error correction itself. I will follow the presentation of Nielsen and Chuang [2].

2.1 Classical Noise

2.1.1 The Classical Bit-Flip

To begin, we consider a classical example of error correction, called the *bit-flip code* to build up some intuition. Suppose Alice wishes to send Bob a bit string - a sequence of 0s and 1s - wirelessly. However, phone signals, radio waves, and all sorts of radiation permeate the air and these may interact with the transmitted bit string, causing some 0s to 'flip' to 1s and vice versa. This is known as *noise*. We model this as a constant probability p for an individual bit to flip before it reaches Bob, and $1 - p$ for the bit to remain the same. Schematically, the process looks like:

FIGURE

To figure out what p is, we need a model for the distribution of electromagnetic radiation in the environment, and a model for how this radiation interacts with Alice's transmitted bits. This basic idea - having a model for both the environment and the system/environment interaction - is the basic starting point for analysing any classical or quantum error process. Let's suppose that p_0 and p_1 are the initial

probabilities that a bit is a 0 or a 1 respectively, and q_0 and q_1 are the corresponding probabilities at the end of the process, after being subjected to noise. If we denote the initial state of an individual bit as A and the final state as B, then we have

$$\mathbb{P}(B = b) = \sum_{a=0}^{1} \mathbb{P}(B = b \,|\, A = a)\mathbb{P}(A = a) \tag{2.1.1}$$

which can be rewritten in matrix form as

$$\begin{pmatrix} q_0 \\ q_1 \end{pmatrix} = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix}. \tag{2.1.2}$$

The quantities $\mathbb{P}(B = b \,|\, A = a)$ are called *transition probabilities*. If further noise occurs after an initial bit-flip, we can model this as being independent of the first flip; this independence is called *markovicity*, and the total noise process is a *Markov process*.
More generally, for a single-state process such as this, we have that the vector of output probabilities $\mathbf{q}$ is related to the vector of input probabilities $\mathbf{p}$ by some matrix E via

$$\mathbf{q} = E\mathbf{p} \tag{2.1.3}$$

where E is known as an *evolution matrix*. Since $E\mathbf{p}$ is a probability distribution, we can instantly extract some conditions on E:

1. All entries of E must be non-negative, so E is a positive matrix,
2. All columns of E sum to 1, called *completeness*.

2.2 Quantum Noise

The basic classical model presented above can be extended to the quantum world. To do so, we need to introduce the theory of *quantum operations*, which can be used to describe the evolution of quantum systems in a vastly more general way than simple unitary evolution. They can describe evolution under quantum noise and Markov processes in particular, so have special importance for our purposes.
Similar to the classical evolution of probabilities (2.0.3), we want a quantum state ρ to evolve as

$$\rho' = \mathcal{E}(\rho). \tag{2.2.1}$$

$\mathcal{E}$ is a quantum operation. A simple example of this is unitary time-evolution, which would be $\mathcal{E}(\rho) = U\rho U^\dagger$. We can think of this unitary evolution as a box which takes a state ρ as an input, and outputs $U\rho U^\dagger$. A *closed* quantum system

is one which evolves like this, but many more realistic systems are *open*, which we think of as arising from the interaction between a *principal system* of interest and an environment, together forming a closed system. In open systems, a state ρ in the principal system evolving as $\mathcal{E}(\rho)$ may not be a unitary evolution.
To be more explicit, suppose the total system-environment system is in the product state $\rho \otimes \rho_{\text{env}}$; then, the total product state evolves under the action of a unitary transformation, so the quantum operation governing the evolution of ρ alone can be found by taking a trace over the environment:

$$\mathcal{E}(\rho) = \text{Tr}_{\text{env}} \left[U(\rho \otimes \rho_{\text{env}}) U^\dagger \right]. \tag{2.2.2}$$

One immediate issue with this definition of quantum operations is that we have to express evolution of the principal system in terms of operators on the environment. How can we rectify this?

2.2.1 Operator-Sum Representation

Suppose $\{|e_k\rangle\}$ is an orthonormal basis for the environment system such that initially, the environment is in the state $\rho_{\text{env}} = |e_0\rangle \langle e_0|$. Then, we can rewrite (??) as

$$\mathcal{E}(\rho) = \sum_k \langle e_k | U \left[\rho \otimes |e_0\rangle \langle e_0| \right] U^\dagger |e_k\rangle \equiv \sum_k E_k \rho E_k^\dagger \tag{2.2.3}$$

where it is implicit that $E_k = \langle e_k | U | e_0 \rangle$. These are operators on $\mathcal{H}$ only, so fix our issue! This defines the *operator-sum representation* of $\mathcal{E}$, and the set $\{E_k\}$ are called *operation elements*.
The operation elements have a completeness property; since $\mathcal{E}(\rho)$ is itself a density operator, we must have $\text{Tr}(\mathcal{E}(\rho)) = 1$, meaning that

$$1 = \text{Tr}\left(\sum_k \rho E_k^\dagger \right) = \text{Tr}\left(\sum_k \rho E_k^\dagger E_k \right) \tag{2.2.4}$$

holds for all ρ, which further implies

$$\sum_k E_k^\dagger E_k = I. \tag{2.2.5}$$

If this equation is satisfied, $\mathcal{E}$ is called a *trace-preserving operation*. We assume all operations from here on out to be trace-preserving, unless stated otherwise.
The operator-sum representation of a quantum operation is not unique. The following theorem characterises this, stated without proof.

Theorem 2.2.1 (Unitary freedom in the operator-sum representation)**.** *Suppose* $\{E_1, \ldots, E_m\}$ *and* $\{F_1, \ldots, F_n\}$ *are operation elements of operations* $\mathcal{E}$ *and* $\mathcal{F}$ *respectively. Append zero operators to the shorter list of elements to ensure* $m = n$*. Then,* $\mathcal{E} = \mathcal{F}$ *if and only if* $E_i = \sum_j u_{ij} F_j$*, where* u_{ij} *are the elements of an* $m \times m$ *unitary matrix.*

2.2.2 Axiomatisation

We can define quantum operations from a purely axiomatic, mathematical perspective, with no reference to an environment at all. Denoting the set of density operators on a Hilbert space $\mathcal{H}$ as $\mathcal{B}(\mathcal{H})$ (note that this notation is often used to denote the set of all operators on a space, rather than **density** operators), a map $\mathcal{E} : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ is called a quantum operation if it satisfies:

1. $\mathrm{Tr}(\mathcal{E}(\rho))$ is the probability that the process represented by $\mathcal{E}$ occurs, so $0 \leq \mathrm{Tr}(\mathcal{E}(\rho)) \leq 1$.

2. $\mathcal{E}$ is *convex-linear*; that is, for a set of probabilities $\{p_i\}$ and density matrices $\{\rho_i\}$, we have
$$\mathcal{E}\left(\sum_i p_i \rho_i\right) = \sum_i p_i \mathcal{E}(\rho_i). \tag{2.2.6}$$

3. $\mathcal{E}$ is a completely positive map; so for any positive operator A, $\mathcal{E}(A)$ is also a positive operator. More generally, if we introduce an auxiliary system R, $(I_R \otimes \mathcal{E})(B)$ is positive for any positive operator B on $R \otimes \mathcal{H}_1$.

We can in fact show that any map $\mathcal{E}$ satisfying these properties has an operator-sum representation, which is formalised by the following theorem, again stated without proof:

Theorem 2.2.2. *A map* $\mathcal{E} : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ *satisfies the above 3 axioms if and only if*
$$\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger \tag{2.2.7}$$
for some set of operators $E_i : \mathcal{H}_1 \to \mathcal{H}_2$*, and* $\sum_i E_i^\dagger E_i \leq I$*.*

2.2.3 The Quantum Bit-Flip

To show how quantum operations can model noise, we turn back to the bit-flip scenario. This time, Alice wishes to send Bob a qubit $|\psi\rangle = a|0\rangle + b|1\rangle$ to Bob,

which has density operator

$$\begin{aligned}\rho &= |\psi\rangle\langle\psi| \\ &= |a|^2 |0\rangle\langle 0| + ab^* |1\rangle\langle 0| + a^* b |0\rangle\langle 1| + |b|^2 |1\rangle\langle 1| \\ &= \begin{pmatrix} |a|^2 & ab^* \\ a^* b & |b|^2 \end{pmatrix}\end{aligned}. \tag{2.2.8}$$

The quantum channel she sends the qubit through acts as a bit-flip exactly analogously to the classical case, flipping a $|0\rangle$ to a $|1\rangle$ and vice versa with probability p, and doing nothing with probability $1-p$. The operation elements are therefore

$$E_0 = \sqrt{1-p}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E_1 = \sqrt{p}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.2.9}$$

This is analogous to (2.0.2) and (2.0.3) in the classical case in a sense, and shows how a quantum operation can model noise. Another simple example is the *phase-flip channel*, which is uniquely quantum. This operation flips the relative phase of the $|0\rangle$ and $|1\rangle$ basis elements; that is, it takes $a|0\rangle + b|1\rangle \rightarrow a|0\rangle - b|1\rangle$ with probability p, and does nothing with probability $1-p$. This operation has operation elements

$$E_0 = \sqrt{1-p}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E_1 = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.2.10}$$

2.3 Error Correction

While we've discussed how to model noise and errors, we have yet to look at how Alice and Bob can actually correct for them. Starting with the bit-flip, suppose Alice wishes to send $|\psi\rangle = a|0\rangle + b|1\rangle$ to Bob across the bit-flip channel, as stated above. Bob only has a probability of $1-p$ of receiving the transmitted qubit as intended by Alice. How can we improve this?

The process of error correcting starts with Alice *encoding* her initial qubit into a *logical* or *code subspace*. She adjoins two ancillary qubits in the $|0\rangle$ state, and applies two $CNOT$ gates conditioned on the first qubit to the second and third qubit. This has the effect of mapping

$$\begin{aligned}|0\rangle &\rightarrow |0_L\rangle \equiv |000\rangle \\ |1\rangle &\rightarrow |1_L\rangle \equiv |111\rangle .\end{aligned} \tag{2.3.1}$$

So overall, the original qubit is mapped as

$$|\psi\rangle = a|0\rangle + b|1\rangle \rightarrow |\psi_L\rangle = a|000\rangle + b|111\rangle . \tag{2.3.2}$$

This is then sent to Bob across the bit-flip channel. The three qubits that Bob receives may have suffered a bit-flip, but he needs to determine both whether this occurred, and if so, on which qubit it occurred on. This is called *error detection* or sometimes *syndrome diagnosis*. To do this, he performs a projective measurement with projectors

$$\begin{aligned} P_0 &= |000\rangle\langle 000| + |111\rangle\langle 111| \qquad \text{(no error)} \\ P_1 &= |100\rangle\langle 100| + |011\rangle\langle 011| \qquad \text{(qubit 1 flipped)} \\ P_2 &= |010\rangle\langle 010| + |101\rangle\langle 101| \qquad \text{(qubit 2 flipped)} \\ P_3 &= |001\rangle\langle 001| + |110\rangle\langle 110| \qquad \text{(qubit 3 flipped).} \end{aligned} \tag{2.3.3}$$

To see why these correspond to the errors stated in brackets, consider the case where only the first qubit flips, so Bob receives

$$|\psi_E\rangle \equiv a\,|100\rangle + b\,|011\rangle\,. \tag{2.3.4}$$

Note that $\langle\psi_E|P_1|\psi_E\rangle = 1$, so the outcome of the measurement is 1 with certainty. Moreover, $P_1\,|\psi_E\rangle = |\psi_E\rangle$, so the syndrome measurement does not change the state. Bob now knows with certainty that a bit-flip occurred on the first qubit, so can apply an X gate to the first qubit to flip it back, and recover the original qubit by measuring off the ancillary qubits.
This procedure works perfectly so long as only one qubit flips; this occurs with probability $(1-p)^3 + 3p(1-p)^2 = 1 - 3p^2 + 2p^3$, so the probability Bob cannot correct the error is $3p^2 - 2p^3$. Since $3p^2 - 2p^3 \leq p \implies p \leq 1/2$, we have increased reliability so long as $p \leq 1/2$.

2.3.1 Generalities

While simple, the bit-flip code provides an example of features common to all error correcting procedures. In general, a state $|\psi\rangle$ is encoded by an isometry (usually consisting of adjoining on some ancillary state and applying a unitary operator) into a code subspace $\mathcal{H}_{\text{code}}$ of some larger Hilbert space $\mathcal{H}$. In the bit-flip code for example, $\mathcal{H}_{\text{code}} = \text{span}\{|0_L\rangle\,, |1_L\rangle\}$. We will often refer to the *code space projector*, denoted P_{code}. After encoding, the state is subjected to noise, and then a syndrome measurement is performed to diagnose the type of error which occurred (if any), and then a recovery operation is performed to obtain the original state. Note that different errors have to correspond to orthogonal subspaces of the full Hilbert space $\mathcal{H}$ in order to be reliably distinguished.
In the general theory of error correction, no assumptions are made about the full recovery procedure - in particular, we do not assume it is necessarily a two-stage detection-recovery process. We just assume that the noise is modelled by

an operation $\mathcal{E}$, and the correction is performed by an operation $\mathcal{R}$. For error correction to be deemed successful, we require that for any state ρ with support on $\mathcal{H}_{\text{code}}$, we have

$$(\mathcal{R} \circ \mathcal{E})(\rho) \propto \rho, \tag{2.3.5}$$

where we have a proportionality constant rather than equality to account for the possibility that the relevant operations may not be trace-preserving.
Not all noise is correctable. This is characterised by the *quantum error correction conditions*, which can be stated as the following theorem:

Theorem 2.3.1 (Quantum error-correction conditions)**.** *Let $\mathcal{H}_{code}$ be a quantum code, and P_{code} the projector onto it. Suppose $\mathcal{E}$ is a quantum operation modelling noise, with elements $\{E_i\}$. An error correction operation $\mathcal{R}$ correcting $\mathcal{E}$ on $\mathcal{H}_{code}$ exists if and only if*

$$P_{code} E_i^\dagger E_j P_{code} = \alpha_{ij} P_{code} \tag{2.3.6}$$

where α_{ij} are the elements of a complex hermitian matrix.

The proof of this theorem can be found in appendix B. The set $\{E_i\}$ are called the *errors*, and if an $\mathcal{R}$ exists then we say they are a *correctable set* of errors.
In general, we may not know the form of the noise precisely, but the error correction conditions can be adapted to characterise an equivalence class of noise which a code $\mathcal{H}_{\text{code}}$ and correction operation $\mathcal{R}$ can correct for.

Theorem 2.3.2. *Suppose $\mathcal{H}_{code}$ is a quantum error correction code, and $\mathcal{R}$ is the full error-correcting operation correcting $\mathcal{E}$ with errors $\{E_i\}$. Then, $\mathcal{R}$ corrects for $\mathcal{F}$ with errors $\{F_i\}$ if*

$$F_i = \sum_i m_{ij} E_j \tag{2.3.7}$$

for all i, and m_{ij} are some the elements of a complex matrix m on $\mathcal{H}_{code}$.

This is a useful statement, as we can talk about a class of errors $\{E_i\}$ which are correctable rather than a class of noises $\mathcal{E}$. For example, if we can find a process satisfying

$$P_{\text{code}} \sigma_i^1 \sigma_j^1 P_{\text{code}} = \alpha_{ij} P_{\text{code}} \tag{2.3.8}$$

for the Pauli matrices, then we can correct for **arbitrary** single qubit errors, since any single qubit operation has operation elements which can be chosen to be the Pauli matrices. A code called the Shor code can do this, for example.

2.3.2 Quantum Erasure

One particularly important class of errors which are relevant for us is that of *quantum erasure*. This is defined as the channel acting to erase a **known** subsystem of the transmitted state. Formally, we suppose that $\mathcal{H}_{\text{code}} \subseteq \mathcal{H}$, where $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\overline{A}}$ has a tensor product structure. The erasure channel (which is non-trace-preserving) can then be modelled by

$$\mathcal{E}(\rho) = \text{Tr}_{\overline{A}}(\rho), \tag{2.3.9}$$

for any ρ with support on $\mathcal{H}$, which erases the $\overline{A}$ system with certainty. This representation is not unique; indeed, erasure can be more loosely defined as simply erasing all information within the subsystem (for example, the decoherence channel in the case of a single qubit), so we no longer have any access to it. To extract the operation elements, we define $\{|a\rangle\}$ and $\{|\overline{a}\rangle\}$ to be orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_{\overline{A}}$ respectively. We can then rewrite (??) as

$$\mathcal{E}(\rho) = \sum_{\overline{a}} \langle \overline{a}|\rho|\overline{a}\rangle \,, \tag{2.3.10}$$

from which we can read off operation elements

$$E_{\overline{a}} \equiv I_A \otimes \langle \overline{a}| = \sum_a |a\rangle \langle a| \otimes \langle \overline{a}| \,. \tag{2.3.11}$$

The natural question to ask is what the quantum error correction conditions reduce to for erasures. To work this out, we can compute

$$E_{\overline{a}}^\dagger E_{\overline{b}} = (I_A \otimes |\overline{a}\rangle)(I_A \otimes \langle \overline{b}|) = I_A \otimes |\overline{a}\rangle \langle \overline{b}| \,, \tag{2.3.12}$$

and so (2.3.6) then reduces to

$$P_{\text{code}} |\overline{a}\rangle \langle \overline{b}| P_{\text{code}} = \alpha_{\overline{a}\overline{b}} P_{\text{code}} \tag{2.3.13}$$

where we drop the I_A for notational simplicity since the $|\overline{a}\rangle$s do not have any action on the A subsystem. Even more simply, we can just write

$$P_{\text{code}} |\overline{a}\rangle \langle \overline{b}| \propto P_{\text{code}}. \tag{2.3.14}$$

This will have an important implication when we come to look at [1]. Note that an arbitrary operator $X_{\overline{A}}$ acting on $\mathcal{H}_{\overline{A}}$ can be decomposed in the $\{|\overline{a}\rangle\}$ basis as

$$X_{\overline{A}} \equiv \sum_{\overline{a},\overline{b}} x_{\overline{a},\overline{b}} |\overline{a}\rangle \langle \overline{b}| \tag{2.3.15}$$

for some $x_{\overline{a},\overline{b}} \in \mathbb{C}$. Therefore, if (2.3.14) holds, we have

$$P_{\text{code}} X_{\overline{A}} P_{\text{code}} \propto P_{\text{code}}. \tag{2.3.16}$$

This will be our starting point for analysing theorem 1 of [1].

An Example

As an example of erasure, we present the three-*qutrit* code of [1]. A qutrit is exactly analogous to a qubit, except the underlying Hilbert space has three basis elements, which we denote $\{|0\rangle, |1\rangle, |2\rangle\}$; the qutrit can then be written

$$|\psi\rangle = \sum_{i=0}^{2} a_i |i\rangle \tag{2.3.17}$$

where $\sum_{i=0}^{2} |a_i|^2 = 1$. Suppose Alice wishes to send this qutrit to Bob through a channel which acts to erase 1 of every three transmitted qutrits with certainty. To protect for this, Alice encodes her qutrit into the logical code subspace $\mathcal{H}_{\text{code}} = \text{span}(|\tilde{0}\rangle, |\tilde{1}\rangle, |\tilde{2}\rangle)$, defined by

$$\begin{aligned} |\tilde{0}\rangle &= \frac{1}{\sqrt{3}}(|000\rangle + |111\rangle + |222\rangle) \\ |\tilde{1}\rangle &= \frac{1}{\sqrt{3}}(|012\rangle + |120\rangle + |201\rangle). \\ |\tilde{2}\rangle &= \frac{1}{\sqrt{3}}(|021\rangle + |102\rangle + |210\rangle) \end{aligned} \tag{2.3.18}$$

Suppose that the erasure acts on the third qutrit. Bob then only has access to the first two qutrits, but he can still recover the original state. Define a unitary operator on the first two qutrits by

$$\begin{aligned} U_{12} \equiv\, &|00\rangle\langle 00| + |01\rangle\langle 11| + |02\rangle\langle 22| + |12\rangle\langle 01| + |10\rangle\langle 12| + |11\rangle\langle 20| \\ &+ |21\rangle\langle 02| + |22\rangle\langle 10| + |20\rangle\langle 21|, \end{aligned} \tag{2.3.19}$$

which does nothing to $|00\rangle$, and permutes the remaining 8 basis states as

$$|11\rangle \to |01\rangle \to |12\rangle \to |10\rangle \to |22\rangle \to |02\rangle \to |21\rangle \to |20\rangle \to |11\rangle. \tag{2.3.20}$$

We can then compute that

$$(U_{12} \otimes I_3)|i_L\rangle = |i\rangle \otimes \frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle), \tag{2.3.21}$$

which explicitly shows state recovery is possible given access to only the first two qutrits, since then

$$(U_{12} \otimes I_3)|\psi_L\rangle = |\psi\rangle \otimes \frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle). \tag{2.3.22}$$

This procedure holds irrelevant of which qutrit has been erased; we can just define a unitary operator with equivalent action to (2.3.19) acting on the remaining two qutrits. This correctability also has an interpretation in terms of operators. Suppose O acts on the initial space of a single qutrit as

$$O\left|i\right\rangle = \sum_{j=0}^{2}(O)_{ji}\left|j\right\rangle . \tag{2.3.23}$$

We can then find a 'logical' $\tilde{O}$ which acts on $\mathcal{H}_{\text{code}}$ similarly:

$$\tilde{O}\left|\tilde{i}\right\rangle = \sum_{j=0}^{2}(O)_{ji}\left|\tilde{j}\right\rangle . \tag{2.3.24}$$

For this example, we can actually find such a logical operator with support on the first two qutrits only:

$$O_{12} \equiv U_{12}^{\dagger} O U_{12} \tag{2.3.25}$$

where we take O to act on the first qutrit only.

2.4 Holography

In this section, I will give a brief summary of what constitutes a holographic theory, and how this can be translated to the language of error correction. We also introduce some holographic terminology which we will frequently refer back to. It should be emphasized that the focus of this dissertation is on error correction rather then holography; this exposition will therefore focus on intuition rather than the full details of holography. The goal is to provide some intuition for the concepts of a *Ryu-Takayanagi (RT) formula*, the *causal and entanglement wedges*, *complementary recovery*, and *radial commutativity*.

2.4.1 Quantum Gravity

General relativity is inherently classical. If we have enough initial data about a gravitational situation (i.e. the position and velocity of all particles, the electric and magnetic fields everywhere, the stress-energy tensor, and the metric), then Einstein's equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{2.4.1}$$

can be used to find a description of the situation at a later time with certainty. Quantum mechanics, while similar, is non-classical. Given a Hamiltonian and initial wavefunction, the Schrodinger equation governs evolution of the wavefunction

in time. We can then recover classical quantities by operating on the wavefunction with the relevant observables; for example, the position or momentum operators. While the expectation values of these observables evolves smoothly, if we actually measure an observable then we project onto one of its eigenstates, and only by repeated measurement of an identical system can we extract the expectation values.

2.4.2 The Holographic Dictionary

A holographic theory is one in which a gravitational theory is in direct correspondence with a non-gravitational theory on its boundary

2.4.3 The RT Formula

The RT formula is one part of the aforementioned holographic dictionary, providing a link between the entanglement entropy of states in the boundary CFT and the geometry of the dual AdS space. The most simple form of the RT formula applies to a CFT state dual to a purely classical geometry:

$$S_A(\rho) = \frac{1}{4G} \min_{\gamma_A} Area(\gamma_A). \tag{2.4.2}$$

On the right hand side, γ_A refers to a geodesic in the bulk (or extremal surface in higher dimensions) which hits the boundary at the edge of some chosen subregion A of the boundary. We then minimise over all such γ_A, choosing the one with smallest length/area. The left hand side refers to the entanglement entropy in subregion A of a state ρ living in the boundary CFT. This explicitly provides a quantitative entry in the dictionary to compute these entropies in terms of purely bulk geometric quantities.
More generally, the RT formula has an extended form which applies to states which have entanglement in the bulk and superpositions of geometries. This form is

$$S_A(\rho) = S_{\text{bulk},A}(\rho) + \text{Tr}(\mathcal{L}\rho). \tag{2.4.3}$$

Here, $\mathcal{L}$ is an operator acting on the quantum-gravitational Hilbert space, with eigenstates dual to classical geometries, and eigenvalues being the areas of the minimal surfaces meeting the boundary subregion A. Loosely, there are multiple states living on the same geometry; the 'bulk entanglement' term $S_{\text{bulk},A}$ identifies which of these states is described by ρ.

2.4.4 The Causal and Entanglement Wedges

Lorentzian metrics describing a spacetime have a causal structure, so an operator at a point in the bulk can only be affected by a boundary subregion which can send or receive a signal to said bulk point. Formally, for a boundary subregion A, the causal wedge $\mathcal{C}[A]$ is the region of the bulk bounded by the boundary region of dependence on A itself, and the set of bulk geodesics which start and end in this region of dependence. (FIGURE)
It's often convenient to work with a spatial slice within the causal wedge. If the RT surface is spacelike, all spatial slices in the causal wedge end on the RT surface itself, but hit the boundary at different times. In this case, we can choose a spatial slice which intersects the boundary precisely at A. In nice situations (for example, a static spacetime), we can pick a spatial slice extending between A and its RT surface, which by causality has all the information we need to reconstruct the full causal wedge. In cases like this, we can draw diagrams suppressing time entirely.
Alternatively, given access to the entropy of a CFT subregion A, the RT formula says we can compute the area of the relevant extremal surface. Intuitively, we expect that if we know the full reduced density operator of a CFT state ρ on A, we can construct the full RT surface itself. We can then use this information to construct the RT surfaces of smaller parts of the subregion A, and so we should be able to find the metric everywhere in the bulk region between A and its RT surface. This idea is formalised by the entanglement wedge $\mathcal{E}[A]$. This is the domain of dependence of the bulk bounded by A, and the RT surface of A. $\mathcal{E}[A]$ is determined by the RT surface, which has area proportional to the von Neumann entropy of the boundary theory contained in A, which motivates the name.
It can be shown that $\mathcal{C}[A] \subseteq \mathcal{E}[A]$. For a pure boundary state, the RT surface of A is the same as the RT surface of $\overline{A}$; the operator $\mathcal{L}$ giving its area is in both the set of operators acting on A and those acting on $\overline{A}$. Causality says that both these sets of operators must mutually commute, so $\mathcal{L}$ must commute with all operators on A and on $\overline{A}$. We say it's in the *centre* of the operators acting on A. One example of such an operator is the identity, but if there is a gauge symmetry in the bulk theory, then there will be non-trivial elements of the centre, and the area operator $\mathcal{L}$ will be one of these. Non-triviality of $\mathcal{L}$ therefore tells us that the bulk is gravitational and has diffeomorphism invariance.

2.4.5 Complementary Recovery and Radial Commutativity

The causal wedge tells us which operators in the bulk can be reconstructed from information in a boundary region A. If we divide the boundary into two disjoint

regions, an operator in an arbitrary point in the bulk must be in the causal wedge of one of the two regions, and only lies in both when it the point is on the corresponding RT surface. This is called *complementary recovery*: given a subregion, we can reconstruct all operators in its causal wedge, but none of the operators outside of it.
If instead we allow the region A to vary, a fixed operator in the bulk lies in the causal wedge of many regions. If we than have access to a boundary region with the property that many of its subregions can individually reconstruct the operator, we don't need the full state on such a region to reconstruct it, and there are multiple possibilities for reconstruction. If this is true, we say the full state has *subregion duality*. If we erase a piece of the boundary A which is much smaller than A itself, almost every bulk operator can still be reconstructed - see the analogy to error correction?
The converse to subregion duality is *radial commutativity*. For sensible spacetimes, the RT surface of a small subregion does not penetrate deep into the bulk; we need large subregions to do this. If a bulk operator lies outside the causal wedge of a subregion, it commutes with every operator inside the causal wedge; in particular, those operators which act on the boundary itself. However, every boundary operator acting at a single point at the boundary lives in the causal wedge of any boundary subregion containing the point, including arbitrarily small subregions around the point. So any bulk operator which does not act at the boundary itself must commute with **every** operator acting at a single boundary point, which is radial commutativity.
This is a problem. In field theories, the operator product expansion states that the product of operators acting at multiple points can be decomposed as a sum of local operators acting only at a single point. Radial commutativity then implies that all of these local operators commutes with the bulk operator - this would mean that any bulk operator commutes with every boundary operator. This conclusion is, however, only true since we treated the bulk Hilbert space as being the same as the boundary CFT Hilbert space. This is not true: the bulk Hilbert space for a given geometry is encoded inside the CFT Hilbert space. The resolution is therefore to map the bulk into the boundary via an isometry.

Chapter 3

Holographic Error Correction

We now present the first theorem of [1]. This is the most basic, non-general formulation of holographic error correction, characterising what makes a conventional quantum erasure correcting code holographic.

Theorem 3.0.1. *Let $V : \mathcal{H}_L \to \mathcal{H}$ be an encoding isometry, where $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\overline{A}}$. Define an orthonormal basis $\{|\tilde{i}\rangle\}$ of $\mathcal{H}_L$, and let $|\phi\rangle \equiv \frac{1}{\sqrt{|R|}} |i\rangle_R (V|\tilde{i}\rangle)_{A\overline{A}}$, where R is an auxiliary system with $\mathcal{H}_R = \mathcal{H}_L$, The following statements are then equivalent:*

1. *$|R| \leq |A|$, and if we decompose $\mathcal{H}_A = (\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}) \oplus \mathcal{H}_{A_3}$ with $|A_1| = |R|$ and $|A_3| < |R|$, then there exists a unitary transformation U_A on $\mathcal{H}_A$ and a state $|\chi\rangle_{A_2\overline{A}} \in \mathcal{H}_{A_2\overline{A}}$ such that*

$$(U_A \otimes I_{\overline{A}}) V |\tilde{i}\rangle_{A\overline{A}} = |i\rangle_{A_1} \otimes |\chi\rangle_{A_2\overline{A}}, \tag{3.0.1}$$

   where $|i\rangle_{A_1}$ is an orthonormal basis for $\mathcal{H}_{A_1}$.

2. *For any operator $\tilde{O}$ acting within $\mathcal{H}_L$, there exists an operator O_A on $\mathcal{H}_A$ such that for any state $|\tilde{\psi}\rangle \in \mathcal{H}_L$, we have*

$$\begin{aligned} O_A V |\tilde{\psi}\rangle &= V\tilde{O} |\tilde{\psi}\rangle \\ O_A^\dagger V |\tilde{\psi}\rangle &= V\tilde{O}^\dagger |\tilde{\psi}\rangle . \end{aligned} \tag{3.0.2}$$

3. *For any operator $X_{\overline{A}}$ on $\mathcal{H}_{\overline{A}}$, we have*

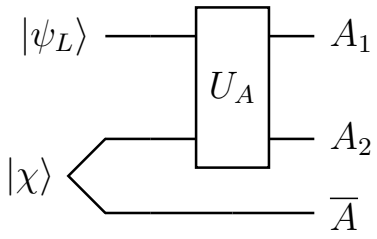
$$P_L V^\dagger X_{\overline{A}} V P_L \propto P_{code} \tag{3.0.3}$$

   where $P_L = \sum_i |\tilde{i}\rangle \langle\tilde{i}|$ is the projector onto $\mathcal{H}_L$.

4. In the state $|\phi\rangle$, we have

$$\rho_{R\overline{A}}[\phi] = \rho_R[\phi] \otimes \rho_{\overline{A}}[\phi]. \tag{3.0.4}$$

Before proving this, let's go over some of the intuition behind each statement, and what all the objects in this theorem refer to. A is the subsystem which is preserved by erasure, and $\overline{A}$ is the erased subsystem. Statement 1 just says that we can recover the full state of the code subspace on the system A_1 by applying some unitary $U_A^\dagger$; we can recover the state with access to the non-erased subsystem only. We can visualise this by means of a circuit diagram:



Statement 2 says that any logical operator on the code space can be equivalently represented by an operator acting on A only. Statement 3 is just equation (2.3.16); the quantum error correction conditions adapted to erasures. In a more physically intuitive way, this says that performing a measurement of any operator on the erased subsystem cannot disturb the encoded information - a plausible condition for erasure to be correctable. Condition 4 states that operators on the auxiliary system R and operators on the erased subsystem $\overline{A}$ are not correlated. The proof of this theorem is as follows:

Proof. (1) $\implies$ (2): Define $O_A \equiv U_A^\dagger O_{A_1} U_A$, where O_{A_1} is an operator on $\mathcal{H}_{A_1}$ with the same matrix elements as $\tilde{O}$ has on $\mathcal{H}_L$; that is

$$\langle \tilde{i}|\tilde{O}|\tilde{j}\rangle_{A\overline{A}} = \langle i|O_{A_1}|j\rangle_{A_1}$$

which is always possible since $|A_1| = |R| = |\mathcal{H}_L|$. (3.0.2) is then immediate.
(2) $\implies$ (3): This implication is by contradiction. Suppose there was some $X_{\overline{A}}$ such that $P_L V^\dagger X_{\overline{A}} P_L \propto P_L$. Now, Schur's lemma in this context states that the only non-trivial operators commuting with all other operators on $\mathcal{H}_L$ are scalar multiples of the identity. Since $V^\dagger X_{\overline{A}} V$ is not the identity, there must be some $\tilde{O}$ on $\mathcal{H}_L$ which doesn't commute with $V^\dagger X_{\overline{A}} V$, and some $|\tilde{\psi}\rangle \in \mathcal{H}_L$ such that:

$$\langle\tilde{\psi}|[P_L V^\dagger X_{\overline{A}} P_L, \tilde{O}]|\tilde{\psi}\rangle = \langle\tilde{\psi}|[V^\dagger X_{\overline{A}} V, \tilde{O}]|\tilde{\psi}\rangle \neq 0. \tag{3.0.5}$$

But such an $\tilde{O}$ cannot have a representation O_A on $\mathcal{H}_A$ as defined in 2, since this would by definition commute with $X_{\overline{A}}$; if it had such an O_A, then $\langle\tilde{\psi}|[V^\dagger X_{\overline{A}} V, \tilde{O}]|\tilde{\psi}\rangle = \langle\tilde{\psi}|[V^\dagger X_{\overline{A}} V, V^\dagger O_A V]|\tilde{\psi}\rangle = \langle\tilde{\psi}|V^\dagger[X_{\overline{A}}, O_A]V|\tilde{\psi}\rangle = 0$, which is a contradiction.
(3) $\implies$ (4): Consider arbitrary operators O_R on $\mathcal{H}_R$ and $X_{\overline{A}}$ on $\mathcal{H}_{\overline{A}}$. If we denote the constant of proportionality in (3.0.3) as $\lambda \in \mathbb{C}$, we have

$$P_{\text{code}} X_{\overline{A}} P_{\text{code}} = \lambda P_{\text{code}}, \tag{3.0.6}$$

so taking the inner product with $|\phi\rangle$:

$$\langle\phi|P_{\text{code}} X_{\overline{A}} P_{\text{code}}|\phi\rangle = \langle\phi|X_{\overline{A}}|\phi\rangle = \lambda \langle\phi|P_{\text{code}}|\phi\rangle = \lambda \langle\phi|\phi\rangle = \lambda, \tag{3.0.7}$$

so $\langle\phi|X_{\overline{A}}|\phi\rangle = \lambda$. But this implies

$$\begin{aligned} \langle\phi|X_{\overline{A}} O_R|\phi\rangle &= \langle\phi|P_{\text{code}} O_R X_{\overline{A}} P_{\text{code}}|\phi\rangle \\ &= \langle\phi|O_R P_{\text{code}} X_{\overline{A}} P_{\text{code}}|\phi\rangle \\ &= \langle\phi|O_R \lambda P_{\text{code}}|\phi\rangle \\ &= \langle\phi|O_R|\phi\rangle \langle\phi|X_{\overline{A}}|\phi\rangle \end{aligned} \tag{3.0.8}$$

since $P_{\text{code}}|\phi\rangle = |\phi\rangle$. Therefore, so long as $\langle\phi|O_R|\phi\rangle$ and $\langle\phi|X_{\overline{A}}|\phi\rangle$ are non-zero for any such O_R and $X_{\overline{A}}$, we have $\rho_{R\overline{A}}[\phi] = \rho_R[\phi] \otimes \rho_{\overline{A}}[\phi]$.
(4) $\implies$ (1): First, note that by definition, $|\phi\rangle$ is a purification of $\rho_{R\overline{A}}[\phi] = \rho_R[\phi] \otimes \rho_{\overline{A}}[\phi]$ on subsystem A. Also note that $|\phi\rangle$ maximally entangles R with A:

$$\rho_R[\phi] = \text{Tr}_{A\overline{A}}\left(\frac{1}{|R|}\sum_{ij}|i\rangle\langle j|_R \left(V|\tilde{i}\rangle\langle\tilde{j}|V^\dagger\right)_{A\overline{A}}\right) = \frac{1}{|R|}\sum_i |i\rangle\langle i|_R = \frac{I_R}{|R|} \tag{3.0.9}$$

so $|\phi\rangle$ maximally entangles R with A (or $\overline{A}$) since $\rho_R[\phi] = I/|R|$ is the maximally mixed state. This means that (3.0.4) becomes

$$\rho_{R\overline{A}}[\phi] = \frac{I_R}{|R|} \otimes \rho_{\overline{A}}[\phi]. \tag{3.0.10}$$

Next, we perform long division on A. Say k is the largest integer such that $|A| = k|R| + r$, and $r < |R|$. Since the R and $\overline{A}$ registers are separable in (3.0.10), we can indeed factorise $\mathcal{H}_A = (\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}) \oplus \mathcal{H}_{A_3}$ such that $|A_1| = |R|$, $|A_2| = k$, and $|A_3| = r$.
We now define the following two states:

$$|\Psi\rangle_{RA_1} \equiv \frac{1}{\sqrt{|R|}}\sum_i |i\rangle_R |i\rangle_{A_1}, \qquad |\chi\rangle_{A_2\overline{A}} \equiv \sum_j \sqrt{p_j}\,|j\rangle_{A_2}|j\rangle_{\overline{A}} \tag{3.0.11}$$

where $\sum_j p_j = 1$, and $\{|j\rangle_{A_2}\}$ and $\{|j\rangle_{\overline{A}}\}$ are bases of $\mathcal{H}_{A_2}$ and $\mathcal{H}_{\overline{A}}$ respectively. Note that the state

$$|\phi'\rangle \equiv |\Psi\rangle_{RA_1} \otimes |\chi\rangle_{A_2\overline{A}} \tag{3.0.12}$$

then purifies $\rho_{R\overline{A}}[\phi]$ on A_1A_2:

$$\begin{aligned} \mathrm{Tr}_{A_1A_2}\left(|\Psi\rangle\langle\Psi|_{RA_1} \otimes |\chi\rangle\langle\chi|_{A_2\overline{A}}\right) &= \mathrm{Tr}_{A_1}\left(|\Psi\rangle\langle\Psi|_{RA_1}\right)\mathrm{Tr}_{A_2}\left(|\chi\rangle\langle\chi|_{A_2\overline{A}}\right) \\ &= \rho_R[\phi] \otimes \rho_{\overline{A}}[\phi], \end{aligned} \tag{3.0.13}$$

since $|\Psi\rangle_{RA_1}$ purifies $\rho_R[\phi]$ on A_1, and $|\chi\rangle_{A_2\overline{A}}$ purifies $\rho_{\overline{A}}[\phi]$ on A_2. In a purification, the dimension of the purifying system A needs to be at least as big as the rank of the state being purified, so we therefore have $|A_1| = |R|$ (since $\rho_R[\phi]$ is maximally mixed), and $\mathrm{rank}\,(\rho_{\overline{A}}[\phi]) \leq |A_2|$.
However, purifications are unitarily equivalent on the purifying system - A in our case - so there exists unitary a U_A on $\mathcal{H}_A$ taking $|\phi\rangle = U_A|\phi'\rangle$. Overall, we therefore have:

$$\begin{aligned} (U_A \otimes I_{\overline{A}})\left(\frac{1}{\sqrt{|R|}}\sum_i |i\rangle_R\,(V|\tilde{i}\rangle)_{A\overline{A}}\right) &= \frac{1}{\sqrt{|R|}}\sum_i |i\rangle_R\,|i\rangle_{A_1} \otimes |\chi\rangle_{A_2\overline{A}} \\ \implies (U_A \otimes I_{\overline{A}})V|\tilde{i}\rangle_{A\overline{A}} &= |i\rangle_{A_1} \otimes |\chi\rangle_{A_2\overline{A}} \end{aligned} \tag{3.0.14}$$

as claimed. □

One important facet of this theorem is that it does not specify the full set of subsystems $\overline{A}$ which can be erased and still corrected. The three-qutrit example above coincidentally can correct for any single-qutrit erasure, but this may not always be true; we need to apply the theorem to each chosen $\overline{A}$ erasure and see if it works.

An RT Formula

From condition 1 of theorem 3.0.1, if the erasure of $\overline{A}$ can be corrected, then for any state $\tilde{\rho}$ on $\mathcal{H}_L$, we have

$$V\tilde{\rho}V^\dagger = U_A^\dagger(\rho_{A_1} \otimes |\chi\rangle\langle\chi|_{A_2\overline{A}})U_A, \tag{3.0.15}$$

where ρ_{A_1} is a density matrix on $\mathcal{H}_{A_1}$ with the same matrix elements as $\tilde{\rho}$ has on $\mathcal{H}_L$. Moreover, we have for the reduced density matrices:

$$\begin{aligned} \tilde{\rho}_A &\equiv \mathrm{Tr}_{\overline{A}}(V\tilde{\rho}V^\dagger) = U_A^\dagger\left(\rho_{A_1} \otimes \mathrm{Tr}_{\overline{A}}(|\chi\rangle\langle\chi|)\right)U_A \\ \tilde{\rho}_{\overline{A}} &\equiv \mathrm{Tr}_A(V\tilde{\rho}V^\dagger) = \mathrm{Tr}_{A_2}(|\chi\rangle\langle\chi|). \end{aligned} \tag{3.0.16}$$

Defining $\chi_{A_2} \equiv \mathrm{Tr}_{\overline{A}} |\chi\rangle \langle\chi|$ and $\chi_{\overline{A}} \equiv \mathrm{Tr}_{A_2} |\chi\rangle \langle\chi|$ and calculating the von Neumann entropies, we find

$$\begin{aligned} S(V\tilde{\rho}V^\dagger) &= S(\rho_{A_1}) + S(|\chi\rangle\langle\chi|) = S(\rho_{A_1}) \\ S(\tilde{\rho}_A) &= S(\rho_{A_1}) + S(\chi_{A_2}) = S(V\tilde{\rho}V^\dagger) + S(\chi_{A_2}) \\ S(\tilde{\rho}_{\overline{A}}) &= S(\chi_{\overline{A}}) = -\mathrm{Tr}_{A_2}\left[\mathrm{Tr}_{\overline{A}} |\chi\rangle\langle\chi|\right] = S(\chi_{A_2}). \end{aligned} \tag{3.0.17}$$

So, if we define 'area operator' $\mathcal{L} \equiv S(\chi_{A_2}) I_L$, we can rewrite these as

$$\begin{aligned} S(\tilde{\rho}_A) &= S(V\tilde{\rho}V^\dagger) + \mathrm{Tr}(\tilde{\rho}\mathcal{L}) \\ S(\tilde{\rho}_{\overline{A}}) &= S(\chi_{A_2}), \end{aligned} \tag{3.0.18}$$

which should be reminiscent of the RT formula!
From a holographic point of view though, this is not good. All of the 'bulk entropy' term $S(V\tilde{\rho}V^\dagger)$ appears in $S(\tilde{\rho}_A)$; it was not symmetric across the A and $\overline{A}$ boundary regions. We therefore need to generalise.

3.0.1 Subsystem Error Correction

A straightforward generalisation of theorem 3.0.1 is the so-called *subsystem error correction*, also of [1]. This theorem is as follows.

Theorem 3.0.2. *Let $V : \mathcal{H}_L \to \mathcal{H}$ be an encoding isometry, where $\mathcal{H}_L = \mathcal{H}_a \otimes \mathcal{H}_{\overline{a}}$ and $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\overline{A}}$. Define orthonormal bases $\{|\tilde{i}\rangle\}$ of $\mathcal{H}_a$ and $\{|\tilde{j}\rangle\}$ of $\mathcal{H}_{\overline{a}}$, and let $|\phi\rangle \equiv \frac{1}{\sqrt{|R||\overline{R}|}} \sum_{i,j} |i\rangle_R |j\rangle_{\overline{R}} V(|\widetilde{ij}\rangle)_{A\overline{A}}$ where R and $\overline{R}$ are auxiliary systems with $\mathcal{H}_R = \mathcal{H}_a$ and $\mathcal{H}_{\overline{R}} = \mathcal{H}_{\overline{a}}$. The following statements are then equivalent:*

1. *$|a| \leq |A$, and if decompose $\mathcal{H}_A = (\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}) \oplus \mathcal{H}_{A_3}$, where $|A_1| = |a|$, and $|A_3| \leq |a|$, then there exists a unitary transformation U_A on $\mathcal{H}_A$ and a set of orthonormal states $|\chi_j\rangle_{A_2\overline{A}} \in \mathcal{H}_{A_2\overline{A}}$ such that*

$$(U_A \otimes U_{\overline{A}}) V |\widetilde{ij}\rangle = |i\rangle_{A_1} \otimes |\chi_j\rangle_{A_2\overline{A}}, \tag{3.0.19}$$

   where $\{|i\rangle_{A_1}\}$ is an orthonormal basis of $\mathcal{H}_{A_1}$.

2. *For any operator $\tilde{O}_a$ acting within $\mathcal{H}_a$, there exists an operator O_A on $\mathcal{H}_A$ such that for any state $|\tilde{\psi}\rangle \in \mathcal{H}_{code}$, we have*

$$\begin{aligned} O_A V |\tilde{\psi}\rangle &= V\tilde{O}_a |\tilde{\psi}\rangle \\ O_A^\dagger V |\tilde{\psi}\rangle &= V\tilde{O}_a^\dagger |\tilde{\psi}\rangle \end{aligned}. \tag{3.0.20}$$

3. *For any operator $X_{\overline{A}}$ on $\mathcal{H}_A$, we have*

$$P_{code} V X_{\overline{A}} V^\dagger P_{code} = (I_a \otimes X_{\overline{a}}) P_{code}, \tag{3.0.21}$$

where P_{code} is the projector onto the image of V; that is, if $P_L = \sum_{i,j} |\widetilde{ij}\rangle \langle\widetilde{ij}|$ is the projector onto $\mathcal{H}_L$, then $P_{code} = V P_L V^\dagger$.

4. *In the state $|\phi\rangle$, we have*

$$\rho_{R\overline{RA}}[\phi] = \rho_R[\phi] \otimes \rho_{\overline{RA}}[\phi]. \tag{3.0.22}$$

The proof of this is virtually identical to that of 3.0.1, only that we must now keep track of the $\mathcal{H}_a$ subsystem. Moreover, it is a straightforward special case of a generalisation we go over in (???) - *operator algebra error correction* - so we do not go through the proof.

Complementary Recovery

In applying this to holography, we also wish to recover arbitrary operators $\tilde{O}_{\overline{a}}$ on $\mathcal{H}_{\overline{a}}$. [1] calls a code having this property a code with *complementary recovery*. Essentially, complementary recovery means that property 1 should also hold for the complementary systems; there should exist a unitary $U_{\overline{A}}$ on $\mathcal{H}_{\overline{A}}$, a decomposition $\mathcal{H}_{\overline{A}} = (\mathcal{H}_{\overline{A}_1} \otimes \mathcal{H}_{\overline{A}_2}) \oplus \mathcal{H}_{\overline{A}_3}$ with $|\overline{A}_1| = |\overline{a}|$ and $|\overline{A}_3 \leq |\overline{a}|$, and a set of orthonormal states $|\overline{\chi}_i\rangle_{\overline{A}_2 A}$ such that

$$(I_A \otimes U_{\overline{A}}) V |\widetilde{ij}\rangle = |j\rangle_{\overline{A}_1} \otimes |\overline{\chi}_i\rangle_{\overline{A}_2 A}\,, \tag{3.0.23}$$

where $\{|j\rangle_{\overline{A}_1}\}$ is an orthonormal basis of $\mathcal{H}_{\overline{A}_1}$. If we apply $U_A \otimes I_{\overline{A}}$ to this, and $I_A \otimes U_{\overline{A}}$ to (3.0.19), we get

$$\begin{aligned} (I_A \otimes U_{\overline{A}})(U_A \otimes I_{\overline{A}}) V |\widetilde{ij}\rangle &= |i\rangle_{A_1} \otimes (I_{A_2} \otimes U_{\overline{A}}) |\chi_j\rangle_{A_2 \overline{A}} \\ (U_A \otimes I_{\overline{A}_2})(I_A \otimes U_{\overline{A}}) V |\widetilde{ij}\rangle &= |j\rangle_{\overline{A}_1} \otimes (U_A \otimes I_{\overline{A}}) |\overline{\chi}_i\rangle_{\overline{A}_2 A}\,. \end{aligned} \tag{3.0.24}$$

In order for both of these to be true simultaneously, there must be a state $|\chi\rangle_{A_2 \overline{A}_2}$ such that $(I_{A_2} \otimes U_{\overline{A}}) |\chi_j\rangle_{A_2 \overline{A}} = |\chi\rangle_{A_2 \overline{A}_2} \otimes |j\rangle_{\overline{A}_1}$, which implies

$$(U_A \otimes U_{\overline{A}}) V |\widetilde{ij}\rangle = |i\rangle_{A_1} |j\rangle_{\overline{A}_1} |\chi\rangle_{A_2 \overline{A}_2}\,. \tag{3.0.25}$$

This expression is what allows us to prove an RT formula.

An RT Formula

Suppose we have a subsystem code with complementary recovery as above. Let $\tilde{\rho}$ be a state on $\mathcal{H}_L$; defining $\chi_{A_2} \equiv \mathrm{Tr}_{\overline{A}_2} |\chi\rangle\langle\chi|$ and $\chi_{\overline{A}_2} \equiv \mathrm{Tr}_{A_2} |\chi\rangle\langle\chi|$, we now have (analogously to (3.0.15) and (3.0.16)):

$$\begin{aligned} V\tilde{\rho}V^\dagger &= U_A^\dagger U_{\overline{A}}^\dagger(\rho_{A_1\overline{A}_1} \otimes |\chi\rangle\langle\chi|)U_A U_{\overline{A}} \\ \tilde{\rho}_A \equiv \mathrm{Tr}_{\overline{A}}(V\tilde{\rho}V^\dagger) &= U_A^\dagger(\rho_{A_1} \otimes \chi_{A_2})U_A \\ \tilde{\rho}_{\overline{A}} \equiv \mathrm{Tr}_A(V\tilde{\rho}V^\dagger) &= U_{\overline{A}}^\dagger(\rho_{\overline{A}_1} \otimes \chi_{\overline{A}_2})U_{\overline{A}}^\dagger, \end{aligned} \tag{3.0.26}$$

where $\rho_{A_1\overline{A}_1}$ has the same matrix elements on $\mathcal{H}_{\text{code}} = V\mathcal{H}_L V^\dagger$ as $\tilde{\rho}$ does on $\mathcal{H}_L$, and ρ_{A_1} and $\rho_{\overline{A}_1}$ have the same matrix elements as $\tilde{\rho}_a$ and $\tilde{\rho}_{\overline{a}}$ do on $\mathcal{H}_a$ and $\mathcal{H}_{\overline{a}}$ respectively. We can therefore define two area operators

$$\begin{aligned} \mathcal{L}_A &\equiv S(\chi_{A_2})I_a \\ \mathcal{L}_{\overline{A}} &\equiv S(\chi_{A_2})I_{\overline{a}}, \end{aligned} \tag{3.0.27}$$

which give us

$$\begin{aligned} S(\tilde{\rho}_A) &= \mathrm{Tr}(\tilde{\rho}_a\mathcal{L}_A) + S(\tilde{\rho}_a) \\ S(\tilde{\rho}_{\overline{A}}) &= \mathrm{Tr}(\tilde{\rho}_{\overline{a}}\mathcal{L}_{\overline{A}}) + S(\tilde{\rho}_{\overline{a}}). \end{aligned} \tag{3.0.28}$$

This gives us two RT formulae with the the information shared between the two systems! In fact, the converse is also true - if (3.0.28) both hold, then so too does condition 2 of the theorem. Since this is a special case of the next theorem anyway, we postpone the proof until then. +

Chapter 4

Generalisations of Holographic Error Correction

Theorem 3.0.1 can be generalised in multiple ways. Two of the more notable ways are that of *operator-algebra error correction* as presented in [1], and the generalisation for *non-isometric codes* as in (???). We present these generalisations here, starting with operator algebra error correction. To do this though, we first need some of the theory of finite-dimensional *von Neumann algebras*. We follow the structure of [1] in presenting the main applications here, with detailed results relegated to an appendix.

4.1 von Neumann Algebras

Definition 4.1.1 (Finite-dimensional von Neumann algebra)**.** A **von Neumann algebra** on a finite dimensional Hilbert space $\mathcal{H}$ is any set of linear operators $M \subseteq \mathcal{L}(\mathcal{H})$ satisfying:

- Contains all scalar multiples of the identity: $\forall \lambda \in \mathbb{C}$, $\lambda I \in M$, where I is the identity operator.
- Closure under Hermitian conjugation: $\forall x \in M$, $x^\dagger \in M$.
- Closure under multiplication: $\forall x, y \in M$, $xy \in M$.
- Closure under addition: $\forall x, y \in M$, $x + y \in M$.

Note the notation of operators written in lower case rather than upper case as is common. This is because we are treating the operators as elements of an algebra,

rather than individual operators in their own right (in some sense). Any von Neumann algebra induces two 'natural' associated algebras: the *commutant* and the *centre*.

Definition 4.1.2 (Commutant)**.** Given a von Neumann algebra M on $\mathcal{H}$, its **commutant**, denoted M' is the set of all operators on $\mathcal{H}$ which commute with M; that is

$$M' \equiv \{y \in \mathcal{L}(\mathcal{H}) \,|\, xy = yx, \forall x \in M\}. \tag{4.1.1}$$

Definition 4.1.3 (Centre)**.** Given a von Neumann algebra M on $\mathcal{H}$, its **centre**, denoted Z_M is the set of all operators on $\mathcal{H}$ in both M and M'; that is

$$Z_M \equiv M \cap M'. \tag{4.1.2}$$

In classifying von Neumann algebras, there is a special role for algebras which have a centre containing only scalar multiples of the identity. Such an algebra is called a *factor*.

Definition 4.1.4 (Factor algebra)**.** A von Neumann algebra M on $\mathcal{H}$ is called a **factor** if Z_M contains only scalar multiples of the identity.

In order to apply the theory of von Neumann algebras to error correction, we need two powerful classification theorems. We first classify factor algebras.

Theorem 4.1.1. *Suppose M is a factor on $\mathcal{H}$. Then there exists a tensor factorisation $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\overline{A}}$ such that $M = \mathcal{L}(\mathcal{H}_A) \otimes I_{\overline{A}}$ and $M' = I_A \otimes \mathcal{L}(\mathcal{H}_{\overline{A}})$.*

In other words, M is the set of all linear operators on the tensor factor $\mathcal{H}_A$ of $\mathcal{H}$. For general von Neumann algebras, this classification generalises:

Theorem 4.1.2. *Suppose M is a von Neumann algebra on $\mathcal{H}$. Then there exists a block-decomposition*

$$\mathcal{H} = \oplus_\alpha (\mathcal{H}_{A_\alpha} \otimes \mathcal{H}_{\overline{A}_\alpha}) \tag{4.1.3}$$

in terms of which M and M' are block-diagonal, with corresponding decompositions

$$M = \oplus_\alpha (\mathcal{L}(\mathcal{H}_{A_\alpha}) \otimes I_{\overline{A}_\alpha}), \quad M' = \oplus_\alpha (I_{A_\alpha} \otimes \mathcal{L}(\mathcal{H}_{\overline{A}_\alpha})). \tag{4.1.4}$$

So far this is all a bit abstract. With a view to linking this to error correction, suppose we have a von Neumann algebra M on $\mathcal{H}_{\text{code}}$. Then we have a decomposition

$$\mathcal{H}_{\text{code}} = \oplus_\alpha (\mathcal{H}_{a_\alpha} \otimes \mathcal{H}_{\bar{a}_\alpha}) \tag{4.1.5}$$

such that M is the set of all operators which are block diagonal in α, and acts as $\tilde{O}_{a_\alpha} \otimes I_{\bar{a}_\alpha}$ within each block, where $\tilde{O}_{a_\alpha}$ is an arbitrary linear operator on $\mathcal{H}_{a_\alpha}$. In matrix form, for some $\tilde{O} \in M$, we can write:

$$\tilde{O} = \begin{pmatrix} \tilde{O}_{a_1} \otimes I_{\bar{a}_1} & 0 & \dots \\ 0 & \tilde{O}_{a_2} \otimes I_{\bar{a}_2} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{4.1.6}$$

The commutant similarly consists of operators $\tilde{O}' \in M'$ which have matrix form:

$$\tilde{O}' = \begin{pmatrix} I_{a_1} \otimes \tilde{O}'_{\bar{a}_1} & 0 & \dots \\ 0 & I_{a_2} \otimes \tilde{O}'_{\bar{a}_2} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{4.1.7}$$

Also in matrix notation, the centre Z_M consists of operators $\tilde{\Lambda}$ of the form:

$$\tilde{\Lambda} = \begin{pmatrix} \lambda_1(I_{a_1} \otimes I_{\bar{a}_1}) & 0 & \dots \\ 0 & \lambda_2(I_{a_2} \otimes I_{\bar{a}_2}) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}, \tag{4.1.8}$$

where $\lambda_\alpha \in \mathbb{C}$. We can also introduce orthonormal bases for $\mathcal{H}_{a_\alpha}$ and $\mathcal{H}_{\bar{a}_\alpha}$ which are in some sense 'compatible' with the decompositions above. We denote these bases $\{|\widetilde{\alpha, i}\rangle\}$ and $\{|\widetilde{\alpha, j}\rangle\}$ respectively, and these induce an orthonormal basis for the full space $\mathcal{H}_{\text{code}}$, given by

$$\{|\widetilde{\alpha, ij}\rangle\} \equiv \{|\widetilde{\alpha, i}\rangle \otimes |\widetilde{\alpha, j}\rangle\}. \tag{4.1.9}$$

We can also extend the standard definition of von Neumann entropy to that of *von Neumann entropy on a von Neumann algebra*. Suppose we have a state $\tilde{\rho}$ with support on $\mathcal{H}_{\text{code}}$, which has diagonal blocks $\tilde{\rho}_{\alpha\alpha}$ with respect to the decomposition (4.1.5). We then define

$$p_\alpha \tilde{\rho}_{a_\alpha} \equiv \text{Tr}_{\bar{a}_\alpha} \tilde{\rho}_{\alpha\alpha}, \tag{4.1.10}$$

where $p_\alpha \in [0, 1]$ are probabilities such that $\text{Tr}_{a_\alpha} \tilde{\rho}_{a_\alpha} = 1$, and so $\sum_\alpha p_\alpha = 1$ as expected. We can then give the following definition for the entropy $S(\tilde{\rho}, M)$ of $\tilde{\rho}$ on M:

$$S(\tilde{\rho}, M) \equiv -\sum_\alpha p_\alpha \log p_\alpha + \sum_\alpha p_\alpha S(\tilde{\rho}_{a_\alpha}). \tag{4.1.11}$$

Note that this definition reduces to the normal von Neumann entropy $S(\tilde{\rho})$ when M is a factor.

We can analogously define entropy of $\tilde{\rho}$ on M'. Define

$$p_\alpha \tilde{\rho}_{\bar{a}_\alpha} \equiv \text{Tr}_{a_\alpha} \tilde{\rho}_{\alpha\alpha}, \tag{4.1.12}$$

and then

$$S(\tilde{\rho}, M') \equiv -\sum_{\alpha} p_\alpha \log p_\alpha + \sum_{\alpha} p_\alpha S(\tilde{\rho}_{\bar{a}_\alpha}). \tag{4.1.13}$$

Intuitively, we view these entropies as consisting of a classical piece, given by the Shannon entropy of the probabilities p_α for the centre Z_M, and a quantum piece given by the von Neumann entropy of each block, averaged over the probabilities.

4.2 Operator-Algebra Error Correction

We can now present theorem 5.1 of [1].

Theorem 4.2.1. *Let $V : \mathcal{H}_L \to \mathcal{H}$ be an encoding isometry with image $\mathcal{H}_{code} \subseteq \mathcal{H}$, where $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\overline{A}}$. Say we have a von Neumann algebra M on $\mathcal{H}_L$. Define orthonormal basis $\{|\widetilde{\alpha, ij}\rangle\}$ of $\mathcal{H}_L$ as in (4.1.9), which is compatible with the decomposition $\mathcal{H}_L = \oplus_\alpha(\mathcal{H}_{a_\alpha} \otimes \mathcal{H}_{\bar{a}_\alpha})$ induced by M. Let $|\phi\rangle \equiv \frac{1}{\sqrt{|R|}} \sum_{\alpha,i,j} |\alpha, ij\rangle_R (V |\widetilde{\alpha, ij}\rangle)_{A\overline{A}}$, where R is an auxiliary system with $\mathcal{H}_R = \mathcal{H}_L$. The following statements are then equivalent:*

1. *$\sum_\alpha |a_\alpha| \leq |A|$, and we can decompose $\mathcal{H}_A = \oplus_\alpha(\mathcal{H}_{A_1^\alpha} \otimes \mathcal{H}_{A_2^\alpha}) \oplus \mathcal{H}_{A_3}$ with $|A_1^\alpha| = |a_\alpha|$ such that there exists a unitary transformation U_A on $\mathcal{H}_A$ and sets of orthonormal states $|\chi_{\alpha,j}\rangle_{A_2^\alpha \overline{A}} \in \mathcal{H}_{A_2^\alpha \overline{A}}$ such that*

$$(U_A \otimes I_{\overline{A}}) V |\widetilde{\alpha, ij}\rangle = |\alpha, i\rangle_{A_1^\alpha} \otimes |\chi_{\alpha,j}\rangle_{A_2^\alpha \overline{A}}, \tag{4.2.1}$$

   where $\{|\alpha, i\rangle_{A_1^\alpha}\}$ is an orthonormal basis for $\mathcal{H}_{A_1^\alpha}$.

2. *For any operator $\tilde{O} \in M$, there exists an operator O_A on $\mathcal{H}_A$ such that for any state $|\tilde{\psi}\rangle \in \mathcal{H}_L$, we have*

$$\begin{aligned} O_A V |\tilde{\psi}\rangle &= V\tilde{O} |\tilde{\psi}\rangle \\ O_A^\dagger V |\tilde{\psi}\rangle &= V\tilde{O}^\dagger |\tilde{\psi}\rangle . \end{aligned} \tag{4.2.2}$$

3. *For any operator $X_{\overline{A}}$ on $\mathcal{H}_{\overline{A}}$, we have*

$$P_{code} X_{\overline{A}} P_{code} = V X' V^\dagger P_{code}, \tag{4.2.3}$$

   where $X' \in M'$ is an element of the commutant, and P_{code} is the image of the projector onto $\mathcal{H}_L$ under V (or the projector onto $\mathcal{H}_{code}$); that is, if $P_L = \sum_{\alpha,i,j} |\widetilde{\alpha, i, j}\rangle \langle\widetilde{\alpha, i, j}|$, then $P_{code} = V P_L V^\dagger$.

4. For any operator $\tilde{O} \in M$, we have

$$[O_R, \rho_{R\overline{A}}[\phi]] = 0, \tag{4.2.4}$$

where O_R is the unique operator on $\mathcal{H}_R$ such that

$$\begin{aligned} O_R |\phi\rangle &= V\tilde{O}V^\dagger |\phi\rangle \\ O_R^\dagger |\phi\rangle &= V\tilde{O}^\dagger V^\dagger |\phi\rangle . \end{aligned} \tag{4.2.5}$$

Similar to the last theorem, this characterises in some sense 'how well' a code subspace can correct a subalgebra M for the erasure of $\overline{A}$. It in fact contains the previous theorem as a special case, reducing to it when M is the full set of linear operators on $\mathcal{H}_{\text{code}}$.

Proof. (1) $\implies$ (2): Define $O_A \equiv U_A^\dagger(\oplus_\alpha(O_{A_1^\alpha} \otimes I_{A_2^\alpha}))U_A$, where $O_{A_1^\alpha}$ is an operator acting on $\mathcal{H}_{A_1^\alpha}$ in the same way as $\tilde{O}_{a_\alpha}$ from (4.1.6) does on $\mathcal{H}_{a_\alpha}$. (4.2.3) is then immediate.
(2) $\implies$ (3): This implication is by contradiction. Suppose that $P_{\text{code}} X_{\overline{A}} P_{\text{code}} = V x' V^\dagger P_{\text{code}}$, where $x' \in \mathcal{L}(\mathcal{H}_L)$ but $x' \notin M'$. Therefore there must be some operator $\tilde{O} \in M$ which does not commute with x', and so there must be a state $|\tilde{\psi}\rangle \in \mathcal{H}_L$ such that

$$\langle\tilde{\psi}|[x', \tilde{O}]|\tilde{\psi}\rangle = \langle\tilde{\psi}|[V^\dagger P_{\text{code}} X_{\overline{A}} P_{\text{code}} V, \tilde{O}]|\tilde{\psi}\rangle = \langle\tilde{\psi}|V^\dagger[X_{\overline{A}}, V\tilde{O}V^\dagger]V|\tilde{\psi}\rangle \neq 0. \tag{4.2.6}$$

However, such an $\tilde{O}$ cannot have a corresponding O_A as this would automatically commute with $X_{\overline{A}}$, which contradicts (2).
(3) $\implies$ (4): Say $\tilde{O} \in M$, and $X_{\overline{A}}$ and Y_R are arbitrary operators on $\mathcal{H}_{\overline{A}}$ and $\mathcal{H}_R$ respectively. We then have:

$$\begin{aligned} \text{Tr}_{R\overline{A}}(O_R \rho_{R\overline{A}}[\phi] X_{\overline{A}} Y_R) &= \langle\phi|X_{\overline{A}} Y_R O_R|\phi\rangle \\ &= \langle\phi|X_{\overline{A}} Y_R V\tilde{O}V^\dagger|\phi\rangle \\ &= \langle\phi|V\tilde{O}V^\dagger X_{\overline{A}} Y_R|\phi\rangle \\ &= \langle\phi|O_R X_{\overline{A}} Y_R|\phi\rangle \\ &= \text{Tr}_{R\overline{A}}(\rho_{R\overline{A}}[\phi] O_R X_{\overline{A}} Y_R), \end{aligned} \tag{4.2.7}$$

where the first equality is by substituting in the definition of $\rho[\phi]$ and expanding, the second is by definition of O_R, the third is due to $V\tilde{O}V^\dagger$ commuting with Y_R trivially and with $X_{\overline{A}}$ by (3), and the last two by similar logic in reverse. This can only hold for arbitrary $X_{\overline{A}}$ and Y_R if $[O_R, \rho_{R\overline{A}}[\phi]] = 0$ as claimed.
(4) $\implies$ (1): Our basis $\{|\widetilde{\alpha, ij}\rangle_R\}$ for $\mathcal{H}_R$ gives a decomposition

$$\mathcal{H}_R = \oplus_\alpha(\mathcal{H}_{R_\alpha} \otimes \mathcal{H}_{\overline{R}_\alpha}) \tag{4.2.8}$$

and so $\mathcal{H}_{R\overline{A}} = \mathcal{H}_R \otimes \mathcal{H}_{\overline{A}}$ can be decomposed as

$$\mathcal{H}_{R\overline{A}} = \oplus_\alpha (\mathcal{H}_{R_\alpha} \otimes \mathcal{H}_{\overline{R}_\alpha} \otimes \mathcal{H}_{\overline{A}}). \tag{4.2.9}$$

From (4), we know that $[O_R, \rho_{R\overline{A}}[\phi]] = 0$ for all O_R as defined in (4.2.6). This means that $\rho_R[\phi] = I_R/|R|$ is the maximally mixed state on R. We therefore have that, in terms of the decomposition (4.2.9)

$$\rho_{R\overline{A}}[\phi] = \oplus_\alpha \left[\frac{|R_\alpha||\overline{R}_\alpha|}{|R|} \left(\frac{I_{R_\alpha}}{|R_\alpha|} \otimes \rho_{\overline{R}_\alpha \overline{A}} \right) \right], \tag{4.2.10}$$

for some states $\rho_{\overline{R}_\alpha \overline{A}}$. The coefficient out the front can be computed by requiring that $\rho_{R\overline{A}}[\phi]$ is a valid density operator tracing to 1. Since $\rho_R[\phi] = I_R/|R|$, we must have also that $\mathrm{Tr}_{\overline{A}}(\rho_{\overline{R}_\alpha \overline{A}}) = I_{\overline{R}_\alpha}/|\overline{R}_\alpha|$.
By definition, $|\phi\rangle_{RA\overline{A}}$ is a purification of $\rho_{R\overline{A}}[\phi]$ on A, and in a purification the dimension of the purifying system is necessarily as big as the rank of the state being purified (this is immediate from the Schmidt decomposition). So, denoting $\mathrm{rank}(\rho_{\overline{R}_\alpha \overline{A}}) \equiv |\rho_{\overline{R}_\alpha \overline{A}}|$, we can write

$$\sum_\alpha |R_\alpha||\rho_{\overline{R}_\alpha \overline{A}}| \leq |A|. \tag{4.2.11}$$

This means that we can indeed decompose

$$\mathcal{H}_A = \oplus_\alpha (\mathcal{H}_{A_1^\alpha} \otimes \mathcal{H}_{A_2^\alpha}) \oplus \mathcal{H}_{A_3} \tag{4.2.12}$$

where $|A_1^\alpha| = |R_\alpha| = |a_\alpha|$ and $|A_2^\alpha| \geq |\rho_{\overline{R}_\alpha \overline{A}}|$ by long division. For each α, we can then purify $\rho_{\overline{R}_\alpha \overline{A}}$ on A_2^α; since $\mathrm{Tr}_{\overline{A}}(\rho_{\overline{R}_\alpha \overline{A}}) = I_{\overline{R}_\alpha}/|\overline{R}_\alpha|$, such a purification has the form

$$|\psi_\alpha\rangle_{\overline{R}_\alpha A_2^\alpha \overline{A}} = \frac{1}{\sqrt{|\overline{R}_\alpha|}} \sum_j |\alpha, j\rangle_{\overline{R}_\alpha} |\chi_{\alpha,j}\rangle_{A_2^\alpha \overline{A}}, \tag{4.2.13}$$

where the $|\chi_{\alpha,j}\rangle_{A_2^\alpha \overline{A}}$ are mutually orthonormal on $A_2^\alpha \overline{A}$. This means we can write a purification for $\rho_{R\overline{A}}$ on the full A system as

$$\begin{aligned} |\phi'\rangle &= \sum_{\alpha,i,j} \frac{1}{\sqrt{|R_\alpha|}} |\alpha, i\rangle_{R_\alpha} |\alpha, i\rangle_{A_1^\alpha} |\psi_\alpha\rangle_{\overline{R}_\alpha A_2^\alpha \overline{A}} \\ &= \frac{1}{\sqrt{|R|}} \sum_{\alpha,i,j} |\alpha, ij\rangle_R |\alpha, i\rangle_{A_1^\alpha} |\chi_{\alpha,j}\rangle_{A_2^\alpha \overline{A}}. \end{aligned} \tag{4.2.14}$$

Finally, since $|\phi\rangle$ and $|\phi'\rangle$ are two different purifications of $\rho_{R\overline{A}}[\phi]$ on A, they must

differ by the action of some unitary U_A. We therefore have

$$(U_A \otimes I_{\overline{A}}) \left(\frac{1}{\sqrt{|R|}} \sum_{\alpha,i,j} |\alpha, ij\rangle_R \, (V \, |\widetilde{\alpha, ij}\rangle)_{A\overline{A}} \right) = \frac{1}{\sqrt{|R|}} \sum_{\alpha,i,j} |\alpha, ij\rangle_R \, |\alpha, i\rangle_{A_1^\alpha} \, |\chi_{\alpha,j}\rangle_{A_2^\alpha \overline{A}}$$
$$\implies (U_A \otimes I_{\overline{A}}) V \, |\widetilde{\alpha, ij}\rangle = |\alpha, i\rangle_{A_1^\alpha} \otimes |\chi_{\alpha,j}\rangle_{A_2^\alpha \overline{A}} \,, \tag{4.2.15}$$

which finishes the proof. □

4.3 Approximate and Non-Isometric Codes

So far, all our theorems have described cases where the error correcting code can perfectly recover information. In real world experiments and quantum computers, this is never truly the case, so we need to be able to talk about *approximate error correction*. In this framework, there are analogous theorems to theorems 3.0.1 and 3.0.2, with the conditions replaced by approximate versions. For example, an approximate version of (3.0.15) is

$$\|\mathrm{Tr}_{A_2\overline{A}}(U_A V \tilde{\rho} V^\dagger) - \tilde{\rho}_{A_1}\|_1 \leq \epsilon \tag{4.3.1}$$

for any $\epsilon > 0$. Here, $\|X\|_1 = \mathrm{Tr}\sqrt{X^\dagger X}$ is called the *trace norm*.
Approximate error correction behaves quite differently to exact error correction. For example, exact error correction is always impossible if the number of errors is greater than 1/4 the number of qubits used for encoding, but approximate error correction is still possible as long as the number of errors is less than 1/2 the number of qubits (CITE).
For erasures, to quantify the difference somewhat, rather than considering the information which is accessible to $\mathcal{H}_A$, consider the information which is inaccessible to $\mathcal{H}_{\overline{A}}$. Since the no-cloning and no-deletion theorems are true in quantum computing, these are exactly equivalent - i.e. exact erasure correction is only possible if $\mathcal{H}_A$ has access to *all* information about the encoded state, or if $\mathcal{H}_{\overline{A}}$ has access to *no* information about it. More concretely, if we trace out the A subsystem from any encoded state $V \, |\tilde{\psi}\rangle \in \mathcal{H}_{\text{code}}$, the resultant state should contain no information about $|\tilde{\psi}\rangle$; for all states $|\tilde{\psi}\rangle \in \mathcal{H}_L$, we have

$$\mathrm{Tr}_A(V \, |\tilde{\psi}\rangle \, \langle\tilde{\psi}| \, V^\dagger) = \omega_{\overline{A}} \tag{4.3.2}$$

for some *fixed* state $\omega_{\overline{A}}$ which is completely independent of $|\tilde{\psi}\rangle$. While this holds for exact erasure correction, for approximate erasure correction we require that for all $|\tilde{\psi}\rangle \in \mathcal{H}_L$

$$\|\mathrm{Tr}_A(V \, |\tilde{\psi}\rangle \, \langle\tilde{\psi}| \, V^\dagger) - \omega_{\overline{A}}\|_1 \leq \epsilon. \tag{4.3.3}$$

Chapter 5

Conclusions

This is the place to put your conclusions about your work. You can split it into different sections if appropriate. You may want to include a section of future work which could be carried out to continue your research.

The conclusion section should be at least one page long, preferably 2 pages, but not much longer.

Appendix A

Quantum Mechanics and Information: Background and Conventions

In this appendix, I outline the background and conventions used for quantum mechanics and quantum information theory throughout this dissertation in detail. I mainly stick to the description of quantum mechanics as given in Nielsen and Chuang's textbook [2] and the notes of David Skinner, and supplement Nielsen and Chuang's description of quantum information and computing with notes by Richard Jozsa and John Preskill.

A.1 Quantum Mechanics

I begin with a discussion of the four defining postulates of quantum mechanics.

A.1.1 State Spaces

Postulate 1. Any isolated quantum system has an associated Hilbert space $\mathcal{H}$ called the *state space* of the system. The system is fully described by its *state vector* (or just *state*), which is a unit vector $|\psi\rangle \in \mathcal{H}$.

Note that this just tells us that the state space for a system exists, and not necessarily what the relevant Hilbert space is. The simplest example of a non-trivial system is that of the two-dimensional *qubit*, with state space denoted $\mathcal{H}_2$. An orthonormal basis of this space is given by the set $\{|0\rangle, |1\rangle\}$, so the state of a qubit

can be written

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle\,, \quad a, b \in \mathbb{C} \tag{A.1.1}$$

with the condition that $|\psi\rangle$ is a unit vector (i.e. $\langle\psi, \psi\rangle = 1$) implying that $|a|^2 + |b|^2 = 1$. We say that $|\psi\rangle$ is in a *superposition* of $|0\rangle$ and $|1\rangle$, with *amplitudes* a and b, which generalises in the obvious way to arbitrary linear combinations of states.

A.1.2 Evolution of States

Postulate 2. The evolution in time of a *closed* quantum system is described by a unitary transformation. Explicitly, if $|\psi, t\rangle$ is the state of a system at time t, and $|\psi', t'\rangle$ is the state at time t', then the two states are related by some unitary operator $U(t, t')$ depending only on times t and t' such that

$$|\psi', t'\rangle = U(t, t')\,|\psi, t\rangle \tag{A.1.2}$$

Again, this does not tell us which unitary operators describe the dynamics of a real-world system. On a simple system such as a qubit, it turns out that any unitary operator can be realised. For example, the Pauli matrix X which takes $|0\rangle \mapsto |1\rangle$ and $|1\rangle \mapsto |0\rangle$ is a unitary operator on a single qubit which can always be physically implemented. Note that this postulate generalises to the case of continuous time, whereupon evolution is governed by the Schrodinger equation, but this is less relevant for the case of quantum information.

A.1.3 Measurements

Postulate 3. A measurement of a quantum system is described by a set $\{M_m\}$ of *measurement operators*, which act on the state space of the system being measured. The index m labels the possible outcomes of the measurement. If the system is in state $|\psi\rangle$ immediately before the measurement is taken, the probability the measurement returns outcome m is given by

$$\mathbb{P}(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle \tag{A.1.3}$$

and the state of the system immediately after the measurement is

$$\frac{M_m\,|\psi\rangle}{\sqrt{\mathbb{P}(m)}}. \tag{A.1.4}$$

Since probabilities sum to 1, we can also derive that the measurement operators obey the *completeness relation*:

$$1 = \sum_m \mathbb{P}(m) = \langle\psi|\sum_m M_m^\dagger M_m|\psi\rangle\,,\ \forall\,|\psi\rangle \iff \sum_m M_m^\dagger M_m = I \tag{A.1.5}$$

which is a necessary condition for the measurement operators to obey.

Projective Measurements

In quantum information, a particularly important class of measurements are the *projective measurements*. These turn out to be exactly equivalent to the more general measurement postulate when augmented with the ability to perform unitary transformations as in postulate 2.

Definition A.1.1 (Projective measurement)**.** A projective measurement is described by a Hermitian operator (or *observable*) M acting on the state space of the system being measured. Since M is Hermitian, it has a spectral decomposition

$$M = \sum_m m P_m \tag{A.1.6}$$

where m indexes the eigenvalues, and P_m is the corresponding projector onto the m-eigenspace. The possible outcomes of the measurement correspond to the eigenvalues.

With this definition, we see that the probability of obtaining outcome m on measuring state $|\psi\rangle$ is

$$\mathbb{P}(m) = \langle\psi|P_m|\psi\rangle \tag{A.1.7}$$

and the state immediately after the measurement is

$$\frac{P_m|\psi\rangle}{\sqrt{\mathbb{P}(m)}}. \tag{A.1.8}$$

Projective measurements have all sorts of nice properties which general measurements do not. One notable such property is how *expectation values* in a probabilistic sense simplify. If we measure observable M on $|\psi\rangle$, the average outcome is

$$\begin{aligned}\mathbb{E}_\psi(M) &\equiv \sum_m m\mathbb{P}_\psi(m) \\ &= \sum_m m\,\langle\psi|P_m|\psi\rangle \\ &= \langle\psi|\sum_m mP_m|\psi\rangle \\ &= \langle\psi|M|\psi\rangle\end{aligned} \tag{A.1.9}$$

which can simplify many calculations. This is often denoted $\langle M\rangle_\psi \equiv \langle\psi|M|\psi\rangle$.
In this dissertation, I use the convention where I define a projective measurement

by just listing the set of projectors $\{P_m\}$ rather than the observable M, where it is implicit that $M = \sum_m mP_m$, $\sum_m P_m = 1$, and $P_m P_n = \delta_{mn} P_m$. I will also say '*perform a measurement in the* $|m\rangle$ *basis*', which refers to projective measurement with projectors $P_m = |m\rangle \langle m|$.

POVM measurements

In quantum mechanics, the post-measurement state is often not particularly relevant; the important item being the outcome probabilities instead, for example where a quantum circuit only performs a measurement at the end of the circuit. *POVM measurements* (standing for 'positive operator valued measure') are particularly well-suited for this sort of application. Suppose we perform a measurement $\{M_m\}$ on a state $|\psi\rangle$, so $\mathbb{P}(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$. If we then define operators

$$E_m \equiv M_m^\dagger M_m, \tag{A.1.10}$$

then the E_m obey $\sum_m E_m = I$ and $\mathbb{P}(m) = \langle\psi|E_m|\psi\rangle$, and so the set $\{E_m\}$ alone is sufficient to determine the outcome probabilities. The operators E_m are called *POVM elements* for the measurement, and the set $\{E_m\}$ is simply called a POVM. It can be shown that any measurement where the measurement operators and the POVM operators are the same is a projective measurement, and moreover that for any POVM $\{E_m\}$, there exists a set of measurement operators $\{M_m\}$ describing an equivalent general measurement.

A.1.4 Composite Systems

Postulate 4. The state space of a composite system AB is the *tensor product* of the state spaces of the individual systems $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. More generally, if we have n systems indexed by i each prepared in the state $|\psi_i\rangle$, the joint state of the total system is $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_n\rangle$.

In practice, we often drop the tensor product symbol $\otimes$, and just write (for example) $|\psi\rangle_A \otimes |\phi\rangle_B \equiv |\psi\rangle_A |\phi\rangle_B$, where the subscript keeps track of which system each state refers to. We can also talk about the composition of operators. For example, if O_A and O_B are operators acting on systems A and B individually, the composition $O_A \otimes O_B$ acts on the joint state $|A\rangle |B\rangle$ as

$$(O_A \otimes O_B) |A\rangle |B\rangle \equiv O_A O_B |A\rangle |B\rangle = (O_A |A\rangle) \otimes (O_B |B\rangle). \tag{A.1.11}$$

We often wish to express operators and states explicitly in a basis. If $\{|a\rangle\}_{a=1}^n$ is an orthonormal basis for system A, then any operator O_A on $\mathcal{H}_A$ can be expressed

as

$$O_A \equiv \sum_{ab} O_{ab} |a\rangle \langle b| = \begin{pmatrix} O_{11} & \dots & O_{1n} \\ \vdots & \ddots & \vdots \\ O_{n1} & \dots & O_{nn} \end{pmatrix}, \quad O_{ab} \in \mathbb{C}. \tag{A.1.12}$$

The composite operator $O_A \otimes O_B$ for some operator O_B on system B with state space $\mathcal{H}_B$ is then the block matrix

$$O_A \otimes O_B \equiv \begin{pmatrix} O_{11}O_B & \dots & O_{1n}O_B \\ \vdots & \ddots & \vdots \\ O_{n1}O_B & \dots & O_{nn}O_B \end{pmatrix}. \tag{A.1.13}$$

Note that this representation of an operator can be used to represent states as a specific case.

A.1.5 Density Operators

For some cases, the description of a state as being just a unit vector $|\psi\rangle$ is insufficient. This motivates the formalism of *density operators*, which find particular use in talking about subsystems of a composite system, systems which we may not know the state with certainty, and in statistical mechanics. Consider a system which is in the state $|\psi_i\rangle$ with probability p_i. The complete set $\{|\psi_i\rangle, p_i\}$ is known as an *ensemble* of pure states. The density operator for this system is then defined as

$$\rho \equiv \sum_i p_i |\psi_i\rangle \langle\psi_i| . \tag{A.1.14}$$

We can reformulate the postulates above in terms of density operators; we use them so frequently, that we do this explicitly.
For postulate 1, we note that the states $|\psi_i\rangle$ define a unique ρ and vice versa, so this remains unchanged. Similarly, the fourth postulate also remains unchanged since it does not explicitly refer to states in its statement. Consider postulate 2, and suppose we have a unitary U governing evolution of the system. ρ then evolves as

$$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| \xrightarrow{U} \sum_i p_i U |\psi_i\rangle \langle\psi_i| U^\dagger = U\rho U^\dagger. \tag{A.1.15}$$

We can similarly reformulate the measurement postulate 3 for density operators. Suppose we perform a measurement with operators $\{M_m\}$. If our system happened to initially be in state $|\psi_i\rangle$, then the probability of outcome m is

$$\mathbb{P}(m|i) = \langle\psi_i|M_m^\dagger M_m|\psi_i\rangle = \mathrm{Tr}(M_m^\dagger M_m |\psi_i\rangle \langle\psi_i|) \tag{A.1.16}$$

where the trace is taken over the entire state space. Therefore, the full probability of outcome m is

$$\mathbb{P}(m) = \sum_i \mathbb{P}(m|i)p_i = \sum_i p_i \mathrm{Tr}(M_m^\dagger M_m \left|\psi_i\right\rangle \left\langle\psi_i\right|) = \mathrm{Tr}(M_m^\dagger M_m \rho). \tag{A.1.17}$$

What is the density operator ρ immediately after the measurement returning m? If the initial state was $|\psi_i\rangle$, then we know that immediately after the measurement it is

$$\left|\psi_i^m\right\rangle = \frac{M_m \left|\psi_i\right\rangle}{\sqrt{\mathbb{P}(m|i)}}, \tag{A.1.18}$$

and so we now have an ensemble of states $\{\left|\psi_i^m\right\rangle, \mathbb{P}(i|m)\}$ immediately post-measurement. Therefore, ρ becomes

$$\rho_m \equiv \sum_i \mathbb{P}(i|m) \left|\psi_i^m\right\rangle \left\langle\psi_i^m\right|. \tag{A.1.19}$$

Probability theory states that $\mathbb{P}(i|m) = \mathbb{P}(m|i)p_i/\mathbb{P}(m)$, so we therefore find

$$\rho_m = \sum_i p_i \frac{M_m \left|\psi_i\right\rangle \left\langle\psi_i\right| M_m^\dagger}{\mathrm{Tr}(M_m^\dagger \rho)} = \frac{M_m \rho M_m^\dagger}{\mathrm{Tr}(M_m^\dagger M_m \rho)}, \tag{A.1.20}$$

which completes the reformulation.
Density operators can also be axiomatised. Indeed, we can show that an arbitrary operator ρ is a density operator for some system if and only if both:

1. $\mathrm{Tr}(\rho) = 1$
2. ρ is a positive operator.

Before concluding, I'll introduce some nomenclature which is used to describe states quite often. A system which is known to be in state $|\psi\rangle$ exactly is called a *pure state*, and the density operator is $\rho = \left|\psi\right\rangle \left\langle\psi\right|$. If not, the state is said to be *mixed*, and the density operator can be written as in (A.1.14). A sufficient criterion for checking whether a state is pure or mixed is calculating $\mathrm{Tr}(\rho^2)$; if this is 1, then the state is pure, and if it is less than 1, it is mixed.

A.2 Quantum Information and Computation

This section will go through the basics of quantum information and computation. We discuss the definitions and conventions used for various distance measures for quantum information (with relevance to approximate error correction), entropy, as well as the basics of the circuit model of quantum computation.

A.2.1 Distance Measures

The point of quantum distance measures is to answer the question 'how close are two states?'. The two main distance measures are *trace distance* and *fidelity*. For our purposes, only fidelity is important, so we run through its definition and properties.

Definition A.2.1 (Fidelity)**.** Given two states (as density operators) ρ and σ, the **fidelity** between them is

$$F(\rho, \sigma) \equiv \mathrm{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}, \tag{A.2.1}$$

where $(\rho^{1/2})^2 = \rho$.

While fidelity is not formally a metric, it does give a useful notion of distance between states. Before going through its properties, we note two cases where fidelity simplifies. The first is when ρ and σ commute, so can be simultaneously diagonalised; in some basis $\{|i\rangle\}$ of the underlying Hilbert space, we have

$$\rho = \sum_i r_i \left|i\right\rangle \left\langle i\right|, \quad \sigma = \sum_i s_i \left|i\right\rangle \left\langle i\right|. \tag{A.2.2}$$

Then, we can compute the fidelity as

$$\begin{aligned} F(\rho, \sigma) &= \mathrm{Tr}\sqrt{\sum_i r_i s_i \left|i\right\rangle \left\langle i\right|} \\ &= \mathrm{Tr}\left(\sum_i \sqrt{r_i s_i} \left|i\right\rangle \left\langle i\right|\right) \\ &= \sum_i \sqrt{r_i s_i}. \end{aligned} \tag{A.2.3}$$

The second special case is the fidelity between a pure state $|\psi\rangle$ and arbitrary ρ. In this case, we have

$$F(\left|\psi\right\rangle, \rho) = \mathrm{Tr}\sqrt{\left|\psi\right\rangle \left\langle\psi|\rho|\psi\right\rangle \left\langle\psi\right|} = \sqrt{\langle\psi|\rho|\psi\rangle}, \tag{A.2.4}$$

so fidelity is just the square root of the overlap between $|\psi\rangle$ and ρ.

Properties

We now present some useful properties of fidelity.

Proposition 1. *Fidelity is invariant under unitary transformations; that is, for unitary U, we have*

$$F(U\rho U^\dagger, U\sigma U^\dagger) = F(\rho, \sigma). \tag{A.2.5}$$

Proof. For any positive operator A, we know that $\sqrt{UAU^\dagger} = U\sqrt{A}U^\dagger$. Density operators are positive, so

$$\begin{aligned} F(U\rho U^\dagger, U\sigma U^\dagger) &= \mathrm{Tr}\sqrt{(U\rho U^\dagger)^{1/2} U\sigma U^\dagger (U\rho U^\dagger)^{1/2}} \\ &= \mathrm{Tr}\sqrt{U\rho^{1/2}U^\dagger U\sigma U^\dagger U\rho^{1/2}U^\dagger} \\ &= \mathrm{Tr}\sqrt{U\rho^{1/2}\sigma\rho^{1/2}U^\dagger} \\ &= \mathrm{Tr}U\sqrt{\rho^{1/2}\sigma\rho^{1/2}}U^\dagger \\ &= F(\rho, \sigma). \end{aligned} \tag{A.2.6}$$

□

Fidelity can be characterised by a theorem called *Uhlmann's theorem*. Before presenting and proving this, we need two basic lemmas.

Lemma A.2.1. *Let A be an arbitrary operator, and U a unitary operator. Then*

$$|\mathit{Tr}(AU)| \leq \mathit{Tr}|A|, \tag{A.2.7}$$

with equality when $U = V^\dagger$, where $A = |A|V$ is the polar decomposition of A.

Proof. Equality is clear under the stated condition. We compute

$$|\mathrm{Tr}(AU)| = |\mathrm{Tr}(|A|VU)| = |\mathrm{Tr}(|A|^{1/2}|A|^{1/2}VU)|, \tag{A.2.8}$$

where we note that the trace term is just the Hilbert-Schmidt inner product $\langle |A|^{1/2}, |A|^{1/2}VU\rangle$. The Cauchy-Schwarz inequality then implies

$$|\mathrm{Tr}(AU)| \leq \sqrt{\mathrm{Tr}|A|\mathrm{Tr}(U^\dagger V^\dagger |A| VU)} = \mathrm{Tr}|A| \tag{A.2.9}$$

as required. □

Lemma A.2.2. *Suppose A and R are two quantum systems with the same Hilbert space, with bases $\{|i_A\rangle\}$ and $\{|i_R\rangle\}$ respectively. Define $|m\rangle = \sum_i |i_R\rangle |i_A\rangle$, and let P be an arbitrary operator on A and Q on R. Then*

$$\mathit{Tr}(Q^\dagger P) = \langle m|Q \otimes P|m\rangle \tag{A.2.10}$$

where the left hand side refers to matrix multiplication, and the matrix elements are taken with respect to the given bases.

Proof. We just compute:

$$\begin{aligned}\mathrm{Tr}(Q^\dagger P) &= \langle i_R|Q|j_R\rangle \langle i_A|P|j_A\rangle \\ &= \langle i_R| \langle i_A| (Q \otimes P) |j_R\rangle |j_A\rangle = \langle m|Q \otimes P|m\rangle ,\end{aligned} \tag{A.2.11}$$

as required, where summation convention is implied. □

Theorem A.2.1 (Uhlmann's Theorem)**.** *Suppose ρ and σ are states of quantum system A. Introduce an auxiliary system R with $\mathcal{H}_R = \mathcal{H}_A$. Then:*

$$F(\rho, \sigma) = \max_{|\psi\rangle, |\phi\rangle} |\langle \psi|\phi\rangle |, \tag{A.2.12}$$

where the maximisation is taken over all purifications $|\psi\rangle$ of ρ and $|\phi\rangle$ of σ, both into RA.

Proof. Choose orthonormal bases $\{|i_A\rangle\}$ and $\{|i_R\rangle\}$ of A and R. Define $|m\rangle \equiv \sum_i |i_R\rangle |i_A\rangle$, and let $|\psi\rangle$ be an arbitrary purification of ρ. The Schmidt decomposition then implies that there exist some unitary operators U_A and U_R on A and R respectively such that

$$|\psi\rangle = (U_R \otimes \sqrt{\rho} U_A) |m\rangle . \tag{A.2.13}$$

We have an identical statement for some other unitaries V_A and V_R, and arbitrary purification $|\phi\rangle$ of σ. Taking an inner product then implies

$$|\langle \psi|\phi\rangle | = |\langle m|(U_R^\dagger V_R \otimes U_A^\dagger \sqrt{\rho}\sqrt{\sigma} V_A)|m\rangle |. \tag{A.2.14}$$

Using lemma A.2.2, we then have

$$|\langle \psi|\phi\rangle | = |\mathrm{Tr}(V_R^\dagger U_R U_A^\dagger \sqrt{\rho}\sqrt{\sigma} V_A)|. \tag{A.2.15}$$

So, defining $U \equiv V_A V_R^\dagger U_R U_A^\dagger$, we have

$$|\langle \psi|\phi\rangle | = |\mathrm{Tr}(\sqrt{\rho}\sqrt{\sigma} U)|. \tag{A.2.16}$$

Then, by lemma A.2.1, we have

$$|\langle \psi|\phi\rangle | \leq \mathrm{Tr}|\sqrt{\rho}\sqrt{\sigma}| = \mathrm{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}} = F(\rho, \sigma). \tag{A.2.17}$$

To attain equality, take the polar decomposition $\sqrt{\rho}\sqrt{\sigma} = |\sqrt{\rho}\sqrt{\sigma}|V$, and choose $U_A = U_R = V_R = I$ and $V_A = V^\dagger$, and we are done. □

While Uhlmann's theorem is not particularly useful in a computational sense, it shows us that the fidelity indeed has properties we would expect from a distance measure. It clearly implies symmetry of inputs $F(\rho, \sigma) = F(\sigma, \rho)$, and that it is bounded between 0 and 1, with equality with 1 if $\rho = \sigma$, and equality with 0 if and only if ρ and σ are supported on orthogonal subspaces. This would make sense: if two states are orthogonal, they are perfectly distinguishable, so we would expect minimum fidelity between them.
Fidelity is not a metric, but we can turn it into one. To do this, we make the following definition:

Definition A.2.2 (Angle)**.** Given two states ρ and σ, the **angle** between them is

$$A(\rho, \sigma) \equiv \arccos F(\rho, \sigma). \tag{A.2.18}$$

The motivation for this comes from the fact that the angle between two points on a sphere is a metric, and Uhlmann's theorem tells us that the fidelity between two states is the maximum possible inner product between purifications of these states. Clearly the angle is non-negative and symmetric in its inputs, and $A(\rho, \sigma) = 0 \iff \rho = \sigma$. The only property left to show angle is a metric is that it obeys the triangle inequality. To do this, consider three states ρ, σ, and τ. Choose purifications $|\phi\rangle$ of σ, $|\psi\rangle$ of ρ, and $|\gamma\rangle$ of τ such that

$$\begin{aligned} F(\rho, \sigma) &= \langle\psi|\phi\rangle \\ F(\sigma, \tau) &= \langle\phi|\gamma\rangle\,, \end{aligned} \tag{A.2.19}$$

and $\langle\psi|\gamma\rangle$ is real and positive. We have in general for arbitrary vectors that

$$\arccos(\langle\psi|\gamma\rangle) \leq \arccos(\langle\psi|\phi\rangle) + \arccos(\phi|\gamma) = A(\rho, \sigma) + A(\sigma, \tau), \tag{A.2.20}$$

but by Uhlmann's theorem, $F(\rho, \tau) \geq \langle\psi|\gamma\rangle$. Therefore $A(\rho, \tau) \leq \arccos(\langle\psi|\gamma\rangle)$. Together with (A.2.20), we then get the triangle equality, so angle is indeed a metric.
Fidelity is also monotonic under trace-preserving quantum operations.

Proposition 2. *Suppose $\mathcal{E}$ is a trace-preserving quantum operation, and ρ and σ are states of system A. Then*

$$F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma). \tag{A.2.21}$$

Proof. Take purifications $|\psi\rangle$ and $|\phi\rangle$ of ρ and σ respectively onto a joint system AR such that $F(\rho, \sigma) = |\langle\psi|\phi\rangle|$. Introduce an environment E for the operation $\mathcal{E}$, which is initially in the state $|0\rangle_E$, and take it to interact with A via a unitary

operation U. $U|\psi\rangle_{RA}|0\rangle_E$ is a purification of $\mathcal{E}(\rho)$, and similarly $U|\phi\rangle_{RA}|0\rangle_E$ is of $\mathcal{E}(\sigma)$. By Uhlmann's theorem, we therefore have

$$\begin{aligned} F(\mathcal{E}(\rho),\mathcal{E}(\sigma)) &\geq |\langle\psi|\langle 0|U^\dagger U|\phi\rangle|0\rangle| \\ &= |\langle\psi|\phi\rangle| \\ &= F(\rho,\sigma) \end{aligned} \tag{A.2.22}$$

as claimed. □

Our last property of fidelity is strong concavity.

Theorem A.2.2. *Let p_i and q_i be probability distributions indexed by i, and ρ_i and σ_i density operators also indexed by i. Then*

$$F\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \geq \sum_i \sqrt{p_i q_i} F(\rho_i,\sigma_i). \tag{A.2.23}$$

Proof. Choose purifications $|\psi_i\rangle$ and $|\phi_i\rangle$ of ρ_i and σ_i respectively such that $F(\rho_i,\sigma_i) = \langle\psi_i|\phi_i\rangle$. Introduce an ancillary system which has orthonormal basis $\{|i\rangle\}$, indexed by the same i as everything else. Define

$$|\psi\rangle \equiv \sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle, \quad |\phi\rangle \equiv \sum_i \sqrt{q_i}|\phi_i\rangle|i\rangle. \tag{A.2.24}$$

Note that $|\psi\rangle$ purifies $\sum_i p_i\rho_i$ and $|\phi\rangle$ purifies $\sum_i q_i\sigma_i$, so by Uhlmann's theorem:

$$\begin{aligned} F\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) &\geq |\langle\psi|\phi\rangle| \\ &= \sum_i \sqrt{p_i q_i}\langle\psi_i|\phi_i\rangle \\ &= \sum_i \sqrt{p_i q_i} F(\rho_i,\sigma_i) \end{aligned} \tag{A.2.25}$$

as claimed. □

Appendix B

Stuff that won't be read by anyone

Some people include in their thesis a lot of detail, particularly lots of tables containing raw results, figures of intermediate results, or computer code which no-one will ever read. You should be careful that anything like this you include should contain some element of uniqueness which justifies its inclusion.

Bibliography

[1] D. Harlow, *The Ryu-Takayanagi Formula from Quantum Error Correction*, Communications in Mathematical Physics, 354 (2017), pp. 865–912.

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 10th ed., 2010.