UNIVERSITY of CALIFORNIA
Santa Barbara

Generation of Three-Qubit Entanglement Using Josephson Phase Qubits

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Physics

by

Matthew Gary Neeley

Committee in charge:

Professor John M. Martinis, Chair
Professor Andrew N. Cleland
Professor Dirk Bouwmeester

December 2010

The dissertation of Matthew Gary Neeley is approved:

Professor Andrew N. Cleland

Professor Dirk Bouwmeester

Professor John M. Martinis, Chair

October 2010

To Mom, Dad, Payjo, Dan the Man, Mia and Jess

Acknowledgements

There are a number of individuals whose support and encouragement during my graduate career have made this work possible. First and foremost, I must acknowledge Professor John Martinis for his guidance and leadership, and for the incredible opportunity to work in his lab. I have learned a tremendous amount, and it has been an amazingly productive and exciting few years.

Of course life in the Martinis group would not have been what it was without the other students and post-docs I've had the pleasure to work with there. Markus Ansmann paved the way for what will no doubt be many graduate students to come, and I'm sure his influence will be felt in the group for a long time, not least because his many contributions to the software infrastructure of the group continue to hum along bug-free. His thesis also paved the way for mine, and if you haven't read it yet, you should.

To Erik Lucero I owe a huge debt of gratitude for reading this work in detail and telling me where I was being high, but even more for his friendship and for countless excellent meals which I'll certainly never be able to repay.

Radek Bialczak and Aaron O'Connell have been superb housemates and sometime business partners, respectively, as well as officemates both of them. They've been a pleasure to work with and an inspiration.

Without Daniel Sank's good cheer, insatiable curiosity and seemingly neverending supply of brainteasers, life in the lab would not have been nearly as enjoyable. He's like a generating function of awesomeness.

The next generation of grad students and postdocs in the Martinis group have been helpful and encouraging, but more than that it is inspiring to see them pick up where I'll be leaving off. I can't wait to see what comes next.

My family deserve more thanks than I could ever give for their unfailing support and generosity throughout my graduate career. I would not be who nor where I am without them, their constant love and patience, and their excellent examples.

Finally, how can I even begin to thank Stacie Furia, Ph.D. (!), for her support and love during this time in graduate school? She gave me more than I deserved and put up with more than she deserved. It's hard to imagine life without her, but it certainly wouldn't be as fun.

Curriculum Vitæ

Matthew Gary Neeley

Education

2010 (exp.)	Doctor of Philosophy, Physics, University of California, Santa Barbara
2007	Master of Arts, Physics, University of California, Santa Barbara
2004	Bachelor of Science, Physics, Stanford University

Honors and Awards

2009-2010	Dean's Fellowship for Graduate Education, University of California, Santa Barbara, California
2005	John Candy Award for Excellence in first-year courses, University of California, Santa Barbara, California
1999	President's Award for Academic Excellence, Stanford University
1999	Booth Writing Prize, Stanford University

Publications

M. Neeley, R. C. Bialczak, M. Lenander, E. Lucero, M. Mariantoni, A. D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, Y. Yin, T. Yamamoto, A. N. Cleland, J. M. Martinis. Generation of Three-Qubit Entangled States using Superconducting Phase Qubits. *Nature* **467**, 570–573 (2010).

R. C. Bialczak, M. Ansmann, M. Hofheinz, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, H. Wang, J. Wenner, M. Steffen, A. N. Cleland, J. M. Martinis. Quantum Process Tomography of a Universal Entangling Gate Implemented with Josephson Phase Qubits. *Nature Physics* **6**, 409–413 (2010).

A. D. O'Connell, M. Hofheinz, M. Ansmann, R. C. Bialczak, M. Lenander, E. Lucero, M. Neeley, D. Sank, H. Wang, M. Weides, J. Wenner, J. M. Martinis,

A. N. Cleland. Quantum ground state and single-phonon control of a mechanical resonator. *Nature* **464**, 697–703 (2010).

H. Wang, M. Hofheinz, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, A. N. Cleland, J. M. Martinis Decoherence Dynamics of Complex Photon States in a Superconducting Circuit. *Physical Review Letters* **103**, 3200404 (2009).

H. Wang, M. Hofheinz, J. Wenner, M. Ansmann, R. C. Bialczak, M. Lenander, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, A. N. Cleland, J. M. Martinis. Improving the Coherence Time of Superconducting Coplanar Resonators. *Applied Physics Letters* **95**, 233508 (2009).

M. Ansmann, H. Wang, R. C. Bialczak, M. Hofheinz, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, A. N. Cleland, J. M. Martinis. Violation of Bell's inequality in Josephson phase qubits. *Nature* **461**, 504–506 (2009).

M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, E. Lucero, A. D. O'Connell, D. Sank, H. Wang, J. Wenner, A. N. Cleland, M. R. Geller, J. M. Martinis. Emulation of a Quantum Spin with a Superconducting Phase Qudit. *Science* **325**, 722 (2009).

M. Hofheinz, H. Wang, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, J. Wenner, J. M. Martinis, A. N. Cleland. Synthesizing arbitrary quantum states in a superconducting resonator. *Nature* **459**, 546–549 (2009).

C. Song, T. W. Heitmann, M. P. DeFeo, K. Yu, R. McDermott, M. Neeley, J. M. Martinis, B. L. T. Plourde. Microwave response of vortices in superconducting thin films of Re and Al. *Physical Review B* **79**, 174512 (2009).

H. Wang, M. Hofheinz, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, J. Wenner, A. N. Cleland, J. M. Martinis. Measurement of the decay of Fock states in a superconducting quantum circuit. *Physical Review Letters* **101**, 240401 (2008).

M. Hofheinz, E. M. Weig, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, H. Wang, J. M. Martinis, A. N. Cleland. Generation of Fock states in a superconducting quantum circuit. *Nature* **454**, 310–314 (2008).

N. Katz, M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, E. Lucero, A. D. O'Connell, H. Wang, A. N. Cleland, J. M. Martinis, A. N. Korotkov. Reversal of the Weak Measurement of a Quantum State in a Superconducting Phase Qubit. *Physical Review Letters* **101**, 200401 (2008).

M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, N. Katz, E. Lucero, A. D. O'Connell, H. Wang, A. N. Cleland, J. M. Martinis. Process tomography of quantum memory in a Josephson-phase qubit coupled to a two-level state. *Nature Physics* **4**, 523–526 (2008).

A. D. O'Connell, M. Ansmann, R. C. Bialczak, M. Hofheinz, N. Katz, E. Lucero, C. McKenney, M. Neeley, H. Wang, E. M. Weig, A. N. Cleland, J. M. Martinis. Microwave Dielectric Loss at Single Photon Energies and milliKelvin Temperatures. *Applied Physics Letters* **92**, 112903 (2008).

E. Lucero, M. Hofheinz, M. Ansmann, R. C. Bialczak, N. Katz, M. Neeley, A. D. O'Connell, H. Wang, A. N. Cleland, J. M. Martinis. High-fidelity gates in a Josephson qubit. *Physical Review Letters* **100**, 247001 (2008).

M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, N. Katz, E. Lucero, A. D. O'Connell, H. Wang, A. N. Cleland, J. M. Martinis. Transformed Dissipation in Superconducting Quantum Circuits. *Physical Review B* **77**, 180508(R) (2008).

R. C. Bialczak, R. McDermott, M. Ansmann, M. Hofheinz, N. Katz, E. Lucero, M. Neeley, A. D. O'Connell, H. Wang, A. N. Cleland, J. M. Martinis. 1/f Flux Noise in Josephson Phase Qubits. *Physical Review Letters* **99**, 187006 (2007).

M. Steffen, M. Ansmann, R. C. Bialczak, N. Katz, E. Lucero, R. McDermott, M. Neeley, E. M. Weig, A. N. Cleland, J. M. Martinis. Measurement of the Entanglement of Two Superconducting Qubits via State Tomography. *Science* **313**, 1423–1425 (2006).

N. Katz, M. Ansmann, R. C. Bialczak, E. Lucero, R. McDermott, M. Neeley, M. Steffen, E. M. Weig, A. N. Cleland, J. M. Martinis, A. N. Korotkov. Coherent state evolution in a superconducting qubit from partial-collapse measurement. *Science* **312**, 1498–1500 (2006).

M. Steffen, M. Ansmann, R. McDermott, N. Katz, R. C. Bialczak, E. Lucero, M. Neeley, E. M. Weig, A. N. Cleland, J. M. Martinis. State tomography of capacitively shunted phase qubits with high fidelity. *Physical Review Letters* **97**, 050502 (2006).

Abstract

Generation of Three-Qubit Entanglement Using Josephson Phase Qubits

by

Matthew Gary Neeley

Phase qubits are a type of superconducting quantum circuit that have numerous applications in quantum computing and quantum information processing. One of the hoped-for advantages of these devices is their potential to be easily scaled up to larger numbers of qubits due to the reliance on conventional microfabrication techniques and the ease with which they can be coupled together with simple wiring.

We describe an experiment in which three phase qubits are coupled together and used to form three-qubit entangled states of two types, namely $|\mathrm{GHZ}\rangle \equiv (|000\rangle + |111\rangle)/\sqrt{2}$ and $|\mathrm{W}\rangle \equiv (|001\rangle + |010\rangle + |100\rangle)/\sqrt{3}$. The physics of three-qubit entanglement is interesting precisely because of the existence of these two inequivalent types of entangled states, unlike the better-known two-qubit case where there is only one type of entanglement. We show that both of the experimentally-created states exhibit genuine three-qubit entanglement, and that the $|\mathrm{GHZ}\rangle$ state

further violates the Mermin-Bell inequality, indicating that the observed measurement correlations cannot be explained by a hidden-variable model.

From a technological standpoint as well, this experiment is interesting because it represents another step along the path of scaling toward larger quantum information processing devices made with superconducting qubits. Significantly, our work was achieved using "off-the-shelf" qubit designs that were coupled together with a simple circuit consisting solely of passive capacitor elements. This underscores the power of the circuit approach and the modularity it enables.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The impact of computing and information technology on our everyday lives is profound and undeniable. Physics too has been influenced by the revolutions of the information age; computers are physical machines after all, and so their computations proceed according to natural laws. Understanding computing as a physical process has led to insights in both directions between computer science and physics, including an emerging view of the very foundational laws of physics as being driven by information. Quantum mechanics of course plays a role in this story because it provides a basic understanding of the laws of the universe; we live in a quantum world, so any computer one might hope to build must of necessity be a quantum computer.

The term "quantum computer", however, is more specifically reserved for those

devices that actually exploit their quantum nature so as to outperform the so-called "classical" computers, which would operate just the same in a world without quantum mechanics. Understanding exactly what advantage quantum effects confer on a quantum computer is an ongoing theoretical question; trying to build such a device to actually demonstrate and exploit these effects is an ongoing experimental problem.

In this thesis we will be concerned with the experimental problem of building a quantum computer[29, 17], and in particular doing so with superconducting quantum circuits of the "phase qubit" variety. Much as classical computers are built up from collections of many simpler components (transistor switches or "bits"), so too a quantum computer would be built up from many quantum bits or "qubits". Superconducting qubits are a family of quantum circuits that could be used as the component parts in a quantum computer. These devices share some of the advantages of the silicon bits in classical computers, namely that they are electrical circuits made on chips using conventional microfabrication techniques, which allows for "easy" scaling by packing more components together on a chip. Of course achieving this scaling with quantum bits of any stripe is an immense challenge, but the first steps are now being taken with superconducting qubits.

Many beautiful experiments have been done with two coupled superconducting qubits[3, 8]. In this thesis we will describe a modest next step, an experiment

in which *three* superconducting phase qubits have been coupled together and controlled to produce entangled states. Entanglement is one of the bizarre features of quantum mechanics, and one of the key resources required for quantum computation, so experimentally creating and measuring entangled states is of crucial importance in the various physical implementations of a quantum computer. Unlike the now well-established two-qubit case where there is only one type of entanglement, three qubits can be entangled in two fundamentally different ways[9], typified by the states $|\mathrm{GHZ}\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$ and $|\mathrm{W}\rangle = (|001\rangle + |010\rangle + |100\rangle)/\sqrt{3}$. We demonstrate the generation of both of these states and show that they meet the criteria for genuine three-qubit entanglement.

This thesis will proceed as follows: after this brief introductory chapter, we turn in Chapter 2 to a description of our superconducting phase qubits and how they are designed, fabricated and operated. Following that, in Chapter 3 we will describe in detail the process by which we go about translating a quantum algorithm into an experimental control sequence that will be applied to the qubits to perform the desired computation; this chapter will potentially be of interest and hopefully of some use to those implementing future experiments on similar systems.

With that introductory material out of the way, the remaining chapters describe our experiment with three coupled phase qubits. We discuss the details of

three-qubit entangled states in Chapter 4, followed by a description of the protocols we will use to generate these states in Chapter 5 and the crucial coupling circuit which generates the entanglement in Chapter 6. We then discuss the subtleties of multi-qubit measurement in our system in Chapter 7, and finally present the experimental results in Chapter 8, in particular the demonstration that the created states are genuine three-qubit entangled states that cannot be explained as mixtures of two-qubit entanglement. This capability to build and operate a multi-qubit device, and demonstrate its truly quantum behavior, represents a step toward scaling to larger numbers of qubits, and eventually a full-blown quantum computer.

In all of this work we are preceded by and greatly indebted to Markus Ansmann whose thesis[2] contains a great deal of information about the UCSB phase qubit, how it is designed, fabricated, operated and measured, how our software control system works, *etc.* Where these issues are discussed here, particularly in Chapter 2, they will be touched on briefly with a great deal of otherwise necessary background material omitted and the curious reader referred to the appropriate places in his work.

1.1 The Qubit and the Bloch Sphere

The simplest nontrivial quantum system is one with just two states. By convention we use the first two non-negative integers as labels and call theses states $|0\rangle$ and $|1\rangle$. By analogy with the classical binary digit or "bit" which can take on the values 0 or 1, such a two-level system is known as a quantum bit or "qubit". The complete state of the system is not simply $|0\rangle$ *or* $|1\rangle$ as for a classical bit, rather the state must be specified as $\alpha|0\rangle + \beta|1\rangle$ where the complex coefficients α and β are the so-called "amplitudes" for the system to be in these two eigenstates. These coefficients satisfy the normalization condition $|\alpha|^2 + |\beta|^2 = 1$ and can thus be interpreted as probabilities: $P_0 = |\alpha|^2$ is the probability that upon measurement the system will be found to be in state $|0\rangle$, while $P_1 = |\beta|^2$ is the probability that the system will be found in state $|1\rangle$.

Because of the normalization constraint $|\alpha|^2 + |\beta|^2 = 1$, we can write an arbitrary qubit state as $e^{i\phi_\alpha}\cos(\theta)|0\rangle + e^{i\phi_\beta}\sin(\theta)|1\rangle$, where the angle θ can vary between 0 and π and parametrizes relative magnitude of the two amplitudes. Then, because global phases are unobservable in quantum mechanics, we can specify without loss of generality that the amplitude for $|0\rangle$ is real, leaving only a relative phase on the other amplitude so that the state can be written as $\cos(\theta)|0\rangle + e^{i\phi}\sin(\theta)|1\rangle$, where the relative phase ϕ can vary from 0 to 2π.

This parametrization in terms of polar angle θ and azimuthal angle ϕ puts

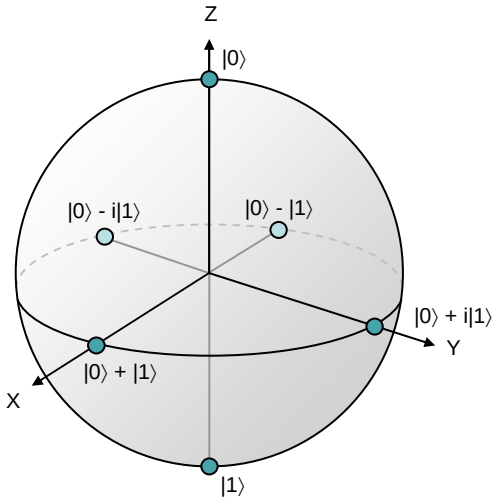

Figure 1.1: The Bloch sphere.

the qubit states in one-to-one correspondence with the points of a unit sphere known as the Bloch sphere, as shown in Figure 1.1. Transformations produced by quantum dynamics correspond to rotations of this Bloch sphere. The simplest rotations to describe are those about the x-, y- and z-axes of the Bloch sphere, which are generated by the Pauli matrices

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.1}$$

For example, a rotation about X by angle θ is given by $R(\theta) = \exp(-(\theta/2)X)$. This is a unitary matrix, as are all quantum operations on pure states. When an operation U is applied to the system starting in state $|\psi\rangle$, the final state is $U|\psi\rangle$. We can build up *any* rotation U from rotations about just two axes, say Y and Z, giving complete control over the qubit state, which is an important requirement

for quantum computation.

In any experimental situation, our qubit will be subject to decoherence, so that the "pure state" description of the system as being in a superposition $\alpha|0\rangle + \beta|1\rangle$ is insufficient. Instead, we can adopt the density matrix formalism which allows for a quantum system to be in a mixed state, a probabilistic mixture of different pure states. The density matrix ρ is defined as $\sum_k p_k |\psi_k\rangle \langle\psi_k|$, where p_k is the probability that the system is in the (pure) state $|\psi_k\rangle$. In terms of the Bloch sphere picture, the effect of allowing mixed states is to 'fill in' the Bloch sphere, with points on the surface corresponding to pure states, and points in the interior corresponding to mixed states.

1.2 Combining Qubits

A single isolated qubit is not particularly exciting; truly interesting things start to happen when many qubits are combined together, as we would like to do in building a universal quantum computer. With N qubits, the allowed states can be labelled by N-digit binary numbers with 0s and 1s corresponding to the states of the individual qubits: $|00\ldots00\rangle$, $|00\ldots01\rangle$, $|00\ldots10\rangle$ and so on up to $|11\ldots11\rangle$. There are a total of 2^N such eigenstates and the complete description of the quantum system requires that we specify amplitudes for each of them. This exponentially-large amount of information carried by the multi-qubit states is part

of what gives a quantum computer its potential power.

In order to make our quantum computer universal we would again like to be able to implement any possible unitary operation in the now 2^N-dimensional space in which the state lives. The ability to perform single-qubit rotations as we described in the previous section is no longer enough. In addition, we require at least one operation that can be performed on two qubits at the same time. Physically, this means that the qubits must be coupled together in such a way that they can interact. This interaction must be controllable so that we can turn it on and off to perform series of one- and two-qubit operations, and hence build up the desired operation on the whole combined system. For many physical systems, coupling can be difficult to achieve and to control; superconducting qubits have a somewhat easier time because as circuits they can simply be wired together. We will see in Chapter 2 some of the numerous possibilities for coupling phase qubits.

1.3 Decoherence

Quantum superposition and entanglement are fragile, particularly when these exist in solid state systems such as a superconducting qubit that is tightly coupled to numerous environmental degrees of freedom. Decoherence refers collectively to the various processes by which the coupling to the environment causes the quantum state to degrade, changing a superposition state into a probabilistic

mixed state in a process known as dephasing; or causing an excited state to give off energy and relax to the ground state in a process known, not surprisingly, as relaxation. In the well-known language of NMR where we consider a two-level system (qubit or spin-1/2), these processes are each governed by exponential decay times T_1, which sets the relaxation time, and T_2, which sets the dephasing time.

For experimentalists, decoherence is perhaps the single most important and troublesome issue to grapple with in trying to realize quantum information processing devices. This is because by definition one would like to create and manipulate quantum states, but in practice the states decay quickly due to decoherence, severely limiting the complexity of states that can be created and the number of manipulations that can be performed. Each physical implementation has its own set of problems related to decoherence. Superconducting qubits are no exception, and in fact because these are large devices coupled strongly to noisy environments with many degrees of freedom, decoherence is a particularly difficult problem.

A great deal of effort has gone into understanding decoherence in superconducting qubits and working to reduce it. We have greatly benefited from this work, and the present state of superconducting qubit performance is good enough for our experiment so that no special effort had to be made to get three qubits to work together. The decoherence we observe in the coupled system is completely due to decoherence of the individual qubits. Thus, ongoing efforts to improve

the performance of single qubits will immediately benefit multi-qubit experiments such as ours.

1.4 Superposition and Entanglement

Superposition and entanglement are two of the most striking features of quantum mechanics. Superposition, as we have seen above, refers to the fact that a physical system at any particular time must be described by its amplitude to be in each of a set of possible eigenstates. These amplitudes are complex numbers whose squared length corresponds to the probability of finding the system in that particular state. When a system is measured, it will be found in one of these eigenstates, but before measurement the entire superposition and the information contained in all the state amplitudes govern the dynamics.

A state is said to be entangled if it cannot be described as a product state, which essentially means that the states of the various constituent systems are linked together. This can lead to interesting results if measurements are performed on the constituent systems, because correlations can be observed that are stronger than any possible classical theory can explain. This is the key to the two-qubit Bell Inequalities, and their various generalization to three qubits, such as the Mermin-Bell inequality, which we discuss in Chapter 4. Entanglement can be considered to be a resource which quantum computers utilize to achieve their

speedups relative to classical computers. Because of this, the measurement and characterization of entanglement is important for any physical system which could be used for quantum information processing.

Chapter 2

Flux-Biased Phase Qubits

Among the many candidate physical systems that might be used as quantum information processing devices, superconducting qubits have emerged in recent years as promising candidates. Rather than relying on fundamental, naturally-occurring quantum systems such as spins, atoms, or photons, these devices are engineered circuits that consist of many constituent atoms, but that exhibit collective quantum behavior which makes them usable as qubits. The two key features are superconductivity, which is a collective quantum behavior of many electrons that allows the entire circuit to be treated quantum mechanically, and the Josephson effect, which gives the strong nonlinearity required to make an effective two-level system or qubit.

The qubits used in our work are phase qubits; a schematic of the circuit is

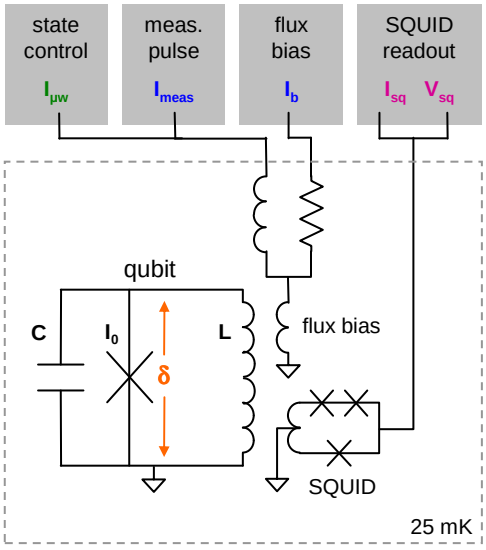

Figure 2.1: Schematic of a phase qubit.

shown in Figure 2.1. The circuit consists of a Josephson junction shunted by an inductor and, typically, an external capacitor as well. The circuit is designed so that the potential energy as a function of the phase difference δ across the junction is a double-well potential, as illustrated in Figure 2.2. The potential is tilted by applying flux to the qubit loop to bias the circuit near the critical current of the junction; this leads to a highly asymmetric potential where one well becomes very shallow and is close to disappearing altogether. This shallower left-hand well serves as the qubit well. More information about this potential energy description of a quantum circuit is found in [2, chap. 2].

Zooming in on just the left-hand qubit well, as in the right hand side of Figure 2.2, we see that the potential well is asymmetric, a nearly-harmonic potential

with a cubic perturbation. The quantum states of the system in this well are hence similar to a perturbed harmonic oscillator. The energy scale $E_{10} \approx 6\,\mathrm{GHz}$ of the oscillator is set by the curvature of the potential due to the inductance, and the "mass" of the system due to the capacitance. The finite barrier on the right side of the well is due to the critical current of the junction, but can be modulated by tilting the potential with the flux bias. Typically when operating a qubit, the bias is adjusted to give a barrier height $\Delta U \sim 5E_{10}$, so that on the order of 5 quantum levels exist in the well.

Due to the cubic perturbation, these levels are not evenly spaced, that is, the energy difference E_{21} between the first and second excited states is different from the qubit energy E_{10} between the ground state and first excited state. This nonlinearity $\Delta \equiv (E_{21} - E_{10})/\hbar$ is a crucial feature of the circuit that allows the system to be operated as an effective two-level system. For typical operating parameters the nonlinearity is on the order of $-2\pi \cdot 200\,\mathrm{MHz}$, small but sufficient for reasonably fast operation. Microwaves can be applied at the qubit frequency $\omega_{10} = E_{10}/\hbar$ to cause transitions between qubit levels. In addition, the flux bias can be adjusted to tilt the potential, causing Z-rotations. Unlike some other types of superconducting qubits, the phase qubit does not have an accessible degeneracy point at which the qubit frequency becomes insensitive to changes in flux bias. This generally results in somewhat reduced phase coherence compared to other

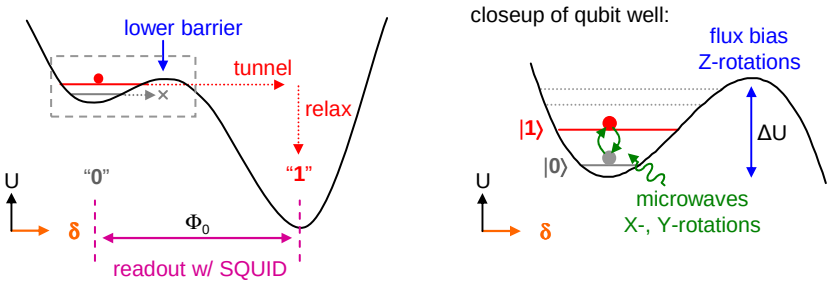

Figure 2.2: Potential energy landscape of a phase qubit.

qubit types. However, phase qubits have a number of other features which are nice advantages, particularly the ease with which they can be coupled together and their relative insensitivity to stray capacitances, both of which are due to their low impedance.

2.1 Single-Qubit Design

A number of considerations must be taken into account to design a phase qubit, which we can broadly divide into three categories: the design of the qubit circuit itself; the design of the measurement circuit; and the "design" of the electromagnetic environment of the qubit to appropriately decouple it from the environment and prevent decoherence. In practice, of course, these three categories are not so cleanly separated and must be considered together. For example, the measurement circuit forms an important part of the qubit's electromagnetic environment, so the effects of the measurement circuit on qubit coherence must be very carefully

considered. The phase qubit designs which we will describe occupy something of a sweet spot where the various design concerns are reasonably well optimized, but it is almost certainly not a global optimum, so there is room for future design improvements. We will describe the basic considerations that went into the current design; additional information is given in [2, chap. 4].

2.1.1 The Qubit Circuit

To design the phase qubit circuit a number of factors are important. Near the critical current, where the device will be biased, the qubit frequency is primarily determined by the junction critical current I_0 and the shunt capacitance C_q, with the inductance L_q and flux bias essentially playing the role of current bias in the tilted washboard model but having very little effect on the actual shape of the qubit potential. To achieve a qubit frequency of $\sim 6\,\mathrm{GHz}$ we take $I_0 \approx 2\,\mu\mathrm{A}$ and $C_q \approx 1\,\mathrm{pF}$. Having chosen I_0 and C_q, the inductance L_q can then be determined by the overall requirement of having two wells in the qubit potential. To achieve the desired potential shape, we choose $L_q \approx 720\,\mathrm{pH}$.

If the inductance L_q were increased, the qubit circuit could be made less sensitive to magnetic flux noise in the qubit loop, which could be a desirable feature. At the same time, however, this would create a potential with more than two wells, which makes it difficult to reset the qubit. Resetting requires that we re-

turn the qubit state to the original well, and doing so is difficult when there are more than two wells. This is one of the many trade-offs that must be considered when designing these circuits, and the full parameter space of different designs has certainly not been fully explored.

Having chosen these component values, we must next lay out a circuit design that will achieve the desired component parameters while being possible to fabricate. Our design is shown in Figure 2.3. A few things to note about this design: first, the inductor loop is designed as a gradiometer to reduce its sensitivity to external flux noise (areas of positive and negative winding are carefully balanced to reduce this sensitivity as much as possible). Second, the shunt capacitor is a parallel-plate design that requires an intermediate layer of dielectric material. Losses due to dielectrics in capacitors such as this have been shown to be the dominant source of relaxation in the phase qubit[20], so a great deal of work has gone into the fabrication of this element, which for which we used hydrogenated amorphous Silicon (a-Si:H) dielectric. Third, the junction is made up of a small triangular overlap between two aluminum layers, which allows a small area of $\sim 1\,\mu\mathrm{m}^2$ to be achieved with our optical lithography capabilities. Two-level state defects in the junction have been found to couple strongly to the qubit[33, 26], producing spectroscopic splittings as well as coherent time-domain oscillations of the qubit state. The number of such defects can be reduced by minimizing the

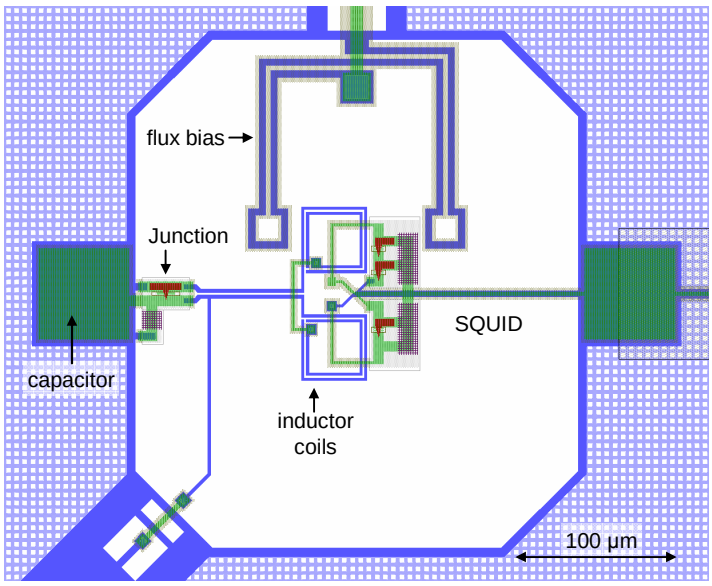

Figure 2.3: Layout of phase qubit circuit.

junction area, which this triangular overlap design allows.

2.1.2 The Measurement Circuit

In order to extract information from our qubits, we must be able to measure whether they are in state $|0\rangle$ or $|1\rangle$. This is accomplished in two stages, which we refer to as "measurement" and "readout". In the measurement stage, a few ns-long pulse is applied to the flux bias line to tilt the qubit potential and briefly lower the barrier in the cubic potential well. The $|1\rangle$ state, which sits at higher energy and hence sees a narrower barrier than the $|0\rangle$ state, tunnels out of the qubit well into the neighboring well. After some time the tunnelled state relaxes to the ground state of the neighboring well, emitting a large amount of energy

in the process, typically on the order of several hundred photons. Because this tunnelling and relaxation process is rather violent, it is unlikely that any quantum coherence that existed in the original qubit state survives, rather the system will be projected into either the left well ("0") or right well ("1") and left in one of these two states. The measurement process is illustrated schematically in the left panel of Figure 2.2, and described in much greater detail in Chapter 7.

At this point the qubit state has been measured, and it remains only for us to readout which final state was chosen. This is rather easy to accomplish due to the fact that the final states at the bottom of the left and right wells produce fluxes through the qubit loop that differ by a flux quantum Φ_0. This large flux difference can be measured easily with an on-chip superconducting quantum interference device (SQUID).

The SQUID used in our readout is a three-junction device with a geometric mutual inductance $M_{sq} \approx 60\,\mathrm{pH}$ between the SQUID and qubit loops. The asymmetric three-junction design allows the device to be biased in such a way that the effective mutual inductance between SQUID and qubit can be tuned by biasing the SQUID, despite the fixed geometric mutual inductance. In particular, by choosing a ratio $\alpha < 2$ between the critical currents of the one small and two large Josephson junctions in the SQUID, it is possible to tune the effective mutual to zero, thereby decoupling the SQUID and qubit; in practice we choose

$\alpha \approx 1.7$. During qubit operation, the SQUID is biased at this insensitive point to prevent fluctuations in the SQUID from causing decoherence in the qubit. Then for readout, the SQUID bias is ramped to the critical current and the mutual inductance turns on. The point at which the SQUID switches into the voltage state will depend on the final qubit state, allowing us to distinguish between qubit states by measuring the time it takes the SQUID to switch for a fixed bias ramp.

The readout SQUID is also designed as a gradiometer to reduce its sensitivity to external fluxes and to properly couple it to the gradiometric qubit loop. The SQUID-qubit mutual inductance M_{sq} is achieved by overlapping the loops, as opposed to previous designs in which the loops were simply placed next to each other. The overlap design allows us to achieve a larger M_{sq} for better readout distinguishability, while also reducing the SQUID self-inductance $L_s \approx 350\,\mathrm{pH}$. Lower self-inductance keeps the SQUID potential single-valued, resulting in more controlled switching behavior. The overlap design does slightly increase the capacitive coupling between the two circuit elements due to the crossovers, but this does not have deleterious effects on qubit performance.

The SQUID junctions are fabricated in the same process step as the qubit junctions, another improvement in this design over previous versions. This reduces the complexity and increases the reliability of the process, and was a driving force behind the redesign. Note, though, that a working SQUID requires three good

junctions to the qubit's one junction, so that readout SQUID failures are the most common failure mode for our phase qubit devices. As a result, junctionless readout schemes that rely for example on coupled resonators are an attractive alternative design, and are currently being developed.

2.1.3 Isolation from the Environment

Although the signal to be measured is large and hence easy to readout using a SQUID, care must be taken to prevent the measurement circuit from adversely affecting qubit performance. This problem was analysed in great detail in a paper and experiment that considered the effect of the SQUID on the qubit[27]. The simple technique outlined in the paper for analysing the effect of the electromagnetic environment on the qubit is to consider circuit elements as impedance transformers coupling environmental noise into the qubit circuit. The measurement SQUID's asymmetric three-junction design allows its effective mutual inductance with the qubit to be dynamically tuned by adjusting the SQUID bias current, thus avoiding decoherence during operation while still allowing strong coupling for readout. Similarly, other circuit elements that couple to the qubit must be designed so that environmental noise is appropriately transformed and will not decohere the qubit.

The other primary circuit element required for the phase qubit is the flux bias

coil which is used to bias the qubit circuit, apply measurement pulses, and, in recent designs, apply microwave pulses for manipulating the qubit state. The mutual inductance between qubit and flux bias coil is chosen to be $M_{fq} \approx 2\,\mathrm{pH}$. This gives an impedance transform such that dissipation from the $50\,\Omega$ environmental noise limits the qubit lifetime to $\sim 10\,\mu\mathrm{s}$, far from being a limiting factor with current devices. At the same time, this mutual inductance allows us to vary the flux bias by $\sim 5\Phi_0$ with the accessible range of bias currents, which gives a sufficient range to access SQUID step edges on several different branches for operating the qubit[2, chap. 8]. The flux coil is also designed as a gradiometer and the readout SQUID is placed along its axis of symmetry to eliminate coupling between the flux bias and SQUID bias. The qubit on the other hand is placed asymmetrically, allowing the flux coil to couple to the qubit as desired. All these mutual inductances were simulated using the FastHenry software package[1], and have been found to be very close to the experimentally observed values in real devices.

As was mentioned above, the SQUID-qubit crossovers lead to some capacitive coupling between these circuit elements in addition to their mutual inductance coupling. In addition, previous designs have included capacitively-coupled microwave lines for performing qubit manipulations, rather than sending these

[1] http://www.fastfieldsolvers.com/

microwave signals through the flux bias coil. In both cases, the same simple framework of impedance transformation can be used to estimate the effects of these capacitively-coupled elements on the qubit coherence, and in both cases the resulting qubit lifetimes are far greater than present limits. This tool allows the electromagnetic environment to be designed to ensure that the qubit is well-isolated from the environment and has the best-possible coherence, presently limited by internal material losses.

2.2 Scaling Up: Multiple Qubits

One of the advantageous features of the phase qubit is its low impedance. Low impedance allows the qubit circuit to be coupled to other circuit elements without being loaded, just as classical electronic sources are designed with low output impedance to allow them to drive a load impedance. In the case of coupling qubits, the qubit impedance should be low compared to the impedance of the coupling circuit, and this is quite easily achieved with the phase qubit. We will discuss briefly the approaches that have been taken to couple phase qubits, and the approach used in this experiment.

2.2.1 Capacitive coupling

The simplest possible coupling scheme is to connect together two qubits with a capacitor. Just like two classical electrical oscillators connected by a capacitance, the coupling strength, which describes the rate at which energy is transferred back and forth between the two circuits, is proportional to the coupling capacitance. This coupling works particularly well with the phase qubit because of its low impedance, as mentioned above, which allows one to run relatively long wires between the qubit and coupling capacitor without unduly loading the qubit due to stray capacitances of the coupling wire to ground. This circuit has the additional advantage of being extremely easy to design and fabricate and extremely reliable to operate.

The interaction produced by coupling two qubits with a capacitor is of the form $(\hbar g/2)(XX+YY)$ where X and Y are Pauli operators on the two qubits, and g is the coupling strength, which is proportional to the ratio of coupling capacitance to qubit capacitance $g \propto C_c/C_q$[34]. This coupling Hamiltonian can alternatively be written as $\hbar g(\sigma_+\sigma_- + \sigma_-\sigma_+)$ where σ_+ and σ_- are operators that respectively create or destroy an excitation in each qubit. From this latter form it is clear that the interaction leads to swapping of excitations back and forth between the two coupled qubits. If the interaction is applied for time $t_{\mathrm{iSWAP}} = \pi/2g$, then the qubit states are transformed as $|01\rangle \rightarrow -i\,|10\rangle$ and $|10\rangle \rightarrow -i\,|01\rangle$ while $|00\rangle$ and $|11\rangle$

are left unchanged. This so-called "iSWAP gate" is the fundamental entangling gate we will use.

It is important to note that the interaction produced by such a fixed coupling capacitor is an "always-on" interaction; the capacitance of the coupling capacitor cannot be changed dynamically to modulate the coupling strength. However, the coupling is a resonant interaction that effectively only acts when the two qubits are at the same frequency, so that we *can* modulate this coupling by detuning the qubits from each other. When the qubits are tuned into resonance, the interaction is turned on, but when they are sufficiently far detuned so that the detuning Δ is much larger than the coupling strength g, the interaction is turned off[2]. While it is not possible to fully turn off the coupling, in practice we can detune far enough to make the errors small. This does set a limit to how strong the coupling can be made while still allowing it to be turned off, since the qubits can only be detuned by a finite amount. In Chapter 8 we will see some of the consequences of using frequency detuning to modulate the coupling between multiple coupled qubits, in particular the problems that arise because of limited bandwidth and the presence of higher excited states in the qubits.

Another problem with the direct capacitive coupling scheme is that it is subject

[2]One way to understand this effect is to compute the eigenstates of the coupled system when the qubits are far detuned. If the second qubit is detuned by $\Delta \gg g$ then the eigenstates are of the form $|\psi_+\rangle \approx |01\rangle + (g/\Delta)\,|10\rangle$ and $|\psi_-\rangle \approx |10\rangle - (g/\Delta)\,|01\rangle$. Hence, the probability of the coupling causing an exchange between the detuned qubits is $\sim (g/\Delta)^2$.

to measurement crosstalk. The tunnelling and relaxation process by which the qubit state is measured results in hundreds of quanta of energy being released into the circuit. The coupling capacitor can transmit enough of this energy to the other qubit to erroneously excite it as well and cause it to tunnel. With typical coupling strengths of $g \sim 2\pi \cdot 20\,\mathrm{MHz}$ we observe crosstalk on the order of $\sim 10 - 15\,\%$. This problem has been known for quite some time, having been observed in the first experiments with capacitively coupled phase qubits[23]. In Chapter 7 we will discuss a new technique for measuring coupled qubits in a way that avoids this sensitivity to measurement crosstalk. However the technique is not scalable to large numbers of qubits, so for future multi-qubit architectures a coupling scheme that does not produce measurement crosstalk will likely be required.

2.2.2 Resonator coupling

A second technique for coupling two qubits involves coupling them not directly by a capacitor, but through an intermediate resonator to which both qubits are capacitively coupled. The resonator can be fabricated as a coplanar waveguide, which means the resonance frequency can be very accurately designed and the films are made low-loss for good quantum coherence. To implement coupled qubit operations, first one qubit is tuned to the resonator frequency to couple the qubit and resonator, then that qubit is detuned and the second qubit is brought into

resonance to couple it. This two-stage process complicates somewhat the control sequences needed to operate such a coupler, but the setup as a whole has a number of key advantages.

First, since multiple qubits can be placed along the resonator cavity, this setup naturally lends itself to a "quantum bus" architecture, where one resonator can serve to pairwise couple any of the desired qubits. In addition, because of their long coherence times resonators can serve as excellent quantum memory elements, enabling architectures in which qubits are used to manipulate quantum information and resonators to store it, much like the separation between logic and memory elements in a classical computer.

The second advantage of resonator coupling over simple capacitive coupling is the greatly reduced effect of measurement crosstalk. One can understand this by considering the resonator to be essentially a band-pass filter that better isolates one qubit from the energy released by a tunnelling event in the other qubit. In experiments, this has been shown to reduce the effect of measurement crosstalk to well below 1%[3]. On balance, the expanded architectural possibilities and elimination of measurement crosstalk more than compensate for the increased complexity of control sequences required to send quantum information through the resonator, making resonator coupling a very promising technique that is rapidly replacing simple capacitive coupling in the next generation of phase qubit exper-

iments.

2.2.3 Tunable coupling

We have seen that with the always-on interactions produced by capacitive coupling (whether directly between qubits or with an intermediate resonator), frequency detuning is required to modulate the coupling interaction as needed for performing complex sequences with multiple gates. However, frequency detuning has various disadvantages, in particular that frequency space can become crowded with many qubits and it can be hard to achieve the appropriate detunings without unwanted level crossings. What's more, even with large detunings, the coupling is never completely "off" as described above; the off/on ratio can be made small but not arbitrarily so. Ideally, one would like to have a coupling element that allows the coupling strength itself to be adjusted so that it could be precisely tuned to zero, and so that this could be done without requiring the qubits to be detuned. This has been achieved with phase qubits by replacing the capacitive coupling with inductive coupling, and then creating a circuit with tunable inductance. The key element is a Josephson junction, which of course has phase-dependent inductance, and so can serve as the needed tunable element. The first realization of this approach[5] shows great promise and will certainly be developed further.

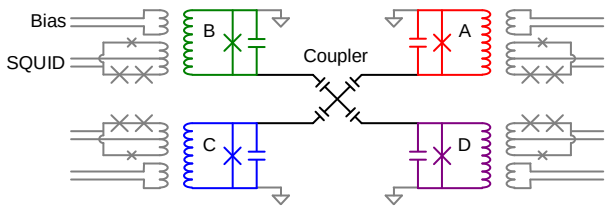

Figure 2.4: Four-qubit design with capacitive coupling.

2.2.4 Multi-qubit coupling

We have now discussed a number of techniques which allow two qubits to be coupled together. How can this be scaled up to larger numbers of qubits? The simplest approach, which we will use in the multi-qubit work described in this thesis, is simply to use more capacitors to create a network of couplings between the various qubits. We designed a device with four superconducting qubits coupled by a central network of capacitors, as shown in Figure 2.4. Chapter 5 discusses the entangling protocols which were designed to be implemented with this sort of coupling circuit; Chapter 6 describes the design of the coupler itself.

This capacitive coupling approach is not scalable very far beyond three or four qubits, mainly due to issues we have already encountered: limited bandwidth for detuning qubits to turn off the coupling, and measurement crosstalk. In Chapter 7 we describe in detail the problems encountered when measuring multiple qubit and a new measurement scheme that allows us to overcome measurement crosstalk, though at the expense of some overhead in extracting information about

the state. The bandwidth constraint can be overcome by changing the network of connectivity, for example coupling the qubits through a "bus" (such as a resonator) rather than coupling all of them simultaneously, so each qubit need only be detuned from the bus, and not all pairs of qubits from each other. The modularity of superconducting circuits allows many such architectures to be realized and explored by piecing together various elements–qubits, capacitors, resonators, tunable couplers. This is one of the hoped-for advantages of superconducting quantum devices, and at the few-qubit level, even with a very simple coupling circuit, we can already see the promise of this modularity.

2.3 Fabrication and Experimental Setup

Superconducting qubits are microfabricated using standard thin-film deposition and etching technologies. The most important components are the superconducting metal (Al), Josephson tunnel junction (Al-AlO_x-Al) and capacitor dielectric (a-Si:H). Our fabrication process is described in detail in the thesis by Ansmann[2, chap. 5]. Of particular importance is the dielectric material, which is crucial for creating a qubit with good coherence[20].

Once devices are fabricated, they are mounted in a dilution refrigerator and cooled to the operating temperature, typically around $20 - 30\,\mathrm{mK}$. The control lines for flux bias, qubit control and readout on each qubit are carefully designed to

be well filtered and impedance matched to 50 Ω to allow for excellent pulse shaping on the high speed lines. Custom-built room temperature electronics generate the various control signals for controlling and measuring the qubits. This cryogenic and wiring setup is also described by Ansmann[2, chap.6], with the main difference here being that we expanded the wiring and electronics to control up to six qubits simultaneously; because the system was designed with expandability in mind, this process was straightforward.

Because of the need for more control lines to operate more qubits, we designed a new sample mounting box for this multiqubit experiment. The sample box is an important piece of infrastructure that must satisfy a number of requirements. The box should serve as a shield against external radiation that might heat the qubit and also against external magnetic flux that might dephase the qubit by causing fluctuations in the qubit frequency. For this purpose a superconducting aluminum box is used. In addition, the box must allow access for microwave and DC control lines to operate the qubit. The microwave lines in particular must be well designed to be impedance matched to 50 Ω so that shaped control pulses applied to the qubit will not be distorted. In addition, these microwave lines should be well isolated from each other to minimize crosstalk between the various control channels. The lines should be isolated not only through the box, but all the way down to the sample die itself. Hence this involves carefully designing the

way the die is mounted into the box. This mounting should also ensure adequate heatsinking so that the die is well thermalized to the refrigerator base temperature.

The mounting box is shown in Figure 2.5, and a closeup of the mounted die is shown in Figure 2.6. The sample is held in place with wirebonds, which are placed densely around the edges to ensure that the die is securely mounted and has good thermal contact with the box. These ground bonds are made as short as possible for rigidity; experience has shown that ~ 5 bonds per side of a 1/4" die is enough to ensure that the die will remain in place, even when the box is inverted and mounted upside down at the bottom of the cryostat. In addition, this number of bonds has been found to be sufficient for heatsinking in the typical experiments we perform. Even so, we typically use more bonds for redundancy and to minimize the crosstalk between the control lines for the various qubits by achieving good grounding at microwave frequencies.

An important component of the box is the design of the microwave feed lines. In previous sample mounts, a PCB structure was used with integrated transmission lines for the control channels, but the PCB surface from which bonds were made to the chip proved to be fragile and unreliable, and the entire design was replaced with a monolithic aluminum box with coaxial feed lines. The first version of such a box used thin coaxial cables threaded through holes in the box as feed lines, which has the advantage that they are nicely matched to $50\,\Omega$, but the

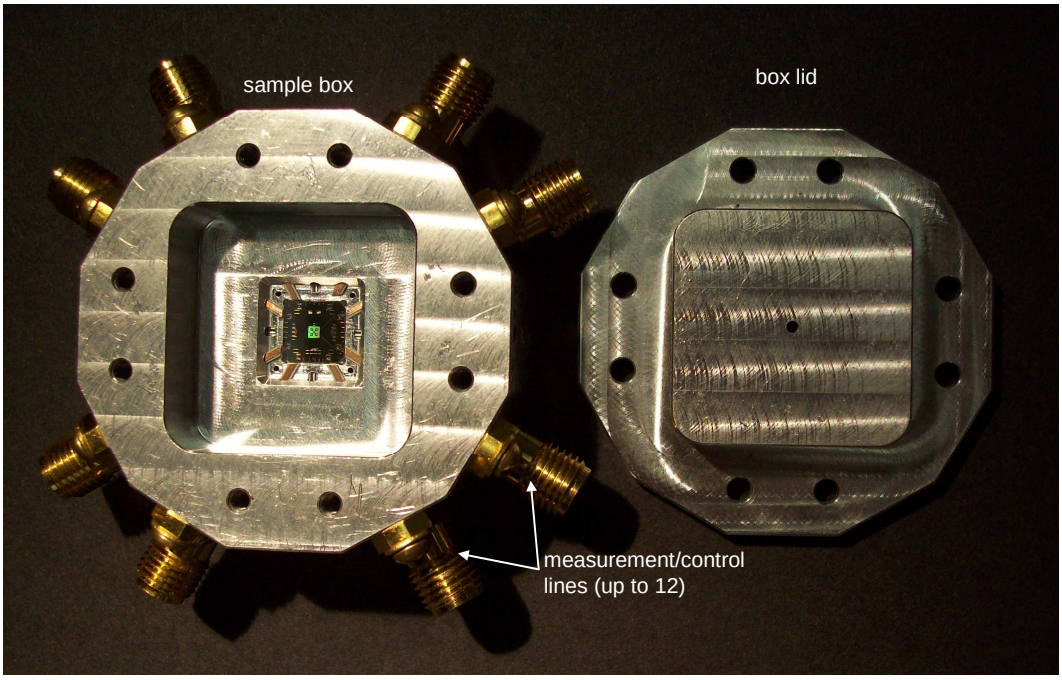

Figure 2.5: Qubit sample mounting box.

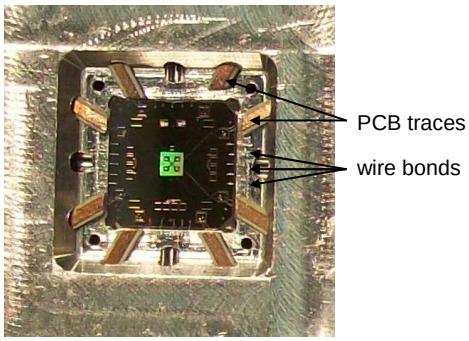

Figure 2.6: Closeup of die mounted in the sample box.

disadvantage that wirebonds must be made from the round central conductor of the coaxial cable to pads on the chip, a process that proved to be difficult and unreliable. Our solution was to replace the coaxial feeds with narrow strips of PCB material. A trace is milled into the upper surface of the copper with the thickness chosen to match the line to $50\,\Omega$ as determined by time-domain reflectometry (TDR) measurements. These PCB strips are soldered to SMA connectors and then inserted into the feed cavities, forming a well-matched $50\,\Omega$ feed. Inside the box, the end of the PCB strip protrudes into the central cavity and rests flat on the ledge surrounding the die, forming a flat, rigid surface that can be readily bonded to in order to connect the control lines to the die, as seen in Figure 2.6.

As part of the design process for the mounting box, we modelled the crosstalk between neighboring microwave control lines both theoretically and using a PCB scale model. This study showed that the box design should work and have acceptably low crosstalk between neighboring microwave control; with 8 wirebonds between neighboring lines the crosstalk is about $-40\,\mathrm{dB}$. In the next section we describe the details of the theoretical and experimental models for those interested.

2.3.1 Modelling Microwave Crosstalk

When a qubit chip is placed in the sample box, there will be a gap between the chip ground plane and the box ground plane that runs around the edge of the chip. This gap can function as a waveguide to transmit signals from one microwave lead to neighboring leads, causing crosstalk. We model this theoretically by considering the wire bonds from the box to the die as inductive shunts between the two ground planes; this ladder of inductors will strongly attenuate any signals that try to propagate along the gap. In addition to the inductive shunts, however, the geometric configuration of the two grounds adds a stray parallel capacitance per unit length. This additional capacitance creates resonances in the circuit, which can produce high transmission around the edge of the die on resonance, resulting in potentially large crosstalk. To reduce the effect of these resonances, we would like to push them to high frequency, which can be done by reducing the stray capacitance per wire bond. This can be achieved by packing the bonds more closely, however if neighboring bonds are too close the direct capacitance between them can kill any gains achieved by reducing the spacing; hence a certain optimum bond spacing is desired.

The model of the interface between the die and box is shown in Figure 2.7. The die has some inductance per unit length ℓ along its edge and wirebonds to ground are spaced by distance d, so that the edge inductance between wirebonds is $L = \ell d$.

Each wirebond has inductance L_W, which we estimate to be $L_W = (1\,\mathrm{nH/mm}) \times g$, where g is the gap between ground and chip. This arrangement leads to a chain of inductor dividers along the interface. Solving for the ratio between voltages at adjacent nodes of the chain, we find

$$\frac{V_1}{V_0} = \left(1 + \frac{1}{2}\frac{L}{L_W}\left(1 + \sqrt{1 + 4\frac{L_W}{L}}\right)\right)^{-1}. \tag{2.1}$$

As the inductance ratio L/L_W is increased, the voltage along the edge will be more quickly suppressed. This implies that we should reduce L_W as much as possible by keeping wirebonds short. The voltage ratio between widely spaced points will be given by a geometric sequence: $V_n/V_0 = (V_1/V_0)^n$. Note that while increasing the distance d between bonds increases L/L_W and thus reduces V_1/V_0, it also reduces the number of bonds n between any two points, and it is this exponent that is most important in giving strong attenuation of the voltage. Table 2.1 gives the predicted attenuation for various inductance ratios and number of bonds.

In addition to the inductances, we must also take into account the capacitance per unit length c between the chip and box in the space between neighboring wirebonds. This chip-to-box capacitance $C = cd$ is parallel to the inductance of the wirebonds themselves and creates a resonance at frequency $\omega_{\mathrm{res}} = 1/\sqrt{L_W C}$. Below this frequency, the wirebond inductances attenuate the signal as described above, but above this frequency the capacitive impedance begins to dominate, and the circuit behaves like a transmission line, propagating signals along the interface

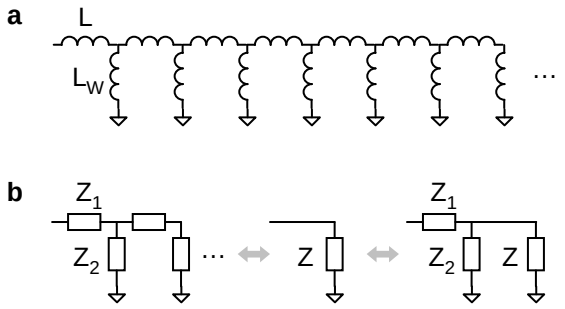

Figure 2.7: Transmission line model of gap between mounting box and sample die.

L/L_W	Attenuation (dB/bond)			
	$n = 1$	$n = 5$	$n = 9$	$n = 14$
0.1	2.74	13.7	24.6	38.3
0.2	3.85	19.3	34.7	53.9
0.3	4.70	23.5	42.3	65.8
0.4	5.41	27.0	48.7	75.7
0.5	6.02	30.1	54.2	84.3
0.6	6.57	32.9	59.1	92.0
0.7	7.07	35.4	63.6	99.0
0.8	7.53	37.7	67.8	105.4
0.9	7.96	39.8	71.6	111.4
1.0	8.36	41.8	75.2	117.0

Table 2.1: Predicted attenuation in wirebond model.

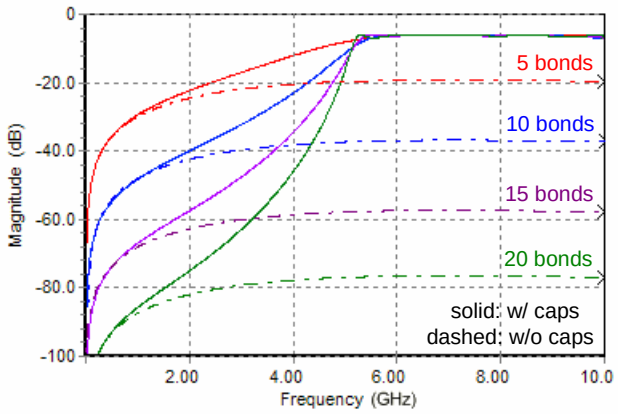

Figure 2.8: Spice simulation of transmission line model.

with little attenuation.

We simulated this circuit with SPICE for various numbers of bonds and ratios of inductance and capacitance and found that these simulations generally confirmed the semiquantitative predictions of the simple model, namely that there would be strong transmission above $\omega_{\rm res} = 1/\sqrt{L_W C}$ and that increasing numbers of bonds would attenuate more strongly at frequencies below the resonance. Figure 2.8 shows a typical simulation result, this one corresponding to model C in Table 2.2. The solid traces include the capacitance, while the dashed traces omit it. The colors are for different numbers of bonds: 5 (red), 10 (blue), 15 (purple) and 20 (green). As predicted, we see greater attenuation for higher bond numbers and L/R rolloff at low frequency, both with and without capacitance. When the capacitance is added (solid traces) we also see high transmission above the resonant frequency.

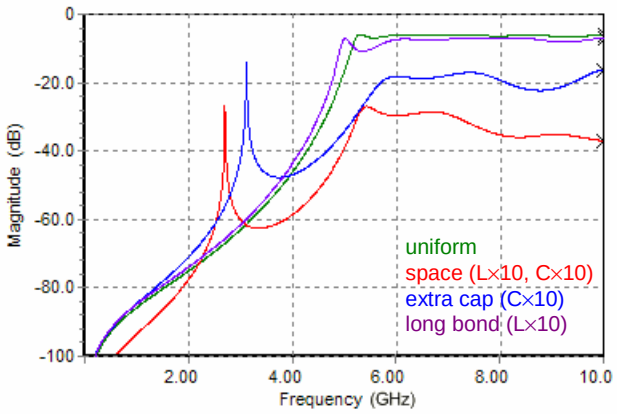

Figure 2.9: Spice simulation of transmission line model with nonuniformity.

We also used SPICE to simulate the effect of non-uniformity in the inductor ladder. Figure 2.9 shows simulation results with the same parameters as in Figure 2.8 with 20 bonds, but with one bond halfway down the line modified to simulate nonuniformity. The green trace is the same in the previous figure for reference; for the red trace both L and C are increased by a factor of 10 to simulate a large space between bonds; for the blue trace the capacitance C alone is increased to simulate localized excess capacitance with no gap in the bonds; finally for the purple trace the inductance L_W is increased by a factor of 10 to simulate a single long bond. Surprisingly, we find that the lower-frequency resonance at this one intermediate location can allow signals to propagate strongly through the entire circuit, but only in the cases where the capacitance is increased. When the inductance alone is increased (purple trace), no additional low-frequency peak is observed. Hence the crosstalk performance is much more sensitive to spaces

between bonds and extra capacitance than it is to nonuniformity in the bonds lengths. However, note that even with a factor of 10 in nonuniformity, the resonance frequency is only reduced by a factor of ~ 2, which can still be kept safely above the qubit frequency.

To test this transmission line model experimentally, we created $10\times$ scale models of the chip-box assembly with variations of several geometric parameters, as shown in Figure 2.10. Each model has three microwave ports which allow us to measure transmission at 0.5, 0.9 and 1.4 inches along the gap. The signal lines at each port are connected to the chip with a $50\,\Omega$ resistor soldered across the gap. The front view shows variations in the wirebond length and spacing, while the back view shows variations in the overlap which creates different amounts of capacitance between chip and ground. In D, for example, the lighter colored region around the edge is the gap between chip and ground on the other side of the board; due to this negative overlap, D has the smallest capacitance per unit length.

The physical distances in the model are g, the gap between the ground plane and chip (the length of the bonds); d, the spacing between bonds; and v, the overlap or distance that the ground plane on the back of the board extends underneath the chip on the top. Here larger v (more overlap) leads to more capacitance. For each model geometry, we calculated the inductance and capacitance per unit

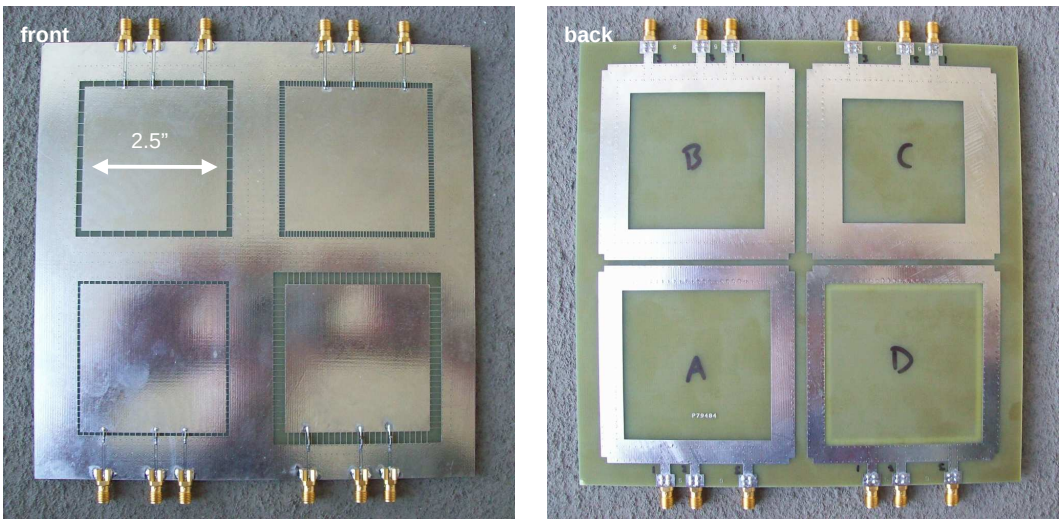

Figure 2.10: Scale model of the wire-bond chip mount.

length along the chip-box interface using the COMSOL multiphysics package[3]. The capacitance per unit length is easy to calculate directly with COMSOL, and from this we can also get the inductance per unit length by repeating the capacitance calculation with all relative dielectric constants set to unity[22]. Table 2.2 lists parameters for all models A-H, as well as the capacitances and inductances calculated with COMSOL. In addition, it shows the L/L_W ratio and the predicted resonant frequency $1/\sqrt{L_W C}$ for each configuration.

For each of the model structures, we used a vector network analyser to measure transmission between various points along the die edge separated by different numbers of bonds. The resulting traces are shown in Figure 2.11. The data show reasonable qualitative agreement with the predictions of the model: we see the

[3] http://www.comsol.com/

	g (in)	d (in)	v (in)	c (fF/mm)	ℓ (nH/mm)	L (nH)	L_W (nH)	C (fF)	L/L_W	f_{res} (GHz)
A	0.1	0.1	0.0	55.9	0.43	1.09	2.54	142	0.43	8.38
B	0.1	0.1	0.1	114.6	0.29	0.74	2.54	291	0.29	5.85
C	0.1	0.1	0.2	174.9	0.22	0.56	2.54	444	0.22	4.74
D	0.1	0.1	-0.1	32.4	0.58	1.47	2.54	82	0.58	11.0
E	0.1	0.2	0.1	114.6	0.29	1.47	2.54	582	0.58	4.14
F	0.05	0.1	0.1	117.2	0.28	0.71	1.27	298	0.56	8.18
G	0.2	0.1	0.1	114.0	0.30	0.76	5.08	290	0.15	4.15
H	0.1	0.05	0.1	114.6	0.29	0.74	2.54	145	0.15	8.28

Table 2.2: Model parameters for chip mount scale model.

inductive rolloff at very low frequency, resonant transmission at frequencies above several GHz, and increased attenuation with increasing bond number when other parameters are kept constant. The vertical lines indicate the predicted resonance frequency f_{res} as given in Table 2.2, again showing reasonable agreement. Note that the FR-4 material of the model does not perform well at high frequencies, so the measurement is expected to break down above several GHz, which is why no resonance is visible on device D, for example.

These experimental results give us confidence that our simple model reproduces the essential physics of crosstalk in the mounting box, so that we can use it to predict the crosstalk levels for the actual geometry of a die mounted in the sample box. We use COMSOL to calculate the inductance and capacitance per unit length for the actual box geometry, where we have a gap between the box edge and the chip ground plane of 200 μm $\approx$ 8 mil and a bond length of 400 μm. With 20 bonds

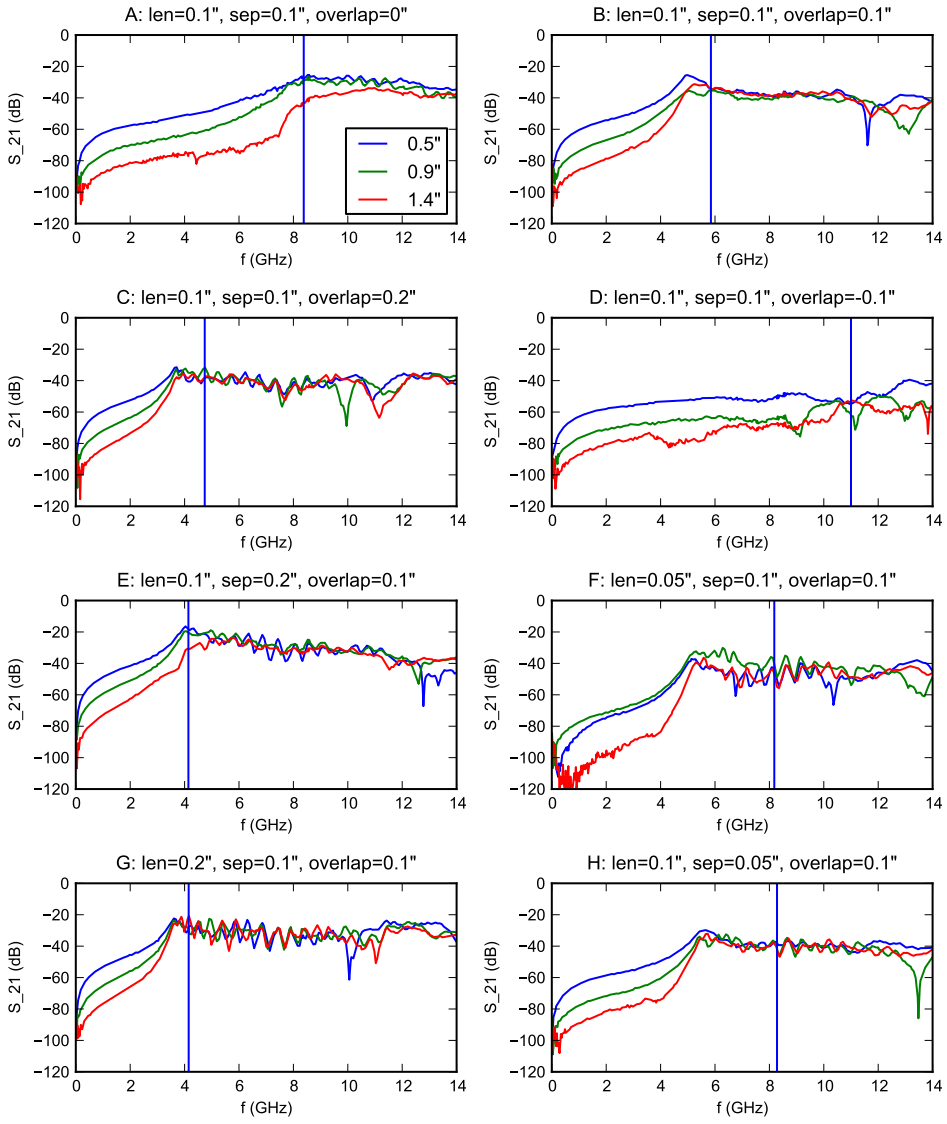

Figure 2.11: Transmission measurements of wirebond model.

per cm, we find $f_{\rm res} = 43\,\mathrm{GHz}$ and $L/L_W = 0.37$ for an attenuation of $5.2\,\mathrm{dB}$ per bond, or $104\,\mathrm{dB/cm}$. In experiments, we use roughly 8 bonds between microwave leads, for an expected attenuation of $\sim 40\,\mathrm{dB}$. This is adequate for our purposes, although one could use even shorter bonds and tighter spacing to achieve even better isolation if needed.

Chapter 3

From Quantum Circuit to Experimental Control

In this chapter we consider the following problem: how do we take an abstract description of a quantum operation and turn this into the currents and voltages that will be applied to our device to implement that quantum operation? An important part of this process is the calibration of the operating parameters of the individual qubits, such as for reset and readout. The procedure for single-qubit calibration has been described in detail previously[2, chap. 8]; here we will focus on a different aspect of the process, namely how to translate sequences of gates—as required for more complicated algorithms—into experimental control signals.

Of particular importance in this process is the concept of a rotating frame. We will discuss how to manipulate individual qubits in particular frames and how to keep track of rotating frames when qubits are subject to time-dependent detunings. Essentially this reduces to keeping track of the phases accumulated as qubits are detuned, and synchronizing the various frames defined by different frequencies. Even at the level of the simple three-qubit experiments we have performed, these phase accumulation and frame synchronization issues are important for properly constructing the qubit control sequences; as the number of qubits increases and the algorithms become more complicated, understanding this translation process becomes more important still.

3.1 Single-qubit Control

Let us start with the problem of controlling a single qubit. In the basis of energy eigenstates, the bare qubit Hamiltonian is extremely simple:

$$H_0 = \begin{pmatrix} 0 & \\ & \hbar\omega_{10} \end{pmatrix} \tag{3.1}$$

where we have assumed without loss of generality that the ground state energy is zero. In terms of time evolution, a system that starts in the ground state remains stationary, while a system that starts in the excited state acquires a time-dependent phase due to the energy difference: $|1(t)\rangle = e^{-i\omega_{10}t}|1\rangle$.

In addition to this bare Hamiltonian, we have a set of control parameters that can be used to control the state evolution by applying time-dependent signals. In the case of our phase qubits, we can tilt the barrier by applying a flux bias, which changes the energy spacing of the qubit states and hence couples to the Z term in the Hamiltonian, and we can also apply microwaves either through a coupling capacitor, in which case we couple to Y, or through a flux coil, in which case we couple to X. It will turn out that these last two give the same results in the end, so let us assume we couple to X such that our time-dependent Hamiltonian with control signals is

$$H = H_0 + H_c = \begin{pmatrix} 0 & \\ & \hbar\omega_{10} \end{pmatrix} + \hbar \begin{pmatrix} 0 & \epsilon(t) \\ \epsilon(t) & \delta(t) \end{pmatrix} = \hbar \begin{pmatrix} 0 & \epsilon(t) \\ \epsilon(t) & \omega_{10} + \delta(t) \end{pmatrix} \quad (3.2)$$

where $\epsilon(t)$ is the applied X control signal and $\delta(t)$ is the applied Z control signal. An important experimental issue is the calibration of these signals, determining for example exactly how an applied voltage $V(t)$ relates to the detuning $\delta(t)$. These calibrations have been described previously[2, chap. 8] so we will here assume that this has been done, and simply work with $\epsilon(t)$ and $\delta(t)$.

It seems rather inconvenient to be working in a basis in which a system starting in an eigenstate will evolve in time, even when no control signals are applied. Instead, we might like to change to a basis where a system starting in an eigenstate will be stationary under the bare Hamiltonian, and only change due to our applied

control signals. This is similar to the "Interaction Picture" often used in quantum mechanics, though for our purposes we will instead think of this as transforming to a rotating frame. The reasoning is simple: we want to take the "rotating" state $|1'\rangle = e^{-i\omega_{10}t}|1\rangle$ to be our basis state.

This basis transformation is a special case of a more general basis transformation that is easy to derive. We start with the standard Schrödinger equation, expressed in terms of the action on some set of basis states $|n\rangle$, namely

$$-i\hbar\frac{d}{dt}|n\rangle = H(t)|n\rangle\,. \tag{3.3}$$

Now, suppose we wish to change to a new basis that is related to the old basis by an arbitrary and possibly time-varying unitary transformation $|n'\rangle = V(t)|n\rangle$. Multiplying by the inverse matrix $V^{-1}(t) = V^\dagger(t)$ and substituting into (3.3) gives

$$-i\hbar\frac{d}{dt}V^\dagger(t)|n'\rangle = H(t)V^\dagger(t)|n'\rangle\,. \tag{3.4}$$

Using the chain rule to evaluate the derivative on the left and then collecting terms, we obtain

$$-i\hbar V^\dagger(t)\frac{d}{dt}|n'\rangle = \left(H(t)V^\dagger(t) + i\hbar\dot{V}^\dagger(t)\right)|n'\rangle\,. \tag{3.5}$$

Finally, by multiplying on the left by the transformation matrix $V(t)$ we obtain

$$-i\hbar\frac{d}{dt}|n'\rangle = \left(V(t)H(t)V^\dagger(t) + i\hbar V(t)\dot{V}^\dagger(t)\right)|n'\rangle\,. \tag{3.6}$$

This result has the same form as the original Schrödinger equation, but now with

a new Hamiltonian given by

$$H' = VHV^\dagger + i\hbar V\dot{V}^\dagger \tag{3.7}$$

where for clarity the time-dependence is left implicit.

The desired transformation to the single-qubit rotating frame can now be found using a simple application of this general result. The transformation matrix is

$$V(t) = \begin{pmatrix} 1 & \\ & e^{-i\omega_{10}t} \end{pmatrix} \tag{3.8}$$

and the new Hamiltonian in the rotating frame $|0'\rangle$, $|1'\rangle = e^{-i\omega_{10}t}|1\rangle$ is

$$\begin{aligned} H' &= VH_0V^\dagger + VH_cV^\dagger + i\hbar V\dot{V}^\dagger \\ &= \begin{pmatrix} 0 & \\ & \hbar\omega_{10} \end{pmatrix} + \hbar \begin{pmatrix} 0 & \epsilon(t)e^{i\omega_{10}t} \\ \epsilon(t)e^{-i\omega_{10}t} & \delta(t) \end{pmatrix} - \begin{pmatrix} 0 & \\ & \hbar\omega_{10} \end{pmatrix} \\ &= \hbar \begin{pmatrix} 0 & \epsilon(t)e^{i\omega_{10}t} \\ \epsilon(t)e^{-i\omega_{10}t} & \delta(t) \end{pmatrix}. \end{aligned} \tag{3.9}$$

Note that H_0 has completely disappeared, and all we are left with is the control terms; the rotating frame has effectively zeroed out all the bare energies, so that the eigenstates are indeed stationary, as desired.

What is the effect of turning on detuning for some time? Suppose we apply a signal $\delta(t)$ between times t_0 and t_1. We can readily compute that the state of the system will undergo a unitary transformation given by

$$U = \begin{pmatrix} 1 & \\ & e^{i\int_{t_0}^{t_1}\delta(t)dt} \end{pmatrix} = \begin{pmatrix} 1 & \\ & e^{-i\theta} \end{pmatrix} \tag{3.10}$$

where $\theta = -\int_{t_0}^{t_1}\delta(t)dt$. This is just the unitary matrix for rotation about the Z

axis of the Bloch sphere. Note: *positive* detuning is a rotation by a *negative* angle about Z, and vice-versa. This is due to our Bloch sphere convention in which $|1\rangle$ is at the south pole.

Now we consider the effect on the state of the off-diagonal terms controlled by $\epsilon(t)$. As we can see in Equation 3.9, there is a factor $e^{i\omega_{10}t}$ multiplying these terms, causing them to rotate. To counter this, we will use a control signal that itself rotates in the opposite direction. Let us take

$$\epsilon(t) = \epsilon_x(t)\cos(-\omega_{10}t) + \epsilon_y(t)\sin(-\omega_{10}t) = \epsilon_x(t)\cos\omega_{10}t - \epsilon_y(t)\sin\omega_{10}t. \quad (3.11)$$

From this we calculate that the terms appearing in the Hamiltonian H' are

$$\epsilon(t)e^{i\omega_{10}t} = \frac{1}{2}\epsilon_x(t)(1+e^{2i\omega_{10}t}) - i\frac{1}{2}\epsilon_y(t)(1-e^{2i\omega_{10}t}) \quad (3.12)$$

$$\epsilon(t)e^{-i\omega_{10}t} = \frac{1}{2}\epsilon_x(t)(1+e^{-2i\omega_{10}t}) + i\frac{1}{2}\epsilon_y(t)(1-e^{-2i\omega_{10}t}). \quad (3.13)$$

Until this point, everything we have done has been exact, but now we make a common approximation known as the Rotating Wave Approximation (RWA) by dropping the terms oscillating at $2\omega_{10}$. What justifies us in making this approximation? Consider that the antiderivative of $\epsilon\cos(2\omega_{10}t)$ is equal to $(\epsilon/2\omega_{10})\sin(2\omega_{10}t)$. This will be small as long as $\epsilon \ll 2\omega_{10}$, which means that the dynamics caused by our applied control signal ϵ should be much slower than the dynamical time scale of the energy difference ω_{10}. In typical experiments the energy difference is

$\omega_{10} \sim 2\pi \cdot 6\,\text{GHz}$ while rotations due to applied control signals take on the order of 10 ns so that $\epsilon \sim 2\pi \cdot 100\,\text{MHz}$. Hence the dynamical effect of the 'counter-rotating' terms at frequencies $\pm 2\omega_{10}$ will indeed be small, and can be ignored. Let us reiterate: the transformation to a rotating frame is exact; the rotating wave approximation is an approximation that assumes the dynamics due to applied controlled signals are slow compared to the characteristic dynamical timescale set by the energy spacing.

Having made the RWA, the Hamiltonian in the rotating frame becomes

$$H' = \hbar \begin{pmatrix} 0 & \frac{1}{2}(\epsilon_x(t) - i\epsilon_y(t)) \\ \frac{1}{2}(\epsilon_x(t) + i\epsilon_y(t)) & \delta(t) \end{pmatrix} \tag{3.14}$$

which means that we have complete control over the qubit Bloch sphere with the capability to rotate about any axis, even though in the original frame of the qubit we only assumed control terms that couple to Z and X. By moving to the rotating frame where X and Y get mixed together, we can adjust the phase of the time-dependent $\epsilon(t)$ control to allow rotations about both axes.

In fact this is also the reason why our qubit-qubit coupling is $(\hbar g/2)(XX+YY)$ as we specified above in §2.2. The actual term in the Hamiltonian is the energy of the coupling capacitor which gives a charge-charge or YY coupling. When we go into the rotating frame of the two qubits, and then make the RWA to cancel counter-rotating terms this coupling term becomes $(\hbar g/2)(XX + YY)$. Just as in the single-qubit case, the RWA is a valid approximation because the coupling

strength g is small $g \ll \omega_{10}$, so that the coupling dynamics are much slower than the qubit frequency and the counter-rotating terms have little effect.

An alternate and often more convenient way to think about the transverse control signal $\epsilon(t)$ is to consider a complex control signal $\epsilon'(t) \equiv \epsilon_x(t) + i\epsilon_y(t)$ with which we can write $\epsilon(t) = \mathrm{Re}(\epsilon'(t)e^{-i\omega_{10}t})$. When $\epsilon'(t)$ is real, this produces rotations around the X axis of the Bloch sphere, when it is imaginary it produces rotations about the Y axis, and by adjusting the phase we can produce rotations about any axis in the X-Y plane of the Bloch sphere. Note that because X and Y are mixed together, there is an overall degree of freedom here that lets us choose the absolute phase reference arbitrarily. Picking a phase reference *defines* our convention for the X axis, and after the choice is made only the relative phase between X and Y or other axes matters.

Given all this, the procedure for generating a sequence of single-qubit gates is straightforward. We simply consider each gate as a rotation by some angle α about a vector direction $\mathbf{r} = (r_x, r_y, r_z)$ in terms of the Bloch sphere, and then for each gate we apply control signals ϵ and δ such that the vector $(\epsilon_x, \epsilon_y, -\delta)$ is parallel to the rotation axis $\mathbf{r}$ and such that the length integrates to the desired rotation angle $\int_{t_0}^{t_1} \sqrt{\epsilon_x^2 + \epsilon_y^2 + \delta^2}dt = \alpha$. Before applying the signals to the qubit, we multiply $\epsilon' = \epsilon_x + i\epsilon_y$ by $e^{-i\omega_{10}t}$ and take the real part to get the actual applied signal, which is done in experiments simply by using an IQ mixer.

3.2 Dealing With Detuning

In the previous section we considered control of a single qubit with one rotating frame of reference given by the qubit frequency ω_{10}. We allowed the system to be detuned away from that reference frame, but only for the purposes of performing Z-rotations. After a Z-rotation we always came back to the original frame, and the fact that the qubit state had rotated was the desired effect. Sometimes however, we want to detune a qubit not to perform a Z rotation, but rather to couple it with another qubit or to move it out of the way so that it does not couple with another qubit. In these cases, we must keep track of the resulting *unwanted* Z-rotations and correct them, either in "hardware" by applying compensating pulses, or in "software" by modifying our subsequent definition of the frame.

In order to know what to correct, we could measure the phase accumulated by the qubit during such a detuning pulse, as shown in Figure 3.1. The measurement is simply a Ramsey fringe experiment where we apply $\pi/2$ rotations before and after the detuning pulse and vary the axis of rotation of the second pulse by changing the phase of ϵ'. This axis of rotation is relative to the fixed frame of reference at the original qubit frequency, but the qubit itself has rotated because it has been detuned. The position of the fringe maximum tells us the rotation angle picked up during the detuning pulse. Instead of measuring the phase angle we could of course calculate it as $\Delta\theta = -\int_{t_0}^{t_1} \delta(t)dt$ as we have seen above, and

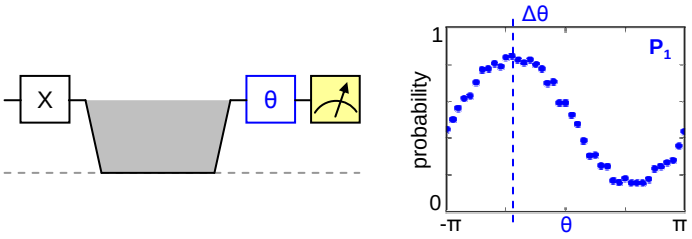

Figure 3.1: Measuring phase accumulated by a detuning pulse.

indeed for more complex sequence construction it will be useful to automate these corrections rather than measuring them for each detuning pulse.

The two techniques for correcting this phase are shown in Figure 3.2. On the left, we have the "hardware" correction, where the pulse shape is modified to produce a net rotation angle that is an integral multiple of 2π and can thus be ignored. Any subsequent pulses will then be programmed in the expected way in the qubit frame. On the other hand we can perform a "software" correction where we simply modify the way subsequent pulses are programmed, offsetting their phase by the angle $\Delta\theta$, as shown on the right. This results in shorter sequences in terms of time, but at the cost of extra bookkeeping, since the phase corrections must be propagated to *all* subsequent pulses, and later detuning pulses will add additional corrections, *etc.* The choice of which correction to choose will thus depend on other factors having to do with the larger sequence of operations of which this one step is a part.

These two corrections really are equivalent; in one case we actually rotate

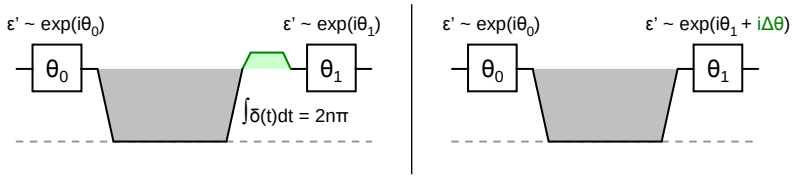

Figure 3.2: Correcting the phase accumulated by a detuning pulse.

the qubit by adding an additional Z-pulse while in the other case we rotate our reference frame to match the qubit by changing the phase of the control signals. It is useful to remember these two complementary approaches and to consider both when implementing complex pulse sequences, so that the most convenient choice can be made in any particular case.

3.3 Multiple Frames of Reference

In addition to the corrections that must be applied to a single qubit when detuned, as described in the previous section, we must sometimes deal with the problem of synchronizing multiple frames at different frequencies. Suppose we have a qubit that starts at frequency ω_{10}, and we apply various pulses in that frame, always multiplying the control parameter by $e^{-i\omega_{10}t}$. At some point in the sequence, the bias is changed to move the qubit to a new frequency ω'_{10}, and we wish to apply gates to it in this new frame. The problem here is to know what phase to use in order to get a particular qubit rotation, say about X, in this new frame.

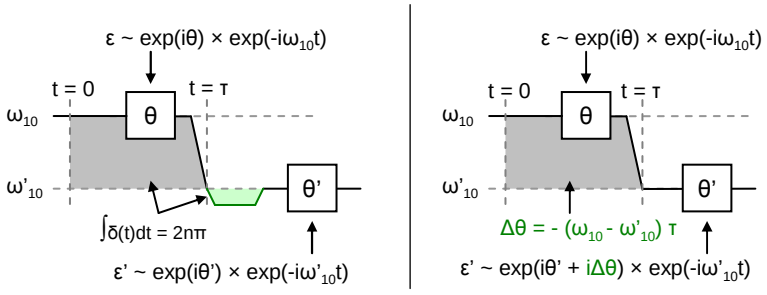

Figure 3.3: Synchronizing two different rotating frames.

In order to solve this, we must have a common time reference for both frames so that we can define a common point $t = 0$. From the point of view of the ω'_{10} frame, the qubit has just been detuned by $\delta = \omega_{10} - \omega'_{10}$ from time $t = 0$ until the time τ at which the qubit is brought into the new frame. This detuning can be measured with a Ramsey fringe as described above, and then corrected via either of the two methods just described, using an actual rotation with an extra Z-pulse, or adjusting the phase of subsequent pulses to account for the detuning. The shift between reference frames and the two techniques for synchronizing the frames are illustrated in Figure 3.3.

We mentioned that this technique requires a global time reference so that we can define $t = 0$ in the two frames at frequencies ω_{10} and ω'_{10}. In our experiments all signal generators share a common clock, so that at the signal generators there is a common time reference. In the case we are considering here, where we have one qubit that is being moved between two reference frames, the control signals

at both frequencies will be produced by one generator via sideband mixing and travel down a single coaxial cable to the qubit, so that the time reference will be the same at the qubit as well.

In other cases however, we may need to synchronize two frames between different qubits, where the signals are produced by different generators or sent to the device in different coaxial cables, so that the path lengths will not be the same, and the common clock at the generators will not be synchronized at the qubits. A technique for calibrating the timing between two qubits when they are on resonance and so at the same frequency has been discussed before[2, p. 213]. This allows to set a common $t = 0$ between the qubits, and then the technique we have just described above can be used to achieve the necessary synchronization when the two qubits are in different frames and only brought into resonance and synchronized at some time τ.

The correction pulse on the left in Figure 3.3 is reminiscent of the "overshoots" that are added to detuning pulses to improve the fidelity of the iSWAP interaction[14], but these two things are very different. The correction pulse in the figure is simply a way to adjust the total integrated detuning for the qubit to be an integral multiple of 2π, and *any* modification of the curve $\delta(t)$ that makes this the case will do. For example this could be done by not changing the shape at all but rather simply adjusting the time τ, or by adding a correction pulse on the

ω_{10} side, *before* the transition to the ω'_{10} frame. The overshoots added to coupling pulses are instead used to correct for the fact that the coupling interaction is not purely on or off, but turns on gradually as the qubits are brought into resonance.

3.4 Higher Qubit Levels

As was discussed in Chapter 2, the phase qubit is not purely a two-level system. Rather, it has several energy levels in the well, but with enough anharmonicity so that it can be treated as an effective two-level system. Typically we need only consider the next higher excited state $|2\rangle$, and any transitions into this state are taken as leakage errors. Various techniques can be used to prevent such errors, the simplest of which is to simply limit the magnitude of control signals so that the dynamics due to driving are slower than the nonlinearity $\Delta \equiv \omega_{21} - \omega_{10}$. More advanced pulse-shaping techniques such as Derivative Reduction by Adiabatic Gates (DRAG) can also be used[25], allowing for faster pulses while still minimizing leakage.

The qudit Hamiltonian is a simple extension of what we saw above for a qubit. The bare Hamiltonian is

$$H_0 = \begin{pmatrix} 0 & & \\ & \hbar\omega_{10} & \\ & & 2\hbar\omega_{10} + \hbar\Delta \end{pmatrix} \tag{3.15}$$

and when we move to a rotating frame defined by $|0'\rangle = |0\rangle$, $|1'\rangle = e^{-i\omega_{10}t}|1\rangle$ and

$|2'\rangle = e^{-2i\omega_{10}t}|2\rangle$ this becomes

$$H_0' = \begin{pmatrix} 0 & & \\ & 0 & \\ & & \hbar\Delta \end{pmatrix}. \tag{3.16}$$

We still couple to the system through the X operator which is now a 3×3 matrix as well. In the case of a purely harmonic system, the X matrix can be easily derived from the raising and lowering operators, and is given by the tridiagonal matrix

$$X_{\text{SHO}} = \begin{pmatrix} 0 & 1 & \\ 1 & 0 & \sqrt{2} \\ & \sqrt{2} & 0 \end{pmatrix}. \tag{3.17}$$

In the weakly-anharmonic qudit case, the matrix elements can all be worked out in various ways, for example using perturbation theory starting from the harmonic oscillator and adding a cubic term[35], or by direct numerical diagonalization of the qubit potential[2, chap. 3]. The upshot of either exercise is that one finds matrices that have essentially the same tridiagonal form as for the harmonic oscillator, although with somewhat different numerical values for the various coefficients. The matrix elements for the other diagonals, connecting states $|n\rangle$ and $|n \pm k\rangle$ where $k \geq 2$ are very small; for typical phase qubit nonlinearities these are down by a factor of 20 for $k = 2$ and even further reduced for larger k. For our purposes these terms are completely negligible, and the interaction term has the form

$$X_{\text{qudit}} = \begin{pmatrix} 0 & 1 & \\ 1 & 0 & \lambda_2 \\ & \lambda_2 & 0 \end{pmatrix} \tag{3.18}$$

where the λ_k matrix elements can be calculated, or just measured experimentally[1]. In addition, of course, we also have the usual Z control by which we can adjust the qudit frequency. For small detunings, we can make the assumption that the nonlinearity Δ is constant.

In the rotating frame, the control Hamiltonian is given by

$$H' = \hbar\epsilon(t)X + \hbar\delta(t)Z = \hbar \begin{pmatrix} 0 & \epsilon(t)e^{i\omega_{10}t} & \\ \epsilon(t)e^{-i\omega_{10}t} & \delta(t) & \lambda_2\epsilon(t)e^{i\omega_{10}t} \\ & \lambda_2\epsilon(t)e^{-i\omega_{10}t} & 2\delta(t)+\Delta \end{pmatrix}. \quad (3.19)$$

If we make the same choice for the time-dependent control as we did in the single-qubit case, namely $\epsilon(t) = \mathrm{Re}(\epsilon'(t)e^{-i\omega_{10}t})$, then this Hamiltonian becomes

$$H' = \hbar \begin{pmatrix} 0 & \epsilon'^* & \\ \epsilon' & \delta & \lambda_2\epsilon'^* \\ & \lambda_2\epsilon' & 2\delta+\Delta \end{pmatrix} \quad (3.20)$$

where we have made the RWA by dropping counter-rotating terms and where we have left the time-dependence implicit. Clearly, whenever we attempt to drive transitions between $|0\rangle$ and $|1\rangle$, we will also be driving transitions between $|1\rangle$ and $|2\rangle$, however the nonlinearity Δ means that our drive at frequency ω_{10} is detuned from this transition, so that as long as the drive is sufficiently weak the leakage will be small. Pulse shaping through DRAG improves on this, as mentioned above.

Instead of considering the states $|2\rangle$ and higher as potential source of er-

[1]This form of the X matrix can be understood as being due to approximate selection rules that prohibit first-order transitions between widely separated energy states of the system. For the Harmonic oscillator such selection rules are exact, while for our weakly anharmonic oscillator, these selection rules are reasonably good approximations.

ror through leakage, we can use the entire set of multiple levels as a quantum system and intentionally drive transitions between the various levels[28]. This is essentially equivalent to working in a rotating frame defined by $|0'\rangle = |0\rangle$, $|1'\rangle = e^{-i\omega_{10}t}|1\rangle$ and $|2'\rangle = e^{-i(\omega_{21}+\omega_{10})t}|2\rangle$. We drive the system at both transition frequencies ω_{10} and ω_{21} simultaneously by using a control signal $\epsilon(t) = \mathrm{Re}(\epsilon_{10}(t)e^{-i\omega_{10}t} + \epsilon_{21}(t)e^{-i\omega_{21}t})$. As this qudit is detuned, we can regard the frames defined by the $0 \rightarrow 1$ detuning $\delta(t)$ and the $1 \rightarrow 2$ detuning $\delta(t) + \Delta$ as two separate frames, and use the various techniques outlined above for building pulse sequences that address each of these transitions when detunings are present.

In the experiments presented here, we used DRAG pulses to minimize leakage to higher qubit levels, but did not otherwise utilize the higher excited levels of the phase qubits. As we will see in Chapter 8 however, these higher levels also play a role in causing errors when we have a system of coupled qubits.

Chapter 4

Three-Qubit Entanglement

Entanglement is usually defined by what it is not: an entangled state is not a product state. A product state of course is one in which each *part* of a composite system is in a well-defined state, and the overall system is just a trivial combination—a "tensor product"—of these components. Such a definition is not particularly enlightening, however, so we will try to unpack this notion a bit in the following sections. We will first consider bipartite entanglement between two qubits, where the theory is very well understood. Next, we will proceed to the case of interest in this thesis, namely that of tripartite entanglement between three qubits, where the story is richer and more interesting than the two-qubit case. Already at the level of three qubits, the theory of entanglement is less well understood, particularly for the experimentally-relevant case of mixed states.

4.1 Just the Two of Us

The idea of entanglement is best illustrated and understood in the context of two qubits. The paradigmatic two-qubit entangled states are the so-called Bell states, named after the British physicist John Bell who first described the nonclassical nature of the correlations that arise when measuring these entangled states[4]. Up to normalization, the four Bell states are $|\Phi^{\pm}\rangle \equiv |00\rangle \pm |11\rangle$ and $|\Psi^{\pm}\rangle \equiv |01\rangle \pm |10\rangle$.

How can we tell that these states are entangled? Consider the first Bell state $|\Phi^{+}\rangle = |00\rangle + |11\rangle$. Let us try to write this as a product state $(\alpha\,|0\rangle + \beta\,|1\rangle)(\gamma\,|0\rangle + \delta\,|1\rangle) = \alpha\gamma\,|00\rangle + \alpha\delta\,|01\rangle + \beta\gamma\,|10\rangle + \beta\delta\,|11\rangle$. Because the coefficients of $|00\rangle$ and $|11\rangle$ are nonzero, we have the condition that none of the complex coefficients in the product states can be zero: $|\alpha| \neq 0$, $|\beta| \neq 0$, $|\gamma| \neq 0$ and $|\delta| \neq 0$. On the other hand, because the coefficient of $|01\rangle$ is zero in the Bell state, we must have either $|\alpha| = 0$ or $|\delta| = 0$, a clear contradiction. This proves that $|\Phi^{+}\rangle$ cannot be expressed as a product state, and so must be an entangled state.

While an explicit proof of entanglement was not difficult in this simple case, it is nevertheless rather tedious. Much more general techniques have been developed to show that various arbitrary states are entangled, and indeed to quantify this entanglement. At first glance it may seem strange to try to quantify entanglement. Note however that the four Bell states mentioned above can be converted into each other by applying "local" rotations (or changes of basis) to the individual

qubits. For example, if we apply a Z-rotation to one qubit so that $|0\rangle \rightarrow |0\rangle$ and $|1\rangle \rightarrow -|1\rangle$, then we can convert $|\Phi^+\rangle$ into $|\Phi^-\rangle$. In this sense the Bell states are interchangeable; modulo Local Operations and Classical Communication (LOCC) they each contain the same *amount* of entanglement.

For two qubits there is only one type of entanglement, as any other maximally-entangled state is also local-unitary equivalent to a Bell State. Various entanglement monotones have been presented that actually measure this entanglement[38]. An entanglement monotone is a function that takes as input a density matrix ρ and outputs a number between 0 and 1. The output is 0 for product states, 1 for maximally-entangled states such as the Bell states, and takes intermediate values for other states that are only partially entangled.

For experiments it is important that these entanglement monotones can be calculated for general density matrices, not just for pure states, because of course experimentally-measured states are always partially mixed states due to decoherence. In general we can take any entanglement measure defined on pure states and extend it to arbitrary density matrices using the so-called "convex-roof" extension[30]. To do this for an entanglement monotone e defined on pure states, we define $e(\rho) = \min \sum p_i e(|\psi_i\rangle)$ where the minimization is taken over all possible pure-state decompositions of the original state $\rho = \sum p_i |\psi_i\rangle \langle\psi_i|$. Clearly this is a tricky procedure to perform explicitly, as we must search over all possible

decompositions, a difficult thing to parametrize let alone actually carry out. Fortunately, explicit formulas for various entanglement monotones have been worked out, so that this explicit procedure is generally not necessary in the two-qubit case.

A famous consequence of this bipartite entanglement is the famous "paradox" put forward by Einstein, Podolsky and Rosen in 1935, in an attempt to show that quantum mechanics was incomplete. In 1964 Bell analysed this thought experiment and derived inequalities that would be satisfied by the classical hidden-variable models proposed by EPR, but could be violated in quantum mechanics. Such violations have since been observed in many systems, providing some of the most convincing evidence of the reality of quantum mechanics. Recently, violations of Bell's inequality have been observed in superconducting phase qubits[3], proving that these macroscopic circuits do indeed behave quantum mechanically.

4.2 Three's a Crowd

As soon as we move from two qubits to three, the physics of entanglement becomes richer and more complicated. In particular, it has been shown that there are two different types of three-qubit entangled states, which unlike the Bell states are not local-unitary equivalent to each other. These two prototypical entangled states are $|\mathrm{GHZ}\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$, named for Greenberger, Horne and Zeilinger,

who first described the nonclassical correlations of this state[13], and the state $|\mathrm{W}\rangle = (|001\rangle + |010\rangle + |100\rangle)/\sqrt{3}$ also known as the Werner state after R. Werner, who has explored families of related states[39]. Dür, Vidal and Cirac have shown that any truly three-qubit state can be converted to one or the other of these two via local operations and classical communication, making this an exhaustive classification of the types of three-qubit entanglement[9].

One way to understand the difference between these two states is to consider what happens when one of the three qubits is lost from the state, or rather, in the case of superconducting circuits, where one qubit is simply ignored. Mathematically, losing or ignoring one qubit is equivalent to performing a "partial trace" over that qubit[29]. Suppose we label the three qubits as A, B and C and then trace over qubit C. The reduced density matrix for qubits A and B is written as $\rho^{AB} \equiv \mathrm{Tr}_C(\rho^{ABC})$. The partial trace operator Tr_C is a linear operator defined by its action on eigenvectors:

$$\begin{aligned} \mathrm{Tr}_C(|a_1b_1\rangle\langle a_2b_2| \otimes |c_1\rangle\langle c_2|) &= |a_1b_1\rangle\langle a_2b_2| \,\mathrm{Tr}(|c_1\rangle\langle c_2|) && (4.1) \\ &= |a_1b_1\rangle\langle a_2b_2| \cdot \langle c_2|c_1\rangle && (4.2) \end{aligned}$$

where the $|a_ib_i\rangle$ are arbitrary eigenvectors in the subspace of qubits A and B and where the $|c_i\rangle$ are arbitrary eigenvectors in the subspace of qubit C. This definition gives a very simple rule for calculating the partial trace for pure states, we simply write out the state as a density matrix with terms of the form $|a_1b_1c_1\rangle\langle a_2b_2c_2|$ and

then cross out all those terms for which $c_1 \neq c_2$, since in that case $\langle c_2|c_1\rangle = 0$.

Following this simple rule, we find the reduced density matrices after tracing out one qubit for $|\mathrm{GHZ}\rangle$:

$$\rho_{\mathrm{GHZ}}^{AB} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) \tag{4.3}$$

$$= \frac{1}{2}\begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix} \tag{4.4}$$

and for $|\mathrm{W}\rangle$:

$$\rho_{\mathrm{W}}^{AB} = \frac{1}{\sqrt{3}}(|00\rangle\langle 00| + |01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|) \tag{4.5}$$

$$= \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & & & \\ & 1 & 1 & \\ & 1 & 1 & \\ & & & 0 \end{pmatrix}. \tag{4.6}$$

As can readily be calculated from these expressions, ρ_{GHZ} is a purely mixed state, meaning that by tracing over one qubit the entanglement of the state is completely destroyed. Hence the entanglement in the $|\mathrm{GHZ}\rangle$ state is purely three-qubit entanglement, and is fragile with respect to loss (or decoherence) in any one of the qubits. On the other hand, ρ_{W}^{AB} is partially entangled, meaning that even when one qubit is lost, some bipartite entanglement remains. The entanglement in $|\mathrm{W}\rangle$ is more robust in this regard. This clearly illustrates that these two states represent different kinds of tripartite entanglement.

4.2.1 Quantifying Three-Qubit Entanglement

Because there are two different types of three-qubit entanglement, the construction of entanglement monotones for three qubits is somewhat more difficult than in the two-qubit case. Some measures of three-qubit entanglement have been developed, such as the three-tangle τ_3[7], however these do not completely characterize the entanglement. For example the $|\mathrm{GHZ}\rangle$ state has $\tau_3 = 1$, but the $|\mathrm{W}\rangle$ state is found to have $\tau_3 = 0$, so this measure fails to identify the entangled $|\mathrm{W}\rangle$-state as being entangled.

Another difficulty with the three-tangle is that it is rather difficult to calculate for mixed states such as are obtained from experimental data. Several works have found analytic expressions for the three-tangle in particular cases of certain types of mixed states[10], but no such expression for general mixed states is known. We saw in §4.1 that the convex roof extension can be used to extend pure-state entanglement measures such as the tangle to all mixed states. Unfortunately, calculating it requires that one explicitly perform the minimization over all possible pure-state decompositions. The difficulty of doing this numerically, even using sophisticated approximation schemes[6], makes the three-tangle an inconvenient measure to apply to experiments.

4.2.2 Witness Operators

A different tool for understanding entanglement of three and more qubits is the use of so-called "entanglement witnesses". An entanglement witness is an operator $\mathcal{W}$ which has the property that $\langle \mathcal{W} \rangle = \mathrm{Tr}(\rho \mathcal{W}) \geq 0$ for any separable state ρ, so that $\mathrm{Tr}(\rho \mathcal{W}) < 0$ implies that ρ is an entangled state. Note that $\mathrm{Tr}(\rho \mathcal{W}) < 0$ is a sufficient but not a necessary condition for identifying the state ρ as an entangled state. In other words, any given entanglement witness operator $\mathcal{W}$ will identify some but not all entangled states, and in general many different entanglement witness operators exist.[1]

Some entanglement witnesses are designed to be easily-measurable, so that for example one can identify entanglement without performing a complete reconstruction of the density matrix ρ, but instead performing a small set of measurements to compute the witness operator $\langle \mathcal{W} \rangle$ directly[36]. Measuring such witnesses can be done very efficiently, but requires particular measurement operators that may not always be available. For our purposes we are not concerned with efficiency, rather with having a simple tool to determine whether a measured density matrix is in fact an entangled state.

Another set of witness operators serves this purpose nicely. Witnesses can be

[1] We saw the same situation with the three-tangle, which identifies $|\mathrm{GHZ}\rangle$ but fails to identify $|\mathrm{W}\rangle$ as an entangled state. The advantage of entanglement witnesses, however, is that they are much easier to compute.

constructed that have the form $\mathcal{W}_\psi = \alpha I - |\psi\rangle\langle\psi|$ where $|\psi\rangle$ is a pure state and α is an appropriately-chosen coefficient. We can then easily compute that $\langle\mathcal{W}_\psi\rangle = \alpha - \langle\psi|\rho|\psi\rangle$ and the condition for witnessing entanglement becomes $F_\psi(\rho) \equiv \langle\psi|\rho|\psi\rangle > \alpha$ where the quantity $F_\psi(\rho)$ is the fidelity of the state ρ compared to the ideal state ψ. In practice the state ψ is typically a maximally-entangled state and the coefficient α is a threshold for the maximum fidelity achieved by a separable state, so that any state with higher fidelity must be entangled. Exactly why such witness operators exist and how they can be constructed—in particular, how the coefficient α is determined—is beyond the scope of the present discussion, but it essentially comes down to the fact that product states form a convex set, and thus there can exist hyperplanes defined by $\langle\psi|\rho|\psi\rangle = \alpha$ such that all product states lie on one side of the plane, so everything on the other side must be entangled.

For the three-qubit states of interest the relevant witnesses are[1]

$$\mathcal{W}_{\mathrm{W}} = \frac{2}{3} - |\mathrm{W}\rangle\langle\mathrm{W}| \tag{4.7}$$

$$\mathcal{W}_{\mathrm{GHZ}_1} = \frac{1}{2} - |\mathrm{GHZ}\rangle\langle\mathrm{GHZ}| \tag{4.8}$$

$$\mathcal{W}_{\mathrm{GHZ}_2} = \frac{3}{4} - |\mathrm{GHZ}\rangle\langle\mathrm{GHZ}|\,. \tag{4.9}$$

We have changed the names of these witnesses from those given in the reference because our emphasis is slightly different. For our purposes, $\mathcal{W}_{\mathrm{W}}$ is a witness that distinguishes $|\mathrm{W}\rangle$-like states from separable states and likewise $\mathcal{W}_{\mathrm{GHZ}_1}$ is

a witness that distinguishes $|\mathrm{GHZ}\rangle$-like states from separable states. Recalling from above that these witnesses are satisfied for fidelities greater than α, we see that in order to prove entanglement we must have the experimental fidelities $F_{\mathrm{W}} = \langle \mathrm{W}| \rho_{\mathrm{W}} |\mathrm{W}\rangle > 2/3$ and $F_{\mathrm{GHZ}} \equiv \langle \mathrm{GHZ}| \rho_{\mathrm{GHZ}} |\mathrm{GHZ}\rangle > 1/2$.

The final witness operator $\mathcal{W}_{\mathrm{GHZ}_2}$ is a stronger condition that distinguishes the class of $|\mathrm{GHZ}\rangle$-like states from the class of $|\mathrm{W}\rangle$-like states. Experimentally this requires a fidelity $F_{\mathrm{GHZ}} > 3/4$ in order to make this separation. Remember though that entanglement witnesses always give sufficient but not necessary conditions for identifying and classifying entanglement. A different measure of the tripartite entanglement that can also distinguish the $|\mathrm{W}\rangle$- and $|\mathrm{GHZ}\rangle$-like states is provided by the Mermin-Bell inequality, which we discuss in the next section.

4.2.3 Classical versus Quantum Correlation

At the beginning of this chapter, we defined entanglement as the property of not being expressible as a product state or a mixture of product states. A more operational way to classify quantum states is to consider the correlations in measurements performed on these states; such measurements are after all the only way to extract information about the state. Werner[39] considered the question of what types of correlations can be reproduced by classical means, versus correlations that are purely quantum. Once we have separated states into "classically corre-

lated" and "quantum correlated", we can then ask whether they can be explained by hidden variable theories, of the type proposed by EPR and considered by Bell. Not surprisingly, all classically correlated states admit hidden variable models. Also not surprisingly, states that do not have hidden variable models (such as the Bell states) are quantum correlated. Werner's interesting result was to show that some states are quantum correlated, but can nevertheless be explained by hidden variable models.

It turns out that the two types of three-qubit entangled states differ from each other in this regard. While both of them are entangled states and hence exhibit "quantum correlations", the $|\mathrm{W}\rangle$ state *can* be explained by a hidden variable model, while the $|\mathrm{GHZ}\rangle$ state cannot. We saw above in Equation 4.8 that the entanglement witness $\mathcal{W}_{\mathrm{GHZ}_1}$ could distinguish $|\mathrm{GHZ}\rangle$-like states from separable states, but could not distinguish them in general from $|\mathrm{W}\rangle$-like states. The non-classicality of the $|\mathrm{GHZ}\rangle$-like states gives us another tool for distinguishing these two classes.

The non-classicality of $|\mathrm{GHZ}\rangle$ was first demonstrated by Greenberger, Horne and Zeilinger, for whom the state is named[13]. Interestingly, they found that with this state it would be possible to rule out hidden variable models with just four measurements, rather than by taking many measurements and looking at correlations among them, as with Bell's Inequality. This leads to a "Bell Violation

without Inequalities" in their words. Unfortunately, producing such a violation would require an ideal pure state and perfect measurement fidelity, neither of which are experimentally feasible. In practice we must still make many measurements, calculate some correlations, and show that they violate an inequality to achieve a violation.

The basic idea behind the GHZ proposal can be very easily understood. Suppose we have three qubits in the state $|000\rangle + |111\rangle$. We split these qubits up and give them to three separate observers, whom we then instruct to make various measurements. We assume that we have several identical copies of the state, so that we can repeat this process as needed. The observers are instructed to make the following measurements in four rounds of the experiment:

1. Each observer measures X.

2. Observer A measures X, B and C measure Y.

3. Observer B measures X, A and C measure Y.

4. Observer C measures X, A and B measure Y.

For the observables X and Y, the possible measurement outcomes for the various observers are the eigenvalues ± 1. In each round, we multiply together the results of the three observers to obtain an overall correlation which itself is equal to ± 1.

A		B		C		M_{XYY}	M_{YXY}	M_{YYX}	Π	M_{XXX}
X	Y	X	Y	X	Y					
+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	-	-	-	+	+	+
+	+	+	-	+	+	-	+	-	+	+
+	+	+	-	+	-	+	-	-	+	+
+	-	+	+	-	+	+	-	+	-	-
+	-	+	+	-	-	-	+	+	-	-
+	-	+	-	-	+	-	-	-	-	-
+	-	+	-	-	-	+	+	-	-	-

Figure 4.1: Hidden variable model for GHZ measurement.

What is the expected outcome with a hidden variable theory? In this case, once the qubits have been separated, the system must have decided what outcome it will give for each possible choice of measurement on each qubit, as determined by the hidden-variable state of the system. In Figure 4.1, we show some of the possible assignments for measurement outcomes. On the left are the measurement results for each qubit A, B or C, and each measurement setting X or Y. Let us denote each such number by m_X^A, where the subscript is the measurement setting, and the superscript is the qubit. In the middle section of the figure, we show several correlations which are just the products of the measurement outcomes: $M_{XYY} = m_X^A m_Y^B m_Y^C$ and so on. The last column of this section gives the product of the three previous correlations $\Pi \equiv M_{XYY} M_{YXY} M_{YYX}$. Finally, the far right column shows the correlation M_{XXX}.

Note that in all cases shown in the figure, we have $\Pi = M_{XXX}$. It is easy to prove that this must be the case no matter what measurement outcomes the system chooses. We have that

$$\begin{aligned}
\Pi &= M_{XYY} \cdot M_{YXY} \cdot M_{YYX} \\
&= m_X^A m_Y^B m_Y^C \cdot m_Y^A m_X^B m_Y^C \cdot m_Y^A m_Y^B m_X^C \\
&= m_X^A (m_Y^A)^2 \cdot m_X^B (m_Y^B)^2 \cdot m_X^C (m_Y^C)^2 \\
&= m_X^A m_X^B m_X^C \\
&= M_{XXX}
\end{aligned} \tag{4.10}$$

where we have used the fact that $m = \pm 1$ so that $m^2 = 1$. Hence in a hidden variable model, the product Π of correlations with two Y measurements must equal the correlation M_{XXX} with all X measurements.

What is the expected outcome with quantum mechanics? Consider the first round in which all observers measure X. Multiplying the results from each observer together to obtain the correlation M_{XXX} is equivalent to measuring the combined operator XXX. We can easily compute that $M_{XXX} = \langle XXX \rangle_{|\mathrm{GHZ}\rangle} = 1$. For the later rounds we find instead that $M_{XYY} = \langle XYY \rangle_{|\mathrm{GHZ}\rangle} = -1$ and by symmetry also $M_{YXY} = M_{YYX} = -1$. Hence we obtain $\Pi = M_{XYY} \cdot M_{YXY} \cdot M_{YYX} = -1 \neq M_{XXX}$, in direct contradiction to the classical case.

The argument as presented here has a few glaring weaknesses. First, note that the four required correlations are not measured at once, but rather in four separate iterations of the experiment. In that case, the hidden variables might

take on one set of values in one iteration, and a different set of values in another iteration, potentially giving a result $\Pi \neq M_{XXX}$ as in the quantum case. Second, and perhaps more important for an experimental test of this result, is the fact that the argument as presented here assumes pure states and perfect measurements, neither of which are experimentally feasible. Fortunately, by reformulating the problem to account for the latter problem, we can solve the former as well.

Mermin[24] reformulated this test of non-classicality for $|\text{GHZ}\rangle$ in terms of inequalities that apply in the realistic mixed-state case. He gave a family of inequalities that are obeyed by hidden variable models, but violated by $|\text{GHZ}\rangle$ and its generalizations to higher numbers of qubits. He also showed that the potential violation grows exponentially with the number of qubits. For the three-qubit case considered here, classical hidden variable models must obey

$$G \equiv \langle XXX \rangle - \langle XYY \rangle - \langle YXY \rangle - \langle YYX \rangle \leq 2 \tag{4.11}$$

where the correlation operator G is defined in terms of the four measured expectation values. In this case, no assumptions are made about what values the hidden variables take in each iteration of the experiment, so that we are not subject to the objections raised in the previous paragraph. Unlike hidden variable models, quantum mechanics allows this inequality to be violated. For the pure $|\text{GHZ}\rangle$ state we have $G_{|\text{GHZ}\rangle} = 4 \not\leq 2$; for our experimental mixed state, the violation will be smaller, but still an indicator of non-classicality, as desired. This inequality is

in fact violated experimentally, as we will see in Chapter 8.

Chapter 5

Design of Entangling Protocols

In the previous chapter we discussed the two types of three-qubit entangled states. We now consider the task of generating these states: what sequence of qubit operations will produce these entangled states, and can these protocols be implemented in a way that is realizable using our phase qubits? The following sections will derive the necessary sequences of gates, and in the next chapter we will see in detail how the coupler has been designed to enable the necessary operations to be performed.

5.1 Multi-step GHZ

5.1.1 Using CNOT gates

The simplest entangling protocol to understand is that for generating the state $|\text{GHZ}\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$. Starting in the ground state, we first perform a $\pi/2$-rotation on one qubit—qubit A, say—to put the system in state $(|000\rangle + |100\rangle)/\sqrt{2}$. Now, we would like to map the second component of this superposition $|100\rangle$ to the state $|111\rangle$ without changing the other component $|000\rangle$. In other words, we need to flip the state of qubits B and C, but only when qubit A is itself in state $|1\rangle$. This is exactly the operation performed by the CNOT gate, so we can generate the desired $|\text{GHZ}\rangle$ state simply by applying two CNOT gates to flip qubits B and C conditioned on the state of qubit A. These CNOTs can be applied in various ways, either using qubit A as the control qubit for both gates, or else using the flipped qubit in the first gate as the control for the second; either protocol results in the same final state. Another variation has both CNOT gates acting on the same control and target qubits, but with a swap gate applied in the middle to ensure that both targets are flipped[21]. A quantum circuit for the basic sequence is shown on the left in Figure 5.1.

This protocol can be readily generalized in two ways. First, as shown on the right in Figure 5.1, we can eliminate the single qubit rotation and instead accept

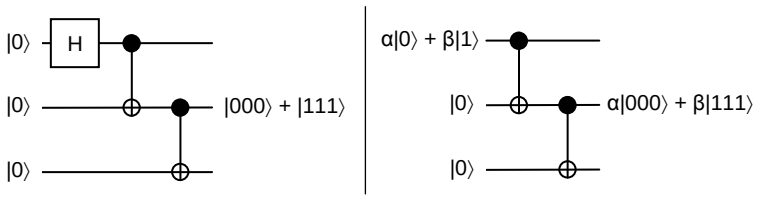

Figure 5.1: GHZ protocol using CNOT gates.

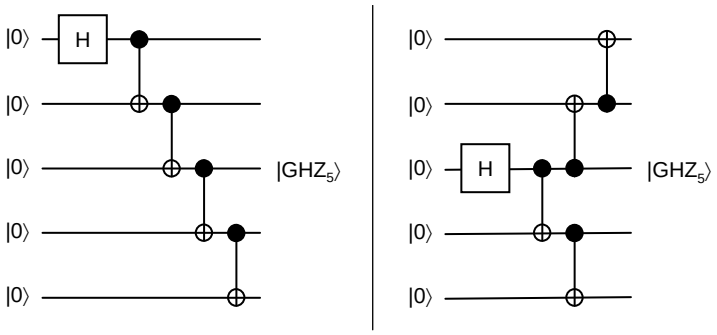

Figure 5.2: Generating GHZ states of an arbitrary number of qubits.

an arbitrary initial state of qubit A: $\alpha\,|0\rangle + \beta\,|1\rangle$. The other two qubits are assumed to start in the ground state so that the collective starting state of the full system is $\alpha\,|000\rangle + \beta\,|100\rangle$. Now, applying two CNOT gates will result in the generalized GHZ state $\alpha\,|000\rangle + \beta\,|111\rangle$. This effectively encodes the single-qubit state in an entangled state that can be used to detect and correct single bit flips using the so-called "repetition code"[29, chap. 10].

The GHZ protocol can also be generalized to arbitrary numbers of qubits simply by adding more CNOT gates, as shown in Figure 5.2. With N qubits a total of $N-1$ CNOT gates must be applied to flip the states of all the other qubits

conditioned on the control qubit that is first rotated. These gates could be applied one after another, taking a total time proportional to the number of qubits N, as shown on the left. Alternatively the gates could be done in parallel, using all flipped qubits as controls for the next stage, in which case the time required would scale as $\log N$, as shown on the right. Such parallelism can often lead to gains in efficiency for quantum circuits, but of course the physical system being used would have to support parallel gates and coupling between potentially widely-separated qubits, and these requirements that may not be satisfied in physical implementations. We will focus on the cases $N = 3$ and $N = 4$ and hence ignore these larger scaling issues.

5.1.2 Recompiling for iSWAP gates

We saw in §2.2 that the natural entangling gate generated by the $XX + YY$ interaction is not the CNOT gate, but rather the iSWAP gate. It can be shown that the iSWAP gate is equivalent to a so-called "CNOT-SWAP" gate, consisting of a CNOT followed by a SWAP[32]. Fortunately, it is easy to recompile the quantum circuits shown in the previous section for $|\text{GHZ}\rangle$ state generation in terms of CNOT-SWAP gates. To see how this is done, note that the $|\text{GHZ}\rangle$ state is invariant under permutations of the qubits, so that any swaps introduced into the circuit after the entangling is accomplished have no effect and can be

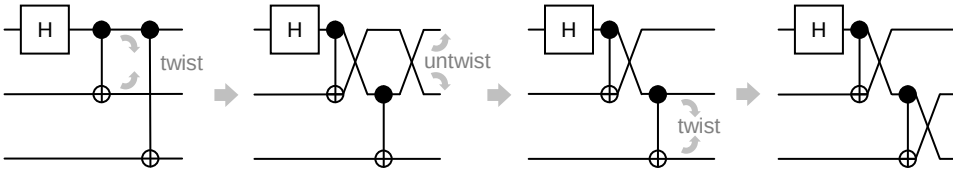

Figure 5.3: Modifying the GHZ protocol to use CNOT-SWAP gates.

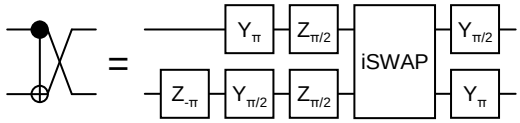

Figure 5.4: CNOT-SWAP gate expressed in terms of iSWAP.

ignored. We can thus twist and bend the wires in our quantum circuit as needed to introduce swaps after the CNOT gates, and then discard any remaining SWAPs at the end as these do not change the final state. This process and the resulting circuit are shown in Figure 5.3.

To complete the transformation into an iSWAP implementation that will be easy to implement, we replace each CNOT-SWAP by an equivalent circuit composed of an iSWAP gate produced by the $XX + YY$ interaction plus single-qubit

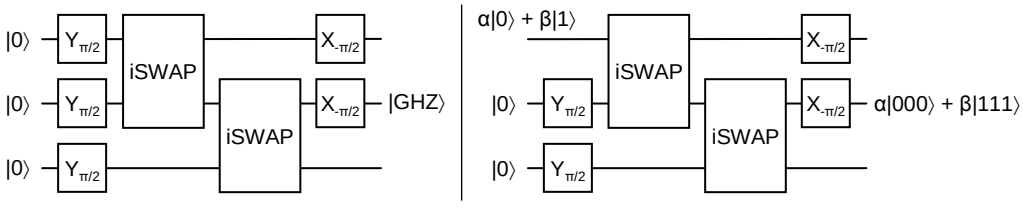

Figure 5.5: Circuit for generating GHZ state with iSWAP gates.

unitaries. The circuit identity, which can be readily derived using the techniques of Makhlin[19], is shown in Figure 5.4. This identity is by no means unique; there are many different ways to decompose the single-qubit unitary operations appearing before and after the iSWAP gate. But if we simply replace the CNOT-SWAP gates in Figure 5.3 with this iSWAP identity, various simplifications can be applied to eliminate neighboring single-qubit unitaries and simplify the overall circuit. In the end, we arrive at the circuit shown in Figure 5.5 for producing $|\text{GHZ}\rangle$ using only operations that are easily realizable experimentally, namely single-qubit rotations and iSWAP gates. As with the straightforward CNOT protocol described in the previous section, this protocol generalizes to arbitrary input states (Figure 5.5, right) and it can also be generalized to larger numbers of qubits.

5.2 Single-step GHZ

The $|\text{GHZ}\rangle$ state-generation protocols discussed in the previous section all utilize single-qubit rotations and two-qubit gates, which are well-defined and easily-realizable building blocks. However, given the symmetry of the start state $|000\rangle$ and end state $|000\rangle + |111\rangle$ with respect to permutations of the qubits, we might wish to find a protocol that explicitly respects the symmetry of the states. For example we might wish to apply the same rotations to all qubits or couple all the qubits simultaneously rather than using just two-qubit gates, potentially reduc-

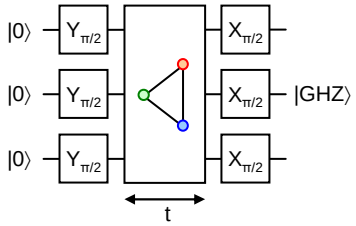

Figure 5.6: GHZ state generation with a single entangling step.

ing the requirements for coherent control of the system. In this particular case of $|\text{GHZ}\rangle$ state generation, a solution was found for $N = 3$ qubits by Galiautdinov and Martinis[12] and then extended to arbitrary N by Galiautdinov *et al.*[11]. The protocol for $N = 3$ is shown in Figure 5.6.

In this protocol, the first step is to create an equal superposition of all qubit states by applying $\pi/2$ rotations about Y to each qubit. The second step is to turn on the interaction to entangle the qubits. The interaction is assumed to be of the form $(g/2)(XX+YY)$, with equal coupling strength g between each pair of qubits[1]. Finally another $\pi/2$ rotation is applied to each qubit to bring the system into the desired $|\text{GHZ}\rangle$ state. The time t for which the coupling is turned on is $t = \pi/2g$, the same time as required to produce a single iSWAP gate between just two of the qubits. Hence this protocol requires half as much time as those presented in the previous section using two separate two-qubit gates.

While we might expect this reduction in the required entangling time to im-

[1]We drop the additional ZZ term considered in the papers, as this is not present in our capacitively-coupled devices.

prove the performance of this protocol compared to those using two-qubit gates, we must consider the fact that this single-step protocol works in a very different way than the two-qubit versions. During the entangling step, the coupling interaction does not actually change the population of the various eigenstates, rather it simply changes the phases of those states. These phase changes have the effect that when the final rotations are applied, $|000\rangle$ and $|111\rangle$ are populated while all other qubit states are depopulated, thus rotating the state into $|\mathrm{GHZ}\rangle$. Proper functioning of the protocol requires that the phases be tuned precisely, so that the final rotations work as desired. Also, note that immediately after the first rotations are applied the system is in an equal superposition of all possible qubit states. Both of these features, the precise phase tuning and the occupation of all qubit states, will prove problematic for the experimental implementation of this protocol with phase qubits, as we will see in Chapter 8.

5.3 Single-step W

We now turn to protocols for generating the $|\mathrm{W}\rangle$ state. While it is straightforward to write down sequences for generating this state with two-qubit gates as in §5.1, it is much more natural to consider the single-step approach of §5.2 that takes into account the symmetries of the state. The state $|\mathrm{W}\rangle = (|100\rangle + |010\rangle + |001\rangle)/\sqrt{3}$ is a superposition of three components, each with one qubit excited. To generate

it we must "share" a single excitation among the three qubits. The transverse $XX + YY$ coupling interaction does exactly this: it causes an excitation to be swapped from one coupled qubit to another, with the excitation shared between the entangled qubits at intermediate times. Thus our technique to create the $|\mathrm{W}\rangle$ state will be as follows: first put one excitation into the system and then couple *all* the qubits together symmetrically to distribute the excitation among them. This simple idea leads to a fast and efficient single-step protocol for generating the $|\mathrm{W}\rangle$ state.

We can work out the details of this protocol in the general case of any number of qubits N. If we label the qubits by integer indices $0, 1, \ldots N-1$ then the generalized $|\mathrm{W}\rangle$-state can be written as

$$|\mathrm{W}_N\rangle \equiv \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |0_0 0_1 \ldots 0_{i-1} 1_i 0_{i+1} \ldots 0_{N-1}\rangle \tag{5.1}$$

$$= \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |\psi_i\rangle \tag{5.2}$$

which is a superposition of N terms, where the ith term has a single excitation in qubit i and where for brevity we write these terms as $|\psi_i\rangle$. Starting from the ground state, the first step in the protocol is to introduce one excitation into the system, so we apply a π-pulse to qubit 0 to create the state $|10\ldots0\rangle = |\psi_0\rangle$.

Next the coupling interaction is turned on between all pairs of qubits. Because of the symmetry of the coupling circuit, the amplitude for the excitation to be in

any of the other qubits $1, \ldots, N-1$ will be equal, so that as a function of time after the coupling is turned on, the system state will have the form

$$a(t)\,|\psi_0\rangle + b(t)\sum_{i=1}^{N-1}|\psi_i\rangle\,, \tag{5.3}$$

where $a(t)$ and $b(t)$ are complex amplitudes to be determined. This argument assumes that after the coupling is turned on the system stays within the subspace of single-excitation states $|\psi_i\rangle$, without straying into the ground state or states with more than one excitation such as $|110\ldots0\rangle$ or $|20\ldots0\rangle$. This constraint is satisfied since the $XX+YY$ swap interaction preserves photon number and simply swaps excitations without creating or destroying them. If the coupling time t_N can be chosen so that $|a(t_N)| = |b(t_N)| = 1/\sqrt{N}$, then the system will be in an equal superposition of single-excitation states, which differs from the desired $|\mathrm{W}_N\rangle$ state only by the possible phase difference between a and b. The phase difference can be corrected by rotating qubit 0 about the Z-axis by an appropriate angle θ_N. Any remaining global phase is of no physical significance, so the final state is equal to $|\mathrm{W}_N\rangle$, as desired.

To find t_N and θ_N we consider the Hamiltonian of the system. As mentioned in the previous paragraph, once in the single-excitation subspace, the system stays there because the coupling Hamiltonian preserves photon number. Hence we can consider only the reduced space of single-excitation states $|\psi_i\rangle$. Because we turn on a coupling interaction $(g/2)(XX+YY)$ between all pairs of qubits, the coupling

Hamiltonian in this restricted subspace has terms for exchanging the excitation between each pair of qubits: $H^1_{\text{int}} = g\sum_{i\neq j} |\psi_i\rangle\langle\psi_j|$, or in matrix form

$$H^1_{\text{int}} = \begin{pmatrix} 0 & g & g & \dots & g \\ g & 0 & g & & g \\ g & g & 0 & & g \\ \vdots & & & \ddots & \vdots \\ g & g & g & \dots & 0 \end{pmatrix}. \tag{5.4}$$

One can easily verify that the eigenstates of this restricted Hamiltonian are $|\phi_+\rangle \equiv (|\psi_0\rangle + |\psi_1\rangle + \ldots + |\psi_{N-1}\rangle)/\sqrt{N}$ with eigenvalue $(N-1)g$, and the set $|\phi_{-k}\rangle \equiv (|\psi_0\rangle - |\psi_k\rangle)/\sqrt{2}, k \in \{1,\ldots,N-1\}$ with eigenvalue $-g$.

In this basis of eigenvectors, the system state at the beginning of the coupling period is $|\phi(t=0)\rangle = |\psi_0\rangle = (\sqrt{N}\,|\phi_+\rangle + \sqrt{2}\sum_{k=1}^{N-1}|\phi_{-k}\rangle)/N$. With the coupling turned on, the eigenvectors simply acquire phase at a rate corresponding to their energy, so the time evolution of this state is

$$|\phi(t)\rangle = (\sqrt{N}e^{-i(N-1)gt}\,|\phi_+\rangle + \sqrt{2}e^{igt}\sum_{k=1}^{N-1}|\phi_{-k}\rangle)/N. \tag{5.5}$$

We can now take the overlap with the single-excitation qubit states to find the occupation probabilities of the various states. Due to symmetry we need only consider two possibilities, the initially-excited qubit and one other qubit since all the rest will have the same occupation. For the initially-excited qubit 0, we have

$$
\begin{aligned}
\langle\psi_0|\phi(t)\rangle &= \frac{1}{N}\left(\sqrt{N}e^{-i(N-1)gt}\langle\psi_0|\phi_+\rangle + \sqrt{2}e^{igt}\sum_{k=1}^{N-1}\langle\psi_0|\phi_{-k}\rangle\right) \\
&= \frac{1}{N}\left(\sqrt{N}e^{-i(N-1)gt}\frac{1}{\sqrt{N}} + \sqrt{2}e^{igt}\sum_{k=1}^{N-1}\frac{1}{\sqrt{2}}\right) \\
&= \frac{1}{N}\left(e^{-i(N-1)gt} + (N-1)e^{igt}\right) \\
&= \frac{e^{igt}}{N}\left((N-1) + e^{-iNgt}\right). \qquad (5.6)
\end{aligned}
$$

From Equation 5.6 we can then calculate the occupation probability of the zeroth qubit:

$$
\begin{aligned}
P_{10\ldots0} &= |\langle\psi_0|\phi(t)\rangle|^2 \\
&= \frac{1}{N^2}\left((N-1) + e^{-iNgt}\right)\left((N-1) + e^{iNgt}\right) \\
&= \frac{1}{N^2}\left((N-1)^2 + 1 + 2(N-1)\cos Ngt\right). \qquad (5.7)
\end{aligned}
$$

The occupation probability $P_{10\ldots0}$ oscillates in time at angular frequency Ng and is bounded from below by $P_{10\ldots0}(t) \geq (N-2)^2/N^2$. In order to create a $|\mathrm{W}\rangle$ state, we must have $P_{10\ldots0}(t_N) = 1/N$, so this bound on the probability immediately tells us that our protocol will not work for arbitrary numbers of qubits, but only when $(1-2/N)^2 \leq 1/N$. This has solutions only for $N = 2$, 3, or 4; in the case $N = 5$ or higher, an equal superposition is never achieved so the $|\mathrm{W}\rangle$-state cannot be created with this protocol.

In a similar fashion, we calculate the amplitude and occupation probability for one of the initially unexcited qubits, say qubit 1. The amplitude is

$$\begin{aligned}
\langle\psi_1|\phi(t)\rangle &= \frac{1}{N}\left(\sqrt{N}e^{-i(N-1)gt}\,\langle\psi_1|\phi_+\rangle + \sqrt{2}e^{igt}\sum_{k=1}^{N-1}\langle\psi_1|\phi_{-k}\rangle\right) \\
&= \frac{1}{N}\left(\sqrt{N}e^{-i(N-1)gt}\frac{1}{\sqrt{N}} - \sqrt{2}e^{igt}\frac{1}{\sqrt{2}}\right) \\
&= \frac{1}{N}\left(e^{-i(N-1)gt} - e^{igt}\right) \\
&= \frac{e^{igt}}{N}\left(e^{-iNgt} - 1\right) \qquad (5.8)
\end{aligned}$$

and the occupation probability is

$$\begin{aligned}
P_{010\ldots0} &= |\,\langle\psi_1|\phi(t)\rangle\,|^2 \\
&= \frac{1}{N^2}\left(e^{-iNgt} - 1\right)\left(e^{iNgt} - 1\right) \\
&= \frac{1}{N^2}\left(2 - 2\cos Ngt\right). \qquad (5.9)
\end{aligned}$$

We now consider the three allowed values of N that will generate $|\mathrm{W}\rangle$ states with this method. For $N = 2$, not surprisingly, we find that an equal superposition is achieved when $2gt = \pi/2$ so that $t_2 = \pi/4g$. This is exactly half of the swap time $t_{\mathrm{iSWAP}} = \pi/2g$ already seen above in the GHZ protocol in §5.2. For $N = 3$, the equal superposition is achieved when $(2/9)(1 - \cos 3gt) = 1/3$, which implies $\cos 3gt = -1/2$. Hence, the equal superposition is obtained when $3gt = 2\pi/3$ giving a time of $t_3 = 2\pi/9g = (4/9)t_{\mathrm{iSWAP}}$. Finally, for $N = 4$ we obtain the equal

N	t_N	θ_N
2	$\pi/4g$	$\pi/2$
3	$2\pi/9g$	$4\pi/3$
4	$\pi/4g$	π

Table 5.1: Parameters for W-state generation protocols.

superposition when $(2/16)(1-\cos 4gt) = 1/4$, giving $\cos 4gt = -1$ with solution $4gt = \pi$. Hence, the equal superposition is obtained at time $t_4 = \pi/4g = t_{\text{iSWAP}}/2$, as in the case of two qubits.

The other quantity that must be calculated to complete the protocol for generating the $|\text{W}\rangle$-state is the amount of Z-rotation needed on the initially-excited qubit to equalize the phases. This rotation is in a sense optional, since local rotations do not change any entanglement measures on a multi-qubit states, as discussed in Chapter 4. However, the required angle is simple to calculate by plugging in the coupling times obtained in the previous paragraph to the amplitudes in Equation 5.6 and Equation 5.8. For the $N = 3$ case of interest here, the amplitudes for qubits 0 and 1 are

$$\alpha_0(t_3) = (e^{4i\pi/9}/3)(e^{-4i\pi/3} + 2) \tag{5.10}$$
$$\alpha_1(t_3) = (e^{4i\pi/9}/3)(e^{-4i\pi/3} - 1). \tag{5.11}$$

Via straightforward algebra, we can verify that a Z-rotation by angle $\theta_3 = 4\pi/3$ will equalize the phase of these two states. Similarly, we can verify that $\theta_2 = \pi/2$ and $\theta_4 = \pi$. All these results are summarized in Table 5.1.

There is a simple geometric picture of the operation of these $|\mathrm{W}\rangle$-state protocols that makes clear how they work and allows the parameters t_N and θ_N to be determined easily. We simply take the amplitudes determined in Equation 5.6 and Equation 5.8 and plot them in the complex plane, ignoring the common prefactor e^{igt}/N. Both amplitudes consist of a stationary part plus a unit vector rotating at angular frequency Ng. These two parts of each amplitude are represented by the colored arrows in the left side in Figure 5.7, with the resultant vectors giving the amplitudes themselves, α_0 in purple and α_1 in blue. The circles represent the full range of variation of the resultant amplitudes. The coupling time t_N is determined by the time at which the two resultant vectors have equal length, as shown. On the right side of the figure, we show only the resultant vectors giving the full amplitudes; the phase correction angle θ_N is determined by the angle between the two vectors at that time. The diagrams confirm the results derived algebraically above, including the fact that no solutions exist for $N \geq 5$.

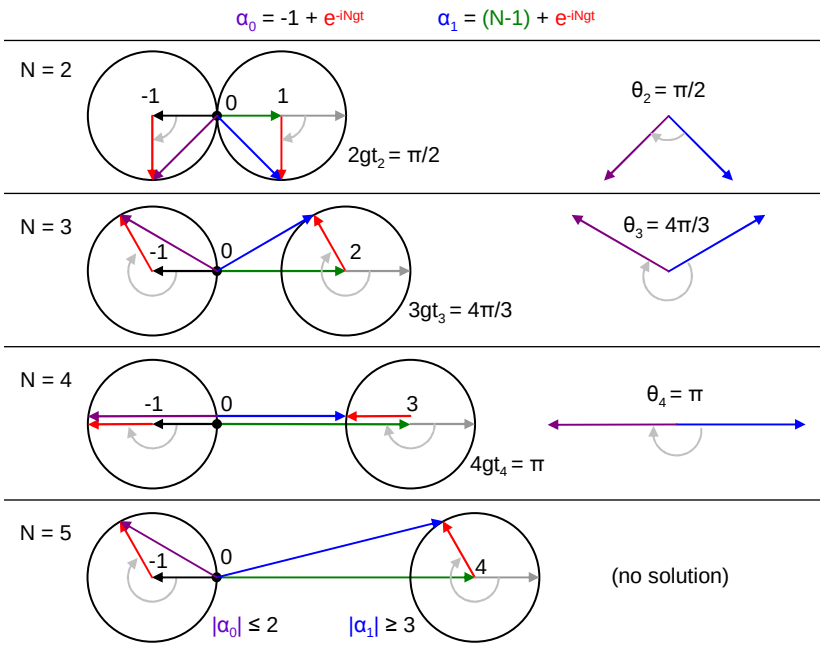

Figure 5.7: Geometric picture of W-state generation protocols.

Chapter 6

Coupling Circuit Design

In this chapter we will consider the design of the coupling circuit for our multi-qubit device. To implement the single-step entangling protocols described in Chapter 5, we require an interaction with equal coupling strength between each pair of qubits, in other words a coupling network that is a complete graph between the qubit nodes. In Chapter 2 we described various schemes for coupling qubits, including the simplest technique, which we will use here, that involves simply connecting the qubits with a capacitor. Our first attempt at a circuit to implement the complete graph coupling, shown in Figure 6.1, is a direct combination of these two ideas: we simply insert a capacitor between each pair of qubits that should be coupled.

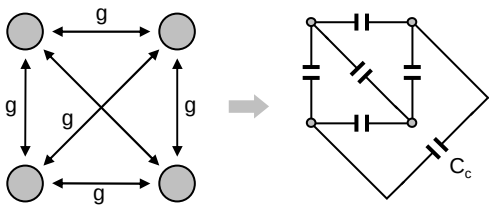

Figure 6.1: Translating a coupling graph into a capacitor network.

6.1 Improving the Coupling Network

The main problem with the direct approach to constructing the coupling network shown in Figure 6.1 is that the symmetry will be lost when we actually try to build the coupling network. A three qubit (triangular) network can be constructed without a problem, but as soon as we have four or more qubits simply laying out the coupling graph will require us to introduce crossovers or, as in the four-qubit case illustrated, to place one of the coupling capacitors on the outside, connected by long leads. Achieving exactly equal coupling strengths as required by the single-step entangling protocols will then only be possible by carefully tuning the coupling capacitors to account for wiring crossovers and other stray parasitic capacitances.

We would prefer a coupling circuit that can be laid out on chip in a symmetric configuration, as this will make it much easier to achieve the desired symmetric coupling. To see how this can be done, consider a well-known trick from conventional electrical engineering known as the "Star-Delta" or "Y-Delta"

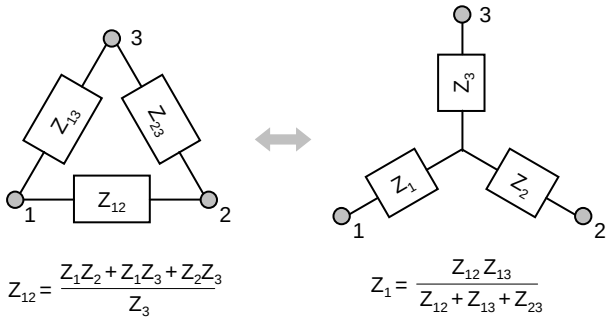

Figure 6.2: Star-Delta transformation of a three-node circuit.

transform[16]. As shown in Figure 6.2, this transforms a three-node circuit with arbitrary impedances between each node (the "Delta" configuration, a complete graph of three nodes) into an equivalent circuit with each node connected to a new internal node in the center (the "Star" configuration). In the specific case of our symmetric coupler with fixed capacitance C_Δ between each node, the equivalent Star configuration has capacitance $C_\star = 3C_\Delta$.

For three nodes, the Star configuration is no simpler than the Delta configuration, but in general the Star configuration for N nodes contains just N impedances, while the complete graph (which we will continue to refer to as the "Delta" configuration) contains $N(N-1)/2$. Hence for $N > 3$ the Star configuration requires fewer coupling links, and in addition it can always be laid out symmetrically as a planar graph with couplers radiating out from a central node, unlike the Delta configuration which becomes increasingly tangled as N increases.

Because there are only N impedances in the Star versus $N(N-1)$ in the

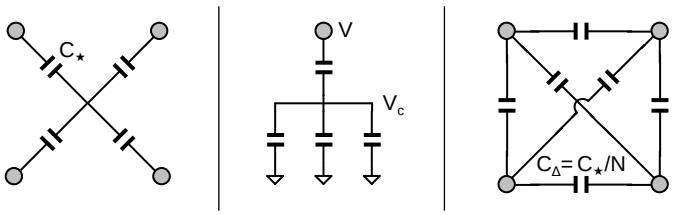

Figure 6.3: Symmetric Star-Delta transformation for arbitrary N.

complete graph, the general Star-Delta transform with arbitrary impedances is not possible for $N > 3$; there are simply not enough degrees of freedom in the Star configuration. But the case we are considering is highly symmetric, with equal coupling between each node. Is it possible to find an appropriate transformation into an equivalent Star circuit in this case?

To answer this question, we must first ask how to tell whether two capacitance networks are equivalent. Let us define the mutual capacitance C_{ij} between two nodes i and j in a network as the ratio $C_{ij} = Q_j/V_i$, where Q_j is the charge that appears on node j when a fixed voltage V_i is applied to node i, with all the other nodes grounded. Because a capacitor network is linear, knowing all the mutual capacitances is sufficient to completely characterize the network, so that if we can find a Star configuration with the same mutual capacitances as the complete graph, we will have found the solution.

For the complete graph network of capacitors, the mutual capacitance between any two nodes is, not surprisingly, simply given by the capacitor $C_{ij} = C_\Delta$

connecting the nodes. Now consider a Star network with N nodes where each node is connected to the central island with a capacitance $C_\star$, as shown in Figure 6.3 (left). By symmetry, it is clear that the mutual capacitances C_{ij} between any pair of nodes must all be the same. Applying a voltage V to one of the nodes with all the other nodes grounded forms a divider with two capacitors, $C_\star$ and $(N-1)C_\star$. The voltage on the floating central island will be $V_c = (C_\star/(C_\star + (N-1)C_\star))V = V/N$ (middle). Hence the charge appearing on any of the other nodes is simply $Q_j = C_\star V_c = C_\star V/N$ and the mutual capacitance is $C_{ij} = Q_j/V = C_\star/N$. Thus, to create a star network that is equivalent to the complete graph, we choose $C_\star = NC_\Delta$ (right), which reproduces the standard Star-Delta result when $N = 3$.

6.2 Effect of Island Loading

In a real device, the central island in our Star network will consist of a chunk of metal that may have some parasitic capacitance to ground. Depending on the configuration of the coupling network and ground planes, this capacitance can in fact be quite large, certainly large enough that it must be taken into account in order to engineer the desired coupling strength. Fortunately, the effect of this capacitance on the coupling network is easily calculated.

We can compute the mutual capacitance between two nodes in the Star net-

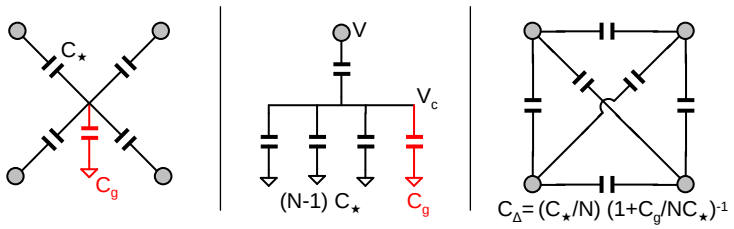

Figure 6.4: Star coupler with island capacitance.

work exactly as we did above, except that now an additional capacitance C_g is added from the island to ground, as shown in Figure 6.4. Applying voltage V to a particular node creates a voltage divider, and the central island voltage is $V_c' = (C_\star/(C_\star + (N-1)C_\star + C_g))V = V/N(1 + C_g/NC_\star)^{-1}$. The mutual capacitance is then

$$C_{ij}' = \frac{C_\star}{N}\left(1 + \frac{C_g}{NC_\star}\right)^{-1}. \tag{6.1}$$

The factor $(1 + C_g/NC_\star)^{-1}$ indicates that the effect of island capacitance is to "load" the circuit, reducing the mutual capacitance and hence the effective coupling strength. When $C_g \ll NC_\star$, the ground capacitance has little effect, but when $C_g \sim NC_\star$, the coupling can be substantially reduced.

If we set the mutual capacitance in Equation 6.1 equal to some desired coupling capacitance C_Δ and solve for $C_\star$, we obtain

$$C_\star = NC_\Delta \cdot \frac{1}{2}\left(1 + \sqrt{1 + \frac{4C_g}{N^2 C_\star}}\right). \tag{6.2}$$

Once the coupling capacitance C_Δ is chosen to give the desired qubit-qubit cou-

pling strength and the island capacitance C_g is determined from the geometry of the coupling circuit, this formula gives the required physical coupling capacitance $C_\star$.

6.3 Physical Layout and Parameter Estimation

The physical geometry of the coupling circuit as it appears on chip is shown in Figure 6.5. For the purpose of estimating the capacitance to ground, we model the island as a coplanar waveguide with a gap of $40\,\mu$m between ground planes and a total length of $200\,\mu$m, the sum of the lengths of the two arms of the cross minus the length of the central crossing region; for this geometry we find $C_g \approx 30\,\text{fF}$. In order to choose the coupling capacitance, we take the target mutual capacitance to be $C_\Delta = (S/f_{10}) \cdot C_q$ where $S = 2g/2\pi$ is the desired qubit-qubit swap frequency, $f_{10} \sim 6\,\text{GHz}$ is the qubit transition frequency and $C_q \sim 1\,\text{pF}$ is the qubit capacitance. This relationship together with Equation 6.2 allows us to determine the required island coupling capacitance $C_\star$ to achieve the desired swap frequency.

Table 6.1 shows various qubit-qubit swap frequencies S along with the mutual capacitance C_Δ and coupling capacitance $C_\star$ needed to produce that swap frequency. Note that the island capacitance to ground has a substantial effect. For all but the largest splitting we have $C_g = 30\,\text{fF} > 4C_\Delta$, so that without taking

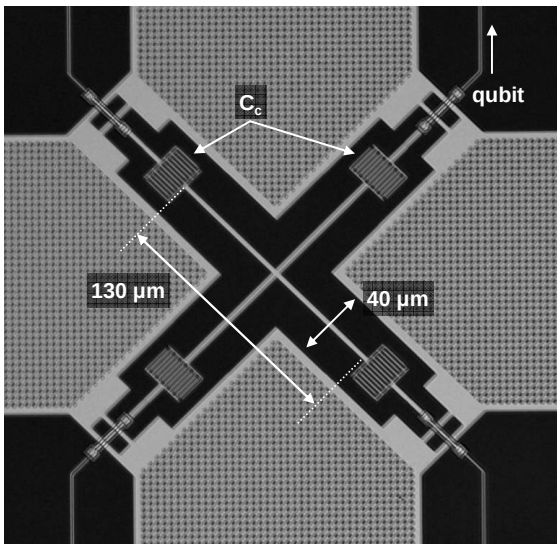

Figure 6.5: Physical geometry of four-qubit coupling circuit.

the loading factor into account the coupling strengths would differ markedly from the expected values. We will see in Chapter 8 that the experimentally observed splitting agrees reasonably well with the theoretical value in a sample designed for $S = 15\,\mathrm{MHz}$. A slightly lower swap frequency of $12.5\,\mathrm{MHz}$ is observed, possibly indicating a slight underestimate of the island ground capacitance.

An important assumption made in our analysis of the coupler up to this point is that the entire central island can be described as having a single voltage, which is valid insofar as the island can be considered a lumped element. To check this assumption, we must consider the frequencies of resonant modes of the island itself; if these are far above the operating frequency of the qubits, then these modes will not be relevant in the operating regime we care about, the lumped-

S (MHz)	C_Δ (fF)	$C_\star$ (fF)
10	1.67	11.2
15	2.50	15.0
20	3.33	18.7
25	4.17	22.3
30	5.00	25.8
40	6.67	32.8
50	8.33	39.6

Table 6.1: Coupling capacitances for 4-qubit star coupler.

element assumption will hold, and the straightforward circuit analysis given above will apply. If we consider one arm of the coupler as a coplanar waveguide of length $130\,\mu$m, then the resonance frequency is found to be $\sim 1\,$THz, far above the qubit operating frequency. Even if there are modes with effective length 20 times longer, for example a slotline mode that wraps around the entire coupling island, this still puts the frequency at 50 GHz. For this reason we feel confident that the simple lumped-element model captures the behavior of the coupling circuit.

Chapter 7

Multi-qubit Measurement

One of the most important requirements for a quantum computer is the ability to extract information from it by measuring the quantum state. In any experimental system such measurements will almost certainly be subject to errors or nonidealities. These errors must be understood and corrected in order to perform reliable quantum information processing. In this chapter we will discuss the errors and nonidealities in our phase qubit measurements and how the raw measurement results are corrected to account for these errors. We will also describe a new measurement scheme that avoids a major potential problem, that of measurement crosstalk. It is this scheme that will be used for all our experimental results.

In Chapter 2 we described the measurement process for our phase qubits, and in particular we separated this process into two stages: the "measurement" stage

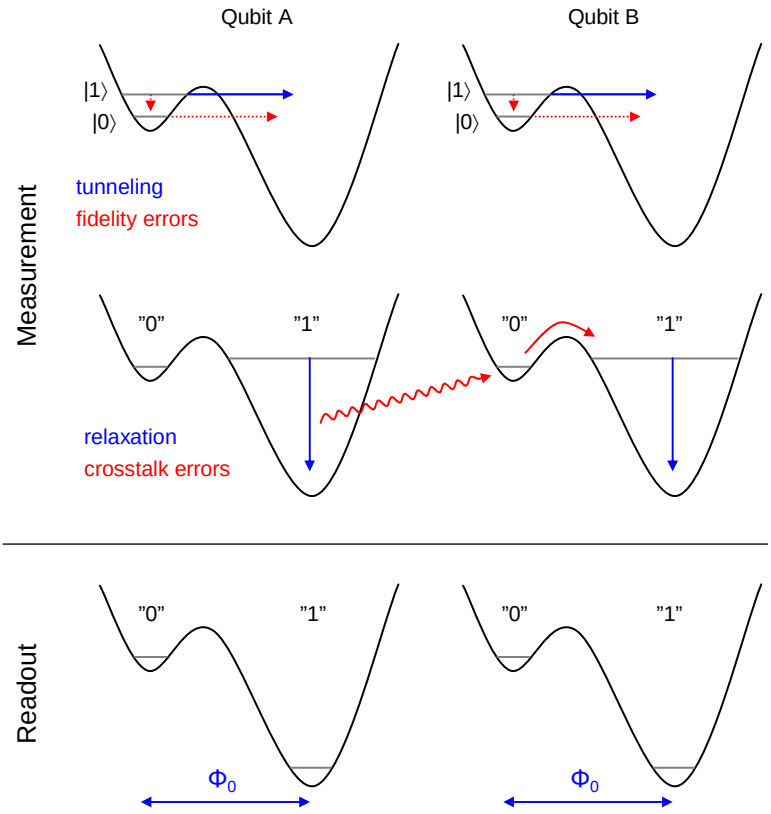

Figure 7.1: Overview of phase qubit measurement and errors.

where the qubit tunnels (or not) and then relaxes to the ground state of either the left or right-hand well, and the "readout" stage in which the on-chip SQUID is used to determine which well the circuit has relaxed into. The entire measurement process is illustrated schematically in Figure 7.1.

For our purposes the readout stage will be assumed to be perfectly ideal; in typical experiments the switching distributions corresponding to the two wells are separated from each other by several standard deviations, so there is no ambiguity

in identifying which well the system is in. Errors can occur during the "measurement" stage however, and it is these errors that will concern us here. It is useful to further divide this stage into two sections, first when the qubit tunnels (or not) out of its well, and second when it relaxes. In the first section of this process, we would like a qubit in state $|0\rangle$ to remain in the qubit well with unit probability, while a qubit in state $|1\rangle$ should tunnel out with unit probability. In practice of course, neither of these probabilities is equal to unity, and such errors give rise to finite measurement fidelity, as shown in the top row in Figure 7.1.

In the second section of the measurement stage, a qubit that has tunnelled out into the right-hand well relaxes into the ground state, emitting hundreds of photons of energy into the surrounding circuit. This is only problematic if other qubits are present in the circuit, as the surge of energy can cause other qubits to be excited and possibly tunnel erroneously instead of staying in the ground state of their qubit well, as shown in the middle row of Figure 7.1. Such errors are known as measurement crosstalk[23].

In our standard treatment of measurement errors[5, supplement], we relate the actual or *intrinsic* probabilities for the various possible qubits states to the *measured* probabilities $\mathbf{P}_m = (P_{m;000}, P_{m;001}, \ldots)^T$ by a matrix equation

$$\mathbf{P}_m = XF\mathbf{P}_i \tag{7.1}$$

where the matrix F describes the fidelity errors, and the matrix X describes

the crosstalk errors. Note the ordering of these two matrices, which implies that fidelity errors happen first followed by crosstalk errors, in accordance with our definition of these two sections of the measurement stage. By inverting this relation, we can extract the intrinsic probabilities $\mathbf{P}_i$ from the measured probabilities $\mathbf{P}_m$.

7.1 Measurement Fidelity

Let us write f_0 to denote the probability that a qubit in its ground state stays in the qubit well (*does not* tunnel), and f_1 for the probability that a qubit in the excited state leaves the qubit well (*does* tunnel). These describe the measurement *fidelity*, and should be close to one. It will sometimes be convenient to think in terms of errors rather than fidelity, so we define $e_0 \equiv 1 - f_0$ and $e_1 \equiv 1 - f_1$ to be the respective probabilities that $|0\rangle$ erroneously tunnels or that $|1\rangle$ erroneously relaxes to the ground state. Given these definitions, we can easily derive the fidelity matrix relating the intrinsic and measured probabilities for one qubit:

$$\mathbf{P}_m = F_1 \cdot \mathbf{P}_i = \begin{pmatrix} f_0 & e_1 \\ e_0 & f_1 \end{pmatrix} \cdot \mathbf{P}_i. \tag{7.2}$$

This fidelity matrix is invertible provided $f_0 + f_1 - 1 = 1 - e_0 - e_1 \neq 0$. Typically, we have measurement errors $e_i \approx 5\%$, so that there is no trouble inverting this relation and correcting for finite measurement fidelity.

Fidelity errors arise from processes local to each qubit, for example relaxation during the measurement pulse due to crossing a TLS[26]. Hence in a multiqubit system these errors are independent for each qubit. This means that we can compute the fidelity matrix for multiple qubits simply by taking a tensor product of the single-qubit fidelity matrices. With three qubits we have

$$F_3 = \begin{pmatrix} f_0 & e_1 \\ e_0 & f_1 \end{pmatrix}_A \otimes \begin{pmatrix} f_0 & e_1 \\ e_0 & f_1 \end{pmatrix}_B \otimes \begin{pmatrix} f_0 & e_1 \\ e_0 & f_1 \end{pmatrix}_C \tag{7.3}$$

where the subscripts A, B and C refer to the different qubits. As with one qubit, this fidelity matrix can be inverted to calculate the intrinsic occupation probabilities $\mathbf{P}_i$ from the measured probabilities $\mathbf{P}_m$.

Determining the fidelities f_0 and f_1 for each individual qubit can be done with a simple "s-curve" experiment: first the qubit is reset to the ground state and we measure the tunnelling probability to find $e_0 = 1 - f_0$. Next, we apply a π-pulse to excite the qubit into state $|1\rangle$ and then measure the tunnelling to find f_1. The measurement of f_1 is less certain because it assumes ideal preparation of the qubit in state $|1\rangle$, however other more detailed experiments can be done to determine the actual fidelity of the $|1\rangle$-state preparation[18] and then appropriately correct the measurement fidelity f_1.

7.2 Measurement Crosstalk

Because measurement crosstalk is fundamentally a multi-qubit effect, the simplest case to consider is for two qubits, as describe by Bialczak[5]. The X matrix for two qubits contains two parameters to describe the crosstalk. These are the probability that when qubit A switches it will cause B to switch (x_{AB}) and the probability that switching on B will cause A to switch (x_{BA}). For typical coupling strengths with direct capacitive coupling, these crosstalk numbers are $\sim 10\%$. The crosstalk matrix is then

$$X_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - x_{BA} & 0 & 0 \\ 0 & 0 & 1 - x_{AB} & 0 \\ 0 & x_{BA} & x_{AB} & 1 \end{pmatrix}. \tag{7.4}$$

Notice that this matrix is lower triangular. This is because the basis states are written going down each column in binary order $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, and because crosstalk can only turn zeros into ones, causing a transition to a lower row. This matrix is completely determined by the two parameters x_{AB} and x_{BA}; we need only determine these two parameters in order to correct for measurement crosstalk. In the matrix equation $\mathbf{P}_m = XF\mathbf{P}_i$ there are three equations (4 - 1 because of normalization). Hence, if we know $\mathbf{P}_m$ (from measurement), F (from s-curves) and $\mathbf{P}_i$ (by initializing to a known state) we can hope to determine X.

The crosstalk matrix does not have a tensor product structure, so it is not immediately clear how to generalize it for N qubits; instead the N-qubit crosstalk

matrix must be constructed manually. First let us consider the structure of the matrix, in particular how many degrees of freedom it can have. Just as in the two-qubit case, the crosstalk matrix will be lower triangular if the kets are listed in the usual binary order because crosstalk can only flip bits on, causing error terms to appear in later rows of the matrix. Because crosstalk can only flip bits on, the error terms will be in positions $X_{i[N-1..0],j[N-1..0]}$ where the binary indices $i[N-1..0]$ and $j[N-1..0]$ are such that all the bits that are on in j must be on in i.[1]

These facts allow us to count the number of degrees of freedom in the general crosstalk matrix as follows: there are $\binom{N}{k}$ different states with k bits turned on; for each such state, crosstalk can flip any of the $N-k$ remaining zero bits, giving 2^{N-k} potential crosstalk terms, but we must subtract one degree of freedom because the crosstalk probabilities for each state must sum to one. In addition, if $k=0$ then no qubits switch so crosstalk can not happen, and similarly if $k=N$ then all qubits switch so there are no more bits for crosstalk to turn on. Hence, the total number of degrees of freedom is $\sum_{k=1}^{N-1}\binom{N}{k}(2^{N-k}-1)$. This sum can be easily evaluated by massaging the limits and noting that $(x+1)^N=\sum_{k=0}^{N}\binom{N}{k}x^{N-k}$. In the end, we find that the number of degrees of freedom is $3^N-2^{N+1}+1$. Plugging

[1] Here one has to be careful about the convention chosen for the crosstalk probabilities. Using the standard convention from the two-qubit case[5], x_{AB} gives the probability that A causes B to switch $A \to B$, but in the X matrix with the usual indexing order the element $X_{i,j}$ denotes crosstalk taking state j to state i, $i \leftarrow j$.

in $N = 2$ we indeed find 2 degrees of freedom, as expected, while for $N = 3$ the general crosstalk matrix is found to have 12 degrees of freedom.

The large number of potential degrees of freedom in a multi-qubit crosstalk matrix is a problem for experimentally determining the crosstalk matrix in order to correct for crosstalk errors. With N qubits there are only $N(N-1)$ terms of the form x_{ij} that describe the crosstalk probability of qubit j affecting qubit i, as in the two-qubit case. But the total number of degrees of freedom in the crosstalk matrix is $3^N - 2^{N+1} + 1 \gg N(N-1)$. The general structure of the crosstalk may be much more complicated than what is found for two qubits. The basic identity $\mathbf{P}_m = XF\mathbf{P}_i$ contains only $2^N - 1$ independent equations, so that it may in general be very hard to determine X by the methods used for two qubits.

To model the multi-qubit crosstalk matrix then, we must make some assumptions about the physics of what actually happens to cause measurement crosstalk, in particular how crosstalk effects can propagate between multiple qubits. We will briefly consider two possible models for crosstalk effects which we'll refer to as "Independent Choice" and "Cascade". Of course, a better solution which bypasses all this complexity is to use a measurement scheme that avoids crosstalk altogether, as outlined in §7.3. For this reason the following two sections should be considered optional.

7.2.1 Crosstalk model: Independent Choice

Suppose that at the moment of measurement, each qubit that tunnels decides to deposit or not to deposit a certain amount of energy in each of the qubits to which it is coupled, and this energy will cause the other qubit to switch with some small probability. Each qubit decides independently of the others, and the qubits that get switched by crosstalk effects do not themselves cause other qubits to switch (the crosstalk effects do not "cascade"). For three qubits, we can model this with $3(3-1) = 6$ parameters x_{AB}, x_{AC}, x_{BA}, x_{BC}, x_{CA}, x_{CB}, analogous to the two-qubit case. For convenience we also define $n_{AB} \equiv 1 - x_{AB}$ to be the probabilities that no crosstalk occurs between the specified qubits.

Consider what happens starting when the system starts in state $|001\rangle$ at the time of measurement. When qubit C switches, it decides with probability x_{CA} to switch qubit A and with probability x_{CB} to switch qubit B. Since these choices are made independently, both qubits will switch with probability $x_{CA}x_{CB}$ and result in the system being measured as $|111\rangle$. The probability that only qubit B switches and we measure $|011\rangle$ is $n_{CA}x_{CB}$ and similarly the probability that only qubit A switches and we measure $|101\rangle$ is $x_{CA}n_{CB}$. Finally, the probability that neither qubit switches and the system is correctly measured as $|001\rangle$ is $n_{CA}n_{CB}$. Similar considerations apply for the starting states $|010\rangle$ and $|100\rangle$.

Next, consider what happens starting from the state $|011\rangle$. In this case, the

only possibilities for measurement crosstalk are that nothing happens, or that qubit A switches resulting in state $|111\rangle$. In order for nothing to happen, qubits B and C must both independently decide not to crosstalk, so that the system ends up in $|011\rangle$ with probability $n_{BA}n_{CA}$. Otherwise qubit A will switch, so that the system is measured as $|111\rangle$ with probability $1 - n_{BA}n_{CA} = x_{BA} + x_{CA} - x_{BA}x_{CA}$. Similar considerations apply for the starting states $|101\rangle$ and $|110\rangle$.

Putting all this together, we find the following crosstalk matrix for the Independent Choice model:

$$X_{3I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & n_{CA}n_{CB} & 0 & 0 \\ 0 & 0 & n_{BA}n_{BC} & 0 \\ 0 & n_{CA}x_{CB} & n_{BA}x_{BC} & n_{BA}n_{CA} \\ 0 & 0 & 0 & 0 \\ 0 & x_{CA}n_{CB} & 0 & 0 \\ 0 & 0 & x_{BA}n_{BC} & 0 \\ 0 & x_{CA}x_{CB} & x_{BA}x_{BC} & 1 - n_{BA}n_{CA} \end{pmatrix} \cdots$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ n_{AB}n_{AC} & 0 & 0 & 0 \\ n_{AB}x_{AC} & n_{AB}n_{CB} & 0 & 0 \\ x_{AB}n_{AC} & 0 & n_{AC}n_{BC} & 0 \\ x_{AB}x_{AC} & 1 - n_{AB}n_{CB} & 1 - n_{AC}n_{BC} & 1 \end{pmatrix}. \quad (7.5)$$

Bialczak[5] outlines a technique for determining X in the two-qubit case by combining various measured probabilities to trace out one or the other of the qubits. We define

$$\begin{aligned} P_{1A} &\equiv P_{10} + P_{11} \\ P_{1B} &\equiv P_{01} + P_{11}, \end{aligned}$$

and then measure these partial probabilities for various initial states, e.g. $|00\rangle$ and $|01\rangle$. This leads to relations between the partial probabilities for various initial states, the measurement fidelity f_0 and measurement error $e_0 = 1 - f_0$ as defined in §7.1, and the crosstalk probabilities x_{AB}. One such relation is

$$\frac{P_{1A}(01) - P_{1A}(00)}{P_{1B}(01) - P_{1B}(00)} = \frac{f_{0A} x_{BA}}{1 - e_{0A} x_{AB}} \approx f_{0A} x_{BA} \tag{7.6}$$

where we drop the denominator since e_{0A} and x_{AB} are both small. In Equation 7.6 we measure the quantities on the left-hand side and the fidelity f_{0A} of qubit A, and from these can then determine the crosstalk probability x_{AB}. This allows *all* crosstalk parameters to be determined from the single-qubit fidelities and simple qubit measurements.

A similar calculation with the independent choice model for the three-qubit crosstalk matrix shows that we can use the same approach to determine the crosstalk probabilities in this case. For example, we calculate the following relations:

$$\begin{aligned}
\frac{P_{1A}(010) - P_{1A}(000)}{P_{1B}(010) - P_{1B}(000)} &= \frac{f_{0A}x_{BA}}{1 - e_{0A}x_{AB}} \frac{1 - e_{0C}x_{CA}}{1 - e_{0C}x_{CB}} \\
&\approx f_{0A}x_{BA} \qquad (7.7) \\
\frac{P_{1A}(001) - P_{1A}(000)}{P_{1C}(001) - P_{1C}(000)} &= \frac{f_{0A}x_{CA}}{1 - e_{0A}x_{AC}} \frac{1 - e_{0B}x_{BA}}{1 - e_{0B}x_{BC}} \\
&\approx f_{0A}x_{CA}, \qquad (7.8)
\end{aligned}$$

and so on for partial probabilities of the other qubits with appropriate combinations of initial states. From these relations one could work out the 6 crosstalk parameters in this model in a way very similar to what can be done with two qubits. Thus we could hope to measure the crosstalk and correct for it. However, this assumes that the Independent Choice model correctly describes the physics of multi-qubit measurement crosstalk. Not all crosstalk models share these properties, as we will see in the next section.

7.2.2 Crosstalk model: Cascade

In the previous section, we stipulated that the qubits which are switched by crosstalk cannot in turn decide to switch other qubits. Here we relax that restriction and allow the crosstalk effects to "cascade" from one qubit to the next. We will again assume that at each stage the 6 crosstalk parameters (x_{AB}, x_{AC} and so on) determine the probability of one qubit switching another, but several stages of crosstalk can happen in succession.

Consider what happens when a system in state $|001\rangle$ is measured. When qubit C switches, it decides with probability x_{CA} to switch qubit A and with probability x_{CB} to switch qubit B. Suppose qubit B switches but A does not, taking the system to state $|011\rangle$ with probability $n_{CA}x_{CB}$. Now, qubit B can in turn decide whether or not to switch qubit A; if it does not switch, the system stays in state $|011\rangle$ with probability $n_{CA}x_{CB}n_{BA}$, otherwise the system switches to $|111\rangle$ with probability $n_{CA}x_{CB}x_{BA}$. Similar considerations apply if crosstalk first switches A but not B, after which A in turn can decide whether or not to switch B. Clearly this model, while based on the same 6 crosstalk parameters as before, will produce a crosstalk matrix with a different structure. Working this out in detail we find:

$$X_{3C} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & & & \\ 0 & n_{CA}n_{CB} & 0 & 0 & & & & \\ 0 & 0 & n_{BA}n_{BC} & 0 & & & & \\ 0 & n_{CA}x_{CB}n_{BA} & n_{BA}x_{BC}n_{CA} & n_{BA}n_{CA} & & & & \\ 0 & 0 & 0 & 0 & & & & \\ 0 & x_{CA}n_{AB}n_{CB} & 0 & 0 & & & & \\ 0 & 0 & x_{BA}n_{AC}n_{BC} & 0 & & & & \\ 0 & 1-\Sigma & 1-\Sigma & 1-n_{BA}n_{CA} & & & & \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ n_{AB}n_{AC} & 0 & 0 & 0 \\ n_{AB}x_{AC}n_{CB} & n_{AB}n_{CB} & 0 & 0 \\ x_{AB}n_{BC}n_{AC} & 0 & n_{AC}n_{BC} & 0 \\ 1-\Sigma & 1-n_{AB}n_{CB} & 1-n_{AC}n_{BC} & 1 \end{pmatrix} \tag{7.9}$$

where the $1-\Sigma$ entries are included for brevity and should be replaced by 1 minus the sum of all other entries in the respective column, to preserve normalization. Note that the columns with two qubits in the excited state are not changed in this model, because after the third qubit decides whether or not to switch the cascade has no effect.

If we attempt to calculate the partial probabilities (e.g. P_{1A}) and their ratios with this model, as in the previous section, the results are complicated and do not show the same simple structure as in the Independent Choice model or the case of two-qubits.

What should be clear at this point is that multi-qubit crosstalk is potentially very complicated and difficult to measure and correct for. If we allow for full generality in the crosstalk matrix, then the number of degrees of freedom is very large and the matrix will be hard to measure. If on the other hand we attempt to reduce the number of degrees of freedom in the crosstalk matrix, then we must rely on a physical model of how crosstalk works in order to produce the appropriate matrix structure. It is possible that accurate measurements of crosstalk (using for example the partial probabilities defined in the previous section) could yield insight into the actual physics. For our purposes, however, we would like to avoid these complexities. Hence we turn next to a new measurement scheme that avoids crosstalk altogether.

7.3 Crosstalk-Free Measurement

As we have seen in the previous section, measurement crosstalk between multiple capacitively-coupled qubits is complicated and hence difficult to measure and correct for. If we look at the form of the various crosstalk matrices in Equations 7.4, 7.5 and 7.9, we notice that all of them share the same first row and first column, a single 1 in the upper left corner with all other entries equal to 0. This reflects two facts about measurement crosstalk: first, if no qubits switch then crosstalk has no effect, so $|000\rangle$ is always measured as $|000\rangle$; second, crosstalk only acts to switch additional qubits, not prevent them from switching, so it can not cause other states to be erroneously measured as $|000\rangle$[2].

This means that the "null-result" probability P_{000} of having no excited qubits is unaffected by crosstalk. The standard measurement scheme is illustrated in Figure 7.2, showing that most of the numbers we extract from a measurement are subject to crosstalk errors. If we can devise an alternate means of measuring the quantum state of the system using only null results, then we can avoid crosstalk errors and hence the need to characterize and correct for them. There are various ways that this can be accomplished, and we will describe two such techniques in this section.

One way to determine the probabilities P_{000}, P_{001}, *etc.* of the various qubit

[2]This is of course true for any number of qubits, but here we consider only the relevant three-qubit case.

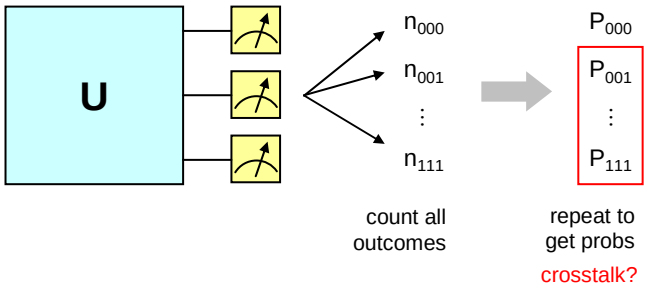

Figure 7.2: Standard measurement, subject to crosstalk.

states using only null-result measurements is illustrated in Figure 7.3. We perform a set of 2^N separate experiments, one for each state, where in each experiment we apply a set of π pulses to map the desired state to the ground state before measurement. We then measure and record the null-result probability, which now corresponds to the probability that the system was originally in the state that was mapped to ground. For example, to measure P_{001} we would apply a π-pulse to qubit C so that $|001\rangle \rightarrow |000\rangle$, as illustrated.

Clearly this measurement technique is inefficient, requiring an exponentially large number of measurements to determine all the state occupation probabilities. However, for three qubits the exponential overhead is still only a manageable factor of 8, so we will not worry about this question of efficiency[3]. More troubling is the fact that this protocol introduces extra pulses into the sequence of qubit operations. These pulses have imperfections so that the mapping to the ground

[3]A quantum state of N qubits contains an exponentially-large amount of information, so *any* protocol that tries to extract it all will run into problems of efficiency.

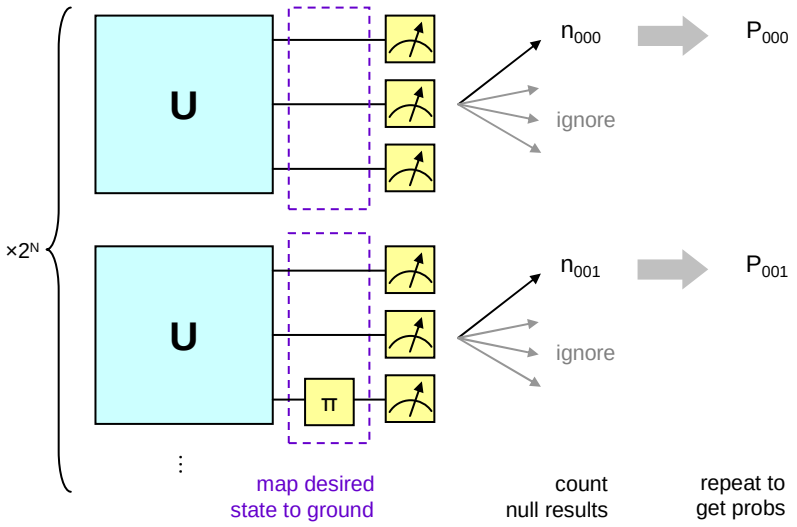

Figure 7.3: Crosstalk-free Measurement by mapping to the ground state.

state is not perfect, and in addition the qubit will be subject to decoherence during the added time providing another potential source of error. If our goal is to understand measurement errors and correct them, then adding extra pulses makes the job more difficult.

A second crosstalk-free measurement scheme is shown in Figure 7.4. To understand this protocol, consider what happens if we only measure qubits A and B, but not qubit C. We obtain four probabilities which we write as $P_{00x}, P_{01x}, P_{10x}, P_{11x}$, where the subscript x indicates that we have no information about the unmeasured qubit C. As before, only the null-result probability P_{00x} is unaffected by crosstalk, but note that $P_{00x} = P_{000} + P_{001}$ because there are two possibilities for the state of the unmeasured qubit C. The two probabilities P_{00x} and P_{000} are

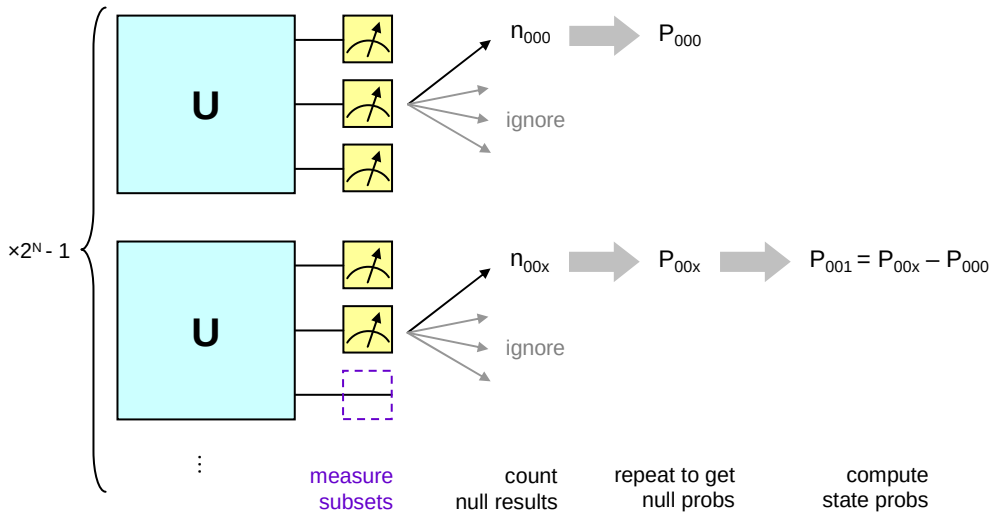

Figure 7.4: Crosstalk-free Measurement using subset measurements.

both null results and so can be measured without crosstalk as we have described; from these the third probability P_{001} can be determined.

Continuing in this manner, we can reconstruct the complete set of occupation probabilities without crosstalk by repeating the experiment $2^N - 1$ times, each time measuring only a certain subset of the qubits and recording only the null-result probability for that subset (for the degenerate case in which no qubits are measured, we have $P_{xxx} = 1$). This gives the following set of measured null-results $\mathbf{P}_{\text{null}} = (P_{000}, P_{00x}, P_{0x0}, P_{0xx}, P_{x00}, P_{x0x}, P_{xx0}, P_{xxx})^T$, which is related to the set of occupation probabilities $\mathbf{P} = (P_{000}, P_{001}, P_{010}, P_{011}, P_{100}, P_{101}, P_{110}, P_{111})^T$ according to

$$\mathbf{P}_{\text{null}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \cdot \mathbf{P}$$
$$= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \mathbf{P}. \tag{7.10}$$

By inverting this equation, we thus obtain the occupation probabilities in a way that is completely insensitive to measurement crosstalk.

These relations can be modified in a straightforward way to account for measurement fidelity as well. Recall that the single-qubit fidelity matrix in Equation 7.2 related the measured probabilities to the intrinsic probabilities for one qubit. With null-result measurement, the ground-state probability P_0 is related to the intrinsic probabilities in the same way, but we do not measure the excited state probability, instead using $P_x = 1$ which is just the sum of the intrinsic probabilities. Hence the appropriate "fidelity matrix" for one qubit with null-result measurement is

$$\mathbf{P}_m = \begin{pmatrix} f_0 & e_1 \\ 1 & 1 \end{pmatrix} \cdot \mathbf{P}_i. \tag{7.11}$$

Using this result with three qubits, Equation 7.10 becomes

$$\mathbf{P}_{\text{null}} = \begin{pmatrix} f_0 & e_1 \\ 1 & 1 \end{pmatrix}_A \otimes \begin{pmatrix} f_0 & e_1 \\ 1 & 1 \end{pmatrix}_B \otimes \begin{pmatrix} f_0 & e_1 \\ 1 & 1 \end{pmatrix}_C \cdot \mathbf{P}. \tag{7.12}$$

This reduces to the form given in Equation 7.10 in the case of perfect fidelity when $f_0 = 1$ and $e_1 = 0$. By inverting this equation we obtain an expression for the occupation probabilities $\mathbf{P}$ in terms of the measured null-result probabilities $\mathbf{P}_{\text{null}}$ with single-qubit measurement fidelity taken into account, and no effect of measurement crosstalk. This is the measurement protocol used in all of our experimental results.

One subtlety in this measurement process is that because each element of $\mathbf{P}_{\text{null}}$ is measured in separate repetitions of the experiment, each is subject to independent statistical noise. As a result, when Equation 7.12 is inverted the elements of $\mathbf{P}$ may not satisfy the requirements for a set of probabilities that they be non-negative and sum to unity. We can use maximum likelihood estimation (MLE) to enforce these constraints and find the probabilities $\mathbf{P}$ which give the measured results $\mathbf{P}_{\text{null}}$ with the highest probability. This maximum likelihood technique is explained in §A.3.1.

Chapter 8

Experimental Results

8.1 Bringup and Calibration

A completed four-qubit device of the type used in our experiments is shown in Figure 8.1. The first step in operating this device is to characterize the individual qubits and then the multi-qubit performance. With the sample we tested, the three qubits A, B and C were working, but no switching signal could be observed from the readout SQUID on qubit D (it appeared to be a short). Thus we were unable to perform any four-qubit measurements on this device. However, the remaining qubits worked well, and in fact were remarkably similar to each other, making it possible to operate all of them and couple them together as desired.

The phase qubits in this device were designed with standard component values,

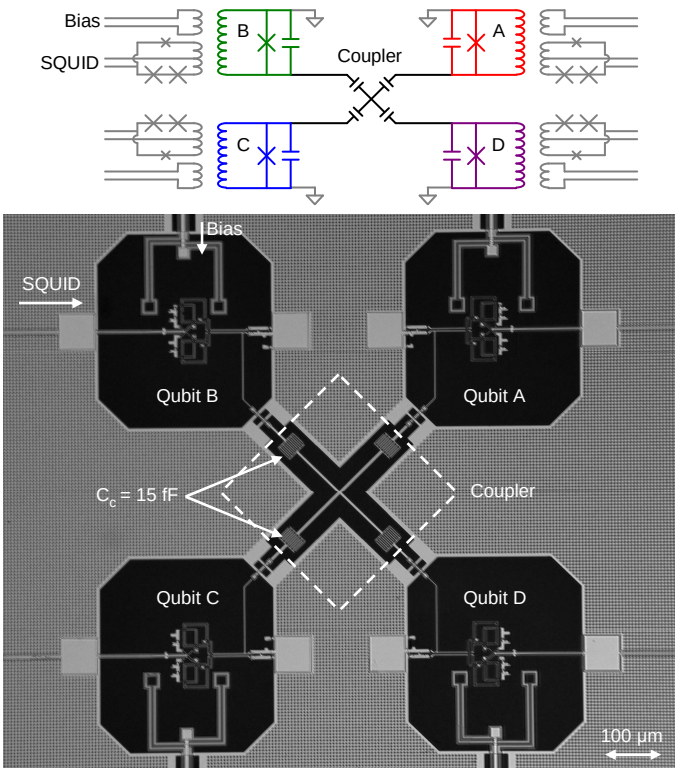

Figure 8.1: Schematic and micrograph of four-qubit device.

as described in Chapter 2: critical current $I_0 \approx 2\ \mu$A, capacitance $C \approx 1$ pF, and inductance $L \approx 720$ pH. The coupling capacitance was $C_c \approx 15$ pF, chosen to give coupling strength $2g/2\pi \approx 15$ MHz at a qubit frequency of 6 GHz, as discussed in Chapter 6. Each qubit has a separate bias control line, and each qubit has a separate measurement SQUID for joint readout.

Because every phase qubit is unique, the first step in operating any qubit device is to calibrate the various qubit parameters for proper reset, operation and readout. This process has been described in detail in the thesis by Ansmann[2].

With our four qubit device, the calibration procedure is essentially the same as for single or coupled qubits samples. The main difference is that even during single-qubit bringup operations, we must operate (though not necessarily measure) all qubits because of small but non-negligible bias crosstalk. If we calibrate one qubit while the others are left unbiased, then later when the other qubits are biased to their operating points, the first qubit will shift and the calibrations will be incorrect. The multi-qubit bringup is thus an iterative process, where we choose a rough operating point for all qubits (based on SQUID steps), then whenever one qubit operating bias is changed, the others are recalibrated and this process repeated until the calibration stabilizes.

In practice this iterative process stabilizes rather quickly because the flux bias crosstalk between the qubits is relatively small. This allows us to change the operating bias on one qubit to choose a new operating frequency, then quickly either recalibrate the other qubits or change their flux bias to compensate and bring them back to the original operating point. We chose to operate the qubits detuned from each other by $\sim \pm 250\,\mathrm{MHz}$ as indicated in Table 8.1, so that the coupling interaction is off since the detuning Δ is much larger than the coupling strength ($\Delta/g \approx 20$). Having chosen operating points for all qubits, we can then characterize the performance of each individual qubit. The relaxation and spin-echo dephasing times found from basic single-qubit experiments are likewise shown

qubit	T_1 (ns)	T_{echo} (ns)	f_{10} (GHz)
A	460	270	6.2995
B	460	300	6.5506
C	450	390	6.7988

Table 8.1: Single-qubit parameters.

in the table.

Turning the coupling interaction on requires tuning the qubits into resonance with each other for some time. The required detuning pulses are calibrated pairwise between the qubits; for example, to calibrate the detuning pulse to couple A with B, we first excite qubit B with a π-pulse and then adjust the amplitude and length of the detuning pulse on A to maximize the transfer of this excitation to A, as shown in Figure 8.2. The time and amplitude for optimum transfer give a pulse that implements an iSWAP gate, as needed for the $|\text{GHZ}\rangle$ protocol. The process is repeated for each pair of qubits, giving swap times and coupling strengths as shown in Table 8.2. The coupling strengths are within 5% of each other and also quite close to the design value of 15 MHz. The disagreement is probably an indication of a slight underestimate of the island ground capacitance, as was discussed in Chapter 6.

For the single-step $|\text{W}\rangle$ and $|\text{GHZ}\rangle$ protocols we must bring all three qubits into resonance simultaneously, and the protocols assume that all coupling strengths are equal. For the $|\text{W}\rangle$ protocol, we make an approximation by taking the interac-

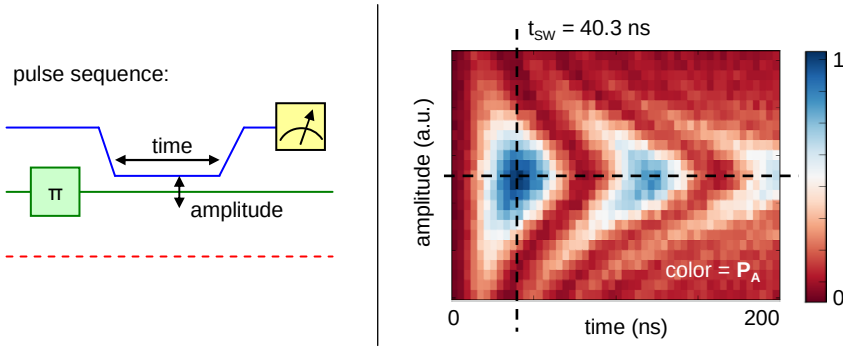

Figure 8.2: Multi-qubit coupling calibration.

qubits	t_{iSWAP} (ns)	$2g/2\pi$ (MHz)
AB	40.3	12.4
AC	40.9	12.2
BC	38.8	12.9

Table 8.2: Qubit-qubit coupling parameters.

tion time to be $t_{\mathrm{W}} = (4/9)\bar{t}_{\mathrm{iSWAP}} \approx 18\,\mathrm{ns}$, where $\bar{t}_{\mathrm{iSWAP}} = 39.8\,\mathrm{ns}$ is the average of the coupling times between the various pairs of qubits, which turns out to work reasonably well. For the single-step $|\mathrm{GHZ}\rangle$ protocol, Galiautdinov has considered techniques for modifying the protocol to account for inhomogeneities in the coupling[11], however this protocol suffered from other problems that make it difficult to implement given the small nonlinearity of phase qubits, as discussed below in §8.4. In the end, the single-step $|\mathrm{GHZ}\rangle$ protocol was abandoned, and only the single-step $|\mathrm{W}\rangle$ and iSWAP $|\mathrm{GHZ}\rangle$ protocols were implemented.

Two final multi-qubit calibrations must be performed in order to implement the desired entangling sequences. These are related to the discussion of reference

frames in Chapter 3. First, we must calibrate the starting reference phase of each qubit's microwave reference frame, so that when the qubits are tuned into resonance they share common X and Y axes[2, p. 240]. In addition, we must perform a Ramsey fringe experiment to measure the change in phase experienced by each qubit during the entire detuning pulse, compared to the reference frame. This allows us to correctly set the phases of later microwave pulses in that frame[1]. These multi-qubit phase calibrations are shown in Figure 8.3.

8.2 Time-evolution of the Qubit State

Having calibrated the single-qubit operation and multi-qubit coupling including phases, we are now ready to implement the entangling protocols, which can be translated directly into applied control signals as outlined in Chapter 3. To begin, we consider the time-evolution of the state occupation probabilities during the entangling protocols, measured using the crosstalk-free measurement and shown in Figure 8.4 and Figure 8.5. In each row of these figures, the left panel shows the pulse sequence with time on the horizontal axis and qubit frequency on the vertical axis, while the right panel shows the measured state occupation probabilities as a function of time during the sequence.

[1] An alternative would be to add a correcting Z-pulse to compensate for this phase change so that later microwave pulses need not be modified at all, though this increases the sequence length and so was not done here.

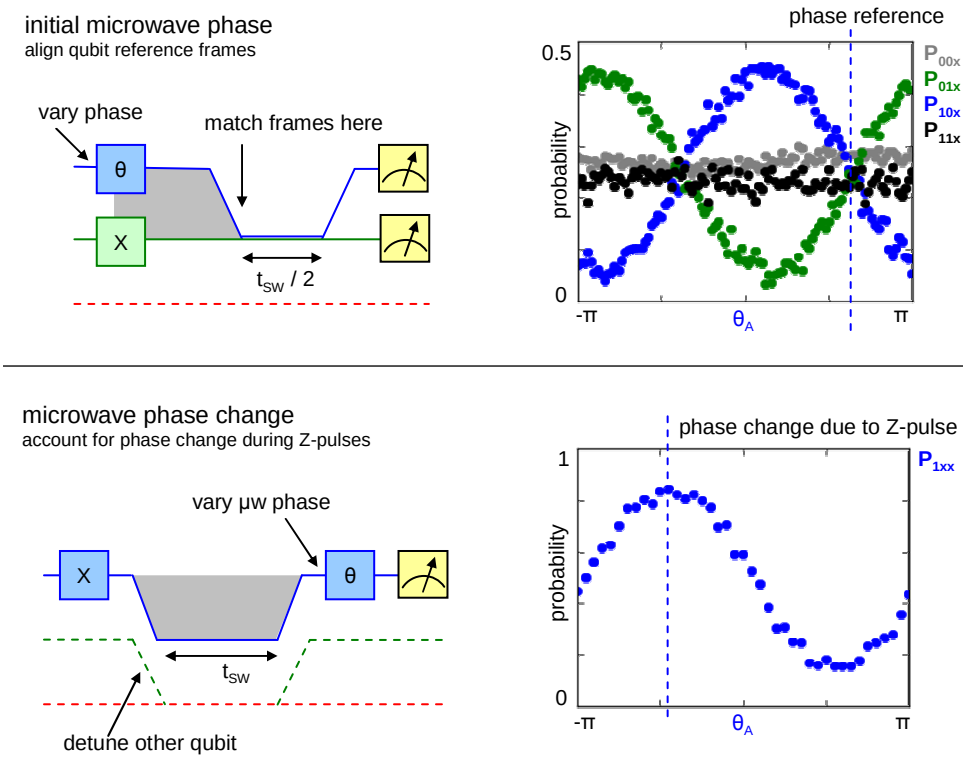

Figure 8.3: Multi-qubit phase calibrations.

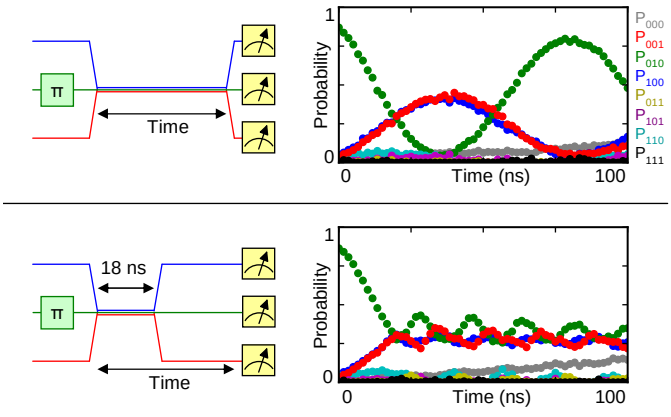

Figure 8.4: Generation of $|W\rangle$ state in the time domain.

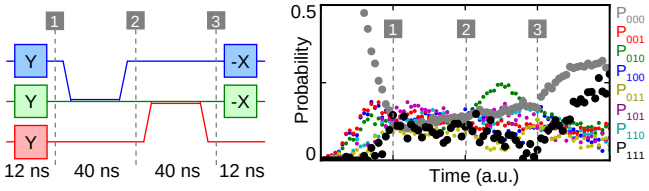

Figure 8.5: Generation of $|\mathrm{GHZ}\rangle$ state in the time domain.

In the $|\mathrm{W}\rangle$ protocol (Figure 8.4, top), one qubit is excited and then the symmetric interaction between all pairs of qubits is used to distribute that excitation among all three. During the interaction, the excitation from qubit B ($|010\rangle$) is swapped to qubits A and C ($|100\rangle$ and $|001\rangle$), then back again. Probabilities P_{100} and P_{001} are nearly equal throughout the entire sequence, indicating that the coupling is nearly symmetric, as desired. When the interaction time is set at 18 ns, the system reaches an equal superposition, and subsequently stays there while the interaction is off (Figure 8.4, bottom). This is a $|\mathrm{W}\rangle$-like state, up to phase rotations due to the detunings. The small residual oscillations visible after the qubits have been detuned are due to the residual coupling from the finite detuning of the qubits; these features are also visible in simulations.

Figure 8.5 shows the state occupation probabilities during the $|\mathrm{GHZ}\rangle$ protocol, plotted in segments corresponding to the stages of the protocol as indicated. The sequence is a direct translation of the circuit shown in Figure 5.5; the iSWAP gates are implemented by tuning the qubits pairwise into resonance for time $t_{\mathrm{iSWAP}} \approx 40$ ns and the single-qubit rotations are implemented by 12 ns-long microwave

pulses, for a total sequence length of 108 ns. The initial rotations create an equal superposition of all qubit states, with all probabilities converging on 1/8 (1). The effect of the two iSWAP gates is then primarily to adjust the phases of the various components of the superpositions (1-2, 2-3), so that in the final rotation constructive interference causes $|000\rangle$ and $|111\rangle$ to be populated, while all other states are depopulated. For an ideal $|\mathrm{GHZ}\rangle$ state, the probabilities P_{000} and P_{111} should approach 50%, though in the experiment this level is reduced due to the effects of decoherence and errors discussed in §8.4.

Note that in the $|\mathrm{W}\rangle$ protocol where one excitation is swapped among the various qubits, the state evolution is clearly visible in the occupation probabilities as they evolve in time. In the $|\mathrm{GHZ}\rangle$ protocol on the other hand, the iSWAP gates do not have much effect on the populations; rather, they modify the phases of the various components of the equal superposition, so that the final rotations bring the system into the desired target state. Most of the state evolution is hidden in the phase information not captured by these probability measurements.

8.3 Tomography of Entangled States

To fully characterize the quantum states created by the entangling protocols, including the phase information, we perform Quantum State Tomography (QST)[15] by applying various combinations of single-qubit rotations before measurement.

The details of this process, including how the density matrix is extracted from the measured data using maximum likelihood estimation (MLE) to satisfy the physicality constraints, are described in Appendix A. The theoretical and experimental density matrices are shown for $|\mathrm{W}\rangle$ in Figure 8.6 and for $|\mathrm{GHZ}\rangle$ in Figure 8.7.

The real part of $\rho_{\mathrm{W}}^{\mathrm{th}} = |\mathrm{W}\rangle\langle\mathrm{W}|$ is shown at the left in Figure 8.6. All nonzero elements are equal to 1/3, and all imaginary parts (not shown) are identically zero. The experimental ρ_{W} real part and imaginary part (middle) and imaginary part (right) compare nicely with the theoretical prediction: the imaginary part is small with $|\mathrm{Im}\,\rho_{\mathrm{W}}| < 0.03$ and the overall fidelity is $F_{\mathrm{W}} = \langle\mathrm{W}|\,\rho_{\mathrm{W}}\,|\mathrm{W}\rangle = 0.78$.

The real part of $\rho_{\mathrm{GHZ}}^{\mathrm{th}} = |\mathrm{GHZ}\rangle\langle\mathrm{GHZ}|$ is shown at the left in Figure 8.7. All nonzero elements are equal to 1/2, and all imaginary parts (not shown) are identically zero. The experimental ρ_{GHZ} real part (middle) and imaginary part (right) again agree nicely with theory, though not quite as well: the imaginary part is somewhat larger $|\mathrm{Im}\,\rho_{\mathrm{GHZ}}| < 0.10$ and the overall fidelity slightly worse at $F_{\mathrm{GHZ}} = \langle\mathrm{GHZ}|\,\rho_{\mathrm{GHZ}}\,|\mathrm{GHZ}\rangle = 0.62$. However, it is clear from the density matrices that the structure of the state is qualitatively correct.

Another way of looking at the state is to consider expectation values of combinations of one-, two- and three-qubit Pauli operators. This is know as the "Pauli set" of the state or, in optics terminology, the generalized Stokes parameters. These are shown for both states $|\mathrm{W}\rangle$ and $|\mathrm{GHZ}\rangle$ in Figure 8.8. The gray

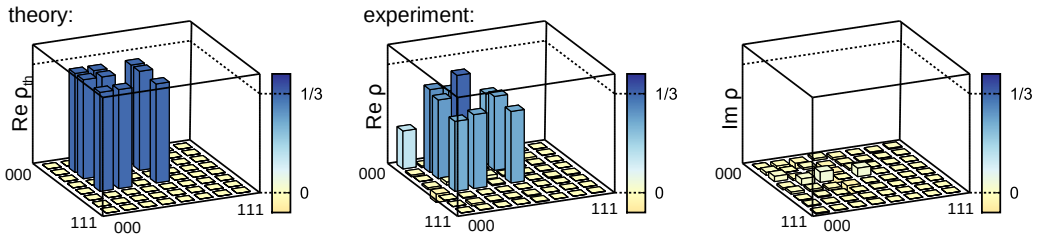

Figure 8.6: Quantum state tomography of $|\mathrm{W}\rangle$.

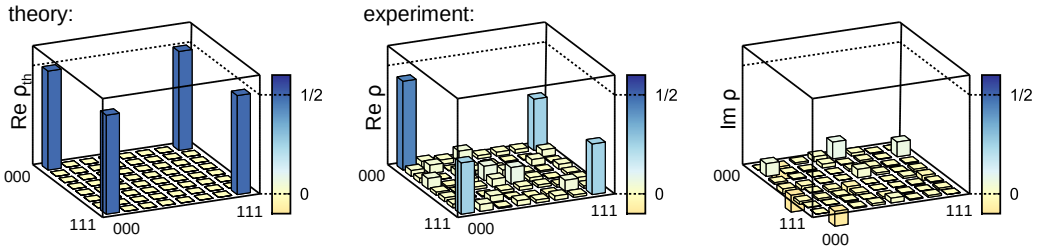

Figure 8.7: Quantum state tomography of $|\mathrm{GHZ}\rangle$.

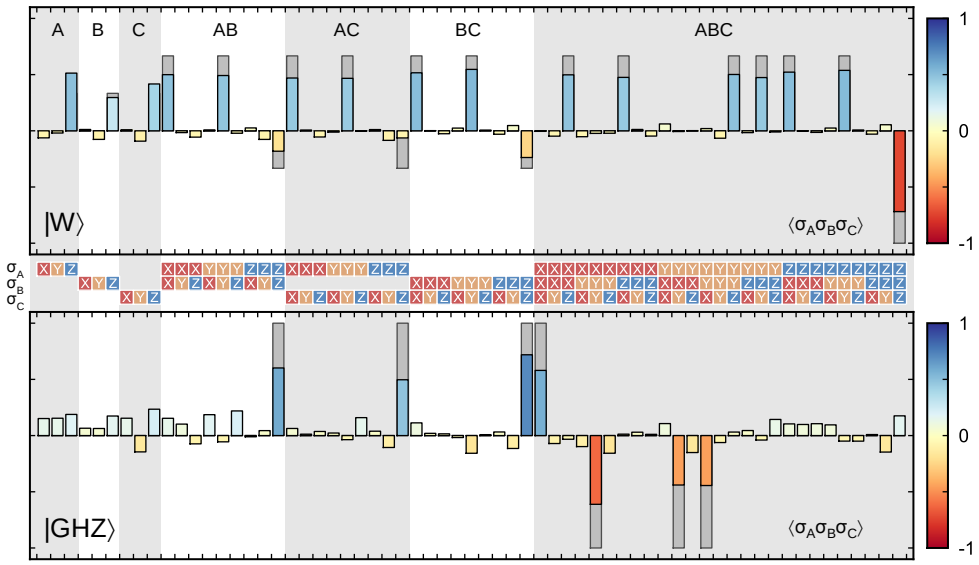

Figure 8.8: Pauli sets for $|\mathrm{W}\rangle$ and $|\mathrm{GHZ}\rangle$ states.

bars show theoretical expectation values, with the experimental results overlaid in color. While the same information is contained in both the density matrix and the Pauli set, looking at the states in this way immediately shows the striking differences between $|\mathrm{W}\rangle$-type and $|\mathrm{GHZ}\rangle$-type entanglement, as discussed above in §4.2. In addition to the three-qubit correlation terms, the $|\mathrm{W}\rangle$-state has two-qubit correlations because tracing out one qubit from a $|\mathrm{W}\rangle$-state still leaves the others partially entangled. For $|\mathrm{GHZ}\rangle$, the two-qubit correlations other than the trivial ZZ-type are absent because tracing out one qubit leaves the others in a completely mixed state.

To understand the significance of the measured fidelities, we compare these results to the entanglement witness operators that detect three-qubit entanglement,

as described in §4.2.2. Three-qubit entanglement is witnessed[1] for the $|\mathrm{W}\rangle$-state provided that $F_{\mathrm{W}} > 2/3$, and for the $|\mathrm{GHZ}\rangle$-state provided that $F_{\mathrm{GHZ}} > 1/2$. Both inequalities are satisfied by the respective measured density matrices, indicating that they are genuine three-qubit entangled states that cannot be decomposed into mixtures of separable states.

While the $|\mathrm{GHZ}\rangle$ fidelity is not high enough to violate the other witness operator $\mathcal{W}_{\mathrm{GHZ}_2}$ which would unambiguously separate it from the class of $|\mathrm{W}\rangle$-like states, the measured state ρ_{GHZ} *is* found to violate the Mermin-Bell inequality given in Equation 4.11. Experimentally we find $G_{\rho_{\mathrm{GHZ}}} = 2.076 \pm 0.029 > 2$; this is far from the ideal value $G_{|\mathrm{GHZ}\rangle} = 4$ for a pure state, but certainly an unambiguous violation, indicating that the state is truly nonclassical and definitely separating it from the $|\mathrm{W}\rangle$ class. The violation is not loophole-free due to use of the crosstalk-free measurement protocol rather than a simultaneous measurement protocol[3], but it is nonetheless an indicator of genuine three-qubit entanglement.

8.4 Sources of Error

The lower fidelity of $|\mathrm{GHZ}\rangle$ compared to $|\mathrm{W}\rangle$ is due to two main factors: first, the $|\mathrm{GHZ}\rangle$ sequence is longer because of the two iSWAP gates; the sequence length of 108 ns is a substantial fraction of the dephasing time T_2 of the qubits, which is particularly harmful because the sequence relies on precise phase adjustment

and interference to populate $|000\rangle$ and $|111\rangle$ while depopulating all other states. Longer coherence times would improve this, as would stronger coupling to reduce the gate time.

Second, the presence of $|2\rangle$ and higher levels and the relatively small nonlinearity of the phase qubit cause errors due to transitions into higher excited states, for example $|110\rangle \rightarrow |200\rangle$. These transitions can be ignored in the $|\mathrm{W}\rangle$ protocol since they are inaccessible with only one excitation in the system, but they cause errors in the $|\mathrm{GHZ}\rangle$ protocol since all qubit states are populated, including those with multiple excitations. These higher levels also make the single-step $|\mathrm{GHZ}\rangle$ protocol very difficult to implement because they perturb the qubit states, leading to phase shifts of the various levels so that the final rotations do not populate $|000\rangle$ and $|111\rangle$ as desired. The effect of higher levels becomes particularly complicated in this experiment when using fixed capacitive coupling with detuning to turn off the interaction, due to spectral crowding from the higher qubit levels.

Spectral crowding greatly constrains the operation of a multiqubit device with fixed coupling. When the detuning between coupled qubits is comparable to the nonlinearity, then unwanted transitions to $|2\rangle$ are possible, resulting in errors. This means detunings must be either kept smaller than the nonlinearity, in which case it is hard to achieve a good coupling on/off ratio, or larger than the nonlinearity, in which case the system will still sweep through avoided-level crossings

with $|2\rangle$ and thus potentially undergo leakage transitions if the sweep is not fast enough. The problem is magnified with more coupled qubits and higher numbers of excitations in the system, as the number of potential leakage transitions quickly becomes unmanageable. This highlights the need to replace frequency detuning with tunable coupling schemes, or to use alternative architectures in the qubits are not all coupled together simultaneously, both of which are currently active areas of research. While the complete graph coupling works well for generating the highly symmetric entangled states we have discussed here, it will not be particularly convenient for a universal quantum computer.

8.5 Conclusion

The experimental data show unambiguously that we have succeeded in creating entangled states of both possible three-qubit varieties. Both states ρ_{W} and ρ_{GHZ} were shown to violate the entanglement witnesses necessary to prove three-qubit entanglement, and ρ_{GHZ} was also shown to violate the Mermin-Bell inequality, which rules out a classical hidden-variable model and hence unambiguously separates this from the $|\mathrm{W}\rangle$ class.

From the standpoint of quantum computation and the future development of superconducting quantum information processing devices, this work strengthens the case for superconducting devices. A multi-qubit device was designed that

took “off-the-shelf” qubit designs and dropped them onto a chip with a simple coupling circuit whose properties could be worked out via straightforward circuit analysis. The device worked as designed, showing the “modularity” that is available with the superconducting circuit approach. As the individual components are improved, they can be easily combined to scale up, toward the eventual goal of a full-fledged quantum computer.

Appendix A

Quantum Tomography

"Tomography", from the Greek *tomos* ("slice") and *graphein* ("to write"), is a method of imaging that involves taking slices and then combining them together to reconstruct a complete image. In quantum mechanics, we know that no single measurement can ever capture all information about the state of a system, because measurement causes the system to collapse into one eigenstate, and hence only reveals a small fraction of the information. By repeating such a measurement on an ensemble of identically-prepared systems we can learn the probabilities for the system to be projected into each eigenstate, but even these probabilities give only a "slice" through the full quantum information contained in the pre-measurement state. A quantum tomographic procedure is a way to reconstruct the complete quantum state by slicing it in different ways before measuring, and then using

the information to deduce the original state. This technique is in principle quite simple, though in practice there are a few subtleties. Here we will describe briefly the procedure we used for quantum state tomography (QST) in the data presented in Chapter 8. We will also briefly describe the related procedure for quantum process tomography (QPT) that has been used in previous works[26].

A.1 State Tomography

Consider first a single qubit. The state of such a system can be described as a point lying in or on the Bloch sphere, with pure states lying on the surface and mixed states lying in the interior. To locate a point in this sphere requires three coordinates; a typical measurement reveals the probability P_0 for the system to be in state $|0\rangle$ and the probability P_1 for the system to be in state $|1\rangle$, which translate into the position along the Z-axis of the Bloch sphere: $\langle Z\rangle = P_0 - P_1$. To obtain the X- and Y-coordinates of the state, we simply rotate the Bloch sphere before measuring: rotating about Y by $-\pi/2$ brings the X axis upright, allowing a measurement of the X-coordinate, while rotating about X by $\pi/2$ brings the Y axis upright, allowing a measurement of the Y-coordinate. These three measurements give the complete location of the point on the Bloch sphere, allowing the full quantum state to be reconstructed.[1]

[1]Notice that we rotate about the Y-axis to obtain the X-coordinate and vice versa. In the Bloch sphere picture it is convenient to think about the axes along which we are measuring,

Now suppose instead that we have a system of N qubits in some unknown state ρ. Performing a standard projective measurement will allow us to determine the occupation probability $P_k = \rho_{kk}$ of each eigenstate $|k\rangle$ of the system. These 2^N numbers have $2^N - 1$ degrees of freedom due to normalization; this represents a small "slice" of the full density matrix which has a total of $(2^N)^2 - 1 = 4^N - 1$ degrees of freedom.

In order to obtain the other needed slices, we perform all possible combinations of rotations on each qubit before measuring the system to obtain the diagonal probabilities. With two qubits, for example, the rotations would be $I \otimes I$, $I \otimes X$, $I \otimes Y$, $X \otimes I$, $X \otimes X$, $X \otimes Y$, $Y \otimes I$, $Y \otimes X$, and $Y \otimes Y$, where I is the identity operation and X and Y represent rotations about the respective axes by $\pi/2$. Similarly, for more qubits we apply all combinations of I, X and Y rotations on all qubits. There are a total of 3^N such combined unitary rotations which we denote by U^j, where the index j runs from 0 to $3^N - 1$[2]. After rotation, the system is in the state $\rho^j = U^j \rho (U^j)^\dagger$ and the measurement reveals the diagonal elements

$$P_k^j = (\rho^j)_{kk} = \langle k | U^j \rho (U^j)^\dagger | k \rangle . \tag{A.1}$$

Altogether, these measurements yield $3^N(2^N - 1) = 6^N - 3^N$ numbers from which

with the rotations simply being used to bring the measurement axis into coincidence with the Z-axis. In general, however, it will be more useful to think in terms of the pre-measurement rotations themselves.

[2]We can think of the index j as an N-digit base-3 number by making the correspondences $I \leftrightarrow 0$, $X \leftrightarrow 1$ and $Y \leftrightarrow 2$.

to determine the density matrix with $4^N - 1$ degrees of freedom. For $N = 1$ these quantities are equal, $6^1 - 3^1 = 4^1 - 1 = 3$, so that we have exactly enough information, as we saw above. For $N \geq 2$ we have $6^N - 3^N > 4^N - 1$ so that this tomographic measurement protocol gives more than enough information to determine the density matrix. In fact the amount of extra information grows exponentially with the number of qubits, so that this protocol is highly inefficient, but for reasonably small numbers of qubits this is of little concern; even if we had a perfectly efficient protocol, the information in the density matrix grows exponentially, so tomography is never going to be a scalable procedure.

Having made the measurements and collected the results, the problem of tomography then becomes to invert the system of equations A.1, solving for ρ in terms of the measurements P_k^j. This can be solved by converting these equations into a single matrix equation; we do this by straightforward but somewhat cryptic manipulation of indices. First, we introduce dummy indices on all the matrix multiplications to obtain

$$\begin{aligned} P_k^j &= (U^j)_{km}\rho_{mn}((U^j)^\dagger)_{nk} \\ &= (U^j)_{km}\rho_{mn}(U^j)^*_{kn} \\ &= (U^j)_{km}(U^j)^*_{kn}\rho_{mn} \end{aligned} \tag{A.2}$$

where there is an implied summation over indices m and n, and where in going to the last line we have simply reordered terms since these three indexed terms

are just c-numbers. The two U terms essentially combine to form a tensor object with 4 indices k, j, m, and n. Instead of solving this tensor relation directly, we can instead convert it into a matrix expression by "reshaping" the objects P_k^j and ρ_{mn} into one-dimensional vectors, and the U tensor accordingly into a matrix. To do this, we introduce two new indices $A(j,k) = 2^N j + k$ and $B(m,n) = 2^N m + n$. These relations are invertible[3] by taking $k(A) = A \mod 2^N$ and $j(A) = \lfloor A/2^N \rfloor$[4], and similarly for m and n. Hence, we can reshape the measured probabilities into a vector $\tilde{P}_A = P^{j(A)}_{k(A)}$ and similarly $\tilde{\rho}_B = \rho_{m(B)n(B)}$. Likewise, the 4-index U tensor becomes a matrix $\tilde{U}_{AB} = (U^{j(A)})_{k(A)m(B)}(U^{j(A)})^*_{k(A)n(B)}$. This finally gives the matrix equation

$$\tilde{P}_A = \tilde{U}_{AB}\tilde{\rho}_B \tag{A.3}$$

which can be inverted to obtain from the measured $\tilde{P}_A$ the desired $\tilde{\rho}_B$, which can then be reshaped one final time to obtain ρ. As mentioned above, the problem is overconstrained since $3^N \cdot 2^N = 6^N$ input measurements are used to find 4^N elements of the density matrix, so that this inversion is typically done as a least-squares optimization to find the closest fit solution. Given the presence of experimental noise on the measured quantities, this overdetermination is quite helpful in achieving a robust solution.

[3]Note that this sort of index manipulation is much more straightforward when indices start with 0, rather than 1. This is one of the primary reasons for choosing this convention, rather than starting indices from 1.

[4]In Python the builtin function `divmod` does what we need: `j, k = divmod(A, 2**N)`.

A.2 Process Tomography

In quantum computing we are often concerned not only with creating particular quantum states, but with creating particular *gates* which are unitary transformations U that take any given input state $|\psi\rangle$ to a transformed output state $|\psi'\rangle = U|\psi\rangle$. Suppose we wish to understand how closely an experimentally-realized gate approximates the desired unitary. To do this, we must recognize first that experimental states are never completely pure states and second that experimental gates are never completely unitary. Hence a more general description of the action of the gate is required, and this is provided by the theory of quantum operations[29, chap. 8]. In this theory, a quantum operation $\mathcal{E}$ is described as a mapping from an input density matrix ρ to an output density matrix $\eta = \mathcal{E}(\rho)$. For our purposes, there are two important ways in which these quantum operations can be expressed. First, in the χ-matrix representation, we pick a basis of unitary operators A^j which span the space of all possible unitaries, and then any operation $\mathcal{E}$ can be expressed as

$$\eta = \mathcal{E}(\rho) = \sum_{jk} \chi_{jk} A^j \rho (A^k)^\dagger. \tag{A.4}$$

The matrix χ_{jk} depends on the choice of basis but otherwise contains all the information about the quantum operation. Second, in the operator-sum representation, a quantum operation $\mathcal{E}$ can be expressed as $\mathcal{E}(\rho) = \sum_k E^k \rho (E^k)^\dagger$, where the summation is over a single index k, and the E^k are unitary matrices known

as the operation elements of $\mathcal{E}$. This representation can be found by diagonalizing the χ matrix; in a sense it gives a set of basis operators for the operation $\mathcal{E}$. Because the operator-sum representation can be found by diagonalizing the χ-matrix, we will discuss only how to determine the χ-matrix.

We begin by preparing a spanning set of input states ρ^i where i is an index labelling the states. These input density matrices can either be assumed to be created perfectly, in which case the ρ^i are taken from theory, or the input states can be measured with state tomography as described in the previous section to determine the input states that were actually created. Generally, we prefer the latter approach. Next, each input state is subjected to the quantum operation we wish to characterize, resulting in the output states η^i, which are likewise measured using quantum state tomography.

Now, from the measured input and output states ρ^i and η^i we must determine χ. To do this, we rewrite Equation A.4 by inserting all matrix indices, as we did above for state tomography, and then reordering the c-number terms to group the unknowns together. We obtain

$$\begin{aligned}(\eta^i)_{mn} &= \chi_{jk}(A^j)_{mp}(\rho^i)_{pq}((A^k)^\dagger)_{qn} \\ &= \chi_{jk}(A^j)_{mp}(\rho^i)_{pq}(A^k)^*_{nq} \\ &= \chi_{jk}(A^j)_{mp}(A^k)^*_{nq}(\rho^i)_{pq}. \end{aligned} \tag{A.5}$$

Now, we reshape the input and output density matrices by introducing two new

indices $A = 2^N m + n$ and $B = 2^N p + q$, recalling that these are invertible so that we can regard m and n as functions of A, and similarly p and q as functions of B. After the reshaping we then have $\tilde{\eta}^i_A = \rho^i_{mn}$, *etc.*, and Equation A.5 becomes

$$(\tilde{\eta}^i)_A = \chi_{jk}(A^j)_{m(A)p(B)}(A^k)^*_{n(A)q(B)}(\tilde{\rho}^i)_B. \tag{A.6}$$

Note that the terms involving χ and the operator basis elements A depend only on the indices A and B. Thus we can define an object we call the "pointer-basis" χ-matrix:

$$\chi^{\text{pointer}}_{AB} \equiv \chi_{jk}(A^j)_{m(A)p(B)}(A^k)^*_{n(A)q(B)}. \tag{A.7}$$

The term "pointer basis" refers to the fact that each element of this object $\chi^{\text{pointer}}_{AB}$ relates the elements ρ_B and η_A in the input and output matrices; in other words, each element of χ^{pointer} "points" at one element each of the input and output matrices, rather than relating them through some operator basis. This matrix is useful primarily because it is easy to calculate; in Equation A.6 we simply regard all the input density matrices as a matrix $\tilde{\rho}_{Bi}$ and similarly regard all the output density matrices as a matrix $\tilde{\eta}_{Ai}$ (we take reshaped input and output matrices as corresponding columns of these two matrices), and then we have

$$\tilde{\eta}_{Ai} = \chi^{\text{pointer}}_{AB}\tilde{\rho}_{Bi} \tag{A.8}$$

which can be easily solved for $\chi^{\text{pointer}}_{AB}$ by a simple least-squares inversion, as we

did for state tomography above.

A.2.1 Change of Basis

Having computed the pointer matrix $\chi_{AB}^{\text{pointer}}$, we must now compute the χ-matrix in the desired operator basis A. To do this, we start from the definition of the pointer matrix in Equation A.7 and perform our reshaping trick one final time. We introduce two new indices $\alpha = 4^N A + B$ and $\beta = 4^N j + k$ and the reshaped matrices $\tilde{\chi}_{\alpha}^{\text{pointer}}$ and $\tilde{\chi}_{\beta}$ and from this we obtain

$$\begin{aligned} \tilde{\chi}_{\alpha}^{\text{pointer}} &= (A^{j(\beta)})_{m(A(\alpha))p(B(\alpha))} (A^{k(\beta)})^*_{n(A(\alpha))q(B(\alpha))} \tilde{\chi}_{\beta} \\ &= T^A_{\alpha\beta} \tilde{\chi}_{\beta}. \end{aligned} \tag{A.9}$$

The 4^N by 4^N matrix $T^A_{\alpha\beta}$ depends only on the operator basis A, and can thus be computed one time and stored. This relation can then be inverted in the usual way to find $\tilde{\chi}$, and this vector is then reshaped to give χ in the operator basis, as desired.

A.3 Enforcing Constraints

As was mentioned above, the matrix inversions used to find density matrices and χ matrices are typically overconstrained so that they can be solved using a least-squares minimization to invert the desired matrix relation. Efficient algorithms are

available in most numerical linear algebra packages for performing this inversion, making the process simple and ensuring that it will be somewhat robust against the inevitable statistical errors in the measured values. However, efficiency and numerical stability are not the only requirements on our procedure for extracting a density matrix or χ-matrix. In addition, these matrices must satisfy certain constraints in order to be physically valid. For example, the density matrix must be Hermitian ($\rho^\dagger = \rho$), positive semidefinite ($\langle\psi|\,\rho\,|\psi\rangle \geq 0$ for all $|\psi\rangle$, or equivalently all $\lambda_i \geq 0$ for all eigenvalues λ_i of ρ) and have unit trace ($\mathrm{Tr}\,\rho = 1$). A physically-valid χ-matrix must satisfy the same set of constraints, Hermiticity, unit trace, and positive semidefiniteness. In order for our tomographic reconstructions to be meaningful, the matrices we extract from the measured data must satisfy these constraints.

The first step in tomography is typically to perform the naive matrix inversion as outlined above. We then check that the constraint violations, if they exist, are small. For typical tomographic reconstructions, we find that $\mathrm{Tr}\,\rho \approx 1$ and $\rho^\dagger \approx \rho$, up to numerical error. The requirement of positive semidefiniteness, on the other hand, is usually not satisfied; the resulting density matrix may have small negative eigenvalues with a magnitude a few percent that of the largest positive eigenvalues. The size of these negative eigenvalues typically corresponds to the binomial uncertainty expected from the number of repetitions of the experiment,

and so such small deviations from physicality are not surprising; they simply indicate the statistical uncertainty of the measurement process and not that the tomography procedure is faulty or the quantum state nonphysical. Having passed these "sanity checks" that the matrix is as close to being physically valid as can be expected, we can then employ one of two techniques to extract a matrix that absolutely satisfies the required constraints.

A.3.1 Maximum Likelihood Estimation

Maximum likelihood estimation (MLE)[15] is a common statistical technique for estimating some quantity based on a set of measurements. For any given value of the quantity, we calculate the "likelihood" that this value would produce the observed measurement results. The estimated quantity is then taken to be that for which the likelihood of the observed measurements is maximized.

Essentially all the qubit measurements we make involve trying to determine a probability by repeatedly performing an experiment that gives a binary answer. Suppose the underlying probability is p and we repeat the measurement n times. The probability of obtaining a result of "1" exactly k times is given by the binomial distribution, and is equal to

$$B(k;n,p) \equiv \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}. \tag{A.10}$$

Normally, we think of p as being fixed and the number of results k varying, but

the trick of maximum likelihood is to reverse this. The number of results k is fixed (this is the number we measure, the input data) and the probability p can vary (this is what we wish to determine based on our measurement). Thus we think of Equation A.10 not as giving the probability of getting k results given underlying probability p, but rather as the *likelihood* that the underlying probability is p *given* that the result "1" was obtained k times. We denote this likelihood by $\mathcal{L}(p)$, and the problem is to find p that will maximize it.

Note that maximizing $\mathcal{L}$ is equivalent to maximizing

$$\log \mathcal{L} = \log(n!/k!(n-k)!) + k \log(p) + (n-k) \log(1-p). \tag{A.11}$$

The first term is independent of p, and so we only need worry about the latter two. The derivative of $\log \mathcal{L}$ with respect to p is

$$\frac{d \log \mathcal{L}}{dp} = \frac{k}{p} - \frac{n-k}{1-p}. \tag{A.12}$$

Setting this derivative equal to zero and solving for p to find the extremum gives $p = k/n$. This result is hardly surprising; it simply says if we perform n trials and get the result "1" exactly k times, the *most likely* value of the underlying probability is k/n, exactly as we would expect.

Now consider the crosstalk free measurement scheme described in Chapter 7. For the moment we will assume there are three qubits with perfect measurement fidelity. We measure each subset of qubits n times and record the number of null results k_{000}, k_{00x}, *etc.* Normally we would immediately convert these to

probabilities $P_{000} = k_{000}/n$, $P_{00x} = k_{00x}/n$, *etc.* and then use Equation 7.10 to solve for the state occupation probabilities themselves. However, when this equation is inverted we obtain

$$\begin{aligned}
\mathbf{P} &= \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\
1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 \\
-1 & 1 & 1 & -1 & 1 & -1 & -1 & 1
\end{pmatrix} \cdot \mathbf{P}_{\text{null}} \\
&= \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \mathbf{P}_{\text{null}}.
\end{aligned} \tag{A.13}$$

The problem with this equation is those minus signs. Because the k's are measured independently, they are each subject to independent statistical noise, so we have no guarantee that, for example, $0 \leq P_{001} = P_{00x} - P_{000} = (k_{00x} - k_{000})/n \leq 1$, as required of probabilities. To solve this we need to do two things: first, we need to constrain the search so that we only consider valid probabilities, and exclude negative numbers; second, we need to maximize the likelihood for probabilities given *all* the measured data, rather than considering individual null counts k independently.

To enforce the constraints, we can reparametrize the problem. Instead of minimizing a likelihood function over occupation probabilities $\mathbf{P} = (P_{000}, P_{001}, \ldots)^T$, we introduce a new set of variables $\mathbf{t} = (t_{000}, t_{001}, \ldots)^T$ and we define $P_{abc} =$

$t^2_{abc}/(\mathbf{t}\cdot\mathbf{t})$. From this definition, it is clear that $P_{abc} \geq 0$ due to the square, and that $\sum P_{abc} = 1$ due to the normalization, so that the components of $\mathbf{P}$ will be a valid set of probabilities, no matter what the value of $\mathbf{t}$. Now, if we rewrite the likelihood function $\mathcal{L}$ in terms of $\mathbf{t}$, then the problem to be solved is still to minimize this function, but the minimization is completely unconstrained with respect to $\mathbf{t}$. Of course the solution is no longer unique in terms of $\mathbf{t}$, but will still give a unique answer when we calculate the probabilities $\mathbf{P}$.

To make the maximum likelihood work for the combined set of measurement data, rather than the individual null counts, note that the counts are taken in independent experiments, so the probability of getting a set of results $(k_{000}, k_{00x}, \ldots)$ is just the product of several binomial distribution factors

$$\left(\begin{array}{c} n \\ k_{000} \end{array} \right) P_{000}^{k_{000}} (1 - P_{000})^{n-k_{000}} \times \left(\begin{array}{c} n \\ k_{00x} \end{array} \right) P_{00x}^{k_{00x}} (1 - P_{00x})^{n-k_{00x}} \times \ldots \qquad \text{(A.14)}$$

where $P_{00x} = P_{001} + P_{000}$, *etc.* We now regard this instead as a likelihood function $\mathcal{L}(\mathbf{P})$, rewrite it in terms of $\mathbf{t}$ as described in the previous paragraph, and then take the logarithm as we saw above to obtain finally

$$\log \mathcal{L}(\mathbf{t}) = C + k_{000} \log P_{000}(\mathbf{t}) + k_{00x} \log P_{00x}(\mathbf{t}) + \ldots \qquad \text{(A.15)}$$

where C is a constant from the various combinatorial factors that can be ignored in the minimization, and where $P_{00x}(\mathbf{t}) = P_{001}(\mathbf{t}) + P_{000}(\mathbf{t}) = (t^2_{001} + t^2_{000})/(\mathbf{t}\cdot\mathbf{t})$ and so on for the other probability terms. To find the actual occupation prob-

abilities, we then simply perform a numerical minimization of this log-likelihood function, for example using `fminsearch`. Including finite measurement fidelity simply changes slightly the equations relating the null result probabilities to the occupation probabilities.

Finally, we discuss briefly the problem of using MLE with state tomography to determine a density matrix given a set of measurement results. Again, the structure of the problem is the same as what we have just seen above. We obtain a set of counts $\mathbf{k}$, each one of which is determined by applying a unitary operation U to the density matrix ρ and then measuring a certain set of qubits, possibly with finite measurement fidelity. The probability we expect for this measurement is $p = \mathbf{F} \cdot \mathrm{diags}(U\rho U^\dagger)$ where $\mathrm{diags}(M)$ is a vector consisting of the diagonal entries of the matrix M, which for a density matrix are the accessible occupation probabilities, and where $\mathbf{F}$ is a "fidelity vector" that determines the relationship between these diagonals and the actual measured result, due to measuring a subset of qubits and/or finite measurement fidelity (given by the rows of the fidelity matrix in Equation 7.12).

The likelihood of obtaining the measured counts is a product of binomial terms, one for each measured count k and expected probability p, and the log-likelihood is a sum of terms, as we have seen above. The only thing remaining is to parametrize the search for the maximum likelihood in such a way that the matrix ρ is always

physical, but the search itself is unconstrained.

It turns out that such a parametrization is quite straightforward[31, chap. 4]. We define a lower-triangular matrix with real diagonals as

$$\Gamma \equiv \begin{pmatrix} t_0 & & & & \\ t_1 + it_2 & t_3 & & & \\ t_4 + it_5 & t_6 + it_7 & t_8 & & \\ t_9 + it_{10} & t_{11} + it_{12} & t_{13} + it_{14} & t_{15} & \\ \vdots & & & & \ddots \end{pmatrix} \tag{A.16}$$

with a total of 64 entries for the 8×8 three-qubit case. We then set $\rho = \Gamma^T\Gamma / \operatorname{Tr}(\Gamma^T\Gamma)$. It can be shown that ρ defined in this way is Hermitian, positive semidefinite, and has unit trace, as required for a physical density matrix. What is more, *every* Hermitian positive semi-definite matrix can be written in this way, so that by rewriting our likelihood in terms of these variables $\mathbf{t} = (t_0, t_1, \ldots)$ through the parametrization of ρ, we can perform an unconstrained minimization of $\mathcal{L}$. This yields the most likely density matrix ρ, given the measured null results.

As might be expected, the minimization of the likelihood function over the 64-dimensional space given by the vector $\mathbf{t}$ can be rather inefficient. The process must be seeded, which we do by taking the output from the simple least-square inversion described above, zeroing out any negative eigenvalues, and renormalizing the resulting matrix to have unit trace. This is a simple though non-rigorous means of finding a "close" physical density matrix that serves as a reasonable initial guess and typically allows the unconstrained search to find a minimum

reasonably quickly. However this entire approach quickly breaks down for larger numbers of qubits due to the exponential growth of the dimensionality of the search space.

A.3.2 Semidefinite Programming

Semidefinite programming (SDP)[37] is another technique that can be used to find a physical density matrix given our measurement data. SDP is one of a family of optimization algorithms that generally fall under the term "convex optimization." A full explanation of these techniques is far beyond the scope of this text; we will note only that the problem of finding a "nearby" matrix that is close to the one derived by the naive matrix inversion and satisfies the physicality constraints on a density matrix falls into this category of optimization problem. As such, we can rely on the well-defined mathematical structure of these problems, and the many techniques that have been developed for solving them efficiently. Using a solver package such as YALMIP[5] for MATLAB allows such a problem to be stated concisely and solved quickly.

Figure A.1 shows a simple MATLAB function that uses YALMIP to enforce physicality constraints. The function accepts a matrix `rho` which typically will be the density matrix determined from a simple least-square inversion as described

[5] http://users.isy.liu.se/johanl/yalmip/

```
function [rhoFit, twoNormDist] = fitTwoNorm(rho)
% find the 2-norm-closest physical density matrix

[m n] = size(rho);

% define matrix to be optimized
rhoVar = sdpvar(n, n, 'hermitian', 'complex');

% build constraints: unit trace and non-negativity
C = [trace(rhoVar) == 1, rhoVar >= 0];

solvesdp(C, trace((rhoVar-rho) * (rhoVar-rho)));

twoNormDist = sqrt(double(real(trace((rhoVar-rho) * (rhoVar-rho)))));
rhoFit = double(rhoVar);
```

Figure A.1: Enforcing physicality constraints with YALMIP.

above. The problem is set up by defining a variable `rhoVar` which we declare to be of the same size as `rho` and to be Hermitian. We then set up a list of constraints, indicating that `rhoVar` should have unit trace and be positive semidefinite. Next this list of constraints is passed to the `solvesdp` function, which solves for the value of `rhoVar` that minimizes the quantity `trace((rhoVar-rho) * (rhoVar-rho))`[6]. Finally, we return the solution matrix `rhoFit` and the distance between the initial matrix and the solution.

There are a few things to note about this semidefinite programming solution for enforcing constraints. First is the fact that we wrote very little code; we essentially just *described* our problem to the YALMIP library and it determined

[6]This is equivalent to minimizing the trace 2-norm $\sqrt{\mathrm{Tr}(A^\dagger A)}$, where the conjugate transpose operation is unnecessary since the matrices involved are Hermitian. Minimizing with or without the square root is equivalent, since the square root function is monotonic.

how to solve the problem in an optimized way. This is possible because of the well-defined mathematical structure of this class of optimization problems, which makes for a very convenient method of solution. Second, note that the choice of matrix norm or distance measure was somewhat arbitrary; we chose the trace 2-norm $\sqrt{\mathrm{Tr}(A^\dagger A)}$ because it is very easy to implement, but in fact there are no good criteria by which to judge which matrix norm is the "right" one. Each matrix norm simply gives a slightly different answer as to what is the "closest" physical density matrix.

Hence, the semidefinite programming approach has a well-defined mathematical structure that can be solved quickly and efficiently, but it requires an arbitrary choice of distance measure and does not take into account any information about how the data were taken, such as the number of repetitions. In addition, the SDP solution assumes that the starting density matrix coming from the least-square fitting procedure is "close" to the true physical density matrix, and then applies a purely mathematical transformation to correct it for physicality. To get the initial guess for input to SDP, we must apply fidelity corrections to the measurements themselves and even apply some other technique such as MLE to correct the probabilities measured in a given run so that they are are non-negative and sum to one. In other words, SDP allows for complete freedom in the way we do the tomography, and then no matter what density matrix it is given, it will

produce a physical density matrix. It is simply a bolt-on final step in the process, but the overall soundness of the tomography procedure depends more on the steps leading up to the application of SDP.

The MLE technique outlined in the last section, on the other hand, is an all-in-one approach. It takes as inputs the raw experimental measurements and directly outputs a physical density matrix, taking into account the statistical aspects of the measurements and requiring no intermediate steps. There is just one correct way to implement the tomography using MLE. However, finding the solution requires an unstructured search over a very high-dimensional space, which can be quite inefficient, and scales poorly as the size of the density matrix increases. However, for the three-qubit density matrices measured in our experiments the MLE technique worked sufficiently quickly and was used for all the tomography presented in Chapter 8.

Bibliography

[1] A. Acín, D. Bruss, M. Lewenstein, and A. Sanpera. Classification of Mixed Three-Qubit States. *Physical Review Letters*, 87(4):040401, Jul 2001.

[2] M. Ansmann. *Benchmarking the Superconducting Josephson Phase Qubit: The Violation of Bell's Inequality*. PhD thesis, University of California, Santa Barbara, 2009.

[3] M. Ansmann, H. Wang, R. C. Bialczak, M. Hofheinz, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, A. N. Cleland, and J. M. Martinis. Violation of Bell's inequality in Josephson phase qubits. *Nature*, 461(7263):504–506, September 2009.

[4] J. S. Bell. On the Einstein Podolsky Rosen Paradox. *Physics*, 1:195–200, 1964.

[5] R. Bialczak, M. Ansmann, M. Hofheinz, E. Lucero, M. Neeley, A. O'Connell, D. Sank, H. Wang, J. Wenner, M. Steffen, A. Cleland, and J. Martinis. Quantum Process Tomography of a Universal Entangling Gate Implemented with Josephson Phase Qubits. *Nature Physics*, 6.

[6] K. Cao, Z. W. Zhou, G. C. Guo, and L. He. Efficient numerical method to calculate the three-tangle of mixed states. *Physical Review A*, 81(3):034302+, Mar 2010.

[7] V. Coffman, J. Kundu, and W. K. Wootters. Distributed entanglement. *Physical Review A*, 61(5):052306+, Apr 2000.

[8] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf.

Demonstration of two-qubit algorithms with a superconducting quantum processor. *Nature*, 460(7252):240–244, June 2009.

[9] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Physical Review A*, 62(6):062314, Nov 2000.

[10] C. Eltschka, A. Osterloh, J. Siewert, and A. Uhlmann. Three-tangle for mixtures of generalized GHZ and generalized W states. *New Journal of Physics*, (4):043014, 2008.

[11] A. Galiautdinov, M. W. Coffey, and R. Deiotte. Greenberger-Horne-Zeilinger state protocols for fully connected qubit networks. *Physical Review A*, 80(6):062302+, Dec 2009.

[12] A. Galiautdinov and J. M. Martinis. Maximally entangling tripartite protocols for Josephson phase qubits. *Physical Review A*, 78(1):010305, Jul 2008.

[13] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger. Bell's theorem without inequalities. *American Journal of Physics*, 58(12):1131–1143, 1990.

[14] M. Hofheinz, H. Wang, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, J. Wenner, J. M. Martinis, and A. N. Cleland. Synthesizing arbitrary quantum states in a superconducting resonator. *Nature*, 459(7246):546–549, May 2009.

[15] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White. Measurement of qubits. *Physical Review A*, 64(5):052312, Oct 2001.

[16] A. E. Kennelly. Equivalence of triangles and stars in conducting networks. *Electrical World and Engineer*, 34:413–414, 1899.

[17] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien. Quantum computers. *Nature*, 464(7285):45–53, March 2010.

[18] E. Lucero, M. Hofheinz, M. Ansmann, R. C. Bialczak, N. Katz, M. Neeley, A. D. O'Connell, H. Wang, A. N. Cleland, and J. M. Martinis. High-Fidelity Gates in a Single Josephson Qubit. *Physical Review Letters*, 100(24):247001, 2008.

[19] Y. Makhlin. Nonlocal properties of two-qubit gates and mixed states, and the optimization of quantum computations. *Quantum Information Processing*, 1(4):243–252, August 2002.

[20] J. M. Martinis, K. B. Cooper, R. Mcdermott, M. Steffen, M. Ansmann, K. D. Osborn, K. Cicak, S. Oh, D. P. Pappas, R. W. Simmonds, and C. C. Yu. Decoherence in josephson qubits from dielectric loss. *Physical Review Letters*, 95(21):210503, 2005.

[21] S. Matsuo, S. Ashhab, T. Fujii, F. Nori, K. Nagai, and N. Hatakenaka. Generation of Macroscopic Entangled States in Coupled Superconducting Phase Qubits. *Journal of the Physical Society of Japan*, 76(5):054802, May 2007.

[22] J. R. Mautz, R. F. Harrington, and C. G. Hsu. The inductance matrix of a multiconductor transmission line in multiple magnetic media. *Microwave Theory and Techniques, IEEE Transactions on*, 36(8):1293–1295, 1988.

[23] R. McDermott, R. W. Simmonds, M. Steffen, K. B. Cooper, K. Cicak, K. D. Osborn, S. Oh, D. P. Pappas, and J. M. Martinis. Simultaneous State Measurement of Coupled Josephson Phase Qubits. *Science*, 307(5713):1299–1302, 2005.

[24] N. D. Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Physical Review Letters*, 65(15):1838–1840, Oct 1990.

[25] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm. Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits. *Physical Review Letters*, 103(11):110501, 2009.

[26] M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, N. Katz, E. Lucero, A. O'Connell, H. Wang, A. N. Cleland, and J. M. Martinis. Process tomography of quantum memory in a josephson-phase qubit coupled to a two-level state. *Nat Phys*, advanced online publication, April 2008.

[27] M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, N. Katz, E. Lucero, A. O'Connell, H. Wang, A. N. Cleland, and J. M. Martinis. Transformed dissipation in superconducting quantum circuits. *Physical Review B*, 77(18):180508+, May 2008.

[28] M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, E. Lucero, A. D. O'Connell, D. Sank, H. Wang, J. Wenner, A. N. Cleland, M. R. Geller, and J. M. Martinis. Emulation of a Quantum Spin with a Superconducting Phase Qudit. *Science*, 325(5941):722–725, August 2009.

[29] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, October 2000.

[30] A. Osterloh, J. Siewert, and A. Uhlmann. Tangles of superpositions and the convex-roof extension. *Physical Review A*, 77(3):032310+, Mar 2008.

[31] M. Paris and J. Rehacek, editors. *Quantum State Estimation*. Springer, 2004.

[32] N. Schuch and J. Siewert. Natural two-qubit gate for quantum computation using the XY interaction. *Phys. Rev. A*, 67(3):032301, Mar 2003.

[33] R. W. Simmonds, K. M. Lang, D. A. Hite, S. Nam, D. P. Pappas, and J. M. Martinis. Decoherence in josephson phase qubits from junction resonators. *Physical Review Letters*, 93(7):077003, 2004.

[34] M. Steffen, M. Ansmann, R. C. Bialczak, N. Katz, E. Lucero, R. McDermott, M. Neeley, E. M. Weig, A. N. Cleland, and J. M. Martinis. Measurement of the Entanglement of Two Superconducting Qubits via State Tomography. *Science*, 313(5792):1423–1425, 2006.

[35] M. Steffen, J. M. Martinis, and I. L. Chuang. Accurate control of josephson phase qubits. *Phys. Rev. B*, 68(22):224518, Dec 2003.

[36] G. Toth, W. Wieczorek, R. Krischek, N. Kiesel, P. Michelberger, and H. Weinfurter. Practical methods for witnessing genuine multi-qubit entanglement in the vicinity of symmetric states. *New Journal of Physics*, 11(8):083002, 2009.

[37] L. Vandenberghe and S. Boyd. Semidefinite Programming. *SIAM Review*, 38(1):49–95, 1996.

[38] G. Vidal. Entanglement monotones. *J. Mod. Opt.*, 47:355, 2000.

[39] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40(8):4277–4281, Oct 1989.