

Extending Ancilla Driven Universal Quantum Computation Beyond Stepwise Determinism

by

Kerem Halil Shah

supervised by

Dr. Daniel Oi



DEPARTMENT OF PHYSICS,
UNIVERSITY OF STRATHCLYDE

A THESIS PRESENTED IN THE FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

August 2016

Copyright declaration

This thesis is the result of the authors original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:

Date:

For Mum and Dad,

It's all for you.

To Meltem and Kadir,

Because I got to go first,

I hope to leave behind something good.

Abstract

A major research goal in the field of quantum computation is the construction of the universal quantum computer (UQC): a device that can implement any quantum algorithm. Several theoretical schemes for implementing UQC have been developed which require different sets of resources and capabilities with varying implications for the optimum experimental implementations. The ancilla driven quantum computation scheme (ADQC) comprises two subsystems: a memory register of qubits on which information is retained and processed and an ancilla system of qubits which couple to the register. This coupling is represented in the ADQC scheme by a fixed quantum gate. By preparing the ancilla in selected states before applying this gate and then measuring it in selected measurement basis afterwards, quantum gates are enacted on the register qubits. ADQC is deterministic in that the probability of the outcome after performing the entire procedure is 1 but we have to apply corrections to the procedure at each step that depend on the probabilistic outcome of the ancilla measurement. An important resource in this model is the availability of a maximally entangling two-qubit gate between the ancilla and register qubits because if the gate is not maximally entangling, the resulting gates on the register can not be selected with stepwise determinism.

It is proven in this thesis that in fact ADQC with non-maximally entangling interaction gates is universal. This requires showing that single- and two-qubit unitary gates can be efficiently implemented probabilistically. We also show a relationship between the expected time of the probabilistic implementation of a gate and the ability to control the ancilla. In the ADQC model, the ancilla is controlled with single qubit unitary gates just before interacting with the register and just before measurement. We show that the increase in time caused by a loss of maximally entangling two-qubit gates can be counteracted by control over the ancilla. This needs not be the ability to perform any single qubit unitary to the ancilla but just the ability to perform a specific small finite set of operations.

This is important because the resource requirements described by a scheme affect the properties of possible experimental implementations. The ADQC scheme was originally designed to be used with physical implementations of quantum computing that involves qubits coming from different physical systems that have different properties. This may restrict the availability of couplings between the register and ancilla systems equivalent to maximally entangling quantum gates. By further focusing on the model under specific restrictions, such as minimal control of the ancilla system or long distance separation between register qubits, we find certain properties of the physical

implementation that may best suit it for ADQC beyond stepwise determinism. Minimal control appears best suited for symmetric ancilla-register interactions; use over long distances suits a transmitter going to an unknown receiver with possible small errors in the receiver's interaction with the ancilla.

Contents

1	Introduction	1
1.0.1	The structure of this document	5
2	Concepts In Computation	7
2.1	Classical Concepts	7
2.1.1	Simulation & Universality	7
2.1.2	Resources	11
2.1.3	Efficiency & Inefficiency	12
2.1.4	Complexity Classes	13
2.2	Quantum Concepts—The impact of quantum mechanics on computation	14
2.2.1	Quantum Mechanics and simulation	14
2.2.2	The Universal Quantum Computer	18
2.3	Schemes for Quantum Computation	24
2.3.1	The Quantum Gate Circuit	24
2.3.2	Universal Finite Gate Sets	29
2.3.3	Measurement Based Quantum Computation	38
2.4	Ancilla Driven Universal Quantum Computation	44
2.4.1	$E_{AR} = (H \otimes H).CZ$	46
2.4.2	Beyond deterministic ADQC	48
3	Dilation theory	50
3.1	Generalised Quantum Mechanics and the Stinespring Dilation Theorem	50
3.1.1	Mixed states	50
3.1.2	Generalised measurement	56
3.1.3	Generalised Evolution	62
3.1.4	Weak Values	65
3.1.5	Iterative implementations of POVMs and binary search trees . .	70

4	Universality	74
4.1	Implementation of unitary channels through projections on an ancilla qubit	74
4.1.1	The Cartan Decomposition of Two-Qubit gates	75
4.1.2	Deriving the conditions for unitarity	77
4.1.3	Available unitary gates	82
4.2	Ancilla Driven Entangling Gates	87
4.2.1	Conditions of the two-qubit entangling gate generation	89
4.2.2	The entangling power of a generated two-qubit gate	100
4.2.3	The local gate decomposition of the Kraus operation	104
4.3	Random walks on a finite group	106
4.3.1	Case 0	107
4.3.2	Case 1	112
4.3.3	Case 2	115
4.3.4	The guaranteed hitting time	117
4.4	Simulating deterministic circuits with probabilistic gate generation . . .	120
4.4.1	Single-qubit unitary gates	121
4.4.2	Simulation of stochastic single-qubit gates	124
4.4.3	Stochastic generation of two-qubit gates: walks on a circle	129
4.4.4	Applying strategies to random gate generation	132
4.4.5	Concluding Commentary	137
5	Extreme conditions: long distances & minimal control	139
5.1	Entangling unitary gates on distant qubits with ancilla feedback	139
5.1.1	Intro	139
5.1.2	The model	140
5.1.3	The one-step strategy	142
5.1.4	Numerical results	143
5.1.5	Asymmetric transmitter/receiver interaction gates	146
5.1.6	Concluding commentary	150
5.2	Minimum Control Ancilla Driven Universal Quantum Computation . . .	152
5.2.1	Fulfilling both sets of unitary conditions simultaneously	154
5.2.2	Single-qubit gate approximation and an entangling gate	156
5.2.3	Replicating other universal gate sets	159
5.2.4	Concluding Commentary	160
6	Outlook	163
6.1	Further avenues of research	163
6.1.1	Characterising single qubit gate walks	163

6.1.2	Noise	166
6.1.3	Using Generalised Measurements	167
6.1.4	Interpretations of post-selection and weak values	168
7	Summary	173
A	Proof of the decomposition of a single qubit unitary into rotations about any two non-parallel axes	178
B	Statistical analysis of numerically computed results	180
B.1	Simulation of random gate hitting times	180
B.1.1	Generating single-qubit gate hitting times.	180
B.1.2	Error bar calculations	183
B.2	Numerical calculations of guided strategy probabilities	184
B.2.1	Errors in the numerical calculation	187

Chapter 1

Introduction

The ability to create experimental devices that exploit the laws of quantum mechanics has grown hand in hand with our theoretical understanding of how to utilise quantum systems to enhance classical technological capabilities in several fields such as quantum cryptography, communications, metrology and computation.

The ability to create experimental devices that exploit the laws of quantum mechanics has grown hand in hand with our theoretical understanding of how to utilise quantum systems to enhance classical technological capabilities in several fields such as quantum cryptography, communications, metrology and computation.

Early theoretical descriptions of computational devices composed of quantum mechanical components were given by Benioff [1] and Deutsch [2] and Feynman [3] in 1980, '82 and '85 respectively. Quantum devices such as SQUIDS [4, 5, 6], scanning tunnelling microscopes [7, 8, 9, 10], quantum dots [11, 12, 13], atom traps [14] and single photon sources [15] had started to arise from the late 1960's to early 1980's. A major part of the theory of quantum computation is understanding which particular capabilities and devices are necessary to create a quantum computer with all possible advantages over a classical computer. The challenge is in creating the language and concepts that can be used to describe an as yet non-existent technology.

The model of the Universal Quantum Computer (UQC) was developed by David Deutsch [2] initially as a way of reconciling the model of the Universal Turing Machine (UTM) with the modern understanding of the physical laws of nature. The idea of a quantum computer or QC had been first inspired by the question of how a computer could accurately simulate a physical universe that is governed by the laws of quantum mechanics. The answer according to early theorists like Deutsch [2] and Feynman [3] is that if a device was to be capable of processing information about the natural world, which runs according to quantum mechanics, then the device would have to run according to the laws of quantum mechanics itself. Feynman had suggested a quantum version of cellular automata to act as a field theory simulator but Deutsch's focus was

on the Church-Turing principle: every “function which would naturally be regarded as computable” can be computed by the Universal Turing machine. David Deutsch believed that to actually turn the Church-Turing principle into an experimentally testable theory, it needed to be related to the issue of physical simulation and created a quantum version of the Turing machine to act as the conceptual bridge between classical computer science and quantum mechanics. What Deutsch and others then found was that such a theorised device would then be capable of performing operations that outperformed any classical algorithm known to date.

This shows the role of theory in science that goes beyond just being a predictor of data; theory can develop new concepts so that we might better understand previous results and can craft a narrative over them that then extends into the future. Quantum computation became a target for future technological development. Not only would a universal quantum computer be capable of the simulations for which it was imagined [16], Deutsch found a novel algorithm that a quantum computer could implement that was more resource efficient than any known classical one [2] but only for a problem that Deutsch had considered for that purpose. Yet after this, further improvements were made by Deutsch & Jozsa [17], Berthiaume & Brassard [18] and Simon [19]. Bernstein and Vazirani [20] applied a complexity class approach to the field so now access to quantum operations could be seen as a novel resource. Eventually Peter Shor developed an algorithm [21] for integer factoring in polynomial time, an example of an already well known and established problem widely studied and assumed to not have an efficient classical solution.

Schemes for quantum computations describe how the goal of UQC can be achieved from a particular set of resources used in a particular sequence. For example, Deutsch [22] developed an analogue of the logic gate model for quantum computers, quantum gate based quantum computing (GBQC), in which a QC is broken up into a finite set of discrete unitary operations on qubits, the quantum gates, which can be put into a selectable arrangement to build a quantum algorithm.

Over the years more schemes have been invented that relate different mathematically defined resources or properties in theoretical models to the issues of implementation: Measurement based quantum computation (MBQC), developed by Raussendorf & Briegel [23], related large entangled states such as cluster states to natural lattice structures and probabilistic operations with photonic systems [24, 25]; Kitaev’s topological approach connected anyons and braid theory to the issue of fault tolerance [26]; adiabatic quantum computation [27] relates the adiabatic theorem and quantum simulation to the class of optimisation problem amenable to annealing [28].

As the field has developed, research has gone into what the resources of each scheme exactly need to be in the face of new considerations about their resources. For example,

Deutsch [22]’s GBQC provided an example of a single gate that could be universal by itself if it could be applied across any trio of qubits; subsequent results provided further single universal gates that applied over two qubits [29, 30, 31] which simplifies inter-qubit interactions; eventually it was found that a set could consist of one two-qubit gate and just two single qubit gates, labelled $\{H, T\}$, that aided in the implementation of error correction [32, 33, 34]. Modern study of universal gate sets and their decompositions has to consider the practicalities of fault tolerant implementation, circuit depth and the number of gates [35, 36].

The Ancilla Driven Universal Quantum Computation (ADQC) scheme proposed by Kashefi *et al* [37, 38] is underpinned by the theory of measurement induced back action under the Stinespring dilation theorem whereby qubits on a register undergo unitary operations by being coupled to ancillary qubits which then undergo a measurement that selects the unitary operation on the register. The finite set of operations needed to perform this process are near identical to those in MBQC but the change in paradigm from the information travelling through a chain of qubits in MBQC to remaining on the same register qubits in ADQC makes the ADQC scheme more suitable for a paradigm of hybrid physical implementations.

In a hybrid design, such as ion trap + photon systems [39] or NV centre spins [40, 41], the register has physical properties optimised for coherence lifetimes while the ancillary system is suited for applying gate operations. The key principle is that the properties of the former and latter usually have to be traded off against each other within a single physical system. There are other schemes involving ancillary systems e.g the quantum bus [42] scheme and the ancilla controlled quantum computation scheme (ACQC) [43], but the reliance on measurement induced back-action emphasises the disparity between the coherence lifetimes of the register and ancilla instead of other properties such as e.g. the relative mobility.

In MBQC, the question of universal resources has been expanded to look at constructions built from non-maximally entangling gates [44, 45]; the problem of characterisation of quantum systems is being explored with weak couplings [46, 47]; future candidates for physical implementation may involve scattering based interactions that may result in a lack of interaction tuning [48, 49]- in this thesis we turn the attention given to resource sets using non-maximally entangling gates onto ADQC. Concurrently with this work, ACQC [43] was developed which also looks at non-maximally entangling gates being used within an ancilla-mediate scheme. This looked at a smaller subset of the ancilla-register interactions and did not utilise measurement-induced back-action but rather used total control over the ancilla and multiple ancilla-register interactions in a manner more similar to the quantum bus. Later, the authors found a minimal control case [50] comparable to that described towards the end of this thesis.

The aforementioned interest of other researchers in applying non-maximally entangling operations or experimental choices that may lend themselves to non-maximally entangling operations to other fields of study, motivates us to do the same with ADQC. In this thesis, we will investigate further the relationship between the model of ancilla driven quantum computation and the set of finite gates it uses as a resource in the construction of a universal quantum computer. We will show

1. Ancilla driven quantum computation with interaction gates of non-maximal entangling strength can be universal
2. In the above case, there is a trade off between the time to implement the scheme and the ability to control the ancilla through the choice of preparation state and measurement basis.

The goal here is not just to provide one or two conditions that demonstrate equivalence to requirements to universality but also to show how the system behaves as we vary the level of control over the ancilla and the strength of the interaction gate. This model of behaviour can then be used to achieve our second goal. It is known from previous results in ADQC that the first goal removes the possibility of stepwise determinism [38, 37]. Getting around this limitation is the major challenge of this thesis. We will find that under some conditions, single qubit operations may be performed deterministically or at least with a small probabilistic number of Pauli operator corrections but we will have to apply a stochastic process to prove universality. For the second goal, the ensuing questions include whether the control of the ancilla is symmetric with respect to whether it is during preparation or measurement and what the trade-off relationship is like and what particular aspects of the ancilla determine it. It is found that control over the ancilla preparation or ancilla measurement can be considered equal, as in deterministic ADQC, except under conditions introduced in this thesis where interaction gates do not maintain a consistent entangling strength across different register qubits. We look at how the trade-off in resources works in terms of how it affects the structure of the stochastic process. Changes in the resource requirements for control over the ancilla or in their behaviour may ultimately reflect changes in the physical properties of a viable implementation.

1.0.1 The structure of this document

In chapter 2, there will be a review of concepts from classical and quantum computation which need repeated reference through this thesis, the aim of which is to bridge the gap between an undergraduate physicist with no computer science background and the introductory level understanding of this thesis. Since the aim of this thesis is to show that a particular model is universal, this chapter will cover the theoretical basis of how a model is established as universal. More advanced readers may wish to skip this chapter but note that some motivations may be expressed with reference to the content of these chapters and some of our choice of notation will be established here. In the last section of chapter 2, section 2.4, the original model of ADQC which this thesis builds on will be described.

In chapter 3, we will cover the mathematical formalism necessary for this work. We cover the applications of the Stinespring and Naimark Dilation theorems that extend basic quantum mechanics foundations to open systems and show how to derive results in background materials that provided inspiration to ADQC model. Knowledge of this chapter's material will provide the reader the ability to reproduce our later results.

In chapter 4, we apply the techniques covered in chapter 3 to the goal of demonstrating that ancilla driven quantum computation with interaction gates of non-maximal entangling strength can be universal. In section 4.1, using dilation theory we can reproduce the claims of earlier work on ADQC about which classes of interaction gate can be used and that the loss of maximally entangling interaction gates causes a loss of stepwise determinism. We go further and show which single qubit gates occur as a result and their relationship to the interaction gate's entangling strength and the ancilla's preparation state and measurement basis. In 4.2, we do likewise with the generation of two-qubit gates on the register. In the rest of the chapter, we then provide the model by which ADQC with non-maximal interaction gates can overcome the loss of stepwise determinism and how it behaves as we vary the level of control over the ancilla and the strength of the interaction gate. The result involves a probabilistic time cost and in section 4.4, there is discussion of the statistics of the time cost. The statistical analysis of the numerically computed results in this section can be found in appendix B.

In chapter 5, the further impact of the extension of the ADQC model on the mode of implementation will be explored by considering two particular extremes: performing entangling operations on qubits in distant nodes of a network and removing control of the ancilla system parameters. Chapter 6 will cover some potential research offshoots for the ancilla driven model.

This thesis provided the basis for the publication and pre-prints:

Ancilla Driven Quantum Computation with arbitrary entangling strength. *Proc. 8th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2013)*, 2013. [51]

Entangling unitary gates on distant qubits with ancilla feedback. *eprint arXiv:quant-ph/1311.3463*, November 2013. [52]

A minimum control ancilla driven quantum computation scheme with repeat-until-success style gate generation. *eprint arXiv:quant-ph/1401.8004*, February 2014. [53]

with the material for the former being covered in chapter 4 and for the latter two in chapter 5.

Chapter 2

Concepts In Computation

2.1 Classical Concepts

2.1.1 Simulation & Universality

Universality will be an often referenced target of this thesis. Generally universality is taken to mean that a device is capable of performing any quantum algorithm. However the proof of universality of the scheme discussed in this thesis will not reference any algorithms. Nor do proofs of many other schemes. In fact it will usually suffice in this thesis to understand it to mean that a device can implement any finite dimensional unitary operation on a finite number of qubits. This chapter provides the background of classical and quantum computation concepts that allow the reader to understand why the latter definition of universality is equivalent to the former and how the notion that universality has been achieved is tested.

In practice the schemes designed to construct and implement universal devices invoke considerable wiggle room in achieving this goal – they achieve a *simulation*, they work to an *approximation* but they do so *efficiently* – and to understand the validity of these accommodations one needs the perspective of the original concept of the Universal Turing Machine.

A Turing machine is an automatic machine devised by Turing in 1936 [54] as an abstraction of a human being calculating a numerical result to a problem. When a human being solves a problem, they read an input, scan each symbol in an order one at a time and then go through a process in which the problem is broken down into steps with each next step depending on the previous result and the description of the problem solving methods. We may have different mathematical techniques or different coding languages but these are the differences of the symbols and configurations. Turing defined the numbers that can result from such a calculation as a *computable number*: “real numbers whose expressions as a decimal are calculable by finite means”

[54]. Turing equated computable to Church's definition of *effectively calculable* [55, 54]. There is a need to specify finite means because human beings themselves have finite memories and access to finite resources and this means that there are numbers which are *definable* numbers which can not be computed.

The machine consists of three parts

1. a memory tape register
2. a read/write head
3. a body that holds internal machine configurations

The tape is divided into square sections that have a single symbol inscribed upon them each and is treated as being infinitely long. The machine head has the ability to read a single symbol at a time which it can then erase and/or write over and then move one square to the left or right. It has no awareness of any other symbol than the scanned symbol at the present time. How the machine will react to each symbol will depend on its internal configuration. Each machine configuration will describe whether the machine will erase and/or what it will write on the square, how it will move and finally what configuration it will be in for the next section. Thus the configurations act as a sort of memory for the machine since its current configuration will depend on the sequence of symbols it has previously scanned and its actions on the current section will be determined by the current symbol and current configuration. If the machine configuration completely described the possible actions of the machine then it will function automatically without need for any further operator input other than the tape. The symbols of the tape could be of any symbol system- the alphabet, mathematical notation, even A,T,G and C- but those that work with binary numbers only are called computing machines.

Recognise that the configurations of the machine detail an algorithm. The machine configurations are represented by a configuration table in which each line shows a starting configuration state and a read symbol and then the corresponding actions and final configuration. For example, the following table represents a machine that scanning rightwards, taking each distinct pair of bit i and j and calculates $i \oplus j$, writing it over the second bit and halting on encountering an empty slot:

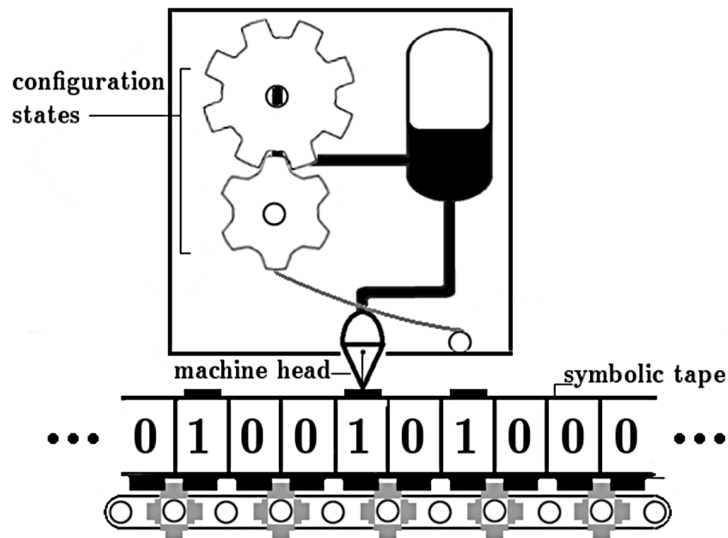


Figure 2.1: Imagining a Turing machine: each of the three main components is given a mechanical expression. Here, the machine moves through configurations in finite steps like a cog, the symbolic language is expressed by raised ink deposits on segments on tape, read by a cantilever head with an attachment to deposit and remove the ink.

start config.	symbol	operation	end config.
a	0		b
	1	R	c
	N/A		halt
b	0/1	R	a
	N/A		halt
c	0	Print[1], R	a
	1	Print[0], R	a
	N/A	R	halt

(2.1)

Note that the line for each configuration splits for different behaviour on different input as a formatting choice. The format of the configuration tables can be simplified by collecting some operations together. Something like multiple moves in one direction by a fixed amount without doing anything, strictly speaking takes several steps and several configuration states that feed directly into the next, but can just be written as a single configuration with multiple steps. Perhaps the multiple options for different symbols for the same configuration can be written on the same line. However by being consistent the configuration table can be written so that each line has the same amount

of slots to write down every available option.

q_i	s_i	Print[s_j], L/R	q_m
-------	-------	-----------------------	-------

For a computing machine, the element q_i may be a string of binary just long enough to have one for each configuration state, the symbol slot may be two binary digits or perhaps just one and a special symbol to represent an empty slot, there needs to be an order chosen for the organisation of the printing actions and the movement actions and how many movements can exist in a single configuration slot. If this is worked out then the configuration table line will map to a single fixed length string of numbers and a sequence of those numbers- the *satisfactory number* and the mapping will define the machine.

Turing was then able to invent a machine, the Universal Turing Machine (UTM), that could be used to compute any computable sequence. The trick is to take the satisfactory number of another machine and feed it into the UTM. The UTM then can take the number and an input for the computation and then be programmed by the number to implement the output of the configuration table.

There is a finite number of options in a single line of a configuration table. Given this, it is conceivable that the UTM can be given the ability to adopt configurations that represent any of the finite range of inputs and resulting operations of another machine. The difficulty and main issue of the UTM configuration table that Turing developed is in moving from the current configuration to the right one next. The UTM has to move along the program input to find the current state and the description for the next one.

The format of the program needs markers to indicate the relevant sections of the significant number. The UTM can have configurations for moving left or right until finding these markers and also copies the current configuration next to the input so the steps of the other machine in the computation are simulated on the tape.

The nature of the UTM is not that it has a configuration table that includes every computation but that it has the ability to read the description of another machine and then simulate it. There are two major implications of such a device. The first is that any computation that can be performed can be performed by the Universal Turing Machine or conversely, any computation that a Universal Turing Machine can not do, can not be done otherwise. Due to Turing's equating to computability and Church's effective calculability, this, when expressed in more formal language, leads to the Church-Turing thesis: any effectively calculable function can be a mechanical computation of the Universal Turing Machine. The second is that any device that can be shown to be equivalent to the Universal Turing Machine, even if it has to use some different mapping or format with its program input, will also be able to simulate

any computation; such a device is referred to as Turing complete or computationally universal.

2.1.2 Resources

In the Turing machine model, the performance of a computation uses up resources. The two most important resources that are considered are those of time and space. The space that the machine uses is the span of the sections on the memory tape register that the machine head visits at least once during the computation while the time is the number of steps taken to perform the computation [56]. The CNOT machine defined by (2.1), for example, will visit each bit once until it reaches an empty slot and so uses a space of $x + 1$ for an input of x bits and uses an equal amount of time. Were it to take the modulo addition of every adjacent bit pair rather treating each bit as part of a unique pair, it would spend more time per slot and so the time resource increases. These resource costs are a function of the machine, M , and the input size, x , $S_M(x)$, $T_M(x)$. There are other resources that come from additions to the Turing machine itself. The ability of access a source of randomness can be a resource. The question of what happens when the Turing machine has access to an input that includes randomly generate bit values was considered by De Leeuw, Moore, Shannon & Shapiro [57], Gill [58] and Santos [59]- the Probabilistic Turing Machine (PTM) (there is also a closely related concept called the Nondeterministic Turing Machine where the configurations themselves are capable of random behaviour). From the work of Solovay & Strassen [60] and Rabin [61] which used algorithms where randomness plays a central role in the construction for the task of testing primality, it was found that a PTM admits a new class of algorithm, *randomised algorithms*, that today find themselves used in various fields such as in number theory, pattern matching, selection, sorting, searching, computational geometry and parallel and distributed computation [62, 63]. Such algorithms can be faster than a deterministic equivalent and may in fact be easier to find than a deterministic equivalent but will now introduce the probability of an erroneous result which either ends up being very small or accounted for by repetition and a probabilistic time cost. Another adjusted Turing machine is the Multitape Turing Machine (MTM) developed by Hartmanis & Stearn [64] where the machine is fed multiple tapes yet all computations that can be performed by a multitape machine, M' , can be performed by a single tape machine, M , just with $T_M \approx (x)\mathcal{O}(T_{M'}(x)^2)$ [64, 56]. Generally, resources can be traded for each other with a loss of capabilities coming at a cost of time and space.

2.1.3 Efficiency & Inefficiency

A function may be implemented by several different machines or their corresponding algorithms which leads to several resource functions $S_M(x)$ and $T_M(x)$. Due to the lack of actual infinite tapes and fleeting nature of time, as well as the intuitive relationship between how long the machine takes and how complicated it is to build, there is a natural need to find the most resource efficient algorithm. This efficiency is given by the function that bounds the value of the resource costs.

For example, the Fourier transform is an operation commonly used in data transmission and analysis that transforms N values into an N entry Fourier series according to:

$$p_k = \sum_{n=0}^{N-1} x_n e^{-i\frac{2\pi k}{N}n}, \quad k = 0, \dots, N-1.$$

Since each of the N terms in the series requires a sum of N terms, a naive algorithm based on the calculation of each $x_n e^{-i\frac{2\pi k}{N}n}$ requires N^2 such operations. However, due to the symmetries of the operation, there is considerable redundancy in all of the terms. It is therefore possible to use an alternative Fast Fourier Transform that takes only $\frac{1}{2}N \cdot \log_2 N$ operations which saves significant resources for large values of N .

Because the concern is based on how the resource costs grow as the input gets large, the resource functions are typically classified by the order of the largest term that will dominate at large sizes. Three different notations exist to indicate the different bounds on the behaviour of the function, \mathcal{O} , Ω and Θ notation [65]:

$$f(n) \text{ is } \mathcal{O}(g(n)) \text{ if } \exists \text{ constants } c, n_0 \text{ s.t. } f(n) \leq cg(n), \forall n > n_0; \quad (2.2)$$

$$f(n) \text{ is } \Omega(g(n)) \text{ if } \exists \text{ constants } c, n_0 \text{ s.t. } f(n) \geq cg(n), \forall n > n_0; \quad (2.3)$$

$$f(n) \text{ is } \Theta(g(n)) \text{ if } \exists \text{ constants } c_0, c_1, n_0 \text{ s.t. } c_0g(n) \leq f(n) \leq c_1g(n), \forall n > n_0. \quad (2.4)$$

Each notation indicates the behaviour of the function as it grows large with \mathcal{O} notation giving an upper bound on the cost of an algorithm while Ω provides a lower bound that is often used more in conjunction with a class of algorithms that apply to a particular problem. Θ notation is essentially being both $\mathcal{O}(g(n))$ and $\Omega(g(n))$, behaving as $g(n)$ asymptotically up to a constant factor.

Errors also induce a resource cost in that resources may be spent trying to minimise them. This may occur with probabilistic failure and repetition to get the right result, it can also occur in approximation computations. Indeed, computability was linked to expression as a decimal in its definition and this decimal may be expressed only by a finite number of decimal points that can be increased by increasing the space of the input or the running times for the output. Here the growth function will be in terms of $\frac{1}{\epsilon}$ for an error ϵ .

There has also been developed a sense in which an algorithm is *inefficient*, that is, it is when the cost of resources grows so high that it just stops being an effective solution. The concept stems from a consideration of Edmonds [66]’s that if some orders of difficulty can be better than others and an algebraic order of difficulty can be good, then an exponential one can be considered bad. Over the past 50 years, this concept appears to have received little extra rigorous justification but survives due to its practicality.

Part of why exponential growth may be such an appealing limit is the exponential behaviour of Moore’s law. Moore’s law originally comes from a projection of the growth in the number of transistors in a chip given by Moore [67] in 1965 to be about double a year. That was a projection about a technology in its early infancy but the statement has grown broader and chimeric in terms of its attribution to become a statement about the doubling of processor speed every 18 months. Moore’s law taps into a view about technological growth itself – that the rate of growth is proportional to the growth achieved. Any algorithm that fails to improve on exponential time may be outstripped by the growth in technology itself.

With the notion of efficiency, the Church-Turing principle itself starts to be adapted to require that the UTM implements the simulation of other algorithms efficiently rather than a straightforward notion of possibility/impossibility. This also means that if an implementation is efficient in an algorithm, it will remain efficient in the UTM and efficient approximations and simulations of a UTM.

2.1.4 Complexity Classes

The model of a Universal Turing Machine where the algorithm of a machine can be simulated by the UTM encourages thinking of the algorithm as a process that can be abstracted from the physical machinery that implements it. In a similar sense, the principle of complexity classes is to abstract the difficulty of a problem from the algorithms used to solve it. The computation problem is characterised by the resources required to solve it.

The principles of complexity classes were established by the first considerations of the space and time resource functions by Hartmanis & Stearns [64]. They provided the relationship between resource costs and the input and how this changed when a multitape Turing machine was used and showed that every computable number could be put into a complexity class by the time to implement on an MTM $T_{MTM}(x)$. They also found that though the addition of extra machine resources such as extra tapes could affect the complexity class, there was a limitation to this effect, supporting the underlying assumption that there is a minimum complexity assignable to the problem. Further support was given by the results of Cook [68] and Karp [69] demonstrating the

existence of problems that could be fixed within the class of nondeterministic polynomial solutions and therefore if one problem of this class could be shown to be of a lower complexity class, it would be true of all the other problems.

In this section, we saw how the Turing machine model establishes the idea of the computations of a machine being implemented by other machines. This may come at a cost of resources but the underlying algorithm and the complexity of the problem it addresses is not changed by an efficient implementation. In the next section, we will describe how access to systems which display properties described by quantum mechanics can be utilised to provide a computational advantage. The lack of them make some problem unsolvable efficiently, the ability to harness them provides the ability to implement previously unknown algorithms and access to certain quantum operations and states become resource costs in the implementation of a machine.

2.2 Quantum Concepts—The impact of quantum mechanics on computation

2.2.1 Quantum Mechanics and simulation

In the modern day, computers play an important role in the modelling of physical systems, especially when the size of the calculation is large or exact analytical methods are just not available. Computers allow us to make the empirical predictions necessary for viable theories from chaotic weather systems, to computational particle physics, to solid state materials science. Naturally then the question of how a computer can simulate a quantum mechanical system has arisen, especially as the focus transitions into smaller and smaller realms where very small parcels of materials become very large numbers of quantum particles. In fact the technology that underpins modern computers depends on such science and the feedback of advancements in the field into the devices that enables them may very well be what drives the exponential growth of Moore’s law.

Bearing this in mind, it is no surprise that Richard Feynman who spoke of “Plenty of Room at the Bottom” [70] and the potential of science at a scale where physics and chemistry find an inter-disciplinary crossover would later be inspired by the work of Ed Fredkin, who at around this time was working on the matter of reversible computation along with Tommaso Toffoli [71], to look at science at the scale where the disciplines of physics and information crossover. In 1982 [3], Feynman asserted that a classical computer would not be able to efficiently simulate an arbitrary quantum system.

To explore why, it is first necessary to establish what one means by “quantum mechanics” and the particular formalism used to describe it. A quantum system is

taken to be one that obeys the laws and results derivable from these five postulates:

1. Physical states are vectors in a complex Hilbert space represented by $|\psi\rangle$. As such, there is an inner product norm $\|\psi\rangle = \sqrt{\langle\psi|\psi\rangle}$ where $\langle\phi|\psi\rangle$ is linear, positive and skew symmetric. This norm provides a distance between vectors in the space under which we can say the space is complete. Every $|\psi\rangle$ can be represented as a linear sum of other vectors in the space and has a unique representation of no more than d vectors in a d dimensional Hilbert space when we use an orthonormal basis $\{|j\rangle\}$ for which $\langle j|k\rangle = \delta_{jk}$.
2. Observables are represented by Hermitian operators \hat{A} that map the Hilbert space to itself and whose eigenvectors $\{|a_m\rangle\}$ for which $\hat{A}|a_m\rangle = a_m|a_m\rangle$ form a basis of the Hilbert space. Being Hermitian, $a_m \in \mathbb{R}$.
3. Measurement results are represented by measurement operators (or “projectors”) $\{|a_m\rangle\langle a_m|\}$, $\sum_m |a_m\rangle\langle a_m| = \mathbb{I}$. The probability of result m is given by $\langle\psi|a_m\rangle\langle a_m|\psi\rangle$.
4. The state after projective measurement is $|a_m\rangle$.
5. The dynamics of a physical state evolve under the Schrödinger equation expressed as

$$\frac{d}{dt}|\psi(t)\rangle = -i\hat{H}|\psi(t)\rangle \quad (2.5)$$

where \hat{H} is the operator representing the Hamiltonian. Therefore physical states evolve as $|\psi(t)\rangle = U(t)|\psi(t_0)\rangle$, $U(t) = e^{-i\hat{H}(t-t_0)}$ for time independent \hat{H} . U is unitary and represents evolution of the system under symmetries that preserve the eigenvectors of the operator \hat{H} . \hat{H} represents the Hamiltonian with energy eigenstates thus U is energy conserving. Since the eigenstates of the Hamiltonian are a basis in which any state can be expressed and a unitary preserves these states up to a phase factor, two states which undergo the same unitary will maintain the same overlap,

$$\langle\psi|U^\dagger U|\phi\rangle = \langle\psi|\phi\rangle, \quad (2.6)$$

$$U^\dagger U = \mathbb{I}. \quad (2.7)$$

Since we can also invert the process in time and consider that undoing a unitary also preserve the norm,

$$UU^\dagger = U^\dagger U = \mathbb{I}. \quad (2.8)$$

In computer science, the smallest workable discrete unit of information is the binary digit or *bit* which exists in one of two values, usually labelled as 0 and 1. The quantum version of the bit -the *qubit*- exists in a complex linear superposition of binary states

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

Following the 1st postulate, it is a vector in a 2 dimensional complex Hilbert space and so $\alpha, \beta \in \mathbb{C}$. However from the 2nd and 3rd, we can see that a global phase applied to the state that transforms $|\psi\rangle \rightarrow e^{i\gamma}|\psi\rangle$ would be unobservable. Any state can be aptly uniquely described by

$$|\psi\rangle = |\alpha||0\rangle + e^{i\phi}|\beta||1\rangle, \phi = \arg(\alpha^*\beta).$$

The inner product of the state with itself $\langle\psi|\psi\rangle = |\alpha|^2 + |\beta|^2$ must equal 1; this condition means that our qubit is characterised by two real parameters that can be expressed as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle.$$

Such two dimensional systems are usually physically associated with spin- $\frac{1}{2}$ fermionic particles with spins in opposite directions on the same axis forming $|0\rangle, |1\rangle$ states (hence the factor of a half in the cosine ¹), or photon polarisation.

Now what happens when the number of qubits is increased? As an extra qubit is added the linearity of quantum systems means that the basis states that describe the Hilbert space undergo a product with each of the $\{|0\rangle, |1\rangle\}$ basis element of the new qubit, doubling the space dimensions. So N qubits are described by 2^N basis states- an exponential increase in the number of real parameters that would have to be calculated in an arbitrary system simulation.

Yet as one would glean from the 2nd and 3rd postulates, the results of quantum mechanics are probabilistic and if one has a system divided into qubits one only has to look at the probabilities of two outcomes at each point. Each of the 2^N basis states can correspond to a result of N two outcome independent probabilistic events. A probabilistic Turing machine is more powerful than a purely deterministic one so it seems conceivable that this is one area where access to a probabilistic device provides another improvement.

But this is in fact not the case, the unitary evolution of a quantum system, a $2^N \times 2^N$ matrix of complex parameters acting on the amplitudes of the basis states, allows interferences that no classical device can attempt. Even with just a two qubit system, a state can be formed whose correlations break the Bell inequalities. Ultimately, Bells' theorem shows that as a classical device, a probabilistic Turing machine can never reproduce all the results of quantum mechanics.

What a classical computer could not do, however, Feynman envisioned a quantum computer, quite unlike a Turing machine, could (a conjecture later confirmed by Lloyd [16]). Naturally, getting one quantum system to behave like another specific system may be possible but particularly useful is the idea that a simulator that is universal in the sense that it can simulate any quantum system. The envisioned simulator would

¹See section 2.2.1

be made up of qubit systems that discretise any continuous dimensions with each qubit representing whether that region of the system state space is occupied or not. Evolution of the entire system would be enacted by the interaction of qubits adjacent to each other, matching the lattice models of field theories and solid state physics.

The Bloch sphere

A useful mathematical property of the qubit is that the two parameters of a single qubit state map to a sphere or the surface of a ball: $|\psi\rangle \rightarrow \hat{\mathbf{n}}(\theta, \phi)$ (see figure 2.2), and the unitary operations on the qubit map to rotations of the sphere. Frequently, we will use the picture of the Bloch sphere to describe them. In this representation, two orthogonal states are placed the furthest they can be at opposite ends of an axis through the centre of the sphere, thus they are at an angle π from each other. The computational basis states that corresponds to $\{0, 1\}$ is placed at the poles. An equal superposition for which $|\langle\psi|i\rangle|^2 = \frac{1}{2}$, $i = 0, 1$ is at an angle $\frac{\pi}{2}$ and so the elevation angle θ maps to a superposition amplitude factor $\cos(\frac{\theta}{2})$. The azimuthal angle ϕ provides the relative phase of the superposition.

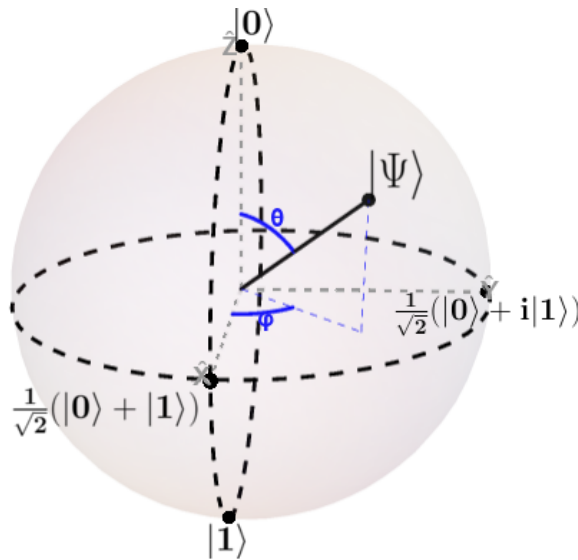


Figure 2.2: $|\Psi\rangle = \cos(\frac{\theta}{2})|0\rangle + e^{i\phi}\sin(\frac{\theta}{2})|1\rangle$ on the Bloch sphere.

This is a handy picture for discussing single qubit unitary gates. As we can ignore global phases on states, we can also ignore scalar phase factors and so represent the group of single qubit unitary gates $U(2)$ with the subgroup of unitaries that have determinant equal to 1, $SU(2)$.

There is a homomorphic map from $SU(2)$ to $SO(3)$, the group of rotations in three dimensions, so we can represent

unitaries as rotations of the Bloch sphere. The norm preserving property of unitaries is represented by the angle preserving properties of rotations; the presentation may also have a physical translation- for a particle spin there is a real space vector that aligns with the field of a measuring device that can be manipulated by rotations of the device or the particle.

The picture also gives insight into how we can decompose a unitary and ascribe a meaning for each of its three parameters. A rotation can be described by a unit vector axis of rotation and a rotation angle. For a unitary operation, the unit vector is given by the pointer from the centre of the Bloch sphere to one of its eigenstates and the angle of rotation by the relative phase imparted on its eigenstates. Operators that commute will be rotations about the same axis.

Given the vector \hat{n} and the angle γ , we can construct the unitary operator

$$e^{-i\frac{\gamma}{2}\hat{\mathbf{n}}\cdot\vec{\sigma}} = \cos\left(\frac{\gamma}{2}\right)\mathbb{I} - i\sin\left(\frac{\gamma}{2}\right)\hat{\mathbf{n}}\cdot\vec{\sigma}, \quad (2.9)$$

where $\vec{\sigma} = \sigma_x\hat{x} + \sigma_y\hat{y} + \sigma_z\hat{z}$, and its matrix representation

$$\mathbf{U} = \begin{pmatrix} \cos\left(\frac{\gamma}{2}\right) - i\sin\left(\frac{\gamma}{2}\right)n_z & -i\sin\left(\frac{\gamma}{2}\right)(n_x - in_y) \\ -i\sin\left(\frac{\gamma}{2}\right)(n_x + in_y) & \cos\left(\frac{\gamma}{2}\right) + i\sin\left(\frac{\gamma}{2}\right)n_z \end{pmatrix}. \quad (2.10)$$

A rotation can also be uniquely decomposed into three rotations where the three axes are fixed by convention and the three Euler angles parametrise the rotation. This means a device only has to actually be able to perform rotations about two axes to implement any arbitrary unitary. Any $U \in U(2)$ can be written as

$$U = e^{i\alpha}R_{\hat{z}}(\beta)R_{\hat{y}}(\gamma)R_{\hat{z}}(\delta) \quad (2.11)$$

for the appropriate choice of $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ where $R_{\hat{n}}(\alpha)$ denotes a rotation about the axis \hat{n} by the angle α . Other pairs of axes can be chosen for a three angle decomposition, including any two non-parallel axes². Occasionally it is important to distinguish gates that are identical save for a global phase, particular when defining control-unitary operations (see the upcoming section 2.2.2), and the choice of notation to distinguish such gates can be confused even within a single textbook, so we will define a notational choice here:

$$R_{\hat{n}}(\gamma) = e^{-i\frac{\gamma}{2}\hat{\mathbf{n}}\cdot\vec{\sigma}} \quad (2.12)$$

$$R_{\hat{z}}(\gamma) = \begin{pmatrix} e^{-i\frac{\gamma}{2}} & 0 \\ 0 & e^{i\frac{\gamma}{2}} \end{pmatrix} \quad (2.13)$$

$$Z^\alpha = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha\pi} \end{pmatrix} \quad (2.14)$$

$$Z_\gamma = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\gamma} \end{pmatrix} \quad (2.15)$$

2.2.2 The Universal Quantum Computer

It had been argued by Feynman that an arbitrary quantum system could not be efficiently simulated by a Universal Turing Machine. This undercuts the idea of the

²See appendix A

Church-Turing hypothesis as a mathematical principle separate from physical considerations of the machinery since the local, classical nature of the machine limits its computational capability.

In 1985, David Deutsch considered that the problem with the hypothesis lay in part in its vagueness about what is computation and its lack of grounding in the physical. Physical principles can be tested and falsified by objective measures. Deutsch rewrote the Church-Turing hypothesis in a more physical principle: *Every finitely realizable physical system can be perfectly simulated by a universal model computing machine operating by finite means* [2].

This is falsifiable: a physical model of a system that solves a problem could turn up a result that disagreed with a mathematically modelled machine. On the other side of the coin, the statement is not so trivial one could just take it as an *a priori* statement; if a conceived mathematical function could not be implemented on a physical device, it would not be possible to compute it in a practical sense and if a physical system was dependent on a function that did not appear possible on a Turing machine, that function could be performed by the physical system. But if the physical Church-Turing (PC-T) principle holds true then this case could not occur.

However this PC-T principle derogates any classical Universal Turing Machine because it can not simulate a quantum system or even technically a continuous system (requiring successive discrete approximation breaks the strictures of Deutsch's principle [2]). It requires that one devise a universal quantum computer that can simulate any physical system (obviously now including the quantum ones) when programmed according to a model included within the computer (this distinguishes it from just a simulator).

Deutsch's Universal Quantum Computer (UQC) model is a quantisation of a Universal Turing Computing Machine with the expected shared features. The infinitely long tape divided into sections is now an infinite sequence of 2 state observables i.e. qubits:

$$\hat{m} = \{\hat{m}_i\}, i \in \mathbb{Z}, \quad (2.16)$$

with eigenstates

$$|\mathbf{m}\rangle = |\dots m_{-1}, m_0, m_1, \dots\rangle, m_i \in \{0, 1\}. \quad (2.17)$$

The processor still has a description as a finite number of configuration states that will now, as with the memory tape, be expressed in terms of M qubits with 2^M configuration states:

$$|\mathbf{n}\rangle = |n_0, n_1, \dots n_{M-1}\rangle, n_i \in \{0, 1\}, i \in \mathbb{Z}_{M-1}. \quad (2.18)$$

The UTM would only address each section of the tape one at a time and could only move by one segment. This limitation is necessary to reflect the "finite means" of the

PC-T but is handled differently in the UQC since the location of any pointer onto the tape can be considered an observable itself $|x\rangle$, $x \in \mathbb{Z}$. So the state of the UQC system is in a Hilbert space spanned by the eigenvalues

$$|x, \mathbf{n}, \mathbf{m}\rangle = |x, n_0, n_1, \dots, m_0, \dots\rangle. \quad (2.19)$$

These eigenvalues give the “computational basis states”. Rather than the configuration of the processor switching conditions on the state of the memory tape at the current address, the states of $|\mathbf{m}\rangle$ and $|\mathbf{n}\rangle$ are coupled by a grand unitary operator U . Since the PC-T principle requires that the system evolve in finite steps, the unitary U is treated as a constant Hamiltonian applied in a fixed time interval T . The time evolution of the system according to the 5th postulate is

$$|\Psi(rT)\rangle = U^r |\Psi(0)\rangle, \quad r \in \mathbb{Z}^+, \quad (2.20)$$

$$|\Psi(0)\rangle = \sum_m \lambda_m |0, \mathbf{0}, \mathbf{m}\rangle. \quad (2.21)$$

The elements of U are also restricted to the form:

$$\langle x', \mathbf{n}', \mathbf{m}' | U | x, \mathbf{n}, \mathbf{m} \rangle = [\delta_{x'}^{x+1} U^+(\mathbf{n}', m'_x | \mathbf{n}, m_x) + \delta_{x'}^{x-1} U^-(\mathbf{n}', m'_x | \mathbf{n}, m_x)] \prod_{y \neq x} \delta_{m'_y}^{m_y}. \quad (2.22)$$

The delta operator terms ensure that only the x th memory qubit interacts in a single step and x only changes by at most one. Other than that, a particular quantum computer is defined by the choice of the $U^\pm(\mathbf{n}', m'_x | \mathbf{n}, m_x)$ terms. Alternatively, visualise the quantum computer as a sequence of discrete operations that connect the processor state $|\mathbf{n}\rangle$ with individual qubits $|m_x\rangle$ (see figure 2.3). The quantum computer is defined by the unitary operations that appear in the sequence and the ordering of the sequence.

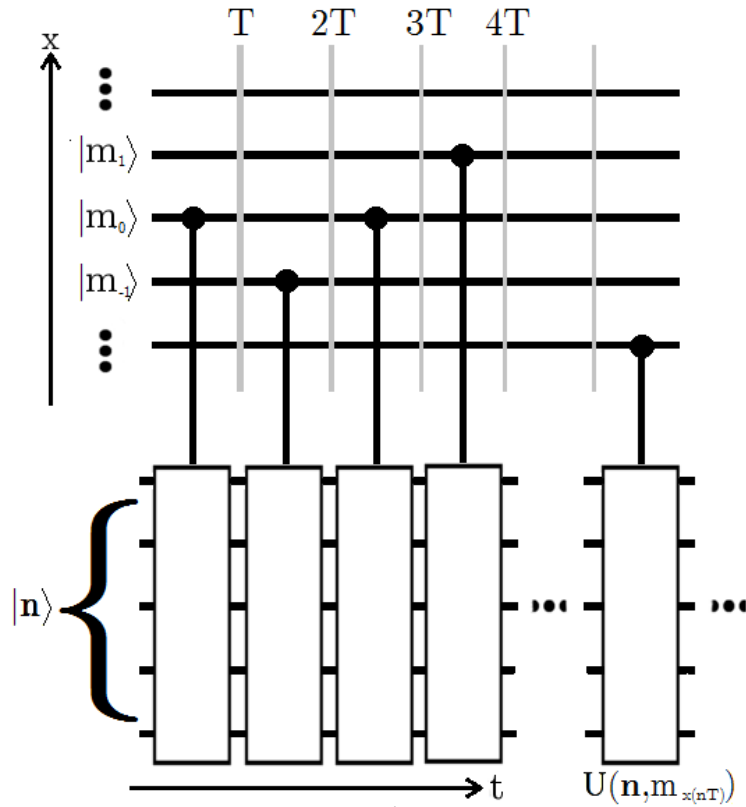


Figure 2.3: A decomposition of Deutsch’s universal quantum computer: A tape register of infinite qubits $\{|m_x\rangle\}$, a finite machine state $|n\rangle$ and grand unitary operator U act “by finite means” therefore (1) the grand unitary acts on a finite subsystem, which Deutsch selects to be a finite machine state and a single register qubit in a single time step (2) the operation depends only on the state of a finite subsystem, also the finite machine state and single register qubit (3) the motion can be broken into finite steps. This is interpreted as a series of controlled unitary operations acting in integers of time units, T .

Unitary evolution is reversible and a consequence of a reversible computer is that two consecutive states after a non-trivial computation are never identical [2]. One would have to observe the system to know if it was done rather than it being halted in a constant state but a measurement would collapse a quantum system. So one qubit in the processor, $|n_0\rangle$, is set aside to signal the end of computation so that ‘not done’ or ‘done’ can be measured on a subsystem that only ever exists in one of $|0\rangle$ or $|1\rangle$ respectively.

To construct a quantum computer that would be universal for all quantum computers, Deutsch [2] first allowed a quantum computer the ability to perform the operations of a UTM. While Deutsch left the specifics up to the reader, it would involve elements of the UQC grand unitary being restricted to binary so these operations are restricted

to transformations between computational basis states. On top of that a universal quantum computer needs to be able to perform operations that produce superpositions of computational basis states.

A classical program on a memory tape will cause the computer to write a function of an input at a designated position on the tape onto another, possibly different, position: the program $\pi(f, a, b)$ is that of a function f which detects state j at slot a and writes $f(j)$ at slot b . Since unitary operations are reversible, a qubit state can not just be written over but the program can be made so that the ‘read’ and ‘write’ qubit is the same and it is just having a single qubit unitary operation applied to it,

$$|m\rangle = |\pi(V, 2), j\rangle \rightarrow \sum_k \langle k|V|j\rangle |\pi(V, 2), k\rangle. \quad (2.23)$$

This is in effect similar to a control unitary. A control unitary is a unitary over two qubits in which a unitary operation is enacted on a qubit (the *target*) if another qubit (the *control*) is in a $|1\rangle$ state but not if the control is in $|0\rangle$:

$$CV = |0\rangle_C \langle 0| \otimes \mathbb{I}_T + |1\rangle_C \langle 1| \otimes V_T. \quad (2.24)$$

This can be extended to many qubit controls where the targeted unitary is performed according to one particular multiqubit state:

$$C^{\otimes n}V = (\mathbb{I} - |m\rangle\langle m|)_C \otimes \mathbb{I}_T + |m\rangle_C \langle m| \otimes V_T. \quad (2.25)$$

Such control unitary operations are usually graphically depicted with a dot on the control with a line leading to the unitary on the target as in figure 2.3. The UQC is performing a controlled operation where the selection of the control and target qubits is mediated through the configuration states of the processor and the grand unitary. Now attribute the UQC the ability to perform a finite set of single qubit unitaries $\{V_i\}$ in this way and that concatenations of gates in this finite set can be used to construct any single qubit unitary (the unitary set Deutsch [2] uses for this will be described in section 2.3.1). Given the ability to program these unitaries in this way, conditioned on computational basis states, and with the ability to permute the computational basis states, Deutsch showed by induction that the UQC could compose an approximation of any finite unitary. By being able to do this, any quantum computer $Q(U^+, U^-)$ could be simulated by performing the finite unitary operations of its step by step decomposition thus proving that the UQC is universal for quantum computation.

Features of the Universal Quantum Computer

Once one has the concept of a quantum computer, the question is what makes it unique? Note the goal of aligning a universal computer with the Physical Church-Turing principle, no real system is perfectly unitary but the UQC exists to enact unitary

operations, even if it does so by approximation. Due to the laws of thermodynamics no real system can be created in a pure quantum state, closed off from interactions from any outside energy systems and a UQC would need to be able to simulate this even though the UQC model is only described through pure unitary evolution. It can use the randomness generated by measurement of its own components. Measurements of a quantum system can produce true random numbers for any arbitrary irrational number probability so it can prepare arbitrary mixed states. Classical computers may use probabilistic systems but using classically generated probabilities to simulate a quantum system may mean that the statistics of it can start to be described by a local hidden variable theory and thus certain features of quantum mechanics, such as the violation of Bell's inequalities, are lost.

Quantum computations may also perform algorithms that are faster than any know classical algorithm for the same problem. Deutsch [2] first considered this possibility from the feature of quantum parallelism. If a memory qubit is first put into a superposition of N possible states $|j\rangle$, $j = 0, 1, 2, \dots, N - 1$, with a program $\pi(f, a)$,

$$|\Psi\rangle = \sum_j \frac{1}{\sqrt{N}} |\pi(f, a), \dots, j\rangle \rightarrow \sum_j \frac{1}{\sqrt{N}} |\pi(f, a), \dots, f(j)\rangle,$$

then the N possible values of $f(j)$ have in some sense been computed. However, they also need to be extracted a process that is probabilistic and was shown to always have an expectation time equivalent to N serial computations. What may be exploitable is the ability of the probabilistic time to have a minimum or maximum lower than the classical equivalent.

Early proof of quantum-assisted speed-up is given by the algorithm Deutsch [2] was able to devise for a quantum computer that made a computation exponentially faster than any deterministic classical algorithm yet it did not beat the possible probabilistic classical algorithm. However, later, Deutsch and Jozsa [17] developed an algorithm for a quantum computer that solved a problem more efficiently than either a deterministic or probabilistic classical computer and gave a result in deterministic time. The Deutsch-Jozsa algorithm solved in deterministic polynomial time what a deterministic classical computer could only solve in exponential time and gave evidence of a difference in the complexity classes of such. Yet the problem the algorithm solved was devised by Deutsch and Jozsa for the purpose of demonstrating this power of a quantum computer. Several other problems in a similar vein were constructed that found further improvements over classical algorithms [18, 19] but when Peter Shor developed an algorithm [21] for integer factoring in polynomial time, it was an example of an already well known and established problem widely studied and assumed to not have an efficient classical solution. Bernstein & Vazirani [72] applied a complexity class approach to the matter to consider that the class of problem with polynomial bounded error solutions

on a UQC, **BQP**, was larger than that for a PTM, suggesting a quantum computer is not just a device for resolving the physical Church-Turing principle with quantum simulation, it underpins a complexity class substantially different than that of a classical Turing machine.

What the Physical Church-Turing principle and the Universal Quantum Computer argue for is that it is the distinguishing physical features of the theory of quantum mechanics that provides differences in the computability of a class of calculations.

2.3 Schemes for Quantum Computation

2.3.1 The Quantum Gate Circuit

The universality of the Universal Quantum Computer is based in its ability to perform any unitary operation on a finite dimensional quantum system. It is possible to create a model that is in a similar sense universal with a description of three basic parts, the input, the unitary operation on the input and the output, without referring to the mechanics of how the machine processor implements the unitary operations.

In the Universal Quantum Computer, a set of quantum unitary operations along with Turing machine operations such as permutations of the computational basis were applied in discrete steps according to the finite means of the model and then the concatenation of these operations built up other unitary operations. In 1989, Deutsch [22] described a quantum computational network as a quantisation of the logic gate model of computation where discretely applied unitary operations acted as quantum logic gates that could be applied in a gate circuit/computational network.

A gate is a computing machine that performs a specific computation on a fixed size input and output in a fixed amount of time. This makes a gate an effective component, a cog in a machine, because they turn in discrete units and halt at a fixed time and applying several gates together across different inputs yields a larger more intricate gate.

A *logic* gate acts on binary inputs and is described by a logic table of what inputs lead to what outputs. This may also be expressed as a matrix acting on a 2^d vector for d bits. A discrete unitary operation has the properties of a gate. A quantum gate can be formed by running a fixed Hamiltonian for a fixed period of time over a finite number of qubits, as with the discretised time evolution of the grand unitary U^{nT} in the Universal Quantum Computer. These form a group with a matrix representation so they can be concatenated and overlaid on multiple qubits to form a larger gate. The differences are that they can only be reversible operations and the inputs and outputs can be linear superpositions of the binary states with complex amplitudes. The Turing machine models discussed previously were not reversible designs but it was shown by Bennett

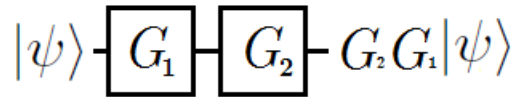


Figure 2.4: A one qubit wire. Gates labelled G_j apply unitary operation G_j successively in left-right order.

[73] that a Turing machine would have a reversible machine equivalent implying that universal classical computing and the classical complexity classes can all be obtained by reversible processes. The development of reversible logic gate networks by Fredkin and Toffoli [71, 74] to that end is a precursor to the quantum computational networks that subsume it.

The gates themselves are connected in a gate circuit or “quantum computational network” where gates are sorted into different computational steps so that the timings are synchronised. This is all equivalent to a single gate until sources and sinks for qubits are introduced to add and remove them without measurement. With these, the circuit can enact the non-unitary operations and mixed states that simulate interaction with an external system.

Graphical depiction of a quantum gate circuit

The circuit is graphically represented by a wire along which the information carried by a qubit passes through. Physically this may actually be occurring with the physical carrier moving through a wire or waveguide but the carrier should not be confused for the qubit itself. The constituent gates will be represented by squares along the wire (see figure 2.4). A gate that acts across multiple qubits will be overlaid on top of the multiple wires. Typically qubits are considered to be laid out in a fixed order and interactions are kept between adjacent qubits and thus only gates on adjacent wires have to be displayed on the circuit diagram. However on occasions where a gate has to be placed over non-adjacent wires, the gate diagram will make use of a connecting line between the wires with a node placed on specific wires to indicate which qubits are actually connected; a bridge loop may also be used to indicate that a qubit is being skipped over (see figure 2.5 and Table 2.3 for examples).

There are certain key gates and types of gate it is important to be familiar with. The Pauli spin operators σ_x , σ_y , and σ_z , being discrete unitary operations with a matrix representation, are naturally quantum gates, normally labelled as X , Y and Z . The X gate is notable for performing the same operation as a bit inverter on simple binary states so it also known as the NOT gate or the bit-flip gate. See table 2.1.

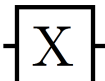


Name	Operation	Matrix	Circuit element
X/NOT	$ 0\rangle\langle 1 + 1\rangle\langle 0 $	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	
Y	$-i 0\rangle\langle 1 + i 1\rangle\langle 0 $	$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$	
Z	$ 0\rangle\langle 0 - 1\rangle\langle 1 $	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	

Table 2.1: Representations of the Pauli operators.

An important class of two-qubit gates are control-unitary gates or just “control gates”. For these gates one qubit is the *control* and the other the *target*. For a gate labelled CU the operation U will be enacted on the target if the control is in the basis state $|1\rangle$ or the identity operation if the control is in the basis state $|0\rangle$. The control is designed by the use of a black node on the wire. These are important because this operation, when used with a control in a superposition of the basis states and with an appropriate U for the state of the target, can produce entanglement. Of these control gates of most note are the $CNOT$ and CZ gates. The CZ is symmetrical with respect to control and target so it actually has a symmetrical gate symbol. See table 2.2.

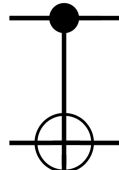
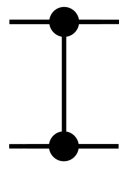
Name	Operation	Matrix	Circuit element
$CNOT$	$ 0\rangle_1\langle 0 _1 \otimes \mathbb{I}_2 + 1\rangle_1\langle 1 _1 \otimes X_2$	$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 0 & 1 \\ & & 1 & 0 \end{pmatrix}$	
CZ	$ 0\rangle_1\langle 0 _1 \otimes \mathbb{I}_2 + 1\rangle_1\langle 1 _1 \otimes Z_2$	$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}$	

Table 2.2: Representations of the $CNOT$ & CZ gates.

The control gates can also be generalised to multiple qubits in two ways. One is that the target can be a multi-qubit system and it enacts a multi-qubit unitary. It can also have multiple control qubits where the control condition is the conjunction of the two qubit systems i.e. each control qubit must be in the $|1\rangle$ state. In a way it is a control gate where the target unitary is also a control gate. A notable example is the Toffoli gate a.k.a. the $CCNOT$ gate that is universal for logic gates. See table 2.4.

Name	Operation	Matrix	Circuit element
$CZ_{13} \otimes \mathbb{I}_2$	$ 0\rangle_1\langle 0 _1 \otimes \mathbb{I}_{23}$ $+ 1\rangle_1\langle 1 _1 \otimes Z_3 \otimes \mathbb{I}_2$	$\begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & -1 & & \\ & & & & & & 1 & \\ & & & & & & & -1 \end{pmatrix}$	

Table 2.3: Representations of the CZ gate on non-adjacent qubits

Name	Operation	Matrix	Circuit element
Toffoli or $CCNOT$	$(\mathbb{I} - 11\rangle\langle 11)_{12} \otimes \mathbb{I}_3$ $+ 11\rangle\langle 11 _{12} \otimes X_3$	$\begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & 0 & 1 \\ & & & & & & 1 & 0 \end{pmatrix}$	

Table 2.4: Representations of the Toffoli gate.

We will also represent the gates CZ_λ in a manner similar to the CZ gate with symmetric black nodes and a label λ . The node and label may also be used on just one qubit to represent Z_θ . See table 2.5. Another important single qubit gate is the Hadamard gate, labelled H . It turns a computational basis state into a equal superposition of basis states with phase ± 1 . It transforms the action of the X gate to a Z gate, $HXH = Z$, and so also can be used to transform a $CNOT$ to a CZ .

Measurements are commonly displayed in different literature using one of two graphical representations. The first is based off of the display of analogue gauges and the second, which is preferred by this thesis, is the D shaped detector symbol based off of the shape of conical detectors. On their own these symbols mean measurement in the computational basis with gates beforehand being used to select the measurement basis however we may label the detector appropriately to indicate a change of basis. See table 2.6.

We will also refer to the $SWAP$ gate. The action of this gate is simply to swap the

Name	Operation	Matrix	Circuit element
CZ_θ	$ 0\rangle_1\langle 0 _1 \otimes \mathbb{I}_2 + 1\rangle_1\langle 1 _1 \otimes Z_{\theta,2}$	$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & e^{i\theta} \end{pmatrix}$	
Hadamard	$ +\rangle\langle 0 + -\rangle\langle 1 $	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$	

Table 2.5: Representations of the CZ_θ and Hadamard gate.

Name	Operation	Matrix	Circuit element
Measurement	$\{ 0\rangle\langle 0 , 1\rangle\langle 1 \}$	$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$	

Table 2.6: Representations of a measurement.

states encoded on two qubits. This might represent an actual physical movement of the qubit carriers, a swapping of information between two carriers (which can be done by performing three consecutive $CNOT$ gates with the target and control swapped for the second one) or just as a way of getting around displaying a gate across non-adjacent wires.

Name	Operation	Matrix	Circuit element
SWAP	$ 01\rangle\langle 10 + 10\rangle\langle 01 + 00\rangle\langle 00 + 11\rangle\langle 11 $	$\begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{pmatrix}$	

Table 2.7: Representations of the SWAP gate.

Quantum Gate Circuit Universality

Gate based quantum computation is a useful picture for considering how to construct a universal quantum computer in terms of a small finite number of different components. Not only can a UQC implement a gate circuit, gate circuits could be used to approximate the UQC up to an arbitrarily large finite size [22] so they are computationally equivalent.

The UQC described by Deutsch [2] evolves according to a particular unitary but

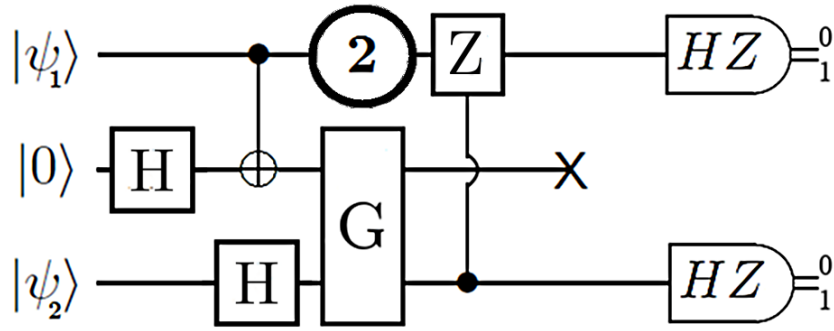


Figure 2.5: An example circuit using the elements of a quantum computational network. A qubit is introduced on each wire on the left with computation running rightwards. A source inputs a $|0\rangle$ state on the centre wire. Gates acting on multiple qubits can cross over adjacent wires. To account for gates taking multiple steps, n steps of wire can be indicated by an encircled n . \times represents a sink where the centre qubit is lost. Measurement of the inputs is indicated by the D style detectors with a classical output. A rotation of the measurement basis may be indicated by either a gate before the measurement or by labelling of the detectors.

what are the other unitary operations would make a universal machine? Well, since the UQC is universal because the unitary enacts some particular unitary gates $\{V_i\}$ on the memory tape when in a particular computational basis state, a quantum gate circuit model is universal if it can implement those unitary gates. If the UQC unitary changes, the gates $\{V_i\}$ change so the question is whether those gates can be combined to produce any arbitrary finite unitary when placed in a network. So the question equates to: what gates can be used to create any finite dimensional unitary when combined in a gate circuit?

2.3.2 Universal Finite Gate Sets

Say there are two sets of logic gates \mathcal{G} and \mathcal{L} , if there exists a logic gate circuit which can construct every element in \mathcal{L} using only gates from \mathcal{G} , then \mathcal{G} can be said to be *adequate* for \mathcal{L} [22]. For quantum gates, a similar property is desirable where there exists a finite set of gates $\{V_i\}$ that can be used to construct any gate in the set of gates of finite dimension; since such a set allows one to construct a universal quantum computer, it can be referred to as a universal set. The primary difference from an adequate set of logic gates is that while the universal set has a finite number of elements, it is being used to generate a continuous group. This will not work if the aim is to exactly produce a gate but it is possible to use a finite set of elements to generate gates that are arbitrarily

close approximations to any gate in a continuous group.

Let us take the NOT gate or X gate as an example desired gate, $\mathcal{L} = \{X\}$, and then a finite set with one element, $G = \{X^\alpha\}$, for $0 < \alpha < 1$. Given that any rational value of α under modulo 1 will just need an integer multiplication to equal 1, the problem we face is an irrational α ;

$$(X^\alpha)^m = X^{m\alpha - 2\lfloor \frac{1}{2}m\alpha \rfloor}. \quad (2.26)$$

There is no integer m such that $m\alpha - 2\lfloor \frac{1}{2}m\alpha \rfloor = 1$ but there is always m for which $|1 - m\alpha - 2\lfloor \frac{1}{2}m\alpha \rfloor| < \epsilon$ for any arbitrarily small real ϵ . However since the laws of thermodynamics prevent a truly completely isolated and unitary system from forming, no real perfect gates, an approximation can be considered equivalent to the real thing.

Given a unitary represented on the Bloch sphere by a rotation $R_{\hat{n}}(\theta)$ where the rotation angle θ is an irrational multiple of π , multiple applications of this unitary can be used to approximate all the gates that correspond to rotations about that same axis (a brief proof of this can be found in [75]). If we look back at the gates used to construct single qubit unitary operations in the Deutsch UQC model, $\{V_i\}$:

$$\begin{aligned} V_0 &= \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} = V_4^\dagger, & V_1 &= \begin{pmatrix} \cos(\alpha) & i\sin(\alpha) \\ i\sin(\alpha) & \cos(\alpha) \end{pmatrix} = V_5^\dagger, \\ V_2 &= \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & 1 \end{pmatrix} = V_6^\dagger, & V_3 &= \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix} = V_7^\dagger, \end{aligned}$$

where α here is also an irrational multiple of π . There are also two gates that correspond to rotations by $\frac{\pi}{2}$ that are not essential additions for universality but are added for convenience:

$$V_8 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad V_9 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}.$$

In this set, there are gates that correspond to rotations by irrational multiples of π about each Cartesian axis, to provide rotations about each axis to angle, as well as their inverse, the ability to apply global phases and exact gates for some rational multiples of π . Clearly, by the above argument this is adequate for single qubit gates but also contains redundancies. In the context of quantum gate circuits, on the other hand, Deutsch [22] considered the other extreme, sets that consisted of only $|0/1\rangle$ bit sources and gates $\{\mathbb{I}, U\}$ for a single U so that U itself could be deemed a universal gate.

The Deutsch gate

A universal gate for logic gates is the Toffoli[76] gate which can also be called the *CCNOT* gate:

$$CCNOT = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & 0 & 1 & \\ & & & & & 1 & 0 & \end{pmatrix}. \quad (2.27)$$

The universal Deutsch gate is effectively a quantisation of the Toffoli gate to $CCR_{\hat{x}}(\pi\alpha)$ where α is irrational [22]:

$$Q = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & i\cos(\frac{1}{2}\pi\alpha) & \sin(\frac{1}{2}\pi\alpha) & \\ & & & & & \sin(\frac{1}{2}\pi\alpha) & i\cos(\frac{1}{2}\pi\alpha) & \end{pmatrix}. \quad (2.28)$$

This provides universality with respect to all gates of the form $CCR_{\hat{x}}(\lambda)$ including the Toffoli gate which provides the ability to perform all logic gates. The logic gates perform permutations of basis states and Deutsch [22] showed that the permutation of different pairs in combination with the Deutsch gate also allowed it to generate gates of the form of $CCR_{\hat{y}}(\lambda)$, $CCR_{\hat{z}}(\lambda)$ and CCZ_{λ} .

With these gates, one can construct a unitary operation that evolves any $|\psi\rangle = \sum_{n=0}^7 c_n |n\rangle$ such that

$$|\psi\rangle \rightarrow \sum_{m=0}^7 c_m = 0^5 c_n |n\rangle + \sqrt{|c_6|^2 + |c_7|^2} |7\rangle, \quad (2.29)$$

thus the coefficient of one basis state has been set to zero. Since permutations of basis states are also possible, this can be done for each basis state until the gate U_{ψ} is formed for which

$$U_{\psi} |\psi\rangle = |7\rangle. \quad (2.30)$$

So by Deutsch's proof, the spectral decomposition of a unitary gate is taken, $\{|\psi_j\rangle\langle\psi_j|\}$ and each $|\psi_j\rangle$ is used to construct $8 U_{\psi_j}$. The unitary gate is then constructed from

$$U = \prod_{n=0}^7 U_{\psi_j} CCZ_{\lambda_j} U_{\psi_j}^{\dagger}. \quad (2.31)$$

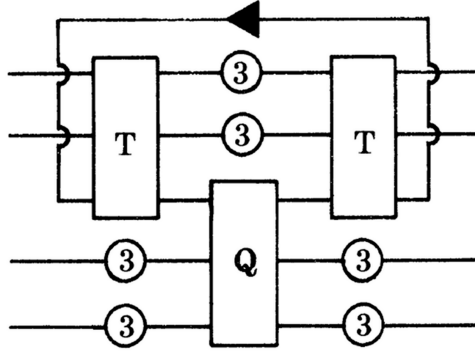


Figure 2.6: The gate circuit for a 4 qubit universal quantum gate from [22] built from the Deutsch gate, Q , (2.28) and Toffoli gate, T , (2.27). The loop represents an ancillary input that inputs and outputs in the state $|0\rangle$ with each use of the gate; alternately consider a fresh $|0\rangle$ state input on the third wire each time.

This constructs the unitary gate $U = \sum_j e^{i\lambda_j} |\psi_j\rangle\langle\psi_j|$ for any $U \in U(8)$, proving that the Deutsch gate is universal for three qubit gates.

Deutsch[22] also showed that it was possible to construct a gate circuit (see figure 2.6) that was equivalent to a gate universal for four qubit gates and to iterate the construction process to higher dimensions. Thus the Deutsch gate is universal for all finite n qubit unitary gates.

Gates that can be turned into the Deutsch gate

Just as the three qubit Deutsch gate can be shown to be capable of constructing gates that are universal in higher dimensions, one can find two qubit gates that will be capable of constructing a Deutsch gate. DiVincenzo [77] found four two qubit gates that can construct the Deutsch gate together. Moreover, DiVincenzo [77] expressed the combination of the four gates in group theoretic language as the use of two generators in a Lie algebra to create a third generator. Four gates in Deutsch's original proof were identified as unitaries of four generators that could, according to the proof, generate the group of three qubit unitary gates. DiVincenzo [77] then showed how these four generators could be constructed from generators of the two qubit gate algebra.

Soon after Barenco [29] and Sleator & Weinfurter [30] showed in independent but highly similar ³ arguments that a single two qubit gate could be universal by reconstructing the Deutsch gate with it. The gates can be thought of as the two qubit equivalent to the Deutsch gate. Rather than a double control gate, it is a single control

³They both suggest the use of cavity QED type interactions as an implementation

gate CU where, using Barenco [29]'s notation,

$$\mathbf{U} = \begin{pmatrix} e^{i\alpha}\cos(\theta) & -i e^{i(\alpha-\phi)}\sin(\theta) \\ -i e^{i(\alpha+\phi)}\sin(\theta) & e^{i\alpha}\cos(\theta) \end{pmatrix} \quad (2.32)$$

where ϕ , α and θ are fixed irrational multiples of π . Note that while a global phase factor of $e^{i\alpha}$ can be pulled out of the description of the unitary $U(\phi, \alpha, \theta)$, in a control-unitary gate this induces a relative phase difference between the computational basis of the control qubit. So the presence of this term is in fact equivalent to a single qubit relative phase gate with irrational multiple of π for a phase which forms a single qubit version of this gate and would be capable of approximating any single qubit relative phase gate. The $CU(\phi, \alpha, \theta)$ gate can be used to construct $CCU(\phi, \alpha, \theta)$ over three qubits and $CCR_{\hat{z}}(\beta)$ where

$$\mathbf{R}_{\hat{z}}(\beta) = \begin{pmatrix} e^{i\beta} & 0 \\ 0 & e^{-i\beta} \end{pmatrix} \quad (2.33)$$

for some β which enables construction of the Deutsch gate. The slightly different argument in [30] found a five gate decomposition of the Deutsch gate (which turns out to be an optimal number of gates [78]) that uses gates in the class $\{CU^m\}$.

A two qubit gate forms a more physically natural approach than a three qubit gate since most physical interactions will occur over two qubit systems. There are also exceptional rational multiples of π that can be used as parameters for $U(\phi, \alpha, \theta)$ such as $(\frac{\pi}{2}, \frac{\pi}{2}, \theta)$ that are also universal and so can be more easily experimentally expressed.

Deutsch *et al* [31] and Lloyd [79] showed that in fact almost any two qubit gate is universal for two qubit unitary gates (and therefore is also universal for n -qubit unitary gates). One can return to the Lie algebra arguments brought up by DiVincenzo and look at a generator \hat{H}_1 , its corresponding unitary operation $e^{-i\hat{H}_1}$ and the effect of the Swap action upon it.

The universality of almost any two qubit gate

First understand how two gates of a unitary group $U(n)$ can be combined to implement any gate of the same dimension n : each gate U relates to a Hamiltonian \hat{H} by $U = e^{i\hat{H}t}$ for some appropriate t . Typically t is a time parameter under our control. Even if not, just as in Deutsch's proof in [22], if one has a unitary that is an irrational root of the identity then one would be able to implement any gate of the same Hamiltonian for any value of t .

Given two different Hamiltonians, one can apply a sequence that generates a gate related to the commutator of two Hamiltonians $\hat{H}_3 = i[\hat{H}_1, \hat{H}_2]$:

$$e^{[\hat{H}_1, \hat{H}_2]t} = \lim_{n \rightarrow \infty} (e^{-i\hat{H}_1\sqrt{\frac{t}{n}}} e^{-i\hat{H}_2\sqrt{\frac{t}{n}}} e^{i\hat{H}_1\sqrt{\frac{t}{n}}} e^{i\hat{H}_2\sqrt{\frac{t}{n}}})^n. \quad (2.34)$$

The same rule can be applied to the commutator $[H_1, H_3]$, $[H_1, H_4]$ etc. until n^2 such Hamiltonians are simulated. If those n^2 H_j are linearly independent, then they can form a decomposition of any n qubit unitary (though not necessarily the most efficient one). This will happen with non-identity H_1 and H_2 unless one of them lies on a submanifold of $U(n)$ of lower dimension which excludes, to paraphrase Lloyd [79], almost all gates.

Deutsch *et al* [31] and Lloyd [79] then point out that if one is given a single gate U_1 and the ability to swap the qubit inputs, represented by the matrix

$$SWAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2.35)$$

U_2 can be created from $SWAP.U_1.SWAP$. Clearly this imposes some restrictions on U_1 , not least an absence of swap symmetry so that $[H_1, SWAP] \neq 0$. It also must not be in the subgroup of unitary gates which can be expressed as single qubit gates on each qubit, $U(2)_1 \otimes U(2)_2 \otimes \dots U(2)_{\log_2 n}$.

Two qubit gates that are not universal for the group of all two qubit unitary gates have been exhaustively characterised by these aspects by Childs *et al* [80]. A two qubit Hamiltonian H is universal for $U(4)$ if and only if it does not satisfy:

- H is $SWAP$ -similar to a local Hamiltonian, $H_{d=2} \otimes H_{d=2}$,
- H shares an eigenvector with $SWAP$,
- $\text{Tr}[H] = 0$ ⁴,

where two matrices A and B are $SWAP$ -similar if there exists a unitary matrix P such that $B = PAP^\dagger$ and $[P, SWAP] = 0$.

This leaves open the question of which gates are universal over more than two qubits since it is known that there are gates which are not universal for two qubit gates but which are universal for three qubit gates [80]. There is also a difference between being universal on just two qubits and being universal on those two qubits when we can temporarily extend the register of qubits with ancillary qubits. If one can apply a two qubit gate over different permutations of three qubits (with the ancilla possibly being prepared in a particular state) then one may be able to produce a gate $U_{12}^{d=4} \otimes U_a$ that implements a universal two qubit gate on the main qubit pair and a local gate on the register (but still may not be universal over three qubits).

⁴For $H' = H - \frac{1}{d}\text{Tr}[H]$, $H'H'^\dagger = H'^\dagger H' = \mathbb{I}$ iff $\text{Tr}[H] = 0$

Two level unitaries are universal

While a single qubit may be the smallest information unit and while quantum computer designs usually look at systems built out of qubits, it is not always best to think of operations as applying over n qubits when dealing with larger dimensional operations. A two qubit gate or three qubit gate may be applied across different permutations of qubits to create an n -qubit gate that does not decompose into two or three qubit subsystems but single qubit gates will only ever make local single qubit operations. However say there is an n dimensional system, $n = 2^k$, with dimensions labelled $|0\rangle, |1\rangle, \dots |n-1\rangle$. There can be a two level unitary on $|1\rangle$ and $|2\rangle$

$$U \sum_{m=1}^n c_m |m\rangle = c_0 |0\rangle + \sum_{m=3}^{n-1} c_m |m\rangle + \sum_{l=1,2} c'_l |l\rangle \quad (2.36)$$

that can be mapped to a one qubit gate yet the two levels may not necessarily be in the same qubit subspace e.g. if the system was built out of k qubits with each dimension number being assigned to a state by binary association: $|0\rangle = |0^{\otimes k}\rangle$, $|1\rangle = |0^{\otimes k-1}, 1\rangle$, $|2\rangle = |0^{\otimes k-2}, 1, 0\rangle, \dots |n\rangle = |1^{\otimes k}\rangle$ then levels 1 and 2 are not in a shared qubit subspace and generally no two non-adjacent levels are on a single qubit.

In photonic systems, rather than introducing multiple qubits, a qu- d -it system of d dimensions can be constructed from a single photon superposition across d modes and a two level unitary can be performed by beam splitters shared by two waveguides. It is in this experimental context that Reck *et al* [81] showed that any $n \times n$ unitary can be decomposed into two level unitaries that can act on any two levels. In practice, switching non-adjacent levels over intermediate waveguides to form a beam splitter is not a simple design feature and does not translate into those experimental systems that are comprised of multiple qubits well. With a two qubit example, levels $|2\rangle$ and $|3\rangle$ correspond to the anti-correlated qubit subspace $|01\rangle, |10\rangle$ so to implement such a two level unitary, the involvement of a two qubit entangling operation is a natural expectation.

CNOT and single qubit unitaries are universal

Barenco, Bennett *et al* [32] explored how a set of gates that consists of all single qubit gates, $U(2)$, and the CNOT gate is universal for all finite unitary gates. Such a set can be used to construct any CU , CCU or n -bit controlled $C^m U$ gate with the number of CNOT gates independent of U . In the extension to $n > 3$, they were able to express the procedure of Reck *et al* [81] in terms of CNOT gates and $U(2)$ gates through Gray codes [32]. This procedure is not efficient due to exponential scaling with n but the universality provokes consideration of the possibility of applying only one two-qubit gate and leaving the problem of adequately approximating an entire group to operations on

only a single qubit. Since the characterisation and implementation of high fidelity gates is costly, reducing the need to a single interaction is useful. Plus, since it does allow for universal two qubit and three qubit gates, future improvements in the decomposition of $U(n)$ are still available to this set.

In fact, any entangling gate, according to Bremmer *et al* [33] can be used in place of CNOT. This proof relies on the fact that CNOT can be transformed by local unitary operations to $e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z}$. An entangling gate V can be transformed to $W = (\mathbb{I}\otimes Z)V(\mathbb{I}\otimes Z)V = e^{i\alpha\sigma_z\otimes\sigma_z}$. Even if α is a *rational* multiple of π , multiple applications of this gate can be used to construct $e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z}$.

With the problem of approximating elements of a group reduced to single qubit unitary gates, established results can be reapplied with ease. For example, recalling DiVincenzo [77], the group can be generated from two gates A and B that provide a non-trivial Lie algebra from their commutator $[A, B]$. Since a single qubit gate can be expressed in an Euler angle decomposition as given by (2.11), an irrational phase gate $R_z(\alpha)$ and the Hadamard gate or some other Cartesian axis switching gate will be universal for single qubit gates. However some of the most important choices for finite gate sets for single qubit universality have their origin in the study of fault tolerance.

Gate sets for fault tolerant quantum computation

So far we have accepted that implementing a gate by approximation up to some error is a valid operation equivalent to implementing a real gate. What has not been discussed is the issue of how the errors may impact the computation and how they are dealt with. This area of research impacts the choice of universality gate sets and so will be briefly elaborated upon. If the gates in a decomposition have some error in their construction then the errors will combine in the final circuit. Fortunately, it was shown [20] that the errors combine linearly so the total error is $\mathcal{O}(n\epsilon)$ for n gates with individual error ϵ . Conversely, to have an error within some tolerable level, ϵ_T , the error per gate must be $\mathcal{O}(\frac{\epsilon_T}{n})$. Shor[82] showed how this $\mathcal{O}(\frac{1}{n})$ bound could be improved to a $\mathcal{O}(\frac{1}{\log^c(n)})$ for some constant c if error correcting codes were utilised.

Error correcting codes work by encoding the quantum information into a state that can be tested for errors by a fixed set of operations, without decoding the information, to correct an error with success probability $\approx 1 - \frac{1}{n}$ using $\mathcal{O}(\log n)$ measurements. The operations used in quantum error correcting codes are $\{\text{CNOT}, H, \sigma_z^{\frac{1}{2}}\}$. Unfortunately these are not a universal gate set. They form the two qubit Clifford group, \mathcal{C}_2 and by the Gottesman-Knill theorem[83], a computation of only computational basis measurements and the Clifford group gates can be efficiently simulated on a probabilistic classical computer. Note that the Clifford gates on a single qubit will only transform Pauli operator eigenstates into other Pauli operator eigenstates since they are rotations

by π and $\frac{\pi}{2}$. Boykin *et al* [34] found a proof that the set $\{\text{CNOT}, H, T\}$ was universal, where a combination of H and T gates could approximate any single qubit gate where $T = \sigma_z^{\frac{1}{4}}$ (also referred to as the “ $\frac{\pi}{8}$ ” gate when in the form $e^{-i\frac{\pi}{8}\sigma_z}$ ⁵; do not confuse with the Toffoli gate; outside of references to [22], T nearly always refers to this single qubit gate). In fact, if one has a single gate outside of the Clifford group then one has a universal set [84]. There are other known finite universal gate sets noted by Boykin *et al*[34] such as Shor’s basis $\{C(\sigma_x), Z^{\frac{1}{2}}, H\}$ [82] and Kitaev [85]’s basis $\{C(Z^{\frac{1}{2}}), H\}$ (equivalent to Shor’s) but the H, T single qubit gates have formed a common standard to this day, especially in the area of resources per gate [35, 36, 86].

The number of gate elements in an approximating decomposition

Just as there is concern about how quickly errors grow with the number of gates in a circuit, there is concern about the number of finite set elements needed to approximate a gate up to a fixed error. If the error per gate needs to be below a certain level to preserve the error rate of the total circuit, then the approximation of that gate needs to be good enough which will require more finite set elements in the approximation. If a gate is being approximated by an irrational angle α multiplying under modulo 1 to create an approximation when $|1 - m\alpha - 2[\frac{1}{2}m\alpha]| < \epsilon$, then we expect that this works with m in the order of $\frac{1}{\epsilon}$, scaling inversely linearly with the size of space on the circle that the error ϵ cuts out. This provides a polynomial scaling but, similarly to Shor[82]’s result, this can be improved to a polylogarithmic scaling. The Solovay-Kitaev theorem [85] provides an algorithm for constructing a sequence of gates from a finite set that will approximate a gate up to a desired error ϵ . The key result is that this algorithm uses a number of gates from the set that grows as a polynomial of $\Theta(\log(\frac{1}{\epsilon}))$.

The area of resources for individual gate implementations shares some conceits with the main topic of this thesis, ancilla driven quantum computation. Since decomposition into a finite gate set is an algorithm, it can also be judged on its efficiency with respect to the number of uses of gates from the finite set. It can then be refined by the cost under some metric of each gate in the set not being equal.

For example the work of Bocharov, Gurevich & Svore [36] uses the gates $V_i = (\mathbb{I} \pm 2iX_i)/\sqrt{5}$ to create shorter decompositions than the $\{H, T\}$ basis and Bocharov & Svore [35] looked at the optimisation of decomposing single qubit gates with respect to the number of T gates since these have a higher cost in fault tolerant implementations. In ancilla driven quantum computation, there is a cost of ancillary resources for the

⁵This derives from the $\frac{\pi}{8}$ phases in its matrix elements. This can become confusing because “ $\frac{\pi}{8}$ ” could refer to several other gates and sometimes the term “ $\frac{\pi}{8}$ ” is used to refer to the $Z^{\frac{\pi}{4}}$ gate which does not have $\frac{\pi}{8}$ in its matrix elements. The author finds it necessary to mention this choice of notation so that future readers are warned but also to discourage anyone else from continuing its use.

creation of a unitary gate and a theme of this thesis will be in how different cost measures can be traded off with each other.

The quantum gate circuit model forms an analogy with the logic gate circuits of classical computation. However while the representation of unitary operations as quantum gates reflects the basic principles of quantum mechanics, there is no reference to any particular physical substrate or implementation beyond the constraints of quantum mechanics and the corresponding conservation of energy for reversible operations. Detailing the combination of resources, that is the finite set of gates, and an instruction set to create any finite unitary comprises the first *scheme* for quantum computation. Yet as experimental developments for quantum computation have come along and knowledge of the nature of the systems that will implement it has improved, other schemes have been created that refer to the nature or characterisation of a particular physical system and detail the relevant costs.

2.3.3 Measurement Based Quantum Computation

In *measurement based quantum computation* (MBQC), also called *one-way quantum computation*, the ability to perform an operation and the system on which it is performed corresponds to the finite state they generate. This state is treated as a resource that the computation uses up. This idea appears in quantum communications where, for example, a maximally entangled pair of particles is called an *e-bit* and protocols are characterised by the number of *e-bits* they use rather than the quantum gates that might be used to create them. The resource state is a way of making different capabilities equivalent. In MBQC, proposed by Raussendorf and Briegel [23], the resource state is constructed in a preparatory stage so that the computation input is put into a large entangled state with many other qubits and then the evolution that matches the computation is driven by projective measurements on the resource state. The resource states of interest are *cluster states* and *graph states*. A graph state takes a graph of edges and vertices and assigns a qubit in the $|+\rangle$ state to each vertex and performs a *CZ* operation between every qubit pair with an edge connecting them [24, 25]. The cluster state is a subclass where the graph can only be comprised of vertices on a cubic lattice with horizontal or vertical edges between nearest neighbour vertices [24, 25, 87] (see figure 2.7). This proposal makes sense when one considers models of Ising type interactions [88] or Heisenberg type interactions [89] that take place over a large field of bodies and induce nearest-neighbour interactions. In this framework, it can be easier to replace the construction of many individual entangling gates into the construction of a single large entangled state through nearest neighbour interactions. It can be shown that this scheme is capable of simulating the gate based scheme and thus is universal.

Consider a qubit in some general state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is entangled using a *CZ*

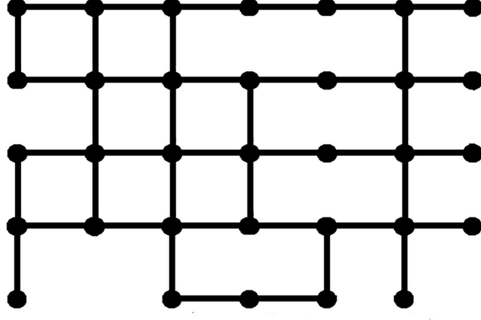


Figure 2.7: Visual representation of a 2d cluster state: Each dot is a qubit prepared in the $|+\rangle$ state, each line is a CZ interaction.

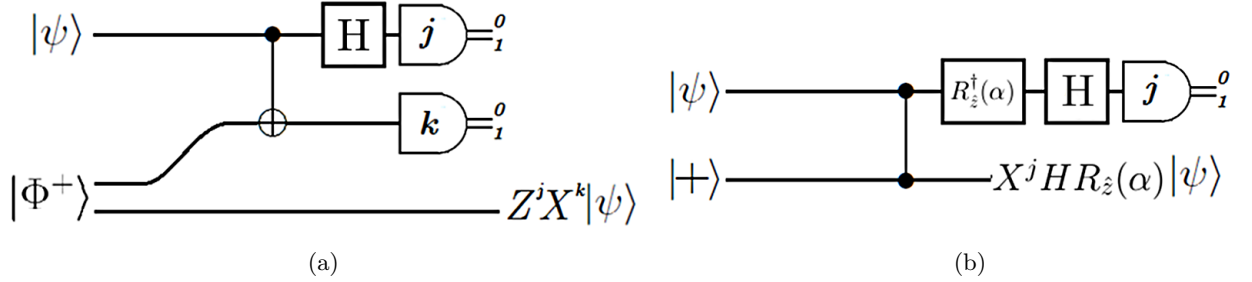


Figure 2.8: Gate circuits of the implementation of a) the teleportation protocol, b) single qubit measurement based QC.

gate with a qubit specially prepared in the state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$:

$$CZ|\psi\rangle_1|+\rangle_2 = \frac{1}{\sqrt{2}}(\alpha|0\rangle_1|+\rangle + \beta|1\rangle_1|-\rangle_2).$$

Now if the first qubit is measured in the $\{|\pm\rangle\}$ basis, the process resembles the teleportation protocol: if measured in the $|+\rangle$ state, the second qubit will be transformed to $\alpha|+\rangle + \beta|-\rangle$ while the equal probability opposing $|-\rangle$ result will transform it to $\alpha|+\rangle - \beta|-\rangle$. A measurement result $\frac{1}{\sqrt{2}}(|0\rangle + (-1)^m|1\rangle)$ teleports the state to the second qubit plus a result dependent unitary evolution $X^m H|\psi\rangle$.

Now say that the measurement basis was changed. Let us apply a phase gate $e^{i\frac{\alpha}{2}\sigma_z}$ to the first qubit or equivalently apply the phase gate $e^{-i\frac{\alpha}{2}\sigma_z} = R_{\hat{z}}(\alpha)$ to the measurement basis. This phase rotation in the measurement can correspond to a physical rotation of the measurement apparatus around a \hat{z} axis where the device alignment \hat{x} selects the $|+\rangle$ state from the spin or polarisation or equivalent system that forms the qubit. Now the corresponding second qubit states will be $X^m H R_{\hat{z}}(\alpha)|\psi\rangle$. Remarkably, one can teleport not only the initial state of the first qubit but also a post-interaction operation through choice of measurement (see figure 2.8b and contrast with 2.8a).

The unitary gate $J(\alpha) = H R_{\hat{z}}(\alpha)$ forms a fundamental part of a decomposition

of any single qubit unitary. For any $U \in U(2)$, $U = e^{i\alpha}J(0)J(\beta)J(\gamma)J(\delta)$ for some $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ [90]. So consider that the above is performed several times. A state is prepared on one qubit and then the procedure is performed to teleport and evolve the state onto a second qubit. This is repeated in a chain of five qubits so that three successive measurements will leave the final qubit in the state

$$X^{m_3}HR_{\hat{z}}(\beta)X^{m_2}HR_{\hat{z}}(\gamma)X^{m_1}HR_{\hat{z}}(\beta)|\psi\rangle.$$

The probabilistic Pauli- X operator can be commuted through;

$$HR_{\hat{z}}(\beta)X = HXR_{\hat{z}}(-\beta) = ZHR_{\hat{z}}(-\beta), \quad (2.37)$$

$$HR_{\hat{z}}(\beta)Z = HZR_{\hat{z}}(\beta) = XHR_{\hat{z}}(\beta), \quad (2.38)$$

$$\begin{aligned} &\rightarrow X^{m_3}HR_{\hat{z}}(\beta)X^{m_2}HR_{\hat{z}}(\gamma)X^{m_1}HR_{\hat{z}}(\delta) = \\ &X^{m_3}Z^{m_2}X^{m_1}HR_{\hat{z}}((-1)^{m_2}\beta)HR_{\hat{z}}((-1)^{m_1}\gamma)HR_{\hat{z}}(\delta), \end{aligned} \quad (2.39)$$

with an additional measurement with no rotation to apply an extra H gate:

$$U = X^{m_4}Z^{m_3}X^{m_2}Z^{m_1}J(0)J((-1)^{m_2}\beta)J((-1)^{m_1}\gamma)J(\delta). \quad (2.40)$$

The probabilistic effects can be countered by adapting the choice of measurement basis of each step according to the information of the measurement result from previous steps, a computation that can be performed by a classical computer constructed only by $CNOT$ and NOT gates [91], and by a Pauli operator correction at the end. The Pauli operator correction does not have to be corrected unitarily but can be corrected by a bit flip in the classical processing of measurements on the final system state [24]. Since this enables any single qubit unitary gate, the actual preparation of the input qubit can occur with such a chain of measurements. Therefore before the actual computation one could construct a 1D wire of qubits, all prepared in the $|+\rangle$ state, and operate a CZ gate between each qubit and its nearest neighbours, performing the computation by the measurement and feed-through of results of each qubit along the line (see figure 2.9).

In the gate circuit model, the ability to implement any single qubit unitary and an entangling gate was needed for universality. Since a 1D wire of connected qubits enables one to simulate the implementation of any single qubit unitary, single qubits states can interact through a CZ gate between the wires. The actual qubits where there is a connection may end up being measured so there must be a measurement that will preserve the CZ gate between the neighbours of the measured qubit.

If one considers a connection point where a qubit with the state information is entangled with its nearest neighbour on each axis, the state of these three points alone would be:

$$\alpha|0\rangle_1|+\rangle_2|+\rangle_3 + \beta|1\rangle_1|-\rangle_2|-\rangle_3. \quad (2.41)$$

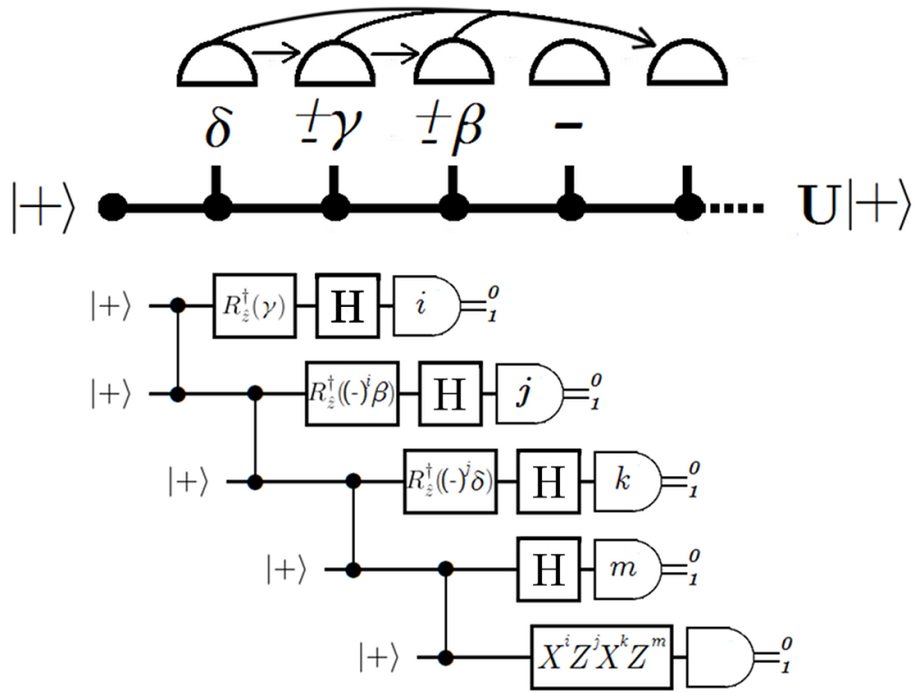


Figure 2.9: A 1d MBQC wire and corresponding gate circuit.

If the first qubit is then measured in the $|\pm\rangle$ basis, the second and third qubit will collapse into the state

$$\begin{aligned} & \alpha|+\rangle|+\rangle \pm \beta|+\rangle|+\rangle \\ & = (X^m \cdot H \otimes \mathbb{I})_{2,3} \cdot CZ \cdot (\alpha|0\rangle + \beta|1\rangle)|+\rangle. \end{aligned}$$

Raussendorf & Briegel [23] demonstrated how the principle can be extended to a more realistic junction in the middle of a wire as in figure 2.10.

Once it has been shown that such a system is computationally universal, further questions can be asked about how it behaves under certain limits and what the most efficient ways to encode the computation are. For example, a square n by n uniform matrix is highly symmetric and a simple structure to form physically. It can be turned into a smaller state with a cross-wire pattern by measuring unneeded qubits in the computational basis (and accounting for Z gate operations). This raises the issue of what the minimal size of square cluster state needed to implement a smaller less uniform structure.

There is also an interesting relationship between the cluster states and the Clifford group operations and error correction codes. One can note that the Clifford group gates are generated by performing measurements in Pauli eigenstate bases and also note that the cluster states can be constructed from Clifford group gates. The Clifford group can also be defined by its property that it maps Pauli operators to Pauli

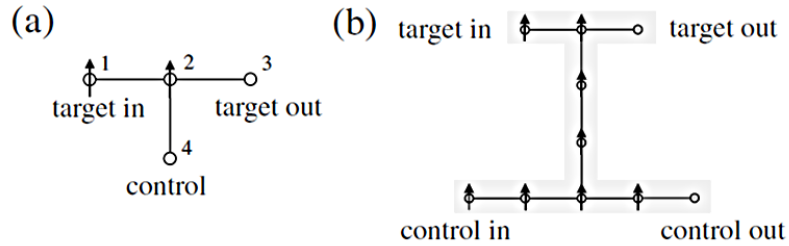


Figure 2.10: Realisation of a CNOT gate by one-particle measurement from Raussendorf & Briegel [23]. In (a) a CNOT gate is applied with the minimum number of qubits, however in practice measurement dependent corrections have to be applied to both control and target and it is more convenient if both qubits are on separate sites from the gate action necessitating the design in (b).

operators so any Pauli corrections from applying Clifford group operations could also be done beforehand so there is no need to adapt measurements[24]. So Clifford group operations are simply equivalent to the Pauli eigenstate measurements that are used to adapt one state into another- the Z basis measurements for removing qubits from the networks and so on. The implementation of Clifford group operations can be reduced to rearranging the graph pattern through which the information will flow which reflects the Gottesman-Knill theorem - this set of operations that are already known to be classically simulatable can in fact be reduced to classical pre-processing of the one-way pattern [24].

Experimental implementations

Physical implementations of MBQC tend to fall into either those that form natural lattices or those that can form a greater range of graphs but with unreliable entanglement operations. The former use physical systems that display a natural arrangement into neighbourhoods that assign themselves to lattice vertices in the graph state with an interaction that can be activated to generate cluster states. These are chosen because of the relative ease with which cluster states can be formed but inherently will tend to be limited to cluster states. The latter class are not usually prone to particular neighbour arrangements, using a highly mobile system whose constituents can be moved around but are suited to MBQC due to relying on an entangling operation that may, due to being non-deterministic or non-unitary, be unsuitable for *in media* application.

An example of the former are ultracold bosonic atoms in optical lattices. Optical lattices are arrays of microscopic potentials induced by the AC Stark effect of interfering laser beams [92] and by controlling the optical potential depth and the transition of the ultracold atoms from superfluid to a Mott insulator each of the lattice sites can

be uniformly loaded [92, 93]. Interactions between the sites can be effected by the use of cold controlled collisions created by adjusting the potential well separation en mass [93], dipole-dipole interactions [94] or spin dependent tunnelling [95]. These techniques allow for the fast generation of cluster states. Focused optical beam addressing may be applied to single atoms for measurements and may possibly be extended to manipulating the atoms to provide single qubit rotation gates [96].

Photonics systems are also more naturally suited to a scheme with a large focus on the upfront creation of an entangled resource. Direct photon-photon interactions through non-linear media such as with the Kerr effect are limited to very weak couplings at the single photon level [97]. Instead success has been found in linear optics based around the use of gate teleportation and ancilla photons [98, 99]. These probabilistic processes can be made to efficiently implement the fault-tolerant two qubit gates needed for universal quantum computation by transforming them into an entangled state preparation problem [100, 99]. Since linear optic for quantum computation will rely on state preparation anyway, by adopting the MBQC scheme, more efficient approaches such as the use of fusion operators like $|0\rangle\langle 00| + |1\rangle\langle 11|$ can be used to generate cluster states [101]. Photonic systems can also be used in hybrid systems with NV-centres in diamond, Pauli-blockade quantum dots with an excess electron or trapped ions with optical transitions trapped ions in cavity QED implementations [102], giving such systems the advantages of the mobility of photon systems. For example, cluster states could now be built using distributed nodes to aid in scaling up the ancillary systems required for error correction [103].

Other schemes and their related implementation systems

GBQC and MBQC are the two major schemes for quantum computation necessary for an appreciation of the scheme at the centre of this work. However, we will spare a passing mention of other schemes in order to emphasise how a scheme can influence our concept of how a universal quantum computer can be implemented and the relationship between a UQC and quantum mechanics.

Topological quantum computation seeks to perform fault tolerant quantum computation by using anyons [26]. Anyons are a class of particle that exists solely in 2 dimensional planes whose statistical properties contrast with fermions and bosons. They are “any”ons because the state of two particles accrues either a phase of +1 (bosons) or -1 (fermions) under exchange except an anyon can have any phase $e^{i\theta}$. This ability to accrue a phase by the exchange of particles can enact quantum operations that lead to computation when the exchanges follow particular braid patterns. Such anyons occur in devices that display the fractional quantum hall effect [104], such as very small Bose-Einstein condensates [105], they can also be created as excitations

in the Hamiltonian formed by certain cluster states [95].

Adiabatic quantum computation is a field of study instigated by Farhi, Goldstone, Gutmann and Sipser [27] which seeks to apply the adiabatic theorem to quantum algorithms. Recall that under the adiabatic theorem if a system is in the eigenstate of a Hamiltonian and then that Hamiltonian is sufficiently slowly perturbed and there is a large enough gap between the original eigenvalue and another eigenvalue of the new Hamiltonian, then the system will remain in the eigenstate of the final Hamiltonian. In particular, if it starts in the ground state, it will end in the ground state. So a system could start in a trivial ground state and then the Hamiltonian slowly changed to one whose ground state is more complex. While Farhi *et al*'s algorithm was found to be inefficient, the overall concept of adiabatic quantum computation was found to be universal [106]. This is functionally the same as the concept of quantum annealing [28] where the aim is to treat an optimisation problem as finding the ground state to a Hamiltonian; the total Hamiltonian is treated as the target Hamiltonian and a time dependent term that adiabatically fluctuates and evolves to zero.

Holonomic quantum computation is an interesting melding of the adiabatic theorem with the gate based model where the Hamiltonian is adiabatically evolved but loops back around to the origin; this imparts a Berry phase on the states of the system which provides non-trivial unitary operations that can form universal quantum computation [107].

2.4 Ancilla Driven Universal Quantum Computation

Different schemes emphasise different properties of potential physical implementations sometimes making explicit requirements, such as the topological scheme's designs for anyons, others implicitly emphasising a quality such as the massive entangled state generation involved in measurement based quantum computation. Quantum computation itself assumes some particular capabilities that are not necessarily involved in general quantum mechanics. Loss and DiVincenzo [108] notably considered five criteria that any quantum computer may need. They have been enumerated many times in different forms but I display them once again in the form in which they appear in [109]:

1. A scalable physical system with well characterised qubits
2. The ability to initialise the state of the qubits to a simple fiducial state
3. Long relevant decoherence times, much longer than the gate operation time
4. A universal set of quantum gates
5. A qubit-specific measurement capability

There is often a trade off between easy access and application of gates or long coherence times. Physical systems that lend themselves to long decoherence times are difficult to manipulate while relatively short-lived systems are more easily controlled and can be quickly initialised and measured. There has been much work that looks at dealing with such properties by hybridising systems that display properties on different sides of the trade-off. Work on optical clocks using aluminium ions have employed magnesium or beryllium ions as an ancillary system to account for a lack of optical accessibility with aluminium ions [110]. Ohshima [111] considered maintaining low decoherence of quantum dots by only activating access through an ancilla qubit in the same cell. It has been proposed that isolated, stable NV centre nuclear spins be used as qubits manipulated by neighbouring electron spins [40, 41] and Bermudez *et al* developed an proposal for nuclear spin interactions mediated by electron spins that effects an Ising type interaction [41]. Ion trap+photon systems such as in [39] and solid state systems with ballistic electrons have been considered for a class of systems that involve generating quantum gates through scattering between a static and flying qubit [48].

A scheme predicated on the idea of using a hybrid implementation was introduced as Ancilla Driven Quantum Computation (ADQC) [38, 37] where the implementation of unitary gates are performed on a system by interacting the system with an ancilla and then measuring the ancilla.

Return to the picture of a universal quantum computer as a memory register of qubits on which the input is written and a separate machine that operates on each qubit one at a time. The register qubits need to maintain a long lived stable coherence to last the length of the computation but in this set up the memory register qubits never had to interact with each other directly and only needed to be accessible at limited finite time intervals. This allows the register to be created in systems highly isolated from external interactions to maintain coherence.

There then needs to be an ancilla system of qubits that will interact with the memory register to implement operations. The ancilla system needs to be able to address every register qubit so it is made from a highly mobile system similar to the use of a hybrid system to create flexible choices of graphs states for MBQC mentioned proposed by Barrett and Kok [102]. However because of the trade off in stability against access and manipulation, it would be difficult to write a program on the memory register and it would be difficult for the memory register to remain coherent long enough to implement the program. So let the programming and single qubit unitary control be restricted to the ancilla system so that the only operation acting on the register qubits is the ancilla-register coupling operation E_{AR} which, as with GBQC and MBQC, keeps characterisation and fidelity costs limited to a single interaction and let the ancilla

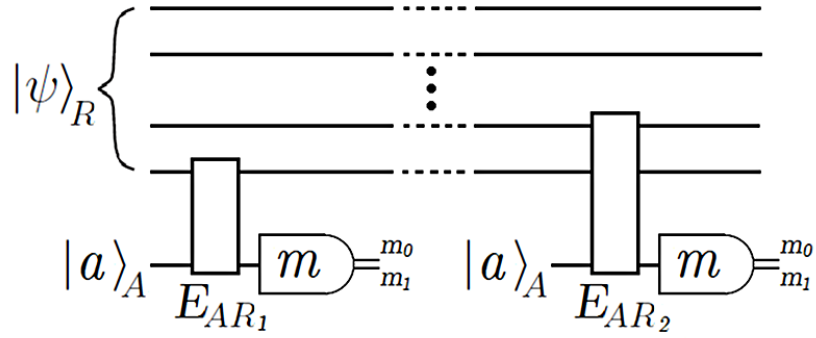


Figure 2.11: Implementing an ancilla driven computation using an unspecified interaction gate E_{AR} : ancilla qubits are created in the state $|a\rangle$ and used once before measurement.

system be restricted to only a single qubit in each time interval before being measured. After that it is replaced by another qubit or recreated in a specific preparation state (see figure 2.11).

To perform ADQC, the ancilla state preparation and final measurement basis must be chosen according to the interaction gate E_{AR} to perform a unitary operation. The scheme deals not only with physical hybrids but also a conceptual hybrid between the quantum gate circuit and the measurement driven evolution of MBQC.

As with MBQC, the generated unitary transformation is expected to depend on the nature of the entangling gate, the ancilla preparation state and the measurement basis. For the scheme to be universal it must be able to implement the gates of a universal set and therefore one must expect to be able to generate an entangling gate between register qubits using an ancilla qubit that has interacted with multiple register qubits. One also needs to be able to account for the randomness of the resulting gate generated by relying on the random measurement result.

Kashefi *et al* [37] found that for the resulting action on the register to be unitary, E has to be *locally equivalent* to an Ising type or Heisenberg-XX type interaction where two gates A, B are “locally equivalent” if $\exists P, Q \in SU(2) \otimes SU(2)$ such that $A = P.B.Q$. Furthermore, should the gates be maximally entangling in those two classes i.e. locally equivalent to CZ or $CZ.SWAP$ then the difference in measurement result can be corrected by the feed through of Pauli operator post-corrections, as with MBQC simulation of a single circuit gate.

2.4.1 $E_{AR} = (H \otimes H).CZ$

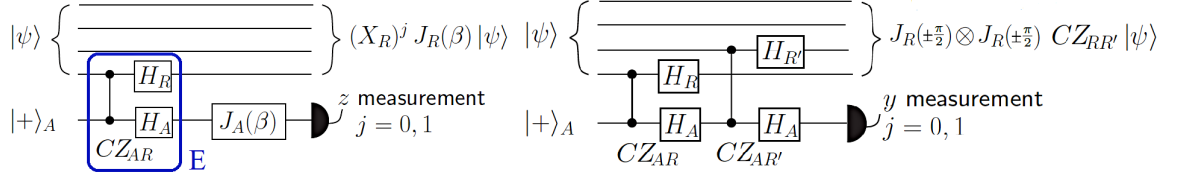


Figure 2.12: Anders *et al.* depiction of ancilla-driven implementation of a single qubit rotation [38]. The ancilla and register qubits are coupled with CZ and the local unitary gates are chosen such that the interaction remains symmetrical with respect to ancilla-register exchange. A rotation $X^j J(\beta)$ is enacted on the register a result $J(\beta)$ is enacted on the ancilla which is then measured in the z basis with a result $j=0,1$.

An example can be performed with $E_{AR} = (H_A \otimes H_R).CZ_{AR}$. The ancilla is prepared in the $|+\rangle$ state, couples with the register qubit with E_{AR} and then undergoes a unitary $J(\beta) = HR_z(\beta)$ before being measured in the computational basis. The action on the register qubit for a measurement result $|j\rangle$ is $X_R^j J_R(\beta)$. The difference between the two resulting unitary operations can be corrected by X_R . Any arbitrary single qubit unitary can be decomposed into four rotations $e^{i\alpha} J(0)J(\beta)J(\gamma)J(\delta)$ so to implement any arbitrary single qubit unitary up to a global phase, four ancilla interaction-measurements are performed changing the parameter of the local unitary on the ancilla to be the Euler angles δ, γ, β and then 0,

$$X^i J(0)X^j J(\beta)X^k J(\gamma)X^l J(\delta). \quad (2.42)$$

As in MBQC, the Pauli corrections can be commuted through each application of $J(\beta)$ if we make adaptations to the local unitary applied to the ancilla measurement basis [23, 24, 37]:

$$X^i J(0)X^j J(\beta)X^k J(\gamma)X^l J(\delta) = X^i Z^j X^k Z^l J(0)J((-1)^k \beta)J((-1)^l \gamma)J(\delta). \quad (2.43)$$

This mirrors the MBQC single qubit simulation but for two details: the measurement is made on the second qubit and the interaction is $(H \otimes H).CZ$ rather than CZ . The latter is necessary because in MBQC the local Hadamard was effected by the preparation of the second qubit in the $|+\rangle$ state while here, since the information remains on the same qubit, the H gate must be part of the interaction.

A two qubit gate can be implemented in the same way, under the same principles of the CZ gate between two wires. The same interaction gate E_{AR} is applied between the ancilla and each register qubit. Since there is only a single target gate, the measurement basis is fixed. If $E_{AR} = (H_A \otimes H_R).CZ_{AR}$ then the resulting gate is equivalent to a CZ gate. Only the local unitary gate corrections $J(\pm \frac{\pi}{2}) = X^m.H.R_z(-\frac{\pi}{2})$ are dependent on the measurement result m so the gate is deterministic up to a local gate correction.

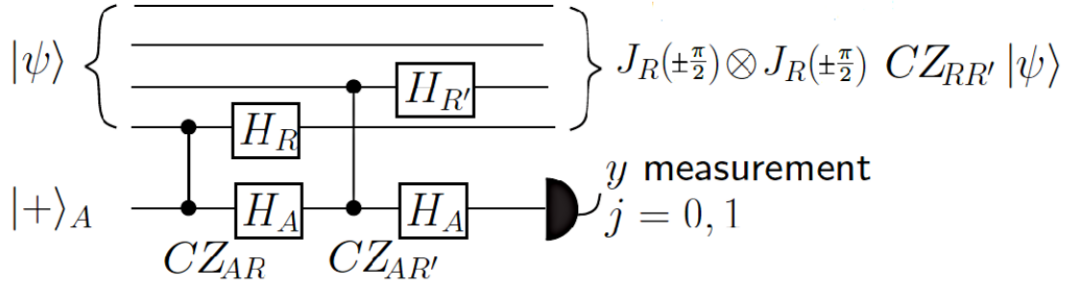


Figure 2.13: Ancilla driven implementation of a CZ gate on two register qubits R and R' using $E_{AR} = (H_A \otimes H_{R/R'}) \cdot CZ_{AR/R'}$ from [38].

Gates of the $E_{AR} = CZ.SWAP$ class can be used in the same way with different local gate corrections but as Proctor & Kendon [43] point out $E_{AR} = CZ.SWAP$ can in fact be performed differently without the use of measurements. This enacts the $CZ.SWAP$ gate so each interaction gate generates a two qubit gate in its own class.

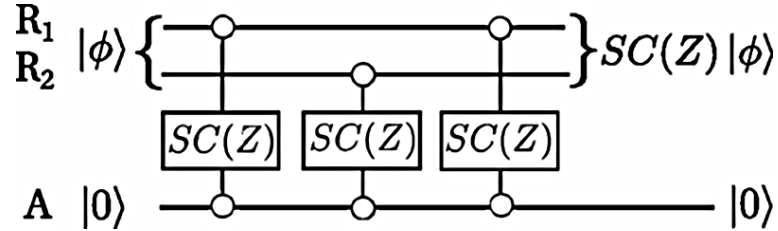


Figure 2.14: Circuit from [43] displaying ADQC two qubit gate generation with a $SC(Z) = CZ.SWAP$ interaction gate without ancilla measurement.

This allows ancilla driven quantum computation to be universal and deterministic up to Pauli corrections. If we describe each ancilla interaction and measurement procedure as a “generation” then it takes four generations to simulate any arbitrary single qubit unitary and up to two generations (including the local unitary corrections) to implement a two qubit entangling gate. All manipulations and unitary control of a qubit are only employed on the ancilla system. It is also possible to perform measurements on register qubits through the ancilla if one leaves off the $J(\beta)$ operation on the ancilla. Moreover, it as proposed by Anders *et al*[38] that other non-unitary operations could be constructed by using a single additional qubit appended to a state in the register, to enact any positive operator valued measurement (POVM) and thus any quantum channel.

2.4.2 Beyond deterministic ADQC

The unitary gates in the resource set of ADQC are equivalent to those in GBQC and MBQC: there is a two qubit gate equivalent to CZ and arbitrary single qubit gates formed by the use of the Hadamard gate and phase rotation gates by any angle. The

difference in the schemes is in the coordination of the application of these gates and different emphasis on the properties of the physical substrate on which they are carried out. We are interested in how this resource set can be altered, particularly in the choice of the interaction gate E_{AR} . In the gate circuit model it was possible to use gates of differing entangling power as a fundamental universal gate and the choice of gate under some limitations had been characterised; an investigation of MBQC with relaxed entangling power requirements has also been performed [44, 45]. So too we are interested in how ADQC can be employed with interaction gates of arbitrary entangling power, how the limitations of ancilla-register structure affect this and what costs in time and efficiency this may inflict. This is particularly of interest due to the possible lack of interaction tuning in some scattering based hybrid interactions [48, 49]. There is also a similarity between ADQC and other proposals for mediating between distant qubits such as the quantum bus or *qubus* [42]. Therefore it is useful to consider ancilla schemes in the context of distributed or networked quantum computation where non-local operations are applied over relatively large separations that inhibit coordination. The original proposal for ADQC required interaction gates of two maximally entangling classes in order to enact deterministic unitary gates up to Pauli post-corrections. In order to broaden the class of gates, we will have to characterise the relationship between the unitary operations enacted and the gate parameters and investigate how they can be used to achieve softer targets than Pauli post-corrected determinism.

Chapter 3

Dilation theory

3.1 Generalised Quantum Mechanics and the Stinespring Dilation Theorem

The mathematical treatment of Ancilla Driven Quantum Computation needs to be able to address the behaviour of a subsystem of a larger quantum system after operations on a separate subsystem. There is a formalism in which the rules of quantum mechanics as per the conventional set of postulates in section 2.2.1 can be generalised to include the mixing of linear quantum states with classical randomness. It is needed to describe the behaviour of imperfect measurements, impure states and the loss of information and energy to external systems that would occur in any realistic experimentation.

One of the results of the generalised formalism is that a system whose quantum mechanical behaviour breaks down can be equated to a subsystem of a larger quantum system into which information of that subsystem is escaping. This aids in consoling the supposedly fundamental nature of quantum mechanics with the classical world around us -the real quantum system is the universe in total and anything we observe is only a subsection- but its chief importance to this work is that this results provides for the opposite. Given a system and a description of the unitary operations and projective measurements on it, the evolution of a subsystem can be described.

In this section, the mathematical results of this formalism that we require will be detailed with some results of papers that helped inspire ADQC. It also demonstrates the choice of notation for this work out of several texts that cover the topic [112, 113, 114].

3.1.1 Mixed states

Imagine that a device is meant to prepare a system in a specific quantum state $|\Psi\rangle$, yet with a probability p the Ph.D. student in charge of the device will forget to flip one of the switches and it instead prepares the orthogonal state $|\Psi_{\perp}\rangle$. If this could

be represented by some state in the Hilbert space, $|\Phi\rangle \in \mathcal{H}$, then according to the postulates of quantum mechanics, one could perform a measurement of the system in the basis corresponding to that state and find a result corresponding to that state with unit probability. Yet the expectation value of the projector $|\Phi\rangle\langle\Phi|$ will be, according to the laws of statistics, the weighted sum of the expectation values of each state:

$$\begin{aligned} p(\Phi) &= \langle P \rangle = (1-p)\langle\Psi|\hat{P}|\Psi\rangle + p\langle\Psi_{\perp}|\hat{P}|\Psi_{\perp}\rangle \\ &= (1-p)|\langle\Psi|\Phi\rangle|^2 + p|\langle\Psi_{\perp}|\Phi\rangle|^2 \\ &= (1-p)\cos^2\left(\frac{\phi}{2}\right) + p\left(1 - \cos^2\left(\frac{\phi}{2}\right)\right). \end{aligned}$$

$p(\Phi)$ can never equal 1 for $0 < p < 1$ so this is not consistent with the description of states as vectors in a Hilbert space.

These issues are not solely an issue of quantum systems interacting outside the prescribed restrictions. Let us consider the case of a specific interaction between two systems.

In the Hilbert space \mathcal{H}_1 , there is a state

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle_1 + |1\rangle_1)$$

while in Hilbert space \mathcal{H}_2 , there is a state

$$|\psi_2\rangle = |0\rangle_2.$$

The product $\mathcal{H}_1 \otimes \mathcal{H}_2$ is also a Hilbert space for the total system in the state $|\psi_{12}\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$. We can choose a specific unitary to evolve the total system

$$U_{12} = CNOT = |0\rangle_1\langle 0|_1 \otimes \mathbb{I}_2 + |1\rangle_1\langle 1|_1 \otimes (|0\rangle_2\langle 1|_2 + |1\rangle_2\langle 0|_2) \quad (3.1)$$

$$\begin{aligned} |\psi_{12}\rangle &= \frac{1}{\sqrt{2}}(|0\rangle_1 + |1\rangle_1) \otimes |0\rangle_2 = \frac{1}{\sqrt{2}}(|0\rangle_1|0\rangle_2 + |1\rangle_1|0\rangle_2) \\ U_{12}|\psi_{12}\rangle &= \frac{1}{\sqrt{2}}(|0\rangle_1|0\rangle_2 + |1\rangle_1|1\rangle_2). \end{aligned}$$

However this new state in the total system Hilbert space can not be expressed as a product of two evolved subsystems $U_{12}|\psi_{12}\rangle = U_1|\psi_1\rangle \otimes U_2|\psi_2\rangle$ because we are able to act on the larger Hilbert space with a unitary that is not expressible as the product of two smaller unitaries. If one had access to only one subsystem \mathcal{H}_1 and was unaware of the other system, one would not be able to describe the evolution of the subsystem according to the five postulates in section 2.2.1. Applying a projector onto a subsystem, $|\Phi\rangle_1\langle\Phi| \otimes \mathbb{I}_2$, can not occur with unity probability for any $|\Phi\rangle$ because the state contains terms for two orthogonal states and no local unitary operation will change that. It would appear as though the device was producing two different pure states with probability $\frac{1}{2}$ and there's no Ph.D. student to blame it on.

These and other effects such as loss in detectors or energy exchange with an outside system lead to situations that are not well described by the five postulates and \mathcal{H}_1 alone. This leads to the formulation of a generalised construction for describing the statistical results of open quantum systems.

The density operator

Previously, a quantum state that could be described as a vector in a Hilbert space defined by the inner product norm was referred to as a *pure* state. When the system is no longer pure and is being produced by a *statistical mixture* of states, as described above, it is referred to as a *mixed* state. These mixed states are then represented, instead of a vector in a Hilbert space, by a *density operator*.

Generally, a density operator ρ can be expressed as a weighted sum of projectors:

$$\rho = \sum_j p_j |\Psi_j\rangle\langle\Psi_j|. \quad (3.2)$$

This is not a linear superposition of a basis of states in a Hilbert space but a point in a convex set whose extreme points are the pure states. A convex set has the property that every point on a line between two points of the set is also in the set and that every point can be expressed as a convex decomposition of other points [115]:

$$\rho = \sum_j p_j \rho_j. \quad (3.3)$$

The coefficients p_j are the probabilities associated with the statistical mixture so in our example of the bumbling Ph.D. student the density operator would be

$$\rho = (1 - p)|\Psi\rangle\langle\Psi| + p|\Psi_\perp\rangle\langle\Psi_\perp|.$$

The expectation value of this operator $\langle\Phi|\rho|\Phi\rangle$ corresponds to the probability of measuring the system in the state $|\Phi\rangle$:

$$\text{prob}(|\Phi\rangle | \rho) = \sum_j p_j \langle\Phi|\Psi_j\rangle\langle\Psi_j|\Phi\rangle = \sum_j p_j |\langle\Phi|\Psi_j\rangle|^2. \quad (3.4)$$

Naturally, the probabilities must be real, positive and sum to 1,

$$0 \leq p_j, \quad \sum_j p_j = 1, \quad (3.5)$$

and these are the eigenvalues of the density operator so the density operator is a self-adjoint, positive operator, diagonalised in an orthogonal basis.

Statistical properties of the density operator are calculated using the trace:

$$\text{Tr}[\rho] = \sum_j \langle j|\rho|j\rangle. \quad (3.6)$$

The trace is a useful property because it is a) cyclically symmetrical:

$$\text{Tr}[ABC] = \text{Tr}[BCA] = \text{Tr}[CAB], \quad (3.7)$$

e.g.

$$\text{prob}(|\Phi\rangle | \rho) = \langle \Phi | \rho | \Phi \rangle = \text{Tr}[\langle \Phi | \rho | \Phi \rangle] = \text{Tr}[|\Phi\rangle\langle \Phi | \rho], \quad (3.8)$$

changing the expression from an observable's average to the action of an operator on the density operator, and also b) linear:

$$\text{Tr}[A + B] = \text{Tr}[A] + \text{Tr}[B], \quad (3.9)$$

e.g.

$$\begin{aligned} \sum_j p_j &= \text{Tr}\left[\sum_j \langle \Psi_j | \rho | \Psi_j \rangle\right], \\ &= \text{Tr}\left[\rho \sum_j |\Psi_j\rangle\langle \Psi_j|\right], \\ &= \text{Tr}[\rho] = 1, \quad \forall \rho \end{aligned} \quad (3.10)$$

which is consistent with (3.5) and demonstrates that the trace of a density operator is always 1.

The properties of cyclical symmetry and linearity together mean that the trace is independent of the choice of diagonal basis:

$$\sum_j \langle j | U^\dagger \rho U | j \rangle = \text{Tr}[U^\dagger \rho U] = \text{Tr}[\rho U U^\dagger] = \text{Tr}[\rho]. \quad (3.11)$$

The expectation value of an observable with a mixed state can thus also be written in terms of a trace:

$$\langle A \rangle = \sum_j p_j \langle a_j | \hat{A} | a_j \rangle$$

for a orthonormal basis $\{|a_j\rangle\}$ that spectrally decomposes \hat{A} . Given any orthonormal basis $\{|k\rangle\}$ for which $\rho = \sum_k |k\rangle\langle k|$ and for which, by definition as an orthonormal basis $\sum_k |k\rangle\langle k| = \mathbb{I}$,

$$\langle A \rangle = \sum_j p_j \langle a_j | \hat{A} \sum_j |k\rangle\langle k | a_j \rangle, \quad (3.12)$$

$$= \sum_{j,k} p_j \langle a_j | \hat{A} | k \rangle \langle k | a_j \rangle = \sum_j \langle a_j | \hat{A} \sum_k p_k |k\rangle\langle k | a_j \rangle, \quad (3.13)$$

$$= \text{Tr} \left[\hat{A} \sum_k p_k |k\rangle\langle k| \right], \quad (3.14)$$

$$= \text{Tr}[\hat{A}\rho]. \quad (3.15)$$

The trace can also be used as a test of purity:

$$\begin{aligned}
\rho^2 &= \sum_j p_j |\Psi_j\rangle\langle\Psi_j| \sum_k p_k |\Psi_k\rangle\langle\Psi_k|, \\
&= \sum_{j,k} p_j p_k |\Psi_j\rangle\langle\Psi_j|\Psi_k\rangle\langle\Psi_k|, \\
&= \sum_j p_j^2 |\Psi_j\rangle\langle\Psi_j|.
\end{aligned} \tag{3.16}$$

$$\text{Tr}[\rho^2] = \sum_j p_j^2, \tag{3.17}$$

therefore $\text{Tr}[\rho^2]=\text{Tr}[\rho]=1$ only when a single p_j is 1 and the rest are 0, i.e. $\rho = |\Psi\rangle\langle\Psi|$. As a corollary, just as unitary evolution does not change the trace of a density operator, it also does not change the trace of its square thus this new formalism maintains unitary operations as a map of pure states to pure states.

Though the density operator is diagonalisable, our description of the statistical ensemble need not be done in an orthogonal basis. For example, let us imagine that in the case of the bumbling Ph.D. student, the machine was alternatively producing the two states $|\Psi\rangle$ and $\frac{1}{\sqrt{2}}(|\Psi\rangle + |\Psi_\perp\rangle)$. Since the density operator is still a diagonalisable self-adjoint positive operator means we can still use the rules of linear addition on the states $|\Psi\rangle$ and $\frac{1}{\sqrt{2}}(|\Psi\rangle + |\Psi_\perp\rangle)$ to find the diagonal description. Since the expectation values of any observable depend only on the trace, there is no way to discern that the state ρ was prepared in the above manner rather than from an ensemble of the orthogonal states. In fact, it is impossible to distinguish the preparation of any state ρ from an infinite number of possible convex combinations unless the state is pure. This can be illuminated by considering how mixed states fit in the Bloch sphere picture.

Mixed states in the Bloch sphere

We can extend the Bloch sphere picture to represent the convex set of states. With the pure states as extreme points on the surface of the sphere, there is no line between pure states that goes outside the sphere. The mixed states are represented by points inside the sphere: in this picture, a mixed state that can be expressed as $\rho = p|\psi\rangle\langle\psi| + (1-p)|\phi\rangle\langle\phi|$ lies on the line between $|\psi\rangle$ and $|\phi\rangle$ at the point that divides it into the ratio $\frac{p}{1-p}$.

This picture demonstrates how a mixed state cannot distinguish between any preparation method for there are an infinite number of lines that pass through any point. The only measurable properties relate to the coordinates of this point. The state ρ is represented by a point $\hat{\mathbf{r}} = (r_x, r_y, r_z)$. If we were to consider one of three axes in isolation, say $\hat{\mathbf{z}}$, any pure state of the same elevation will have the same expectation value for σ_z , $\langle\sigma_z\rangle = \cos(\theta)$, so any statistical ensemble of only those states will clearly

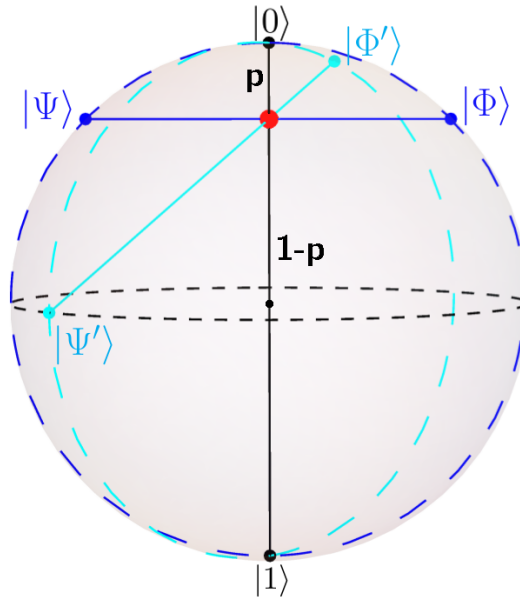


Figure 3.1: For $|\Psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)|1\rangle$, $|\Phi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle - \sin\left(\frac{\theta}{2}\right)|1\rangle$, the state $\rho = \frac{1}{2}(|\psi\rangle\langle\psi| + |\phi\rangle\langle\phi|)$ is equal to $p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ for $p = 1 - \cos(\theta)$. There are infinitely many alternative pairs of states, $|\Psi'\rangle$ & $|\Phi'\rangle$, for which ρ lies on the line but the state is diagonalised by the states on the line which passes through the origin.

have the same expectation value $\text{Tr}[\sigma_z \rho] = \cos(\theta)$ as seen in figure 3.1. By the Bloch sphere symmetry this clearly applies for each axis and each coordinate can be found from

$$r_j = \text{Tr}[\sigma_j \rho]. \quad (3.18)$$

The coordinates of a point in a unit radius Bloch sphere can also be used to construct the density operator:

$$\rho = \frac{1}{2}(\mathbb{I} + \hat{\mathbf{r}} \cdot \vec{\sigma}). \quad (3.19)$$

Partial Trace

Given the probabilities of a statistical mixture of states, a density operator may be constructed, but when the operator is a subsystem of a larger pure or mixed state it is found by use of the partial trace. The partial trace of a density operator in a product space $\mathcal{H}_A \otimes \mathcal{H}_B$ is a trace utilising only the basis of only one of the spaces A or B . To understand the partial trace, start with the process of extracting a pure state from a product of pure states. Say there are many pure states and a large dimensionality for each subsystem and thus many coefficient terms, no matter what since they are all

pure subsystems it could be rearranged into the form

$$|\psi\rangle = \sum_i c_i |i\rangle_A \sum_j c'_j |j\rangle_B \sum_k c''_k |k\rangle_C \dots \quad (3.20)$$

and if substate B was to be extracted, one could apply the sum of projection operators on all the other subsystems:

$$\sum_{i'} \langle i'|_A \sum_{k'} \langle k'|_C \dots |\psi\rangle = \sum_{i'} c_i \langle i'|_i \sum_j c'_j |j\rangle_B \sum_{kk'} c''_k \langle k'|_k \dots \quad (3.21)$$

$$= \sum_i c_i \sum_j c'_j |j\rangle_B \sum_k c''_k \dots \quad (3.22)$$

Renormalising the result would then provide the state of subsystem B but if instead the density operator of $|\psi\rangle$ was used then from (3.21), the end result would be normalised,

$$\sum_{i'} {}_A \langle i' | \psi \rangle \langle \psi | i' \rangle_A \dots = \sum c_i \langle i' | i \rangle_A c_{i''}^* \langle i'' | i \rangle_A \sum_j c'_j |j\rangle_B \sum_j c_{j'}^* \langle j' | \dots = \sum_{jj'} c'_j c_{j'}^* |j\rangle \langle j'|_B. \quad (3.23)$$

So if this is applied to a general density operator, for example with two subspaces,

$$\rho_{AB} = \sum_{jj'kk'} c_{jj'kk'} |j\rangle_A |k\rangle_B \langle j'|_B \langle k'|_A, \quad (3.24)$$

the partial trace with respect to B yields the density operator of subsystem A :

$$\rho_A = \text{Tr}_B[\rho_{AB}] = \sum_{k''} \langle k'' |_B \rho_{AB} |k''\rangle_B \quad (3.25)$$

$$= \sum_{jj',k''} c_{jj'k''k''} |j\rangle_A \langle j'|_A. \quad (3.26)$$

3.1.2 Generalised measurement

Any measurement of a density operator ρ can be represented by a positive operator value measure (POVM) which is a set of elements $\{M_j\}$ for which $\sum_j M_j = \mathbb{I}$ and $M_j \geq 0$. The outcome j , represented by M_j occurs with probability $p_j = \text{Tr}[M_j \rho]$.

The idea behind this generalisation is that to be completely general in describing a measurement, one would look for a map between the density operator and the set of possible probabilities. Recall that with a projective measurement, the outcome probabilities are the expectation values of the projector operator, $\text{Tr}[\hat{P}_j \rho]$ and these always return a real, positive number less than or equal to 1 because of the conditions $\text{Tr}[\rho] = 1$, $\rho \geq 0$ and the properties of the trace operation assure that $\sum_j \text{Tr}[\hat{P}_j \rho] = 1$, for any choice of basis to form the set $\{P_j\}$, which guarantees that all probabilities sum to one.

Now consider that we have another operator M_j , to map from the density operator to the set of possible probabilities, that is real and positive semi-definite so that $\text{Tr}[M_j \rho]$

is also real and positive. We would then want

$$\sum_j \text{Tr} [M_j \rho] = 1, \forall \rho. \quad (3.27)$$

It follows from this that

$$\begin{aligned} \text{Tr} \left[\sum_j M_j \rho \right] &= 1 = \text{Tr}[\rho], \forall \rho \\ &\rightarrow \sum_j M_j = \mathbb{I}. \end{aligned} \quad (3.28)$$

It was previously shown how the probability of a projective measurement could be calculated with a density operator in (3.4). However a projective measurement may not accurately represent the informational understanding we gain after a measurement when dealing with realistic experimental designs and/or interference from outside environments with a presumed closed system. A projective measurement indicates that we absolutely know the current state of the system and that all future measurements will consistent with that state and the probabilities of that measurement describe the before state but some simple thought experiments can lead us to cases where that is not true.

Example: Mixed channels and Weak measurements

The measurement of the spin of a particle may be performed with a Stern-Gerlach experiment. The particle passes through a magnetic field to which their magnetic moment couples and induces an attractive potential to either side of the plate depending on whether the spin is aligned parallel or anti-parallel with the field. So the two spin states are distinguished by whether the particle exits the field drawn to one side or the other. There is a notion, as Hans Margeneau may have put it [116], that all measurements are ultimately measurements of position and so from that the natural pointer, position, is naturally continuous and the state of the pointer has some spread.

When it enters the Stern-Gerlach devices, the particle has a Gaussian distribution of trajectories. When it emerges on either side, each spin will be correlated with a spread with separated averages but whose total width may have some significant overlap (see figure 3.2). Depending on where the measurement of the particle is made, a tail of the distribution representing a probability p of the particles will be on the other side of the natural division of the line that symmetrises the distribution centres.

If the spin was prepared in a specific, known alignment 100 times and then sent through we would notice that for the spin state $|0\rangle = |\uparrow\rangle$, there would be p measurements missing from the side correlated with “0” and p appearing on the “1” side and

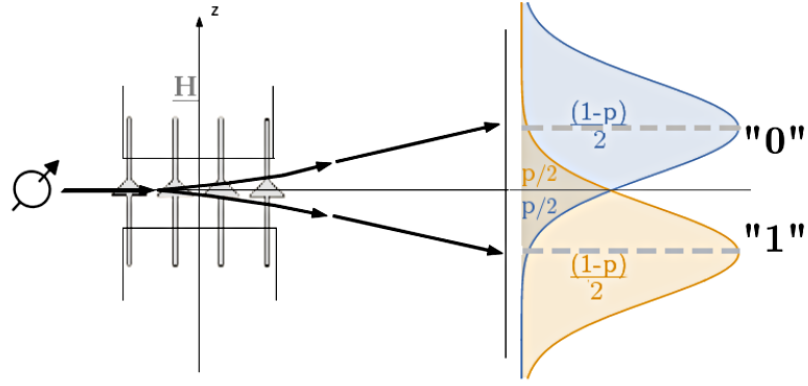


Figure 3.2: A Stern-Gerlach experiment on an equal superposition of spin states resulting in overlapping normal distributions of results.

vice versa for $|1\rangle = |\downarrow\rangle$ states, symmetric with respect to the spins and plane symmetry. So this can be represented by the two element, non-projective POVM where each result includes a contribution from both projectors of the two spin states:

$$\mathbf{M}_0 = \begin{pmatrix} 1-p & 0 \\ 0 & p \end{pmatrix}, \quad \mathbf{M}_1 = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}. \quad (3.29)$$

If p is small relative to one, this may work as an approximation to a projective measurement with some uncertainty in the result but if the spreads of the distributions were broadened or the interaction was so weak that it did not significantly move the two possible distributions away from each other, the measure elements are close to proportional to identity. The increased uncertainty in a single measurement can be compensated for by taking all 100 measurements and taking the average. This loosely defines the concept of weak measurements- the change of the quantum system due to the interaction is negligible but with the trade off that the measured variable can only be determined from an average over an ensemble of pointers [117, 118].

A technical definition can be made if one has a model of the measurement system. The measurement involves an interaction $\hat{\mathcal{E}}$ between the system state $|\Phi\rangle_S$ and the probe state $|\Psi\rangle_P$. When the system is in an observable's eigenstate, $|e_i\rangle$, the interaction reduces to an operator \hat{E}_i on the probe:

$$\hat{\mathcal{E}}|\Phi\rangle_S|\Psi\rangle_P = \sum_i c_i |e_i\rangle_S \hat{E}_i |\Psi\rangle_P. \quad (3.30)$$

\hat{E}_i evolves $|\Psi\rangle$ such that a pointer parameter of the probe changes $p \rightarrow p_i$ and becomes correlated with a system eigenstate. By characterising the pointer parameter we deduce the state of the system in the corresponding eigenstate. In the case where

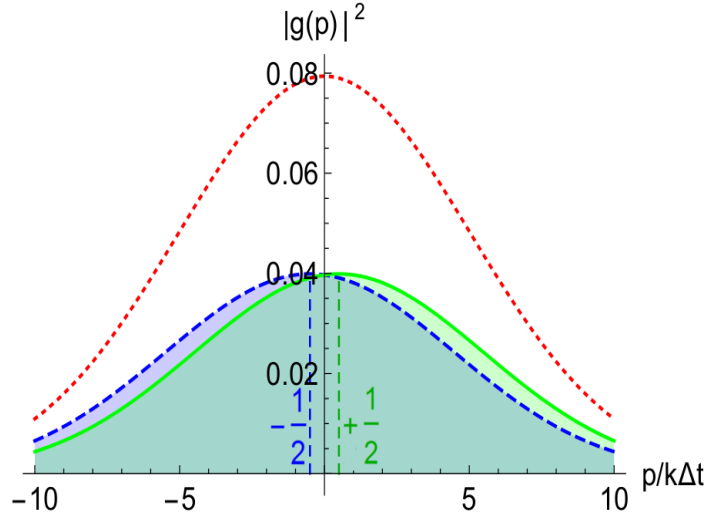


Figure 3.3: The dashed(blue) and solid(green) curves represent the distributions of the pointer after interaction with the system post-selected in the $-\frac{1}{2}$ and $+\frac{1}{2}$ spin eigenstates respectively. Dotted(red) curve represents the average distribution of the pointer for an equal superposition system. The large overlap between the post-interaction pointer distributions is a necessary condition for the weak measurement; only after a large ensemble of measurements have been taken can the state of the system be inferred.

$\{\hat{E}_i|\Psi\rangle\}$ form an orthonormal basis, the system and probe are entangled, the correlations between p and e_i are strong and the measurement corresponds to a projective measurement. The system state can be described by a single measurement result of the probe. This is the strong measurement regime.

Aharonov, Albert and Vaidman [117] supposed that measurements could be treated as a case where the interaction $\hat{\mathcal{E}}$ has a Hamiltonian $\hat{H} = k'(t)\hat{A}\hat{x}$ over a time Δt and the probe is in a Gaussian state $g(x) = e^{-x^2/2\sigma^2}$. Given that $k'(t)$ is required to have compact support around the time of measurement, we can treat it as roughly linear and the interaction as

$$\hat{\mathcal{E}} = e^{-ik\Delta t\hat{A}\hat{x}}. \quad (3.31)$$

A system in the eigenstate of \hat{A} , $|A = a_i\rangle$, results in a phase operator $e^{-ik\Delta t a_i \hat{x}}$ applied to the probe. This phase in the x representation results in a shift in the canonical momentum p representation which is likewise Gaussian, $\tilde{g}(p) \rightarrow \tilde{g}(p - k\Delta t a_i)$. The mean of the momentum forms the pointer parameter. The probe has an associated spread $\frac{1}{\sigma}$. If the spread is small relative to the shift then $|\langle \tilde{g}(p - k\Delta t a_i) | \tilde{g}(p - k\Delta t a_i) \rangle|^2 \approx \delta_{ij}$ and the measurement is strong but when $k\Delta t |a_i - a_j| \ll \frac{1}{\sigma}$ (visualised in fig. 3.3) the measurement is weak.

Example: Biased Measurements

POVM elements may be of different rank. The rank refers to the minimum number of state projectors needed to describe the element. A rank one element may be expressed as $|\tilde{\Psi}\rangle\langle\tilde{\Psi}| = p|\Psi\rangle\langle\Psi|$ for normalised $|\Psi\rangle$. An example of a POVM with elements of different ranks is the biased measurement.

In this example, imagine that an experiment is performed to couple a particle's spin to the polarisation of an emitted photon. The two-dimensional system state of light polarisation manages to easily avoid the issue of non-orthogonal probe states in section 3.1.2. However the experiment fails to couple the spin and polarisation states strongly. Ideally the two systems would become maximally entangled

$$U(\alpha|0\rangle + \beta|1\rangle)_s |H\rangle_p \rightarrow \alpha|0\rangle|H\rangle + \beta|1\rangle|V\rangle,$$

$|H/V\rangle$ represents the horizontal/vertical state of polarisation. Instead, the experiment achieves something more like:

$$U(\alpha|0\rangle + \beta|1\rangle)_s |H\rangle_p \rightarrow \alpha|0\rangle|H\rangle + \beta|1\rangle|D\rangle$$

$$|D\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle).$$

The problem is that now if the photon is sent through a horizontal filter, one can not immediately distinguish between the non-orthogonal states $|H\rangle$ and $|D\rangle$. On the other hand, if the light passes through a vertical filter it must have been in the $|D\rangle$ polarisation because that provides the only contribution of vertical polarisation. This provides some certain knowledge and therefore projects the particle spin into the $|1\rangle$ state. The POVM on the spin is represented by:

$$\mathbf{M}_0 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \quad \mathbf{M}_1 = \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.32)$$

If instead of $|D\rangle$ we described this with a general state $|\Theta\rangle = \cos\left(\frac{\theta}{2}\right)|H\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|V\rangle$, it would be:

$$\mathbf{M}_0 = \begin{pmatrix} 1 & 0 \\ 0 & \cos^2\left(\frac{\theta}{2}\right) \end{pmatrix}, \quad \mathbf{M}_1 = \sin^2\left(\frac{\theta}{2}\right) \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.33)$$

The POVM can in fact be generated by the use of projective measurements with classical post processing in part because the two elements share a basis. It can be constructed by performing a projective measurement in the computational basis and then if the result is $|1\rangle$ tossing a coin (or performing a random number generation in the general case) and marking the result as $|0\rangle$ instead if it turns up heads. The effect of classical systems can also be seen in the Stern-Gerlach example: the initial distribution of trajectories may be classical noise but the same effect on measurement probabilities arises.

Example: Unambiguous state discrimination

POVMs and the framework of generalised measurements is important for not only describing what may go wrong but also for providing measurements that achieve properties not possible by projective measurements and classical post processing alone. An example of this is unambiguous state discrimination.

Say that we have returned to a case like the biased measurement example of a spin through non-orthogonal polarisation states or even the bumbling Ph.D. student. A problem has arisen because we are unable to absolutely discriminate between two non-orthogonal states.

We are able to give an absolute statement on certain measurement results- when we measure a state orthogonal to one of the two states we know that it excludes one of the options. Using a POVM we can prioritise this unambiguity in our result so that we are at least able to state some of the time that the system is in a specific state.

For two non-orthogonal states in a 2 dimensional system, $|\Psi\rangle$ and $|\Phi\rangle$, we want a POVM with elements proportional to the projectors of their orthogonal states:

$$M_0 = a|\Phi_\perp\rangle\langle\Phi_\perp|, \quad M_1 = b|\Psi_\perp\rangle\langle\Psi_\perp| \quad a, b < 1. \quad (3.34)$$

These alone will not fulfil the condition $\sum_j M_j = \mathbb{I}$ so we must add a third element

$$M_2 = \mathbb{I} - a|\Phi_\perp\rangle\langle\Phi_\perp| - b|\Psi_\perp\rangle\langle\Psi_\perp|. \quad (3.35)$$

To simplify the example, set $a = b$ and this will turn out to provide the optimal solution. For $0 \leq a \leq 1$, $M_0, M_1 > 0$ but we must also set a s.t. $M_2 > 0$. Since the probability of getting a conclusive result increases with a , we would want to maximise a while under the positivity constraint. This gives

$$a = \frac{1}{1 + |\langle\Psi|\Phi\rangle|}. \quad (3.36)$$

Now if we are at an extreme case where the two non-orthogonal states occur with equal probability, $\rho = \frac{1}{2}(|\Psi\rangle\langle\Psi| + |\Phi\rangle\langle\Phi|)$ then the probability of an inconclusive result is

$$p_2 = \text{Tr}[M_2\rho], \quad (3.37)$$

$$= 1 - a\text{Tr}[|\Psi_\perp\rangle\langle\Psi_\perp|\rho] - a\text{Tr}[|\Phi_\perp\rangle\langle\Phi_\perp|\rho], \quad (3.38)$$

$$= 1 - a|\langle\Psi_\perp|\Phi\rangle|^2, \quad (3.39)$$

$$= 1 - \frac{1 - |\langle\Phi|\Psi\rangle|^2}{1 + |\langle\Phi|\Psi\rangle|} = |\langle\Phi|\Psi\rangle|. \quad (3.40)$$

This turns out to be the optimal failure rate [119, 120, 121, 122, 123]. On the other hand a naive approach with just projective measurement where we measure in a fixed basis including one of the states, the inconclusive result occurs with probability

$$\text{Tr}[|\Psi\rangle\langle\Psi|\rho] = \frac{1}{2}[1 + |\langle\Psi|\Phi\rangle|^2]. \quad (3.41)$$

A complete generalisation of measurement of quantum states also requires a description of the post-measurement state. To go into this, we will need to approach the topic of generalised evolution of states.

3.1.3 Generalised Evolution

Unitary evolution can map any pure state to another pure state but is not sufficiently general to describe the mapping of any density operator (pure or mixed) to all other density operators. We require a generalised formalism for evolution to describe the maps $\Lambda : \rho \rightarrow \rho'$. The trace and positivity conditions on the density operators, $\text{Tr}[\rho] = 1$ and $\rho \geq 0$, enforce conditions on these maps. If $\Lambda(\rho) = \rho'$ and $\text{Tr}[\rho'] = 1$ and $\rho' \geq 0$ then Λ must be a trace preserving and positive map.

Physical intuition guides us to consider other conditions on the map, particularly within the theme of subsystems within a larger system. If a map occurs on a subsystem and this map is a good representation of physical reality then we would not expect it to alter the physical reality of the rest of the system. Therefore we also expect the stronger condition that the map is “completely positive” i.e. for $\Lambda : \mathcal{H} \rightarrow \mathcal{H}'$, there is the map that expands to a larger space with d times as many dimensions

$$\Lambda \otimes \mathbb{I}_d : \mathcal{H} \otimes \mathcal{H}_d \rightarrow \mathcal{H}' \otimes \mathcal{H}_d \quad (3.42)$$

for which

$$\Lambda \otimes \mathbb{I}_d(\rho_1 \otimes \rho_d) \geq 0, \forall d > 0 \text{ [124, 125]}. \quad (3.43)$$

Choi’s theorem [125] informs us that a completely positive trace preserving (CPTP) map if and only if it is of the form

$$\Lambda(\rho) = \sum_j K_j \rho K_j^\dagger, \quad (3.44)$$

$$\sum_j K_j^\dagger K_j = \mathbb{I}. \quad (3.45)$$

The elements K_j are Kraus operators. The product $K_j^\dagger K_j$ will be positive and with the trace preserving condition (3.45) shows that $\{K_j^\dagger K_j\}$ forms a POVM. To complete the description of the post measurement state in the POVM formalism, a result M_j can be associated with a state evolution $\rho \rightarrow K_j \rho K_j^\dagger$ where $K_j^\dagger K_j = M_j$. Since the trace preserving condition is only a constraint on the total set and not the individual Kraus operators, there needs to be a normalisation factor which will be the measurement probability:

$$\rho \rightarrow \rho' = \frac{K_j \rho K_j^\dagger}{\text{Tr}[M_j \rho]}. \quad (3.46)$$

So one can naturally express a channel linearly as a probabilistic implementation of other channels. Our earlier “bumbling Ph.D. student” example could be seen as

the random implementation of a unitary gate e.g. if the system is prepared in the $|0\rangle$ state but the mistake is that after preparation a random bit flip, an X gate, is implemented. On its own a unitary corresponds to a POVM, $M = U^\dagger U = \mathbb{I}$, because a unitary operation is independent of the state it is implemented on. The bumbling Ph.D. student enacts the channel

$$\Lambda(\rho) = (1 - p)\rho + pU\rho U^\dagger$$

where the Kraus operators are of the form

$$\begin{aligned} K_j &= \sqrt{p_j}U_j, \quad U_j \in U, \\ K_j K_j^\dagger &= K_j^\dagger K_j = p_j \mathbb{I}. \end{aligned} \quad (3.47)$$

To more directly tie a non-unitary channel into the issue of measurement, consider the dephasing channel:

$$\Lambda(\rho) = \sum_j |j\rangle\langle j| \rho |j\rangle\langle j|. \quad (3.48)$$

This can be seen as the measurement of the state in the computational basis but without a recording of the result thus the weightings of each sub-channel are state dependent. The dephasing — the loss of information about the phase of the initial state — projects states from the Bloch sphere onto the polar axis so there is still a general description in terms of an operation on the Bloch sphere.

With the relationship between evolution and measurement established, it is then necessary to relate these generalisations to the original postulates of quantum mechanics. In addition, for the study of ancilla driven operations, it is necessary to know how to construct a specific channel in a controlled manner.

When the above result of Choi's theorem is combined with the Stinespring dilation theorem, it is shown that any CPTP map can be constructed by applying a coupling between the system with an ancillary system[126][127]:

$$\Lambda(\rho_s) = \sum_j K_j \rho_s K_j^\dagger = \text{Tr}_a [U_{sa}(\rho_s \otimes |a\rangle\langle a|)U_{sa}]. \quad (3.49)$$

From this the Kraus operators can be related to a coupling U_{sa} and ancilla state $|a\rangle\langle a|$ and a free choice of orthonormal basis $\{|j\rangle\}$;

$$\text{Tr}_a [U_{sa}(\rho_s \otimes |a\rangle\langle a|)U_{sa}] = \sum_j \mathbb{I}_s \otimes \langle j_a|U_{sa}(\mathbb{I}_s \otimes |a\rangle) (\rho_s \otimes \mathbb{I}_a) (\langle a| \otimes \mathbb{I}_s) U_{sa}^\dagger |j_a\rangle \otimes \mathbb{I}_s.$$

The notation can be simplified

$$K_j = \langle j_a|U_{sa}|a\rangle := (\mathbb{I}_s \otimes \langle j_a|)U_{sa}(\mathbb{I}_s \otimes |a\rangle) \quad (3.50)$$

for the operation of the vectors in the ancilla space reduce the rank of the unitary to an action on the subspace \mathcal{H}_s :

$$\begin{aligned}
U_{sa} &= \sum_{kk'mm'} u_{kk'mm'} |k\rangle\langle k'|_s \otimes |m\rangle\langle m'|_a, \\
|a\rangle &= \sum_n \alpha_n |n_a\rangle, \\
\langle j_a|U_{sa}|a\rangle &= \sum_{kk'mm'} \sum_n u_{kk'mm'} \langle j|m\rangle \alpha_n \langle m'|n\rangle |k\rangle\langle k'|_s \\
&= \sum_{kk'n} u_{kk'jn} \alpha_n |k\rangle\langle k'|_s.
\end{aligned}$$

The Stinespring dilation theorem means that we can maintain the axiomatic nature of pure states and unitary evolution by treating any physical map as unitary in a larger Hilbert space. It also provides an instruction for how we can construct any channel if we have access to the appropriate coupling and ancilla system.

Since a basis $\{|j\rangle\}$ is not unique to the description of the trace, a channel can be described by any unitarily related basis. However if the ancilla is measured after the coupling then the projection selects an individual state $|i\rangle$ and reduces the evolution to a single Kraus operator description.

$$\begin{aligned}
\text{Tr}_a \left[|i_a\rangle\langle i_a| U_{sa} (\rho_s \otimes |a\rangle\langle a|) U_{sa}^\dagger |i_a\rangle\langle i_a| \right], &= \langle i_a| U_{sa} (\rho_s \otimes |a\rangle\langle a|) U_{sa}^\dagger |i_a\rangle, \\
&= \langle i_a| U_{sa} |a\rangle \rho_s \langle a| U_{sa}^\dagger |i_a\rangle, \\
&= K_i \rho_s K_i^\dagger.
\end{aligned}$$

This returns the result of Naimark's dilation theorem that a POVM can always be expressed as a projective measurement on a larger Hilbert space [113].

Example: Unambiguous state discrimination

One can look back to the example of unambiguous state discrimination to show how we can use Naimark's dilation theorem. A system that is either in state $|\Psi\rangle$ or $|\Phi\rangle$ for $\langle\Psi|\Phi\rangle > 0$ can be coupled to an ancilla state $|a\rangle$ by a unitary operation U that performs the transformations:

$$U|\Psi\rangle|a\rangle \rightarrow \alpha|0\rangle|a'\rangle + \beta|\theta\rangle|a'_\perp\rangle, \quad (3.51)$$

$$U|\Phi\rangle|a\rangle \rightarrow \alpha|1\rangle|a'\rangle + \beta|\theta\rangle|a'_\perp\rangle, \quad (3.52)$$

$$\langle a'|a'_\perp\rangle = 0. \quad (3.53)$$

The ancilla is measured in the basis of $\{|a'\rangle, |a'_\perp\rangle\}$ and if the $|a'_\perp\rangle$ result is returned then the system has collapsed into the state $|\theta\rangle$ independent of the initial state and

thus the measurement is inconclusive. However if the $|a'\rangle$ result is returned then a computational basis measurement of the system returns a result associated with a particular initial state. The probability of a $|a'_\perp\rangle$ result, $|\beta|^2$, is constrained by the norm preserving condition thus $|\beta|^2 = |\langle\Psi|\Phi\rangle|$.

The combined system-ancilla Hilbert space is at least of dimension 4 but we only required a 3 element POVM. In this case, two potential results are collected together in a rank 2 POVM element.

In some cases, we may in fact not even want to perform measurements directly upon the system such as when the measurement is physically destructive. So an ancilla may be chosen with dimensions equal to the number of outputs and all measurements performed upon it. For a d dimensional system and n outputs, the total system dimension is dn .

However we may not always want such “excessive dimensionality”. For some systems, such as linear optics [99], it is easier to create a direct sum extension of a Hilbert space, $\mathcal{H}_d \oplus \mathcal{H}_{d'}$, with dimension $d + d'$ rather than a product $\mathcal{H}_d \otimes \mathcal{H}_{d'}$ with dd' . When implementing a POVM with optical waveguides, it is easier to add another waveguide path than induce an interaction between two photons.

Unambiguous state discrimination can be easily expressed with a direct sum approach. If the initial states are in the Hilbert space spanned by the $\{|0\rangle, |1\rangle\}$ basis, add another orthogonal state $|2\rangle$ and perform the unitary:

$$U|\Psi\rangle \rightarrow \alpha|0\rangle + \beta|2\rangle, \tag{3.54}$$

$$U|\Phi\rangle \rightarrow \alpha|1\rangle + \beta|2\rangle. \tag{3.55}$$

3.1.4 Weak Values

A key feature of quantum mechanics is the lack of commutation between measurement operators and questioning what information can be extracted from successive strong measurements in different bases is the foundation of such problems as e.g. “The Mean King” [128]. When a weak measurement is followed by a strong measurement in a different basis, the result of the weak measurement correlated with one of the strong measurement results is known as the weak value and may display an amplification outside of the bounds of the eigenstates of the weak measurement.

Weak value amplification has been applied to the detection of small effects such as polarization dependent deflections and the optical Spin Hall effect[129, 130] and improving signal to noise ratios in interferometric experiments [131] as well as fundamental tests of quantum mechanics such as the generalized Leggett-Garg inequalities [132] and Hardy’s paradox [133, 134]. Weak value amplification has typically been demonstrated

using an additional degree of freedom of the system as the probe. Previous descriptions of weak values also heavily rely on the probe state being continuous, parametrised by a spread and highly classical such as a Gaussian distribution in position or momentum [117, 135]. Feizpour *et al.* [136] considered a coupling between two separate systems (two distinct optical beams) but the probe was still a classical coherent field that could be parametrised by a spread $|\alpha|$.

This counter-intuitive amplification of observables has been treated with significant skepticism [137, 138, 139] and derogation [140]. More recently, work has been done to extend into regimes where the weak value is not valid [141]. Here we will describe the model of weak values and how and when they arise. In the generalised form of quantum mechanics, measurement and evolution start to align in form and some evolutions can be seen as the processing of the information of measurement (e.g. just forgetting it!). ADQC is a special use of the dilation theorem where the information of measurements is retained and fed forward to maintain unitarity of evolution. Weak values are an interesting case for us since they involve post-selections and so are architecturally similar to ADQC (see fig.3.4) except ADQC deals with interactions between finite quantum states on different physical systems [38]. Thus weak values make an approximate description of the expected behaviour of ADQC as interaction couplings get weak while ADQC can reflect new light upon the interpretation of weak values. While the effects of WVA have in the past been described by such astounding statements as “the measurement of a component of a spin- $\frac{1}{2}$ particle can turn out to be 100”, a view based in the ADQC framework sees it as the transfer of a parameter of unitary transformation between two weakly entangled systems.

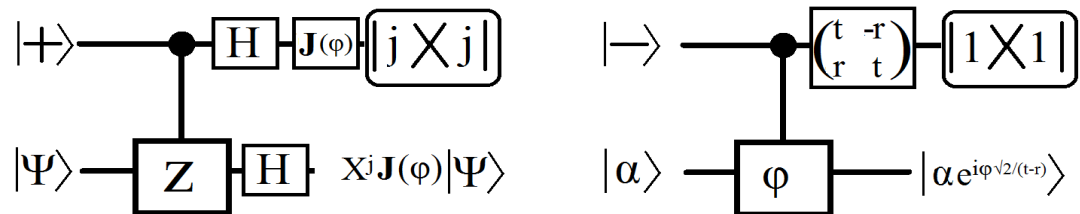


Figure 3.4: A comparison of the circuit diagrams that describe a) phase gates in ADQC and b) Weak value amplification. Both schemes involve a control-unitary (Control-Z /Control-Phase) and a choice of measurement basis defined by a local unitary parameter ($\phi/(t, r)$). The evolution of the register/probe depends on the parameter of the local unitary and the final measurement result.

The conditions for weak value amplification

Weak values rely on post-selection; after the weak measurement interaction a strong measurement with a different probe is performed then the individual pieces of data from the weak measurement are collected according to which strong measurement result they went on to give. If the strong measurement were just in the same basis as the weak one then the weak results would just be found to form the same distributions shifted by δtka_j (see fig. 3.3) that correspond to the measurement eigenstates. But for weak value amplification the strong measurement is performed in a different, even complementary, basis. It is one thing to apply a post-selection to the system being weakly measured but the formalism of weak value amplification only works under certain conditions. Consider expressing a $2d$ system measurement in another basis:

$$\begin{aligned}\hat{\mathcal{E}}|\Phi\rangle_S|\Psi\rangle_P &= c_0|0\rangle_S\hat{E}_0|\Psi\rangle_P + c_1|1\rangle_S\hat{E}_1|\Psi\rangle_P, \\ &= c_0\frac{1}{\sqrt{2}}(|+\rangle_S + |-\rangle_S)\hat{E}_0|\Psi\rangle_P + c_1\frac{1}{\sqrt{2}}(|+\rangle_S - |-\rangle_S)\hat{E}_1|\Psi\rangle_P, \\ &= c_+|+\rangle_S\hat{E}_+|\Psi\rangle_P + c_-|-\rangle_S\hat{E}_-|\Psi\rangle_P\end{aligned}$$

where $c_{\pm} = \frac{c_0 \pm c_1}{\sqrt{2}}$, $\hat{E}_{\pm} = \frac{1}{c_0 \pm c_1}(c_0\hat{E}_0 \pm c_1\hat{E}_1)$. There are two things to note. First, if $c_0 \approx c_1$ i.e. $|\Phi\rangle \approx |+\rangle$, then the amplitude $c_+ \gg c_-$. Second, the evolution of the probe caused by \hat{E}_{\pm} is not necessarily akin to that of \hat{E}_i . For example, if $\hat{E}_i = U_i$, $U_i U_i^\dagger = U_i^\dagger U_i = \mathbb{I}$, $U_+ = c'_0 U_0 + c'_1 U_1$:

$$U_+ U_+^\dagger = |c'_0|^2 \mathbb{I} + |c'_1|^2 \mathbb{I} + c'_0 c'_1{}^* U_0 U_1^\dagger + c'_1 c'_0{}^* U_1 U_0^\dagger, \quad (3.56)$$

$$= \mathbb{I} + \Gamma, \quad (3.57)$$

where

$$\Gamma = c'_0 c'_1{}^* U_0 U_1^\dagger + c'_1 c'_0{}^* U_1 U_0^\dagger. \quad (3.58)$$

The linear sums of unitary operators are not necessarily unitary. So now consider the action of a post-selection state $|\Psi_f\rangle = \sum_j \alpha'_j |A = a_j\rangle$. The unnormalised state of the probe would be

$$\langle \Psi_f | e^{i\Delta tk \hat{A} \hat{x}} | \Psi_i \rangle | P \rangle. \quad (3.59)$$

The probe evolves as

$$\langle \Psi_f | e^{i\Delta tk \hat{A} \hat{x}} | \Psi_i \rangle = \sum_n \langle \Psi_f | \frac{(i\Delta tk \hat{A} \hat{x})^n}{n!} | \Psi_i \rangle, \quad (3.60)$$

$$= (\langle \Psi_f | \Psi_i \rangle + (i\Delta tk \hat{x}) \langle \Psi_f | \hat{A} | \Psi_i \rangle + \dots) \quad (3.61)$$

$$\approx \langle \Psi_f | \Psi_i \rangle \left(1 + (i\Delta tk \hat{x}) \frac{\langle \Psi_f | \hat{A} | \Psi_i \rangle}{\langle \Psi_f | \Psi_i \rangle} \right). \quad (3.62)$$

$$\langle \Psi_f | \Psi_i \rangle \left(1 + (i\Delta tk\hat{x}) \frac{\langle \Psi_f | \hat{A} | \Psi_i \rangle}{\langle \Psi_f | \Psi_i \rangle} \right) \approx \langle \Psi_f | \Psi_i \rangle e^{i\Delta tk A_W \hat{x}} \quad (3.63)$$

where

$$A_W = \frac{\langle \Psi_f | \hat{A} | \Psi_i \rangle}{\langle \Psi_f | \Psi_i \rangle}. \quad (3.64)$$

A_W is the “weak value”. In a weak measurement, the original system is not greatly disturbed by the interaction and so the post-selection probability is close to the square of the overlap of the input and post-selected state. So the unnormalised factor of $\langle \Psi_f | \Psi_i \rangle$ reflects this. The weak value itself is inverse to the overlap and so can get arbitrarily large as the overlap tends to zero.

Note that the weak value can be complex. Real values and the real parts of complex values correspond to the original consideration of shifts in the x representation in the continuous case while the imaginary parts induce a shift in the p representation [142].

However for the approximations in (3.62) and (3.63), we have restrictions related to the weakness of the interaction[137].

(3.62) \rightarrow

$$\begin{aligned} \frac{|\hat{x}^n \langle \Psi_f | \hat{A}^n | \Psi_i \rangle|}{|\langle \Psi_f | \Psi_i \rangle|} &\ll 1, \\ |\hat{x}^n \langle \Psi_f | \hat{A}^n | \Psi_i \rangle| &\ll |\hat{x} \langle \Psi_f | \hat{A} | \Psi_i \rangle|, \end{aligned} \quad (3.65)$$

for $n \geq 2$ and (3.63) \rightarrow

$$|\Delta tk A_W \hat{x}| \ll 1. \quad (3.66)$$

If we consider a Gaussian probe state then a spread of σ for x will mean a spread of $\frac{1}{\sigma}$ for p and this spread will govern the x terms in our condition. So we can substitute $x \rightarrow \sigma$ into the above conditions [137]. The condition (3.65) for WVA is to ensure that a linear sum of unitary displacement operators E_i , will approximate a unitary displacement operator itself while condition (3.66) is specifically for when the simple form of the weak value is valid. The physical meaning of (3.65) & (3.66) is that the spread σ must be much greater than the range of the shifts in p from eigenstate measurements and the the weak value shift $\Delta tk A_W$ is limited by the scale of order of $\frac{1}{\sigma}$ (see figure 3.5).

This is because weak value amplification relies on the interference effects between the two Gaussian distributions split by the weak measurement. By the definition of a weak measurement, these distributions are still close enough to have their combined probabilities still approximate a Gaussian so their edges can interfere with each other in the strong measurement. The results after the post-selection must still be post-selected out from that combined distribution while still approximately Gaussian and thus well within its spread.

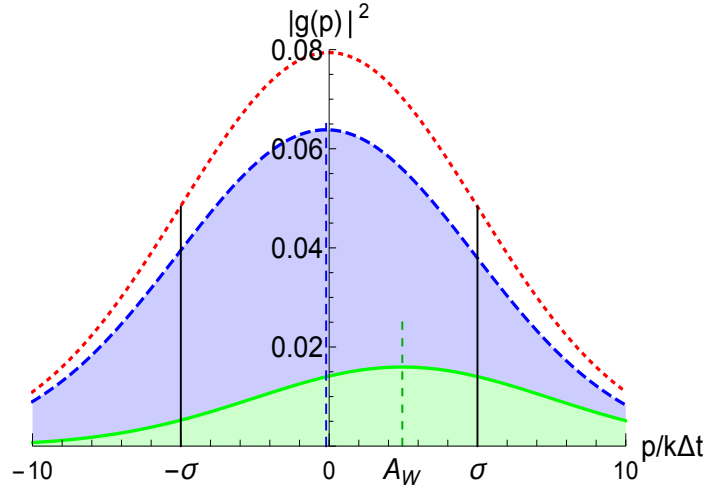


Figure 3.5: Interference between the amplitudes of the probe state after post-selection produces a distribution with an amplified shift, the solid curve. The shift amplification factor is of the order of the amplitude reduction factor, with a squared reduction in the probability distribution. Most measurement results will lie in the dashed curve. The amplified shift will remain within the spread of the initial distribution, the dotted curve.

Characterising continuous and high dimensional systems

More recent work has focused on the application of applying weak value measurements to higher dimensional and continuous systems such as a spatial wavefunction [143, 46, 47]. For such a system, it is simpler to use a measurement operator that divides the system into two subspaces so the resulting procedure is independent of the dimensionality of the system.

Considers a state $|\psi\rangle$ with a spatial wavefunction $\psi(x)$ such that

$$|\psi\rangle = \int dx \psi(x) |x\rangle \langle x|. \quad (3.67)$$

The probability of the system being detected at a point x_0 can be treated as the expectation value of the projector $\pi_{x_0} = |x_0\rangle \langle x_0|$. On the other hand, if one instead uses a weak measurement of the projector after a post-selection in the state $|p\rangle$:

$$\langle \pi_{x_0} \rangle_W = \frac{\langle p | x_0 \rangle \langle x_0 | \psi \rangle}{\langle p | \psi \rangle}. \quad (3.68)$$

Post-selections are made in eigenstates of complementary variables in bases unitarily equivalent to the weak measurement so the final state $|p\rangle$ will be a state of momentum p . From (3.67) and its momentum space equivalent, $\langle x_0 | \psi \rangle$ and $\langle p | \psi \rangle$ are equal to $\psi(x_0)$ and $\Phi(p)$. Therefore the weak value is proportional to the wavefunction at x_0 and when

$p = 0$ is equal to:

$$\langle \pi_{x_0} \rangle_W = \frac{e^{\frac{ipx}{\hbar}} \psi(x_0)}{\Phi(p)} = \frac{\psi(x_0)}{\Phi(0)}. \quad (3.69)$$

This term is still complex and so includes the phase of the wavefunction, unlike the square $|\psi(x_0)|^2$, and so the real and imaginary parts can be extracted from measurements of complementary variables of the pointer system.

The pointer can be qubit system. In a way, this inverts the original description of weak value amplification with a continuous probe and a qubit system while still performing the same basic operations. If the qubit is prepared in the state $|0\rangle$ and then couples with the system by $e^{i\alpha\sigma_y|x_0\rangle\langle x_0|}$, the weak value can be calculated from the expectation values of the final probe state $|s\rangle$ [143]:

$$\langle \pi_{x_0} \rangle_W = \frac{1}{\sin(\alpha)} [\langle s|\sigma_x|s\rangle - i \langle s|\sigma_y|s\rangle]. \quad (3.70)$$

By performing this measurement over a range of many projectors $\{|x_i\rangle\langle x_i|\}$ one can find an approximation $|\tilde{\psi}\rangle = \sum_i \Phi(0) \langle \pi_{x_i} \rangle_W |x_i\rangle\langle x_i|$.

It is also possible to calculate the density matrix of a mixed state system if one uses a basis of post-selection states [46]. The procedure given by Lundeen and Bamber [46] calculates the elements of a discrete version of the Dirac distribution given by Chaturvedi *et al* [144]. The elements of the distribution are proportional to the weak value of an operator $|a\rangle\langle a|$ and a post-selection $|b\rangle$:

$$\mathbf{S}_a b \equiv |b\rangle\langle b|a\rangle\langle a|, \quad (3.71)$$

$$\langle \mathbf{S}_{ab} \rangle = \text{Tr}[\mathbf{S}_{ab}\rho] = \langle b|a\rangle\langle a|\rho|b\rangle, \quad (3.72)$$

$$\langle \mathbf{S}_{ab} \rangle = p_b \langle \pi_a \rangle_{W,b}. \quad (3.73)$$

To have a complete finite Dirac distribution, the value must be calculated over two mutually unbiased bases $\{|a_i\rangle\}, \{|b_j\rangle\}$ so there must be multiple choices of post-selection with different normalisation factors. From (3.72) one can calculate the density matrix from the Dirac distribution [47].

3.1.5 Iterative implementations of POVMs and binary search trees

A question arises about how many additional dimensions are required to implement a POVM and the relative efficacy of the direct sum extension and the product space extension. The dimensionality of the extended space will depend on the number of outputs and their rank. An N output projective measurement on a N dimensional space will produce elements of rank 1. If the elements M_j have varied rank r_j then a d dimensional system would need to be extended by $\sum_j^N r_j - d$ under the sum extension which forms a lower bound on the product extension [145]. However by employing an iterative approach in which a smaller output measurement is performed first and then

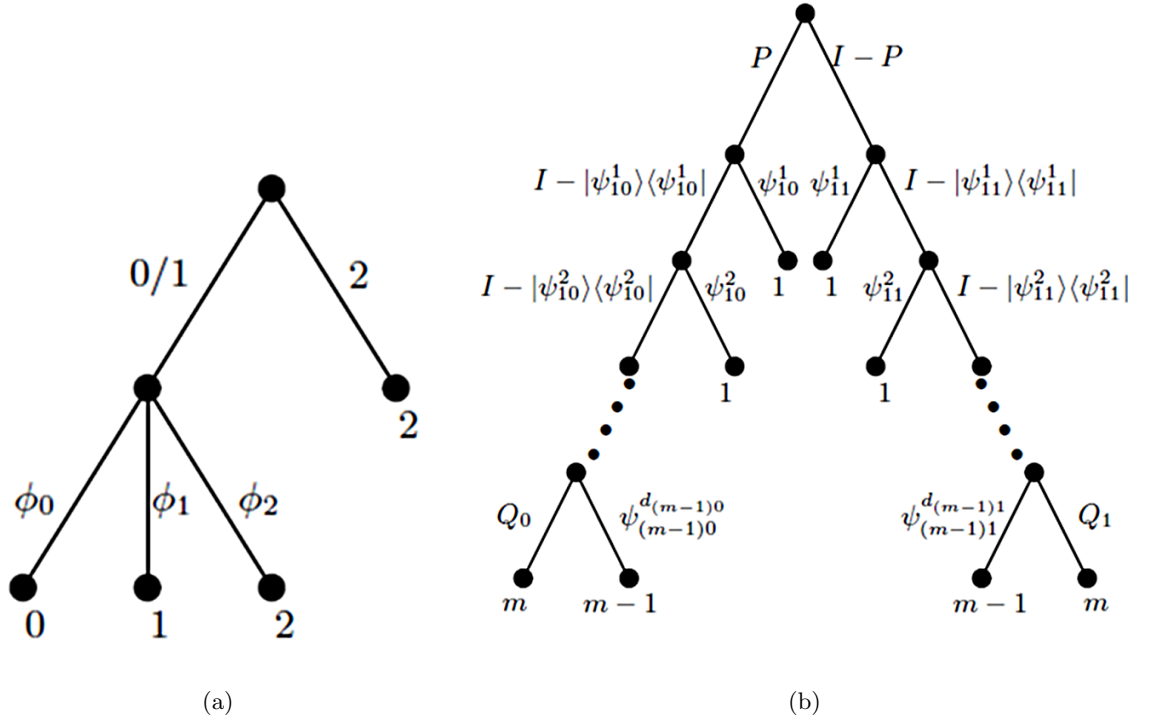


Figure 3.6: (a) A tree depiction of the example of the protocol from Wang & Ying [146]. The numbers 0/1,2 and the symbols ϕ_i next to each branch represents the immediate measurement result while the numbers at the end of each node represents the cumulative result of the preceding chain. (b) A binary search tree representation of the general procedure to realise POVMs by projective measurements without introducing ancillary dimensions from [146].

other measurements conditional on the first performed in succession, we can reduce the extension size and thus also the size of the accompanying unitary gate.

It is in fact possible to perform some POVMs without the use of any ancillary dimensions. Wang & Ying [146] considered the realisation of POVMs without ancillary dimensions and found a constraint under which a POVM could be implemented with iterated projective measurements and classical processing.

Consider, as per the example from Wang & Ying [146], a three dimensional system with a basis $\{|0\rangle, |1\rangle, |2\rangle\}$. It is first measured under the projective measurement $\{|0\rangle\langle 0| + |1\rangle\langle 1|, |2\rangle\langle 2|\}$. If it is projected into the $\{|0\rangle, |1\rangle\}$ subspace, another measurement is then performed in another basis $\{|\phi_0\rangle\langle \phi_0|, |\phi_1\rangle\langle \phi_1|, |\phi_2\rangle\langle \phi_2|\}$.

After the second step, we have four possible outcomes but the results $|2\rangle\langle 2|$ and $|\phi_2\rangle\langle \phi_2|$ can be coarse grained and output as the same. This has in effect created the

POVM $\{E_0, E_1, E_2\}$,

$$E_0 = (\mathbb{I} - |2\rangle\langle 2|)|\phi_0\rangle\langle\phi_0|(\mathbb{I} - |2\rangle\langle 2|), \quad (3.74)$$

$$E_1 = (\mathbb{I} - |2\rangle\langle 2|)|\phi_1\rangle\langle\phi_1|(\mathbb{I} - |2\rangle\langle 2|), \quad (3.75)$$

$$E_2 = (\mathbb{I} - |2\rangle\langle 2|)|\phi_2\rangle\langle\phi_2|(\mathbb{I} - |2\rangle\langle 2|) + |2\rangle\langle 2|. \quad (3.76)$$

To continue the example, if one selects

$$|\phi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle + |2\rangle), \quad (3.77)$$

$$|\phi_1\rangle = \frac{1}{\sqrt{14}}(|0\rangle + 2|1\rangle - 3|2\rangle), \quad (3.78)$$

$$|\phi_2\rangle = \frac{1}{\sqrt{42}}(5|0\rangle - 4|1\rangle - |2\rangle), \quad (3.79)$$

then one can show that the resulting POVM elements are

$$E_0 = \frac{2}{3}|\psi_0\rangle\langle\psi_0|, \quad E_1 = \frac{5}{14}|\psi_1\rangle\langle\psi_1|, \quad E_2 = \mathbb{I} - E_1 - E_0 \quad (3.80)$$

where

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |\psi_1\rangle = \frac{1}{\sqrt{5}}(|0\rangle + 2|1\rangle). \quad (3.81)$$

In a general treatment, the protocol is to apply a projective measurement $\{P_0, P_1, \dots, P_n\}$ and then after a result i , implementing projector P_i , to choose another measurement conditioned on i , $\{P_1^i, \dots, P_{n(i)}^i\}$, and then to repeat to build up the operation $P_{i_t}^{(i, i_1, \dots, i_{t-1})} \dots P_{i_2}^{(i, i_1)} P_{i_1}^i P_i$. The results of several of these operators are then summed together to create a higher rank result.

Wang & Ying [146] found a necessary and sufficient condition for a POVM to be implemented by this protocol. Each choice of projective measurement is conditioned on the results in the prior stages except of course the first unconditioned projective measurement. It also follows that the POVM elements must be a sum of the element with the projectors of an orthogonal basis:

$$E_k = \sum_j P_j E_k P_j. \quad (3.82)$$

Since every operator is a product $K = P_{i_t}^{(i, i_1, \dots, i_{t-1})} \dots P_{i_1}^i P_i$, the POVM elements will be

$$K^\dagger K = P_j \Pi P_j \quad (3.83)$$

for positive operator Π and so combing the conditions (3.82) and (3.83) show that each POVM element is a sum

$$E_k = \sum_j P_j \Omega_{kj} P_j, \quad \Omega > 0. \quad (3.84)$$

Therefore there must be an initial projection that commutes with all the elements E_k ,

$$[E_k, P_0] = 0, P_0 \neq \mathbb{I}, \quad (3.85)$$

since we must be able to perform this projective measurement and then still be able to perform the general measurement.

This condition becomes trivial if we consider implementing a POVM when we have access to just one extra dimension described by $|\Psi\rangle\langle\Psi|$, leading to a corollary of the above result. A POVM, $\{E_0, E_1, \dots, E_{d-1}\}$ on a d dimensional Hilbert space can be mapped to $\{E_0, E_1, \dots, E_{d-1}, |\Psi\rangle\langle\Psi|\}$ on a $d + 1$ dimensional space so the commuting projector is just the projector onto the ancilla dimension. Since the ancilla dimension will be initially empty one does not even have to enact the projective measurement. Therefore the corollary states that an arbitrary POVM on d dimensions can be realised with just a one-dimensional direct sum extension using an iterated sequence of projective measurements.

Chapter 4

Universality

4.1 Implementation of unitary channels through projections on an ancilla qubit

In light of the work on binary search trees for generalised measurements, ancilla driven quantum computation can be seen as an extension to a special case where the ultimate aim is not to implement state measures but to implement state independent unitary gates as the Kraus operators. Otherwise, several elements stay the same: an operation is performed by extending to a larger Hilbert space by a product extension; the additional dimensions are limited to a product with a qubit system and then performing projective measurements solely on the ancilla system; the target operation must be built up over several steps. In addition, just as the iterative POVM implementations had to be shown by their authors to be capable of implementing any arbitrary POVM, to be counted as ancilla driven *universal* quantum computation, the technique must be capable of implementing any arbitrary unitary gate or at least up to an efficient approximation.

Because unitary gates over any number of dimensions can be decomposed into single-qubit gates and an entangling two-qubit gate, we only have to consider implementing unitary operations on single-qubit and two-qubit systems and therefore limit the investigation to the results of enacting two-qubit unitaries and three-qubit unitaries on register qubits and the ancilla. In fact, because it is desirable to be able to use the same interaction between all qubits and because we are limited to no interactions between register qubits, the three qubit unitaries can be limited to cases where they can be expressed as an overlap of two two-qubit unitary gates that themselves are appropriate for enacting universal single-qubit gates on a register qubit.

4.1.1 The Cartan Decomposition of Two-Qubit gates

We will now derive the restrictions on the parameters for the ancilla preparation state, unitary gate coupling and measurement basis that allow one to enact a unitary single-qubit operation. Given the parameters of an initial ancilla state, $|a\rangle$, a post-measurement state $|m\rangle$ and a unitary $E \in U(4)$, we can construct the Kraus operator $K_m = \langle m|E|a\rangle$ and then evaluate whether the resulting operation is unitary from $KK^\dagger = K^\dagger K = p_m \mathbb{I}$ as per (3.47). However the unitarity of the resulting operation is unaffected by applying any pre- or post-corrective unitary gates to the register qubit. In addition, any unitary operations applied to the ancilla qubit before or after the measurement could be treated as a change in the preparation state or measurement basis. For a unitary E that fulfils the conditions under some specific choice of $|a\rangle$ and $|m\rangle$, there may be a class of unitary $\tilde{E} = (V_s \otimes V_a).E.(U_s \otimes U_a)$, $V, U \in U(2)$ which fulfils the conditions over a range of $|a\rangle$ and $|m\rangle$. Because of the symmetry of such a class under local unitary gates, the class can be described by a smaller set of parameters rather than having to deal with the 15 parameters of a two-qubit unitary.

Any two-qubit gate can be decomposed into the form $(V_s \otimes V_a).\Delta_{\alpha_x, \alpha_y, \alpha_z}.(U_s \otimes U_a)$ where

$$\Delta_{\alpha_x, \alpha_y, \alpha_z} = e^{-i(\alpha_x \sigma_x \otimes \sigma_x + \alpha_y \sigma_y \otimes \sigma_y + \alpha_z \sigma_z \otimes \sigma_z)} \quad (4.1)$$

which can also be diagonalised in the Magic State basis [147]:

$$\Delta = \sum_{k=1}^4 e^{-i\lambda_k} |\Phi_k\rangle \langle \Phi_k|; \quad (4.2)$$

this is a basis derived from the Bell states,

$$|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \quad (4.3)$$

$$|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle), \quad (4.4)$$

with

$$|\Phi_1\rangle = |\Phi^+\rangle, \quad \lambda_1 = \alpha_x - \alpha_y + \alpha_z, \quad (4.5)$$

$$|\Phi_2\rangle = -i|\Phi^-\rangle, \quad \lambda_2 = -\alpha_x + \alpha_y + \alpha_z, \quad (4.6)$$

$$|\Phi_3\rangle = |\Psi^-\rangle, \quad \lambda_3 = -\alpha_x - \alpha_y - \alpha_z, \quad (4.7)$$

$$|\Phi_4\rangle = -i|\Psi^+\rangle, \quad \lambda_4 = \alpha_x + \alpha_y - \alpha_z. \quad (4.8)$$

Two-qubit unitary gates that are equivalent up to local unitary gates can be characterised by just the three parameters $[\alpha_x, \alpha_y, \alpha_z]$. Certain operations on the three parameters cause only local unitary gate differences.

A sign change on a pair of parameters $\{\alpha_i, \alpha_j\}$ is equivalent to a pre- and post-application of the local unitary Pauli operator α_k where (i, j, k) are any permutation

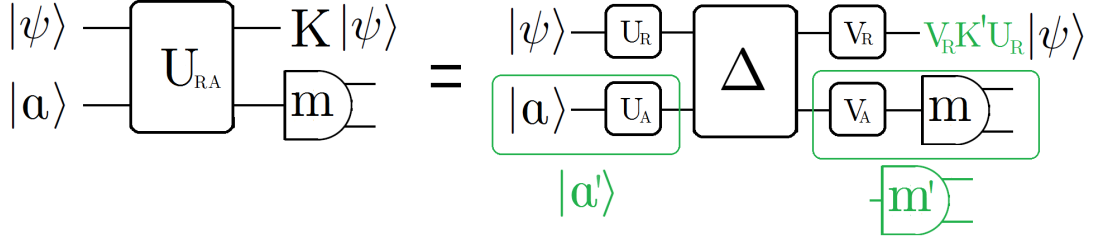


Figure 4.1: The Cartan decomposition of a two-qubit unitary gate and the allocation of local unitary gates to the ancilla preparation state and measurement basis. A preparation state of $|a\rangle$ for an ancilla state with the use of an arbitrary two-qubit gate U_{RA} can be treated as the preparation of $|a'\rangle = U_A|a\rangle$ with the use of the Δ component and likewise with the measurement basis $\{|m\rangle\} = \{V_A^\dagger|m\rangle\}$. Given the freedom of choice of ancilla preparation and measurement basis, therefore, the resulting action on the register system can be completely characterised by the Δ component.

of (x, y, z) . For example $e^{-i(\alpha_y\sigma_y\otimes\sigma_y+\alpha_z\sigma_z\otimes\sigma_z)} = (\mathbb{I} \otimes \sigma_x).e^{i(\alpha_y\sigma_y\otimes\sigma_y+\alpha_z\sigma_z\otimes\sigma_z)}.(\mathbb{I} \otimes \sigma_x)$. Since $e^{i\frac{\pi}{2}\sigma_i\otimes\sigma_i}\equiv\sigma_i\otimes\sigma_i$ the parameters will display a symmetry with respect to $\frac{\pi}{2}$ period shifts. In addition, just as the Hadamard gate transforms between the Pauli X and Z operators according to:

$$\begin{aligned} H &= \sigma_z.e^{-i\frac{\pi}{4}\sigma_y}, \\ H\sigma_z &= \sigma_x H, \\ H\sigma_y H &= -\sigma_y, \end{aligned}$$

the pre- and post-application of a pair of local Hadamard gates, $(H\otimes H).e^{i\alpha_x\sigma_x\otimes\sigma_x+\alpha_y\sigma_y\otimes\sigma_y+\alpha_z\sigma_z\otimes\sigma_z}.(H\otimes H)$ will swap the parameters $(\alpha_x, \alpha_y, \alpha_x) \rightarrow (\alpha_z, \alpha_y, \alpha_x)$. We can envisage a set of Hadamard gate like operators

$$H_x = \sigma_z e^{-i\frac{\pi}{4}\sigma_x}, \quad (4.9)$$

$$H_z = \sigma_x e^{-i\frac{\pi}{4}\sigma_z}, \quad (4.10)$$

$$H_y := H \quad (4.11)$$

that will enable us to equate any permutation of $(\alpha_x, \alpha_y, \alpha_z)$ to local unitary gate pre-and post-corrections.

The combination of these operations means that if $(\alpha_x, \alpha_y, \alpha_z)$ is part of a class of locally equivalent unitary gates then $(\alpha_i, \alpha_j, \alpha_k), (\frac{\pi}{2} - \alpha_i, \frac{\pi}{2} - \alpha_j, \alpha_k), (\frac{\pi}{2} - \alpha_i, \alpha_j, \frac{\pi}{2} - \alpha_k)$ and $(\alpha_i, \frac{\pi}{2} - \alpha_j, \frac{\pi}{2} - \alpha_k)$ where (i, j, k) are permutations of (x, y, z) are also part of the same class. These transformation are representations of the Weyl group so we can map these three parameters to co-ordinates in a 3 dimensional real space, $\hat{\mathbf{w}} = \alpha_x\hat{\mathbf{x}} + \alpha_y\hat{\mathbf{y}} + \alpha_z\hat{\mathbf{z}}$, and represent all the local equivalent classes as points in a Weyl chamber

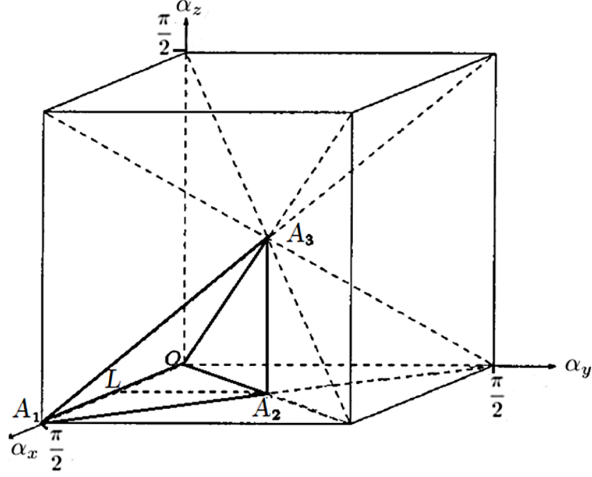


Figure 4.2: Graphical representation of the Weyl chamber. The permutations and shifts in values of the three parameters that lead to locally equivalent gates map to reflection in 2d planes in a 3d space [148]; the planes define a tetrahedron $OA_1A_2A_3$ which can represent the space of all non-equivalent $(\alpha_x, \alpha_y, \alpha_z)$.

[148]. The chamber is a tetrahedron $OA_1A_2A_3$ where $A_1 = (\frac{\pi}{2}, 0, 0)$, $A_2 = (\frac{\pi}{4}, \frac{\pi}{4}, 0)$ and $A_3 = (\frac{\pi}{4}, \frac{\pi}{4}, \frac{\pi}{4})$ except for the plane LA_2A_1 (for $L = (\frac{\pi}{4}, 0, 0)$) which is equivalent to OLA_2 [148, 149] (see figure 4.2).

4.1.2 Deriving the conditions for unitarity

For our purposes, the utility of the decomposition of two-qubit unitary gates is that if one element of a local equivalence class $E \in \tilde{\Delta}_{\mathbf{w}}$ has an ancilla preparation state and measurement basis such that $K_m = \langle m|E|a\rangle$ is proportional to unitary then all elements of $\tilde{\Delta}_{\mathbf{a}}$ will also have such a preparation state and basis. The preparation state and basis of other elements in the class can be found from the local unitary gates that transform between them (see figure 4.1).

The ancilla preparation state and measurement basis can be parametrised by

$$|a\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle, \quad (4.12)$$

$$|m\rangle = \cos\left(\frac{\zeta}{2}\right)|0\rangle + e^{i\xi}\sin\left(\frac{\zeta}{2}\right)|1\rangle, \quad (4.13)$$

$$|m_{\perp}\rangle = \sin\left(\frac{\zeta}{2}\right)|0\rangle - e^{i\xi}\sin\left(\frac{\zeta}{2}\right)|1\rangle. \quad (4.14)$$

The conditions for unitarity

α_x	α_y	α_z	θ	ϕ
0	0	—	—	—
0	—	0	—	—
—	0	0	—	—
0	—	—	—	$\frac{\pi}{2}$
—	0	—	—	0
—	—	0	$\frac{\pi}{2}$	—

Table 4.1: Parameter restrictions for an ancilla qubit driven unital map. The two-qubit gates are limited to the plane OLA_2 in the Weyl chamber with one free parameter on the interaction or two. Note that the plane condition given by ϕ subsumes some conditions on θ .

the requirement that the trace preserving channel is also *unital* [150, 151]. A channel is unital if it preserves the maximally mixed state i.e. the identity operator:

$$\Lambda(\mathbb{I}) = \mathbb{I} \rightarrow \sum_j K_j K_j^\dagger = \mathbb{I}. \quad (4.15)$$

From this, the condition can be expressed in a matrix representation:

$$\sum_j \mathbf{K}_j \mathbf{K}_j^\dagger = \begin{pmatrix} 1 + \delta & \epsilon \\ \epsilon^* & 1 - \delta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (4.16)$$

where

$$\delta = \cos(\theta) \sin(2\alpha_x) \sin(2\alpha_y) = 0, \quad (4.17)$$

$$\epsilon = \sin(2\alpha_z) \sin(\theta) [\sin(2\alpha_y) \cos(\phi) - i \sin(2\alpha_x) \sin(\phi)] = 0. \quad (4.18)$$

The restrictions that allow for a unital map are displayed in Table 4.1. Because of the limits on the degrees of freedom there is a trade off between the number of available free parameters of the two-qubit interaction gate $\Delta_{\hat{\mathbf{w}}}$. A Δ described by one parameter places no restrictions on the ancilla preparation state and all these permutations will be of the class $\tilde{\Delta}_\alpha$ so we consider these to be a single case. The two-parameter gates restrict the ancilla preparation state to a plane described by which of the parameters are free: $\alpha_i = 0$ restricts the ancilla state to the $\hat{\mathbf{j}} - \hat{\mathbf{k}}$ plane of the Bloch sphere. So the two-parameter gates can be taken as a single class whose transformations preserve the conditions on the ancilla. However there are no ancilla state parameters that would

The total number of parameters will be seven and after applying the unitary condition we expect to have three free parameters remaining. The task of narrowing down the free options can be split up into two stages; the first step is to fix the measurement basis to the computational basis $\{|j\rangle\}$, $j = 0, 1$ and to not in fact apply the unitary condition. Because of the unitary freedom in the choice of the basis of the trace when defining a channel according to (3.49), any basis can be used to examine conditions on the pre-measurement channel. At that point, the condition to be applied is that the channel can be expressed as a random unitary channel. Fortunately for a qubit system, this can be equated to

specify a state that allows for a three parameter gate to implement a unital channel, indicating that this type of interaction leaks information unavoidably.

The conditions for unitarity

To derive the conditions for unitarity, we construct the Kraus operators for the measurement result $|m\rangle$ from linear addition of the Kraus operators derived from the computational basis K_0, K_1

$$K_m = \cos\left(\frac{\zeta}{2}\right) K_0 + e^{-i\xi} \sin\left(\frac{\zeta}{2}\right) K_1. \quad (4.19)$$

Applying the unitary condition (3.47) gives

$$K_m K_m^\dagger = \cos^2\left(\frac{\zeta}{2}\right) K_{00} + \sin^2\left(\frac{\zeta}{2}\right) K_{11} + e^{i\xi} \cos\left(\frac{\zeta}{2}\right) \sin\left(\frac{\zeta}{2}\right) K_{01} + e^{-i\xi} \cos\left(\frac{\zeta}{2}\right) \sin\left(\frac{\zeta}{2}\right) K_{10} = \mathbb{I} \quad (4.20)$$

and its adjoint, where $K_{ij} = K_i K_j^\dagger$.

Consider the one-parameter interaction case $[\alpha_x = 0, \alpha_y = 0]$. The computational basis Kraus operators are

$$\mathbf{K}_0 = \cos\left(\frac{\theta}{2}\right) \begin{pmatrix} e^{-i\alpha_z} & 0 \\ 0 & e^{i\alpha_z} \end{pmatrix}, \quad \mathbf{K}_1 = e^{i\phi} \sin\left(\frac{\theta}{2}\right) \begin{pmatrix} e^{i\alpha_z} & 0 \\ 0 & e^{-i\alpha_z} \end{pmatrix}. \quad (4.21)$$

Applying the condition (4.20) and equating the elements of the identity operator with the elements of the Kraus operator matrix representation yields the condition

$$\cos(2\alpha_z - (\xi - \phi)) = \cos(2\alpha_z + (\xi - \phi)), \quad (4.22)$$

given arbitrary α_z , the solution to this is

$$(\xi - \phi) \bmod \pi = 0, \quad (4.23)$$

in other words, the ancilla preparation state and measurement basis must be in the same vertical plane.

When the two-qubit gate is transformed to other members of the one-parameter class, the local unitary gate transformations can be corrected by applying Hadamard set gates to the ancilla preparation state and measurement basis. Thus if there was any one-parameter gate $e^{-i\alpha\sigma_{\hat{n}}}$ associated with a general vector on the Bloch sphere $\hat{\mathbf{n}}$, the measurement basis would have to lie in the same plane as the vector $\hat{\mathbf{n}}$ and the initial state $|a\rangle$.

If we now consider a two-parameter interaction case $[y = 0, \phi = 0]$, the computational basis Kraus operators are

$$\begin{aligned} \mathbf{K}_0 &= \begin{pmatrix} \cos(\alpha_x) \cos\left(\frac{\theta}{2}\right) e^{-i\alpha_z} & -i \sin(\alpha_x) \sin\left(\frac{\theta}{2}\right) e^{-i\alpha_z} \\ -i \sin(\alpha_x) \sin\left(\frac{\theta}{2}\right) e^{i\alpha_z} & \cos(\alpha_x) \cos\left(\frac{\theta}{2}\right) e^{i\alpha_z} \end{pmatrix}, \\ \mathbf{K}_1 &= \begin{pmatrix} \cos(\alpha_x) \sin\left(\frac{\theta}{2}\right) e^{i\alpha_z} & -i \sin(\alpha_x) \cos\left(\frac{\theta}{2}\right) e^{i\alpha_z} \\ -i \sin(\alpha_x) \cos\left(\frac{\theta}{2}\right) e^{-i\alpha_z} & \cos(\alpha_x) \sin\left(\frac{\theta}{2}\right) e^{-i\alpha_z} \end{pmatrix}. \end{aligned} \quad (4.24)$$

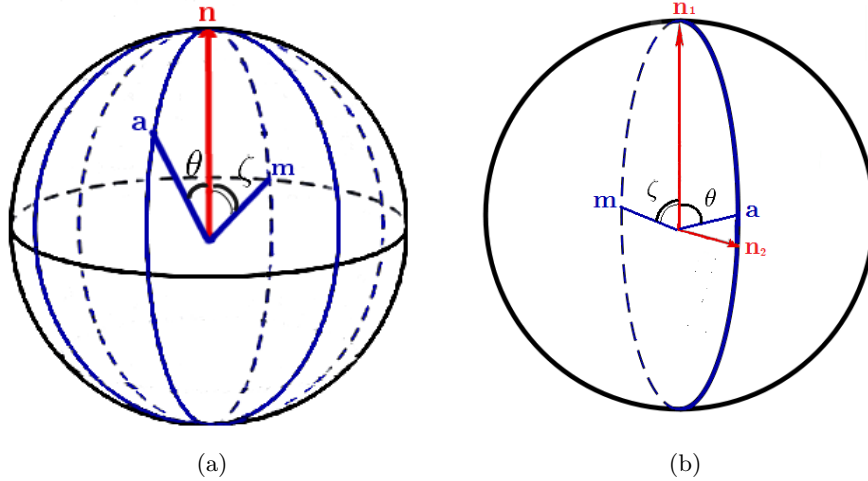


Figure 4.3: A Bloch sphere representation of the conditions on the preparation and measurement state, at positions \hat{a} and \hat{m} respectively, given the vector of interaction \hat{n} for (a) a single interaction parameter (b) two interaction parameters.

This generates the parameter condition $\xi = 0$ (including the case where $\zeta = 0$) thereby restricting the measurement basis to the same plane as the preparation state again but with the added restriction that the plane must be the one that intersects the vectors of the two free interaction parameters (see figure 4.3 for a Bloch sphere representation of these conditions) which is consistent as an intersection of the conditions of the two-parameters individually. Again, we found that these conditions rotate around the Bloch sphere along with the axis of the free parameters.

Understanding the unitary condition

To understand the unitary conditions, recall that the time symmetric and norm preserving nature of unitary operations means that they extract no information about the system they are performed on. Yet in section 3.1.2 we have seen how the ancilla model can function just as well as a measurement device. When ancilla states are correlated with states or subspaces of the original system, measuring the ancilla in that state projects the system into the correlated state or subspace. Even when the ancilla states are non-orthogonal the measurement of one state applies a biased projector as in (3.33) but since it has not completely projected the system into one state, one could send a second ancilla and measure in the basis of the other state then one would get the same POVM as (3.33) but with the basis flipped. If one measured both states the product is in fact just the identity operator and we have learnt nothing. It is natural then to consider that a measurement that is symmetric with respect to both ancilla states will also not extract any more information.

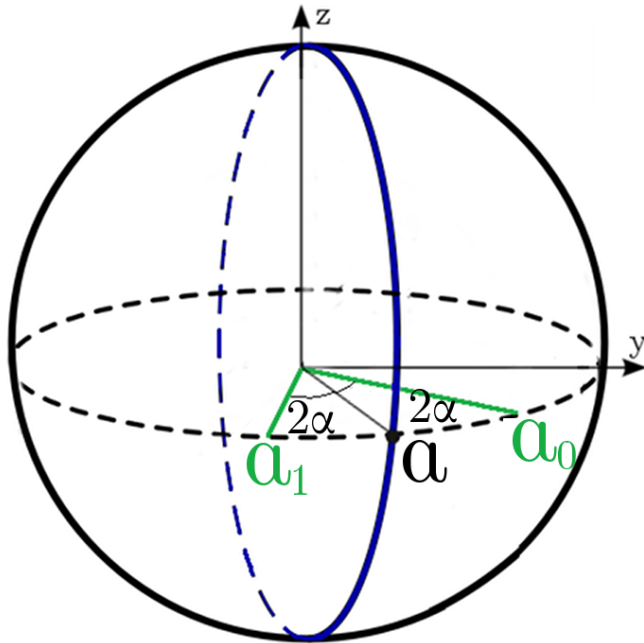


Figure 4.4: Representation of the intermediary ancilla states $|a_j\rangle$ on the Bloch sphere as \hat{a}_j with the symmetric plane of measurement (blue).

This is proven mathematically rather simply when dealing with one-parameter interaction gates. When the ancilla interacts with a register qubit via a symmetric interaction $e^{i\alpha\sigma_i\otimes\sigma_i}$, the ancilla or the register being in a pre-determined eigenstate of σ_i enacts a deterministic unitary operation on the other subsystem. So for a general register state, we can describe the evolution of the pair in terms of the back-action on the ancilla:

$$|a\rangle \sum_j c_j |j\rangle \rightarrow \sum_j c_j |a_j\rangle |j\rangle. \quad (4.25)$$

Because $e^{i\alpha\sigma_i\otimes\sigma_i}$ is symmetric with respect to the $j = 0 \leftrightarrow 1$ exchange, the two ancilla states $|a_j\rangle$ are symmetric about the initial state $|a\rangle$ and are just given by rotations about the principle axis of the interaction:

$$|a_j\rangle = R_{\hat{z}}(\pm 2\alpha)|a\rangle. \quad (4.26)$$

When the ancilla is then measured in the state $|m\rangle$ the resulting Kraus operator is

$$K_m = \sum_j \langle m|a_j\rangle |j\rangle\langle j|. \quad (4.27)$$

The condition (3.47), yields the condition $|\langle m|a_j\rangle|^2 = p_m$, constant over j and so $|m\rangle$ must lie on the plane on which was $|a\rangle$ originally (see figure 4.4).

For the two-parameter interactions of the class $\Delta_{(\alpha_i, \alpha_j)}$, there are no eigenstates that result in a deterministic unitary action but we know that the conditions must match

the conditions for the single parameter α_i as $\alpha_j \rightarrow 0$. Since the conditions for the one-parameter case are independent of the interaction parameters coupling strength, the two-parameter conditions must be the overlap of the individual one-parameter conditions. Hence the ancilla preparation and measurement must be on the plane of both principle axes.

α_i	α_j	α_k	$ a\rangle, m\rangle$
0	0	—	$\langle m \sigma_k a\rangle \in \mathbb{R}$
0	—	—	$\langle m \sigma_j a\rangle, \langle m \sigma_k a\rangle \in \mathbb{R}$

Table 4.2: Parameter restrictions for an ancilla qubit driven unitary channel. The available interaction gates are generalised to any one or pair of orthogonal spin-spin product terms. If given an arbitrary axis to associate the spin-spin products, the condition of whether the measurement and preparation states lie in the appropriate plane on the Bloch sphere can be evaluated by the realness of the overlap between measurement state and the spin operator on the preparation state.

4.1.3 Available unitary gates

This insight into how the unitary conditions are formed also provides for how we derive the available unitary gates. As per (4.27), the Kraus operator is diagonal in the eigenstate basis of the operator associated with the principle axis of the interaction gate, σ_i , of the single parameter gate $e^{-i\alpha\sigma_i\otimes\sigma_i}$. Using $\hat{\mathbf{z}}$ to represent the principle axis of all single parameter interaction gates from now on, the unitary gate for a single parameter must be of the form $e^{-i\frac{\gamma}{2}\sigma_z}$ (alternatively represented as $R_{\hat{\mathbf{z}}}(\gamma)$). Also, as per (4.19), as a result of Choi's theorem and the Stinespring dilation theorem, a Kraus operator resulting from a particular measurement (or preparation) can be constructed from the linear addition of the Kraus operators of a measurement (or preparation) in a unitarily related basis. For one-parameter interaction gates, preparing or measuring in eigenstates of σ_z lead to operations $e^{\pm i\alpha\sigma_z}$ with terms of proportionality being the amplitudes of the initial (or measurement) state $|a\rangle$:

$$\begin{aligned}
\text{Eqn.(3.50)} \rightarrow K_j &= \langle j|e^{-i\alpha\sigma_z\otimes\sigma_z}|a\rangle = \langle j|e^{-i\alpha\sigma_z\otimes\sigma_z} \sum_k c_k |k\rangle \\
&= \sum_k c_k \langle j|k\rangle e^{-i(-1)^k \alpha\sigma_z} \\
&= c_j e^{-i(-1)^j \alpha\sigma_z}.
\end{aligned} \tag{4.28}$$

Therefore any Kraus operator is equal under diagonalisation to

$$K_m = \cos\left(\frac{\zeta}{2}\right) \cos\left(\frac{\theta}{2}\right) e^{-i\alpha\sigma_z} + e^{i(\phi-\xi)} \sin\left(\frac{\zeta}{2}\right) \sin\left(\frac{\theta}{2}\right) e^{i\alpha\sigma_z} \tag{4.29}$$

or in a matrix representation, with $\phi - \xi = 0$ to fulfil the unitary condition,

$$\mathbf{K}_m = \begin{pmatrix} \cos\left(\frac{\zeta}{2}\right)\cos\left(\frac{\theta}{2}\right)e^{-i\alpha} + \sin\left(\frac{\zeta}{2}\right)\sin\left(\frac{\theta}{2}\right)e^{i\alpha} & 0 \\ 0 & \cos\left(\frac{\zeta}{2}\right)\cos\left(\frac{\theta}{2}\right)e^{i\alpha} + \sin\left(\frac{\zeta}{2}\right)\sin\left(\frac{\theta}{2}\right)e^{-i\alpha} \end{pmatrix} \quad (4.30)$$

$$= \sqrt{p_m} \begin{pmatrix} e^{-i\gamma} & 0 \\ 0 & e^{i\gamma} \end{pmatrix}. \quad (4.31)$$

The probability of measurement is

$$\begin{aligned} p_m &= \left| \cos\left(\frac{\zeta}{2}\right)\cos\left(\frac{\theta}{2}\right)e^{i\alpha} + \sin\left(\frac{\zeta}{2}\right)\sin\left(\frac{\theta}{2}\right)e^{-i\alpha} \right|^2 \\ &= \cos^2\left(\frac{\zeta - \theta}{2}\right) - \sin(\zeta)\sin(\theta)\sin^2(\alpha) \end{aligned} \quad (4.32)$$

and the relative phase imparted is

$$\Gamma = 2\gamma = 2\text{Arg} \left[\cos\left(\frac{\zeta}{2}\right)\cos\left(\frac{\theta}{2}\right)e^{i\alpha} + \sin\left(\frac{\zeta}{2}\right)\sin\left(\frac{\theta}{2}\right)e^{-i\alpha} \right]. \quad (4.33)$$

There are a couple of ways we can rewrite this phase expression:

$$\text{Arg} \left[\cos\left(\frac{\zeta}{2}\right)\cos\left(\frac{\theta}{2}\right)e^{i\alpha} + \sin\left(\frac{\zeta}{2}\right)\sin\left(\frac{\theta}{2}\right)e^{-i\alpha} \right] = \arctan \left[\tan(\alpha) \frac{\cos\left(\frac{\zeta+\theta}{2}\right)}{\cos\left(\frac{\zeta-\theta}{2}\right)} \right] \quad (4.34)$$

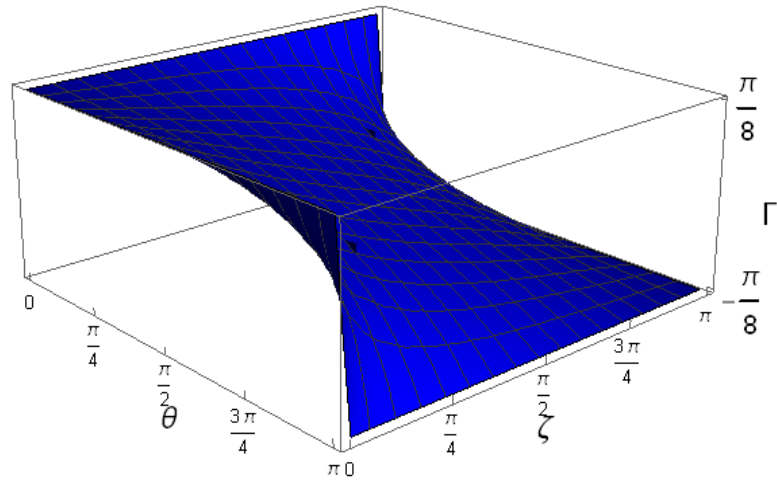
$$= \arctan \left[\tan(\alpha) \frac{\langle m | \sigma_z | a \rangle}{\langle m | a \rangle} \right]. \quad (4.35)$$

The orthogonal measurement state $|m_\perp\rangle$ that provides the other measurement result is $\sin\left(\frac{\zeta}{2}\right)|0\rangle - \cos\left(\frac{\zeta}{2}\right)|1\rangle$ whose results can also be calculated applying a shift $\zeta \rightarrow \zeta + \pi$. As can be observed in figures 4.5a and 4.5b and as can be inferred from (4.34), the features of the Γ values are dominated by the relative value of ζ and θ rather than the absolute value so they can be easily be studied by fixing the preparation state as an equal superposition and then varying only the measurement basis as in figure 4.6.

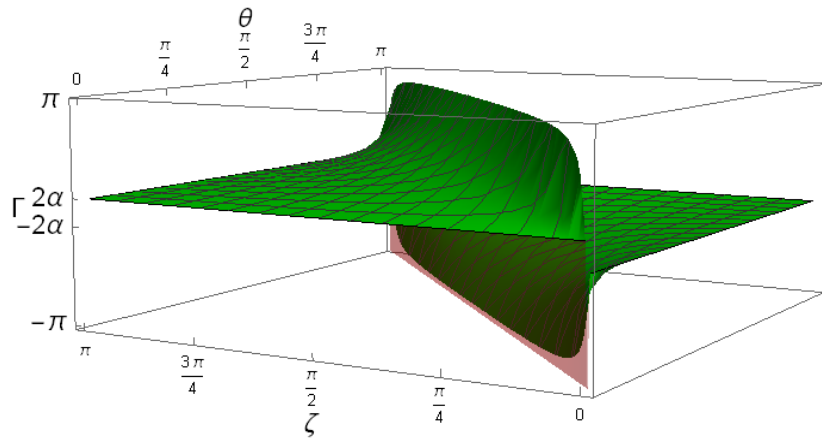
Special case: orthogonal measurement to preparation

If the ancilla is measured in a state orthogonal to its initial state so that $\langle m | a \rangle = 0$ then we can see from (4.33) and (4.35) that the relative phase imparted will be -1 and the unitary operation a Pauli gate, σ_z since as $\langle m | a \rangle \rightarrow 0$, (4.35) $\rightarrow \frac{\pi}{2}$. The measurement occurs with a probability of

$$p_m = \sin^2(\theta)\sin^2(\alpha). \quad (4.36)$$



(a)



(b)

Figure 4.5: Plot of the relative phase of the single-qubit output gate, Γ , against ancilla parameters for preparation, θ , and measurement, ζ , from (a) the more likely result (b) the less likely result, given $\alpha = \frac{\pi}{16}$

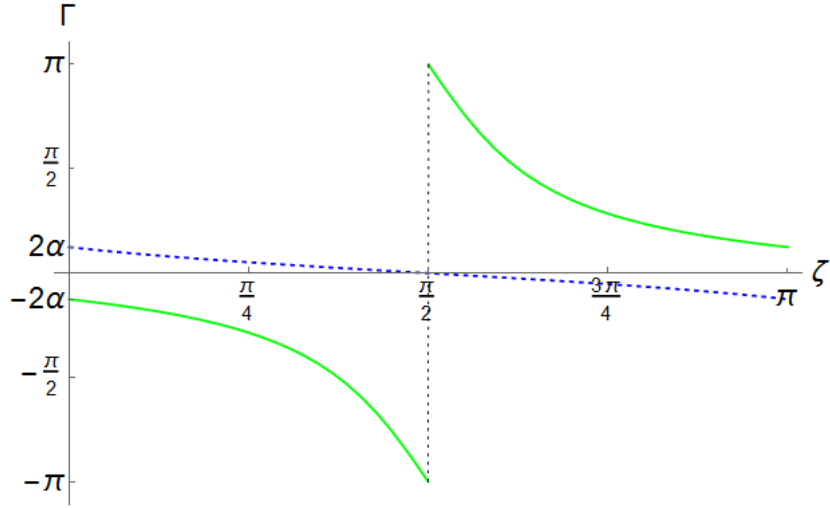


Figure 4.6: Plot of the relative phase of the single-qubit output gate, Γ , against ancilla parameters for measurement, θ , for fixed preparation in the $|+\rangle$ state, given $\alpha = \frac{\pi}{16}$, from the more likely result (the dotted blue curve) and the less likely result (the solid green curve).

If we consider the other possible measurement, measuring the ancilla in its initial state, then note that the phase depends on the property $\langle a|\sigma_z|a\rangle$. So consider preparing the ancilla in the $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ equal superposition where $\langle a|\sigma_z|a\rangle = 0$. This will actually return us the identity operations:

$$K_+ = \cos^2(\alpha)\mathbb{I}, \quad (4.37)$$

$$K_- = \sin^2(\alpha)\sigma_z. \quad (4.38)$$

Preparing and measuring the ancilla in the $|\pm\rangle$ basis thereby provides a guaranteed set of operations, independent of α , with the parameter of the interaction gate only affecting the probability of generation.

Unitary gates of two-parameter interactions

Unlike the one-parameter classes of gates, the two-parameter classes of gates locally equivalent to $e^{-i(\alpha_z\sigma_z\otimes\sigma_z+\alpha_x\sigma_x\otimes\sigma_x)}$ for $\alpha_x < \alpha_z < \frac{\pi}{4}$ do not have an ancilla preparation state which can induce a deterministic unitary operation. As a result there is no fixed basis of states in which the resulting Kraus operators can be diagonalised. An operation that is generated of the general form $e^{-i\frac{\gamma}{2}\hat{n}\cdot\hat{\sigma}}$ may have γ and $\hat{\sigma}$ as functions of (α_x, α_z) , functions which we must find.

Using the case where $[y = 0, \phi = 0]$ to represent the two-parameter class, we take the computational basis results from (4.24) and apply (4.19) to give a matrix representation

of the Kraus operators for a general measurement:

$$\mathbf{K}_m = \begin{pmatrix} \cos(\alpha_x)k_{00} & -i\sin(\alpha_x)k_{01} \\ -i\sin(\alpha_x)k_{01}^* & \cos(\alpha_x)k_{00}^* \end{pmatrix} \quad (4.39)$$

where

$$k_{00} = \cos\left(\frac{\zeta}{2}\right) \cos\left(\frac{\theta}{2}\right) e^{-i\alpha} + \sin\left(\frac{\zeta}{2}\right) \sin\left(\frac{\theta}{2}\right) e^{i\alpha}, \quad (4.40)$$

$$k_{01} = \cos\left(\frac{\zeta}{2}\right) \sin\left(\frac{\theta}{2}\right) e^{-i\alpha} + \sin\left(\frac{\zeta}{2}\right) \cos\left(\frac{\theta}{2}\right) e^{i\alpha}. \quad (4.41)$$

The Pauli product operators $\{\sigma_i \otimes \sigma_i\}$, unlike the individual Pauli operators, commute with each other so a two-parameter interaction gate can be split into a product of one-parameter gates $e^{-i(\alpha_z \sigma_z \otimes \sigma_z + \alpha_x \sigma_x \otimes \sigma_x)} = e^{-i\alpha_z \sigma_z \otimes \sigma_z} e^{-i\alpha_x \sigma_x \otimes \sigma_x}$. Because of this, there are the special cases where the ancilla is prepared or measured in the eigenstate of one of the factors:

$$\langle m | e^{-i(\alpha_z \sigma_z \otimes \sigma_z + \alpha_x \sigma_x \otimes \sigma_x)} | \pm \rangle = \langle m | e^{-i\alpha_z \sigma_z \otimes \sigma_z} | \pm \rangle e^{\mp i\alpha_x \sigma_x} = R_{\hat{z}}(\gamma) R_{\hat{x}}(\pm 2\alpha_x), \quad (4.42)$$

$$\langle \pm | e^{-i(\alpha_z \sigma_z \otimes \sigma_z + \alpha_x \sigma_x \otimes \sigma_x)} | a \rangle = e^{\mp i\alpha_x \sigma_x} \langle \pm | e^{-i\alpha_z \sigma_z \otimes \sigma_z} | a \rangle = R_{\hat{x}}(\pm 2\alpha_x) R_{\hat{z}}(\gamma), \quad (4.43)$$

$$\langle m | e^{-i(\alpha_z \sigma_z \otimes \sigma_z + \alpha_x \sigma_x \otimes \sigma_x)} | j \rangle = \langle m | e^{-i\alpha_x \sigma_x \otimes \sigma_x} | j \rangle e^{(-1)^j i\alpha_z \sigma_z} = R_{\hat{x}}(\gamma) R_{\hat{z}}(\pm 2\alpha_x), \quad (4.44)$$

$$\langle j | e^{-i(\alpha_z \sigma_z \otimes \sigma_z + \alpha_x \sigma_x \otimes \sigma_x)} | a \rangle = e^{(-1)^j i\alpha_z \sigma_z} \langle j | e^{-i\alpha_x \sigma_x \otimes \sigma_x} | a \rangle = R_{\hat{z}}(\pm 2\alpha_x) R_{\hat{x}}(\gamma). \quad (4.45)$$

While the factorisation of the two-parameter gate has no order preference, $e^{-i\alpha_z \sigma_z \otimes \sigma_z} e^{-i\alpha_x \sigma_x \otimes \sigma_x} = e^{-i\alpha_x \sigma_x \otimes \sigma_x} e^{-i\alpha_z \sigma_z \otimes \sigma_z}$, choosing whether the ancilla preparation state or measurement basis is the eigenstate selects the order in which rotations about the principle axes are applied. Since only one of the preparation or measurement has been fixed, one can still use the other to find the special cases found in the one-parameter case. It is possible to both prepare and measure in an eigenstate basis, one for each parameter, so that the unitary operation is equivalent to

$$R_{\hat{x}}(\pm 2\alpha_x) R_{\hat{z}}(2\alpha_z).$$

One can also select a measurement in the basis that includes the preparation state yielding unitary operations equivalent to:

$$R_{\hat{z}}(2\alpha_z) \& \sigma_x R_{\hat{z}}(2\alpha_z).$$

However outside of these special cases, one has to rely on the elements of the unitary matrix composition in (2.10):

$$R_{\hat{n}}(\gamma) = e^{-i\frac{\gamma}{2} \hat{n} \cdot \hat{\sigma}} = \cos\left(\frac{\gamma}{2}\right) \mathbb{I} - i \sin\left(\frac{\gamma}{2}\right) \hat{n} \cdot \hat{\sigma},$$

$$\mathbf{R}_{\hat{n}}(\gamma) = \begin{pmatrix} \cos\left(\frac{\gamma}{2}\right) - i \sin\left(\frac{\gamma}{2}\right) n_z & -i \sin\left(\frac{\gamma}{2}\right) (n_x - i n_y) \\ -i \sin\left(\frac{\gamma}{2}\right) (n_x + i n_y) & \cos\left(\frac{\gamma}{2}\right) + i \sin\left(\frac{\gamma}{2}\right) n_z \end{pmatrix}.$$

From the elements of the Kraus operator, we can find the relative terms

$$\frac{n_x}{n_z} = \frac{\langle 0|K_m|1\rangle + \langle 1|K_m|0\rangle}{\langle 0|K_m|0\rangle - \langle 1|K_m|1\rangle}, \quad (4.46)$$

$$\frac{n_y}{n_z} = -i \frac{\langle 0|K_m|1\rangle - \langle 1|K_m|0\rangle}{\langle 0|K_m|0\rangle - \langle 1|K_m|1\rangle}. \quad (4.47)$$

Absolute values for n_x , n_y and n_z can be found by applying normalisation s.t. $|\hat{\mathbf{n}}|^2 = 1$.

Using the elements of the Kraus operator and the relationships

$$\cos\left(\frac{\zeta}{2}\right)\cos\left(\frac{\theta}{2}\right) - \sin\left(\frac{\zeta}{2}\right)\sin\left(\frac{\theta}{2}\right) = \langle m|\sigma_z|a\rangle, \quad (4.48)$$

$$\cos\left(\frac{\zeta}{2}\right)\sin\left(\frac{\theta}{2}\right) + \sin\left(\frac{\zeta}{2}\right)\cos\left(\frac{\theta}{2}\right) = \langle m|\sigma_x|a\rangle, \quad (4.49)$$

$$\cos\left(\frac{\zeta}{2}\right)\sin\left(\frac{\theta}{2}\right) - \sin\left(\frac{\zeta}{2}\right)\cos\left(\frac{\theta}{2}\right) = \langle m|\sigma_z\sigma_x|a\rangle = i\langle m|\sigma_y|a\rangle, \quad (4.50)$$

the vector of the axis of rotation can be found to depend upon the interaction gate parameters, the ancilla preparation state, and ancilla measurement basis. These parameters produce rotations about different axes for each of the two ancilla measurement results:

$$\frac{n_x}{n_z} = \frac{\tan(\alpha_x)\langle m|\sigma_x|a\rangle}{\tan(\alpha_z)\langle m|\sigma_z|a\rangle}, \quad (4.51)$$

$$\frac{n_y}{n_z} = \tan(\alpha_x) \frac{-i\langle m|\sigma_y|a\rangle}{\langle m|\sigma_z|a\rangle}. \quad (4.52)$$

$$(4.53)$$

The rotation angle is found from

$$\cos\left(\frac{\gamma}{2}\right) = \frac{1}{2\sqrt{p_m}}[\langle 0|K_m|0\rangle + \langle 1|K_m|1\rangle] = \frac{1}{\sqrt{p_m}}\cos(\alpha_x)\cos(\alpha_z)\langle m|a\rangle \quad (4.54)$$

where the probability of the result, p_m is

$$p_m = \frac{1}{2}[1 + \cos(\zeta)\cos(\theta)\cos(2\alpha_x) + \sin(\zeta)\sin(\theta)\cos(2\alpha_z)]. \quad (4.55)$$

If we return to the case where the measurement basis includes the preparation state then when $|m\rangle = |a\rangle$ these terms will depend on the expectation values $\langle \sigma_i \rangle_a$ which are the coordinates of the state $|a\rangle$ on the Bloch sphere. A result of $|m\rangle$ orthogonal to $|a\rangle$ enforces $\cos\left(\frac{\gamma}{2}\right) = 0$ as with the single parameter case. If one then goes a step further and restricts $\alpha_x = \alpha_z$ then we see that the measurement $|m\rangle = |a\rangle$ also enacts identity.

4.2 Ancilla Driven Entangling Gates

In the previous chapter, we covered the ancilla driven generation of single-qubit gates and replicated the description of the conditions and resulting gates given by [38, 37].

We saw that the range of two-qubit gates available as resources for probabilistic single-qubit gate generations was larger than the set of two local equivalence classes that allow for determinism up to Pauli corrections. If the large range of interaction gates is to be viable for universal quantum computation, they must also be capable of enacting an entangling two-qubit gate generation on the register.

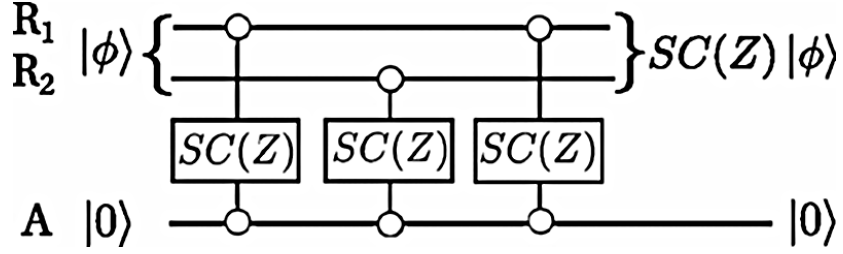
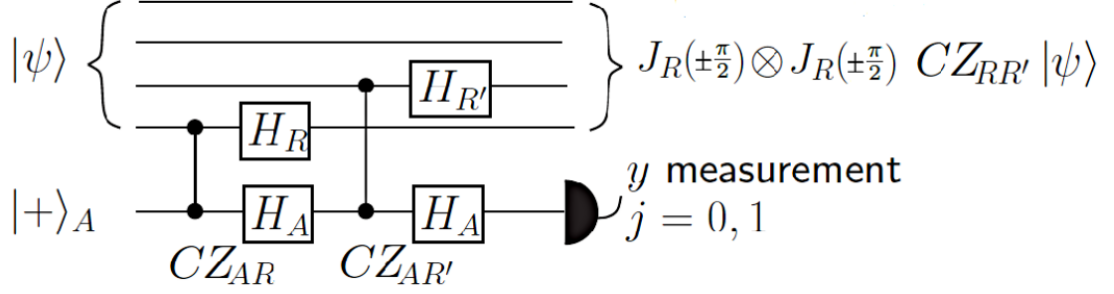


Figure 4.7: Circuit implementing a two-qubit gate on a register using (a) a one-parameter class gate ($H \otimes HCZ$) [38], (b) a two-parameter class gate $SC(Z)$ [43].

In the original ADQC model, a fixed interaction gate in the class $\Delta_{(\frac{\pi}{4}, 0, 0)}$ (locally equivalent to CZ) would couple the ancilla with one register qubit, local unitary gates would be applied and then the ancilla would undergo the same interaction gate with the second register qubit before being measured (see figure 4.7a). This approach with a single interaction gate per qubit per ancilla is in keeping with the treatment of the individual interaction gates as the basic building block of any three qubit gate and using the most efficient number of applications of the CZ gate. On the other hand, if we wanted to explore all possibilities of what could be done over the full range of gates we could consider different circuit patterns. For example, the CZ.SWAP (or $SC(Z)$) gate which is of a two-parameter equivalence class can be used without an ancilla measurement if a second interaction gate with the first qubit is applied (see figure 4.7b).

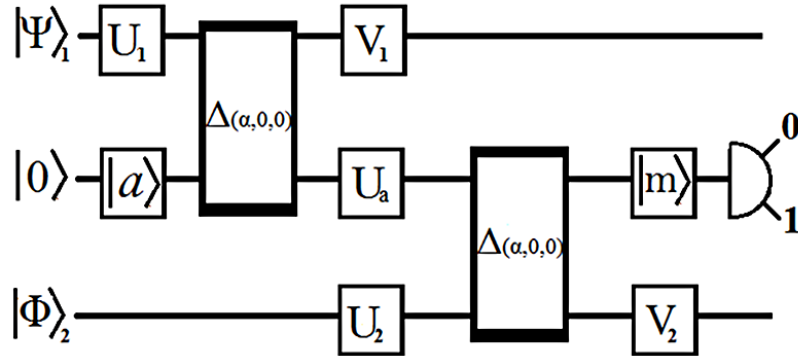


Figure 4.8: The general form of the ADQC two-qubit gate model. Differences between elements of the class $\tilde{\Delta}_\alpha$ can be absorbed into the ancilla preparation and measurement plus an intermediate local unitary U_a and into local unitary gate pre- and post-corrections on the register qubits.

The latter model was explored for the two-parameter interaction class concurrently with this work by Proctor & Kendon [43]. We will confine ourselves to the one-interaction gate per register qubit per ancilla rule (let us call it the *one-shot* restriction) with fixed interaction gate because it allows for a like to like comparison of resources with the original ADQC model. It also reflects some physical aspects of the physical paradigm ADQC is meant to represent. A fixed interaction gate equivalent to $e^{-i\alpha\sigma_z\otimes\sigma_z}$ with arbitrary α can represent the low-tuning, fast scattering hybrid physical approaches discussed in section 2.4.2. The one-shot restriction can reflect the passage of a flying qubit that travels over long distances between register qubits. In that context, the circuit pattern of figure 4.8 distinguished ADQC from the models of [152, 42, 153] and we wish to maintain this distinction.

4.2.1 Conditions of the two-qubit entangling gate generation

Techniques from the single-qubit gate generation can be used for the two-qubit generations. We can start by considering the one-parameter interaction gates before using the resulting conditions to explore the two-parameter case and thus we can look at the unitary condition given by $|\langle m|a_j\rangle|^2 = p_m$ for a computational basis j of the register.

Because the ancilla interacts with each register qubit individually with an interaction gate that is diagonal in the basis of each register subsystem, the back-action of the ancilla can be described by four states that correspond to the product of the computational bases of the two register qubits:

$$|a\rangle|\Psi\rangle_{12} = |a\rangle \sum_{jk} c_{jk}|j\rangle_1|k\rangle_2 \rightarrow \sum_{jk} |a_{jk}\rangle|j\rangle_1|k\rangle_2. \quad (4.56)$$

Now that there are four points on the Bloch sphere, $|a_{jk}\rangle$, the measurement basis $\{|m\rangle\}$

is more restricted. To ensure unitary backaction, there must be a measurement basis corresponding to two points on the surface of the Bloch sphere which are equidistant from those four points. Thus the four points must describe a plane that cuts the Bloch sphere into two caps which leaves only two points equidistant from those four points—the summits of the minor and major caps—to ensure unitary backaction.

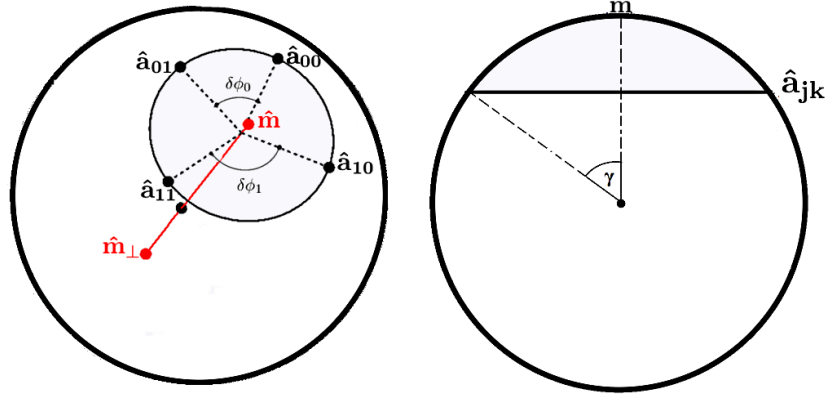


Figure 4.9: For four states that have the same value $|\langle m|a_{jk}\rangle|$, there are four points on the Bloch sphere that define a ring that encircle and thus define the state $|m\rangle$. $|\langle m|a_{jk}\rangle|^2 = \cos^2(\frac{\gamma}{2})$.

The Kraus operator matrix representation will be $\text{diag}(\langle m|a_{00}\rangle, \langle m|a_{01}\rangle, \langle m|a_{10}\rangle, \langle m|a_{11}\rangle)$ with

$$\langle m|a_{jk}\rangle = |\langle m|a_{jk}\rangle|e^{i\phi_{jk}} = \sqrt{p_m}e^{i\phi_{jk}} \quad (4.57)$$

therefore the resulting unitary is

$$U_m = \text{diag}(e^{i\phi_{00}}, e^{i\phi_{01}}, e^{i\phi_{10}}, e^{i\phi_{11}}). \quad (4.58)$$

This is in a one-parameter class too. In the Bloch sphere geometric picture, the phases $e^{i\phi_{ij}}$ will correspond to angles of the points $|a_{ij}\rangle$ on the cap (figure 4.9). We require not only that the gate is unitary but also that it is entangling. The gates of a $\tilde{\Delta}_{(\alpha,0,0)}$ class are all equivalent to a Control-Unitary gate $C_{\hat{n}}(\gamma) = |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes e^{-i\frac{\gamma}{2}\hat{n}\cdot\hat{\sigma}}$. This allows a very simple measurement of entangling power that will be unique in the $\alpha_i < \frac{\pi}{4}$ subspace of the Weyl chamber of interaction gates [147].

Given a two-qubit unitary of the form $\text{diag}(e^{i\phi_{00}}, e^{i\phi_{01}}, e^{i\phi_{10}}, e^{i\phi_{11}})$ we can multiply it by local unitary gates

$$\begin{pmatrix} e^{-ia_1} & \\ & e^{-ia_2} \end{pmatrix} \otimes \begin{pmatrix} e^{-ib_1} & \\ & e^{-ib_2} \end{pmatrix} = \text{diag}(e^{-i(a_1+b_1)}, e^{-i(a_1+b_2)}, e^{-i(a_2+b_1)}, e^{-i(a_2+b_2)}) \quad (4.59)$$

to give $\text{diag}(e^{i\phi_{00}-(a_1+b_1)}, e^{i\phi_{01}-(a_1+b_2)}, e^{i\phi_{10}-(a_2+b_1)}, e^{i\phi_{11}-(a_2+b_2)})$

We can choose a_1, b_1, a_2, b_2 such that

$$\phi_{00} - (a_1 + b_1) = 0, \quad (4.60)$$

$$\phi_{01} - (a_1 + b_2) = 0, \quad (4.61)$$

$$\phi_{10} - (a_2 + b_1) = 0 \quad (4.62)$$

$$\rightarrow a_2 + b_2 = a_2 + b_1 - (a_1 + b_1) + (a_1 + b_2) = \phi_{10} - \phi_{00} + \phi_{01}, \quad (4.63)$$

the resulting gate (4.58) must therefore be equivalent to the form

$$\tilde{\mathbf{U}} = \text{diag}(1, 1, 1, e^{i((\phi_{11}-\phi_{10})-(\phi_{01}-\phi_{00}))}). \quad (4.64)$$

We therefore will use

$$\Phi = \delta\phi_1 - \delta\phi_0 = (\phi_{11} - \phi_{10}) - (\phi_{01} - \phi_{00}), \quad 0 \leq \Phi \leq \pi, \quad (4.65)$$

to characterise the entangling power of each gate by equating them to a control gate of the form $\text{diag}(1, 1, 1, \Phi)$ (and label them as $C(\Phi)$ for convenience). Therefore $\Phi = 0$ means that the gate has no entangling power and $\Phi = \pi$ is a maximally entangling gate. From this and the geometric Bloch sphere picture, we have some hints of what will be necessary conditions for entangling gate generations.

A geometric picture of the entanglement condition

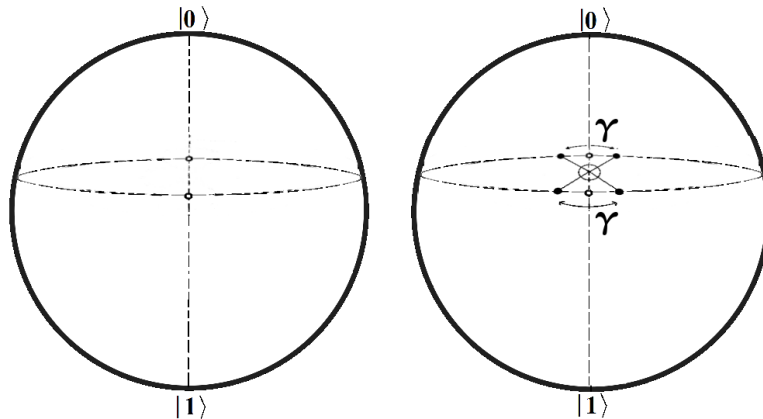


Figure 4.10: If the intermediate states $|a_j\rangle$ start on the same elevation the four back-action states $|a_{jk}\rangle$ will lie on the same horizontal plane. All horizontal planes through a sphere create a spherical cap with a midpoint at the poles thus, as one can see in comparison with figure 4.9, $|m\rangle = |0\rangle$. However, the angle between states marked by different k in the pairs $|a_{0k}\rangle$ and $|a_{1k}\rangle$ would be the same for each j . To be entangling $\delta\phi_1 - \delta\phi_0$ must be non-zero but on the horizontal plane both $\delta\phi$ are equal.

After the first interaction, but before the second, the ancilla will be in one of two states $|a_j\rangle$ corresponding to $|\Psi\rangle_1 = |j\rangle_1$, $j = 0, 1$. The second interaction induces

a unitary on the second register qubit that is given by $\langle m|\Delta|a_j\rangle$. For entanglement between the first and second register qubit, this unitary must be distinguishable by j . Therefore, each $|a_j\rangle$ must have a distinct value of $\langle\sigma_z\rangle_j$.

In the geometric picture, two points given by a_j that are on the same horizontal plane, before the second interaction, will produce four a_{jk} on the same horizontal plane afterwards (see figure 4.10). This means that the measurement basis will have to be the computational basis. By the definition of the control gates, a computational basis measurement corresponds to a local single qubit unitary on the register. Similarly an ancilla state prepared in the computational basis corresponds to the same constant unitary enacted with each interaction and extracts no information from the first interaction to transmit in the intermediate stage.

The geometry of the unitary condition

The measurement of the ancilla is fixed by the four back-action states $|a_{jk}\rangle$ which we know from the circuit pattern in figure 4.8 will be of the form

$$|a_{jk}\rangle = R_z((-1)^k 2\alpha) U_a R_z((-1)^j 2\alpha) |a\rangle. \quad (4.66)$$

We can find restrictions on the preparation state and intermediate unitary by working backwards from the measurement.

Each state can be represented by the Bloch sphere vector $\vec{a}_{jk} = (\theta_{jk}, \phi_{jk})$ and we can find restrictions on co-ordinates from the fact that they must lie in the same plane to form a cap on the Bloch sphere with a point equidistant from all of them. Each point can be expressed in Cartesian coordinates by the relationships:

$$|\vec{a}_{jk} \cdot \vec{x}| = \sin(\theta_{jk}) \cos(\phi_{jk}), \quad (4.67)$$

$$|\vec{a}_{jk} \cdot \vec{y}| = \sin(\theta_{jk}) \sin(\phi_{jk}), \quad (4.68)$$

$$|\vec{a}_{jk} \cdot \vec{z}| = \cos(\theta_{jk}). \quad (4.69)$$

Three points alone can always be found to be on the same plane. We will define a plane from three points and then find the expression for the distance from the fourth point to that plane. Thus we will find the conditions for the fourth point to be in the same plane as the other three. The equation for a plane defined by three points is

$$a.x + b.y + c.z + d = 0, \quad (4.70)$$

$$a = \frac{-d}{D} \begin{vmatrix} 1 & y_1 & z_1 \\ 1 & y_2 & z_2 \\ 1 & y_3 & z_3 \end{vmatrix}, b = \frac{-d}{D} \begin{vmatrix} x_1 & 1 & z_1 \\ x_2 & 1 & z_2 \\ x_3 & 1 & z_3 \end{vmatrix}, c = \frac{-d}{D} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}, D = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}. \quad (4.71)$$

There is a freedom of choice of d so it can be set $d = D$. The distance of point 4 to the plane is given by

$$\text{distance} = \frac{|a.x_4 + b.y_4 + c.z_4 + d|}{\sqrt{a^2 + b^2 + c^2}}. \quad (4.72)$$

As we are only interested in the case where the distance is zero, we can ignore the normalisation factor and simply examine

$$\text{distance}' = |a.x_4 + b.y_4 + c.z_4 + d|. \quad (4.73)$$

We now make use of some restrictions on the formation of these four points. From (4.66), states of the same label j are related by

$$|a_{j0}\rangle = R_{\hat{z}}(4\alpha)|a_{j1}\rangle.$$

This means that the second interaction gate has no effect on their relative elevations and fixes the difference in azimuthal angle and we can set

$$\theta_{00} = \theta_{01}, \theta_{10} = \theta_{11} \quad (4.74)$$

and

$$\phi_{01} - \phi_{00} = \phi_{11} - \phi_{10} = 4\alpha. \quad (4.75)$$

Using (4.74), the distance between the fourth point and the plane can be simplified to

$$2(\cos(\theta_{01}) - \cos(\theta_{11})) \left[\cos\left(\phi_{01} - \frac{\phi_{10} + \phi_{11}}{2}\right) - \cos\left(\phi_{00} - \frac{\phi_{10} + \phi_{11}}{2}\right) \right] \sin(\theta_{01})\sin(\theta_{11})\sin\left(\frac{\phi_{10} - \phi_{11}}{2}\right). \quad (4.76)$$

This generates several possible conditions for the fourth point to lie in the plane, some more trivial than others. If $\cos(\theta_{01}) = \cos(\theta_{11})$ then all four points must lie on the same horizontal plane which means that there is no entangling power since then $\delta\phi_1 = \delta\phi_0 = 4\alpha$. $\sin(\theta_{01}) = 0$ and $\sin(\theta_{11}) = 0$ would mean that there are only three distinct points with one being at the pole i.e. the $|0\rangle$ state. Due to the construction of these four points it is not possible for $\phi_{10} - \phi_{11} = 0$ to be true. The final condition is that

$$\cos\left(\phi_{01} - \frac{\phi_{10} + \phi_{11}}{2}\right) = \cos\left(\phi_{00} - \frac{\phi_{10} + \phi_{11}}{2}\right). \quad (4.77)$$

If (4.75) is then substituted in, this becomes

$$\cos\left(\phi_{00} + 4\alpha - \frac{2\phi_{10} + 4\alpha}{2}\right) = \cos\left(\phi_{00} - \frac{2\phi_{10} + 4\alpha}{2}\right), \quad (4.78)$$

$$\cos(\phi_{00} - \phi_{10} + 2\alpha) = \cos(\phi_{00} - \phi_{10} - 2\alpha). \quad (4.79)$$

Since α is non-zero, this requires $|\phi_{00} - \phi_{10}| \bmod \pi = 0$ i.e. the two points are in the same vertical plane.

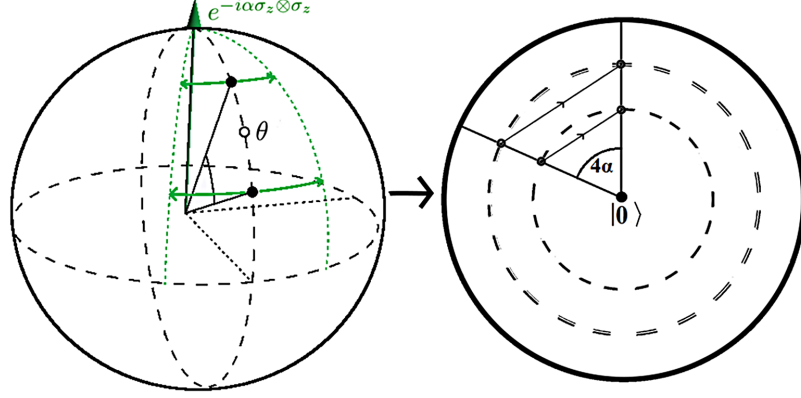


Figure 4.11: A 2d projection of the construction of the four points of equation (4.66), $|a_{jk}\rangle$, on the Bloch sphere. Constructing the intermediate back-action states in the same vertical plane (left) restricts the points to only one of two polar and one of two azimuthal angles. The vectors that connect two points of the same polar angle will thus be parallel. This guarantees that all four points lie on the same plane.

Finding the measurement basis

If the points of the same label j must be in the same vertical plane then it must be that the intermediate back-action states that describe the ancilla just after the intermediate unitary

$$|a_j\rangle = U_a R_{\hat{z}}((-1)^j 2\alpha) |a\rangle \quad (4.80)$$

must be in the same vertical plane. Under the interaction gate $e^{-i\alpha\sigma_z \otimes \sigma_z}$, there is a gauge symmetry of rotations about the \hat{z} axis so we can without any loss of generality consider that they are aligned on the $\hat{x} - \hat{z}$ plane. Labelling the average elevation of the two points as θ and the angle between them as 4β for some as yet unknown β (as opposed to α), they can be written as

$$|a_j\rangle = \cos\left(\frac{\theta - (-1)^j 2\beta}{2}\right) |0\rangle + \sin\left(\frac{\theta - (-1)^j 2\beta}{2}\right) |1\rangle. \quad (4.81)$$

The four points after the second interaction gate can then be rewritten in terms of (θ, β) :

$$|a_{jk}\rangle = R_{\hat{z}}((-1)^k 2\alpha) |a_j\rangle = e^{-i(-1)^k \alpha} \cos\left(\frac{\theta - (-1)^j 2\beta}{2}\right) |0\rangle + e^{i(-1)^k \alpha} \sin\left(\frac{\theta - (-1)^j 2\beta}{2}\right) |1\rangle. \quad (4.82)$$

Now say that the ancilla is measured in the state

$$|m\rangle = e^{-i\frac{\xi}{2}} \cos\left(\frac{\zeta}{2}\right) |0\rangle + e^{i\frac{\xi}{2}} \sin\left(\frac{\zeta}{2}\right) |1\rangle \quad (4.83)$$

this means that the unitary constraint can be expressed as

$$\begin{aligned} & \cos^2\left(\frac{\zeta}{2}\right)\cos^2\left(\frac{\theta - (-1)^j 2\beta}{2}\right) + \sin^2\left(\frac{\zeta}{2}\right)\sin^2\left(\frac{\theta - (-1)^j 2\beta}{2}\right) \\ & + 2\cos\left(\frac{\zeta}{2}\right)\cos\left(\frac{\theta - (-1)^j 2\beta}{2}\right)\sin\left(\frac{\zeta}{2}\right)\sin\left(\frac{\theta - (-1)^j 2\beta}{2}\right)\cos\left(\xi - (-1)^k 2\alpha\right) = p_m. \end{aligned} \quad (4.84)$$

The four points are symmetric about the $\hat{\mathbf{x}} - \hat{\mathbf{z}}$ plane so it is expected that the measurement basis should lie in that plane. This is quickly confirmed by considering the only term in (4.84) that is dependent on the second index k :

$$\cos\left(\xi - (-1)^k 2\alpha\right)$$

which is independent of k e.g. $|\langle m|a_00\rangle|^2 = |\langle m|a_01\rangle|^2$ when

$$\cos(2\alpha - \xi) = \cos(2\alpha + \xi). \quad (4.85)$$

To find the parameter ζ we use the equation (4.84) for two vertically separated points. We can simplify the equation using the trigonometric relationships

$$\begin{aligned} \cos^2\left(\frac{\zeta}{2}\right) &= \frac{1}{2}[1 + \cos(\zeta)], \\ \sin^2\left(\frac{\zeta}{2}\right) &= \frac{1}{2}[1 - \cos(\zeta)]. \end{aligned}$$

From this one can show that

$$\begin{aligned} & \cos^2\left(\frac{\zeta}{2}\right)\cos^2\left(\frac{\theta - (-1)^j 2\beta}{2}\right) + \sin^2\left(\frac{\zeta}{2}\right)\sin^2\left(\frac{\theta - (-1)^j 2\beta}{2}\right) \\ &= \frac{1}{2}[1 + \cos(\zeta)\cos(\theta - (-1)^j 2\beta)]. \end{aligned}$$

From (4.85) one can set $\xi = 0$ or π :

$$\begin{aligned} & \rightarrow \cos(\zeta)\cos(\theta + 2\beta) + \sin(\zeta)\sin(\theta + 2\beta)\cos(2\alpha) \\ &= \cos(\zeta)\cos(\theta - 2\beta) + \sin(\zeta)\sin(\theta - 2\beta)\cos(2\alpha) \end{aligned} \quad (4.86)$$

$$\begin{aligned} & \rightarrow \sin(\zeta)\cos(2\alpha)[\sin(\theta + 2\beta) - \sin(\theta - 2\beta)] \\ &= \cos(\zeta)[\cos(\theta - 2\beta) - \cos(\theta + 2\beta)] \end{aligned}$$

$$\sin(\zeta)\cos(2\alpha)\cos(\theta)\sin(\beta) = \cos(\zeta)\sin(\theta)\sin(\beta),$$

$$\beta = 0 \text{ or } \sin(\zeta)\cos(2\alpha) = \cos(\zeta)\sin(\theta),$$

$$\zeta = \theta = \frac{\pi}{2} \text{ or } \tan(\zeta) = \frac{\tan(\theta)}{\cos(2\alpha)}. \quad (4.87)$$

So either the vertical split is centred on the equator and thus the ancilla is measured at that same point on the equator or there is an adjustment dependent on α but interestingly not on β . So however β depends on the initial ancilla state and the intermediate unitary, it can be set independently of the measurement basis.

The intermediate unitary

From the above relationship between the measurement basis and the parameters (θ, β, α) , what can we deduce about the necessary initial ancilla state and the intermediate unitary? Suppose the measurement basis was to be fixed and thus θ is to be fixed too, the only degree of freedom of the ancilla preparation and intermediate unitary is represented by the parameter β . What is β ?

If there is an initial ancilla state

$$|a\rangle = \cos\left(\frac{\theta'}{2}\right)|0\rangle + \sin\left(\frac{\theta'}{2}\right)|1\rangle, \quad (4.88)$$

the first interaction gate creates the two back-action states

$$|a'_j\rangle = e^{-i(-1)^j\alpha}\cos\left(\frac{\theta'}{2}\right)|0\rangle + e^{i(-1)^j\alpha}\sin\left(\frac{\theta'}{2}\right)|1\rangle. \quad (4.89)$$

These will lie on some horizontal plane which the intermediate unitary must then rotate into a vertical plane. The most straightforward action that will do so is to rotate the ancilla about the axis of intersection between the $\hat{\mathbf{x}} - \hat{\mathbf{z}}$ and the plane of the Bloch vectors of the two back-action states, $\hat{\mathbf{a}}_1$. Because of the symmetry of the two vectors about $\hat{\mathbf{x}} - \hat{\mathbf{z}}$ the perpendicular norm of the plane of the two vectors will in fact lie in $\hat{\mathbf{x}} - \hat{\mathbf{z}}$ so the angle of rotation will be $\pm\frac{\pi}{2}$ independent of θ' .

The two points are connected to the origin of the Bloch sphere by the vectors

$$\hat{\mathbf{a}}_j = \sin(\theta')\cos(2\alpha)\hat{\mathbf{x}} + (-1)^j\sin(\theta')\sin(2\alpha)\hat{\mathbf{y}} + \cos(\theta')\hat{\mathbf{z}}. \quad (4.90)$$

The normalised vector of the plane will be $\hat{\mathbf{b}} = (\hat{\mathbf{a}}_0 \times \hat{\mathbf{a}}_1)/\sin(\beta')$ where β' is the great circle angle between the two points on the Bloch sphere. Unnormalised

$$\begin{aligned} \hat{\mathbf{b}} \propto (\hat{\mathbf{a}}_0 \times \hat{\mathbf{a}}_1) &= 2\sin(\theta')\sin(2\alpha)\cos(\theta')\hat{\mathbf{x}} - 2\sin(\theta')\sin(2\alpha)\sin(\theta')\cos(2\alpha)\hat{\mathbf{z}} \\ &\propto \cos(\theta')\hat{\mathbf{x}} - \sin(\theta')\cos(2\alpha)\hat{\mathbf{z}}. \end{aligned} \quad (4.91)$$

The norm of $(\hat{\mathbf{a}}_0 \times \hat{\mathbf{a}}_1)$ provides us with

$$\sin^2(\beta') = 4\sin^2(2\alpha)\sin^2(\theta')(1 - \sin^2(2\alpha)\sin^2(\theta')), \quad (4.92)$$

$$\begin{aligned} \sin^2(\beta') &= 4\sin^2\left(\frac{\beta'}{2}\right)\sqrt{1 - \sin^2\left(\frac{\beta'}{2}\right)} \\ &\rightarrow \sin^2\left(\frac{\beta'}{2}\right) = \sin^2(2\alpha)\sin^2(\theta'). \end{aligned} \quad (4.93)$$

Recall the equation for the probability for measuring the ancilla in its preparation basis as this point from (4.36). We see here that $\hat{\mathbf{b}}$ has no $\hat{\mathbf{y}}$ component and thus lies in the $\hat{\mathbf{x}} - \hat{\mathbf{z}}$ plane. So the axis of intersection will be given by the vector also within that plane, perpendicular to $\hat{\mathbf{b}}$:

$$\hat{\mathbf{n}} = \hat{\mathbf{b}} \times \hat{\mathbf{y}} \propto \sin(\theta') \cos(2\alpha) \hat{\mathbf{x}} + \cos(\theta') \hat{\mathbf{z}}. \quad (4.94)$$

When we related this to the form

$$\hat{\mathbf{n}} = \sin(\theta'') \hat{\mathbf{x}} + \cos(\theta'') \hat{\mathbf{z}}, \quad (4.95)$$

we can find the elevation of the axis of intersection from

$$\begin{aligned} \tan(\theta'') &= \frac{\sin(\theta'')}{\cos(\theta'')} = \frac{\sin(\theta') \cos(2\alpha)}{\cos(\theta')} \\ &= \cos(2\alpha) \tan(\theta'). \end{aligned} \quad (4.96)$$

Once the two points have been rotated about $\hat{\mathbf{n}}$ into the $\hat{\mathbf{x}} - \hat{\mathbf{z}}$ plane, they are separated vertically by the great circle angle β' , therefore in (4.81) $\beta = \frac{\beta'}{4}$ and

$$\sin^2(2\beta) = \sin^2(2\alpha) \sin^2(\theta'). \quad (4.97)$$

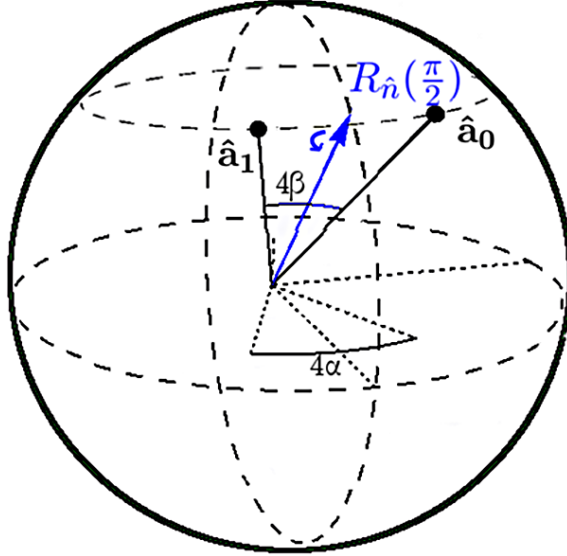


Figure 4.12: Representation of the intermediate unitary gate on the ancilla in the Bloch sphere picture. The rotation axis passes through the great angle between the intermediate ancilla states, not through the horizontal plane of elevation.

If no further operation takes place then we can use (4.81) to equate $\theta = \theta''$. Using the ζ, θ relationship (4.87) we find

$$\tan(\zeta) = \frac{\tan(\theta'')}{\cos(2\alpha)} = \tan(\theta'). \quad (4.98)$$

The measurement would be in the basis of the initial ancilla state. However, the unitary and entangling condition only required that the two points lie in the same vertical plane. Therefore the intermediate unitary allows for an additional degree of freedom. After aligning the two points vertically, they can be rotated about the \hat{y} axis to set any θ we choose i.e.

$$U_a = R_{\hat{y}}(\theta - \theta'')R_{\hat{n}}(\pm\frac{\pi}{2}). \quad (4.99)$$

The degree of freedom of the initial state preparation is effectively in the parameter β according to (4.93) while the intermediate unitary selects the measurement basis with θ . There is an additional degree of freedom in the intermediate unitary that rotates the intermediate ancilla states $|a_j\rangle$ into the $\hat{x} - \hat{z}$ plane-the direction of rotation. If $R_{\hat{n}}(-\frac{\pi}{2})$ were to be chosen over $R_{\hat{n}}(\frac{\pi}{2})$ in (4.99) then the intermediate states would have their orientation in the vertical plane flipped. The association of $|a_j\rangle$ in (4.81) to the computational basis state $|j\rangle$ would be flipped so that the final Kraus operator would exchange elements $\langle m|a_{0k}\rangle \leftrightarrow \langle m|a_{1k}\rangle$ and thus too the terms $\delta\phi_j$ in (4.65). Therefore $\Phi \rightarrow -\Phi$.

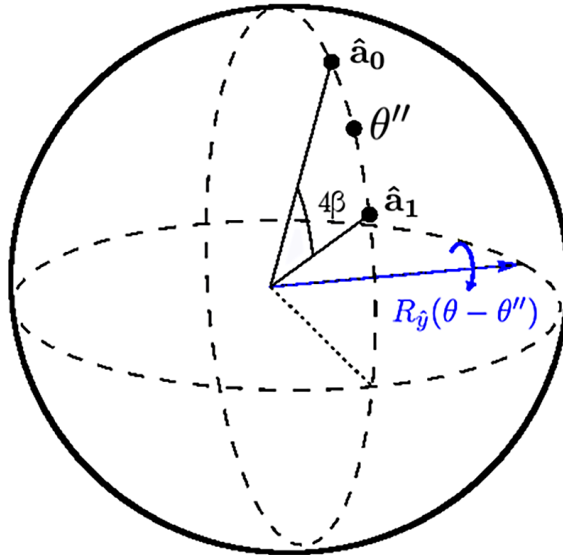


Figure 4.13: Representation of the choice of measurement basis in the intermediate unitary gate on the ancilla in the Bloch sphere picture.

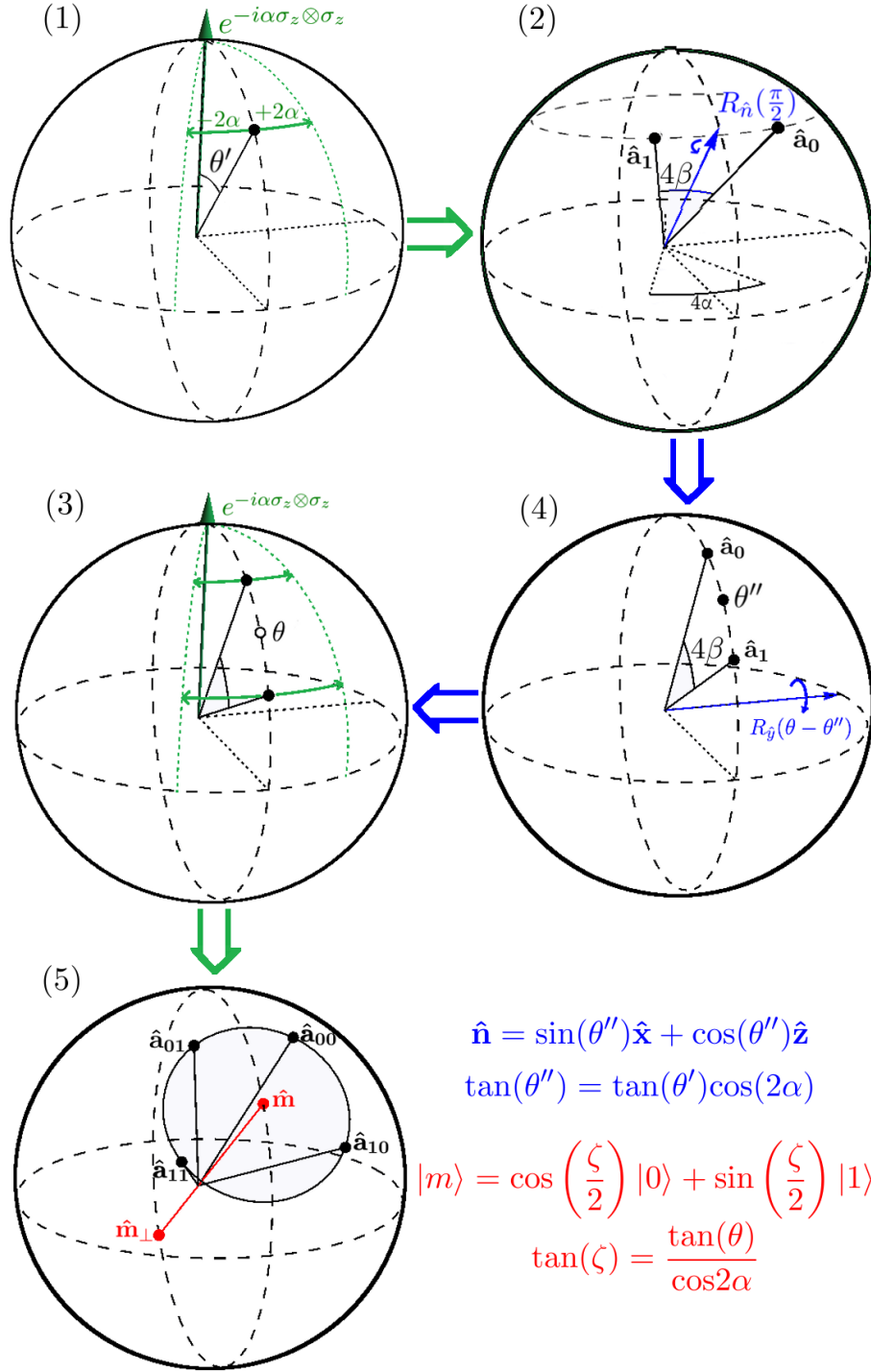


Figure 4.14: A geometric picture of the operations on the ancilla during the two-qubit entangling gate generation: (1) The ancilla couples to the first register qubit with interaction gate $e^{-i\alpha\sigma_z \otimes \sigma_z}$. (2) The ancilla qubit undergoes single qubit gate $R_{\hat{n}}(\frac{\pi}{2})$. (3) The ancilla measurement basis is chosen by a single-qubit gate rotation $R_{\hat{y}}(\theta - \theta'')$. (4) The ancilla undergoes coupling with the second register qubit with interaction gate $e^{-i\alpha\sigma_z \otimes \sigma_z}$. (5) The entangling gate generation is now fully armed and operational upon a measurement of the ancilla in the basis $\{|m\rangle, |m_{\perp}\rangle\}$ for $|m\rangle = \cos\left(\frac{\zeta}{2}\right)|0\rangle + \sin\left(\frac{\zeta}{2}\right)|1\rangle$.

4.2.2 The entangling power of a generated two-qubit gate

Given α , the entangling power measure, Φ , can be calculated for any (θ, β) using the four points from (4.82), the relationship between the measurement basis and θ from (4.87) and combining (4.65) and (4.57) to express Φ and the probability of obtaining Φ as

$$\Phi = \text{Arg}\left[\frac{\langle m|a_{11}\rangle\langle m|a_{00}\rangle}{\langle m|a_{10}\rangle\langle m|a_{01}\rangle}\right], \quad (4.100)$$

$$p_\Phi = |\langle m|a_{00}\rangle|^2. \quad (4.101)$$

An example of the relative entangling powers of the Kraus operators

As an example, take the intermediate state to have $\theta = \frac{\pi}{2}$ so that the $|a_i\rangle$ are vertically split about the $|+\rangle$ state. It is helpful to think of the intermediate state as $R_{\hat{x}}(\frac{\pi}{2})R_{\hat{z}}(\pm 2\beta)|+\rangle$ while the ancilla was prepared in $|+\rangle$. The effect of the preparation choice and the reduced solid angle are treated like an effective reduction in the interaction strength of the first interaction while the intermediate $U_a = R_{\hat{x}}(\frac{\pi}{2})$. The four states of $|a_{jk}\rangle = e^{-i(-1)^k\alpha\sigma_z}R_{\hat{x}}(\frac{\pi}{2})e^{-i(-1)^j\beta\sigma_z}|+\rangle$ will be all symmetrically placed around $|+\rangle$ so we can also say $|m\rangle = |+\rangle$ and measure in the $\{|+\rangle, |-\rangle\}$ basis.

$$\langle m|a_{jk}\rangle = \langle +|e^{-i(-1)^k\alpha\sigma_z}R_{\hat{x}}(\frac{\pi}{2})e^{-i(-1)^j\beta\sigma_z}|+\rangle. \quad (4.102)$$

$$\mathbf{R}_{\hat{z}}(2\alpha)\mathbf{R}_{\hat{x}}\left(\frac{\pi}{2}\right)\mathbf{R}_{\hat{z}}(2\beta) = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-iA} & -ie^{iB} \\ -ie^{-iB} & e^{iA} \end{pmatrix}, \quad (4.103)$$

$$\rightarrow \langle +|a_{00}\rangle = \frac{1}{2\sqrt{2}}(e^{-iA} - ie^{iB} - ie^{-iB} + e^{iA}) = \frac{1}{\sqrt{2}}(\cos(A) - i\cos(B)), \quad (4.104)$$

$$\langle -|a_{00}\rangle = \frac{1}{2\sqrt{2}}(e^{-iA} - ie^{iB} + ie^{-iB} - e^{iA}) = \frac{1}{\sqrt{2}}(\sin(B) - i\sin(A)), \quad (4.105)$$

where we define $A = \alpha + \beta$, $B = \beta - \alpha$.

Each other element of the two-qubit evolution operator will just be a transformation of A and B (which thus fulfils the unitary condition) and because of the \pm symmetries of sine and cosine we will be able to express the final Kraus operator and Φ in terms of just ϕ_{00} .

$$\Phi_+ = \delta\phi_1 - \delta\phi_0 = 4\phi_{00}^+ + \pi, \quad (4.106)$$

$$\Phi_- = 4\phi_{00}^- + \pi. \quad (4.107)$$

j	k	A→	B→	ϕ_{jk}^+	ϕ_{jk}^-
0	0	A	B	ϕ_{00}^+	ϕ_{00}^-
0	1	B	A	$-\phi_{00}^+ - \frac{\pi}{2}$	$-\phi_{00}^- - \frac{\pi}{2}$
1	0	-B	-A	$-\phi_{00}^+ - \frac{\pi}{2}$	$-\phi_{00}^- + \frac{\pi}{2}$
1	1	-A	-B	ϕ_{00}^+	$\phi_{00}^- + \pi$

Table 4.3: Table of transformations of A and B for different computational states of the register

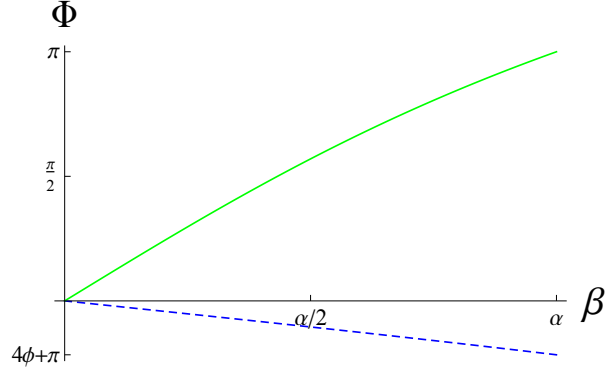


Figure 4.15: For $\alpha = \frac{\pi}{8}$, the two relative entangling phases of the measurement in the $|\pm\rangle$ basis are plotted with the $|+\rangle$ result given by the blue (dashed) curve and the less likely $|-\rangle$ result by the (solid) green curve. By manipulating the effective coupling strength, β , with the ancilla state preparation, the difference in Φ for the two possible operator outputs can be adjusted.

If $\beta = \alpha$ the four points are a square around the initial preparation and measurement basis. In this case $B = 0$ and so

$$\langle +|a_{00}\rangle = \frac{1}{\sqrt{2}}(\cos(A) - i), \quad \Phi^+ = 4\arctan\left(\frac{-1}{\cos(2\alpha)}\right) + \pi, \quad (4.108)$$

$$\langle -|a_{00}\rangle = \frac{1}{\sqrt{2}} - i\sin(A), \quad \Phi^- \bmod 2\pi = \pi. \quad (4.109)$$

Φ^- is in fact independent of α and provides a gate locally equivalent to the maximally entangling CZ gate.

Given some likely values of α , the form of Φ_0 in (4.108) can be compared to the other potential result π . Let $\arctan\left(\frac{-1}{\cos(2\alpha)}\right) = q\pi$, if

$$\cos(q\pi) = \frac{1}{\sqrt{n}}, \quad n \geq 3 \quad (4.110)$$

then q is an irrational number [154]. If the values of α were to be rational fractions of π which often have rational or root of rational cosines (e.g. $\cos\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}$), then we know from (4.108) that

$$\tan^2(q\pi) = m, \quad (4.111)$$

$$\frac{1 - \cos^2(q\pi)}{\cos^2(q\pi)} = m, \quad (4.112)$$

$$\cos^2(q\pi) = \frac{1}{m+1}, \quad (4.113)$$

which will fulfill (4.110) if $m \geq 2$.

Plotting the entangling power outputs

As we see in the previous example, the output of each measurement result matches the greater entangling power with the lower probability result. This mirrors the correlation of greater rotation angle with lower probability in the single-qubit gate generation. Using (4.100), the relationship of the entangling power output with both parameters can be plotted. In figures 4.16 and 4.17, when compared with figure 4.20, we see that the power/probability relationship is maintained as both parameters are varied. The entangling power of both outputs increases with both β and θ so the previously considered example of $\theta = \frac{\pi}{2}$ was an upper limit. While β was considered before to be limited by α as according to (4.93), if we look at extending β to $\frac{\pi}{4}$, we can cover cases where the first interaction gate is not equal to the second. The case where $\beta = \frac{\pi}{4}$ actually restores many of the advantages of ADQC with a maximally entangling interaction gate- the output entangling gate power is deterministic, with only possibly local gate differences between the two outputs and equal probabilities and the gate power matches the gate power of the second interaction gate.

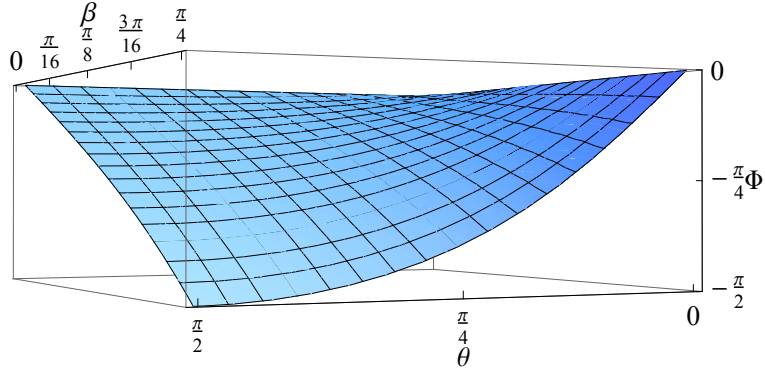


Figure 4.16: For $\alpha = \frac{\pi}{8}$, the entangling power of greater probability, Φ_0 , plotted against β and θ over the range $0 \leq \theta \leq \frac{\pi}{2}$, $0 \leq \beta \leq \frac{\pi}{4}$. $|\Phi_0|$ is limited by the entangling power of the original gate, 4α , in this case $\frac{\pi}{2}$.

Note the substantial difference in curvature between Φ_0 for fixed θ and for fixed β . In a contour plot in figure 4.19 that demonstrates β against θ for fixed Φ , the two curves have different gradient signs with β decreasing as θ increases for Φ_0 while it increases for Φ_1 . Pairs of outputs are unique. Since figure 4.20 shows us that the probability increases with both parameters, the maximum probability of a specific value of Φ_1 occurs with one parameter at its maximum so the earlier case where β is varied for $\theta = \frac{\pi}{2}$ is probability optimised.

Sticking to the case where $\theta = \frac{\pi}{2}$, as β increases past α , Φ_1 becomes negative before returning to the limit of $|\Phi| = 4\alpha$ (see figure 4.18). Together, Φ_1 and Φ_0 will cover the range $[-\pi, \pi]$ enabling any effective value of Φ to be generated.

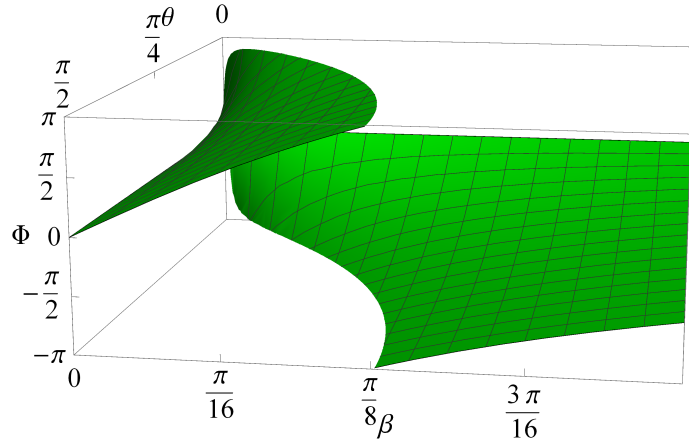


Figure 4.17: For $\alpha = \frac{\pi}{8}$, the minority entangling power Φ_1 plotted against β and θ over the range $0 \leq \theta \leq \frac{\pi}{2}$, $0 \leq \beta \leq \frac{\pi}{4}$. The 2π periodicity means values of $\pi + \epsilon$ equate to $-\pi + \epsilon$ so the plot here includes a discontinuity. Reducing θ makes the distributions of the four ancilla state points more eccentric, reducing the β necessary for $\Phi = \pi$, yet an eigenstate measurement causes no entanglement so the value drops off rapidly as the parameters approach 0.

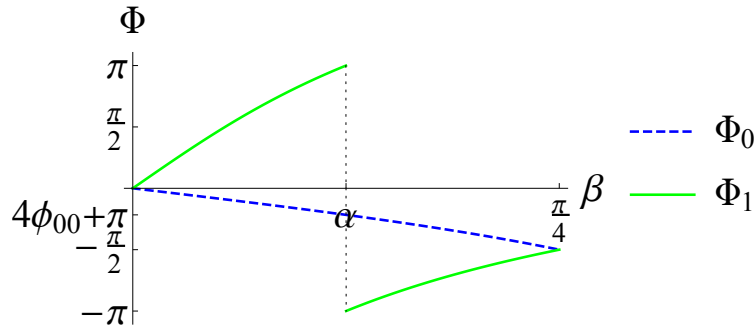


Figure 4.18: For $\alpha = \frac{\pi}{8}$, the two relative entangling phases of the measurement in the $|\pm\rangle$ basis are plotted with the $|+\rangle$ result given by the blue (dashed) curve and the less likely $|-\rangle$ result by the (solid) green curve over the extended range $0 \leq \beta \leq \frac{\pi}{4}$. As β becomes greater than α , Φ_1 loops round to becoming the extreme negative end of its possible range before increasing to equal $\Phi_0 = -4\alpha$ at $\beta = \frac{\pi}{4}$.

The negativity of Φ simply indicates that the gate is equivalent to $C(-|\Phi\rangle) = C(|\Phi\rangle)^\dagger$ since the sign of such can not be fixed by post-correction alone. So say two generations produces the results $\Phi_0 < 0$ and $\Phi_1 > 0$, the resulting product must be locally equivalent to $C(\Phi_1 - |\Phi_0|)$. However the sign of the output for a given (β, θ) is not truly fixed. Since the sign of Φ can be flipped by a change of sign in the intermediate unitary in (4.99), values of Φ in the range $[4\alpha, \pi]$ can be generated by values of β greater or less than α but since the probability of success of the minor

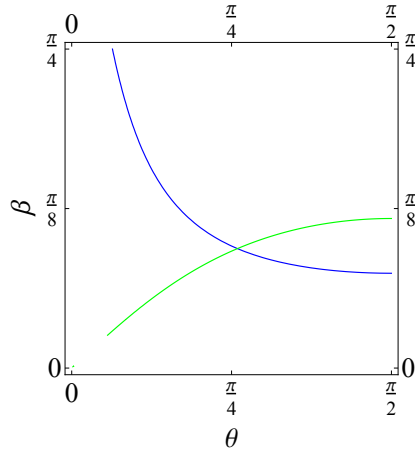


Figure 4.19: For $\alpha = \frac{\pi}{8}$, two curves of constant Φ are plotted where the blue negative gradient curve represents $\Phi_0 = \frac{\pi}{8}$ and the green positive gradient curve represents $\Phi_1 = \pi$.

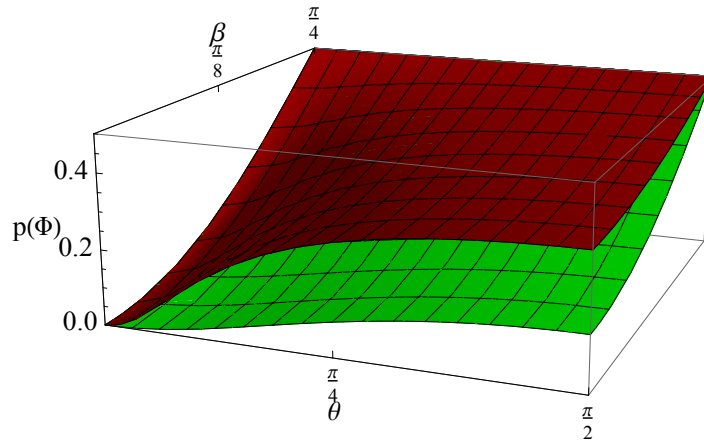


Figure 4.20: The probability of the minor output Φ_1 plotted against (β, θ) for two values of alpha- the higher probability, red, plot represents $\alpha = \frac{\pi}{5}$, the green, lower probability, plot is $\alpha = \frac{\pi}{8}$. The probability of the minor output can be represented visually by the size of the minor cap formed by the four ancilla states on the Bloch sphere. Increasing, any one of β , θ or α increases the size of this cap. Ultimately, the limit is a hemispherical cap representing $p = \frac{1}{2}$.

output continues to increase with β , the probability of success is optimised by taking values greater than α .

4.2.3 The local gate decomposition of the Kraus operation

Φ is calculated independently of local unitary gate effects. On the one hand, this means one does not know the necessary local gates that allows one to correct the generated gate into the locally equivalent $C(\Phi)$ gate. On the other hand, any calculation of the

local gates will be highly susceptible to the convention used to equate gates up to global phases. For example, an intermediate state $|a_i\rangle$ may be expressed as

$$\begin{aligned} |a_0\rangle &= e^{-i\beta} \cos\left(\frac{\theta'}{2}\right) |0\rangle + e^{i\beta} \sin\left(\frac{\theta'}{2}\right) |1\rangle = e^{-i\beta} \left(\cos\left(\frac{\theta'}{2}\right) |0\rangle + e^{i2\beta} \sin\left(\frac{\theta'}{2}\right) |1\rangle \right), \\ |a_1\rangle &= e^{i\beta} \cos\left(\frac{\theta'}{2}\right) |0\rangle + e^{-i\beta} \sin\left(\frac{\theta'}{2}\right) |1\rangle = e^{i\beta} \left(\cos\left(\frac{\theta'}{2}\right) |0\rangle + e^{-i2\beta} \sin\left(\frac{\theta'}{2}\right) |1\rangle \right). \end{aligned}$$

If the intermediate gate then corresponds to the transformation

$$\cos\left(\frac{\theta'}{2}\right) |0\rangle + e^{\pm i2\beta} \sin\left(\frac{\theta'}{2}\right) |1\rangle \rightarrow \cos\left(\frac{\theta \mp \beta}{2}\right) |0\rangle + \sin\left(\frac{\theta \mp \beta}{2}\right) |1\rangle,$$

the Kraus operator elements will accrue an additional $e^{\pm i\beta}$ phase compared to our earlier description and the total Kraus operator is multiplied by a local phase operation of 2β on the first qubit:

$$\begin{pmatrix} e^{i\beta} & & & \\ & e^{i\beta} & & \\ & & e^{-i\beta} & \\ & & & e^{-i\beta} \end{pmatrix} = \begin{pmatrix} e^{i\beta} & \\ & e^{-i\beta} \end{pmatrix} \otimes \mathbb{I}.$$

One must be careful in considering how relative phases are represented in a subsystem as they are not truly “global” phases. These choices of convention must be consistent with the form of the fundamental interaction and the experimental implementation of the unitary gates – for example, the distinction between implementing a phase difference between two paths by inducing a phase shift in one or both paths. The measure Φ is beneficially independent of these issues for gates locally equivalent to $e^{-i\alpha\sigma_z \otimes \sigma_z}$.

Nevertheless, given a consistent convention or an exact experimental model, one can calculate the local gate effects from a few simple relationships.

$$\begin{aligned} \begin{pmatrix} e^{i\phi_{00}} & & & \\ & e^{i\phi_{01}} & & \\ & & e^{i\phi_{10}} & \\ & & & e^{i\phi_{11}} \end{pmatrix} &= \begin{pmatrix} e^{i\frac{\Phi}{4}} & & & \\ & e^{-i\frac{\Phi}{4}} & & \\ & & e^{-i\frac{\Phi}{4}} & \\ & & & e^{i\frac{\Phi}{4}} \end{pmatrix} \cdot \begin{pmatrix} e^{ia} & \\ & e^{ib} \end{pmatrix} \otimes \begin{pmatrix} e^{ic} & \\ & e^{id} \end{pmatrix}, \\ \begin{pmatrix} e^{ia} & \\ & e^{ib} \end{pmatrix} \otimes \begin{pmatrix} e^{ic} & \\ & e^{id} \end{pmatrix} &= \begin{pmatrix} e^{i(a+c)} & & & \\ & e^{i(a+d)} & & \\ & & e^{i(b+c)} & \\ & & & e^{i(b+d)} \end{pmatrix}, \end{aligned} \quad (4.114)$$

$$\delta\phi_1 - \delta\phi_0 = \Phi,$$

$$\delta\phi_1 + \delta\phi_0 = 2(d - c), \quad (4.115)$$

$$\phi_{11} + \phi_{10} - \phi_{01} - \phi_{00} = 2(b - a). \quad (4.116)$$

Since these corrections must be diagonal in the same basis as the interaction gate, their effects will commute through every generation if there are no other local gate effects or if they are compensated by a fixed correction. They could then be corrected by a single pair of local gate operations after multiple generations. During these generations the cumulative value of Φ will randomly walk through the range $[-\pi, \pi]$. This raises the question of how a specific target gate, at least just one entangling gate for universality, can be achieved and how long this random walk process will take to achieve it.

4.3 An example of a random walk on a finite subgroup of two-qubit entangling gates

The coupling strength of the interaction gate determines the determinism of ancilla driven quantum computation up to Pauli corrections. We are interested in exploring how weaker interaction gates can still be used in ADQC which means having to adapt the model for non-deterministic gate generations and most likely incurring additional costs in terms of other resources. Any strategy to generate a particular result will take several steps; any adaptive strategy will have to feed forward results and post-correction gates may have variable resource costs. The trade off of resources becomes a significant issue.

The output of each generation is characterised by a single continuous parameter Φ governed by several other parameters but the issues outlined above can first be approached by a simpler toy model. We will start by first focusing on the probabilistic generation of gates in a finite group and using this to devise some strategies by which one can aim to achieve a specific target gate.

Consider a device meant to generate a *CNOT* gate over two qubits but only achieves this with a probability p . It otherwise implements a *CZ* gate. Since $CZ = (\mathbb{I} \otimes H).CNOT.(\mathbb{I} \otimes H)$, the device can also be viewed as a device that implements a *CNOT* gate with a $q = 1 - p$ probability of a local unitary error on the target. This error does not have the convenient property that it can be fixed by a local unitary post-correction. The ultimate aim is to achieve a *CNOT* gate up to local unitary post-corrections.

We will describe some context that we assume in order to refine our aims. We will assume an inability to discard our current qubits and the actions implemented on them in order to reflect the intention to use static and highly isolated physical systems for memory qubits in ADQC. If we are enacting gates as part of a computational algorithm, then our control and target qubits will be in transitory states of a possibly very long process that we do not wish to discard and which have a high cost to replace.

Single-qubit gates will be treated as deterministic or that, if they are also proba-

bilistically generated, the expected time to implemented is known and the time needed to implement a single-qubit gate with near unity probability is fixed and of the order of the time to implement the $CZ/CNOT$ generation.

The current state of the gate will be known and stopping the generation of two-qubit gates is under our control. This “active controller” assumption is based on the control of the ancilla input and the measurement readouts obtained in ADQC.

4.3.1 Case 0

First look at the basic case, case 0, where the two-qubit gate generation is implemented repetitively without interference until the final product of all generations is a $CNOT$ gate. For two control unitary gates CU_0 and CU_1 ,

$$\begin{aligned} CU_i &= |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes U_i, \\ CU_0.CU_1 &= |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes U_0.U_1 = C(U_0.U_1). \end{aligned}$$

In this case, $U_0 = X$, $U_1 = Z$. The group properties of the single-qubit unitary gates will thus determine that of the control-unitaries. Since X and Z are generators of the Pauli group— $G_1 = i^n.\{\mathbb{I}, X, Y, Z\}$, $n = 0, 1, 2, 3$ —whose element’s global phases i^n are caused by the different orders of multiplication and the commutation relations of the Pauli operators. For the single-qubit gates a global phase could be ignored but since our attention is on the Control-Pauli group, the control unitary gates which enact identity or Pauli operators on the target with $CNOT$ equalling CX , one must note

$$|0\rangle\langle 0| \otimes \mathbb{I} + e^{i\alpha}|1\rangle\langle 1| \otimes U = (Z^{\frac{\alpha}{\pi}} \otimes \mathbb{I}).CU \quad (4.117)$$

so the two-qubit gate Control-Pauli group is

$$\mathcal{G}_c = (Z^{\frac{\pi}{4}} \otimes \mathbb{I}).\{\mathbb{I}, CNOT, CY, CZ\}. \quad (4.118)$$

Fortunately the $Z^{\frac{\pi}{4}}$ local gate on the control commutes with the Control-Unitary gates and so the local gate effects from different multiplication orders can be treated as a single local gate post-correction. For $C\sigma_j \in \mathcal{G}_c$

$$C\sigma_j C\sigma_{j'} = C\sigma_{j'} C\sigma_j.(V \otimes \mathbb{I}). \quad (4.119)$$

Now a tree diagram can be easily constructed (see figure 4.21) but since the group has a finite number of elements, it can also be represented by a graph in figure 4.22.

With a passive controller who can only halt the generator, a $CNOT$ gate can not be guaranteed to be generated but since there are a finite number of points on the graph, the expectation time and its variance can be calculated through recursion. The

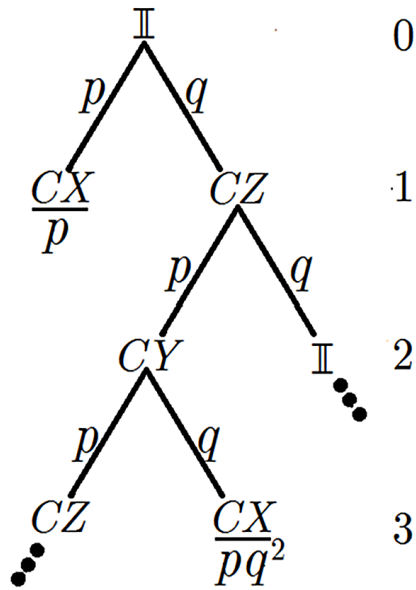


Figure 4.21: The probability tree of the random gate generator of case 0. Each level number marks the number of gate generations implemented. A simple tree with only two branches per level but some branches return to points further up the tree and a success may be found on the left or right branch depending on the level.

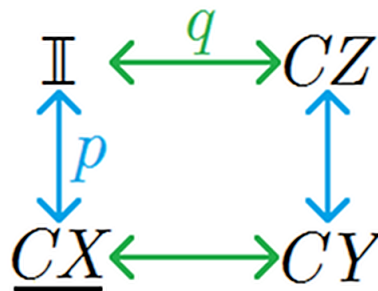


Figure 4.22: A graph of the random generator gates. The generator moves the gate from one point on the graph to another along lines that represent the different results. The green horizontal lines represent the CZ generation, the blue vertical ones the CX generation.

expectation time from each level can be expressed as the probabilistically weighted sum of the times of the lower branches plus one for the extra level.

$$\bar{n}_j = \sum_k p_k (\bar{n}_{j+1,k} + 1) \quad (4.120)$$

$$\rightarrow \bar{n}_0 = p + q \cdot (\bar{n}_1 + 1). \quad (4.121)$$

CX is the success condition so its own expectation time is 0 so the only unknown at the 1st level is that of CZ . In the next level, one of the results returns the gate to a previous point and so this branch must have the same expectation time as the starting point. The recursion relations extend to each unique level thus:

$$\bar{n}_0 = p + q \cdot (\bar{n}_1 + 1), \quad (4.122)$$

$$\bar{n}_1 = p(\bar{n}_2 + 1) + q \cdot (\bar{n}_0 + 1) = p\bar{n}_2 + q\bar{n}_0 + 1, \quad (4.123)$$

$$\bar{n}_2 = p(\bar{n}_1 + 1) + q. \quad (4.124)$$

\bar{n}_0 and \bar{n}_2 is are expressed in terms of \bar{n}_1 so let them be substituted into the term for \bar{n}_1 .

$$\begin{aligned} \bar{n}_1 &= p^2(\bar{n}_1 + 1) + pq + pq + q^2(\bar{n}_1 + 1) + 1 \\ &= (p^2 + q^2) \cdot (\bar{n}_1 + 1) + 2pq + 1, \\ p^2 + q^2 + 2pq &= (p + q)^2 = 1 \\ \rightarrow \bar{n}_1 &= (1 - 2pq) \cdot (\bar{n}_1 + 1) + 2pq + 1 \\ &= (1 - 2pq)\bar{n}_1 + 2, \\ \frac{2}{\bar{n}_1} + (1 - 2pq) &= 1, \\ \frac{2}{\bar{n}_1} &= 2pq, \\ \bar{n}_1 &= \frac{1}{pq}. \end{aligned} \quad (4.125)$$

Let this then be substituted back into \bar{n}_0 .

$$\bar{n}_0 = p + q \left(\frac{1}{pq} + 1 \right) = 1 + \frac{1}{p}. \quad (4.126)$$

Notice that there is no probability of success in level 2. One might think therefore that there may be a way to improve the recursion relations by simplifying out the need to write the terms for this level. We can simplify the above result and more later by applying results from the geometric distribution to the recursion relations.

The geometric distribution

The geometric distribution deals with a basic scenario very similar to the major problem of this section but simplified in one major aspect: let each result be wiped clean after each attempt so that each trial is attempted with only success and failure as outcomes with the same probability of success in each step. For the resulting distribution of the time until success, the probability of first success occurring on the n th step is $p.q^{n-1}$ for $0 < p \leq 1$, $q = 1 - p$. The expected number of steps is

$$\bar{n} = \sum_n n.p.q^{n-1} \quad (4.127)$$

which can be shown to be equal to $\frac{1}{p}$ [155]. The distribution also has a simple cumulative distribution function which simply subtracts the probability of n independent failures from unity,

$$C(n) = 1 - q^n, \quad (4.128)$$

so a perfect probability of success can be approximated by $1 - \epsilon$ by guaranteeing enough resources to perform n attempts where n is logarithmic with respect to ϵ .

The cumulative distribution function of the geometric distribution brings it nicely into line with the issues of resources for generating quantum gates. If one has a random gate generator that produces an ideal gate, U or identity, \mathbb{I} , then this random generator can efficiently approximate an ideal gate. This idea is indeed in place in schemes for quantum computation with linear optical systems where the trick is to force any probabilistic procedures that may destroy the input quantum state into an offline preparation so that the process can be recycled [99].

The key feature of the geometric distribution, the “success/failure” condition, seems absent from our scenario. There are three possible unsuccessful results which each result in different probabilities of success in the next step. However the structure of the graph in figure 4.22 is very limited. If two elements are opposite each other, to get from one to the other requires one horizontal and one vertical path, regardless of whether the path is clockwise or anti-clockwise. Also, the intermediate steps can only return success or the original starting point. So if we now consider a two step time interval, a gate that starts as CZ has either moved to CX or returned to its starting point.

We can see this if we take the probability tree in figure 4.21 and examine the sub-tree between levels 1 and 3 as in figure 4.23.

The sub-tree can be considered to consist of a “success/fail” trail but performed over two steps. Instead of a probability of success p and expectation time $\frac{1}{p}$, p is substituted by $p' = 2pq$ with an expectation time of $\frac{2}{p'} = \frac{1}{pq}$. This provides an alternative derivation of the result in (4.125).

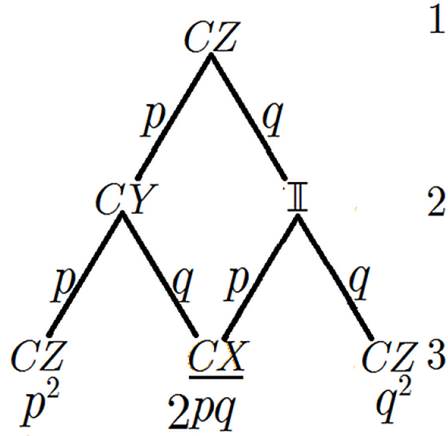


Figure 4.23: The sub-tree of levels 1 to 3: Due to the limited graph, the four branches of two steps collect into only two results, the initial gate and the success condition. Overall is it in effect a Bernoulli trial with probability of success of $2pq$.

One minor adjustment can be made to improve this behaviour, provided awareness of the value of p . If the target was switched from CX to CZ the probability tree would simple form a reflection about the first level split and the statistics of CZ hitting time match the earlier results with the exchange $p \leftrightarrow 1 - p$. The expected hitting time for CZ is $1 + \frac{1}{1-p}$ which is an improvement on $1 + \frac{1}{p}$ for $p < \frac{1}{2}$. For $p < \frac{1}{2}$, one could make use of the knowledge of the probability by preparing the system with a local Hadamard gate on the target and implementing another Hadamard gate as a post-correction to exchange the $CZ \leftrightarrow CX$ results. Now the expectation time can be defined as

$$\bar{n}_0 = \begin{cases} 1 + \frac{1}{1-p} & : 0 < p \leq \frac{1}{2} \\ 1 + \frac{1}{p} & : \frac{1}{2} \leq p < 1 \end{cases} \quad (4.129)$$

However an additional consideration that may have to be made with such a technique is the additional time taken by the implementation of the local Hadamard pre- and post-corrections. The time for the local gate operations may be relatively far smaller yet still significant to the time needed for one random gate generation, resulting in a small adjustment to the matching condition:

$$\frac{1}{p} - \frac{1}{1-p} - \bar{n}_H = 0 \quad (4.130)$$

where \bar{n}_H is the time for the local operations which may be an average itself. This is a quadratic equation with real solutions for all real \bar{n}_H but for small \bar{n}_H where $p \approx \frac{1}{2} + \epsilon$,

a rough estimation can be made

$$\begin{aligned} \frac{1-2p}{p(1-p)} - \bar{n}_H &= 0, \\ 1-2p &= \bar{n}_H \cdot p(1-p), \\ &= \bar{n}_H \cdot \left(\frac{1}{2} - \epsilon\right) \left(\frac{1}{2} + \epsilon\right), \\ &\approx \frac{\bar{n}_H}{4}, \\ p &\approx \frac{1}{2} - \frac{\bar{n}_H}{8}. \end{aligned}$$

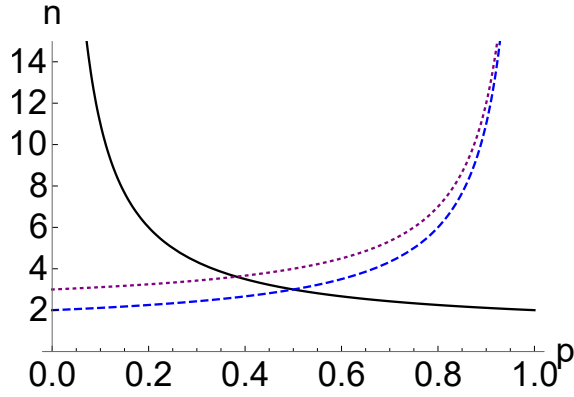


Figure 4.24: A plot of the expectation time for the implementation of a CX gate against p . The solid black curve represents the expectation time if the controller does nothing, the dashed blue line is the expectation time of the CZ gate in the same case and thus also the expectation time for when the gate is prepared with a local Hadamard gate. The dotted purple line represents an adjustment to take into account the time to implement the local gate corrections; the intersection of the black curve shifts to the left and results in a larger maximum time.

4.3.2 Case 1

Now let us introduce the ability to implement local gate post-corrections in between each random gate generation. This allows one to construct a protocol to improve CX achievement times.

In the unaltered case, the statistics of the expected time to achieve a CX gate were dominated by the tail of the distribution of hitting times after the first step which behaved as a geometric distribution with probability of success per step of $2pq$ but with the average time extended by a factor of two by the doubling of the number of steps to achieve a “success/failure” trial instead of a random gate generation. Therefore a natural focus for any protocol would be the reduction of the number steps between

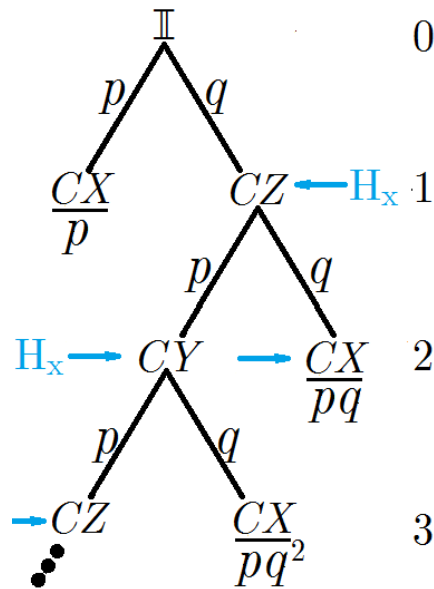


Figure 4.25: The probability tree of the random gate generator of case 1 with applications of the local gate H_x after each arrival at a CZ and after the subsequent generation. After the first step, every successful generation occurs on a q branch split from the level above.

each possibility of hitting CX by assuring that there is one branch that will result in CX directly at every level.

What local gate post-corrections are to be available to us? In fact, there are a limited number of gates that are relevant to the scenario: the only gate we are interested in are those that convert Control-Pauli gates into each other and they are the two-qubit gates of the Clifford group so the only other gates we need to generate one from another are a pair that will generate the single-qubit Clifford group gates. We will use two gates from the Hadamard set as defined in (4.9) to create two new cases where a possibility of success in the second step is created- one case where the probability exists on the q probability branch and one where it exists on the p probability branch.

In case 1, before the second step, the gate H_y is applied to the target qubit, the random gate generation occurs and then a post-corrective H_y gate is applied. Because H_x affects the Pauli operators thus

$$H_x Z H_x = Y,$$

$$H_y X H_y = -X,$$

it will preserve the CX gate generation upto local gate commutable post-corrections but convert the CZ generation to CY . If this is done every time the cumulative gate product is at CZ the q branch will produce CX (see figure 4.25).

Now after the first step the tail of the distribution will behave like a geometric distribution with a probability of success $1 - p$ for a trial every one step. $\bar{n}_1 = \frac{1}{1-p}$ is substituted into (4.122) to give

$$\bar{n}_0 = p + q \cdot \left(\frac{1}{q} + 1\right) = p + q + 1 = 2. \quad (4.131)$$

\bar{n}_0 is now constant for all p . The probability of success in the first step now balances against the probability of success in the later steps so if the value of p is very high then conditioned on a failure in the first step the expected time to hit CX afterwards is very large while if p is very low, fairly on the first step is most likely but the number of steps is then very unlikely to get past 2.

Since 2 is the lower bound on $1 + \frac{1}{p}$ in the range $0 < p < 1$, case 1 provides an improvement on case 0 for all values of p . Except we have not yet taken into account the additional time needed to implement the local unitary gates. Once again, there is sometimes the additional implementation time of both pre- and post-correction, \bar{n}_H , but furthermore this only occurs at certain points in the procedure for case 1. The local gate pre- and post-corrections are used to set up successes on the 2nd, 4th and further even numbered steps. This is after the first level q branch from which point on it behaves as a geometric distribution so the contributions will come from every other element of a geometric series:

$$\bar{n}_0 = 2 + T_H \quad (4.132)$$

where

$$T_H = \bar{n}_H \cdot [q^2 + 2p^2q^2 + 3p^4q^2 + \dots] = \bar{n}_H \cdot q^2 \sum_{n=0}^{\infty} (n+1)p^{2n}. \quad (4.133)$$

$$T_H = \bar{n}_H \cdot q^2 \frac{1}{(1-p^2)^2} \quad (4.134)$$

$$= \bar{n}_H \frac{1}{(1+p)^2}. \quad (4.135)$$

Since this value is now p dependent, one can consider applying the trick of switching the targets $CX \leftrightarrow CZ$. Because the switching is done by another gate in the Hadamard set, the crossover point will in fact be independent of \bar{n}_H

$$\begin{aligned} 2 + \bar{n}_H \frac{1}{(1+p)^2} &= 2 + \bar{n}_H \frac{1}{(2-p)^2} + \bar{n}_H, \\ (1+p)^2 &= (1+p)^2(1+(2-p)^2), \\ (2-p)^2 &= 2+2p \\ \rightarrow 1 + 10p - 3p^2 - 2p^3 + p^4 &= 0. \end{aligned}$$

This has no real positive solution so the original gate target always wins out (see figure 4.26).

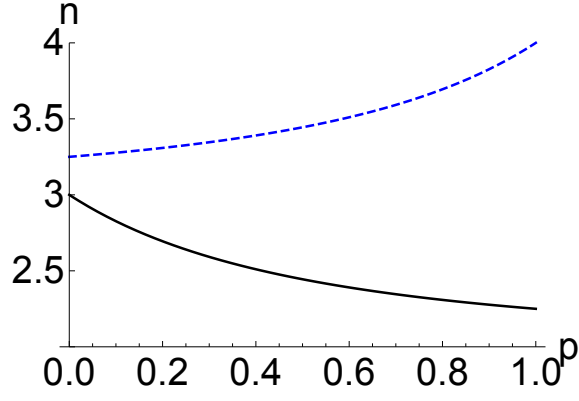


Figure 4.26: The expectation times of Case 1 plotted against p , taking into account the time to implement local gate operations, \bar{n}_H , with $\bar{n}_H = 1$. The dashed blue curve represents the expectation time when trying to enact CX by enacting CZ with local gate corrections. The meeting point is independent of \bar{n}_H and thus the latter case gives no advantage.

4.3.3 Case 2

The alternative to Case 1 is to correct the p branch of the second step so that its product with CZ produces CX . Therefore the CX gate must be converted to a CY by pre- and post-applications of the H_z gate. This also preserves the CZ gate generated by the q branch so in the probability tree in figure 4.27, we recognise “success/failure” condition of the geometric distribution.

The resulting fundamental $\frac{1}{p}$ expectation time is superior to the $1 + \frac{1}{p}$ value of case 0 for all p and one can again apply the same trick of switching the target gate with a preparation gate as in case 0 but again, there is the issue of the time added due to the local gate operations. Like with case 1, the additional cost \bar{n}_H is incurred on alternating steps, after the 1st, 3rd and other $2n + 1$ failures:

$$\begin{aligned}
 \bar{n}_0 &= \sum_n n \cdot p \cdot q^{n-1} + \bar{n}_H \cdot pq \sum_{n=0}^{\infty} (n+1) q^{2n} \\
 &= \frac{1}{p} + \bar{n}_H \cdot \frac{pq}{(1-q^2)^2} \\
 &= \frac{1}{p} + \bar{n}_H \cdot \frac{1-p}{p(2-p)^2}.
 \end{aligned} \tag{4.136}$$

For case 2 to improve upon case 0, it must be true that

$$\begin{aligned}
 \bar{n}_H \cdot \frac{1-p}{p(2-p)^2} &< 1, \\
 \bar{n}_H &< \frac{p(2-p)^2}{1-p}.
 \end{aligned} \tag{4.137}$$

Given a value of \bar{n}_H , one could calculate a range of values of p for which this is valid. On the other hand, one can make the value of p at which case 2 switches from

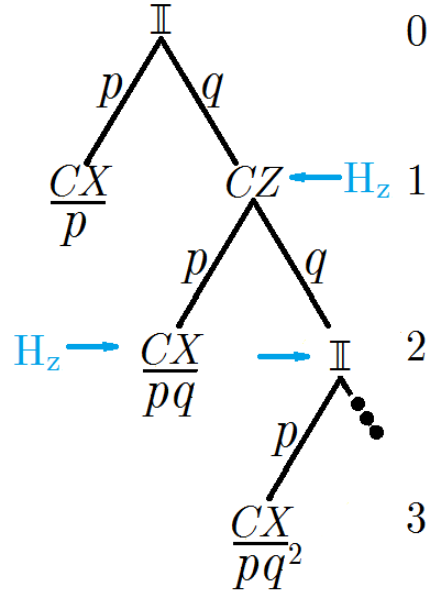


Figure 4.27: The probability tree of the random gate generator of case 2 with applications of the local gate H_z after each arrival at a CZ and after the subsequent generation. Each level number marks the number of gate generations implemented. All successful generations occurs on a p branch split from the above level.

targeting CX to CZ a lower bound, to provide an upper bound for \bar{n}_H below which it will always be valid.

Cases 1 and 2 are compared in figure 4.28a (including the exchange of target gates for case 2). Without including the time for the local gate corrections, the $\frac{1}{p}$ and $\frac{1}{1-p}$ curves would meet at $p = \frac{1}{2}$ and never be greater than the constant 2 of case 1. Including the correction times, using (4.132) and (4.136), case 1 meets case 2 at $p = \frac{1}{2}$ independent of \bar{n}_H and each case has its own region in which it provides the shortest expectation time. Given an active controller aware of the probability of generator, the correct strategy for the given value of p can be chosen and the ultimate expectation time be the minimum of all available cases (as in figure 4.28b).

The appropriate time to adopt a strategy: ignoring the first step

However if there is the ability to select a particular strategy over the others on the basis of foreknowledge of the probability, note that the first step of each case is always the same. So actually the issue is then which has the smallest expectation time from level 1, \bar{n}_1 at which point each case is a repeat-until-success trial with geometric probability distribution. Case 0 has

$$\bar{n}_1 = \frac{1}{pq}, \tag{4.138}$$

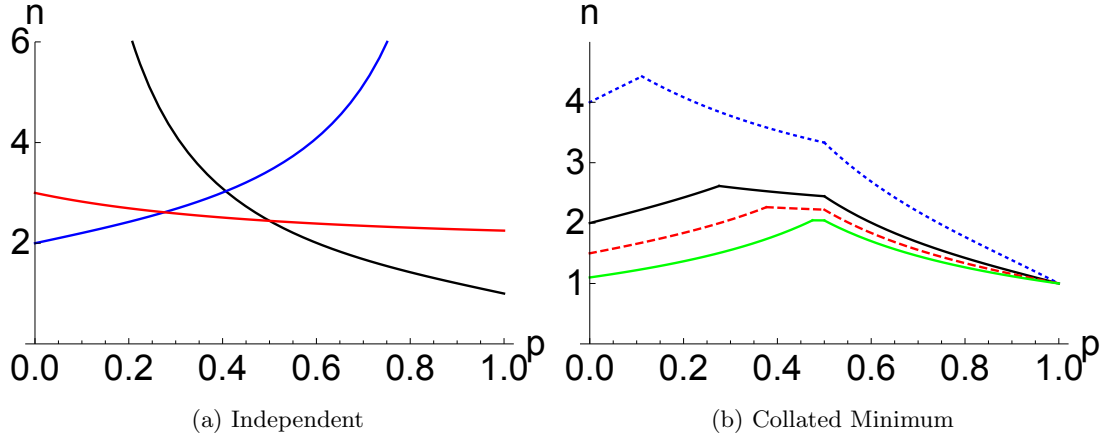


Figure 4.28: (a) The plot of cases 1 (red) and cases 2 (black) and 2-switched (blue) plotting expected number of steps against probability of generating CX . The time for local gate corrections is set at $\bar{n}_H = 1$ for all curves. (b) The minimum value of cases 1 and 2 across $0 < p < 1$ for several values of \bar{n}_H : $\bar{n}_H = 1$ —the solid black curve; $\bar{n}_H = 3$ —the dotted blue curve; $\bar{n}_H = 0.5$ —the dashed red curve; $\bar{n}_H = 0.1$ —the solid green curve

Case 1 has

$$\begin{aligned} \bar{n}_1 &= \frac{1}{q} + \bar{n}_H \cdot q \sum_{n=0}^{\infty} (n+1)p^{2n} \\ &= \frac{1}{q} + \bar{n}_H \cdot \frac{1}{(1+p)^2(1-p)} \end{aligned} \quad (4.139)$$

and Case 2 is symmetric to Case 1 under $p \leftrightarrow 1-p$ and will always meet it at $p = \frac{1}{2}$. A change of strategies will occur at the point when (4.138) and (4.139) equate; from those one derives the inequality

$$\bar{n}_H \leq \frac{(1+p)^2(1-p)}{p}. \quad (4.140)$$

As we see in figure 4.29a, the minimum point of the curve of case 0 is at $p = \frac{1}{2}$ so evaluating the inequality at the point gives $\bar{n}_H = 2.25$ as the maximum value for which a corrective strategy always wins. As \bar{n}_H increases, (4.140) and its reflection in $p = \frac{1}{2}$ can be used to mark out the region in which case 0 has a lower expectation time.

4.3.4 The guaranteed hitting time

The expectation time is only one of several statistical properties that may be optimised and may not produce the best choice to discriminate the cases under some circumstances. Consider instead that there needs to be a fixed limit on the number of steps taken to implement the $CNOT/CX$ gate. The random gate generator may be part

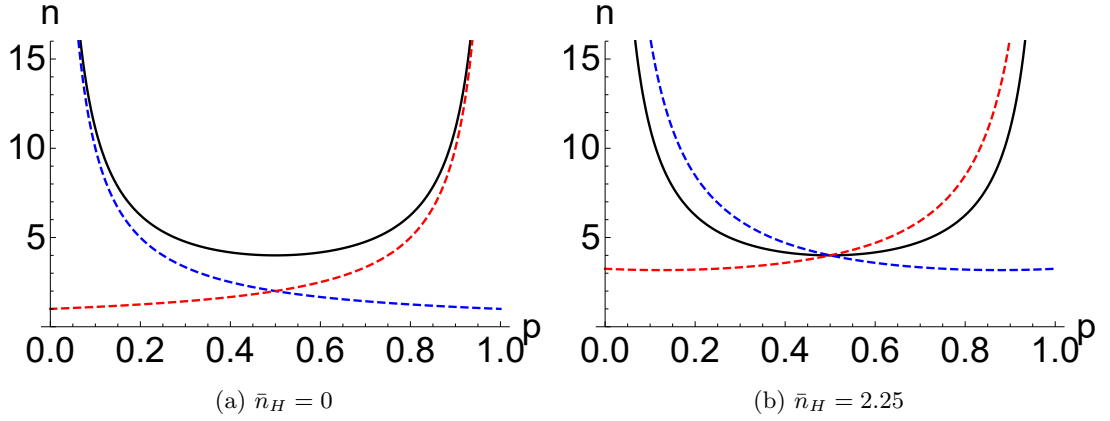


Figure 4.29: A plot of the expected time for each strategy to succeed given a failure in the first step for (a) no correction time (b) correction time $\bar{n}_H = 2.25$, Case 0 is plotted in solid black, case 1 in dashed red (with positive gradient), case 2 in dashed blue (negative gradient). Since case 0 has a minimum at the meeting point of cases 1 and 2, it takes longer than both cases for all p until the correction time is increased to the value in (b).

of a circuit or architecture that relies on the timing of different components or the emphasis may be on the total number of component resources that need to be prepared before implementing the generator and that can not be generated easily during the process. We now need an awareness of how many steps are necessary in order to create a guarantee that the desired gate will be generated to a very high probability. Then the random gate generator can be treated as a device that produces a *CNOT* gate to a specified fidelity that reflects the occasional failure.

The probability of “success/failure” operation *not* succeeding in n steps is simply $(1 - p)^n$ and all that is needed is for this to be smaller than a small fidelity limit ϵ so that

$$n \geq \frac{\ln \epsilon}{\ln(1 - p)}. \quad (4.141)$$

If one sticks by the principle of the above section 4.3.3, each case will obey (4.141) exactly but with different substitutions for p . However, cases 1 and 2 will take longer and/or take additional resources due to the local gate corrections which happen every other step after the first so for each, there is a correction

$$n \rightarrow n + \bar{n}_H \cdot \left(\frac{n}{2}\right) = n \left(1 + \frac{\bar{n}_H}{2}\right). \quad (4.142)$$

Case 0 should be treated as having a success probability of $2pq$ and then have its time doubled. As one can see in figure 4.30, the behaviour is very similar to the expectation

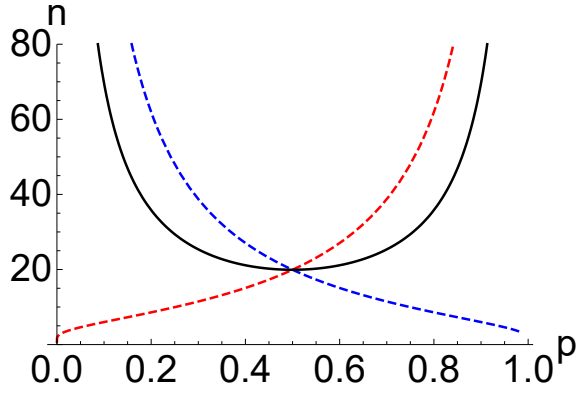


Figure 4.30: A plot of the time to guarantee success for each strategy to succeed given a failure in the first step with correction time $\bar{n}_H = 2$ and allowed probability of error $\epsilon = 0.001$. Case 0 is plotted in solid black, case 1 in dashed red (with positive gradient), case 2 in dashed blue (negative gradient). The curves meet at $p = \frac{1}{2}$ for $\bar{n}_H = 2$, independent of ϵ .

time. However, if we look at the condition at which all of the curves meet

$$2 \frac{\ln \epsilon}{\ln(1-2pq)} = \frac{\ln \epsilon}{\ln(1-p)} \left(1 + \frac{\bar{n}_H}{2}\right), \quad (4.143)$$

$$\left(1 + \frac{\bar{n}_H}{2}\right) = 2 \frac{\ln(1-p)}{\ln(1-2pq)}, \quad (4.144)$$

$$p = \frac{1}{2} \rightarrow 1 + \frac{\bar{n}_H}{2} = 2, \quad (4.145)$$

$$\bar{n}_H = 2. \quad (4.146)$$

There is a different constraint on the correction time, it is stricter than for the expectation time but fortunately independent of ϵ .

The actual effect of the first step can be included by scaling the error by the probability of failure on the first step

$$\epsilon \rightarrow \frac{\epsilon}{1-p} \quad (4.147)$$

and adding one to each value so for case 0

$$n \geq 2 \frac{\ln \frac{\epsilon}{q}}{\ln(1-2pq)} + 1, \quad (4.148)$$

for case 1

$$n \geq \frac{\ln \frac{\epsilon}{q}}{\ln p} \left(1 + \frac{\bar{n}_H}{2}\right) + 1 \quad (4.149)$$

and for case 2

$$n \geq \frac{\ln \frac{\epsilon}{q}}{\ln p} \left(1 + \frac{\bar{n}_H}{2}\right) + 1. \quad (4.150)$$

As seen in figure 4.31 this just means that the mean times are shorter for higher probabilities, much more so as ϵ is increased, but the crossing points of the curves remains the same.

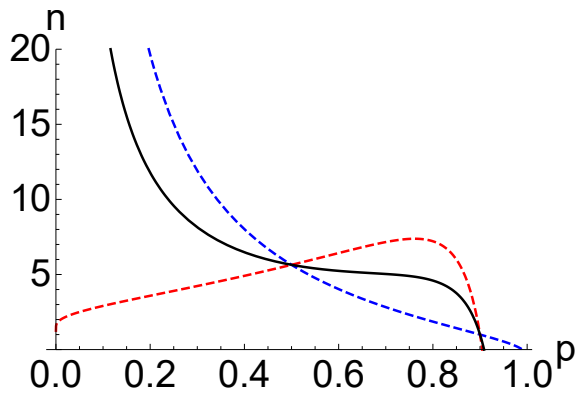


Figure 4.31: A plot of the time to guarantee success for each strategy to succeed with correction time $\bar{n}_H = 2$ and allowed probability of error $\epsilon = 0.1$. Taking into account the probability of success on the first step skews the curve since each will go to 1 at $p = 1 - \epsilon$ but the major features will be preserved.

4.4 Simulating deterministic circuits with probabilistic gate generation

In this section, we will show that interaction gates with arbitrary coupling parameters can be used to implement probabilistic ADQC to simulate a universal quantum gate circuit. To be specific, it is the simulation of any single-qubit gate and a fixed entangling gate.

Doing this requires an explanation of how the results from the generation of single-qubit gates can be universal for single-qubit gates and of the random behaviour that probabilistic ADQC may exhibit. Unlike the toy models of section 4.3, probabilistic ADQC generates a random walk on continuous groups of single-qubit and two-qubit gates and won't be able to recreate an exact gate. It will have to rely on efficiently creating an approximation within an error bound. This raises the question of how the number of generations relates to the the probability error bound and the approximation error.

It would also be beneficial to be able to guide the behaviour according to a strategy that utilises the resources available in the model. We will demonstrate that strategies inspired by the discussion in section 4.3 that use a feedback of measurement results are capable of parleying a greater coupling strength in the interaction gate into shorter expectation times to implement a gate.

4.4.1 Single-qubit unitary gates

Recomposing the interaction gate

While considering only the non-local part Δ_a of a decomposition of two-qubit interaction gates was helpful in finding the conditions for unitarity and the variation in generated gates, understanding the universality of the gates generated requires consideration of the local unitary gates that distinguish elements of the class $\tilde{\Delta}_a$. After all, it is plain to see that any gates generated by $\Delta_{(\alpha,0,0)}$ alone, since they all commute and correspond to rotations about the same axis on the Bloch sphere, are incapable of generating the entire group of single-qubit gates. In the original model of ADQC with maximally entangling gates, universality depended upon the generation of the class of gates $J(\beta) = H.e^{-i\frac{\beta}{2}\sigma_z}$ which occurred after applying the gate $(H_r \otimes H_a).CZ$. If one undoes the local register Hadamard gate, one loses the class $J(\beta)$.

For a gate $(V_s \otimes V_a).\Delta_{(\alpha_x, \alpha_y, \alpha_z)}(U_s \otimes U_a)$, the enacted gate is simply $V_s.U_m.U_s$. If the interaction gate has only one non-zero parameter then one can still use (4.19) and a pair of unitary gates related to the ancilla eigenstates to construct all other actions $U'_j = V_s.U_j.U_s$. If $V_s = U_s^\dagger$ then the local unitary gates will just apply a common transformation of the basis to all U_0 , U_1 and U_m and every generation will still commute with each other, insufficient for universal single-qubit gates.

How a one-parameter interaction gate of any coupling parameter is used to implement universal single-qubit gates on the register

Let us say there is no symmetry between the two unitary gates U_s, V_s then there can be $U'_0 = V_s.U_0.U_s$ and $U'_1 = V_s.U_1.U_s$ which may be two non-commuting operations whose Lie algebra closure covers the group of single-qubit unitary gates. These two gates themselves would be able to form a universal set for approximating single-qubit unitary gates [34]. It is therefore possible to implement deterministic arbitrary single-qubit gates by approximation by preparing ancilla states in the eigenstate basis and sending them in a string to implement a decomposition of a single-qubit gate as a string of $\{U'_0, U'_1\}$.

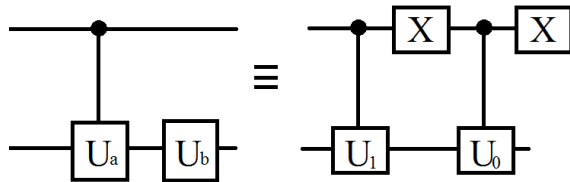


Figure 4.32: A circuit by which members of the interaction class $\Delta(\alpha < \frac{\pi}{4}, 0, 0)$ can generate two unitary gates whose commutators cover $SU(2)$. $U_0 = U_b$ and $U_1 = U_b U_a$

To demonstrate this proposal, consider the interaction gate $E = (H \otimes H).CZ^{\frac{1}{4}}$ (which corresponds to $\alpha = \frac{\pi}{16}$). This specific gate is chosen for two reasons; 1) It is directly comparable with the gate $E_{AR} = (H \otimes H).CZ$ from [38] but with a smaller rotation angle parameter of the control unitary, 2) It will generate $\{T, HT\}$ which generates the same group as the finite set $\{H, T\}$ which is well studied and proven to be universal [34] and of interest for its applications in fault tolerant quantum computation [34, 156]. The method can then be generalised to arbitrary coupling strengths and local unitary gate products. For nearly all values of α , $e^{\pm i\alpha\sigma_z}$ is outside of the Clifford group so one can use nearly all values of α in conjunction with non-Pauli operator local unitary gates in the Clifford group such as H and $\{\sigma_j^{\frac{1}{4}}\}$ to fulfil the condition for single-qubit unitary gates of the Clifford group and one outside of it.

With this method, the circuit is programmed by the ancilla state preparation and requires no further manipulation nor measurement of the ancilla. This may have benefits for particular physical systems depending on the lifetime and robustness of the ancilla.

How do we use the two-parameter class?

Unlike the one-parameter class of interactions, the set of gates generated from a two-parameter interaction may be non-trivially non-commuting even before one considers local unitary gates. The unitary gates implemented by this class, as seen in (4.51), have an axis of rotation dependent on the ancilla measurement result which generally gives a non-trivial commutation relationship between the two results. One can demonstrate by using the special case of complementary eigenstate preparation and measurement. The unitary operation can be equivalent to $R_{\hat{x}}(\pm 2\alpha_x)R_z(2\alpha_z)$. Say then that $\alpha_x = \alpha_z = \frac{\pi}{8}$, the two gates are $\sigma_x^{\frac{1}{4}}.\sigma_z^{\frac{1}{4}}$ and $\sigma_x^{-\frac{1}{4}}\sigma_z^{\frac{1}{4}}$. From Boykin *et al* [34], it is known that the latter will produce a unitary equivalent to a rotation by an angle that is an irrational multiple of π ;

$$U_{-+} = e^{-i\lambda\hat{n}.\hat{\sigma}}, \quad (4.151)$$

$$\cos(\lambda\pi) = \frac{1}{2} \left(1 + \frac{1}{\sqrt{2}} \right), \quad (4.152)$$

$$\hat{n} \propto \cot\left(\frac{\pi}{8}\right) (\hat{z} - \hat{x}) + \hat{y} \quad (4.153)$$

and then the difference in sign of the former gate only results in a change to a different rotation axis, non-parallel to the latter, by the same irrational multiple. Together these two gates could be used to create an approximation to any single-qubit unitary using a two axis decomposition. However the gates are being generated randomly rather than deterministically.

In fact, even with random gate generations, it may be possible to generate any

single-qubit gate up to some approximation with polynomial time with respect to the approximation error using concepts of stochastic recurrence. Given a universal single-qubit gate set such as $\{H, T\}$ or $\{R_{\hat{x}}(\pm 2\alpha_x)R_{\hat{z}}(2\alpha_z)\}$ and a fixed approximation error ϵ , $\exists N_m$ s.t. $\forall U_T \in U(2)$, a sequence can be found that approximates the unitary gate target U_T , $\|\prod_{j=1}^l \{H, T\}_{i(j)} - U_T\| < \epsilon$, for $l \leq N_m$.

The random gate generations perform a random walk over the 3-sphere space of single-qubit gates, since the gate generations are independent of each other, the process is memoryless and the present unitary action on the register at a given time forms a Markov chain. The error bound ϵ creates a target volume around the target unitary, \tilde{U}_T , so we can imagine that the 3 dimensional space is divided into volumes of size proportional to ϵ^3 of which there are a finite number that the product of the generated sequence jumps between. If there are 2^k sub-volumes then N_m is at least equal to k and naturally really much larger because the random walk can visit the same place more than once.

For an N_m number of generations, there is a finite probability that the sequence will have landed in the target volume \tilde{U}_T at some point. Now if all values of l for all U_T are treated as equal to N_m , ignoring the possibilities of multiple shorter sequences that can hit the target in fewer than N_m steps and assume for sufficiently large N_m an approximately uniform probability distribution over the entire space of single-qubit unitaries, then the probability of being within \tilde{U}_T is at a lower bound of 2^{-N_m} . For a probability of success close to one, $p_s = 1 - \delta$, one wants to generate $k \cdot N_m$ gates s.t. $(1 - 2^{-N_m})^k \leq \delta$.

This leaves open questions about the relationship between N_m and ϵ and the mixing time for a finite gate set to produce a uniform distribution over the space of single-qubit gates. It requires assumptions about the large size of N_m which suggests that 2^{-N_m} will be very small and thus it would create a large overhead but this is a very loose bound used to create the principles on which the procedure is guaranteed to work. A probability of hitting the target can exist well before the uniform mixing time and it may be improved if the procedure can be guided by the control of the parameters of the ancilla preparation and measurement which may possibly allow us to involve the Solovay-Kitaev algorithm [157, 85] to create a sequence of targets, one that may adapt to different measurement results, that would be limited to a polylogarithmic number of intermediate targets with respect to the error bound around the final target. One can also find a lower bound on the time to hit all targets by first considering the expected time to hit targets in a universal gate set and then multiplying by the number of times the set elements appear in a decomposition.

Issues about guiding random gate generations will be dealt with in a later section. In the upcoming section, we will use computational simulations of random gate generation

to verify some expected features. With the space of single-qubit unitary gates divided into finite volumes proportional to the cube of the error bound, it is expected that the probability of hitting the target in a fixed time would scale with that volume and thus the expected hitting time would scale with ϵ^3 . The success time according to $(1 - 2^{-N_m})^k \leq \delta$ should also form an exponential tail distribution.

4.4.2 Simulation of stochastic single-qubit gates

There is a range of parameter choices for the ancilla preparation, measurement basis and target unitary for simulating a stochastic method of generating a gate. It should be noted that while the two-parameter interaction gates are necessarily stochastic in their generations, it is also possible for a single-parameter interaction gate to be so. Such a case may arise under a scenario in which there is limited control over the state generation and measurement basis or where the initial ancilla state is not consistent and so the ancilla needs to be measured in the computational basis to produce the same U_0, U_1 actions (though this scenario introduces another variable in the probabilities of the results).

As a preliminary investigation, we simulated the random generation of gates using the interactions $(H \otimes H)\Delta(0, 0, \frac{\pi}{16})$ and $\Delta(\frac{\pi}{16}, 0, \frac{\pi}{16})$, ancilla preparation state $|+\rangle$ and measurement in the computational state basis. For the one-parameter class interaction the resulting unitary gates are $U_0 = HR_z(\frac{\pi}{8})$, $U_1 = HR_z(-\frac{\pi}{8})$, $p_0 = p_1 = \frac{1}{2}$; for the two-parameter class, $U_0 = R_z(\frac{\pi}{8})R_x(\frac{\pi}{8})$, $U_1 = R_z(-\frac{\pi}{8})R_x(\frac{\pi}{8})$, $p_0 = p_1 = \frac{1}{2}$. The target unitary was $U_T = R_x(\frac{\pi}{2})$. At each step a gate corresponding to the $\{U_0, U_1\}$ of each interaction was multiplied to the product of the previous step starting with the identity operator. The normalised trace distance of the product, $V = \prod_k U_{i(k)}$, and the target unitary was evaluated at each step until within a chosen error size $\epsilon < 0.01$;

$$\|V - U_T\| = \sqrt{\frac{2 - |\text{Tr}[V^\dagger U_T]|}{2}} \leq \epsilon. \quad (4.154)$$

The number of gates required for this to occur was collected 1000 times and used to create a probability distribution for the gate count required to reach the target unitary (see figure 4.33).

Some predicted features start to appear. While the target is not achievable in a single step and there is no finite probability for success per step, the aggregate behaviour over many steps, taken over a large number of simulations, can be modelled as a geometric or discretised exponential distribution. This is true for both one-parameter and two-parameter interactions. Particular target unitaries may cause anomalous effects; the two-parameter case is able to produce an exact solution of the target unitary in 4 steps which causes a large peak in the distribution and then suppresses the probability of a result for several steps after (see figure 4.34). With a large enough bin size this is

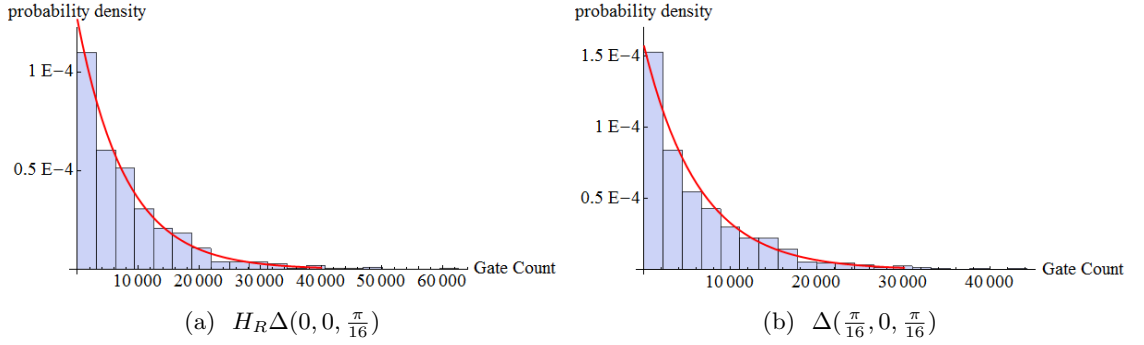


Figure 4.33: Probability distribution of required gate count to achieve target $U_T = R_{\hat{x}}(\frac{\pi}{2})$. (a) Use of a single-parameter interaction in a 20 bin histogram. (b) Use of a two-parameter interaction in a 20 bin histogram. The probability distribution corresponding to the exponential distribution parametrised by the mean of the results is displayed by the solid red curve with decay parameter (a) $\lambda = \frac{1}{7914}$ (b) $\lambda = \frac{1}{6372}$ (with the mean rounded to no decimal places).

not apparent but we see the effect in figure 4.36 in the discrepancy between the mean and the standard deviation of the distribution.

Error scaling

From the logarithmic plot of the average number of gates needed in the simulation against the trace distance error in 4.35, it can be seen that they share a polynomial relationship and so the random generation of gates is not inefficient with respect to the error. The power of the polynomial is affected by the choice of the trace distance.

To be clear, the textbook definition of the trace distance as per [112] is given by

$$D(\rho, \sigma) = \frac{1}{2} \text{Tr}[|\rho - \sigma|] \tag{4.155}$$

where

$$|A| = \sqrt{A^\dagger A} \tag{4.156}$$

but this is sensitive to global phases on ρ and σ so we use the form of (4.154), as used by Bocharov & Svore [35] for single-qubit unitary gates, which is a maximum over all global phases, thereby guaranteeing the the inequality in (4.154) holds for all phases, and is normalised to a maximum value of 1 for when $V = \mathbb{I}$. However the trace distance is normally described in the context of quantum density operators where it corresponds to half of the Euclidean distance between the vectors of the quantum states on the Bloch sphere. It would be useful to match that for the vector that corresponds to the unitary gates to examine the relationship between the expectation time and the volume defined by the error in the space of the unitary gate vectors since, by our assumptions, we make

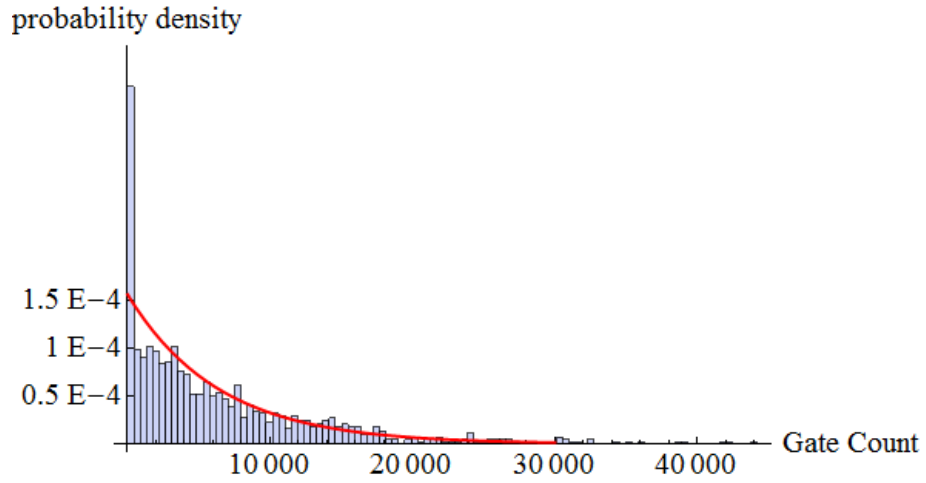


Figure 4.34: Probability distribution of required gate count to achieve target for a two-parameter interaction in a 100 bin histogram compared with an exponential distribution parametrised by the mean plotted in the red curve.

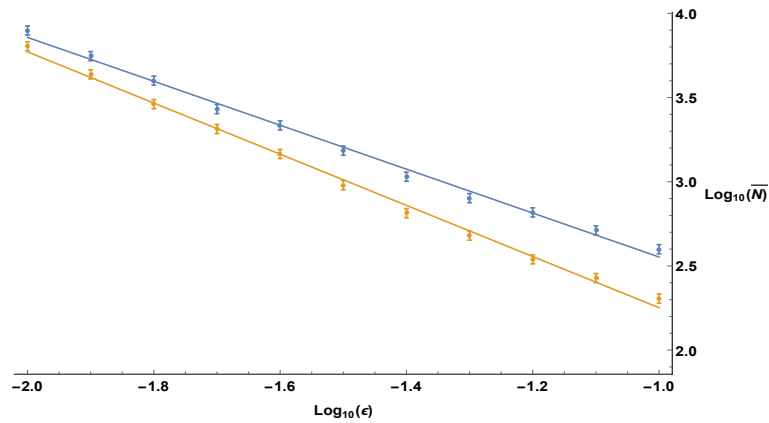


Figure 4.35: A logarithmic plot of the average number of gates over 1000 simulations to hit a target unitary $R_{\hat{x}}(\frac{\pi}{2})$ against the target error, ϵ for the one-parameter case (in blue) and the two-parameter case (in gold). A linear curve was deduced for each data set by a least squares fitting, with a gradient of ≈ -1.3 for the one-parameter case and -1.5 for the two-parameter case. For further details, see appendix B.

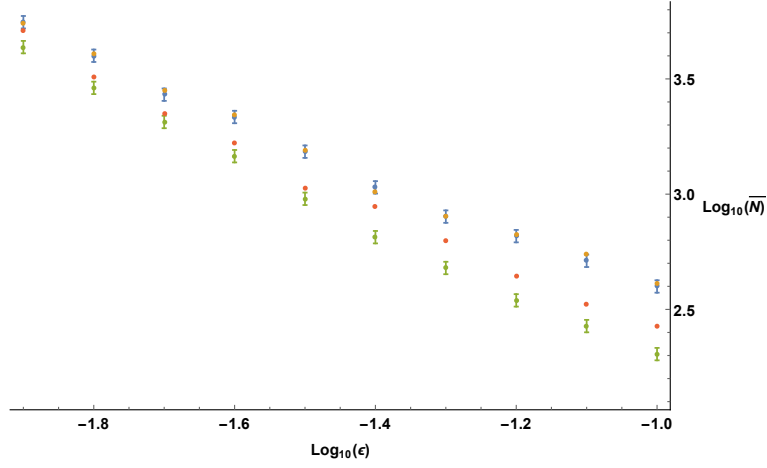


Figure 4.36: A logarithmic plot of the average number of gates over 1000 simulations to hit a target unitary $R_{\hat{x}}(\frac{\pi}{2})$ against the target error, ϵ for the one-parameter case (in blue) and the two-parameter case (in green). The estimated standard deviations are also plotted with the one-parameter case in orange and the two-parameter case in red. Under the exponential distributions these would be equal to the mean. We see that for the one-parameter case they lie close to the mean but for the two-parameter case, they are significantly larger.

the prediction that the general trend for further simulations is for the expectation time to scale with the volume.

Yet, by the composition of single-qubit unitary matrices (2.10), for

$$U = c_1 \mathbb{I} - i s_1 \hat{\mathbf{m}} \cdot \hat{\sigma}, \quad (4.157)$$

$$V = c_2 \mathbb{I} - i s_2 \hat{\mathbf{n}} \cdot \hat{\sigma}, \quad (4.158)$$

$$\frac{1}{2} |\text{Tr}[V^\dagger U]| = c_1 c_2 + s_1 s_2 \hat{\mathbf{m}} \cdot \hat{\mathbf{n}}. \quad (4.159)$$

As an alternative, consider the measure

$$\delta(U_T, V) = |\mathbf{m} - \mathbf{n}| = \sqrt{\sum_j (m_{x_j} - n_{x_j})^2} \quad (4.160)$$

where

$$\mathbf{m} = s_1 \hat{\mathbf{m}}. \quad (4.161)$$

This means that for each vector, the magnitude corresponds to $s_{1/2}$, the sine of the half of the rotation angle of the unitary operation. $\text{Sin}(\frac{\gamma}{2})$ is a one to one map of γ from $[-\pi, \pi]$ to $[-1, 1]$ as a continuous function and forms an appropriate measure of magnitude for the vector. For $U \in SU(2)$, the term can be calculated from the trace

operation:

$$s_1 m_{x_j} = \frac{i}{2} \text{Tr}[\sigma_{x_j} U] \quad (4.162)$$

and the measure is calculated

$$\delta(U_T, V)^2 = -\frac{1}{4} \sum_j \text{Tr}[\sigma_{x_j} (U_T - V)]^2. \quad (4.163)$$

This measure matches the Euclidean distance between vectors on a spherical representation of $SU(2)$. We therefore expect that using it will cause the time to scale with the third power of the error, proportional with the volume it designates in the sphere. Using this, we ran the same simulation for figure 4.35 but with the measure (4.160)¹ to create the plot in figure 4.37 which confirms those expectations.

Interpretation of the simulation

The simulations demonstrate a behaviour that is close to our expectations. However we notice that the logarithmic plot of the scaling of the expectation times against the error bound does not fit with the results and their error bars. This can be explained by the noted effects in the distribution of the hitting times. The ability to hit the target has to in reality be split into finite time intervals and there is a supression of the probability to succeed around particular times when the random walk gets very close to or hits exactly on the target. The fact that these times are dependent on the gates being randomly generated by the model explains the differences in the curves between the two-parameter and one-parameter cases. However, we have developed methods that will aid in extending the scope of the simulations such as the use of the Euclidean distance method which did provide the expected increase in the gradient of the logarithmic plots in line with the predicted scaling behaviour.

To be really sure of this scaling one needs to find the scaling for a large sample of gates but if in fact such a large step up in the number of calculations is to be done, it may be more effective to use it to confirm the assumed underlying behaviour by looking at the gates generated at every step and see if over time the random walk forms a uniform distribution over the spherical representation of unitary gates. If the assumption turns out not to be true then a set of generator gates may have some gates which they favor. The target gate could be decomposed into a finite gate set built out of favoured gates. This would scale up the time and/or the error by a polylogarithmic function from the additional gates in the decomposition but may overall be advantageous depending on the scale of the bias towards the gates in the decomposition. If the assumption is true, then exploring the times needed for a finite gate set may be used to derive an upper

¹The factor in the measure relies on the unitary being in $SU(2)$ so the generating elements have to be in $SU(2)$ or a phase correction calculated from $\arg(\text{Tr}[V])$.

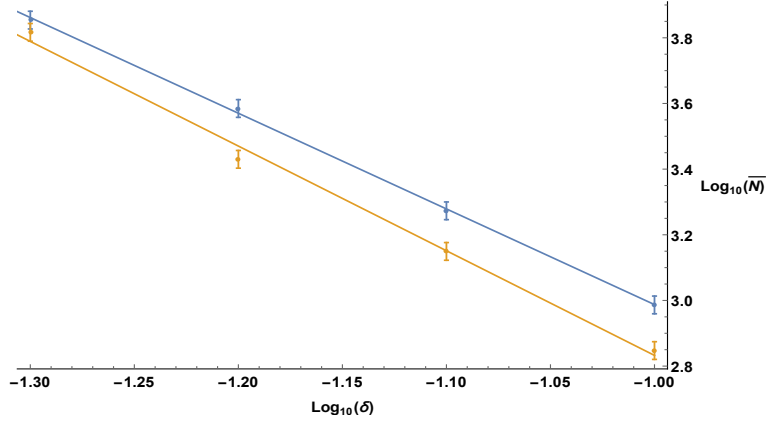


Figure 4.37: A logarithmic plot of the mean number of gates over 1000 simulations to hit a target unitary $R_{\hat{x}}(\frac{\pi}{2})$ against the target error, δ , defined by the Euclidean distance for the one-parameter case (in blue) and the two-parameter case (in gold). A linear curve was deduced for each data set by a least squares fitting, with a gradient of ≈ -2.9 for the one-parameter case and -3.2 for the two-parameter case. For further details, see appendix B.

bound for all gates based on the expected time needed to generate each gate in that set multiplied by the number of uses of each gate in the decomposition.

4.4.3 Stochastic generation of two-qubit gates: walks on a circle

Any arbitrary two-qubit gate may require three parameters for the non-local part of the decomposition but we have seen how the gates generated from ancilla driven computation must be of the same commuting classes of one-parameter gates. These gates are represented by a line in the Weyl chamber that closes under symmetry. Since this is the line that maps to the control unitary gates, we can represent them with a circle, $C(\Phi)$, each unique gate being represent by an angle Φ that denotes its entangling power. The angles are randomly generated and several generations of gates will execute a random walk on the circle. It is therefore expected that one can perform a two-qubit gate stochastically as per the one-qubit gate but with the advantage that a one dimensional curve will have a linear scaling with the error, $\gamma = |\Phi - \Phi_T|$, for target gate $C(\Phi_T)$.

Simulation of stochastic two-qubit gates with one-parameter interaction gates

We simulated the creation of a CZ equivalent gate by the random generation of entangling gates. The gate generation was based on an interaction gate with coupling parameter $\alpha = \frac{\pi}{16}$, preparation parameter $\beta = \frac{\pi}{32}$ and measurement parameter $\theta = \frac{\pi}{2}$.

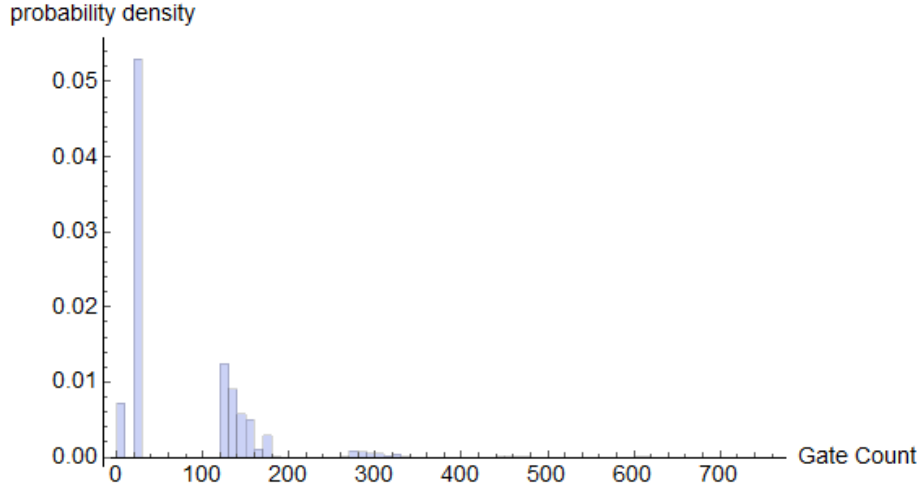


Figure 4.38: The distribution of the target region hitting times given by the simulation of the gate generation with an interaction with $\alpha = \frac{\pi}{16}$ out of 10000 simulations placed in 100 bins. The mean is 73.1 and the standard deviation is 75.1 (to 3 s.f.).

This results in $U_0 \equiv C(\Phi_0)$, $U_1 \equiv C(\Phi_1)$ where, according to (4.104),(4.105),(4.106) and (4.107):

$$\Phi_0 = 4\arctan\left(\frac{-\cos(\frac{\pi}{32})}{\cos(\frac{3\pi}{32})}\right) + \pi, \quad (4.164)$$

$$\Phi_1 = 4\arctan\left(\frac{\sin(\frac{3\pi}{32})}{\sin(\frac{\pi}{32})}\right) + \pi. \quad (4.165)$$

The target gate is $C(\Phi_T) = C(\frac{\pi}{2})$, so that the choice of target and the gate generation results exclude the CZ gate, and the measure of distance is the angle between the product of the generations and the target on the circle representing the class:

$$\gamma(C(\Phi), C(\Phi_T)) = |\Phi - \Phi_T|. \quad (4.166)$$

This was performed 10,000 times, with an error bound, γ , of $\frac{\pi}{100}$ and the resulting distribution is displayed in figure 4.38.

The distribution of hitting times drops off similarly to the previous test but the gap between time results is far more pronounced. One of the outputs is an irregular angle that depends on the coupling strength while the other is π for all couplings. The latter result creates a probability of success in the first step while the former can be used to reach the target within the bound if applied a large number of times scaling with the error according to the dimensions of the space, in this case, linearly but creates the large gaps in the distributions as it needs a large number of steps to get close to the target. One can view the distribution as the effect of the small angle being able to get close to the target after $N_0, N_1, N_2 \dots$ steps but with a random number of π angles

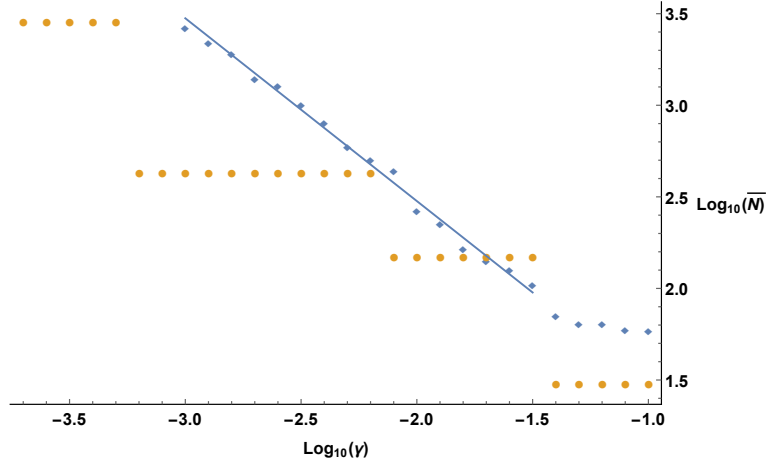


Figure 4.39: A logarithmic plot of the mean number of gates over 10,000 simulations required to hit a target gate against the error bound γ . The target is a gate equivalent to $C(\frac{\pi}{2})$ and the interaction has $\alpha = \frac{\pi}{16}$ and preparation parameter $\beta = \frac{\alpha}{2}$, chosen to eliminate the skew caused by the constant probability of success and the periodicity of a π angle generation. The orange plot with circle points are the times for a deterministic output using the minor result, the blue plot with diamond points is the random gate generation. The latter behaves similarly to the first at large error sizes when the first probability of success dominates. After that, a linear relationship dominates and the points can be fit to a gradient of 1.00 (to 2.d.p.). For further details, see appendix B

generated in between the steps and ruining an opportunity unless an even number of them have cancelled out; that could then apply to any pair of results where each one requires its own number of steps to approximate a π angle.

If the gate generation occurred deterministically, making steps equal to the minor output each time, then for a specific target, the walk can stop at specific points at $N_0, N_1, N_2 \dots$ steps that will be within error bounds $\gamma_0, \gamma_1, \gamma_2 \dots$ so that a single value of N will be valid in the range $\gamma_{j+1} < \gamma < \gamma_j$. This is reflected in discontinuities in the error-step scaling relationship (see figure 4.39). However, generally and as $\frac{1}{\epsilon}$ grows, it is expected to tend to a linear relationship [75]. When the random gate generation is introduced, the random walk may miss the position at N_0 steps but hit later steps and, since an irrational step size can generate a position arbitrarily close to the target, the set of valid end points is of order of the space in between the error bound. This provides a more explicit picture for why the expected scaling for the average time to hit an approximation of a target, even with randomness, is still the dimension of the space of the random walk and can apply to any pair of step sizes as long as one is an irrational multiple of π .

While some small simulations indicate support the expectations of behaviour, a

robust characterisation requires significantly more computational time for a greater number of targets over a greater range of error size (just compare our ranges to those in reference [35]). Our solution is to enforce such properties by devising strategies in which the behaviour of the system is regulated according to a strategy that records results and feeds them back into an adjustment of the ancilla parameters.

Unregulated random gate generation makes no use of the resources of control over the ancilla preparation and measurement and is unoptimised for implementation times. The expectation is that it is the most time-costly method of implementation of universal ADQC with arbitrary coupling parameters and that by introducing some method of control, the additional resources trade off for time.

4.4.4 Applying strategies to random gate generation

We will apply strategies to generating a two-qubit gate. There are several advantages in the case of two-qubit gates. For one thing, it is only necessary to be able to achieve a single entangling gate for universal quantum computation. So while a walk with irrational step sizes can always approximate it, if the target is a rational multiple of π itself then of course a walk over steps of rational multiples of π (for certain factors depending on the target) will be able to hit it. The most likely choice of target will be the maximally entangling CZ equivalent gates. The multiplication of a group of gates that all commute is also simpler, making simple rotations on the circle and allowing us to adjust the parameters to change directions and step sizes. These advantages can be combined if we focus on adjusting the parameters so that the random walk is performed only on a finite space of very few points.

Repeat-until-success

In probabilistic gate proposals for linear optical systems [158, 159, 99], entangling gate operations rely on measurements that will either result in the desired action or if failed, produce a state onto which the exact same action can be repeated in the next step. Effectively, a walk is performed over only two points and the time to generate a finite probability of hitting the target is a single operation step. The statistics of the hitting time then faithfully obey a geometric distribution with an expectation time of $\frac{1}{p}$ where p is the probability of success. The probabilistic gate generation then benefits from an efficient relationship to the probability error bound as per (4.141), relying on a single self contained process in repeated steps rather than having to feed forward results and a simple, consistent failure event- the identity operation.

We provide two method that allows one to enact a “success/fail” scenario based on the strategies covered in section 4.3. The first method ensures that failure corresponds

to generation of identity by replicating the structure observed in the basic case in section 4.3.1.

In this protocol gate generation is enacted twice: once as described before and once again with parameters that convert $\Phi \rightarrow -\Phi$. The latter can be done either by local Pauli-X gate pre- and post-corrections on the register or, if there is access to the ancilla between interactions, by making a correction to the intermediate ancilla unitary to exchange the resulting intermediate states $|a_0\rangle \rightarrow |a_1\rangle$. The change in the second implementation enacts a change of the resulting gate $C(\gamma) \rightarrow C(-\gamma)$. If these two output gates are (under local operator corrections) $C(\Phi_0)$ and $C(\Phi_1)$ with probabilities p_0 and p_1 then after two generations the Control-Unitary is one of $C(\Phi_0 - \Phi_1)$, $C(\Phi_1 - \Phi_0)$ or \mathbb{I} . If $\Phi_0 - \Phi_1 = \pi$ then we have performed probabilistic CZ generation with a success probability of $2p_0p_1$.

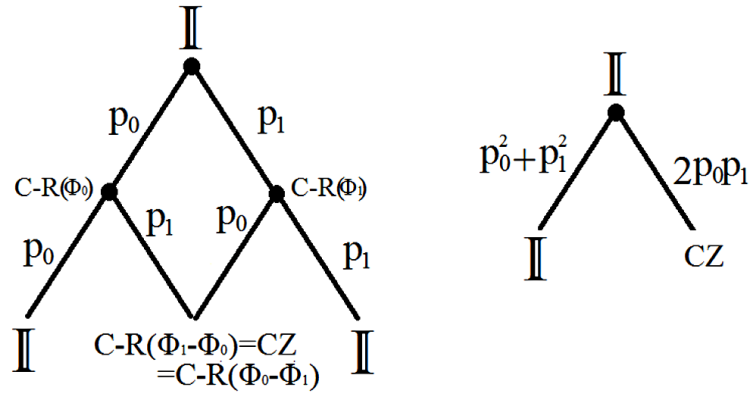


Figure 4.40: Since all enacted gates are of the same group, Control- Z_Φ gates, multiple gates can be easily combined. Random gate production is a Markovian process. By manipulating the relative values, the output can be limited to a finite probability tree (left), including a case equivalent to a single outcome “success/fail” gate suitable for a repeat-until-success method (right).

For example, in figure 4.15, we see that we can lower β to match such conditions. In figure 4.41, the value of $2p_0p_1$ is calculated for $\alpha = \frac{\pi}{16}$, giving us a value of $p \approx 0.128$. This can be calculated for every value of α that characterises the interaction gate and provides a smooth relationship between the coupling strength and the probability of success as can be seen in figure 4.42

In this strategy, every ancilla parameter can be set up beforehand and the results only have to be fed through in order to cease the process: the flip only occurs on odd steps but it occurs on every odd step. If there is a cost to adjusting the ancilla parameters then that cost is likely a constant for α .

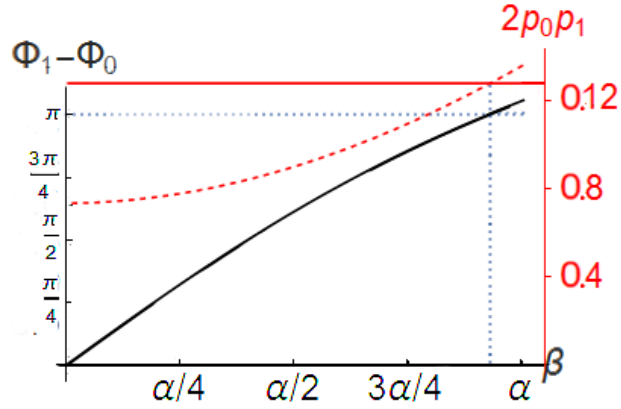


Figure 4.41: Plot of the probability of the probability of achieving two opposing ancilla measurements, $2p_0p_1$, against β , displayed on a red dashed curve. The difference between the two outputs of the two-qubit gate generation is plotted against β on the solid black curve. The solution to $\Phi_1 - \Phi_0 = \pi$ can be found through numerical minimisation methods and the resulting β value used to calculate the resulting probability. In this case where $\delta\Phi = \pi$, $p = 0.128$ (to 3 s.f.).

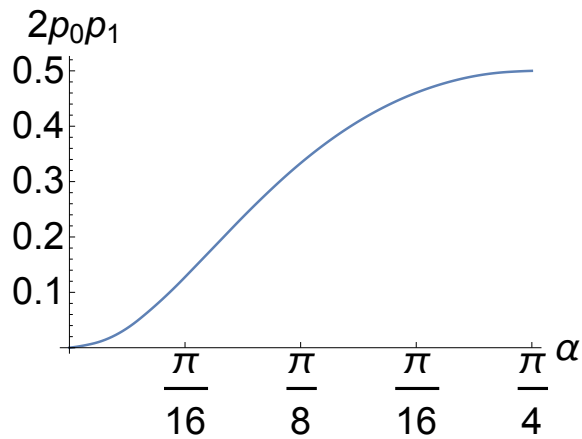


Figure 4.42: The probability of success of the Repeat-until-success strategy in a pair of steps, equal to $2p_0p_1$, against the interaction gate coupling strength $0 \leq \alpha < \frac{\pi}{4}$.

Flip-undo

The next method loses the consistent relationship between failure and the identity operation but it maintains the behaviour of a walk on a familiar finite graph and simplifies the implementation in the ADQC scheme.

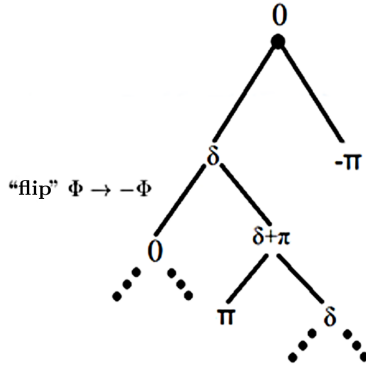


Figure 4.43: The probability tree of the flip-undo strategy receives all possible points in the strategy after 2 steps. After the first step, it can be seen as a repeat-until-success strategy where the time to repeat is 2 gate generations.

Up until now we have discussed the ability to manipulate the ancilla with the assumption that we can exercise any arbitrary single-qubit unitary gate. This is in line with the requirements of the ancilla-driven and ancilla-control schemes. However we have also developed a scheme for exploring what can be done with as simple an action on the ancilla as possible: we have only available to us a fixed preparation state, a fixed measurement basis and the choice of whether or not to implement a bit flip gate, X - specifically the bit flip required to change the sense of direction of the ports. What this provides is the ability to attempt to undo a previous action hence the designation of the “flip undo” strategy.

After attempting to generate a $C(\pi)$ gate in a single step, if the result failed, attempt to go back to the origin. Whether you have arrived at the origin or not, attempt to generate a $C(\pi)$ gate with the product of the next generation. If one fails again, repeat the process from the second step. Repeat until success.

It should seem obvious that in a case where either $C(\pi)$ or $C(\gamma)$ is generated and the target is $CZ = C(\pi)$ that it is preferable to label a result $C(\gamma)$ a failure and attempt to undo it in order to try again to generate CZ directly. Yet that then creates a possible result where the sequence product is $C^\dagger(\pi)C(\gamma) = C(\gamma + \pi)$. Again the apparent best decision is to attempt to undo $C(\gamma)$ as this will now immediately lead to the target gate. In the following step, the only two possible product sequences must result in $C(\pi)$ or $C(\gamma)$ which makes employing this strategy form a closed loop.

Now that there is a description and probability tree for a finite number of points on the circle, we can find an exact description of the time statistics using the recursive relationships between the expectation times at different points, if we take the probability of generating $C(\pi)$ in the first step to be p :

Note that this description provides a probability tree, figure 4.43, that has the same structure as the tree in figure 4.22 in section 4.3.1. If we take the probability of

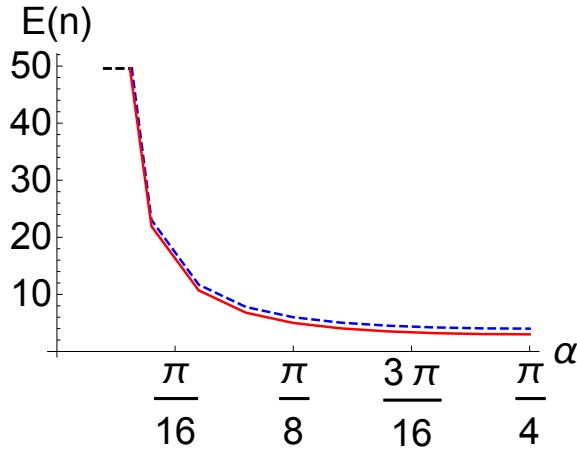


Figure 4.44: Plot of the expected number of steps for the repeat-until-success strategy (the dashed blue curve) and the flip-undo strategy (solid red) against the interaction gate coupling parameter α .

generating $C(\pi)$ in the first step to be p , the solution to the expected time for success is provided by (4.126):

$$\bar{n} = 1 + \frac{1}{p}$$

and the time to guarantee a result is (4.148):

$$n \geq 2 \frac{\ln \frac{\epsilon}{q}}{\ln(1 - 2pq)} + 1.$$

The flip is conditional on certain results in the sub-tree so the results in this case need to be fed forward and the average additional cost with the flip will be dependent on the branch probabilities and thus the coupling parameter. Notice in figure 4.44 how relatively small the difference between the two strategies is, tending towards a maximum difference of one step at maximum coupling parameter. The greater advantages comes with the possible increase in difficulty in setting the ancilla parameters of the repeat-until-success strategy so let this be represented by \bar{n}_U per step and the cost of the flip is \bar{n}_f , the repeat-until-success expectation time is now

$$\bar{n} = \frac{1}{p_0 p_1} \left(1 + \frac{1}{2} (\bar{n}_f + 2\bar{n}_U) \right) \quad (4.167)$$

and for the flip-undo approach it is

$$\bar{n} = 1 + \frac{1}{p} \left(1 + \frac{\bar{n}_f}{2} (2 - p) \right). \quad (4.168)$$

By applying different relative and absolute values for \bar{n}_f and \bar{n}_U as in figure 4.45, a quantitative impact of the different resource costs can be determined.

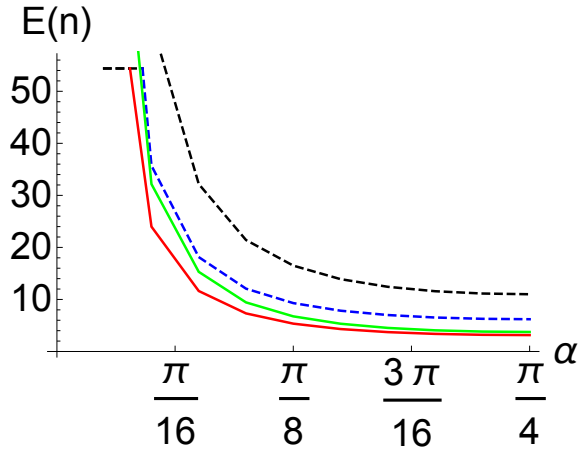


Figure 4.45: Plot of the expected number of steps for a scheme against interaction gate coupling parameter α with (a) $\bar{n}_f = 0.1$, $\bar{n}_U = 0.5$, the flip-undo strategy is given by the solid red curve, the repeat-until-success strategy the dashed blue (b) $\bar{n}_f = 0.5$, $\bar{n}_U = 1.5$, the flip-undo strategy is given by the solid green curve, the repeat-until-success strategy the dashed black

4.4.5 Concluding Commentary

Ancilla driven quantum computation with arbitrary coupling parameters can implement universal quantum gates, generally, with a cost of probabilistic implementation times. The successive application of the ADQC process with arbitrary ancilla parameters will execute a random walk over the group of the generated results. A two-parameter interaction gate or single-parameter interaction gate with appropriate local gate effects will execute a random walk over all single-qubit gates or single-parameter two-qubit gates. We suggest that the time to reach a particular target follows an exponential distribution and that the average time scales with the error of approximation of the gate up to the power of the dimension of the gate group.

This was tested with a simulation for specific target gates with results supporting our hypothesis. However in addition to the lack of a comprehensive simulation for the times to achieve a large sample of gates, an exact expression of the target gate in terms of the gates generating each step led to some off-model behaviour. This provides some features to test when extending the average over a large number of target unitaries. An average over many random target unitaries may smooth out such effects and verify the general applicability of the exponential distribution model or one may look at the distributions of locations of the random walk over time to see if it grows more uniform over the space. With the ability to record the results of ancilla measurement and so know the resulting random walk and the ability to stop the process when the walk has hit the desired target to within the desired error, this allows ADQC with non-maximally

entangling gates to achieve universality with a probabilistic, polynomial time overhead compared to deterministic ADQC.

When generating two-qubit gates, the results can be restricted to the single-parameter class so that the random walk only has to occur over that single parameter but this requires compensation for local unitary effects which cannot be avoided if using a two-parameter interaction gate. An advantage of generating gates restricted to the single-parameter class is that we can adopt simple strategies similar to those in section 4.3 or in probabilistic gate proposals for linear optical systems. These strategies also form a relationship between the coupling parameter and the expected time so additional coupling parameter strength becomes a resource that can be used to reduce time costs. Both example strategies require the ability to control the ancilla parameters for preparation and measurement but to different degrees and so have their own associated additional costs.

The ability to guide the random behaviour of the two-qubit gates raises the question of how we can do so for the single-qubit gate. The ability to do so with two-qubit gates is related to the ability to probabilistically generate any element of the relevant group in a single step thus naturally including the inverses of previous results. For single-qubit gates, the class of gates generated in a single step that is caused by local gate effects may be universal but does not itself include the entire group of single-qubit gates and excludes the inverses of nearly all its members.

Chapter 5

Extreme conditions: long distances & minimal control

In the next chapter, we will discuss the behaviour of probabilistic ADQC under two scenarios that extend key features of the mode to extremes. The first relates to the dichotomy of mobility between the ancilla and register that makes hybrid systems of interest for scaling up quantum computation and in networked models. The second relates to an implication of the use of non-maximally entangling gates- a loss of control over the ancilla as the number of ancilla qubits grows and the lifetime or engineering flexibility decreases.

5.1 Entangling unitary gates on distant qubits with ancilla feedback

5.1.1 Intro

One use of the ADQC scheme is to aid in scaling up a stable memory register. If a physical system that produces stable memory qubits is not easily scaled up to handle many qubits, one solution may be to replicate the system many times and then connect each one up through an ancilla system. There are other proposals that aim to improve scalability in similar ways. In proposals for networked quantum computation a memory qubit may be placed in a node of qubits along with the several other qubits that are needed per memory qubit to perform error-correction and fault tolerance codes [103]. In these cases different local nodes of qubits may have to be connected by a different physical medium, such as a photon or coherent beam system [42]. A distributed design may also aid in parallelising circuit design where computational tasks are split up and done in parallel for a time speed up or in aiding scalability of a physical implementation [158, 160].

This principle can be looked at in a larger scale in proposals such as Blind Quantum Computation [161, 162] where it is proposed that a client may send a computation to a server which has access to greater quantum computation resources. As quantum computers actually start to become technology feasible and start to be created, it will likely start with a handful of quantum computers in the world with limited memories and perhaps not capable of universal quantum computation. In that scenario, one may wish to share these rare resources between centres around the world. ADQC may be a good fit for this scenario because then there is no guarantee that each quantum computer nor the network that connects them are made from the same physical implementations.

The problem of entangling a physically separated pair has been considered before with methods such as the Barrett-Kok double heralding approach [102] where entangled states are generated through projecting the system via photon pair measurements or Lim, Barrett *et al* [152]’s repeat-until-success method through Bell basis measurements. A particular protocol may be optimised for particular resource requirements such as with Eisert *et al* [153]’s optimisation of entangling gate operations and classical communication.

Some of the features of ancilla driven quantum computation distinguish it when applied to the problem of long distance operations. It is capable of using a single qubit ancilla and can consider arbitrary entangling power between the register and the ancilla which may be more relevant to distributed designs that don’t use an optical medium. Conversely, some of the limitations of the model such as the restriction to one interaction gate per register per ancilla qubit make sense when relatively long distances may restrict repeated transmission back and forth between two nodes.

The problem also raises some questions for ADQC and the limits of the model. Applying control over both parameters of the ancilla would have to be coordinated by two ends of a distant connection. What speed up in establishing a maximally entangling gate can occur if we utilise full control in a walk between entangling gates? If only one side can perform any optimisation or feedback of the parameters, what would the slow down be? How is the fidelity of the output affected by the increased exposure to transmission noise? Do the two register devices have to interact the same way or with the same coupling parameter?

5.1.2 The model

For two parties to coordinate an ADQC CZ gate on a pair of qubits that they share between themselves, an ancilla qubit can be prepared, interacted with a register qubit and then transmitted by the first party to the receiver who then interacts the ancilla with their qubit and then measures the ancilla..

For a generation of two-qubit entangling gates, we have found there to be two

free variables, labelled (β, θ) , which set the probabilities and entangling power of the generated gates. Since β is fixed by the elevation of the initial ancilla state before interaction while θ fixes what the measurement basis must be, the former is under control of the transmitter and the latter the receiver. However the preparation and measurement determine two parameters of the intermediate unitary that must be performed on the ancilla qubit in between each interaction gate with the register. This intermediate unitary has been decomposed into $U_a = R_{\hat{y}}(\theta - \theta')R_{\hat{n}}(\frac{\pi}{2})$ where \hat{n} and θ' depend upon the preparation and θ sets the measurement. It can be decomposed again into $U_a = R_{\hat{y}}(\theta - \Theta)R_{\hat{y}}(\Theta - \theta')R_{\hat{n}}(\frac{\pi}{2})$ so the transmitter can perform the operation $R_{\hat{y}}(\Theta - \theta')R_{\hat{n}}(\frac{\pi}{2})$ before sending the ancilla off while the receiver performs $U_a = R_{\hat{y}}(\theta - \Theta)$ on receipt of the ancilla. As long as there is a pre-agreed Θ they do not have to coordinate their actions during the transmission.

After a measurement, the result can be sent from the receiver back to the transmitter by classical communications. The transmitter can also send a set of instructions for each measurement classically thus there can be feedback of each result into the next set of parameters if there is a one-way classical communication channel for each parameter and there is time to send those communications in between each new ancilla qubit. See figure 5.1.

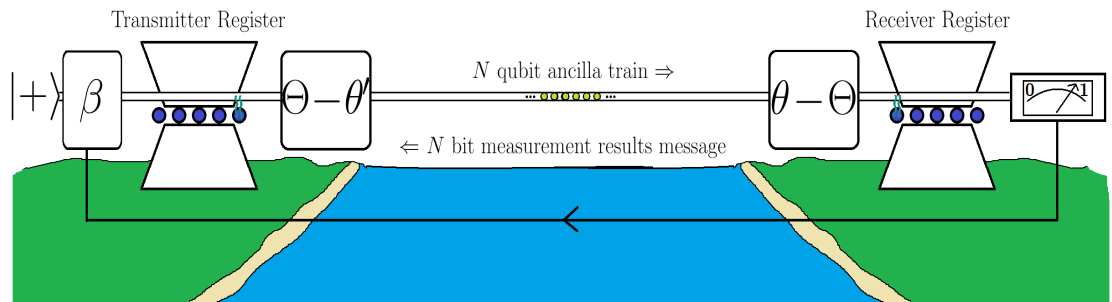


Figure 5.1: Representation of the arrangement of a long distance entangling gate over two separated registers.

On the other hand, imagine that the amount of communication is to be minimised and so the transmitter sends a package of N ancilla qubits, preceded by a set of instructions. The instructions may detail a strategy that the receiver must follow and then once they make the measurements, the results are sent back. Since the strategies have been probabilistic the package may be large enough to ensure a very high probability of success but the actual number of qubits needed in the package will be smaller. The receiver would need to be able to recognise when it has achieved success and then in order to preserve the entangling unitary operation on the two registers, the excess ancilla qubits must be intercepted and measured before being interacted

with the second register qubit. If they are measured in the computational basis then the receiver can send back the results for the transmitter to apply the necessary local unitary corrections.

5.1.3 The one-step strategy

Previously, in section 4.4, it was argued how the random generation of gates should be loosely bounded by the time distribution of a gate being generated within a fixed time which leads to a broadly exponential distribution of times (see section 4.4.2 specifically). Also strategies were concocted in section 4.4.4 that exactly enacted a random gate under such principles where the steps of a procedure could simply be repeated in the next time interval if it had not previously succeeded. The “repeat until success” and “flip undo” strategies take 2 or 3 steps to attempt to implement the desired gate before returning to an earlier starting point.

If the time interval before one could re-enact generating the gate was reduced, the average time to do so would be reduced. The natural conclusion is to look for a strategy in which there is a probability of success in every step.

So at each step, we set the conditions so that one measurement result generates a gate which corresponds to the angle difference between the present point on the circle and the point π . If this measurement does not occur, find the distance between the present point and the point π and attempt to generate that gate. Upon every failure, find the new distance between the target and the current point and attempt to generate that gate. This is the “one step” strategy (see figure 5.2).

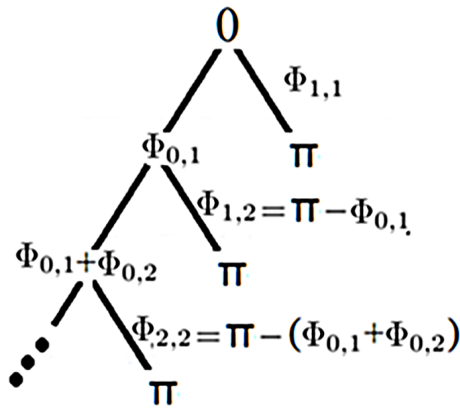


Figure 5.2: The probability tree of the “one step” strategy. The first step will always require generating a CZ equivalent gate, the other result will be dependent on the coupling and will then dictate all future conditions. The conditions are reset at each step with each new ancilla.

Given any coupling strength, at the start of the strategy, the first attempt to generate CZ is performed the same way: the ancilla is prepared in the $+$ state and then measured in the $|\pm\rangle$ basis with the “ $-$ ” (port 1) result generating a gate equivalent to CZ ($C(\pi)$). The “ $+$ ” (port 0) result would generate a blow back gate $C(\Phi_0)$. There

is a sense of direction with the gate generation; one port gives $C(-|\Phi_0|)$, the other $C(+|\Phi_1|)$, clockwise or anticlockwise around the circle that represents the $C(\gamma)$ group (see figure 5.4). One can simply switch the direction association of the ports by performing a bit flip either immediately before or after transmission from Alice to Bob, which does not change the probability of success, so we will ignore the exact sign requirements in the notation from here on and simply note the need to flip. Having travelled “clockwise”, the best next step is to continue in that direction and generate $C(\pi - |\Phi_0|)$. If Φ_0 is small then $\pi - \Phi_0$ will be large enough that it can only be generated from port 1, the port with larger Φ but smaller probabilities upper-bounded by $\frac{1}{2}$. Another feature of port 1 is that the probability increases as the preparation and measurement variables (β, θ) are increased and for a fixed Φ_1 , θ increases with β . Therefore for optimal probability, it is only needed to fix one of these parameters to the maximum and vary the other.

At every step n , there is one gate that matches success $C(\Phi_{1,n})$ and a failure gate $C(\Phi_{0,n})$, therefore to be at step n , the current action on the register system is the product of previous failures $C(-|\Phi_{0,1}| + \Phi_{0,2} + \dots + \Phi_{0,n-1})$. The next gate to be generated for success must be $C(\pi - (|\Phi_{0,1}| - \Phi_{0,2} - \dots - \Phi_{0,n-1}))$.

The magnitude of the angle Φ of both ports increases with the probability of success of Φ_1 (see figure 4.20). So because the largest angle to be generated is π in the first step, the first step has the highest probability of success and also the highest value of the failure gate Φ_0 . Therefore $\pi - |\Phi_{0,1}|$ is the smallest value and has the smallest probability of success. These two first values provide a bound on the behaviour of the strategy (see figure 5.3 for an example of the spread of probabilities). The cumulative distribution function based measure, $P(n < N)$, can be compared to the CDF of constant probability for each step using the extreme probabilities: $1 - (1 - p_2)^n < P(n < N) < 1 - (1 - p_1)^n$.

This also means that the first step provides the threshold for when a two-port strategy is viable: when is $\pi - |\Phi_{0,1}|$ small enough that it can be generated from port 0? Since $\Phi_{0,1}$ is also the maximum Φ_0 , it must be when $\Phi_{0,1} = \frac{\pi}{2}$. Port 0 has a different (β, θ) for fixed Φ_0 relationship and it’s probabilities are optimised away from the fixed measurement conditions so this is also the threshold for when a 2 degree of freedom strategy can be involved.

5.1.4 Numerical results

We found the parameters and resulting probabilities for continuing with the one-step strategy for 500 steps for a range of coupling strengths of the connection interaction. The full range for $0 < \alpha < \frac{\pi}{4}$ was covered for the 1-port 1-degree-of-freedom strategy where the degree of freedom was represented by the preparation parameter β (plotted

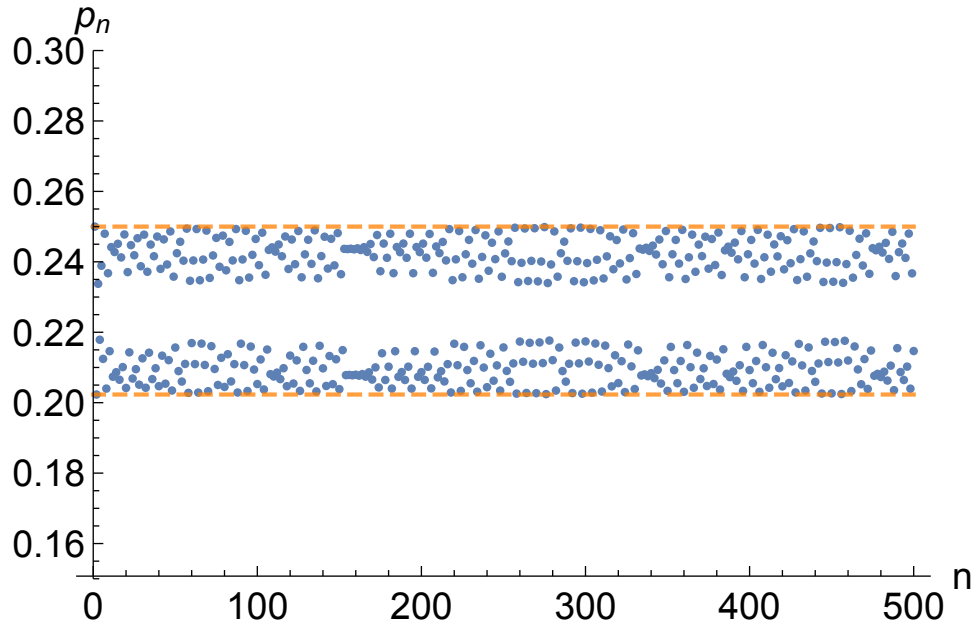


Figure 5.3: For $\alpha = \frac{\pi}{8}$, the values of the probability p_n of a successful two-qubit entangling gate generation on step n for the first 500 steps of the one-step strategy. The dashed orange lines at 0.25 and 0.202314 match the probabilities of the first two steps and bind all subsequent probabilities.

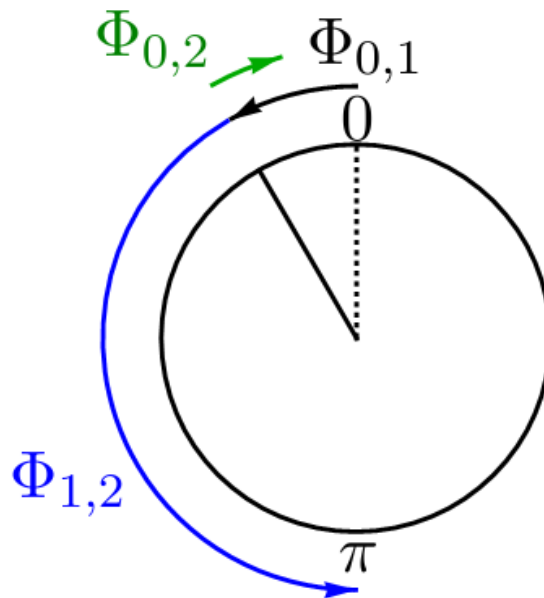


Figure 5.4: Representation of the random walk of the gate generation over a circle, demonstrating the change of direction of the port 0 output.

in figures 5.5 and 5.6). We then found more values for the range of coupling strengths that starts just before the threshold for the two-port strategy (see figure 5.7). In this range we then found the values for a two-port strategy where one could only vary β and perform no optimisation of port 0 and then found them again for when optimisation over β & θ is allowed. Finally we checked for just the one-port strategy, the probabilities for each step when the measurement parameter is the allowed degree of freedom rather than the preparation and this did turn out to give the exact same results.

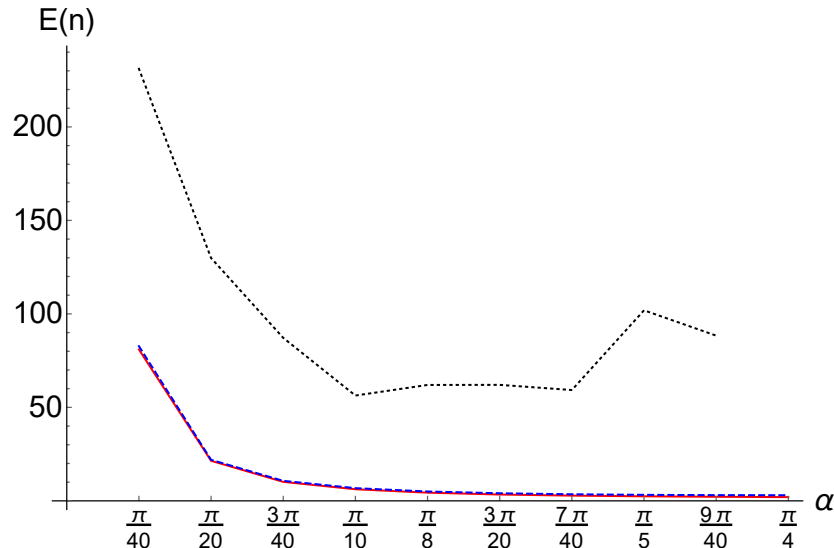


Figure 5.5: A comparison of the unguided (dotted black) generation against the one-step (solid red) and flip-undo (dashed blue) strategy with their expected hitting times against coupling strength. Relative to the unguided approach, the flip-undo strategy is nearly as completely effective as the one-step strategy while required less ancilla control. The unguided expectation times are not well correlated with the coupling strength: while significant differences in step size from large coupling differences do impact on the number of steps, the ability to approach arbitrarily close to the target relates to the difference of the step size with rational divisions of π making for small scale chaotic behaviour.

The order of improvement between different one-step strategies is less than a single step in the expectation time. The 1 port strategy tends towards an expectation number of 2 while the 2 port/ 2 degree of freedom strategy tends toward 1.5 ; the 2 port/ 1 d.o.f. approach has a peak in improvement near the middle of the range but at very close to the maximum coupling returns back to the 1 port strategy. This scale of improvement can be expected from the behaviour of the probability of either port. As $\alpha \rightarrow \frac{\pi}{4}$, $\Phi_{0,1} \rightarrow -\pi$, $\pi - |\Phi_{0,1}|$ becomes very small and thus the probability out of port 0 in the second step tends to 1. Any consideration of multi-step strategies in this range is

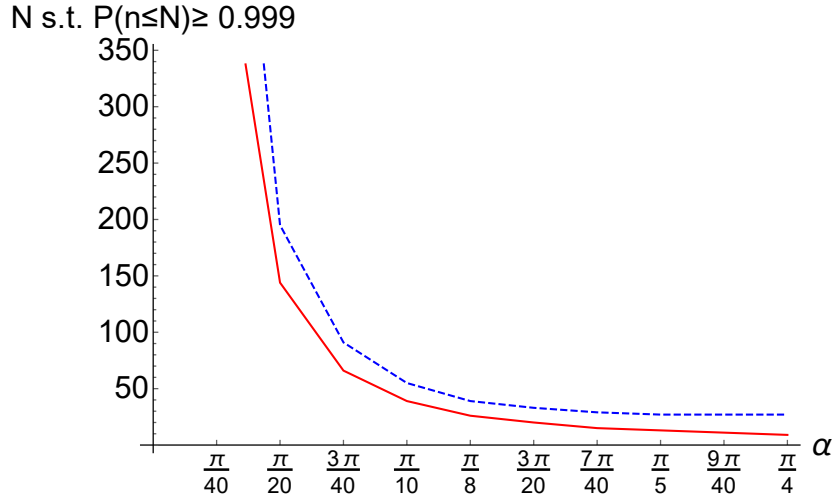


Figure 5.6: A comparison of the one-step (solid red) strategy against the flip-undo (dashed blue) strategy: the number of ancilla that need to be prepared to guarantee a 99.9% chance of success.

therefore of limited advantage – the behaviour where failure in the first step improves the probability of success in the second step which we would expect to be a feature of any two step strategy is already a feature of the one-step strategy with two ports and access to both parameters.

The viability of a multi-step strategy at the lower coupling step range can be examined using the lower bound on the probabilities of success in each step found from the probability in the second step. This value describes the behaviour of a geometric distribution that bounds the behaviour of the one-step strategy; for a multi-step strategy to be effective it must at least improve upon this and since a multi-step strategy takes place over n steps, it must improve the probability by at least a factor of n yet this will be limited by the maximum value of 1. In figure 5.8 we can see what the maximum possible number of steps for a multi-step strategy could be for any improvement to be possible.

The most striking result is that the flip-undo strategy has very little cost in the expectation time compared to the one-step strategy. The gap between it and the 1 port approach only approaches a maximum of one step however the effect is more significant when considering the minimum number of ancilla required to secure $P(n \leq N) \geq 0.999$ but the relative effect is diluted as the coupling strength gets weaker.

5.1.5 Asymmetric transmitter/receiver interaction gates

Up until now the focus has been on each side of the transmission interacting with the ancilla in the same way, as per the original ADQC paradigm. However, should the two

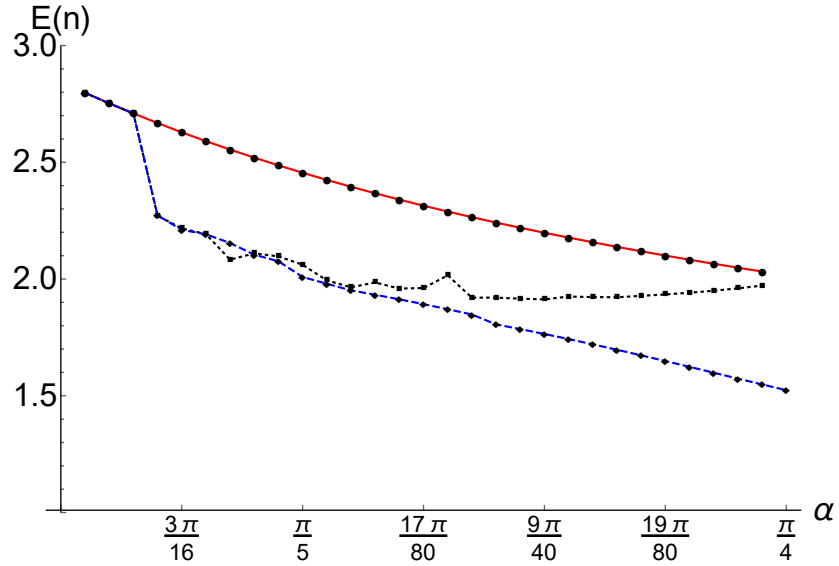


Figure 5.7: A comparison of strategies for different numbers of degrees of freedom at high coupling strengths. The solid red curve represents the 1 port/1 d.o.f. approach, the dotted black curve is the 2 port/ 1 d.o.f. strategy and the dashed blue curve is the 2 port/2 d.o.f. strategy. The threshold occurs at approximately $0.73\frac{\pi}{4}$.

qubits not be part of the same register nor distributed nodes of the same design but of two distinct devices of different physical implementations, they may not have the same interaction gate available.

One might consider that different quantum computers at different locations are being networked to share computational capability. For example, there is Blind Quantum computation [161] where a central server can carry out computations on behalf of a client without learning the client's inputs. It may be that one system is transmitting to a variety of receivers without knowledge of its system parameters. We will look at a some potential scenarios for a small starting investigation into how ADQC fits into such a case.

The suboptimal receiver

Consider that the interaction gate of the transmitter is now characterised by $e^{i\gamma\sigma_z\otimes\sigma_z}$ while the receiver's gate is still $e^{i\alpha\sigma_z\otimes\sigma_z}$. Control of the ancilla remains the same so the transmission can still be parametrised by β but with the range $0 \leq \beta \leq \gamma$ where γ is not equal to α .

If $\gamma < \alpha$ then this disrupts the ability to perform the previously considered strategies since some values of Φ in the range $[-\pi, \pi]$ will no longer be possible. For example, the one-step strategy would be categorically impossible because success on the first step

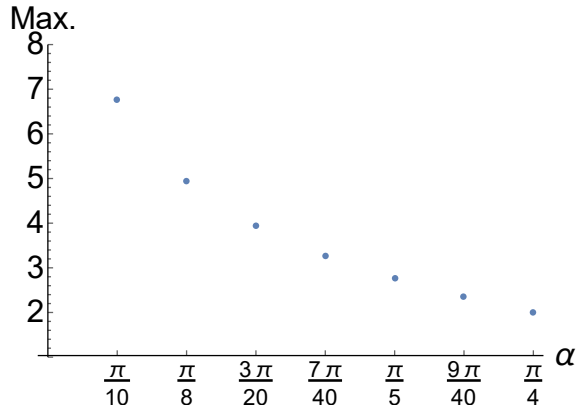


Figure 5.8: The maximum number of steps in a strategy as allowed by the hard limit of $\frac{1}{p}$ for a given coupling, displayed over the top half of the range of coupling strengths.

relies upon $\beta = \alpha$. Attempting to replicate the strategy would take multiple steps for a success condition and as β goes to lower values, the probabilities drop.

On the other hand, looking at figure 4.18 in section 4.2, when $\beta > \alpha$, the same $|\Phi\rangle$ can be achieved with greater probability. The relative signs of Φ_0 and Φ_1 also changes so that Φ_0 is no longer a backwards action to Φ_1 and now reduces the magnitude of $|\Phi\rangle$ needed in the next step. The positions of the random walk are no longer bound and since lower $|\Phi\rangle$ can now also be generated by greater values of β , this increases the probability of success in the next step and the positions of the random walk are no longer bound. The walk will tend to and then reverberate around the target.

To examine the potential of this advantage, the two parameter optimisation of the one-step strategy for $\gamma = \alpha$ was compared to the same for γ taken up to the limit of $\frac{\pi}{4}$; see figure 5.9.

The deterministic gate

Say that a transmitter device is optimally designed for such a task so it can produce an ancilla fully coupled to its register qubit. The exact process may vary but in our model it is reflected by having a first interaction gate locally equivalent to $e^{i\frac{\pi}{4}\sigma_z \otimes \sigma_z}$ with a $|\pm\rangle$ ancilla preparation. It then transmits the ancilla to another device which was not designed with optimal interactions with this ancilla in mind. One might consider almost any interaction gate but under the limits of our present model, we can examine the variability of the second gate with the parameter α in the gate $e^{i\alpha\sigma_z \otimes \sigma_z}$.

Recall from figure 4.18 that as β is increased to $\frac{\pi}{4}$ both entangling outputs Φ match. A perfect transmitter returns some of the behaviour of the two-qubit gate in the original ADQC model: deterministic entangling power with probabilistic local gate Pauli post-corrections. The gate that is generated has the same entangling power as the second

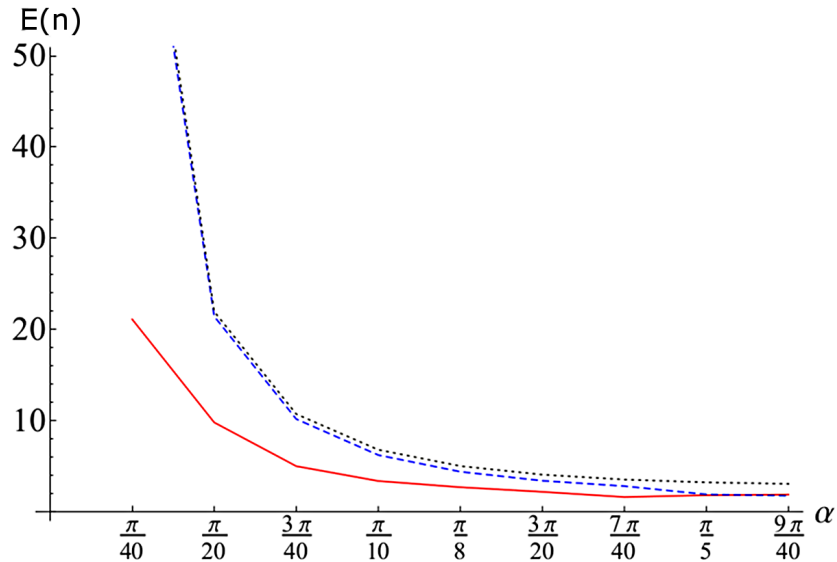


Figure 5.9: Plot of the expectation times of the one-step strategy with suboptimal receiver against coupling strength (the solid red curve), alongside the prior results of the case with equal transmitter and receiver (dashed blue curve) and flip-undo strategy (dotted black).

gate though the sign of Φ is dependent on the intermediate gate.

This is very useful when the interaction gate power is one of certain rational fraction of the target gate. Several deterministic gate generations will quickly hit the target if it is $\frac{1}{2}$ or $\frac{1}{5}$ of the target gate and the ability to apply the inverse provides the ability to efficiently select other targets in the group of multiples of the interaction gate.

On the other hand consider that the interaction gate provides a multiple of $1 - \frac{1}{n}$ for a very large rational n or possibly even some irrational multiple close to the target of a $1 - \epsilon$ multiple. To take $\pi \cdot (1 - \frac{1}{n})$ to π under modulo 2π would take n steps. If the distance from the target is irrational then one must use the technique of $\Theta(\frac{1}{\epsilon})$ applications of the gate to reach an approximation of the target. When the distance of the generated gate to the target is smaller than the generated gate, the time to implement the target gate is larger. The primary issue is not the magnitude of α but how it fits into a walk on the group that contains the target gate.

However, by reintroducing a degree of control, this can be improved upon. A $e^{i\frac{\pi}{4}\sigma_z \otimes \sigma_z}$ transmitter interaction gate results in intermediate ancilla states in the computational basis states, $|a_j\rangle = |j\rangle$ (or the bit flip thereof). The gate generation must produce a deterministic value of Φ because the final four ancilla states $|a_{ij}\rangle$ are all going to lie on a great circle with a hemispherical cap and thus the four points are symmetric with respect to an exchange of the measurement states at the centre of the caps. This applies even if the greater circle is titled by an elevation angle to the vertical. If they

were at the equator, no entangling power is expected (refer back to section 4.2 for more details). As one would then expect, as we can see looking back at the figures in section 4.2.2 and as we isolate the case when $\beta = \frac{\pi}{4}$ to create figure 5.10, one can apply a rotation in the intermediate stage, a tilt of the plane of the intermediate ancilla states (see figure 5.11) to reduce the strength of the deterministic gate result.

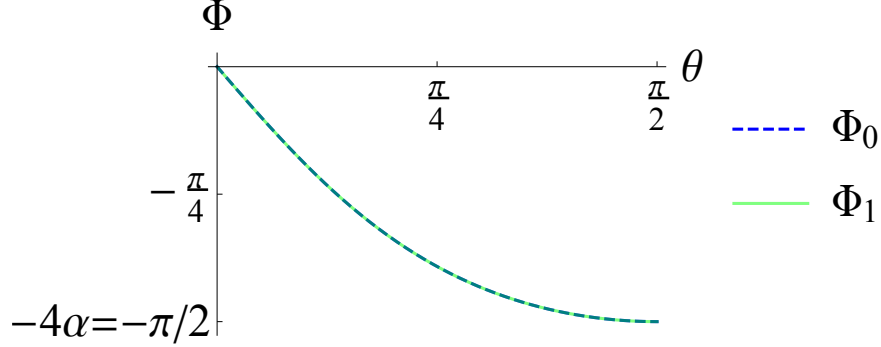


Figure 5.10: The entangling power of a two-qubit gate generation, Φ , against the tilt of the intermediate states, θ , when the transmission gate interaction is fixed as $\frac{\pi}{4}$, using an example receiver interaction gate strength $\alpha = \frac{\pi}{8}$. The output is the same for both measurement results but the entangling strength can be varied by varying the size of the tilt.

Now if the maximal deterministic gate was applied and the gap between the current gate product and the target gate is less than the interaction gate, it can be reached exactly in the next step. So a CZ equivalent can be implemented in $\lceil \frac{\pi}{4\alpha} \rceil + 1$ steps.

5.1.6 Concluding commentary

By focusing on one specific aspect of ADQC, the generation of two-qubit entangling gates, we were able to test some of the features and resource trade offs involved in performing ADQC with arbitrary coupling strength interaction gates using stochastic simulation of circuit gates, some of which may be used to guide further study into the stochastic generation of any single qubit gate. Here, we were particularly focused on how the control of the preparation and the measurement of the ancilla traded off on each other, each being associated with control over a single continuous variable.

If there was control over only one of the parameters, it did not matter which variable was controlled, they both implemented the chosen strategy with the same time distributions. When two parameters could be controlled, there was a threshold effect- it only produced an improvement beyond a particular coupling strength. This threshold effect is a result of the combination of the one-step strategy and the particular curves of the preparation and measurement basis relationship with the entangling power out-

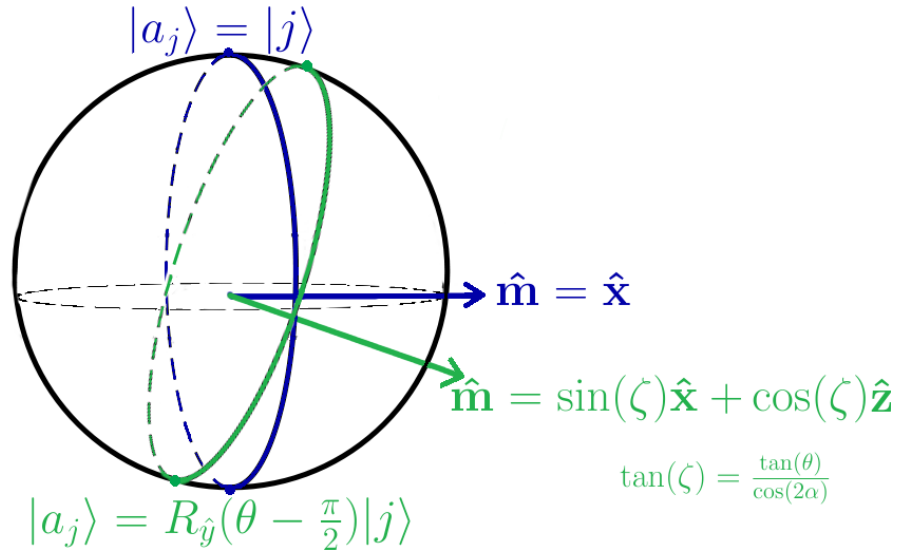


Figure 5.11: The tilting of the measurement basis of the two-qubit gate generation when using a maximally entangling primary interaction gate, pictured on the Bloch sphere.

put. It is a peculiarity of the two-qubit entangling gate problem since the single qubit gate generation only has a single real parameter and can not guarantee a probability of hitting a target in a fixed number of steps.

The approach required the use of some strategy for which the parameters could be optimised. The one-step strategy in which the best possible probability of success in the immediate next step is taken was a viable deduction from earlier examination of strategies. We did not find a numerical comparison of the one-step strategy to optimisation over a greater number of steps but the behaviour of the probability in individual steps suggests that any such advantage given by an increase in the number of steps is minimal to non-existent for higher coupling strength interaction gates.

However the advantage given by full control of ancilla parameters can not be attributed all of the speed up over unregulated gate generation. For a limited range of coupling values, just including both possible results in a generation as candidates for optimisation produced an equal advantage as including both ancilla parameters for optimisation over one. Also, when the control of the ancilla parameters was limited to adding a bit flip or not in the flip-undo strategy, it cost only an additional one step on average. The loss of control over the ancilla parameters was countered by the ability to limit the range of possible results and the strategy only required a record of the current product of the generations out of three possible options and no knowledge of the coupling parameter α nor the history of results.

By interpreting this step of ADQC as a potential long range operation between

distant nodes in a network, it opened up the possibility of having the ancilla undergo different operations with the sender and receiver nodes. If the transmitter interaction gate is weaker than the receiver then we may be forced into a multi-step strategy. If it were stronger, then the trade off between transmitter control and receiver control is no longer equal- it is advantageous to have the transmitter control the ancilla parameter if only one can do so. If both do so then this means that the stronger interaction gate is an additional resource that reduces the time needed. If the transmitter is a maximally entangling gate then ADQC provides a strategy for a suboptimal receiver to adapt its operation to establish a maximally entangling long distance operation more quickly.

Further research may explore the avenue of using the ancilla driven model and description of interaction gates as a way of characterising the suboptimal generation of entanglement over long distances. ADQC had its best advantage over naive repetition when interaction gates were just of suboptimal coupling strength.

5.2 Minimum Control Ancilla Driven Universal Quantum Computation

We have seen how, in our stochastic approach to implementing a universal quantum computer, a use of control over the parameters of the initial state and measurement basis of the ancilla can be used to improve the expected hitting time and spread of times.

Conversely then, one expects that applying no feedback or strategy into the ancilla parameters would result in the longest implementation times. We can look at this case as an extreme in the trade off of resources where we look at the cost of reducing the parameter control to its minimum. To do this requires removing several capabilities.

Even if there is no feedback, different choices of $|a\rangle$ and $\{|m\rangle\}$ may have different implementation times for different targets. There is also control of the unitary conditions when changing between single-qubit and two-qubit gate generations; we have to consider how the intermediate gate U_a is to be applied and if it has to adapt to local unitary effects on the ancilla in the decomposition of the interaction gate.

If every ancilla had a random preparation state and random measurement basis, there would be no way of assuring the operation enacted fulfilled the unitary conditions. If one is lucky to have a system which produces random ancilla states that all fit into the unitary conditions for some generation and then if one had at least the option to choose to adapt the register to implement a single- or two-qubit gate depending on the ancilla, then the register operations would perform a random walk through the entire group of two-qubit unitary gates. The problem is that this would require a way

of knowing what ancilla state was produced and reacting to potentially your fastest component system¹. So it has to be assumed that the parameters can at least be fixed.

The actual practical minimum unitary control would be the case where there is only the application of a single fundamental interaction gate E , whose local unitary effects match unitary conditions for both single- and two-qubit generations with fixed ancilla preparation and measurement. The only control would be whether it is a one- or two-qubit generation.

There is a possible practical and a foundational benefit to considering this mode of implementation. First, we note that if the ancilla parameters can be fixed one may benefit from the engineering of many ancillae with greater precision. Another natural trade off is that between the ability to perform many different operations and to perform an operation well. The ancilla driving process can be seen as the inner workings of a quantum gate and the minimum control mode makes a single gate the universal set and this one gate just has to be performed really well. The second benefit is that this idea can then be used to compare the interaction gate E to other solo-universal gates. Some of the first proofs of universality involved a single quantum gate operation: the three-qubit Deutsch gate [22] was a generalisation of the classically universal Toffoli gate; later, gates that could be applied to only two qubits were shown to be universal by Barenco [29] and Sleator & Weinfurter [30] by showing that they could be used to recreate the Deutsch gate by successive applications to different pairs, with a minimum of five applications [78]. One could create single-qubit or two-qubit gates by using ancillary qubits and generating the subgroups $U_2 \otimes \mathbb{I}_4$ and $U_4 \otimes \mathbb{I}_2$. DiVincenzo [77] showed that it was possible to decompose any unitary operation into two-qubit operations on different pairs and that the ability to perform any arbitrary two-qubit unitary was universal. Deutsch *et al* [31] and Lloyd [79] showed that in fact almost any two-qubit gate is universal for two-qubit unitary gates; all those that aren't have been characterised [80] though those which are not universal with ancilla have not been completely characterised. A key part of using a single fixed gate is that one should be able to swap the qubits with respect to the interaction and then rely upon a lack of SWAP symmetry in the two qubit gate interaction. One would not expect to be able to create a gate that decomposes into different local single-qubit gates if one did not have this swapping ability and swap asymmetry.

On the other hand, being focused on the hybrid paradigm, ancilla schemes do not allow nor require capabilities assumed in previous proofs of universality of single unitaries such as swapping the qubits with respect to the interaction or being able to apply interactions between any pair of qubits. So this motivates looking at the capability of weaker interaction gates to minimise control and see if we can find cases

¹and a massive redefinition of the word “lucky”.

where the interaction gate lacks a swap asymmetry.

The problem is the difficulty in aligning the single-qubit and two-qubit unitary conditions. We will have to find the restrictions that allow for both processes to be unitary and then see what gate generations are possible under them. The conditions for universal quantum computation then have to be considered. For single-qubit gates, this will depend mostly on the local unitary effects of the interaction gate on the register. Since the non-local effects are always symmetric under the Cartan decomposition, looking at these local effects will also determine the swap symmetry of the interaction gate.

5.2.1 Fulfilling both sets of unitary conditions simultaneously

If the interaction gate consisted of only its non-local component, the single-qubit generation requires that the preparation state $|a\rangle$ and measurement basis $\{|m\rangle\}$ form a plane that includes the primary axis of the interaction gate (treated as \hat{z}) in the Bloch sphere. The choice of plane does not affect the result generated so it is always treated as \hat{x} and the results are controlled by the relative angle between the preparation and measurement state given by their elevations.

For the two-qubit generation, it is required that there is an operation in between the two uses of the interaction gate of the form $U_a = R_{\hat{y}}(\theta - \theta'')R_{\hat{n}}(\frac{\pi}{2})$. Any effects that would rotate the plane of measurement and preparation before and after the interaction would also affect this operation but the two-qubit generation is also invariant under rotations about the primary axis and so it can be treated as being prepared and measured in the $\hat{x} - \hat{z}$ plane with \hat{n} in that plane.

However since in this scenario, there is no control over the ancilla in the intermediate stage, local effects on the ancilla can not be considered compensated for by control over the operation U_a . Instead, as we see in figure 5.12, it must be that there are local pre- and post-interaction effects on the ancilla, U_A and V_A , that combine to make

$$U_a = U_A V_A. \quad (5.1)$$

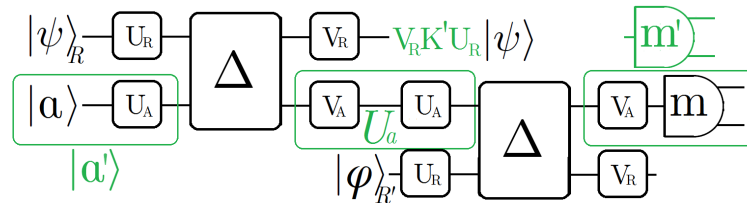


Figure 5.12: Circuit diagram of the Cartan decomposition of the two-qubit gate generator in the minimal control scheme.

These local effects will now have to be considered separately from the ancilla state preparation and measurement. This raises a problem because the local effects, U_A and V_A may change the required $|a\rangle$ and $|m\rangle$ of the single-qubit generations while enforcing the form of (5.1) to maintain the preparation and measurement states conditions of the two-qubit generation. To deal with this, we will treat $|a\rangle$ and $|m\rangle$ as already being fixed in the $\hat{\mathbf{x}} - \hat{\mathbf{z}}$ plane and then see what parameters for the local effects will maintain this property. Then from these parameters, the results of the generations can be calculated.

In order to preserve the vertical plane of the preparation and measurement basis for the single-qubit case, only certain actions can be taken: rotations about the axis perpendicular to the plane, $R_{\hat{y}}(\gamma)$, rotations about the axis of preparation/measurement itself, $R_{\hat{a}/\hat{m}}(\gamma)$, and rotations by π about an axis in the $\hat{\mathbf{x}} - \hat{\mathbf{z}}$ plane, $R_{\hat{x}-\hat{z}}(\pi)$.

For the two-qubit case, U_a is characterised by parameters θ and θ'' (recall this is the elevation of the vector $\hat{\mathbf{n}}$) which correspond to a specific measurement basis and preparation state respectively. For a fixed U_a , there is only one appropriate measurement basis and preparation state so the local unitary effects must leave them unchanged. This allows only the rotations about the preparation and measurement states $R_{\hat{a}/\hat{m}}(\gamma)$ and, if one can account for an exchange of the results corresponding to the measurement basis, rotations about the perpendicular axis but only by π , $R_{\hat{y}}(\pi)$:

$$U_a = Y^k R_{\hat{a}}(\delta) Y^{k'} R_{\hat{m}}(\gamma) Y^{k''} = R_{\hat{y}}(\theta - \theta'') R_{\hat{n}}\left(\frac{\pi}{2}\right), \quad (5.2)$$

$k = 0, 1$. While we have not checked by a decomposition of both forms of U_a to make an exhaustive comparison to see whether there are any other values for which rotations about $\hat{\mathbf{a}}$ and $\hat{\mathbf{m}}$ will exactly match the parameters of θ and $\hat{\mathbf{n}}$ that the axes of rotation require, we assume an absence of any case which is not a result of setting $\theta - \theta'' = 0$.

When $\theta - \theta'' = 0$, recall that from (4.98), $\hat{\mathbf{m}} = \hat{\mathbf{a}}$ so

$$U_a = R_{\hat{a}}(\delta + \gamma) = R_{\hat{n}}\left(\frac{\pi}{2}\right) \quad (5.3)$$

but since $\hat{\mathbf{n}} \neq \hat{\mathbf{a}}$ except when prepared on the equator of the Bloch sphere as per (4.87), the ancilla is prepared and measured in the $\{|\pm\rangle\}$ basis. $U_a = R_{\hat{x}}\left(\frac{\pi}{2}\right)$ with some possible Pauli operator corrections.

$$U_A = X_i R_{\hat{x}}(\delta) X_{i'}, \quad (5.4)$$

$$V_A = X_j R_{\hat{x}}(\gamma) X_{j'}, \quad (5.5)$$

$$U_a = X_k R_{\hat{x}}\left(\frac{\pi}{2}\right) X_{k'}. \quad (5.6)$$

Any Pauli operations before U_A or after V_A can just be commuted through with a possible sign change for δ or γ .

5.2.2 Single-qubit gate approximation and an entangling gate

The conditions for minimal control over the ancilla for both single- and two-qubit gate generation require an ancilla preparation state in the measurement basis of equal superposition states. We know from the previous result (4.37) that this will generate identity or σ_z on a single register qubit up to local unitary effects so that the two results are

$$U_+ = V_R U_R, \quad (5.7)$$

$$U_- = V_R \sigma_z U_R. \quad (5.8)$$

Similarly it has been shown that the results of the two-qubit gate generation will be local equivalents to $e^{i(\phi+\frac{\pi}{4})\sigma_z \otimes \sigma_z}$ and $e^{-i\frac{\pi}{4}\sigma_z \otimes \sigma_z}$ where $\tan(\phi) = \frac{-1}{\cos(2\alpha)}$.

If the results U_+ and U_- provide a universal gate set for single-qubit gates then they can be used for the non-deterministic approximation of single-qubit gates and they can correct any local unitary effects in between two-qubit gate generations so that there can be a non-deterministic approximation of a target two-qubit unitary gate by random walk on a 1-d curve.

Example with an asymmetric interaction gate

For example, consider a case that mirrors the use of a control unitary with Hadamard gates local effects as the interaction gates. There is a simplification since then local effects can be considered entirely post effects with $V_A = U_a$. That then dictates that $V_A = R_{\hat{x}}(\frac{\pi}{2})$ up to a possible additional Pauli operation which allows the comparison to be extended by using the gate H_y as defined by (4.9). By commuting through rotations about the \hat{z} axis, this can be rewritten as $|a\rangle = |+\ i\rangle$, $\{|m\rangle\} = \{|\pm\ i\rangle\}$ and $U_a = H$. We can now make an example that mirrors the original ADQC interaction gate. $E = (H_A \otimes H_R).C(\sigma_z^{\frac{1}{2}})$:

$$C(\sigma_z^{\frac{1}{2}}) = \text{diag}(1, 1, e^{-i\frac{\pi}{4}}, e^{-i\frac{\pi}{4}}). \quad (5.9)$$

The eigenstate basis single-qubit generations are identity and $\sigma_z^{\frac{1}{2}} \equiv \text{diag}(e^{-i\frac{\pi}{4}}, e^{i\frac{\pi}{4}})$, following from the rule of linear addition:

$$K_{+i} = \frac{1}{2} \left(\mathbb{I} + \sigma_z^{\frac{1}{2}} \right) = \cos\left(\frac{\pi}{8}\right) \sigma_z^{\frac{1}{4}}, \quad (5.10)$$

$$K_{-i} = \frac{1}{2} \left(\mathbb{I} - \sigma_z^{\frac{1}{2}} \right) = i \sin\left(\frac{\pi}{8}\right) Z. \sigma_z^{\frac{1}{4}}. \quad (5.11)$$

Including the local Hadamard gate effects, the normalised unitary gates are $X^j H \sigma_z^{\frac{1}{4}}$. If the gate was $(H \otimes H).C(\sigma_z^{\frac{1}{2^n}})$, the resulting gate would be $X^j H \sigma_z^{\frac{1}{2^n}}$. The two potential actions may form a universal set for single-qubit gates depending on the value of n .

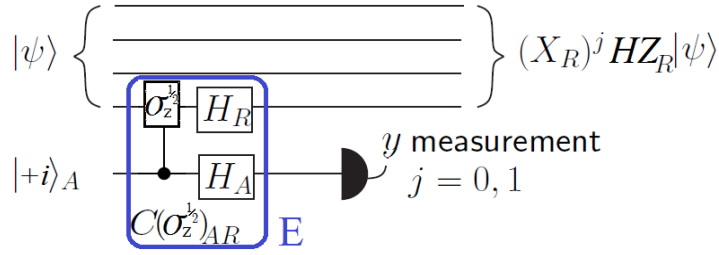


Figure 5.13: The single step on a single qubit in the minimal control scheme; it follows the same process as ancilla driven quantum computation and the same limitation to one interaction per qubit per ancilla but with a different preparation and measurement basis.

With $n = 2$, two applications of the process will generate either $H\sigma_z^{\frac{1}{4}}H\sigma_z^{\frac{1}{4}} = \sigma_x^{\frac{1}{4}}\sigma_z^{\frac{1}{4}}$ or $\sigma_x^{-\frac{1}{4}}\sigma_z^{\frac{1}{4}}$ up to some Pauli operator correction that will affect future pairs of operations and we can use the proof from Boykin *et al* [34] used in section 4.4.1.

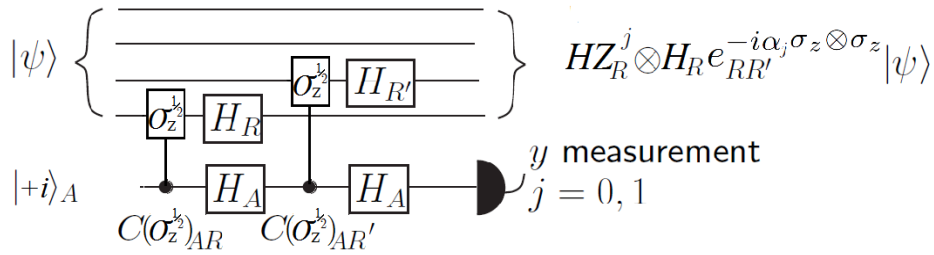


Figure 5.14: The schematic for performing a two-qubit gate under minimal control. An identical interaction with a second register qubit is turned on but all other processes remain the same.

For the two-qubit gate generation, the $+i, -i$ results will correspond to the generation of the unitary operations $e^{i(\phi+\frac{\pi}{4})\sigma_z\otimes\sigma_z}$ and $e^{-i\frac{\pi}{4}\sigma_z\otimes\sigma_z}$ where $\tan(\phi) = \frac{-1}{\cos(\frac{\pi}{4})} = -\sqrt{2}$. ϕ is an irrational multiple of π , provable from the theorem that $\frac{1}{\pi}\arccos(\frac{1}{\sqrt{n}})$ is irrational for odd $n \geq 3$ [154]. Again, a target unitary can be chosen such as $e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z}$, equivalent to CZ, and the process repeated until the product of repeated generations approaches within a distance of the target.

The decomposition of the interaction gate can be written as a symmetric non-local action $e^{-i\frac{\pi}{8}\sigma_z\otimes\sigma_z}$ and an asymmetric pair of local actions on each subsystem. This asymmetry is key to the result. When the gate has been prepared in a Pauli operator eigenstate, this will mean that the local gate actions on the ancilla after the symmetric

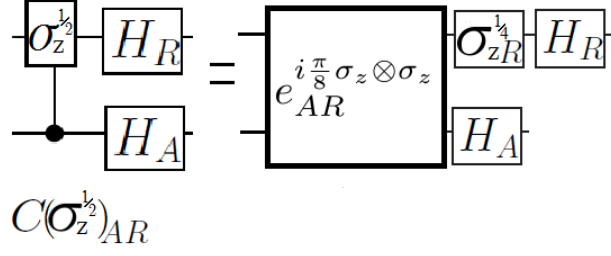


Figure 5.15: The Cartan decomposition of the gate in the minimal control scheme reveals that it is asymmetric in its local gate actions by a $\sigma_z^{\frac{1}{4}}$.

interaction will be Clifford group single-qubit gates. However, universality for the single-qubit gates requires that there is a non-Clifford gate acting on the register qubit side, making an asymmetry in the gate construction. It also means that the process includes a universal gate set for direct register manipulation. Another possibility is that since the ancilla is being prepared and measured in an eigenstate of a Pauli operator, this process could be performed with a two parameter gate with a component $e^{i\alpha_y \sigma_y \otimes \sigma_y}$ whose effect will be reduced to additional local gate effects $R_{\hat{y}}(\pm 2\alpha_y)$. What we will demonstrate in the following section is that it is also possible to use a one parameter symmetric interaction with Clifford group local gates.

Ancilla-Register symmetry in the interaction gate

To consider the use of an interaction gate with swap symmetry, the local register effects U_R, V_R must be set equal to U_A, V_A respectively. From (5.7), the results of the single-qubit generation will be

$$U_+ = X_j R_{\hat{x}}(\gamma) X_{j'} X_i R_{\hat{x}}(\delta) X_{i'}, \quad (5.12)$$

$$U_- = X_j R_{\hat{x}}(\gamma) X_{j'} Z X_i R_{\hat{x}}(\delta) X_{i'}. \quad (5.13)$$

To focus on discriminating between the two results, ignore X_j and $X_{i'}$ and account for those central Pauli terms with a single gate X_m :

$$U_{\pm} = U_m = R_{\hat{x}}(\gamma) X_m R_{\hat{x}}(\delta). \quad (5.14)$$

X_m can be a result from two possible pairs $\{\mathbb{I}, Z\}$ or $\{X, Y\}$. \mathbb{I} and X will both commute with $R_{\hat{x}}(\gamma)$ while Z and Y will commute through with a change in sign, $\gamma \rightarrow -\gamma$. So the results for either set will be

$$U_m = X_m R_{\hat{x}}(\delta \pm \gamma). \quad (5.15)$$

One of the results is $R_{\hat{x}}(\frac{\pi}{2})$ with some possible Pauli pre- and post-corrections and thus a member of the Clifford group. The other will depend on the difference $\delta - \gamma$. For

this to be universal for single-qubit gates, the second result just needs to be outside of the Clifford group and then have some n for which U_m^n is equal to or approximates a member of the Clifford group that forms a distinct generator with the first result.

A successful example would be $ZR_{\hat{x}}(\frac{\pi}{2})$ and $R_{\hat{x}}(\delta - \gamma)$ for $\delta - \gamma = \frac{\pi}{4}$, similar to the $\{H, T\}$ gate set. On the other hand, if either δ or γ were zero i.e. local effects are confined to pre- or post-interaction gate as in the $(H \otimes H)C(\sigma_z^{\frac{1}{2}})$ example then the results can only generate the single-qubit Clifford group and will not produce a universal gate set.

5.2.3 Replicating other universal gate sets

Given some symmetric ancilla-register gates, we can not use the universality of a fixed two-qubit entangling gate with arbitrary single-qubit gates to demonstrate the universality of a classical control ancilla scheme as with other schemes. However it is still possible to achieve universality by constructing other finite gate sets on the register.

If the local qubit gate U_a and the σ_z gate only generate a finite group then there is at least the capacity to generate U_a^\dagger . This can also be done exactly and most likely more quickly than approximating arbitrary gates through random walks through a three dimensional continuous space. This also means that the local gate corrections can be performed that will allow the random walk through one parameter entangling gates by the gates $e^{-i(\phi + \frac{\pi}{4})\sigma_z \otimes \sigma_z}$ and $e^{-i\frac{\pi}{4}\sigma_z \otimes \sigma_z}$ ($\tan(\phi) = \frac{-1}{\cos(2\alpha)}$) until it has reached an appropriate target. The chosen target does not have to be CZ equivalent which means that one is free to consider generating a universal gate set on the register which includes only single-qubit gates of a finite group. For example, there is Shor [82]'s set $\{C(\sigma_x), Z^{\frac{1}{2}}, H\}$ and Kitaev[85]'s set $\{C(Z^{\frac{1}{2}}), H\}$ (known to be equivalent to Shor's basis [34]). Generally though, by having the ability to select a single-qubit gate out of the finite set, one can apply them asymmetrically to register qubits along with the entangling gate to create an asymmetric gate that fulfils the conditions in section 2.3.2 and is mono-universal.

For a specific example consider $(H \otimes H).e^{i\frac{\pi}{8}\sigma_z \otimes \sigma_z}$ with preparation and measurement of the ancilla in the states $\{|\pm i\rangle\}$ which is equivalent to our earlier example with the asymmetry removed. The single-qubit generations will cover the group generated by $\{H, HZ\}$ (see figure 5.16) while the two-qubit generations will be able to approximate any one parameter entangling gate to arbitrary error.

In fact, if the register qubits were all initialised in the state $|0\rangle$ then there is enough local unitary control to place them in either $|0\rangle$ or $|1\rangle$. Then using a secondary register qubit and generating an interaction gate $e^{i\gamma\sigma_z \otimes \sigma_z}$ can substitute for generating single-qubit unitary gates $J(\gamma) = H.e^{i\frac{\gamma}{2}\sigma_z}$. This could be used either with just the result $\gamma = \phi + \frac{\pi}{4}$ created from one use of the minimum control two-qubit gate generation and

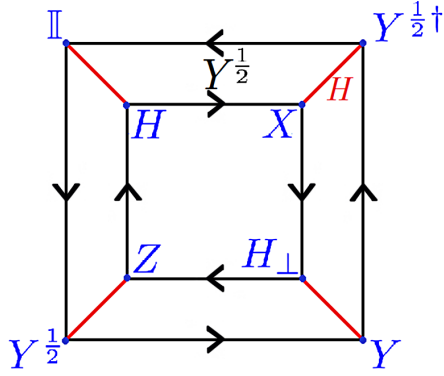


Figure 5.16: Graph of the group generated by H and $HZ = Y^{\frac{1}{2}}$ up to global phases. The eight group members are represented by the blue vertices with each edge representing the gate operation that moves from one to another. The red diagonal edges are the H operation and the black straight lines the HZ .

then making a decomposition in terms of $J(\pm 2\gamma) = H.e^{i\pm\gamma\sigma_z}$ as described in section 4.4.1 (using also the occasional H correction and/or feed-through of a Pauli correction) or even using the ancilla multiple times to try to create a specific $e^{i\alpha\sigma_z\otimes\sigma_z}$ of any coupling parameter α in order to be able to generate the entire class $J(\alpha)$.

This forms another trade off in resources with on the one hand, the differing circuit depths or time costs of the choice targeted gate set against on the other, the addition of secondary register qubits to act as a pseudo-ancilla system. This trade off will be mediated by the actual time to implement the members of the finite gate set. For our example set, we can find a computed expected time to implement different members of the group (figure 5.17). The small order of the group compared to all single-qubit unitaries makes for a short expected time for a target gate and short mixing time. In the future, analytical answers on how different groups may behave may possibly be found in the further study of random walks on finite groups as done with \mathbb{Z}_\times and S_n groups [163, 164, 165, 166].

5.2.4 Concluding Commentary

We have demonstrated the possibility of creating a minimal control scheme for ancilla driven quantum computation where “minimal control” is interpreted as the need to control only the number of times an ancilla interaction with the register occurs with ancilla preparation state, measurement basis and interaction gate being fixed.

Because of the restrictions necessary for all conditions to match for both the single-qubit and two-qubit gate generation, the preparation and measurement of the ancilla must occur in the same basis. This limits the random component of the single-qubit

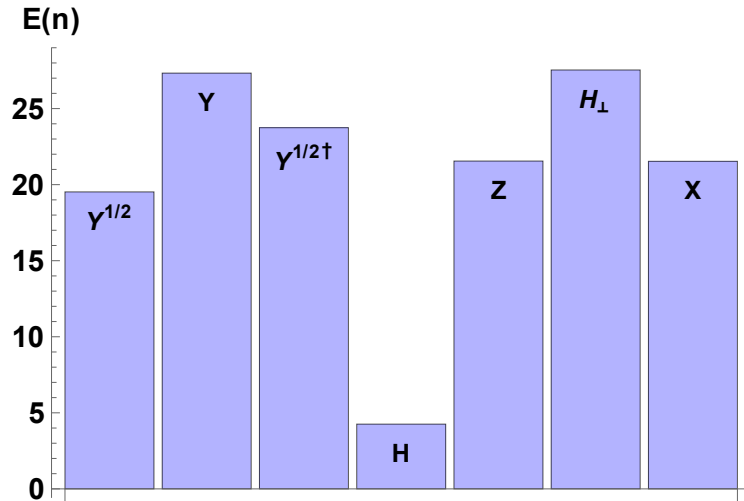


Figure 5.17: The expected number of steps to reach elements of the group generated by H and $HZ = Y^{\frac{1}{2}}$, calculated from 10000 simulations with $\alpha = \frac{\pi}{8}$ ($p(H) \approx 0.15$).

gate generation to a Pauli operator and so the ability to create any arbitrary single-qubit gate from single-qubit generations alone will depend entirely on the local effects on the register available.

If we consider the swap symmetry of the register and restrict the local effects on the register to equal those on the ancilla, the scheme will generate only a finite group of single-qubit gates. However if that is combined with the two-qubit gate generation to nullify the local gate effects in the latter case so that it can perform a walk over a 1d space of two-qubit gates to select an additional gate, it will form a universal gate set with the local gate generation results.

This opens up the possibility that this may be a faster way to reach an arbitrary unitary gate target; a question that may be cleared up by further investigation into the behaviour of random walks on finite group graphs. In this scenario, the model would be aiming to achieve a previously discussed upper bound on random behaviour— finding the time to generate gates in a finite set and then multiplying the expectation time by the number of gates from that set in a decomposition for arbitrary unitary gates. Having a far smaller time to hit a target gate in a finite group may make the product of this time with the number of gates in a decomposition less than the actual average time to generate any arbitrary single-qubit gate.

Minimal control may be best suited to a system where qubits of the same type as the register act as a pseudo-ancilla. An ancilla and register of the same implementation may be a cause of symmetry in the interaction gate. Since we have associated register qubits with stable and well isolated systems, this may also be the cause of the small coupling strength parameter. If the cause of minimal control may be the difficulty of

manipulating the register rather than the short lifetime of the qubits, then the short life of the ancilla qubits may be attributable to other factors such as the exposure that makes the interaction possible or the destructive nature of the measurements.

Chapter 6

Outlook

6.1 Further avenues of research

Beyond the results previously demonstrated, we present topics that may be addressed in future research. We start with topics which stick to the same model of ADQC with non-maximally entangling interaction gates as used in this thesis but which we did not have enough time to fully explore. The later topics will cover possible changes to the ADQC model.

6.1.1 Characterising single qubit gate walks

In the main body of this thesis, we have not fully addressed the issue of characterising the random walk through the group of single qubit gates performed by random gate generation in favour of focusing on the trade off of resources that occurs when attempting to guide a walk through a strategy. In section 4.4, a theory of the expected behaviour with a single gate target example was provided. This leaves the question of whether the behaviour can be really said to apply to every gate target and if this fits our suggested theory. For a single target, the distribution of hitting times can be more appropriately analysed if several possible distribution models are considered but it may be more efficient to look at the underlying behaviour by simulating the distribution of locations the random walk visits.

In keeping with the study of guided strategies in generating two qubit gates, there is also the issue of how one can guide the walk through single qubit gates to effect a trade off of resources. This is a more difficult question than the two qubit gate because the single qubit gates can not be restricted to a one dimensional curve where the multiplication of gates maps to a simple addition of the dimension parameter. Instead the single qubit gate group maps to a 3-sphere: the set of points of equal distance from a central point in 4d space (analogous to the circle or sphere in lower dimensions) and the motion is not analogous to a simple left/right or up/down translation which

becomes a problem when trying to incorporate where the gate will be upon a failure during the strategy and what sort of “distance” from the target that makes.

We can consider it as a decomposition problem- given $\{U_j\}$ results, one has an n length decomposition for a gate $\prod_{t=1}^n U_{j_t}$ for which $|U_T - \prod_{t=1}^n U_{j_t}| \leq \epsilon$ for some error bound ϵ but then there may be multiple decomposition strings with different n and different probabilities of success. After a failure to get a particular U_{j_t} result, it will then be necessary to apply a new p length decomposition which raises the question of how p relates to the initial n and the generator set $\{U_j\}$. A choice of decomposition will have to be determined based on the statistics of the probabilities of many different branches of differing lengths.

Guiding the single qubit gate walks through short chains of decomposition

In section 4.4, it was mentioned that the Solovay-Kitaev theorem may be utilised here. A decomposition could be broken into short sections which correct the previous section to bring the product into a smaller error boundary of the target. The smaller sections would be an easier optimisation problem and would have individually greater probabilities of success and the length of the subsequent sections would be limited by the Solovay-Kitaev theorem. The question that results from this is whether a failure after a short section has been implemented will be limited to either moving closer to the target or keeping the “distance” approximately the same so that the new correction section has roughly the same length.

Random exact decompositions

A number of the complications of the single qubit gate decomposition may be removed by attempting an exact decomposition, this requires being able to completely control one of the preparation or measurement parameters.

If the interaction gate has Hadamard local effects or an equivalent similar to the example gates in sections 4.4 and 5.2 then the gates generated are of the class $J(\alpha) = H.e^{-i\frac{\alpha}{2}\sigma_z}$ which provides an exact decomposition for any single qubit unitary up to a global phase $U = J(0)J(\beta)J(\gamma)J(\delta)$. If the gate is prepared (or measured) in an equal superposition then the possible results can provide all possible values for β etc. by choosing the appropriate measurement (or preparation) basis. This relies on there not being any other local effects that prevent a minimum value of 0 being generated like with the gate in figure 5.15 in section 5.2. This then limits the length of a decomposition to four components, including the length of any corrections made after a failed result, so there is an upper bound characterised by a success/failure trial that takes four time steps. The final gate $J(0)$ i.e. H can be produced from the finite group walk generated by measurement and preparation in the same plane for any coupling parameter so the

four generations can be reduced to three components and then a correction expectation time, n_H dependent on the coupling parameter.

This methods simplifies into a question of how long term a strategy is efficient. A step-at-a-time approach can look simply to create the next component of the shortest three step chain but an alternative is to look at the probabilities of succeeding on the 4th, 5th and successive steps and create a decomposition with a greater number of components as a trade off for improving the probability of success on enough possible branches to improve the expectation time and its variance.

Walking in confined spaces

One may also combine the notion of finding an upper limit to the expected time by looking at the expected time of some elements of a decomposition, mentioned in section 4.4, with the results of sections 5.1 and 5.2 that show an advantage to confining the possible range the generations may walk over, specifically the advantages of being confined to a circle or a finite group.

The decomposition of $U \in SU(2)$ can also be written as $U = R_{\hat{z}}(\beta)H.R_{\hat{z}}(\gamma)H.R_{\hat{z}}(\delta)$. The class $R_{\hat{z}}(\alpha)$ are confined to a circle and H can be found on a finite group walk. The Hadamard gate and the finite group that contains it, we know can be produced from using the $J(\alpha)$ class as above but the $R_{\hat{z}}(\alpha)$ gates would then have to be generated by generating $J(\alpha)$ and then the Hadamard correction. Having done so, one can then reproduce the one step strategy of section 5.1 but with a multiplying factor for the correction time for the Hadamard gate (minus one for the last step). Say that an angle α is generated in the one step strategy by an expectation time \hat{n}_α the expected time for the unitary target is upper limited to $(\hat{n}_\beta + \hat{n}_\gamma + \hat{n}_\delta - 2).\hat{n}_H$.

To put it in terms of our earlier context, this is equal to attempting to apply the gate $J(\beta)$, generating $J(\beta')$ instead and then making the next gate in the decomposition $J(\beta)J(\beta')^\dagger = J(\beta - \beta')H$, a gate which is not in the class $J(\alpha)$.

Since this is actually a different technique for generating the components of the decomposition, it may not be in fact an absolute upper limit but may be an improvement on strategies of different depths. One avenue to consider is how using a confined walking space is able to improve on a strategy that optimises over different numbers of steps because the confinement may eliminate some long chains in very short term strategies that contribute heavily to the variance.

Strategies that minimise communication

In chapter 4, section 4.1.3 we noted that in the special case where the measurement basis of the ancilla is the same as the basis in which the ancilla was prepared, when the measurement result is the state perpendicular to the preparation state i.e. $|m\rangle = |a_\perp\rangle$,

the single-qubit gate generated will always be equivalent to Z (plus the local gate effects on the register). On the other result, the gate generated depends on the preparation state.

Interestingly, this creates a scenario where one can attempt to generate a specific gate through measuring the ancilla in its specific initial state and then the result of failure would be limited to a Z gate. One could repeat this n times until one finally succeeds to measure the ancilla in its initial state and then the final result would be $R_{\hat{z}}(\gamma)Z^{n-1}$ which is only either $R_{\hat{z}}(\gamma)$ or $R_{\hat{z}}(\gamma + \pi)$.

This is somewhat similar to the case in the original ADQC proposal where the two gates generated upon measurement would differ only by a Z gate. The difference is that it takes also a probabilistic number of steps, n , to achieve this result. This is complicated further by the need to have local gate effects on the register that ensure universality. Most likely, these local gate effects would instead be something like a Hadamard gate so that the results would be $J(\gamma) = HR_{\hat{z}}(\gamma)$ or $J(\gamma)(HZ)^n - 1$. In this case one can not as easily commute through the Z gate and apply a sign correction in the next stage of the decomposition as in the original ADQC proposal or similarly MBQC. However the additional effects of $(HZ)^{n-1}$ are limited since it will only ever be a member of the single-qubit Clifford group which cause two effects if commuted through: a change of sign in the rotation angle γ or a change of axis of rotation to another orthogonal axis.

The above scenario may then be of interest if looking to keep the amount of information that has to be fed forward into later steps as low as possible. In the original ADQC proposal, one just had to feed forward one bit of information, the measurement result, into the next step in order to perform one possible type of correction, the sign change and it was independent of what the actual desired gate to be implemented was. In our proposal it may be possible to create strategies that only rely on sending forward between different stages two bits of information (for the H, X, Y, Z Clifford group members) and then still be independent of the target gate. This may be desirable in research relating to long distance communications such as Blind ADQC [162] or perhaps investigations into the classical computing power associated with the feed forward process such as in Browne & Anders [91].

6.1.2 Noise

The noise created by errors in the ancilla measurement was studied by Morimae & Kahn [167] for the original ADQC model to produce a relationship between the register entanglement and its state fidelity. This could be replicated for probabilistic ADQC.

Entangling gates over long distances also raises the issue of noise that may occur during the transmission and of loss of the ancilla. One could study the impact of noise

at the intermediate stage. We have seen how the unitary conditions during this step can allow for X and Z flips so if the transmission was only affect by random flips, the results would still be unitary operators enacting randomly on the register thus keeping the effects on the register a unital operation. There may be other ways in which using some knowledge of the noise of the channel could allow negative effects to be mitigated with an aim to keeping the errors of the type and within a threshold of magnitude to be ultimately corrected by quantum error-correcting codes. Network codes such as Nickerson *et al*[168]'s protocols for dealing with noisy links in a network can adapt to exploit characteristics of the noise in the links.

6.1.3 Using Generalised Measurements

Something that may be necessary to consider when operating on the ancilla to mitigate noise effects is utilising higher dimensions of the ancilla. The parameters of the interaction gate and ancilla state needed to fulfil the unitary conditions have lent themselves to qubit ancillae. However additional ancilla dimensions could be used to implement some basic error correction or enact generalised measurements that filter some other events like a lost ancilla. Once the possibility of different POVM elements acting on the ancilla is opened up, the model can be substantially changed. First from the point of view of the properties of the system an ancilla of higher dimension may be of multiple qubits or it may be of a qudit system that has to be longer lived to survive many operations or it may even be more comparable to having an additional register. Then there are the effects from the different measurement on the register.

As an example, consider the two qubit gate generation with preparation in an equal superposition. A POVM could be created where the elements correspond to a projector on the opposing state and then each of the four computational basis correlated states $|a_{ij}\rangle$.

$$M_0 = a|-\rangle\langle-|, \quad (6.1)$$

$$M_{1-4} = M_{ij} = b|a_{ij}\rangle\langle a_{ij}|, \quad (6.2)$$

$$|a_{ij}\rangle = \cos(\theta)e^{-i\gamma_{ij}}|+\rangle + \sin(\theta)e^{i\gamma_{ij}}|-\rangle. \quad (6.3)$$

Because the four points can be created symmetrically around the initial state $|+\rangle$ so that

$$\gamma_{1j} = -\gamma_{0j}, \quad (6.4)$$

$$\gamma_{i1} = \pi - \gamma_{i0}, \quad (6.5)$$

these can fulfill the condition $\sum_k M_k = \mathbb{I}$ when $b = \frac{1}{4\sin^2(\theta)}$ and $a = 1 - \frac{1}{\tan^2(\theta)}$.

The M_0 result returns the maximally entangling gate generation but since the other four results are correlated with the computational basis of the register they will project the register into separable states. Strict unitarity is abandoned in exchange for failures being confined to a finite set of separable states that could easily be returned to a ready state for attempting the process again. With a loss of information upon failure and higher dimensional ancillae, this model becomes more like approaches using Bell state measurements on ancillae but with a characterisation of the entangling process as a unitary given by parameter α . The model may even be rephrased as an entanglement distillation like the Procrustean method [169] where the entangling power α is converted into a maximally entangling operation.

Looking at the single gate generation, one can also note that it is possible to create more than a two output measurement on the symmetric plane.

Another possibility is to use multiple ancilla qubits prepared or measured in an entangled state in order to make multiple operations on the register correlate with each other upon a single ancilla measurement. For a simple example, consider an ancilla pair prepared in the state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ but kept by all operations within the two dimensional subspace of $\{|00\rangle, |11\rangle\}$ so that it behaves as a qubit with double the strength of the interaction gate.

6.1.4 Interpretations of post-selection and weak values

The interaction model for ADQC can also be used as a weak measurement system to which weak values can be applied. Our model uses gates of the class $e^{i\alpha\sigma_z\otimes\sigma_z}$, equivalent to $C(4\alpha)$, whose strongest member $e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z}$ can be used as a measurement of σ_Z on system S by preparing the probe system P in the $|+\rangle$ state and the measuring it in the $\{|\pm i\rangle\}$ basis:

$$\begin{aligned} e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z}|+\rangle_P|\psi\rangle_S &= e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z}(\alpha|+\rangle_P|0\rangle_S + \beta|+\rangle_P|1\rangle_S) \\ &= \alpha|-i\rangle_P|0\rangle_S + \beta|i\rangle_P|1\rangle_S. \end{aligned}$$

If the interaction gate is weaker, the computational basis states $|j\rangle_S$ correlate with non-orthogonal states on the probe $R_z(-1^{j+1}2\alpha)|j\rangle_S$ but the expectation value of the probe, $\langle\sigma_y\rangle_P$ can be used to calculate the expectation value of the system;

$$\langle\sigma_y\rangle_P = \langle\sigma_z\rangle_S \sin(2\alpha). \quad (6.6)$$

The parameter α is the equivalent to the spread parameter of the measurement probe; α determines the orthogonality of the probe states and a smaller α increases the number of results needed to determine $\langle\sigma_z\rangle_S$ to a high statistical accuracy.

Based on this model, the weak value for this system is

$$\langle\sigma_z\rangle_W = \frac{\langle\Psi_f|\sigma_z\otimes\Psi_i\rangle}{\langle\Psi_f|\Psi_i\rangle} \quad (6.7)$$

and then the probe evolves

$$e^{i\alpha\langle\sigma_z\rangle_W\sigma_{z,S}}|+\rangle_P. \quad (6.8)$$

If we associate the probe with a register qubit in the ADQC model and the system with the ancilla for which $|\Psi_Y\rangle = |a\rangle$, $|\Psi_f\rangle = |m\rangle$ then we can compare the above with the result (4.35) from section 4.1:

$$|\Psi\rangle_R \rightarrow e^{-i\gamma\sigma_z}|\Psi\rangle$$

$$\gamma = \arctan \left[\tan(\alpha) \frac{\langle m|\sigma_z|a\rangle}{\langle m|a\rangle} \right].$$

The weak value formalism requires that the shift of both the normal weak measurement and the weak value are both small so we would require that both α and γ are small and thus

$$\tan(\alpha) \approx \alpha,$$

$$\tan(\gamma) \approx \gamma,$$

so the weak value aligns with our results.

The condition for an imaginary or complex weak value matches the conditions for a loss of unitarity in the ADQC model as seen in table 4.2. The non-unitary operation will be of the specific form

$$e^{-\alpha\text{Im}(\langle\sigma_z\rangle_W)\sigma_{z,S}}. \quad (6.9)$$

It enacts an attenuating operation diagonal in the basis of the primary axis of the interaction gate which is always going to be orthogonal to the plane of probe states under unitary action so it changes the expectation value of a complementary value to that of the real weak value.

Interpreting weak values in two system models

It is not just that weak values can be used as an approximation for the ADQC model but that the ADQC model can also provide a description for the features of arrangements used by those working in the weak values formalism. Consider the set up in the work of Feizpour, Xing & Steinberg [136] who examined amplifying a single photon interaction through a cross-phase Kerr medium to observable levels. They considered a set up where a single photon system is coupled to light in a coherent state passing through a channel (see fig.6.1). The cross-phase shift from passing through the Kerr medium was modelled as $\exp(i\phi\hat{n}_b\hat{n}_c)$ with b and c indicated a dependence on photon path. The induced phase is then measured by interferometry between the coherent state that passed through a channel that entered the Kerr medium and one that did not.

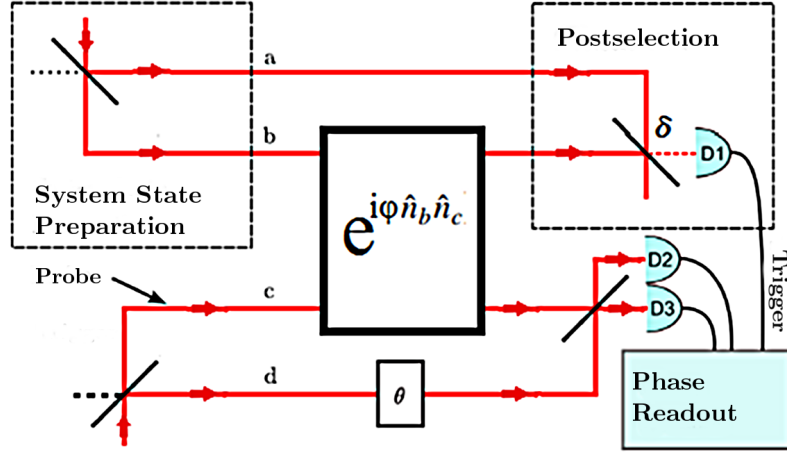


Figure 6.1: Schematic of two photon interaction apparatus from [136]. A single photon system is prepared in an equal superposition of channels a and b . A probe system is created in a superposition of channels c and d . A cross-phase-modulation interaction between channels b and c replicates a Control-Phase between the two systems. Postselection on the system is performed by an imbalanced beam splitter parametrised by δ . The probe system phase is read out by the lower interferometer with a phase shift θ used to maximise the sensitivity.

This set up and their description of the interaction can be treated as a Control-Phase gate when restricted to only single photon-single photon interactions. We consider each photon to be in a 2 dimensional state with $|0\rangle$ corresponding to a path not through the Kerr medium and $|1\rangle$ to the one that does i.e. $|0\rangle_a \equiv |a\rangle, |1\rangle_a \equiv |b\rangle, |0\rangle_p \equiv |d\rangle, |1\rangle_p \equiv |c\rangle$. With the pre-selection and post-selection this is similar to ADQC where the interaction $E = C(\phi) = |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes Z_\phi$ and what is referred to as the ancilla in ADQC equates to the measured system in WVA, prepared in $|a\rangle = |+\rangle$. The probe system of WVA corresponds to the register in ADQC.

As we have seen in previous results, the local gate differences between $C(\phi)$ and $e^{i\frac{\phi}{4}\sigma_z \otimes \sigma_z}$ shifts the unitary condition away from the plane of preparation. Given

$$|\Psi_f\rangle = v|0\rangle + \omega|1\rangle, \quad (6.10)$$

$$\begin{aligned} K_f &= \langle \Psi_f | C(\phi) | \Psi_i \rangle, \\ &= v^* \langle 0 | C(\phi) | \Psi_i \rangle + \omega^* \langle 1 | C(\phi) | \Psi_i \rangle, \\ &= v^* K_0 + \omega^* K_1 \end{aligned} \quad (6.11)$$

and for $E = C(\phi)$ and $|\Psi_i\rangle = |+\rangle$,

$$K_0 = \frac{1}{\sqrt{2}}\mathbb{I}, K_1 = Z_\phi.$$

Let $v = \cos(\frac{u}{2})$ and $\omega = \sin(\frac{u}{2})e^{iv}$ so we can write

$$\mathbf{K}_f = \frac{\cos(\frac{u}{2}) + e^{-iv}\sin(\frac{u}{2})}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & \frac{\cos(\frac{u}{2}) + e^{-i(v-\phi)}\sin(\frac{u}{2})}{\cos(\frac{u}{2}) + e^{-iv}\sin(\frac{u}{2})} \end{pmatrix}, \quad (6.12)$$

$$= \langle \Psi_f | \Psi_i \rangle \mathbf{U}_f, \quad (6.13)$$

where

$$\mathbf{U}_f = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\Phi} \end{pmatrix} \quad (6.14)$$

when

$$\cos(v) = \cos(v - \phi), \quad (6.15)$$

$$v = m\pi + \frac{\phi}{2}. \quad (6.16)$$

The unitary condition is now sensitive to the size of the phase imparted on the ancilla system by the register probe. This captures the behaviour of Feizpour *et al*'s system in response to the back action of the probe onto the system. The probe imparts $|\alpha|^2\phi$ (the strength parameter amplified by the average photon number of the probe) on the system before post-selection which would have to be compensated by a parameter $\epsilon = |\alpha|^2\phi - 2\pi n$ for integer n .

In Feizpour *et al*'s model, the post-selection is controlled by the parameter δ of the second beamsplitter of transmission t and reflectance r : $\delta = \langle \Psi_f | \Psi_i \rangle = \frac{(t-r)}{\sqrt{2}}$. This will correspond to $\cos(\frac{u}{2})$ in the ADQC model.

There is some difference in the model because the ϵ parameter implies a phase kick of 0 under modulo 2π while ADQC expects a slight compensation. This raises some questions of how the small value approximations affect the model. But on the other hand, δ contributes to a shift in the imaginary quadrature, the phase of the final coherent probe state, and the ϵ contributes to a shift in the photon number which aligns with the previously made association with the parameters of the ADQC and the phase rotation action and a dephasing action. The phase shift in the post-selection v controls the appearance of dephasing effects but it depends on ϕ which may be multiplied by a photon number n so a superposition of photon number states will have each number state discriminated by a dephasing effect and in a regime of first order approximations this may look like a weak coherent state shifted in average photon number.

In this model, one might look at how the optimised parameters for largest amplification change in between the weak value and ADQC formalisms; in the latter effects can be plotted from exact values (such as with figure 6.2), the levels of approximation in weak values may benefit from some scepticism and may be effected by the levels of computational power available in 1988.

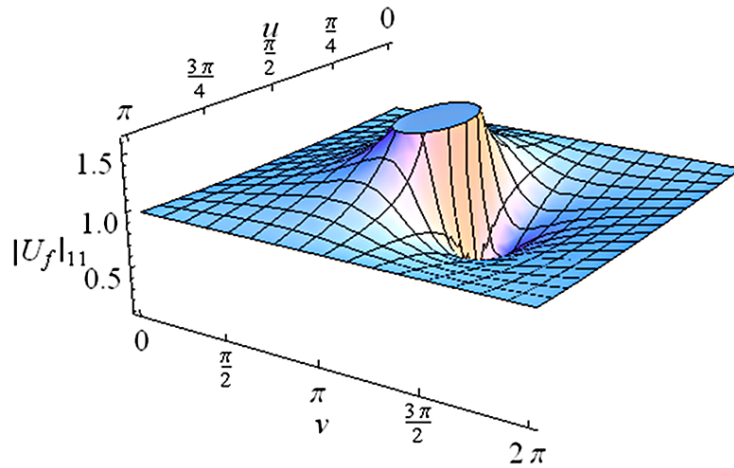


Figure 6.2: A surface plot of the absolute value of the matrix element $|U_f|_{11} \exp i\Phi$ against the parameters of the final ancilla state. For our Kraus operator to perform a phase shift, a necessary condition is that the absolute value of $|U_f|_{11}$ is 1. Here this condition is only satisfied when $v = m\pi - \frac{\phi}{2}$ (Figure uses $\phi = 0.5$).

We question whether some of the proposed uses of weak values in characterising systems could not be replaced with using finite dimensional systems like the qubit in the ADQC model. One can start by looking at what POVMs occur on the ancilla after we measure the register. Since our model provides a way of calculating what the parameters of the ancilla were from the operation enacted on the ancilla, it should be possible to reproduce the ability to characterise the wavefunction as in the work of Lundeen & Bamber[46], Lundeen *et al*[143] and Salvail *et al* [47]. One can then compare how well it performs to the weak values formalism: Is there information in the measurement results that we can retain instead of losing? Are there cases when the approximations of weak values are too broad? Does the ability to use a wider range of interaction strengths provide an advantage?

It appears that the advocates for use of the weak values formalism hope for it to be used to describe interesting quantum phenomena through a realist interpretation [170, 171, 172]. UQC is clearly an interesting quantum phenomenon and we see in the ADQC model how certain features of an efficient universal quantum computer can be assigned weak values. However, if we were to actually apply the weak value approximations instead of the exact calculations used in our model, would the errors introduced destroy the efficiency of the scheme? If the evolution of a register qubit was described by a series of weak values, would its description start to drift away from the un-approximated calculation? If this technique is meant to advance a particular interpretation of quantum mechanics, it needs to be valid for all quantum mechanical phenomena.

Chapter 7

Summary

We set out to show two things:

1. Ancilla driven quantum computation with interaction gates of non-maximal entangling strength can be universal. To achieve this we developed a model thus
 - limited ourselves to a qubit ancilla which can interact with a register qubit only once and interacts with no more than two register qubits
 - showed how a system under these limits will probabilistically generate gates on the register
 - showed how this probabilistic generating of gates can be modelled as a random walk, where the walker jumps through a space that corresponds to the group of single-qubit or one-parameter two-qubit gates, which can then be guaranteed to arrive at the desired gate
 - tested, through simulations, the possibility that these walks, when unguided, are guaranteed to land at a desired target and created an error measure to test the scaling of the walk times that may be used in later research to expand the scope of the simulations.
2. There is a trade off between the time to implement this scheme and the ability to control the ancilla through the choice of preparation state and measurement basis. To that end:
 - We created strategies for reducing the number of steps and ancilla qubits needed to implement the scheme based around reducing the range of the random walk in the random walk model of generating gates.
 - We tested the symmetry between preparation control and measurement control which we found to be, in most cases, equal.
 - We placed the model under extreme conditions: a long distance separation between register qubits and minimal control of the ancilla system.

In chapter 2, we reviewed the concepts of classical and quantum computation that this thesis relies upon. In section 2.1, we described the Turing machine and explained the ideas of universality, efficiency and the resources of the machine, time and space. In section 2.2, we explained Deutsch's quantisation of the Universal Turing Machine, the Universal Quantum Computer (UQC), which implements unitary operations over finite dimensions in discrete steps. We see schemes for quantum computation, as described in section 2.3, as a method for implementing a simulation of the Universal Quantum Computer. They describe how to take a set of operations and perform them according to some ordering to create the finite unitary operations of a UQC. These schemes require their own set of resources, for example the finite gate sets of circuit based computation or the entangled cluster states of measurement based computation, and then they will perform their simulation of the UQC finite operations with some cost of time and space. A scheme may be apt for a particular physical implementation whether because of the way it is performed or the resources it requires though these two issues are not strictly separable. This thesis is in particular concerned with the ancilla driven universal quantum computation scheme (ADQC) which uses a set of finite group operations that can be used for universal gate circuits or universal measurement based quantum computation. ADQC is distinguished by the way in which it discriminates between qubits used for a main register and qubits used as ancillae that makes it suitable for hybridised physical implementations that mix systems with long coherence lifetimes and short lived but highly manipulable systems. An ancilla qubit in a fixed state interacts with one or two register qubits with a single fixed interaction before being measured in a variable basis and is discarded with the measurement result being fed forward into the next measurement. ADQC has a restricted choice of two qubit operations in its resource set, requiring maximally entangling operations equivalent to a CZ or $CZ.SWAP$ but we have seen in gate based quantum computation, measurement based quantum computation and a range of hybrid physical implementation research look at discrete coupling operations with non-maximal entangling power.

Chapter 3 details the mathematical formalism used to describe the evolution of a system undergoing an ancilla driven operation. This formalism has a use in describing a number of different measurement operations, some of which are detailed in this chapter. In particular, we discuss some examples of measurements that operate by applying multiple iterations of the same process to build up a desired product operation in a guided stochastic process.

In chapter 4, we discuss the results of attempting to implement a unitary operation by an ancilla driven operation with interaction gates with arbitrary coupling parameters. In sections 4.1 and 4.2, we see that the restrictions on the two qubit interaction gates in ADQC are for ensuring stepwise determinism with post-corrections. As the

interactions become non-maximal, ancilla driven operations display a stochasticity that cannot be corrected by local gate operations after the interaction. In the rest of the chapter we show how to use strategies and the control of the ancilla preparation and measurement basis to overcome this.

Our results mostly focused on the use of interaction gates defined by a single parameter since solutions for two parameter gates under our restrictions behaved as a special sub-case of the one parameter case while outside of our restrictions, the question was being addressed elsewhere [43]. The gates that are generated by an ancilla driven step with arbitrary coupling parameters are parametrised by an angle from a random result of two angles whose size and probabilities are determined by the ancilla preparation state, measurement basis and coupling strength. In section 4.4, we show how the gates generated may achieve universal quantum computation by executing a random walk through the group of single and two qubit gates. This forms a probabilistic bounded-error polynomial time cost with respect to the error of approximation of the circuit gate. There was discussion, inspired by toy models in section 4.3, on how confining the space of the random walk leads to a speed up, exchanging the scheme resources of ancilla preparation and measurement for time.

In chapter 5, the implications of the probabilistic ADQC model on physical implementations are explored by placing the model under specific extreme circumstances. These conditions related the model to earlier research on long distance entanglement generation and research on universal gates in circuit based computation respectively and found that under the strictures of the ADQC model, the former favours asymmetry between the different register qubits and the latter favours ancilla systems with properties more similar to the register. In section 5.1, we focus on a model with long distance separation between register qubits where the preparation and measurement of ancilla qubits has to be performed by separate parties coordinated by classical communication. This adds another aspect to the question of trading off ancilla resources for time; there may be the ability to control the measurement of the ancilla or the preparation or both. The notion of confining the walk space was further explored when we used a simple confinement strategy that optimised the probability of success over one step. Under symmetrical conditions the ability to use both ancilla parameters was limited to a certain range of coupling strengths and had a small improvement. Comparing the one step strategy to another simple confinement strategy that could be performed by just the receiver end of the operation yielded a small improvement. The use of both parameters became more relevant when the interactions of the transmitter and receiver end were asymmetric. A possible distinguishing feature of probabilistic ADQC compared to other schemes for utilising mobile ancillae is the characterisation of weaker interactions and asymmetry between transmitter and receiver.

In section 5.2, we showed that probabilistic ADQC can be universal when the ancilla parameters for preparation and measurement are fixed. For single parameter interaction gates the universality depends on the local unitary effects. The universality of an interaction gate was compared to that of a single two qubit gate in the gate circuit scheme. Some gates that are not suitable for the gate circuit scheme may be suitable for minimal control ADQC and may in fact work faster because they can take advantage of confinement of the random walk over gates.

In the final chapter, we discussed some potential avenues for further research involving ancilla driven processes with non-maximal interactions: looking at guiding the random walk process, performing optimisations over several steps and the advantages of confinement; the impact of noise on the model, especially during the generation of a two-qubit gate; expanding the resource set further by including generalised measurements to implement non-unitary operations on the register; using the model to evaluate how much measurement information is lost in the characterisation of quantum systems using the weak value approximation.

Appendices

Appendix A

Proof of the decomposition of a single qubit unitary into rotations about any two non-parallel axes

In this thesis, we use the following theorem:

Theorem. *Suppose U is a unitary operations on a single qubit and $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$ are non-parallel real unit vectors in three dimensions. Then \exists real numbers α, β, γ and δ s.t.*

$$U = e^{i\alpha} R_{\hat{\mathbf{n}}}(\beta) R_{\hat{\mathbf{m}}}(\gamma) R_{\hat{\mathbf{n}}}(\delta) \quad (\text{A.1})$$

This is not a surprising result given the proofs of the universality of almost any two qubit gate given in 2.3.2 since one would expect that two unitary operations $R_{\hat{\mathbf{m}}}(x)$ and $R_{\hat{\mathbf{n}}}(y)$ for irrational x and y will correspond to Hamiltonians \hat{H}_m and \hat{H}_n with non-trivial commutation relations. It appears to be a fairly taken for granted result given it appears in [173] as a minor exercise. Since this theorem is useful in generalising some of the results of this thesis, we will provide a proof here.

Any $U \in U(2)$ can be written as $U = e^{i\alpha} R_{\hat{\mathbf{z}}}(\beta) R_{\hat{\mathbf{y}}}(\gamma) R_{\hat{\mathbf{z}}}(\delta)$ for the appropriate choice of $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ where $R_{\hat{\mathbf{n}}}(\alpha)$ denotes a rotation about the axis $\hat{\mathbf{n}}$ by the angle α . We can ignore the global phase factor and discuss just the factor $U_S \in SU(2)$ for which

$$U = e^{i\alpha} U_S. \quad (\text{A.2})$$

Given the non-parallel real unit vectors $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$ in three dimensions, define a unitary \widetilde{U}_S :

$$\widetilde{U}_S = R_{\hat{\mathbf{n}}}(\beta') R_{\hat{\mathbf{m}}}(\gamma') R_{\hat{\mathbf{n}}}(\delta') \quad (\text{A.3})$$

where β', γ' and δ' are variable real numbers. By the symmetry of the Bloch sphere, the result should be independent of rotations of the coordinate system. It is expected

that the axis $\hat{\mathbf{n}}$ could be set as equal to $\hat{\mathbf{z}}$. To confirm this, consider the unitary V that rotates the $\hat{\mathbf{n}}$ axis into the $\hat{\mathbf{z}}$ axis:

$$R_{\hat{\mathbf{n}}}(\beta') = V^\dagger R_{\hat{\mathbf{z}}}(\beta') V, \quad (\text{A.4})$$

$$R_{\hat{\mathbf{m}}'}(\gamma') := V R_{\hat{\mathbf{m}}}(\gamma') V^\dagger, \quad (\text{A.5})$$

$$\widetilde{U}_S = V^\dagger R_{\hat{\mathbf{z}}}(\beta') V V^\dagger R_{\hat{\mathbf{m}}'}(\gamma') V V^\dagger V^\dagger R_{\hat{\mathbf{z}}}(\delta') V, \quad (\text{A.6})$$

$$= V^\dagger R_{\hat{\mathbf{z}}}(\beta') R_{\hat{\mathbf{m}}'}(\gamma') R_{\hat{\mathbf{z}}}(\delta') V. \quad (\text{A.7})$$

The rotation about the $\hat{\mathbf{m}}'$ axis can also be decomposed in the Z-Y decomposition:

$$R_{\hat{\mathbf{m}}'}(\gamma') = R_{\hat{\mathbf{z}}}(\beta'') R_{\hat{\mathbf{y}}}(\gamma'') R_{\hat{\mathbf{z}}}(\delta''). \quad (\text{A.8})$$

For fixed $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$, which determine $\hat{\mathbf{m}}'$, the angles β'', γ'' and δ'' are set by γ' alone.

$$\widetilde{U}_S = V^\dagger R_{\hat{\mathbf{z}}}(\beta' + \beta'') R_{\hat{\mathbf{y}}}(\gamma'') R_{\hat{\mathbf{z}}}(\delta' + \delta'') V \quad (\text{A.9})$$

so we can now choose β', γ' and δ' for a given $\hat{\mathbf{m}}$ & $\hat{\mathbf{n}}$, s.t.

$$R_{\hat{\mathbf{z}}}(\beta' + \beta'') R_{\hat{\mathbf{y}}}(\gamma'') R_{\hat{\mathbf{z}}}(\delta' + \delta'') = V U V^\dagger \quad (\text{A.10})$$

$$\Rightarrow \widetilde{U}_S = U_S. \quad (\text{A.11})$$

Q.E.D.

Appendix B

Statistical analysis of numerically computed results

B.1 Simulation of random gate hitting times

In section 4.4, results were displayed from a simulation of the time taken to arrive at a target gate through the random generation of unitary gates. In this appendix, an analysis of the sources of error of the simulation will be performed.

B.1.1 Generating single-qubit gate hitting times.

We simulated the random generation of gates in two cases (1) the random unitary gates are $U_0 = HR_{\hat{z}}(\frac{\pi}{8}), U_1 = HR_{\hat{z}}(-\frac{\pi}{8})$ with probabilities $p_0 = p_1 = \frac{1}{2}$ (2) $U_0 = R_{\hat{z}}(\frac{\pi}{8})R_{\hat{x}}(\frac{\pi}{8}), U_1 = R_{\hat{z}}(-\frac{\pi}{8})R_{\hat{x}}(\frac{\pi}{8})$, $p_0 = p_1 = \frac{1}{2}$. The target unitary was $U_T = R_{\hat{x}}(\frac{\pi}{2})$. At each step a gate corresponding to the $\{U_0, U_1\}$ of each interaction was multiplied to the product of the previous step starting with the identity operator. The total product was compared to the target gate by a distance measure $\|V - U_T\|$ and the simulation stopped when the distance was smaller than a given error.

The simulation was repeated 1000 times for a given error size; furthermore, there were two possible measures and so the batch of 1000 simulations was collected for first the trace distance measure,

$$\|V - U_T\| = \epsilon(U_T, V) = \sqrt{\frac{2 - |\text{Tr}[V^\dagger U_T]|}{2}},$$

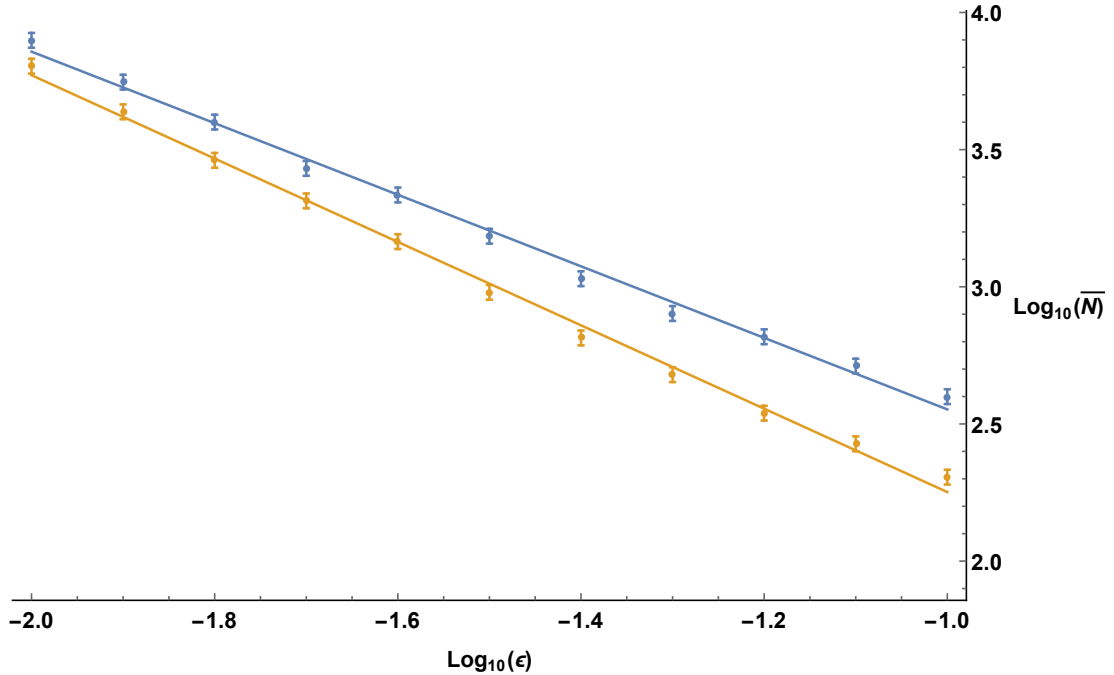
and then again for the Euclidean distance measure,

$$\|V - U_T\| = \delta(U_T, V) = \sqrt{-\frac{1}{4} \sum_j \text{Tr}[\sigma_{x_j}(U_T - V)]^2}$$

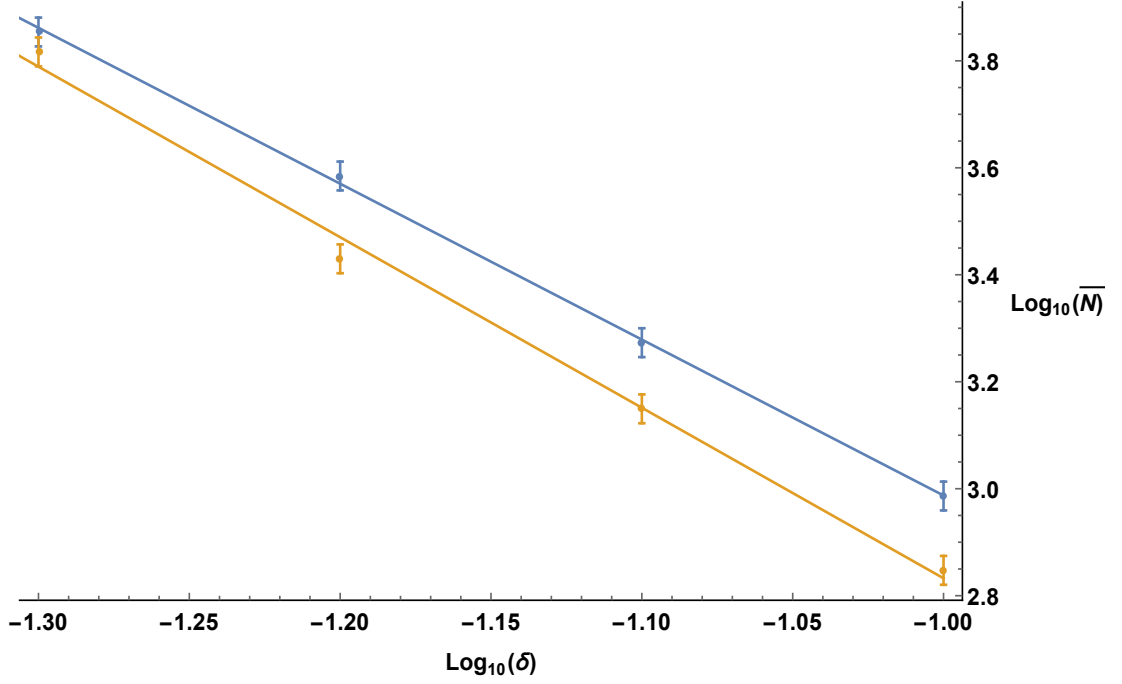
To find the scaling relationship between the mean hitting times and the error size, the simulation was repeated with the error bound successively multiplied by a factor of

$10^{-0.1}$ and then the logarithms of the mean times with base 10 were used to construct a linear model fit using the least-squares method.

The results of the simulation using the trace distance measure are displayed in figure 4.35 and the results with the Euclidean distance measure in figure 4.37 in section 4.4.2. For the trace distance error, the one-parameter case was fitted with the line $1.25 - 1.30x$ (in blue), with $R^2 = 0.994$ (all numbers rounded to 3 s.f.), and the two-parameter case gave $0.7322 - 1.52x$ (in gold), $R^2 = 0.996$:



For the Euclidean distance error, the one-parameter case was fit with the line (in gold) $40.400 - 2.89x$, $R^2 = 1.000$, and the two-parameter case with $-0.0381 - 3.18x$ (in blue), $R^2 = 0.992$. Because of the additional computational time needed to calculate the Euclidean distance measure, as well as the increased average hitting times, the range of values of δ was not extended as far as ϵ :



Two-qubit gate hitting times

A similar process was repeated for the simulation of a two-qubit gate generation randomly generating gates locally equivalent to the class $C(\gamma)$. The gate generation was based on an interaction gate with coupling parameter $\alpha = \frac{\pi}{16}$, preparation parameter $\beta = \frac{\pi}{32}$ and measurement parameter $\theta = \frac{\pi}{2}$. This results in $U_0 \equiv C(\Phi_0)$, $U_1 \equiv C(\Phi_1)$ where, according to (4.104),(4.105),(4.106) and (4.107):

$$\Phi_0 = 4\arctan\left(\frac{-\cos(\frac{\pi}{32})}{\cos(\frac{3\pi}{32})}\right) + \pi, \quad (\text{B.1})$$

$$\Phi_1 = 4\arctan\left(\frac{\sin(\frac{3\pi}{32})}{\sin(\frac{\pi}{32})}\right) + \pi. \quad (\text{B.2})$$

The target gate is $C(\Phi_T) = C(\frac{\pi}{2})$ and the measure of distance is the angle between the product of the generations and the target on the circle representing the class:

$$\gamma(C(\Phi), C(\Phi_T)) = |\Phi - \Phi_T|. \quad (\text{B.3})$$

A deterministic comparison was also simulated by running the simulation with $p_0 = 1$. The results are displayed in figure 4.39 in section 4.4.2.

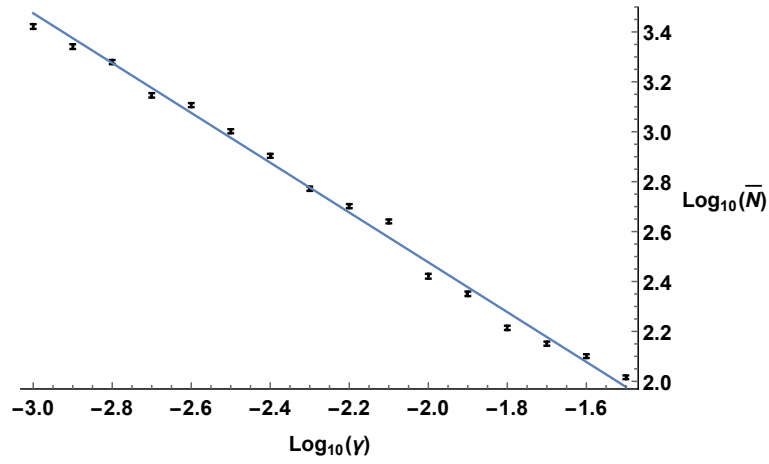
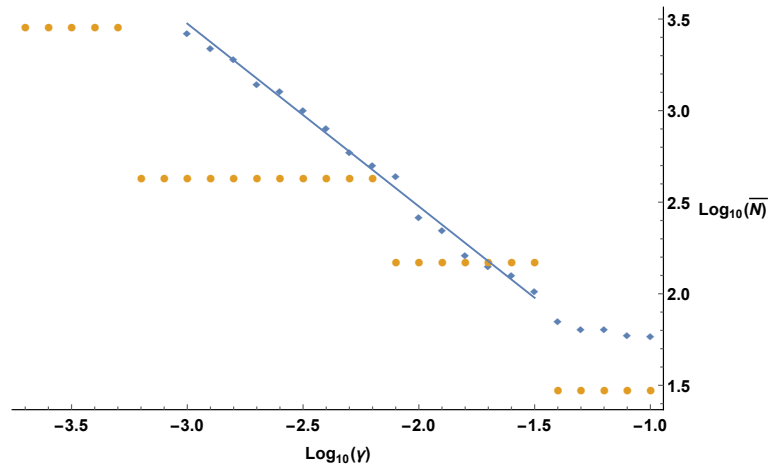


Figure B.1: A logarithmic plot of the mean number of gates over 10,000 simulations required to hit a target gate against the error bound γ , across the range $10^{-3} \leq \gamma \leq 10^{-1.5}$. The error bars, based on a supposed exponential distribution of hitting times, are plot in black.



Here the line is $0.481 - 0.998x$ (to 3 s.f.) with $R^2 = 0.993$ and there are 10,000 results and the corresponding reduction of errors makes the error bars difficult to see on the scale of the full range of the results so figure B.1 redisplayes the results over the range of the data points used to construct the line and excludes the data from the deterministic case.

B.1.2 Error bar calculations

Since we are treating the hitting times as an exponentially distributed population, we use this distribution to calculate a confidence interval for the mean hitting times which are placed into the plots as error bars. From [174], we can say that if $x_1, x_2 \dots x_n$ are independent random variables from an exponential distribution with mean θ then a

confidence interval of $1 - \alpha$ can be given by the chi-squared distribution using

$$\text{prob} \left(\frac{2n\bar{x}}{\chi_{\alpha/2, 2n}^2} < \theta < \frac{2n\bar{x}}{\chi_{1-\alpha/2, 2n}^2} \right) = 1 - \alpha \quad (\text{B.4})$$

where the maximum likelihood mean estimator is $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

If n is sufficiently large, the chi-squared distribution of $2n$ degrees of freedom will be approximately equal to a normal distribution with mean $2n$ and variance $4n$. Then one can use 1.96 times the standard deviation of the distribution for a 95% confidence interval:

$$\text{prob} \left(\frac{2n\bar{x}}{2n + 1.96\sqrt{4n}} < \theta < \frac{2n\bar{x}}{2n - 1.96\sqrt{4n}} \right) = 0.95 \quad (\text{B.5})$$

$$\text{prob} \left(\frac{\bar{x}}{1 + \frac{1.96}{\sqrt{n}}} < \theta < \frac{\bar{x}}{1 - \frac{1.96}{\sqrt{n}}} \right) = 0.95 \quad (\text{B.6})$$

B.2 Numerical calculations of guided strategy probabilities

In sections 4.4 and 5.1, we discussed strategies in which the random gate generation of two-qubit entangling gates was guided to provide a relationship between the implementation time and the coupling parameter of the interaction gate. For three guided strategies, the expected number of steps needed were plotted against the coupling parameter, α , and these expectation times are functions of probabilities that are in turn functions of the parameters of the two-qubit entangling gate generation. The calculation of these probabilities required numerical methods whose details we will go into here.

For the flip-undo strategy in section 4.4.4, the only probability needed to be calculated is the probability of generating a CZ equivalent gate in a single step which occurs under only one set of parameters as detailed in section 4.2. In this case the probability could be calculated directly from the function of those parameters for a given α but on the other hand, for the repeat-until-success strategy, the parameters depend upon α . It is necessary to find ancilla preparation parameter, β , (and/or ancilla measurement parameter θ in a more general case) such that for the two entangling power outputs of the entangling gate generation, Φ_0 and Φ_1 ,

$$|\Phi_0(\alpha, \beta, \theta) - \Phi_1(\alpha, \beta, \theta)| = \pi. \quad (\text{B.7})$$

For the figure 4.44, this was calculated using Wolfram Mathematica 10's FindRoot function [175] which employs Newton's root finding method. It attempts to find root approximations x with a numerical error less than

$$10^{-a} + |x|10^{-p} \quad (\text{B.8})$$

for a user specified accuracy goal, a , and use specified precision goal p . For our simulations, $a = p = 8$.

Unlike the other two guided strategies, the calculation of the one-step strategy expectation times was more extensive. Each step n requires the generation of a different entangling power output, $\Phi_{1,n}$ with an associated step dependent probability $p_{1,n}$. To calculate the expected value requires the value of $p_{1,n}$ over all n .

In practise, $p_{1,n}$ is calculated up to a finite number of steps, C , that fulfils a convergence criteria. For example, when the cumulative density function, a property that can be calculated at n without knowledge of any $p_{n'}$ for $n' > n$, is within a small threshold of one:

$$1 - CDF(C) \leq \epsilon, \epsilon \in \mathbb{R}, \epsilon \ll 1. \quad (\text{B.9})$$

An early test run using a proxy number of 500 steps found that 500 steps equated to a threshold of the order of 10^{-3} for the smallest coupling parameter in the range plotted and quickly dropped to the order of 10^{-11} for the next highest value and so the evaluation of the convergence criterion was left out to simplify code and a standard constant of 500 steps used.

The necessary entangling power output at each step n , $\Phi_{1,n}$, is a function of the output of the previous steps:

$$\Phi_{1,n} = \pi - \sum_{k=1}^{n-1} \Phi_{0,k} \quad (\text{B.10})$$

therefore calculating each step requires starting at the first step, finding the probability of generating $\Phi_{1,1}$, $p_{1,1}$, and the output of the failure result $\Phi_{0,1}$ to calculate the requirement for the next step. Then the requirement must be used to calculate the necessary parameters and the resulting probability. This is then iterated over all steps:

1. Set $\beta_1 = \alpha$, $\theta = \frac{\pi}{2}$,
2. Calculate $\Phi_{0,1}$, $p_{1,1}$,
3. Set at $n = 2$,
4. Find β_n (and/or θ_n) s.t. $\Phi_1(\alpha, \beta_n, \theta_n) - (\pi - \sum_{k=1}^{n-1} \Phi_{0,k}) = 0$,
5. Calculate $\Phi_0(\alpha, \beta_n, \theta_n)$ and $p_1(\alpha, \beta_n, \theta_n)$,
6. Increase $k \rightarrow k + 1$, if $k < C$ go to 4, else stop.

However under the variations of the one-step strategy some particular issues arise.

The one-port, one-degree-of-freedom case

This is the simplest case where only the β value is varied and the same result of the two in the fixed measurement basis is associated with the success in every step. Like the the repeat-until-success strategy, the values of β in step 4 of the calculation were found using Wolfram Mathematica 10's FindRoot function [175]. Because the direction associated with the angle corresponding to the required entangling power output keeps changing, yet does not affect the probabilities of success, a function of the absolute distance is created, d , for which

$$d_n = |d_n + \Phi_{0,n-1}|. \quad (\text{B.11})$$

Then the requirement is

$$\Phi_{1,n+1} - (\pi - d_n) = 0. \quad (\text{B.12})$$

The two-port, one-d.o.f. case

In this case, only the β value is varied but there is a choice of adding the ability to switch between using either of the two outputs from the measurement. Essentially this means that there is another curve against β for which a limited range of Φ occurs with greater probability. It is only viable under the one-step strategy when $|\Phi_{0,1}| = |\Phi_{0,n}|_{\max} \leq \frac{\pi}{2}$. This condition depends on α and so is evaluated before the calculation of the probabilities. During the calculation of the probabilities it is then necessary to insert test the following condition:

$$\pi - d_n \leq |\Phi_0|_{\max}. \quad (\text{B.13})$$

If false, then we proceed as with the one-port case, if true then we substitute the calculation for finding β_n s.t. $\Phi_0(\alpha, \beta_n, \theta_n) - (\pi - d_{n-1}) = 0$ into step 4 and exchange $0 \leftrightarrow 1$ for that step.

The two-port, two-d.o.f. case

In this case, both the preparation parameter, β , and the measurement parameter, θ , may be varied. For port 1, the probability of a given Φ_1 is maximised by taking the measurement on the equator of the Bloch sphere but for the 0 port, the optimum probability for a given Φ_0 can occur for different (β, θ) . The task is a constrained optimisation problem to find the maximum probability along a contour of fixed Φ_0 as in figure 4.19. For this, we used Wolfram Mathematica 10's FindMaximum function [176] which uses the interior point algorithm for constrained optimisation [177], appropriate for the smooth curves seen in figures 4.16 and 4.20. The convergence criteria on the numerical

calculation of x for maximised function $f(x)$ is that $\|x_k - x\| \leq \max(10^{-a}, \|x_k\|10^{-p})$ and $\nabla f(x_k) \leq 10^{-a}$ for accuracy goal a and precision goal p .

For the calculations of the two-port, two-d.o.f. case displayed in figure 5.7 in section 5.1.4, to accommodate the additional computation time of the two parameter optimisation, the number of steps in the one-step strategy was limited to 50 which, for this smaller range with higher probabilities, means that the cumulative density function has converged to within 10^{-9} of 1, and the accuracy and precision goals were set to $a = p = 3$. We will treat the β and θ errors equally leading to a doubling of the square of the error ϵ (see below).

B.2.1 Errors in the numerical calculation

The numerical method of estimating the probabilities of each step introduce a source of error into the calculations and due to the way in which each step's calculation depends upon previous results, these errors propagate. In this section, we will analyse these errors.

For an expected number of steps,

$$\bar{n} = \sum_{n=1}^C n\tau_n, \quad (\text{B.14})$$

for which τ_n is the probability of halting on step n , the errors on the expected number add in quadrature

$$\epsilon_{\bar{n}}^2 = \sum_n^C n^2 \epsilon_{\tau_n}^2. \quad (\text{B.15})$$

The probability of *halting on* step n depends on the probability of *success on* step n , p_n ,

$$\tau_n = \prod_j^{n-1} (1 - p_j)p_n \quad (\text{B.16})$$

and the errors depend thus:

$$\epsilon_{\tau_n} = \sum_k^n \frac{\partial \tau_n}{\partial p_k} \epsilon_{p_k}^2 \quad (\text{B.17})$$

$$\frac{\partial \tau_n}{\partial p_k} = \begin{cases} -\frac{p_n}{1-p_k} \prod_j^{n-1} (1 - p_j) & \text{if } k < n \\ \prod_j^{n-1} (1 - p_j) & \text{if } k = n \end{cases} \quad (\text{B.18})$$

We will simplify the calculation of the propagation of errors in two ways. First, we will treat each calculation of p_k as independent i.e. as though we ran the calculation up to p_k and then stopped to run it again independently for p_{k+1} . This means that we can treat each ϵ_{p_k} as independent in the addition of quadratures in (B.17) as a trade-off for overestimating the contribution of each error introduced at step k .

$\alpha \cdot \frac{\pi}{40}^{-1}$	ϵ (to 3 s.f.)
1	5.58×10^{-4}
2	1.91×10^{-5}
3	2.89×10^{-6}
4	8.30×10^{-7}
5	3.44×10^{-7}
6	1.82×10^{-7}
7	1.15×10^{-7}
8	8.25×10^{-8}
9	6.43×10^{-8}

Table B.1: Table of the numerical calculation errors against coupling parameter α as according to (B.21).

Second, we will make a linear approximation to the relationship between β and p_k , and thus also between the probabilities at different steps, justified by the shape of the curve in 4.20, so that

$$p_k \approx \sum_j^{k-1} m_j p_j + c_j. \quad (\text{B.19})$$

Also, we will make an order approximation that $(O)(|m_j|) = 1$. This results in treating ϵ_{p_k} simply as

$$\epsilon_{p_k}^2 = \sum_j^k \epsilon_j^2 \approx k\epsilon^2 \quad (\text{B.20})$$

given a similar size of error introduced at each step. Recall that the probabilities of the one-step strategy are mostly bound by the values of the first two steps (see figure 5.3) which makes similar sized probabilities, that are already bound within the range $(0, 0.5)$. Using an approximately linear relationship between β and the probabilities to relate the error sizes, we use (B.8) with $|x| < 0.5$ to make an upper bound estimate of $\epsilon = 1.5 \times 10^{-8}$. Our evaluation of the error on the expected number of steps is

$$\epsilon_{\hat{n}}^2 = \sum_n^C n^2 \sum_k^n \frac{\partial \tau_n}{\partial p_k}^2 k \epsilon^2 \quad (\text{B.21})$$

The results of (B.21) are displayed in tables B.1 and B.2. The small size of the numerical errors shows that the dominant factor in the quality of the calculation results lies not likely in the error sizes of the calculations but in the quality of the choice of calculation method itself. Efforts to improve upon this choice we leave to future researchers.

$\alpha \cdot \frac{\pi}{40}^{-1}$	ϵ (to 3 s.f.)
7.5	5.96×10^{-3}
8	4.39×10^{-3}
8.5	4.11×10^{-3}
9	3.84×10^{-3}
9.5	3.82×10^{-3}

Table B.2: Table of the numerical calculation errors against coupling parameter α as according to (B.21), for calculations of the two degree of freedom case.

Bibliography

- [1] Paul Benioff. The computer as a physical system: A microscopic quantum mechanical hamiltonian model of computers as represented by turing machines. *Journal of Statistical Physics*, 22(5):563–591, May 1980.
- [2] David Deutsch. Quantum theory, the church-turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117, July 1985.
- [3] Richard P. Feynman. Simulating Physics with Computers. *International Journal of Theoretical Physics*, 21:467–488, June 1982.
- [4] Brian D. Josephson. Possible new effects in superconductive tunnelling. *Physics Letters*, 1(7):251 – 253, July 1962.
- [5] J.M. Rowell. Magnetic field dependence of the phson tunnel current. *Phys. Rev. Lett.*, 11:200–202, September 1963.
- [6] R. Jaklevic, John Lambe, A. Silver, and J. Mercereau. Quantum interference effects in josephson tunneling. *Phys. Rev. Lett.*, 12:159–160, February 1964.
- [7] G. Binnig and H. Rohrer. Surface studies by scanning tunneling microscopy. *Helvetica Physics Acta*, 55:726–735, December 1982.
- [8] G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel. Surface studies by scanning tunneling microscopy. *Phys. Rev. Lett.*, 49:57–61, July 1982.
- [9] G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel. Surface studies by scanning tunneling microscopy. *Phys. Rev. Lett.*, 49:57–61, July 1982.
- [10] Gerd Binnig and Heinrich Rohrer. Scanning tunneling microscopy- from birth to adolescence. *Rev. Mod. Phys.*, 59:615–625, July 1987.
- [11] R. Rossetti, J. L. Ellison, J. M. Gibson, and L. E. Brus. Size effects in the excited electronic states of small colloidal cds crystallites. *The Journal of Chemical Physics*, 80(9):4464–4469, May 1984.

- [12] A.I. Ekimov, A.L. Efros, and A.A. Onushchenko. Quantum size effect in semiconductor microcrystals. *Solid State Communications*, 56(11):921 – 924, December 1985.
- [13] Arnim Henglein. Small-particle research: physicochemical properties of extremely small colloidal metal and semiconductor particles. *Chemical Reviews*, 89(8):1861–1873, December 1989.
- [14] Stig Stenholm. The semiclassical theory of laser cooling. *Rev. Mod. Phys.*, 58:699–739, July 1986.
- [15] Shigeki Takeuchi. Recent progress in single-photon and entangled-photon generation and applications. *Japanese Journal of Applied Physics*, 53(3):030101, February 2014.
- [16] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, August 1996.
- [17] David Deutsch and Richard Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 439(1907):553–558, December 1992.
- [18] André Berthiaume and Gilles Brassard. Oracle quantum computing. *Journal of Modern Optics*, 41(12):2521–2535, December 1994.
- [19] D.R. Simon. On the power of quantum computation. In *35th Annual Symposium on Foundations of Computer Science, 1994 Proceedings*, pages 116–123, November 1994.
- [20] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. In *in Proc. 25th Annual ACM Symposium on Theory of Computing, ACM*, pages 11–20, October 1993.
- [21] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, October 1997.
- [22] David Deutsch. Quantum computational networks. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 425(1868):73–90, September 1989.
- [23] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001.

- [24] Dan Browne and Hans J. Briegel. One-way quantum computation. In *Lectures on Quantum Information*, pages 359–358. Wiley-VCH, Weinheim, Germany, 2007.
- [25] Marc Hein, Wolfgang Dür, Jens Eisert, Robert Raussendorf, Maarten Van Den Nest, and Hans Jürgen Briegel. Entanglement in Graph States and its Applications. *eprint arXiv:quant-ph/0602096v1*, February 2006.
- [26] A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2 – 30, January 2003.
- [27] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Quantum Computation By Adiabatic Evolution. *eprint arXiv:quant-ph/0001106*, January 2000.
- [28] Giuseppe E Santoro and Erio Tosatti. Optimization using quantum mechanics: quantum annealing through adiabatic evolution. *Journal of Physics A: Mathematical and General*, 39(36):R393, August 2006.
- [29] Adriano Barenco. A universal two-bit gate for quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 449(1937):679–683, June 1995.
- [30] Tycho Sleator and Harald Weinfurter. Realizable universal quantum logic gates. *Phys. Rev. Lett.*, 74:4087–4090, May 1995.
- [31] David Deutsch, Adriano Barenco, and Artur Ekert. Universality in quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 449(1937):669–677, June 1995.
- [32] Adriano Barenco, Charles H. Bennett, Richard Cleve, David P. DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A. Smolin, and Harald Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, 52:3457–3467, November 1995.
- [33] Michael J. Bremner, Christopher M. Dawson, Jennifer L. Dodd, Alexei Gilchrist, Aram W. Harrow, Duncan Mortimer, Michael A. Nielsen, and Tobias J. Osborne. Practical scheme for quantum computation with any two-qubit entangling gate. *Phys. Rev. Lett.*, 89:247902, November 2002.
- [34] P. Oscar Boykin, Tal Mor, Matthew Pulver, Vwani Roychowdhury, and Farrokh Vatan. On universal and fault-tolerant quantum computing. *eprint arXiv:quant-ph/9906054*, June 1999.

- [35] Alex Bocharov and Krysta Svore. Resource-optimal single-qubit quantum circuits. *Phys. Rev. Lett.*, 109:190501, November 2012.
- [36] Alex Bocharov, Yuri Gurevich, and Krysta Svore. Efficient decomposition of single-qubit gates into v basis circuits. *Phys. Rev. A*, 88:012313, July 2013.
- [37] E. Kashefi, D.K.L. Oi, D. Browne, J. Anders, and E. Andersson. Twisted graph states for ancilla-driven universal quantum computation. *Electronic Notes in Theoretical Computer Science*, 249(0):307 – 331, August 2009. Proceedings of the 25th Conference on Mathematical Foundations of Programming Semantics (MFPS 2009).
- [38] Janet Anders, Daniel K. L. Oi, Elham Kashefi, Dan E. Browne, and Erika Andersson. Ancilla-driven universal quantum computation. *Phys. Rev. A*, 82(2):020301, August 2010.
- [39] B. B. Blinov, D. L. Moehring, L.-M. Duan, and C. Monroe. Observation of entanglement between a single trapped atom and a single photon. *Nature*, 428:153–157, March 2004.
- [40] M. V. Gurudev Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A. S. Zibrov, P. R. Hemmer, and M. D. Lukin. Quantum register based on individual electronic and nuclear spin qubits in diamond. *Science*, 316(5829):1312–1316, June 2007.
- [41] A. Bermudez, F. Jelezko, M. B. Plenio, and A. Retzker. Electron-mediated nuclear-spin interactions between distant nitrogen-vacancy centers. *Phys. Rev. Lett.*, 107:150503, October 2011.
- [42] Simon J. Devitt, Andrew D. Greentree, and Lloyd C.L. Hollenberg. Information free quantum bus for generating stabiliser states. *Quantum Information Processing*, 6(4):229–242, August 2007.
- [43] Timothy J. Proctor, Erika Andersson, and Viv Kendon. Universal quantum computation by the unitary control of ancilla qubits and using a fixed ancilla-register interaction. *Phys. Rev. A*, 88:042330, October 2013.
- [44] D. Gross and J. Eisert. Novel schemes for measurement-based quantum computation. *Phys. Rev. Lett.*, 98:220503, May 2007.
- [45] D. Gross, J. Eisert, N. Schuch, and D. Perez-Garcia. Measurement-based quantum computation beyond the one-way model. *Phys. Rev. A*, 76:052315, November 2007.

- [46] Jeff S. Lundeen and Charles Bamber. Procedure for direct measurement of general quantum states using weak measurement. *Phys. Rev. Lett.*, 108:070402, February 2012.
- [47] Jeff Z Salvail, Megan Agnew, Johnson Allan S., Bolduc Eliot, Jonathan Leach, and Robert W. Boyd. Full characterization of polarization states of light via direct measurement. *Nat Photon*, 7:316–321, March 2013.
- [48] G. Cordourier-Maruri, F. Ciccarello, Y. Omar, M. Zarccone, R. de Coss, and S. Bose. Implementing quantum gates through scattering between a static and a flying qubit. *Phys. Rev. A*, 82:052313, November 2010.
- [49] G. Cordourier-Maruri, Y. Omar, R. de Coss, and S. Bose. Graphene Enabled Low-Control Quantum Gates between Static and Mobile Spins. *eprint arXiv:quant-ph/1307.0217*, June 2013.
- [50] Timothy J. Proctor and Viv Kendon. Minimal ancilla mediated quantum computation. *EPJ Quantum Technology*, 1(1):1–11, September 2014.
- [51] Kerem Halil Shah and Daniel K.L. Oi. Ancilla Driven Quantum Computation with Arbitrary Entangling Strength. In Simone Severini and Fernando Brandao, editors, *8th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2013)*, volume 22 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 1–19, Dagstuhl, Germany, November 2013. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- [52] Kerem Halil Shah and Daniel K.L. Oi. Entangling unitary gates on distant qubits with ancilla feedback. *eprint arXiv:quant-ph/1311.3463*, November 2013.
- [53] Kerem Halil-Shah and Daniel K.L. Oi. A minimum control ancilla driven quantum computation scheme with repeat-until-success style gate generation. *eprint arXiv:quant-ph/1401.8004*, January 2014.
- [54] A. M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1):230–265, 1937.
- [55] Alonzo Church. An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 58(2):pp. 345–363, April 1936.
- [56] Ding-Zhu Du and Ker-I Ko. *Models of Computation and Complexity Classes*, pages 1–44. John Wiley & Sons, Inc., 2014.

- [57] K. de Leeuw, E. F. Moore, C. E. Shannon, and N. Shapiro. Computability by probabilistic machines. In *Automata studies*, Annals of mathematics studies, no. 34, pages 183–212. Princeton University Press, Princeton, N. J., 1956.
- [58] John T. Gill, III. Computational complexity of probabilistic turing machines. In *Proceedings of the Sixth Annual ACM Symposium on Theory of Computing*, STOC '74, pages 91–95, New York, NY, USA, May 1974. ACM.
- [59] Eugene S. Santos. Probabilistic Turing machines and computability. *Proc. Amer. Math. Soc.*, 22:704–710, September 1969.
- [60] R. Solovay and V. Strassen. A fast monte-carlo test for primality. *SIAM Journal on Computing*, 6(1):84–85, 1977.
- [61] Michael O Rabin. Probabilistic algorithm for testing primality. *Journal of Number Theory*, 12(1):128 – 138, February 1980.
- [62] D.J.A. Welsh. Randomised algorithms. *Discrete Applied Mathematics*, 5(1):133 – 145, January 1983.
- [63] Richard M. Karp. An introduction to randomized algorithms. *Discrete Applied Mathematics*, 34(13):165 – 201, November 1991.
- [64] J. Hartmanis and R. E. Stearns. On the computational complexity of algorithms. *Trans. Amer. Math. Soc.*, 117:285–306, May 1965.
- [65] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*, chapter 3, pages 136–137. Cambridge University Press, Cambridge, United Kingdom, 2000.
- [66] Jack Edmonds. Paths, trees, and flowers. *Canadian Journal of Mathematics*, 17:449–467, February 1965.
- [67] G. E. Moore. Cramming More Components onto Integrated Circuits. *Electronics*, 38(8):114–117, April 1965.
- [68] Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, STOC '71, pages 151–158, New York, NY, USA, May 1971. ACM.
- [69] Richard M. Karp. Reducibility among combinatorial problems. In Raymond E. Miller, James W. Thatcher, and Jean D. Bohlinger, editors, *Complexity of Computer Computations*, The IBM Research Symposia Series, pages 85–103. Springer US, 1972.

- [70] Richard P. Feynman. There's plenty of room at the bottom. *Journal of Microelectromechanical Systems*, 1(1):60–66, March 1992.
- [71] Edward Fredkin and Tommaso Toffoli. Conservative logic. *International Journal of Theoretical Physics*, 21(3-4):219–253, April 1982.
- [72] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. In *in Proc. 25th Annual ACM Symposium on Theory of Computing, ACM*, pages 11–20, October 1993.
- [73] C.H. Bennett. Logical reversibility of computation. *IBM Journal of Research and Development*, 17(6):525–532, November 1973.
- [74] Tommaso Toffoli. Reversible computing. In Jaco de Bakker and Jan van Leeuwen, editors, *Automata, Languages and Programming*, volume 85 of *Lecture Notes in Computer Science*, pages 632–644. Springer Berlin Heidelberg, 1980.
- [75] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*, chapter 4, page 196. Cambridge University Press, Cambridge, United Kingdom, 2000.
- [76] Tommaso Toffoli. Bicontinuous extensions of invertible combinatorial functions. *Mathematical systems theory*, 14(1):13–23, December 1981.
- [77] David P. DiVincenzo. Two-bit gates are universal for quantum computation. *Phys. Rev. A*, 51:1015–1022, February 1995.
- [78] Nengkun Yu, Runyao Duan, and Mingsheng Ying. Five two-qubit gates are necessary for implementing the toffoli gate. *Phys. Rev. A*, 88:010304, July 2013.
- [79] Seth Lloyd. Almost any quantum logic gate is universal. *Phys. Rev. Lett.*, 75:346–349, July 1995.
- [80] Andrew M. Childs, Debbie W. Leung, Laura Mancinska, and Maris Ozols. Characterization of universal two-qubit hamiltonians. *Quantum Information & Computation*, 11(1&2):19–39, January 2011.
- [81] Michael Reck, Anton Zeilinger, Herbert J. Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.*, 73:58–61, July 1994.
- [82] P. W. Shor. Fault-tolerant quantum computation. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science, FOCS '96*, pages 56–65, Washington, DC, USA, October 1996. IEEE Computer Society.

- [83] Daniel Gottesman. The Heisenberg Representation of Quantum Computers. *eprint arXiv:quant-ph/9807006v1*, July 1998.
- [84] Gabriele Nebe, E.M. Rains, and N.J.A. Sloane. The Invariants of the Clifford Group. *eprint arXiv:math/0001038v2*, January 2000.
- [85] A. Yu. Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191, 1997.
- [86] Adam Paetznick and Krysta M. Svore. Repeat-Until-Success: Non-deterministic decomposition of single-qubit unitaries. *eprint arXiv:quant-ph/1311.1074*, November 2013.
- [87] Michael A. Nielsen. Cluster-state quantum computation. *Reports on Mathematical Physics*, 57(1):147 – 161, February 2006.
- [88] Hans J. Briegel and Robert Raussendorf. Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.*, 86:910–913, January 2001.
- [89] Massoud Borhani and Daniel Loss. Cluster states from heisenberg interactions. *Phys. Rev. A*, 71:034308, March 2005.
- [90] Vincent Danos, Elham Kashefi, and Prakash Panangaden. Parsimonious and robust realizations of unitary maps in the one-way model. *Phys. Rev. A*, 72:064301, December 2005.
- [91] Dan Browne and Janet Anders. The role of classical computation in measurement-based quantum computation. In *Proceedings of the 4th Conference on Computability in Europe: Logic and Theory of Algorithms*, CiE '08, pages 94–99, Berlin, Heidelberg, July 2008. Springer-Verlag.
- [92] D. Jaksch, C. Bruder, J. Cirac, C. Gardiner, and P. Zoller. Cold bosonic atoms in optical lattices. *Phys. Rev. Lett.*, 81:3108–3111, October 1998.
- [93] D. Jaksch, H.-J. Briegel, J. Cirac, C. Gardiner, and P. Zoller. Entanglement of atoms via cold controlled collisions. *Phys. Rev. Lett.*, 82:1975–1978, March 1999.
- [94] Gavin Brennen, Carlton Caves, Poul Jessen, and Ivan Deutsch. Quantum logic gates in optical lattices. *Phys. Rev. Lett.*, 82:1060–1063, February 1999.
- [95] L.-M. Duan, E. Demler, and M. Lukin. Controlling spin exchange interactions of ultracold atoms in optical lattices. *Phys. Rev. Lett.*, 91:090402, August 2003.
- [96] Christof Weitenberg, Manuel Endres, Jacob F. Sherson, Marc Cheneau, Peter Schausz, Takeshi Fukuhara, Immanuel Bloch, and Stefan Kuhr. Single-spin addressing in an atomic mott insulator. *Nature*, 471:319–324, March 2011.

- [97] Monika Bartkowiak, Lian-Ao Wu, and Adam Miranowicz. Quantum circuits for amplification of kerr nonlinearity via quadrature squeezing. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 47(14):145501, July 2014.
- [98] Daniel Gottesman and Isaac L. Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402:390–393, November 1999.
- [99] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling, and G. J. Milburn. Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.*, 79:135–174, January 2007.
- [100] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:46–52, January 2001.
- [101] Daniel Browne and Terry Rudolph. Resource-efficient linear optical quantum computation. *Phys. Rev. Lett.*, 95:010501, June 2005.
- [102] Sean D. Barrett and Pieter Kok. Efficient high-fidelity quantum computation using matter qubits and linear optics. *Phys. Rev. A*, 71:060310, June 2005.
- [103] Daniel K. L. Oi, Simon J. Devitt, and Lloyd C. L. Hollenberg. Scalable error correction in distributed ion trap computers. *Phys. Rev. A*, 74:052313, November 2006.
- [104] Chetan Nayak, Steven Simon, Ady Stern, Michael Freedman, and Sankar Das Sarma. Non-abelian anyons and topological quantum computation. *Rev. Mod. Phys.*, 80:1083–1159, September 2008.
- [105] B. Paredes, P. Fedichev, J. Cirac, and P. Zoller. $\frac{1}{2}$ -anyons in small atomic bose-einstein condensates. *Phys. Rev. Lett.*, 87:010402, June 2001.
- [106] Dorit Aharonov, Wim van Dam, Julia Kempe, Zeph Landau, Seth Lloyd, and Oded Regev. Adiabatic Quantum Computation is Equivalent to Standard Quantum Computation. *SIAM J. Comput.*, 37(1):166–194, April 2007.
- [107] Paolo Zanardi and Mario Rasetti. Holonomic quantum computation. *Physics Letters A*, 264(23):94 – 99, December 1999.
- [108] Daniel Loss and David DiVincenzo. Quantum computation with quantum dots. *Phys. Rev. A*, 57:120–126, January 1998.
- [109] David P. DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik*, 48:771, February 2000.

- [110] C. W. Chou, D. B. Hume, J. C. J. Koelemeij, D. J. Wineland, and T. Rosenband. Frequency comparison of two high-accuracy Al^+ optical clocks. *Phys. Rev. Lett.*, 104:070802, February 2010.
- [111] Toshio Ohshima. All-optical electron spin quantum computer with ancilla bits for operations in each coupled-dot cell. *Phys. Rev. A*, 62:062316, November 2000.
- [112] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, United Kingdom, 2000.
- [113] John Preskill. Foundations of quantum theory ii: Measurement and evolution, December 2010.
- [114] Stephen Barnett. *Quantum Information*. Oxford University Press, Oxford, United Kingdom, 2009.
- [115] Victor Klee. What is a convex set? *The American Mathematical Monthly*, 78(6):pp. 616–631, June 1971.
- [116] Nancy Cartwright. The only real probabilities in quantum mechanics. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1978:54–59, 1978.
- [117] Yakir Aharonov, David Z. Albert, and Lev Vaidman. How the result of a measurement of a component of the spin of a spin- $\frac{1}{2}$ particle can turn out to be 100. *Phys. Rev. Lett.*, 60:1351–1354, April 1988.
- [118] Y. Aharonov and L. Vaidman. The Two-State Vector Formalism of Quantum Mechanics. *eprint arXiv:quant-ph/0105101*, May 2001.
- [119] Anthony Chefles. Quantum state discrimination. *Contemporary Physics*, 41(6):401–424, November 2000.
- [120] Stephen M. Barnett and Sarah Croke. Quantum state discrimination. *Adv. Opt. Photon.*, 1(2):238–278, April 2009.
- [121] I.D. Ivanovic. How to differentiate between non-orthogonal states. *Physics Letters A*, 123(6):257 – 259, August 1987.
- [122] D. Dieks. Overlap and distinguishability of quantum states. *Physics Letters A*, 126(56):303 – 306, January 1988.
- [123] Asher Peres. How to differentiate between non-orthogonal states. *Physics Letters A*, 128(12):19, March 1988.

- [124] Man-Duen Choi. A schwarz inequality for positive linear maps on c^* -algebras. *Illinois J. Math.*, 18(4):565–574, December 1974.
- [125] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285 – 290, June 1975.
- [126] M. Keyl and R.F. Werner. Channels and maps. In *Lectures on Quantum Information*, pages 73–86. Wiley-VCH, Weinheim, Germany, 2007.
- [127] W. Forrest Stinespring. Positive functions on c^* -algebras. *Proceedings of the American Mathematical Society*, 6(2):pp. 211–216, April 1955.
- [128] Lev Vaidman, Yakir Aharonov, and David Z. Albert. How to ascertain the values of σ_x , σ_y , and σ_z of a spin- $1/2$ particle. *Phys. Rev. Lett.*, 58:1385–1387, April 1987.
- [129] Onur Hosten and Paul Kwiat. Observation of the spin hall effect of light via weak measurements. *Science*, 319(5864):787–790, 2008.
- [130] P. Ben Dixon, David J. Starling, Andrew N. Jordan, and John C. Howell. Ultra-sensitive beam deflection measurement via interferometric weak value amplification. *Phys. Rev. Lett.*, 102:173601, April 2009.
- [131] David J. Starling, P. Ben Dixon, Andrew N. Jordan, and John C. Howell. Optimizing the signal-to-noise ratio of a beam-deflection measurement with interferometric weak values. *Phys. Rev. A*, 80:041803, October 2009.
- [132] Nathan S. Williams and Andrew N. Jordan. Weak values and the leggett-garg inequality in solid-state qubits. *Phys. Rev. Lett.*, 100:026804, January 2008.
- [133] Yakir Aharonov, Alonso Botero, Sandu Popescu, Benni Reznik, and Jeff Tollaksen. Revisiting hardy’s paradox: counterfactual statements, real measurements, entanglement and weak values. *Physics Letters A*, 301(34):130 – 138, 2002.
- [134] Jeffrey S. Lundeen and Aephraim M. Steinberg. Experimental joint weak measurement on a photon pair as a probe of hardy’s paradox. *Phys. Rev. Lett.*, 102:020404, January 2009.
- [135] Christoph Simon and Eugene S. Polzik. Fock-state view of weak-value measurements and implementation with photons and atomic ensembles. *Phys. Rev. A*, 83:040101, April 2011.
- [136] Amir Feizpour, Xingxing Xing, and Aephraim M. Steinberg. Amplifying single-photon nonlinearity using weak measurements. *Phys. Rev. Lett.*, 107:133603, September 2011.

- [137] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan. The sense in which a "weak measurement" of a spin- $\frac{1}{2}$ particle's spin component yields a value 100. *Phys. Rev. D*, 40:2112–2117, September 1989.
- [138] Paul Busch. Surprising features of unsharp quantum measurements. *Physics Letters A*, 130(67):323 – 329, 1988.
- [139] Wayne Hu. The curious quantum mechanics of pre- and post-selected ensembles. *Foundations of Physics*, 20:447–458, 1990. 10.1007/BF00731712.
- [140] A. J. Leggett. Comment on "how the result of a measurement of a component of the spin of a spin-(1/2) particle can turn out to be 100". *Phys. Rev. Lett.*, 62:2325–2325, May 1989.
- [141] Shengjun Wu and Yang Li. Weak measurements beyond the aharonov-albert-vaidman formalism. *Phys. Rev. A*, 83:052106, May 2011.
- [142] Nicolas Brunner and Christoph Simon. Measuring small longitudinal phase shifts: Weak measurements or standard interferometry? *Phys. Rev. Lett.*, 105:010405, July 2010.
- [143] Aabid Patel Jeff S. Lundeen, Brandon Sutherland, Corey Stewart, and Charles Bamber. Direct measurement of the quantum wavefunction. *Nature*, 474:188–191, June 2011.
- [144] S Chaturvedi, E Ercolessi, G Marmo, G Morandi, N Mukunda, and R Simon. Wignerweyl correspondence in quantum mechanics for continuous and discrete systemsa dirac-inspired view. *Journal of Physics A: Mathematical and General*, 39(6):1405, 2006.
- [145] Ping-Xing Chen, János A. Bergou, Shi-Yao Zhu, and Guang-Can Guo. Ancilla dimensions needed to carry out positive-operator-valued measurement. *Phys. Rev. A*, 76:060303, December 2007.
- [146] Guoming Wang and Mingsheng Ying. Realization of positive-operator-valued measures by projective measurements without introducing ancillary dimensions. *eprint arXiv:quant-ph/0608235*, August 2006.
- [147] B. Kraus and J. I. Cirac. Optimal creation of entanglement using a two-qubit gate. *Phys. Rev. A*, 63:062309, May 2001.
- [148] Jun Zhang, Jiri Vala, Shankar Sastry, and K. Birgitta Whaley. Geometric theory of nonlocal two-qubit operations. *Phys. Rev. A*, 67:042313, April 2003.

- [149] A. T. Rezakhani. Characterization of two-qubit perfect entanglers. *Phys. Rev. A*, 70:052313, November 2004.
- [150] L.J. Landau and R.F. Streater. On Birkhoff's theorem for doubly stochastic completely positive maps of matrix algebras. *Linear Algebra and its Applications*, 193(0):107 – 127, November 1993.
- [151] Koenraad M R Audenaert and Stefan Scheel. On random unitary channels. *New Journal of Physics*, 10(2):023011, February 2008.
- [152] Yuan Liang Lim, Sean D. Barrett, Almut Beige, Pieter Kok, and Leong Chuan Kwek. Repeat-until-success quantum computing using stationary and flying qubits. *Phys. Rev. A*, 73:012304, January 2006.
- [153] J. Eisert, K. Jacobs, P. Papadopoulos, and M. B. Plenio. Optimal local implementation of nonlocal quantum gates. *Phys. Rev. A*, 62:052317, October 2000.
- [154] Martin Aigner and Günter M. Ziegler. *Proofs from The Book*, page 31. Springer-Verlag, Berlin, Germany, 1998.
- [155] Eric W. Weisstein. Geometric distribution. <http://mathworld.wolfram.com/GeometricDistribution.html>, Accessed on 13/04/12. From MathWorld—A Wolfram Web Resource.
- [156] Panos Aliferis, Daniel Gottesman, and John Preskill. Quantum accuracy threshold for concatenated distance-3 codes. *eprint arXiv:quant-ph/0504218*, April 2005.
- [157] Michael A. Nielsen and Isaac L. Chuang. The Solovay–Kitaev theorem. In *Quantum Computation and Quantum Information*, pages 617–624. Cambridge University Press, Cambridge, United Kingdom, 2000.
- [158] Yuan Liang Lim, Almut Beige, and Leong Chuan Kwek. Repeat-until-success linear optics distributed quantum computing. *Phys. Rev. Lett.*, 95:030505, July 2005.
- [159] D. Gross, K. Kieling, and J. Eisert. Potential and limits to cluster-state quantum computing using probabilistic gates. *Phys. Rev. A*, 74:042343, October 2006.
- [160] J. I. Cirac, A. K. Ekert, S. F. Huelga, and C. Macchiavello. Distributed quantum computation over noisy channels. *Phys. Rev. A*, 59:4249–4254, June 1999.
- [161] A. Broadbent, J. Fitzsimons, and E. Kashefi. Universal blind quantum computation. In *Foundations of Computer Science, 2009. FOCS '09. 50th Annual IEEE Symposium on*, pages 517–526, October 2009.

- [162] Takahiro Sueki, Takeshi Koshihara, and Tomoyuki Morimae. Ancilla-driven universal blind quantum computation. *Phys. Rev. A*, 87:060301, June 2013.
- [163] Persi Diaconis. Chapter 3: Random walks on groups. In Shanti S. Gupta, editor, *Group representations in probability and statistics*, volume 11 of *Lecture Notes–Monograph Series*, pages 17–68. Institute of Mathematical Statistics, Hayward, CA, 1988.
- [164] Jeffrey S. Rosenthal. Random walks on discrete and continuous circles. *Journal of Applied Probability*, 30(4):780–789, December 1993.
- [165] Carl Dou and Martin Hildebrand. Enumeration and random random walks on finite groups. *Ann. Probab.*, 24(2):987–1000, April 1996.
- [166] Martin Hildebrand. A survey of results on random random walks on finite groups. *Probab. Surveys*, 2:33–63, 2005.
- [167] Tomoyuki Morimae and Jonas Kahn. Entanglement-fidelity relations for inaccurate ancilla-driven quantum computation. *Phys. Rev. A*, 82:052314, November 2010.
- [168] Naomi H. Nickerson, Ying Li, and Simon C. Benjamin. Topological quantum computing with a very noisy network and local error rates approaching one percent. *Nat Commun*, 4(1756), April 2013.
- [169] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53:2046–2052, April 1996.
- [170] R. Mir, Jeffrey S Lundeen, Morgan W. Mitchell, Aephraim M. Steinberg, Joshua L Garretson, and Howard M. Wiseman. A double-slit ‘which-way’ experiment on the complementarityuncertainty debate. *New Journal of Physics*, 9(8):287, August 2007.
- [171] J. S. Lundeen and A. M. Steinberg. Experimental joint weak measurement on a photon pair as a probe of hardy’s paradox. *Phys. Rev. Lett.*, 102:020404, January 2009.
- [172] Lee A. Rozema, Ardavan Darabi, Dylan H. Mahler, Alex Hayat, Yasaman Soudagar, and Aephraim M. Steinberg. Violation of heisenberg’s measurement-disturbance relationship by weak measurements. *Phys. Rev. Lett.*, 109:100404, September 2012.

- [173] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*, chapter 4, page 176. Cambridge University Press, Cambridge, United Kingdom, 2000.
- [174] Sheldon M. Ross. *Introduction to probability and statistics for engineers and scientists*, chapter 7, page 267. Elsevier Academic Press, Boston, 2009.
- [175] Wolfram Research Inc. Find root. <https://reference.wolfram.com/language/ref/FindRoot.html>, (accessed Dec 2015).
- [176] Wolfram Research Inc. Find maximum. <https://reference.wolfram.com/language/ref/FindMaximum.html>, (accessed Dec 2015).
- [177] Wolfram Research Inc. Find maximum. <https://reference.wolfram.com/language/tutorial/ConstrainedOptimizationLocalNumerical.html#interiorpointalgorithm>, (accessed Dec 2015).