Abstract

Error-Detected Networking for 3D Circuit Quantum Electrodynamics

Luke David Burkhart

2020

Quantum machines have the potential to serve as groundbreaking tools for scientific discovery in the coming decades. As the complexity of these devices increases, it may be necessary to borrow ideas from complex classical systems, and build them in a modular fashion, with independently designed, optimized, and tested components, networked together into a functioning whole. To build a modular machine from superconducting circuits requires the ability to perform operations between quantum bits housed in separate modules. For this, we must be able to move qubits between modules, or generate entanglement across the network, conveying information in the form of photons. In all implementations to date, photon loss in the links between modules is a dominant source of error, which must be overcome in order to build a scalable modular machine. We demonstrate two approaches for rapid and faithful quantum communication and entanglement between modules in a superconducting quantum network. Encoding information in cavity resonators allows application of strategies for error mitigation in harmonic oscillators to detect photon loss in the communication path. Using a low-loss communication bus, we transfer a qubit in a multi-photon encoding and track loss events to improve the fidelity. Furthermore, generating entanglement with two-photon interference and post-selection against loss errors produces a Bell state with half the error obtained in the single photon case. We discuss several routes towards high-fidelity operations in superconducting quantum networks based off these tools.

Error-Detected Networking for 3D Circuit Quantum Electrodynamics

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Luke David Burkhart

Dissertation Director: Robert J. Schoelkopf

May 2020

Contents

List of Figures

List of Tables

Acknowledgments

Like many things in life, this thesis and the experiments described within would not have been possible without the work, assistance, guidance, and love of many individuals.

I'd like to begin by thanking some of the partners and mentors I have had during the past six years. WOLFGANG PFAFF has been one of the most formative influences on my development as a scientist. His ability to cut through the mental clutter and ask the question "yeah, but so what?" is something I truly admire. I am deeply indebted to him for taking me on as a still-young grad student and pointing me at a fun and relevant problem. I hope I have also absorbed some of his unparalleled presentational and aesthetic style. Working alongside CHRIS AXLINE was also a crucial part of my growth as a physicist. His attention to detail and creative problem-solving approaches are impressive, and his work ethic is inspiring. I owe a huge fraction of my technical knowledge to Chris' patience and willingness to pass on his mental database. Nearly two years after he has left the lab, the answer to most questions is "I think Chris tried that already." Wolfgang and Chris have become dear friends, for which I am very thankful.

It has been such an unexpected pleasure to work among the scientific and social communities of the fourth floors of Becton and 17 Hillhouse. I must thank CHEN WANG and YVONNE GAO for taking responsibility for me when I first arrived, before I knew how much I didn't know. I am grateful for their patience and support.

NAT COTTET, and Lev make in the future.

I am also grateful for the real-life friendships that have come from my time in Becton. I would like to thank the contingent of climbing enthusiasts, which has included Wolfgang, Chris A., Yiwen, Katrina, Nick, Ben, Narly, ALEX GRIMM, ANGELA KOU, URI VOOL, VIJAY JAIN, CHRISTA FLÜHMANN, ANIKET MAITI, and many others, for helping me find and keep a hobby I love. I especially want to thank Narly, Alex, Yiwen, Vijay, Katrina, Ben, Chris A., Jacob, Kevin, and BRIAN LESTER for their friendship, generosity, and outstanding taste in beverages[1].

The litany of Becton saints is not complete without thanking GISELLE MAILLET, MARIA RAO, NUCH GRAVES, and TERRI EVANGELISTE for keeping the wheels from falling off the bus, and for providing a source of laughter, assistance, and leftover pizza. Likewise, thanks to FLORIAN CARLE and RACQUEL MILLER for making the Yale Quantum institute the wonderful place that it is to learn, explore, and share. Thanks also to Florian and SPENCER TOPEL for tricking me into a brief return to performing with Quantum Sound. I am thankful for that unique experience, and especially grateful to Kyle for keeping me sane through it.

LUIGI FRUNZIO has been an invaluable source of insight, laughter, and friendship. His tireless effort keeping our fabrication facilities in order would be enough, but his contributions to the scientific output of the group are equally essential. I thank him for years of training, knowledge, advice, and friendship. His attention to detail reviewing the text of this thesis was irreplaceable. Thanks also to Chris W., Sal, James, Ben, and Vijay for catching the few mistakes Luigi missed. Any typographical errors herein must have been added after they read it.

I would also like to thank the community of senior scientists for their support, insight, and guidance, especially my thesis committee, MICHEL DEVORET, STEVE GIRVIN, and JACK HARRIS, for their time and input. I am always impressed by

---

1. Except for Yiwen.

Chapter 1

Introduction

1.1 Quantum computing

The advent of electrical[1] computers has had an impact on the world that cannot be overstated. The effect on daily life is obvious to anyone with a cell phone, and is beyond the scope of this thesis. More close to home is the impact of computing power on technological and scientific development. It is the domain of a physicist to reduce a complicated physical system to one which can be understood on a whiteboard, but the ability to produce detailed numerical simulations for nonlinear systems, or to check the validity of simplifying assumptions, is indispensable. Moreover, there are seemingly simple problems on which one can make no headway beyond educated guesses until modeling numerically. This makes the laptop, or the computing cluster, an essential scientific tool.

The continual advancement of speed, memory and parallelism, without a corresponding increase in size or cost, is a remarkable achievement. One of the drivers of this progress is the well-known Moore's Law, an observation made in 1965 by Gordon Moore [1]. The trend Moore noted, which has persisted for 6 decades, is that

---

1. as a successor to organic and mechanical

new integrated circuits tend to double in their transistor count, and hence number of operations per cycle, every 2 years or so. This growth has largely been driven by miniaturization, as the minimum feature size of commercial semiconductor fabrication has shrunk by a factor of two roughly every 4 years[2], from 10 μm in 1971 [2] to 10 nm in 2016 [3]. A steady exponential over six order of magnitude is even more amazing when one considers that Moore made his observation based on five data points. In addition to improving device density, smaller transistors are faster, since their reduced resistance and capacitance improves switching times, and hence, clock rates.

While Moore-like progress has driven every advance in computing since the start of the Space Race, it cannot continue indefinitely. The simplest reason is that atoms have a finite size, and it is impossible to make metallic features smaller than a few Å that still act like a metal. Of course, there are practical limits before this, and diminishing returns on pushing further miniaturization before this hard lower bound.

A concern one might have is that, as the number of transistors per area increases, the heat dissipated into the chip will rise with it, leading to a thermal runaway. Fortunately, Moore's law has a cousin, Dennard scaling, the 1974 observation [4] that the power consumption of a transistor is *proportional to its area*. This leads to a constant power usage per area on chip as the transistor density increases. This is essential for continued performance improvements, not merely due to the cost of wall-power consumption, but so the processor does not simply melt. The problem is that Dennard scaling started to break down a little more than ten years ago [5], which is partly responsible for the stagnation in processor clock speeds over the last few years. The end result is that, while it may still be possible to push for smaller transistors and denser chips, it may not be a *good* idea.

---

2. The increase in count is twice as fast as the increase in linear dimension because the count follows an area law.

The eventual end, or at least saturation, of traditional transistor density and speed has prompted many alternative routes for the future of computing. New kinds of transistor technology may provide different scaling constraints. However, the emergence of quantum technologies prompts a more drastic paradigm shift.

For a compelling argument for why quantum mechanical hardware might be of interest and use, one can look to the *first* such argument, made by Richard Feynman in 1982 [6] in a lecture about simulating or imitating physical systems on a computer. In particular, he describes a kind of computer which would be adept at predicting probabilistic natural processes, like quantum mechanics. This is an important scientific task in order to gain more understanding and predictive power over real materials. Feynman gives up on the idea of efficiently *calculating* the probability of each possible outcomes, since the number of probabilities to compute is exponentially large in the system size. Instead, one can build a computer which is itself probabilistic, and *simulate* the system and record the output n times, yielding an estimate of the probabilities with relative accuracy proportional to $1/\sqrt{n}$. The crucial insight is that classical machines have a difficult time emulating quantum ones, since quantum mechanics deals with *probability amplitudes*, which do not add the same way classical probabilities — they *interfere*. This makes a classical machine rather bad at simulating a quantum one. One instead considers a computer with elements that behave quantum mechanically themselves.

Feynman's insights spurred significant interest in the new field of quantum simulation. But in the next decade the view of what a quantum computer could do expanded with the discovery of several algorithms which use this interference property to solve computational tasks which have nothing to do with quantum physics at all. Perhaps the first and most-often quoted was Peter Shor's algorithm for integer factorization in sub-linear time, a substantial improvement over the best known classical algorithm [7]. Another, Grover's search algorithm [8], makes clear use of the interference prop-

erties of quantum states to enhance the probability of getting correct answer and suppress the wrong ones. The past two decades has seen continual growth in the quantum algorithms toolbox, as well as in classification and understanding thereof; see for example [9]. Other, more recent, developments have turned back to the original quantum simulation idea, with applications in quantum chemistry. Some of these methods, such as the variational quantum eigensolver [10], lend themselves to the mapping of a wide variety of optimization problems onto the problem of finding the ground state of a quantum mechanical Hamiltonian, and can do this very efficiently.

The ever-growing interest in programmable quantum simulators and universal quantum computers leads naturally to the question of *how* to build such a machine. This thesis aims to carve out a small part of that question.

1.2   Modularity as a design principle

When considering how one might put together a quantum computer, it is worth making some comments about how classical computers are built. The question of computer architectures is far beyond the scope of this thesis, but I will make a few general observations which serve to motivate similar ideas in quantum machines. The central point is that most classical computers[3] are constructed out of simpler building blocks. The idea of such a modular construction has many implications, and is present at all scales of design. Here we comment on a few of the ways in which classical computers are modular in construction, and some of the advantages.

Modular processors and memory

At the heart of a modern computer is a memory unit, which stores data, and a processing unit, which performs operations on said data. But both of these parts

---

3. Most complex machines in general, really.

have many subdivisions, and hierarchical structure. For instance, modern processors have several levels of local memory ("caches"), with reduced size and increased speed as they get closer to the heart. This hierarchy has huge advantages over an isotropic memory bank, which would either be extremely expensive or very slow. Additionally, newer processors may contain several interconnected "cores," which can function as independent processors with their own local caches. At some point, CPU designers realized it was more feasible to connect four cores, rather than make one core four times larger. This allows for a kind of tiling of functional blocks, which provide an organized route towards increased complexity.

Modular assembly

A layer up from the modular design of individual components, there is a modularity in the assembly of those components. The fact that one can buy a processor, RAM, and hard drive from three different vendors and plug them in to a motherboard from a fourth, all before lunchtime, without any advanced technical experience, is a remarkable achievement. This degree of interchangeability was made possible in part by competition in the market, but it also makes sense from a design-oriented perspective. Having these essential parts of a computer be independent in their design, optimization, and manufacture allows progress to occur in parallel. Advances in memory can be made by researchers who are not experts in processor design, for example.

Replaceable parts are also an essential feature for both emerging and mature technologies. Computers in the 1940's had to contend with the high failure rates of vacuum tubes. ENIAC, one of the first general purpose electronic computers, had 18,000 vacuum tubes, and the operators had to replace one about every two days [11]. While modern hardware is much more reliable, failures still happen, especially in large-scale operations. According to a report by Google in 2007 [12], the failure

rate for disk drives in their server farms was a few percent per year; in the population of 100,000 drives included in their study, this is several failures per day[4]. The notion that reliability allows scale is clear here, but it is certain that neither of these examples would be possible without replaceable parts.

The next step: modular networking

The final level of modularity relevant here is the idea of networks. Where do we go once we have built the largest or densest processor we know how to build, but still need more computing power? The answer in classical computing is unquestionably to build a network of smaller machines which can work in parallel. Distributed computing is responsible for a huge variety of large-scale scientific and industrial tasks which are out of reach for single computers, because of time or memory roadblocks. But this idea of networking is especially relevant for early-stage hardware, when it is not yet obvious how one might double the complexity of a processor without encountering a host of new design problems and constraints. Freely-scalable quantum integrated circuits are an active research goal, but many of the current approaches require industrial-level fabrication capabilities, which are simply not widely available to the scientific community. This bottleneck means developing methods for scaling networks of less-complex systems is a valuable approach for the academic community to continue to drive innovation.

Modularity for quantum machines

Modularity is at the heart of design, assembly, and scaling of classical computers, but it is not without challenges. A central task in a modular or networked configuration is communication between modules. This comes with issues of latency, bottlenecks,

---

4. This is even after screening for faulty hardware before installation. Reliability of manufacture is an entirely separate issue. I had to replace brand-new desktop RAM when running simulations for the sample hardware in Chapter 7.

routing, information loss, and error correction. A particularly challenging scaling problem is the issue of getting signals on and off of a chip. As the chip size increases, the number of bits can scale as the area, but the space for interconnects along the edge goes as the perimeter, and hence grows more slowly. This necessitates the use of the third dimension for signal routing. Communication between quantum modules faces all of these problems, as well as some new ones. Namely, the problem of communication errors requires a new set of tools in the quantum domain.

1.3 The structure of this thesis

This thesis describes two experiments which represent complementary routes towards error-corrected quantum networking in a circuit quantum electrodynamic (cQED) platform.

Chapter 2 will lay out some basic formalisms and concepts of quantum information that will run throughout this work, including quantum bits (qubits), measurement, and entanglement. A central theme of this work is quantum errors and error correction, since information loss between modules in cQED networks appears to be a major bottleneck. We will give the basic language for talking about quantum error correction in this chapter. We also present a few envisioned architectures for modular quantum machines and their relation to the experiments which make up this thesis.

Chapter 3 will introduce a vehicle for storing quantum information, the harmonic oscillator, which is well-suited for the task of distributing information in a cQED network. This chapter will end with a discussion of how to detect and correct energy loss errors in a harmonic oscillator. We will begin with an encoding strategy which fails to correct more errors than in introduces. Motivated by insight from this example, we will discuss a few encodings which *can* improve the performance of the system.

Chapter 4 will consider what kind of interactions and mediators we need to net-

work harmonic oscillators in cQED. The central theme will be building links which use microwave transmission line to carry information with photons. Uni- and bi-directional channels for communication and entanglement will be discussed, and the effect of photon loss in the link will be considered in a few different regimes. We will discuss strategies for suppressing the effects of this loss, as well as differences between linking qubits versus oscillators.

Chapter 5 will turn to the implementation of oscillators, qubits, transmission line links, and the couplings discussed in the previous chapter, in a three-dimensional cQED hardware platform. I have really stood on the shoulders of giants when it comes to this, so I will defer much of the technical discussion to other theses [13, 14, 15], and mainly include details which are non-standard, or not covered elsewhere. This includes the design and measurement of coaxial transmission line resonators for the experiment in Chapter 7.

In Chapter 6 we will discuss the *pitch and catch* experiment that I worked on with Chris Axline and Wolfgang Pfaff [16], in which we implemented the proposal in [17] to send quantum states and generate entanglement between two modules with propagating photons in a unidirectional transmission channel. Chris's thesis contains a very complete description of the experiment ([15], Chapter 7), so I will not reproduce the technical details, but rather summarize the results and focus on the advantages and challenges of this approach, and what I learned from this experiment. I will try to focus on aspects of the approach which are not discussed elsewhere.

The last technical chapter will be the *quantum network bus*, on which I worked closely with James Teoh, and for which a manuscript is in preparation [18]. This experiment uses a bidirectional, standing-wave channel to accomplish some of the same goals as pitch and catch, but the use of a higher-quality link, as well as independent readout of the modules, allowed us to implement the error-correction state transfer protocol envisioned in Chapter 6. The bidirectional link also admits new features

based on interference of photons from the two modules, which we use to implement a beamsplitter transformation between microwave cavities in separate modules. We use the beamsplitter, and associated Hong-Ou-Mandel interference [19], to generate an entangled state across the network which we can check for errors locally, allowing us to generate higher-fidelity entanglement with multiple photons.

Finally, Chapter 8 will try to provide an outlook for these two approaches. I'll discuss how these tools can be combined and expanded to build networks with more modules and higher quality communication channels. I'll also explain some broader applications for the quantum network bus, including how it can be used to realize operations over the network, a novel approach which may simplify the networking apparatus.

Chapter 2

Modular Quantum Processors

Given our interest in building a quantum computer, and having discussed some of the appeals of modular design principles, this chapter will consider these two goals together. In Section 2.1, we will lay out some of the essential properties of quantum bits and their manipulation, before turning to interactions between qubits in Section 2.2. Section 2.3 will discuss some of the things that can go wrong. The concept (and prevalence) of quantum errors will bring us to a discussion of quantum error correction by grouping multiple qubits together into a so-called "logical qubit" (Section 2.4. Finally, in Section 2.5 we will consider a few different ideas for modular quantum machines, and looks at how integrating traditional error correction with them may prove difficult. This will motivate the idea of a special kind of logical qubit made of only one object, rather than multiple, which may prove more amenable to essential operations in a modular quantum computer.

2.1 Quantum bits

2.1.1 Bits and Qubits

We begin by describing the elementary building blocks of quantum computer, the quantum bit (abbreviated "qubit"), and how it differs from a classical bit (simply, "bit"). One statement of the difference between bits and qubits that I like (and is some derivative of something I remember first hearing from Michel Devoret) is that bits have a unique description — their state is labeled by either 0 or 1. For qubits the story is rather different, as there is no such unique labeling, but rather a continuous family of labels, or bases, with which to describe the state, all of which are equally good and equivalent.

This lack of a preferred basis is equivalent to the notion that a qubit state can be written as a vector in a two-dimensional vector space (Hilbert space), fully described by two complex numbers. Again, there is a continuous family of equally valid basis states, but the most common way to see the state written is

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

with $\alpha, \beta \in \mathbb{C}$.

We use a handy notation for the inner product, a measure of the length and similarity of two vectors:

$$\langle 0|0\rangle = \langle 1|1\rangle = 1$$
$$\langle 0|1\rangle = \langle 1|0\rangle = 0 \tag{2.2}$$

which means that $|0\rangle$ and $|1\rangle$ are orthogonal to one another, and have unit length. Orthogonality is a requirement for these two states to comprise a basis. Finally, we consider the length of $|\psi\rangle$, and find that it is equal to one as long as $|\alpha|^2 + |\beta|^2 = 1$, a condition referred to as normalization. We will often (read: immediately) cheat

and suppress this normalization when it is clear from the context that the state vector should have unit length. Another simplification is that the overall phase of the complex numbers α and β is irrelevant; that is, multiplying the entire state vector by any unit complex number changes no physical observable quantity. So there are really only two *real* parameters needed to completely describe this state.

2.1.2 Projective measurement and operators

To get a better sense of what this is all for, we will introduce the phenomenon of measurement, which is of central importance to quantum mechanics and quantum information machines. After all, if we perform a computation, we need to read out the answer at the end.

When measured, a single qubit always [1] gives a binary outcome. Essentially, measurement turns a qubit back into a bit, at least for a time. The meaning of the coefficients α and β in this context lies in their relation to the probabilities with which the two outcomes of this measurement occur. When performing such a measurement, one finds the answer 0 with probability $p_0 = |\alpha|^2$ and 1 with probability $p_1 = |\beta|^2$. This is the motivation for working with normalized vectors, since the length of ψ is the sum of these two probabilities. If there are only two outcomes, these probabilities had better add up to one. Note that in this imagined experiment, there is no difference between, for example the two states

$$|{+}x\rangle \equiv |0\rangle + |1\rangle$$

$$|{-}x\rangle \equiv |0\rangle - |1\rangle$$

(2.3)

to which we give special names because of their particular importance, to be seen

---

1. When we speak of measurement in this thesis, we consider it to be instantaneous and projective. For some experiments demonstrating the continuous interplay of measurement and unitary dynamics, see [20, 21, 22, 23, 24, 25].

momentarily. Both of the above have $p_0 = p_1 = 0.5$, so they will appear identical to our experiment.

So far, a qubit seems to behave the same as a bit, but with some extra randomness thrown in. But by remembering our earlier comments about the equivalence of different bases, we can note that the above states, which are orthogonal by construction, can be used to describe any other states; for example,

$$|0\rangle \equiv |+z\rangle = |+x\rangle + |-x\rangle$$
$$|1\rangle \equiv |-z\rangle = |+x\rangle - |-x\rangle \tag{2.4}$$
$$|\psi\rangle = \alpha_x |+x\rangle + \beta_x |-x\rangle \qquad \alpha_x = \alpha + \beta, \quad \beta_x = \alpha - \beta.$$

We call this the "x basis." Note the introductions of new names for $|0\rangle$ and $|1\rangle$, which we will use interchangeably, so that we can call it the "z basis" for symmetry. More on those names in a moment. Since there is no preferred basis, anything we can do in one we can do in another. This includes measurement — we can measure a qubit in the x basis just as well as in the z basis. If we do this on the state $|+x\rangle$, we'll find outcome $+1$ every time, and conclude there is no randomness at all. Moreover, if we measure the state $|-x\rangle$ in the x basis, we'll find -1 every time. In this basis, these two states are *maximally* distinguishable. This basis-dependent apparent randomness, which has no classical analogue, is a central feature of qubits.

Any basis can be defined by an operator (matrix) for which the basis states are eigenvectors with distinct eigenvalues. These eigenvalues represent the measurement outcomes, and are sometimes just labels, but can correspond to physically measured values. For the bases discussed so far, these operators are called the Pauli matrices,

and the correspondence is:

$$\sigma_z \left|+z\right\rangle = +1 \left|+z\right\rangle \qquad \sigma_x \left|+x\right\rangle = +1 \left|+x\right\rangle \qquad \sigma_y \left|+y\right\rangle = +1 \left|+y\right\rangle$$

$$\sigma_z \left|-z\right\rangle = -1 \left|+z\right\rangle \qquad \sigma_x \left|-x\right\rangle = -1 \left|+x\right\rangle \qquad \sigma_y \left|-y\right\rangle = -1 \left|+y\right\rangle$$

$$\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{2.5}$$

Note the addition of the y basis, with basis states $\left|\pm y\right\rangle = \left|+z\right\rangle \pm i \left|-z\right\rangle$. We refer to the collection $\{\left|\pm z\right\rangle, \left|\pm x\right\rangle, \left|\pm y\right\rangle\}$ as the cardinal states.

Given an arbitrary basis $\left|\pm n\right\rangle$ in which to make a measurement, the probabilities of the two outcomes when measuring a state $\left|\psi\right\rangle$ are

$$p_+ = |\left\langle +n|\psi\right\rangle|^2 = \left\langle +n|\psi\right\rangle \left\langle \psi|+n\right\rangle \qquad p_- = |\left\langle -n|\psi\right\rangle|^2, \tag{2.6}$$

which is the general case of the z basis measurement discussed above.

To complete this discussion of measurement probabilities, we define the expectation value of a measurement, which is the average of all the possible outcomes, weighted by the probabilities of those outcomes. For a general operator $\mathbb{O}$ with eigenvectors $\{\left|o_j\right\rangle\}$ and eigenvalues $\{\lambda_j\}$, the expectation value is

$$\left\langle \mathbb{O} \right\rangle = \sum_j \lambda_j p_j = \sum_j \lambda_j \left\langle \psi|o_j\right\rangle \left\langle o_j|\psi\right\rangle = \left\langle \psi| \mathbb{O} |\psi\right\rangle, \tag{2.7}$$

where we have used the fact that $\mathbb{O}$ can be written in its eigenbasis as $\mathbb{O} = \sum_j \lambda_j \left|o_j\right\rangle\!\left\langle o_j\right|$.

2.1.3 The Bloch sphere

The discussion of the essential features of single qubits benefits from a geometric visualization. A common way to represent the Hilbert space of a qubit is the Bloch sphere, where the state of the qubit is a vector anchored at the center of a unit sphere
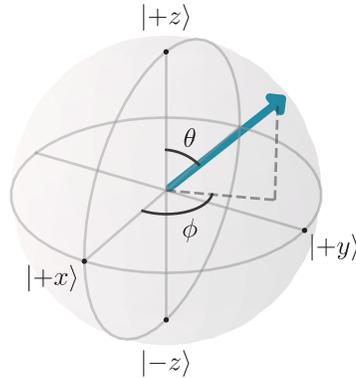
Figure 2.1: The Bloch Sphere. An arbitrary pure qubit state can be parametrized by polar angle θ and azimuthal angle ϕ. A mixed state requires also a length, or can be parametrized by Cartesian coordinates, $(x, y, z) = (\langle \sigma_x \rangle, \langle \sigma_y \rangle, \langle \sigma_z \rangle)$.

and whose tip is on the surface. We can then interpret the magnitude and complex phases of the coefficients of the qubit state as geometric quantities, namely

$$|\psi\rangle = \alpha |+z\rangle + \beta |-z\rangle = \cos\left(\frac{\theta}{2}\right) |+z\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |-z\rangle, \qquad (2.8)$$

with $\phi \in [0, 2\pi]$ the azimuthal angle, and $\theta \in [0, \pi]$ the declination from the north pole. As mentioned in Subsection 2.1.1, these two real numbers are sufficient to describe the state vector. In this representation, shown in Figure 2.1, the pairs of cardinal states introduced above point to opposite sides of the sphere, with $|\pm z\rangle$ pointing to the north and south poles, and $|\pm x\rangle$ and $|\pm y\rangle$ on the equator.

2.1.4 Evolution and gates

So far we have discussed different qubit states on the Bloch sphere, but we haven't discussed how one gets their qubit to go from one state to another. The general prescription is to apply some control to the system, which causes the state to continuously evolve. In this section we will give an example of how one can rotate the qubit state from an eigenvector in the z basis to one in the x basis. A much more complete and pedagogical discussion is given, for instance, in [26]. We describe the

time evolution under a Hamiltonian with the Schrödinger equation

$$i\frac{\partial}{\partial t}\ket{\psi} = \mathcal{H}\ket{\psi} \qquad \mathcal{H} = \Omega\sigma_y, \tag{2.9}$$

which, for a qubit, is just a simple matrix equation. Note that we are using a convention where the Hamiltonian has units of frequency rather than energy, which corresponds to setting $\hbar = 1$. The quantity Ω is called the Rabi frequency, or Rabi rate. This is an eigenvalue equation in the y basis — the time evolution of the states $\ket{\pm y}$ is just a phase factor $e^{\mp i\Omega t}$. However, for superposition of these two states (so, any other state), the evolution is of the form

$$\ket{\psi(t)} = \alpha_y e^{-i\Omega t}\ket{+y} + \beta_y e^{+i\Omega t}\ket{-y} \tag{2.10}$$

By writing any initial state in the y basis and applying this formula, the time evolution can be computed. But an often-more-useful formalism is to integrate the Schrödinger equation and calculate the time-dependent operator which corresponds to this evolution, which is called the *propagator*. For this kind of time-independent Hamiltonian, the propagator is defined as

$$\ket{\psi(t)} = U(t)\ket{\psi(t=0)} = e^{-i\mathcal{H}t}\ket{\psi(t=0)} = e^{-i\Omega t\sigma_y}\ket{\psi(t=0)} \tag{2.11}$$

Through a handy identity (true for any of the Pauli matrices),

$$e^{-i\Omega t\sigma_y} = \cos\Omega t\hat{\mathbb{I}} - i\sin\Omega t\sigma_y \tag{2.12}$$

which will continuously rotate the state about the y axis on the Block sphere, interchanging z and x eigenstates.

This kind of continuous rotation is a tool used in most any quantum information

platform to enact discrete gates. For instance, the above propagator, when $\Omega t = \pi$, will swap the position of the $|\pm z\rangle$ states, as well as the $|\pm x\rangle$ states, and is called a Y gate, or a π pulse about the y axis. The same evolution for $\Omega t = \pi/2$ has the effect

$$|+z\rangle \rightarrow |+x\rangle \rightarrow |-z\rangle \rightarrow -|-x\rangle \rightarrow -|+z\rangle \qquad \mathrm{Y}_{\pi/2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \qquad (2.13)$$

called a $\pi/2$ pulse, often denoted $\mathrm{Y}_{\pi/2}$ to distinguish it from the "full" Y gate, sometime written as Y_π for clarity. We will continue to use this notation of Pauli matrices σ_i when describing Hamiltonians, and capital letters when describing the unitary operations that result.

These types of gates are essential parts of quantum algorithms. They can also be very useful in practical implementations thereof, since not all physical hardware platforms have the same native capabilities. For instance, above we discussed measurement in arbitrary bases, but this is not typically possible. In our platform of choice, cQED, we can usually only measure in the z basis[2]. So how then do we measure in, say, the x basis? The easiest and most common method is to apply a $\mathrm{Y}_{\pi/2}$ rotation on the qubit before measuring in the z basis; this exchanges x for z and effects an x measurement, and is shown in Figure 2.2. Note, however, that there is a phase acquired in Equation 2.13 as one goes around the Bloch sphere. This phase is not visible in this measurement, but it may not be desired when the gate occurs in the middle of some sequence of operations. A more faithful exchange of the x and z bases is the Hadamard gate, defined as

$$\begin{array}{c} |+z\rangle \leftrightarrow |+x\rangle \\ \\ |-z\rangle \leftrightarrow |-x\rangle \end{array} \qquad H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad (2.14)$$

---

2. With certain tricks, one can measure directly in other bases. For examples see [27, 28]

which cannot be implemented directly with a single Pauli Hamiltonian, but can be generated with $\mathcal{H} = \Omega(\sigma_x + \sigma_z)$, or with two consecutive rotations of the kind we have discussed.
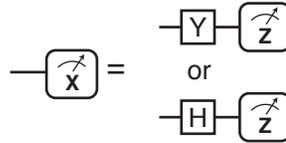
Figure 2.2: Single Qubit Basis Change. A qubit can be measured in the x basis by rotating with a Hadamard or Y gate, then measuring in the z basis.

2.1.5 Impure states and the environment

So far we have discussed quantum states which can be fully described as a state vector living on the surface of the Bloch sphere, which are called *pure states*. We saw that these states can exhibit randomness in measurement, but there is always a basis in which they will give a deterministic result. How does quantum mechanics describe a truly random outcome, that is equally random in any basis? These kinds of processes we can emulate classically by flipping a coin, for instance. These states must be "far" from a pure state in some sense — they must be equally far from any pure state to be equally random in all bases. The point in Figure 2.1 which is equidistant from any point on the surface is the center of the Bloch sphere, so we use vectors inside the Bloch sphere to represent these so-called *mixed states*.

We cannot describe mixed states with a state vector, but instead with a matrix, called the density matrix, for which we use the symbol ρ. Properly, ρ is an operator which can be represented by a matrix. This matrix can still be written in any basis in Hilbert space, and we usually use the z basis like we did for state vectors. The simplest example is the density matrix for a pure state $|\psi\rangle$, which we write is

$$\rho = |\psi\rangle\langle\psi|, \tag{2.15}$$

which seems rather pointless, but is fine. The real use comes when we want to write down our coin flip example. The density matrix for a fair coin is

$$\rho = \frac{1}{2}\left|+z\right\rangle\!\left\langle+z\right| + \frac{1}{2}\left|-z\right\rangle\!\left\langle-z\right| = \begin{bmatrix} \frac{1}{2} & 0 \\[4pt] 0 & \frac{1}{2} \end{bmatrix} \tag{2.16}$$

which we can interpret in words as "equally likely to be in $\left|+z\right\rangle$ as in $\left|-z\right\rangle$." This looks like a superposition, but it is decidedly not. When written as a sum of terms corresponding to orthogonal states, the coefficients are probabilities, and are strictly positive (and sum to one), so they do not interfere the way that amplitudes in a superposition can. This maximally-mixed density matrix is the same in any basis:

$$\begin{aligned} \rho &= \frac{1}{2}\left|+z\right\rangle\!\left\langle+z\right| + \frac{1}{2}\left|-z\right\rangle\!\left\langle-z\right| \\ &= \frac{1}{2}\left|+x\right\rangle\!\left\langle+x\right| + \frac{1}{2}\left|-x\right\rangle\!\left\langle-x\right| \\ &= \frac{1}{2}\left|+y\right\rangle\!\left\langle+y\right| + \frac{1}{2}\left|-y\right\rangle\!\left\langle-y\right|, \end{aligned} \tag{2.17}$$

This non-obvious fact makes clear the basis-independence of the measurement probabilities.

The density matrix lends itself to being described concisely in terms of expectation values of measurements, which is excellent news, since that is what we measure when we take an ensemble average over many copies of the state. The expectation value of an operator $\mathbb{O}$ is generalized from Equation 2.7 using the trace:

$$\left\langle\mathbb{O}\right\rangle = \mathrm{Tr}\left(\rho\mathbb{O}\right) = \sum_{\left|\phi\right\rangle\in\{\left|+z\right\rangle,\left|-z\right\rangle\}} \left\langle\phi\right|\rho\mathbb{O}\left|\phi\right\rangle = \left\langle+z\right|\rho\mathbb{O}\left|+z\right\rangle + \left\langle-z\right|\rho\mathbb{O}\left|-z\right\rangle, \tag{2.18}$$

which gives, in the above coin flip example,

$$\left\langle\sigma_z\right\rangle = \mathrm{Tr}\left(\rho\sigma_z\right) = \frac{1}{2} - \frac{1}{2} = 0 \tag{2.19}$$

In the general case, one writes

$$\rho = \frac{1}{2}\left(\hat{\mathbb{I}} + \langle\sigma_z\rangle\,\sigma_z + \langle\sigma_x\rangle\,\sigma_x + \langle\sigma_y\rangle\,\sigma_y\right) \tag{2.20}$$

Returning to our graphical picture of Hilbert space, the coordinates of the state vector in the Bloch sphere are $(x, y, z) = (\langle\sigma_x\rangle, \langle\sigma_y\rangle, \langle\sigma_z\rangle)$. In contrast to a pure state, the length of this vector is no longer required to be equal to one; for instance, the maximally mixed state is located at $(x, y, z) = (0, 0, 0)$.

But we haven't yet said under what situation this kind of classical randomness appears in a quantum experiment. How might we find a qubit which seems to carry no information at all? Let's say you prepare a qubit in the state $|+x\rangle$ and then measure it in one of the three Pauli bases. You would expect a deterministic answer of $+1$ in the x basis, and randomness in the z and y bases. Imagine, however, that in between the time you prepare the state and measure it, your lab-mate measures it in the z basis, and discards her random answer. In doing this, she will project the qubit into $|+z\rangle$ or $|-z\rangle$. When *you* measure the state, you'll find randomness in the x or y bases, since you're now measuring in the wrong basis. But you'll *also* find randomness in the z basis, because even though the qubit is in a z eigenstate, you don't know which one. You will get the same answer she did, but if she doesn't tell you what it was, you will have nothing to compare your result to. If you repeat this experiment many times, preparing $|+x\rangle$, allowing your lab-mate to measure in the z basis, and then also measuring in the z basis, the statistics of your experiment will suggest the qubit is not prepared in a z eigenstate. Likewise, if you sometimes measure in the x or y bases and look at those results too, the qubit will not appear to be in an eigenstate of *any* basis. This set of experiments is enough to infer the density matrix, since we measure all of the expectation values of Equation 2.20, but they will all be zero, and the qubit will sit at the center of the Bloch sphere — the

maximally mixed state.

This is one way to prepare a mixed state. While it seems a rather contrived example, it is exactly how one can describe what happens when your qubit has un-intended interactions with the environment, which tries to measure the qubit state faster than you can manipulate it. We will discuss this further in Section 2.3. But first, we will investigate the case where your lab-mate is replaced with another quantum system. This situation will produce measurement results on your qubit which are indistinguishable from the case we have just discussed — until we investigate the correlations.

2.2   Interacting qubits and entanglement

In Subsection 2.1.5 we discussed a situation where a qubit will have a completely random measurement outcome in any basis; in stark contrast with a pure state, this *maximally mixed* state contains no quantum information at all. We interpreted this as the information being lost to the environment. But there is another case we must consider, which is when that information is not lost, but is simply stored non-locally. What we mean by this is that two bits of quantum information can be stored in two qubits, but in an inseparable way, where the information is stored in the correlations between the qubits, not in the state of one or the other. When we look at only one of the two qubits, it will look maximally mixed because it contains no information However, when taken together, the correlations between the two qubits will reveal the underlying state. This phenomenon is called *entanglement*, and its properties underly (and in some cases, allow) much of the behavior of a quantum machine.

We will begin this discussion in Subsection 2.2.1 with a brief description of how we write down states for multiple qubits. In Subsection 2.2.3 we will define some entangled states and describe their properties. Subsection 2.2.4 will explain how

one makes entangled states, with gates which act on multiple qubits, and how more complicated quantum circuits can be build from some of these building blocks.

2.2.1 Two-qubit states

We begin by describing two-qubit states in the general. We will still write a state vector, but instead of indicating the state of a particular qubit, it will give the state of the *system*, or collection of qubits. With d qubits, the state lives in a 2^d-dimensional vector space, so we can always describe the state in some basis of that space. A convenient basis is to use strings of 0's and 1's, where each bit in the string corresponds to one of the qubits. For example, for two qubits, one can write

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle = \begin{bmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{bmatrix}. \qquad (2.21)$$

Strictly, these strings are a shorthand: $|01\rangle = |0\rangle \otimes |1\rangle$ (sometimes $|0\rangle|1\rangle$ when the meaning is clear) is called a tensor product, which means that we assume some underlying divisibility to the system Hilbert space. The tensor product just takes two vectors from different Hilbert spaces and glues them together into a vector which lives in this new, larger, space.

A special kind of two-qubit state is one for which Equation 2.21 factors; for instance,

$$\begin{aligned}
|\psi\rangle &= \frac{1}{2}|00\rangle + \frac{1}{2}|01\rangle + \frac{1}{2}|10\rangle + \frac{1}{2}|11\rangle \\
&= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad (2.22) \\
&= |{+}x\rangle \otimes |{+}x\rangle.
\end{aligned}$$

This kind of state, which can be written as a product of two pure states, is called a

separable state. When the system is in such a state, the behavior of the two qubits is completely uncorrelated, and can be prepared by applying independent gates to each qubit: $|+x\rangle \otimes |+x\rangle = (H\,|0\rangle) \otimes (H\,|0\rangle) = (H \otimes H)(|0\rangle \otimes |0\rangle)$. However, there are (many more) states which do not admit such factorization; for example,

$$|\psi\rangle = \frac{1}{2}\,|00\rangle + \frac{1}{2}\,|01\rangle + \frac{1}{2}\,|10\rangle - \frac{1}{2}\,|11\rangle \tag{2.23}$$

This is an example of an entangled state, where the two qubits' states are too linked to be separated out. We will explore the special properties of entangled states beginning in Subsection 2.2.3, and further discuss their use in a quantum computer in Subsection 2.5.3.

2.2.2   How to ignore a qubit

To discuss the properties of entanglement and two-qubit states at large, we need a mechanism for describing measurement of a particular qubit in a two-qubit system. Let's say you are measuring the first qubit (qubit 1) in the z basis, but you can't, or don't want to, measure qubit 2. The framework we use to "ignore" one or more qubits in a multi-qubit state is the partial trace, which takes a state vector and returns a density matrix.[3] The partial trace for a two-qubit system in state $|\psi\rangle$ is defined as

$$\rho_1 = \mathrm{Tr}_2(\rho) = \sum_{|\phi_2\rangle \in \{|+z\rangle, |-z\rangle\}} \langle\phi_2|\,\rho\,|\phi_2\rangle = \langle+z_2|\rho|+z_2\rangle + \langle-z_2|\rho|-z_2\rangle \tag{2.24}$$

where $\rho = |\psi\rangle\langle\psi|$ is the density matrix for the original two-qubit state, and $|\pm z_2\rangle \equiv \hat{\mathbb{I}} \otimes |\pm z\rangle$. Here the notation $\mathrm{Tr}_2(\rho)$ means "trace out qubit 2."

---

3. The fact that tracing out qubits from the system requires the same formalism as Subsection 2.1.5 is a clue that entanglement and the environment have something to do with one another.

Let's give an example. The somewhat trivial case is a separable state like $|00\rangle$:

$$
\begin{aligned}
\rho_1 &= \text{Tr}_2(|00\rangle\langle00|) \\
&= \langle0_2|00\rangle\,\langle00|0_2\rangle + \langle1_2|00\rangle\,\langle00|1_2\rangle \\
&= \langle0_2|0_2\rangle\,|0_1\rangle\langle0_1|\,\langle0_2|0_2\rangle + \cancel{\langle1_2|0_2\rangle}\,|0_1\rangle\langle0_1|\,\cancel{\langle0_2|1_2\rangle} \\
&= |0_1\rangle\langle0_1|
\end{aligned}
\tag{2.25}
$$

which is a pure state! The interpretation of this is that the two qubits knew nothing of one another to begin with, qubit 1 doesn't lose anything by ignoring qubit 2. Now that we have reduced ourselves to a single-qubit density matrix, we can use Equation 2.18 to calculate expectation values of any operator for qubit 1.

2.2.3 Entanglement

What are the properties that define an entangled state? To answer this, we consider an example of a particularly useful and easy-to-work-with entangled state, called a Bell state:

$$
|O_+\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right). \tag{2.26}
$$

There's actually nothing terribly special about this state as compared to any other entangled state. In fact, any so-called maximally entangled state can be turned into any other with the right set of single-qubit gates. But Equation 2.26, along with its cousins $|O_-\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right)$ and $|E_\pm\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle \pm |11\rangle\right)$, form a basis for the Hilbert space and have some readily apparent properties which make them useful[4]. The task of preparing $|O_+\rangle$ will be the subject of much of this thesis, and we'll use this state as an example for several of the properties of entanglement.

The Bell state $|O_+\rangle$ has some obvious properties, and some not-so-obvious ones.

---

4. Here, O and E stand for "odd" and "even", respectively, referring to the parity of the number of excitations in these states.

First we consider the individual measurement probabilities of qubit 1 in the Pauli bases. Tracing out qubit 2

$$\rho = \frac{1}{2}\left(|01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|\right)$$
$$\rho_1 = \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right)$$

(2.27)

yields the maximally mixed state in Subsection 2.1.5. We know that this state has random measurement outcomes in any basis, so it seems this entangled state I told you was special looks just like a random coin.

To see the difference between an entangled state and a mixed state, we need to measure *both* qubits. To do this we define a measurement operator on the two-qubit space: $\sigma_z \otimes \hat{\mathbb{I}}$. This corresponds to measuring qubit 1 and doing nothing to qubit 2, but doesn't require us to trace out the second qubit. The expectation value of this operator is zero (we already showed measuring σ_z gives a random answer), but if after the measurement, the two-qubit state is projected into one of two pure states, $|01\rangle$ or $|10\rangle$, with equal probability depending on the measurement outcome. Let's say we take all the cases where we measured qubit 1 to be in $|0\rangle$. If we proceed to measure qubit 2, we will *always* find it in $|1\rangle$. Likewise if we take the cases where qubit 1 was in $|1\rangle$, qubit 2 is always in $|0\rangle$. We say that the qubits are perfectly anticorrelated in the z basis — there are only two joint outcomes in this basis, which are shown in Figure 2.3. Another way to say this is that qubit 2 has a definite state if and only if you condition the measurement on the state of qubit 1.

This is all well and good, but this anticorrelation in the z basis would also be observed for the mixed state $\rho = \frac{1}{2}\left(|01\rangle\langle 01| + |10\rangle\langle 10|\right)$, which we could have if someone were randomly preparing either separable state $|01\rangle$ or $|10\rangle$ and not telling us which. How do we distinguish the entangled state from this mixed one? We do this in the same way as we did for one qubit in Subsection 2.1.5 — by preparing it many times and measuring in one of a few bases on in each preparation. The entanglement will

a

b $\rho = |01\rangle\langle01| + |10\rangle\langle10|$

c $|\psi\rangle = |01\rangle + |10\rangle$
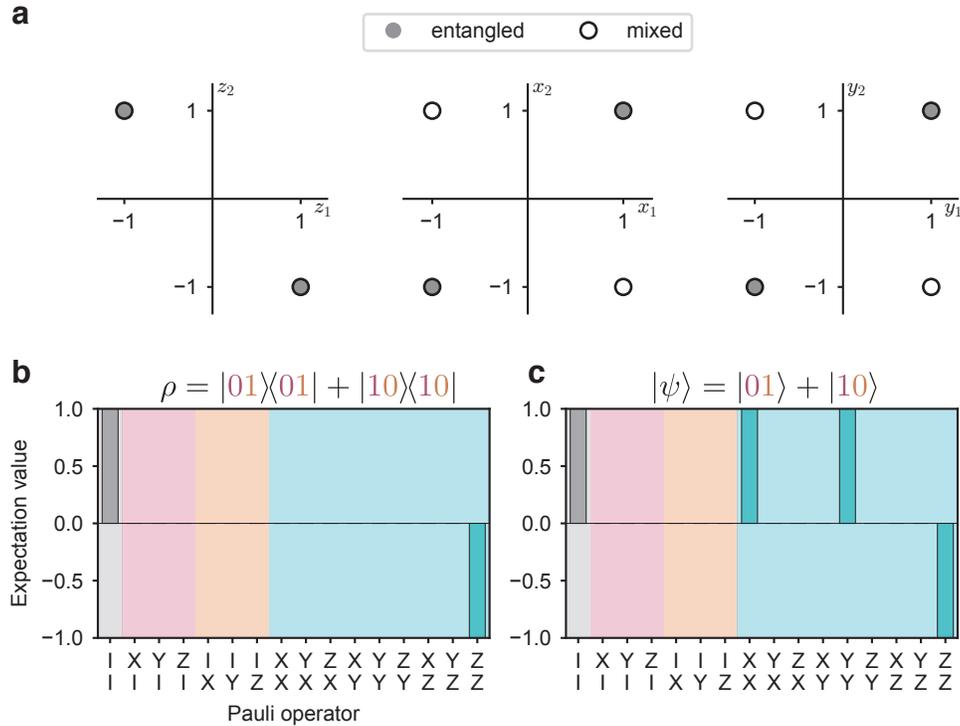
Figure 2.3: Correlations of Entangled States. a) Measurement correlations of two qubits in three bases for a maximally entangled state (filled circles) and a maximally mixed state (open circles). Correlations in z basis are maximal for both, but measurement in x and y bases reveal more correlations for the entangled state. b) Joint Pauli expectation values for mixed (left) and entangled (right) states. Horizontal axis labels denote which pair of Pauli operators are measured.

reveal itself in an overabundance of correlations between the qubits. If we measure both qubits in the x basis, for instance, the mixed state will yield totally uncorrelated results, with equal probabilities for $|+x\rangle\,|+x\rangle$, $|+x\rangle\,|-x\rangle$, $|-x\rangle\,|+x\rangle$, and $|-x\rangle\,|-x\rangle$. However, the entangled state $|O_+\rangle$ will yield *correlation* in the x bases — only the outcomes $|+x\rangle\,|+x\rangle$ and $|-x\rangle\,|-x\rangle$ will occur; likewise for the y basis. These results are summarized in Figure 2.3.

Observations of these correlations are necessary and sufficient for demonstrating entanglement. In analogy to Equation 2.20, the two-qubit state can be written in

terms of all the joint Pauli expectation values:

$$\rho = \frac{1}{2} \sum_{\sigma_1, \sigma_2 \in \{\hat{\mathbb{I}}, \sigma_x, \sigma_y, \sigma_z\}} \langle \sigma_1 \otimes \sigma_2 \rangle \, \sigma_1 \otimes \sigma_2 \tag{2.28}$$

The 16 expectation values in the above expression fully characterize any two-qubit state, and we use them to graphically display the state. Figure 2.3b shows the joint Pauli expectation values for both the maximally mixed state and the maximally entangled state $|O_+\rangle$. They differ clearly in the amount of two-qubit correlations, but both show zero single-qubit expectation values, denoting that there is no information stored locally in either case. A separable pure state like $|+x\rangle \, |+x\rangle$ would show obvious single-qubit Pauli expectation values for the $\hat{\mathbb{I}} \otimes \sigma_x$ and $\sigma_x \otimes \hat{\mathbb{I}}$ operators.

2.2.4 Two-qubit gates

We have given a flavor of some of the properties of entangled states, and in Section 2.5 we will discuss some of their uses, but we haven't yet described how one *makes* an entangled state. This thesis is largely concerned with this question, so we introduce several operations between multiple qubits, which will appear in different contexts throughout.

Entangling gates

Two-qubit gates can be used to generate entanglement, and are important ingredients in quantum algorithms. There is a broad class of such gates which look like a single-qubit gate, controlled on the state of another qubit. The circuit symbols for some important ones are shown in Figure 2.4. A simple one, which can take the separable state in Equation 2.22 into the entangled state in Equation 2.23, is called the ZZ, or CPHASE, gate. The name ZZ is read to mean "a Z gate on qubit 1, if qubit 2 is in the $-z$ eigenstate." Since the Z gate flips the sign of the $|1\rangle$ component, this gate flips

the sign on the $|11\rangle$ state, leaving the others untouched. This gate is symmetric on permutation of the two qubits (it reads equivalently as "a Z gate on qubit 2, if qubit 1 is in the $-z$ eigenstate") and the circuit symbol reflects this notion. The effect of the gate in operator notation is

$$
\text{CPHASE} = \hat{\mathbb{I}} - 2\,|11\rangle\langle 11| =
\begin{array}{cccc}
00 & 01 & 10 & 11
\end{array}
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & -1
\end{bmatrix}
\tag{2.29}
$$

where the column headers on the matrix are a reminder of the basis ordering we are using (as per Equation 2.21). Note that the ZZ gate is *not* the same as the unitary $Z \otimes Z = (Z \otimes I)(I \otimes Z)$. This is just two single-qubit gates applied simultaneously, which does not generate entanglement. However, a Hamiltonian of the form $\mathcal{H}_{\text{CPHASE}} = g\sigma_z \otimes \sigma_z$ does generate a CPHASE-like[5] gate, and is a native interaction in cQED.

Another common two-qubit gate, which is also defined for classical computers, is the XZ or $\text{CNOT}_{1,2}$, gate. Following the same nomenclature, it is "an X gate on qubit 1, if qubit 2 is in the $-z$ eigenstate." We name qubit 1 the "target", and qubit 2 the "control." Its dual, the ZX, or $\text{CNOT}_{2,1}$, gate, has the opposite sense of target and control. In matrix-land,

$$
\text{CNOT}_{1,2} =
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0
\end{bmatrix}
\qquad
\text{CNOT}_{2,1} =
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0
\end{bmatrix}
\tag{2.30}
$$

---

5. Specifically, this Hamiltonian generates a CPHASE followed by some single-qubit Z gates.
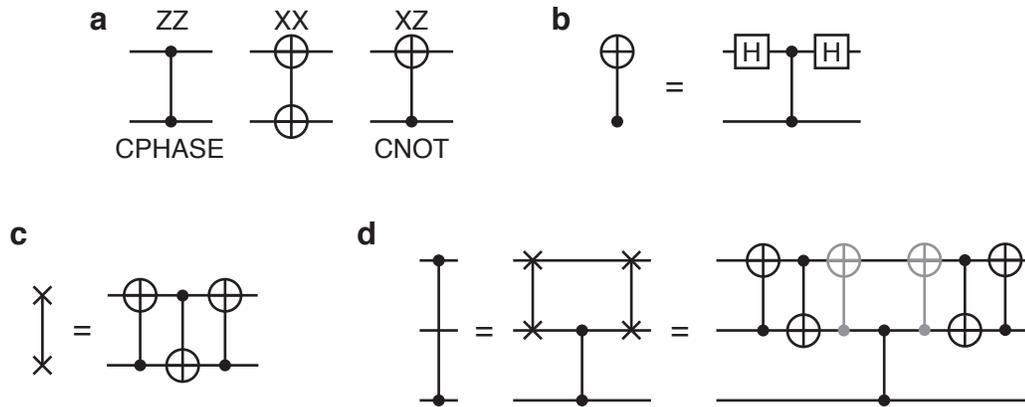
Figure 2.4:   Two-qubit Gates. a) Three universal two-qubit controlled gates. b) A CPHASE gate can be transformed into a CNOT gate with two single-qubit gates on the target (top) qubit. c) A SWAP gate constructed from three CNOT gates. d) A CPHASE gate between two qubits (top, bottom) using a third (middle) as a mediator. SWAP gates between top and middle can be constructed from CNOT gates, two of which (gray) cancel out and are not needed. Implementation with native SWAP can be more efficient.

The circuit symbol in Figure 2.4 reflects the asymmetry of this gate. A non-obvious feature of these gates is that you can change basis, and consider the above XZ gate as "a Z gate on qubit 2, if qubit 1 is in the $-x$ eigenstate" — it is invariant under exchange of the z and x basis *and* the target and control. Another symmetric two-qubit gate is the XX gate, which is an X rotation controlled on the x basis, but this is less commonly used as primitive gate[6].

Each of these two-qubit gates can be maximally-entangling — there are separable states to which they can be applied to produce maximally entangled states. For CPHASE, Equation 2.23 is such an input state. They also have the property of *universality* — any one of them is sufficient for universal control over two-qubits, when coupled with a complete set of single-qubit gates. This makes one or more high-quality two-qubit gates essential for any quantum computing platform. Furthermore, since any of these gates is universal, the others can be synthesized from it with some

---

6. The XX gate is a native operation in trapped ion systems called the the Mølmer-Sørensen gate [29].

single-qubit rotations. Figure 2.4b shows an example of how to construct a CNOT from a CPHASE and Hadamard gates.

SWAP-like gates

There is another class of two-qubit gates which can be used to interchange the state of the qubits. This may be useful in a system with reduced connectivity, where it is necessary to move qubit states around to perform gates between neighbors, for instance. The simplest is the SWAP gate, which does what you might expect:

$$\text{SWAP} : |\psi\rangle |\phi\rangle \rightarrow |\phi\rangle |\psi\rangle \qquad \text{SWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.31}$$

While its matrix form looks similar to $\text{CNOT}_{2,1}$, SWAP cannot produce an entangled state from separable states. Interestingly, SWAP can be constructed from three CNOT gates, as shown in Figure 2.4c, an identity which is also true for classical bits.

The potential utility of SWAP is shown in Figure 2.4d. Imagine you have three qubits {1,2,3}, in a chain with only nearest-neighbor gates. If you want to implement a CPHASE gate between qubits 1 and 3, the conceptually simplest way is the SWAP 1 and 2, perform the CPHASE between 2 and 3, then swap 1 and 2 again. Such an operation can be constructed from an additional four CNOT gates, but native SWAP gates would be much more efficient. This is also of use in a star- or tree-like structure, where some "hub" qubits might be well connected to many other "spoke" qubits, but the spoke qubits might only talk to one hub. In this way the hub can be used to effect gates between spokes.

Finally, there are variations on SWAP which we call SWAP-like, because they

have some similar properties. First is $\sqrt{\text{SWAP}}$, which is defined as

$$\sqrt{\text{SWAP}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{e^{i\pi/4}}{\sqrt{2}} & \frac{e^{-i\pi/4}}{\sqrt{2}} & 0 \\ 0 & \frac{e^{-i\pi/4}}{\sqrt{2}} & \frac{e^{i\pi/4}}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.32}$$

which, unlike SWAP, is a maximally entangling gate. As the name suggests, applying $\sqrt{\text{SWAP}}$ twice yields SWAP.

Somewhat more strangely, there is iSWAP, which looks like SWAP with some single-qubit phases, but is decidedly not.

$$i\text{SWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.33}$$

Though it may not look like it, iSWAP is also a maximally entangling gate. Surprisingly, it is equivalent (up to single-qubit $Z_\pi/2$ gates) to CPHASE×SWAP, as described by Schuch and Siewert [30]. These authors point out that, in a chain of qubits with only nearest-neighbor couplings, it is necessary to swap qubits around, and give a few examples where the ability to do this in a single step while entangling is advantageous. On the other hand, in a system with dense connectivity, SWAP is largely a bookkeeping matter, so iSWAP is just as powerful as CPHASE or CNOT.

The SWAP-type interaction can be obtained with

$$\mathcal{H}_{\text{SWAP−like}} = g\sigma_+ \otimes \sigma_- + g\sigma_- \otimes \sigma_+, \tag{2.34}$$

where we have introduced the raising and lowering operators $\sigma_+ = |1\rangle\langle0|$ and $\sigma_- =$

$|0\rangle\langle 1|$, respectively. Interactions of this type are readily available in cQED, and we will make use of them. As we will see in Chapter 4, when we apply these interactions to harmonic oscillators, we will get something like SWAP and $\sqrt{\text{SWAP}}$, but with some interesting wrinkles.

2.3   Quantum errors

So far we have discussed the ideal operation of gates and measurements, but quantum systems are far from ideal. The power that originates from continuous rotations on the Bloch sphere also make qubits very sensitive to small amounts of noise. This inherent fragility of quantum information, along with the great care that must be taken to protect it, make errors a central challenge in quantum computing.

In this section we will introduce common types of errors which occur on qubits, as well as a common framework for discussing them, the Kraus map, which we will employ in Chapter 3 to discuss errors in oscillators. We will turn in the next section to the topic of how to correct these errors.

2.3.1   Types of quantum errors

In this section we will lay out some error types, but the exact mathematical formalism will be saved for Subsection 2.3.2. In classical information systems there are a few kinds of errors which can occur, most of which have some parallel in quantum machines. A simple one which we will use as an example is the *bit-flip* error, which, well, flips the bit from 0 to 1 and from 1 to 0. In the quantum case, there are two kinds of bit flips, X and Y, but we'll consider just X errors for simplicity. As usual, we'll represent the occurrence of an error with an operator: $A_{\text{bf}} = A_{\text{bf}}^{\dagger} = \sigma_x$. So the occurrence of the error is described by applying this operator to the initial state ρ_{i},

resulting in final state ρ_f:

$$\rho_\text{f} = A_\text{bf}\rho_\text{i}A_\text{bf}^\dagger = \sigma_x\rho_\text{i}\sigma_x. \tag{2.35}$$

A similar looking error type is the *phase-flip*, which is a rotation around the z axis: $A_\text{pf} = \sigma_z$.

Another kind of error is decay of probability from $|-z\rangle$ to $|+z\rangle$. This is *very* common because most physical qubits use an energetically excited state as $|-z\rangle$, and nature likes to absorb that energy and put it somewhere else. Formally called *amplitude damping*, we will sometime call it a *loss error*, or a T_1 process, which is what we call the time scale on which this decay happens. It looks a little different: $A_{T_1} = \sigma_-$. It's different because σ_- isn't unitary, so $\sigma_-\rho\sigma_-^\dagger = \sigma_-\rho\sigma_+$ doesn't preserve the trace of ρ. This operation yields zero on the 0 state, because it can't decay anywhere. This is a bit odd, and leads to some interesting results; namely, the absence of the error also has an effect on the state. Let's look into that, with math, because it's necessary for understanding properties of error-correcting encodings used in this thesis.

2.3.2   Describing errors with Kraus maps

In Equation 2.35, we said the action of the bit flip error was $\rho_\text{i} = \sigma_x\rho_\text{f}\sigma_x$. Of course, this must be something which happens only some of the time, with probability p. If the error always happened, it wouldn't really be an error, it would just relabel the states. So we also need to describe the case where the error doesn't happen. This type of stochastic process will result in a mixed state, so it can't be described with a unitary operator. We use instead a more general process, called the Kraus map, which is application of several operators with various weights. The set of Kraus operators

(sometimes "jump operators") is labeled $\{E_k\}$, and their action effects a process

$$\rho_f = \sum_k E_k \rho_i E_k^\dagger. \tag{2.36}$$

Since the diagonals of the density matrix represent probabilities for each basis state, their sum, $\text{Tr}(\rho)$, should always be equal to one[7]. The condition for trace-preservation is $\sum_k E_k^\dagger E_k = \hat{\mathbb{I}}$.

Equation 2.36 can clearly take a pure state into a mixed state. But what are the Kraus operators, and what is the resulting mixed state? If the error occurs with probability p, then this number should appear on the error term in the sum. In the notation of the operators in Subsection 2.3.1, the operator denoting the error should be something like $E_k = \sqrt{p}\, A_k$. Let's examine this for a few of the error channels described above.

Bit-flip

For bit-flip errors we have $E_1 = \sqrt{p}\,\sigma_x$. For trace-preservation, $E_0 = \sqrt{1-p}\,\hat{\mathbb{I}}$, since $\sigma_x^2 = \hat{\mathbb{I}}$. The clear interpretation of this Kraus operator is "with probability $1-p$, nothing happens." As an example, starting from the pure state $\rho_i = |0\rangle\langle 0|$, the error channel results in

$$\rho_f = (1-p)\,|0\rangle\langle 0| + p\,|1\rangle\langle 1|, \tag{2.37}$$

which is mixed to the extent that the error probability p is close to one-half. As we mentioned above, $p = 1$ is not really an error at all, and in this case the purity of the state is preserved.

---

7. This is not always true. When the qubit lives in a larger Hilbert space, and there is probability to go outside the computational subspace into the rest of the space, this can reduce the trace. This is called code-space leakage. Physics is still probability-conserving. This is just a side-effect of ignoring the rest of Hilbert space.

Amplitude damping

Now for something a bit more interesting. As we said, the amplitude damping operator is not unitary, which means $E_1^\dagger E_1$ is not proportional to $\hat{\mathbb{I}}$. This means that the rest of the Kraus operators will similarly not square to something proportional to $\hat{\mathbb{I}}$. The surprising conclusion is that the "no-error" case E_0 cannot be written as $\sqrt{1-p}\,\hat{\mathbb{I}}$ as it was for the bit-flip error. So "no-error" has a backaction action on the state, which is in general not unitary, and therefore not completely reversible. This will be developed for a more extensive case when we talk about multi-level systems in Chapter 3, but for a qubit it is straightforward.

We start with $E_1 = \sqrt{p}\,\sigma_-$. Since $E_1^\dagger E_1 = p\,|1\rangle\langle1|$, this implies $E_0 = |0\rangle\langle0| - \sqrt{1-p}\,|1\rangle\langle1|$. This is indeed not proportional to the identity. It has the effect of reducing the population in the 1 state, even when the decay error didn't happen. We interpret this as a sort of Bayesian update — if an excitation wasn't lost, it means that is is more likely that there was no excitation to begin with, and the population in 1 decreases continuously with p. Earlier we discussed mixed states as something which result from interactions with the environment. In this case, the environment likes to eat the excitation. If it sits there trying to take bites of the state and comes up hungry, it will eventually conclude that there was no excitation to eat in the first place. This no-error backaction results in an effective dephasing of the state, to be discussed more in Chapter 3.

2.3.3   Discrete versus continuous errors

Here we make some final words of conclusion about quantum errors. It is a bit strange that we began the topic of errors by claiming that qubits are continuously sensitive to noise, but then formally introduced errors as discrete events which happen with some probability. There is a dual to the Kraus map, a continuous description using the Lindblad master equation (LME), the extension of the Schroödinger equation to

density matrices.

The LME describes continuous evolution under a Hamiltonian $\mathcal{H}$, which gives unitary dynamics, and a Lindbladian L, which yields non-unitary evolution[8]. The LME reads

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -i[\mathcal{H}, \rho] + \gamma L\rho L^\dagger - \frac{1}{2}\rho L^\dagger L - \frac{1}{2}L^\dagger L\rho. \tag{2.38}$$

In this formalism, the coefficient γ denote the rates at which the error occurs. The Lindbladian is related to the Kraus operators — integrating the LME yields a sum of terms which look like powers of L^k. Specifically, the term $\gamma L\rho L^\dagger$ in the LME gives rise to the sum of terms $E_k\rho E_k^\dagger$ in the Kraus map (Equation 2.36). This sum terminates for qubits, but in Chapter 3 we'll see an infinite series for amplitude damping on oscillators. The error probability after a time t is something like $p = 1 - e^{-\gamma t}$, which completes the correspondence between the discrete and continuous pictures.

These two descriptions predict the same density matrix, so you may wonder why we bother with the Kraus map, and indeed, why it is even valid to treat continuous error processes as discrete ones. The real use of this approach will become clear in Section 2.4, where we will have the ability to measure whether errors have occurred. Consider the bit-flip channel. If the underlying process is some small noisy angle $\approx \theta$ of rotation about the x axis, then the Bloch vector will never really flip entirely. However, if we have the ability to projectively measure whether a flip has occurred, without measuring the state itself, then with a probability $p \sim \sin^2\theta$, we will find a bit flip. In a sense, a correctly tailored measurement forces one of the Kraus operators E_k to be the one that occurred. When detecting errors in this way, the Kraus map is the natural formulation of the problem because it closely corresponds to the resultant measurement records. As we mentioned earlier, we are only going to concern ourselves with strong, projective measurements. The dynamics when the strengths of the error

8. There can be multiple Lindbladians, but for simplicity we consider only one type of error here

and measurement processes are similar is beyond the scope of this thesis.

2.4   Error correction and logical qubits

In the previous section we introduced a few types of quantum errors, and developed a formalism, the Kraus map, for treating the effect of errors on a single qubit. But what is the effect of errors on a *computation*, and what do we do about it?

Loosely speaking, any quantum algorithm will involve preparing an input state, applying some (large) number of single- and multi-qubit gates, and then measuring some subset of the qubits to extract an answer. But if an error occurs on any of the qubits during the algorithm, it will probably give the wrong answer at the end. If the probability of all the errors anywhere in the system adds to more than one-half, then the algorithm will give the wrong answer *most* of the time. Given that many interesting quantum algorithms which promise an advantage over classical ones require many (dozens–hundreds) qubits, and even more (hundreds–thousands) gates, the probabilities of individual errors need to be extremely small to not build up with increasing qubit- and gate-count. Clearly, we need a way to recover from errors so that a single one doesn't completely derail the computation.

In this section, we will begin with a brief discussion of classical error correction, then extend these ideas to the simplest quantum error correction code, the bit-flip code. This is a prototypical example of a *logical qubit*, one made from multiple physical qubits. We will describe the concept of a syndrome measurement, and explain how it can be used to improve the performance of a logical qubit over physical ones. This will bring us to the crucial concepts of overhead and fault-tolerance in error correction, and some of the challenges therein.

2.4.1 Classical error detection and correction

As we mentioned in Section 2.3, there are a few types of errors in classical computers. The simplest is the bit-flip, which we continue to use as an example here. It's somewhat simpler to describe errors which occur in memory, but the ideas also can be extended to errors during operations. Classical RAM, for instance, stores a number as a string of bits. For example, $6 = (110)$. But if something causes the most significant (leftmost) bit to flip, then this string will read $(010) = 2$, and any proceeding computation using this data will fail. A common way to deal with this in early computers, before hardware was as reliable as it is today, was to add an extra bit, which was set to 0 or 1, depending on whether the parity (the number of ones) of the data string was even or odd, respectively. This is called a parity-check bit, and in this example, we represent the number $6 = (\underline{0}110)$, with the parity bit underlined. Now, before any computation, we check that the parity bit agrees with the parity of the data. If it does not, something terrible has happened, and the computation usually aborts to avoid further corruption of data or an erroneous result.

The parity-check bit is an example of a simple error-detection protocol. It is limited in scope, as it gives no way of knowing which bit flipped, and so there is no way to correct the error. Additionally, it only works if an odd number of bits have flipped. In many applications, this is not sufficient. We may want our computer to detect errors and then *correct* them, rather than just stalling. In order to do this, we need to know which bit flipped so we can flip it back. One way to do this is to add redundancy. By storing the same information in a larger number of bits, the information is delocalized, and local errors cannot completely corrupt it. A simple example is a repetition code, where we use three (or any odd number) of physical bits per logical piece of information — for example, $6 = ([111][111][000])$. Within each subgroup of three, each physical bit is initialized to the same value. This way, if the middle bit flips, we have $6 \stackrel{?}{=} ([111][101][000])$. By checking each subgroup to

see if all physical bits are the same, we find an error in the middle subgroup. Two of the three bits read 1, so we assume this is the correct value, and we reset the middle bit to restore the correct string. While this encoding still only allows us to detect a single error per subgroup, it has the advantage over the parity-check bit that we can actually fix the error. The cost is tripling the size of the data string. Through this section, we will quantify the advantages of this kind of scheme, as well as the cost.

2.4.2 Quantum bit-flip code

The quantum analogue of the classical bit-flip code looks very similar on the face of it. However, there are a few subtleties which have to do with the properties of quantum measurement.

Encoding

Figure 2.5a shows the structure of the three-qubit bit-flip code. We begin with a quantum state $|\psi\rangle$ in the first data qubit (D1). Classically, we would measure D1, and then set each of D2 and D3 equal to this value, but this would collapse an initial superposition in D1. Instead, we must use a unitary procedure which flips D2 and D3 to 1 if D1 is in the 1 state. This is the CNOT gate described in Subsection 2.2.4. This series of two CNOT gate prepares the entangled state

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \rightarrow |\psi_{\mathrm{L}}\rangle = \alpha\,|000\rangle + \beta\,|111\rangle\,. \tag{2.39}$$

We use the notation $|\psi_{\mathrm{L}}\rangle$ to denote a *logical qubit* state — one that is not prepared in a two-dimensional Hilbert space, but in some larger (in this case, $2^3 = 8$-dimensional[9]) one. The subspace spanned by $|\pm z_{\mathrm{L}}\rangle = \{|000\rangle, |111\rangle\}$ is called the code space, and $|\pm z_{\mathrm{L}}\rangle$ the code words.

---

9. More generally, the dimension is 2^n for n physical qubits

Error syndrome measurement

How do we check for errors in this encoding? We cannot in general simply measure each qubit and compare, as we did in the classical case. If this is the end of the algorithm and we simply want to learn the value of the logical qubit, then it is sufficient to measure and majority vote in this way. But this collapses the superposition, so if we are continuing the algorithm, this is no-go. So we need another way to extract information from the system, while preserving the superposition. We are going to ask the question: "which error occurred?" There are four possible outcomes: "no error," "D1 flipped," "D2 flipped," and "D3 flipped," corresponding to error operators $\{\hat{\mathbb{I}}, \sigma_x^{(1)}, \sigma_x^{(2)}, \sigma_x^{(3)}\}$. If we are to distinguish these four cases, we need to measure two bits of classical information. To do this, we add two extra *ancilla* qubits, A1 and A2, to the mix.

Each ancilla will interact with two data qubits, as shown in Figure 2.5b,c, and then be measured. This measurement will tell us "did either of these two qubits flip?" without telling us which one flipped, or what its state is. When no errors have occurred, the qubit states are the same; but if an error has happened, they will differ. We apply a CNOT gate with D1 as control and A1 as target, then another between D2 and A1. Each will flip A1 if the data qubit is in the 1 state. If they are both 0, A1 will not flip; if both are 1, A1 will flip twice, returning to 0. If D1 and D2 are different, A1 will flip once and be measured in 1. This is a kind of parity measurement, just like the classical parity check bit. As in that case, it doesn't tell us which bit flipped, just that an error occurred. The same process is repeated, or carried out simultaneously, with D2, D3, and A2. The two measured bits are the *error syndrome*, and together they tell us which of the data bits flipped, if any. The mapping between error operators and error syndromes is

$$\hat{\mathbb{I}} \longleftrightarrow 00 \qquad \sigma_x^1 \longleftrightarrow 10 \qquad \sigma_x^2 \longleftrightarrow 11 \qquad \sigma_x^3 \longleftrightarrow 01 \qquad (2.40)$$
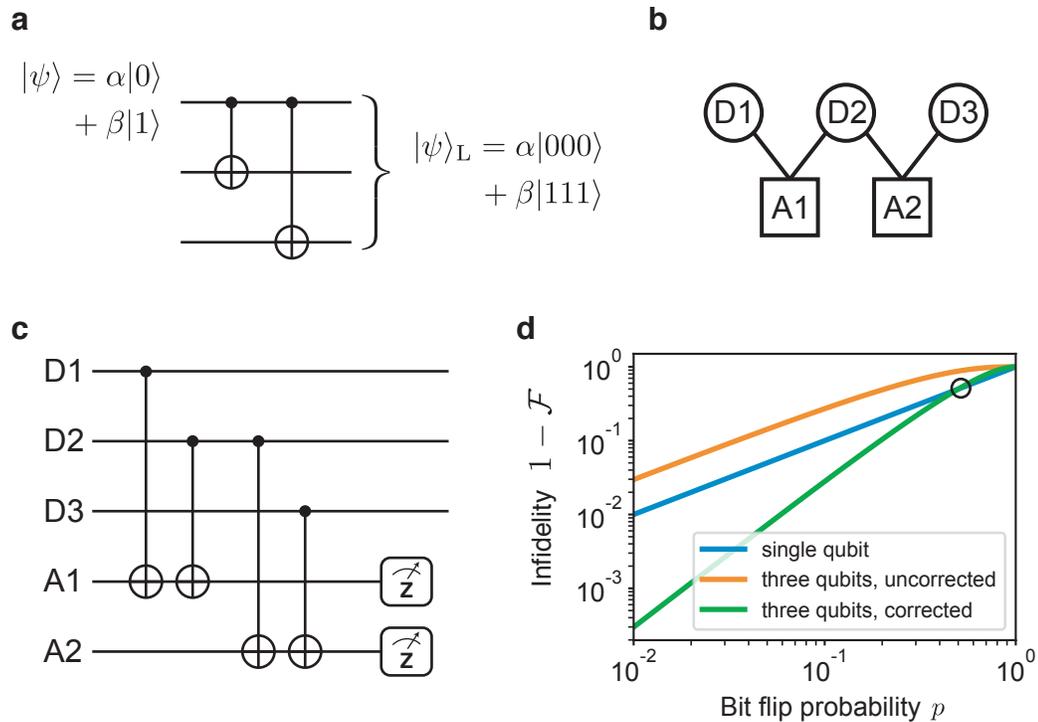
Figure 2.5:   Three qubit Bit-flip Code.   a) Encoding operation initializes logical qubit into code space. b) Physical connectivity for ancilla gates for syndrome measurements. c) Gates and measurements for error syndrome measurements. d) Infidelity versus single-qubit error probability. Performance for single qubit, three qubits without syndrome measurement, and three qubits with perfect error detection. Break even point $p=0.5$, where corrected fidelity exceeds single-qubit fidelity, is circled.

Now we can unambiguously determine which qubit flipped *without collapsing the superposition*, and apply a corrective X gate to restore it.

In many cases, knowledge of the error is sufficient, and it is not strictly necessary to correct it. In this example, if we found a bit-flip occurred on D1, we could correct it. However, we could just let it be, allowing the code space to be deformed, with new code words $\{|100\rangle, |011\rangle\}$. These two states are just as distant from one another (in sense to be defined in Subsection 2.4.3) as the original code words, so this is just as good a code to use. We simply need to update all our future operations, as well as our interpretation of the results at the end of the experiment. We call this *error tracking*. If we repeat the syndrome measurement several times and catch a few errors, the code space may deform several times.

2.4.3 Conditions for error correction

We can formalize and generalize the example above to give a set of conditions for a "good" error-correction code. There is a set of necessary and sufficient requirements, called the Knill-Laflamme (K-L) conditions [31]. Given set of Kraus error operators $\{E_k\}$, logical codewords $|\pm z_\mathrm{L}\rangle$ allow recovery if and only if

$$
\begin{aligned}
\langle +z_\mathrm{L}|\, E_k^\dagger E_j \,|-z_\mathrm{L}\rangle = 0 \\
\langle +z_\mathrm{L}|\, E_k^\dagger E_j \,|+z_\mathrm{L}\rangle = \langle -z_\mathrm{L}|\, E_k^\dagger E_j \,|-z_\mathrm{L}\rangle.
\end{aligned}
\tag{2.41}
$$

The first of these says that no two errors can map different logical codewords onto overlapping states. For $k \neq j$, this ensures that we can unambiguously distinguish different errors. For $k = j$, it states that error k maintains orthogonality of the codewords, thus preserving the information content. The second K-L condition states that the probability of a given error occurring is the same for both code words. This ensures that the error syndrome itself contains no information about the state. This is part of what makes amplitude damping a difficult error to correct. In Chapter 3 we'll look at a *bad* error-correcting code which does not satisfy this second condition.

2.4.4 Error correction performance

It is essential to be able to characterize how well an error-correction code performs. For instance, even in ideal conditions the bit flip code only allows approximate recovery from errors. This is because the code only corrects a single error, but it is possible that two of the three bits can flip. This is represented by additional Kraus operators (e.g. $\sigma_x \otimes \sigma_x \otimes \hat{\mathbb{I}}$) for which the K-L conditions are not satisfied. This kind of error happens with a probability $\sim p^2$, so it is less likely than a single bit flip, so long as p is small. We can hope to enhance the overall quality by detecting and correcting the dominant errors. We'll continue to use this code as an example; in Chapter 3 we

will consider other approaches.

Mean state fidelity

There are many metrics for quantifying how well a bit of quantum information is preserved as it passes through an error channel, but the one used throughout this work will be the mean state fidelity. The Uhlmann fidelity [32] of one quantum state to another is

$$\mathcal{F}(\rho_1, \rho_2) = \text{Tr} \left(\sqrt{\sqrt{\rho_1} \rho_2 \sqrt{\rho_1}} \right)^2 = \text{Tr} \left(\sqrt{\sqrt{\rho_2} \rho_1 \sqrt{\rho_2}} \right)^2 \tag{2.42}$$

which for pure states reduces to

$$\mathcal{F}(|\psi_1\rangle, |\psi_1\rangle) = |\langle\psi_1|\psi_2\rangle|^2 \tag{2.43}$$

We use this as a way to assess how close to an ideal state $|\psi_1\rangle$ a measured state $|\psi_2\rangle$ is. We can imagine measuring "is the qubit in $|\psi_1\rangle$?" and the fidelity is the probability with which we find the answer to be "yes." Some sources like [26] use the square root of this definition, but I prefer the physical interpretation as a probability of ending up in the right state. The square root definition is also always a higher number, and so sometimes looks overly optimistic.

So far this is just for a given pair of states. But to understand an error channel, we can imagine running all possible input states through it and measuring the average fidelity to the input state. This is of course impractical, but we can choose a uniform and over-complete set of states as a proxy. For a representative set of N initial states $\{\rho_i^{(k)}\}$, one can compute the state fidelity for each final state $\{\rho_f^{(k)}\}$ and average:

$$\bar{\mathcal{F}} = \frac{1}{N} \sum_{k=1}^{N} \mathcal{F} \left(\rho_i^{(k)}, \rho_f^{(k)} \right) \tag{2.44}$$

which we call the mean state fidelity. We typically use as input states the six cardinal points on the Bloch sphere: $\{|\pm z\rangle, |\pm x\rangle, |\pm y\rangle\}$.

Code overhead

With a metric in hand that we can easily calculate or measure, we can talk about the quality of an error channel. There are three relevant cases we should consider. The first is the fidelity as a function of error probability p for the simplest encoding we can make use of. In the bit-flip error channel, this is a single qubit. The second case is the fully error-corrected logical encoding, with all syndrome measurements. This is the encoding we hope will work best. Finally, it is instructive to consider the performance of the logical encoding *without* the syndrome measurement or correction. More qubits exposes us to more errors, so this uncorrected fidelity is always worse than the single qubit, but it's useful to see how far we've come by adding the measurements and correction. For our bit-flip example, the infidelity for these three cases is plotted in Figure 2.5d, calculated as described below.

As discussed in Subsection 2.4.2, the bit-flip code corrects at most one error. But we should consider how likely these errors are, and how likely multiple errors are. For one qubit, the error probability is p. When this error occurs, it inverts the qubit state. With probability $1-p$, no error occurs. Simple? Yes, but the last subtlety is that not all states are affected equally by the error. We need to consider the infidelity resulting from the error, which is state-dependent (hence why we average over all input states). The states $|\pm z\rangle, |\pm y\rangle$ have zero fidelity to themselves after a bit flip, but $|\pm x\rangle$ are unaffected up to a global phase. So the resulting *average* infidelity $1-\mathcal{F}_{\text{single}} = \frac{2}{3}p$, because only two-thirds of the states are corrupted. Note the difference from the classical bit flip, which always corrupts the bit completely.

When we add more qubits, the probability of errors increases, and we have the possibility of up to three errors. The probability of n errors p_n is given, as well as the

asymptotic probability for small p:

$$p_0 = (1-p)^3 \sim 1 - 3p \qquad p_1 = 3p(1-p)^2 \sim 3p$$
$$p_2 = 3p^2(1-p) \sim 3p^2 \qquad p_3 = p^3 \sim p^3 \tag{2.45}$$

The factor of 3 that appears on p_1 is combinatorial in nature, because there are three possible ways one error can occur. The same goes for p_2. In general, the probability for n errors on k qubits is k choose n.

It is clear from Equation 2.45 that the probability of a single error $p_1 \approx 3p$ is about three times higher than in the single-qubit case. We refer to this increase as the *code overhead*. It is the price we pay for redundancy — introducing a larger probability of errors. But wait, there's more. With no error detection, one or two errors completely corrupts *any* state — that is, the uncorrected fidelity given n errors $\mathcal{F}_{n,\text{uncorr}} = 0$ for $n = 1, 2$. This is also worse than the single-qubit encoding, since one or two errors takes us out of the code space, with no symmetry points left invariant under error. All told, the uncorrected infidelity is

$$1 - \mathcal{F}_{\text{uncorr}} = \sum_{n=0}^{3} p_n \times (1 - \mathcal{F}_{n,\text{uncorr}})$$
$$= \underbrace{(1-p)^3 \times 0}_{\text{no error}} + \underbrace{3p(1-p)^2 \times 1}_{\text{one error}} + \underbrace{3p^2(1-p) \times 1}_{\text{two errors}} + \underbrace{p^3 \times \frac{2}{3}}_{\text{three errors}} \tag{2.46}$$
$$= 3p - 3p^2 + \frac{2}{3}p^3.$$

Note that $1 - \mathcal{F}_{3,\text{uncorr}} = 2/3$. This has the same origin as the same factor of $2/3$ in the single-qubit case, that $|\pm x_\text{L}\rangle$ are invariant under the error. For small p, $1 - \mathcal{F}_{\text{uncorr}} \approx 3p$ is a constant factor of $9/2$ times worse than the single-qubit infidelity.

This all sounds terrible, but we can save the day with the error syndrome measurement. When we introduce the syndrome measurement, assuming it is perfect, we completely eliminate the infidelity due to single errors, and we're left with an

uncorrectable error probability $p_2 + p_3 = 3p^2(1-p) + p^3 = 3p^2 - 2p^3$, which is much reduced. Accounting for both the error probabilities and infidelities as above, the resulting infidelity is

$$
\begin{aligned}
1 - \mathcal{F}_{\text{corr}} &= \sum_{n=0}^{3} p_n \times (1 - \mathcal{F}_{n,\text{corr}}) \\
&= \underbrace{(1-p)^3 \times 0}_{\text{no error}} + \underbrace{3p(1-p)^2 \times 0}_{\text{one error}} + \underbrace{3p^2(1-p) \times \frac{2}{3}}_{\text{two errors}} + \underbrace{p^3 \times \frac{2}{3}}_{\text{three errors}} \qquad (2.47) \\
&= 2p^2 - \frac{4}{3}p^3.
\end{aligned}
$$

Here, $\mathcal{F}_{2,\text{corr}} = 2/3$ because, after restoration to the code space, $|\pm x_{\text{L}}\rangle$ will again be invariant.

Error suppression and break-even

For small error probability, the corrected infidelity $1 - \mathcal{F}_{\text{uncorr}}$ is a factor of $3p$ smaller than the trivial single-qubit encoding fidelity $1 - \mathcal{F}_{\text{single}}$. This is the whole point — by introducing a larger Hilbert space with which we can check for errors, we add errors but reduce their overall impact. The remarkable fact is that the gain is not a constant factor, but rather a scaling advantage. The corrected infidelity is only second-order in the error rate p. This error-suppression means that, on a logarithmic scale, the advantage of error correction increases as the error is made smaller, as we see clearly in Figure 2.5d. We call this a first-order error-correction code, because it suppresses the infidelity by one order of p.

The consequence of this scaling suppression is that there is a critical error rate at which we begin to see the advantage. This is where the lines corresponding to the single-qubit and corrected three-qubit infidelities cross in Figure 2.5d, at $p = 0.5$. We interpret this as the point where the error suppression balances out the overhead, and it is called the *break-even point*. The goal is to work reasonably below this point to

reap the advantages of our hard work.

2.4.5 Continuing to win

What if we go through all this trouble and find the resulting error rate is better, but still not good enough to reliably carry out our computation? We could use a higher-order code which suppresses the infidelity from p to order p^n, for some $n > 2$, but in practical cases this may have diminishing returns. For instance, we may find that some other uncorrected error, like phase flips, becomes dominant at this point. What then?

One approach is to use concatenation — to use the resultant logical qubit as the underlying "physical" qubit in another sort of repetition code. In this simple case, this would involve using three groups of three qubits to form a new, second-level, logical qubit. In some cases, the second order error on the physical qubit can manifest as an error of a different type, so the concatenation may not be uniform. Another approach is to use a larger code. For instance, a five-qubit bit-flip code can correct *two* single-qubit errors, and is therefore a second-order error-correction code. In general one can build the best practical low-order code, characterize the errors of the logical qubit, and then choose a second level encoding tailored to correct these errors, rinse, and repeat.

2.4.6 Fault tolerance

One critical consideration we have not discussed yet are imperfections in the error-correction procedure. There are many such errors that can occur, and by increasing the number of qubits from one to five (including the ancilla), we have exposed ourselves to a lot of them. We discussed double bit-flip errors, which are the minimum cost. But the scary thing is that errors in the ancillae can really hurt our logical information. There are benign errors, like incorrect measurements, which lead us to

think an error occurred, and introduce a bit flip by "correcting" this false syndrome. This is not so bad, because if we run the correction circuit repeatedly, we can catch this introduced error the next time around. More problematically, certain ancilla errors which occur during the interaction with the data qubits can corrupt the logical qubit. For instance, a bit flip on ancilla 1 between the two CNOT gates will introduce a phase error on the logical qubit, which is *uncorrectable*. This is horrific, and would require introducing an entirely new circuit to detect and correct phase errors, even if they don't occur on the data qubits to begin with.

This idea of error-propagation from ancilla to data is one of the things which make practical error correction extremely difficult. The fact that the additional complexity can not only amplify errors but introduce new kinds as well is related to the idea of *fault-tolerance* — adding complexity in a way which does not make things worse. This is a phrase with many meanings, but in this context it means "not adding enough new errors to nullify the quantum error-correction gain." In the context of concatenated error-correction codes, it means that the error rate will continue to decrease as we add more layers. For repetition code, it means that seven qubits is better than five is better than three. Remarkably, there are code-dependent *fault-tolerance thresholds* on memory, gate, and measurement errors such that, if everything operates below these thresholds, the logical error rate can be made arbitrarily low by adding more layers, or more qubits within the layers. Thus getting operations "below threshold" is a bit of a holy grail in quantum error correction.

2.5 Modular quantum information processors

Having laid out some of the basic ideas which we will be using throughout this thesis, we turn now to a difficult question to answer. It seems we will need a large number of independently-controlled and measured qubits to do anything of use. Thanks to the

inevitability of quantum errors, we may even need a lot more physical qubits than we have logical bits of computing power. So a question we can and should begin to ask is, how can we build such a machine where qubits are sufficiently reliable, independent, and connected?

These three requirements can to some extent be traded off on one another, and the balance will absolutely be platform-dependent. A quantum computer made of superconducting circuits may look very different, architecturally speaking, from one based on trapped ions or color centers. It is possible that the structures of competing technologies will on some level converge in the future, but differences will likely remain based on their relative strengths and challenges. Certainly in the near term these candidate platforms will have their similarities and their differences. But one core idea which is being pursued across the board is that of a modular quantum machine [33, 34, 35, 36].

The concept of a modular architecture for quantum computing means a lot of different things to different people, and is difficult to precisely define. But the overarching principle is the attempt to break this exceedingly complex problem into smaller pieces, or *modules*, which can then be wired together into a larger device. Modules which can be independently conceived of, designed, manufactured, tested, and assembled offer a massive simplification as compared to creating one monolithic component which is far more complex. Modularity in this sense may have additional benefits in the early days of quantum computers, when it is still difficult to make the modules in a reliable way. The ability to pre-screen or test modules before assembly, or replace ones which fail or under-perform, has been critical for prototyping devices for the experiments in this thesis.

It is unclear exactly what a modular quantum machine will look like. This thesis will not answer this question. I will simply provide a few concepts for structures which might be of use, and the rest of the thesis will be about providing some tools

for these architectures. Specifically, we will be concerned with the problem of quantum communication and networking between modules.

2.5.1 Basic requirements

To begin with, if we will have qubits in separate modules, there are a few things we will need:

– **Quantum storage**: At least some of the modules must have qubits where information can live, probably with some level of quantum error correction.

– **Quantum processing**: Some or all of the modules will have qubits on which gates are applied. This may be completely overlapping with the storage qubits, or totally disjoint, with the storage serving only as a memory bank. Error correction is probably needed here too.

– **Classical I/O**: Each module will need inputs from the classical control apparatus. These inputs will be needed to effect gates, and, as we will discuss, to actuate commutation with the rest of the computer in some way. Additionally, some or all modules will need input and output for measurement.

– **Quantum I/O**: In addition to classical data, the modules must have some way for quantum information to get in and/or out. This may take many forms, and we will discuss three possibilities in the rest of this section. The common feature is that some quantum link is needed to make the individual modules act like a unified whole. Broadly speaking, this relies on the ability to generate entanglement which spans the entire modular machine.

The different kinds of quantum I/O links we can have motivate a few different architectures. This is a non-specific and non-exhaustive list. The three architectures described here are not even mutually exclusive. Different approaches may have advantages at different scales, and we can imagine combining or concatenating them.
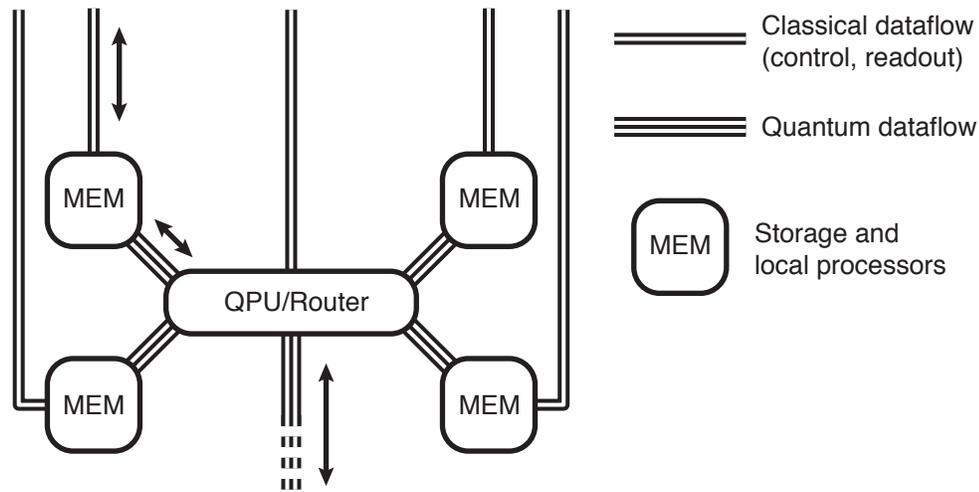
Figure 2.6: Shuttling-based Modular Architecture Multiple storage and processing elements are connected to a router via bi-directional channels which allow transfer of qubit states. Router can serve to shuttle qubits between storage banks, or to perform multi-qubit gates between qubits brought from disparate storage elements.

2.5.2 Shuttling architecture

The first architecture we'll describe, sketched in Figure 2.6, is perhaps conceptually simplest. We envision some number of memory modules containing multiple qubits, connected in a star- or tree-like manner about one or more routing elements. The general prescription is that qubits stay in the memory modules as much as possible, but move between modules for interactions. This is somewhat analogous to a classical cluster computer, but also looks loosely like a von Neumann architecture.

Links

In this paradigm, we imagine that the modules are connected to the router with links which allow qubits to be sent in both directions. In some platforms, like trapped ions or neutral atoms, it may be possible to physically move the qubits [37, 38]. For solid-state qubits, this is likely not viable, so we instead map the qubit onto the quantum state of light, which is sent over microwave transmission line, waveguide, or optical

fiber.

Role of the router

The router in this architecture may have different uses. It may be a kind of switch-yard which is mostly passive, just directing the qubits through. It may have way-stations where some error correction is performed to compensate for errors in transit. Or it may even be more of a processing unit, as indicated in Figure 2.6. One could imagine two qubits being brought in from different modules, interacting through a two-qubit gate, and then being sent back to their modules. It could even be the full analog of a von Neumann machine, where all operations happen in this quantum processing unit (or units), and the outer modules are just memory elements, like RAM.

One challenge in this approach is that the links and the router must be of extremely high quality that moving a qubit from one module to another does not significantly degrade its state. Of course there will be some errors, so we need them to be of a form and magnitude that can be handled by our error-correction scheme. As we will discuss in Section 2.6, some amount of dedicated error-correction on the transmitted qubits is likely important.

Place in this thesis

The kind of links needed for this shuttling architecture are discussed in Chapter 4 in theory, and in practice in Chapter 7. If the router can provide switchable non-reciprocity, or the module-link interface is of sufficient bandwidth, then the propagating photonic link in Chapter 6 could also be used to transport qubits in this architecture.

2.5.3 Entanglement-based architecture

A different kind of architecture may not require the ability to send arbitrary qubit states. It is sufficient to generate entanglement between modules and use that as a resource for all computation. This is known thanks to some pioneering work by Gottesman and Chuang [39].

Entanglement as a resource

There are two ways to compute without direct communication between modules, both using quantum teleportation and an entangled state between qubits in separate modules. It is possible to move the state of a third qubit from one module to another, using only *local* operations and *classical* commutation. An extension of this scheme, diagrammed in Figure 2.7, allows an arbitrary controlled unitary — a two-qubit gate — to be performed between two qubits which never interact directly. This means that if we can efficiently prepare entangled states which span modules, we never need to send quantum states through physical channels. We can teleport states between modules and perform local gates, or we can perform teleported gates directly. As pointed out in the original proposal, this scheme has advantages in terms of error propagation. Since the link is only classical, the extent to which errors can spread between modules is much reduced as compared to with direct two-qubit gates. This may reduce error propagating within or between error-correction codes, which makes it easier to compute fault tolerantly. Teleported gates have recently been demonstrated in cQED [40] and in a trapped ion system [41].

Entanglement through uni-directional channels

This entanglement-based architecture of course requires the ability to efficiently prepare many entangled states between modules. One approach is to prepare entanglement locally, and send half of the Bell pair through a quantum channel to another
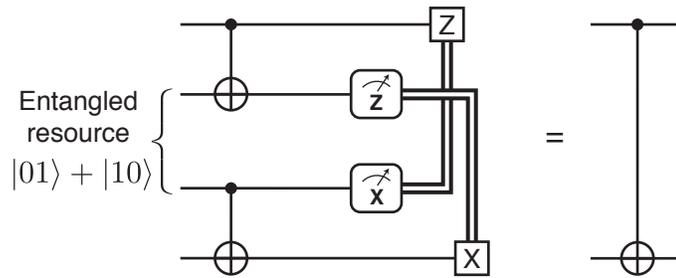
Figure 2.7: Teleported Gate. Top pair and bottom pair of qubits envisioned to be in separate modules. An inter-module resource Bell, intra-module gates, measurements, and classical feed-forward, effects a CNOT gate between modules.

module. This sounds like it requires the same kind of link as the shuttling-based architecture in Subsection 2.5.2, but it is a bit more flexible, since the interface only needs to allow transport in one direction. For instance, we can entangle a qubit in one module with light in the channel, then route and absorb the light in another module, mapping it onto a qubit and generating intermodule entanglement. This lends itself to an architecture with a bank of "sender" modules and a bank of "receiver" modules, with a router in between, as in Figure 2.8a. This distinction is not strictly necessary if the router is bidirectional.

Entanglement with interference and measurement

A complementary approach to preparing entanglement relies on the use of interference, and detectors which cannot distinguish the source of a photon [42, 43, 44]. The use of beamsplitters for interference lets us treat the two modules on equal footing, and is diagrammed in Figure 2.8b. Instead of entangling and sending from one to another, a communication qubit in each module is entangled with a "flying" qubit encoded in the state of light. The two flying qubits are routed to, then interfered on a beamsplitter to become mutually entangled, and then measured at a set of detectors. The act of measurement, if done correctly, projects the communication qubits into
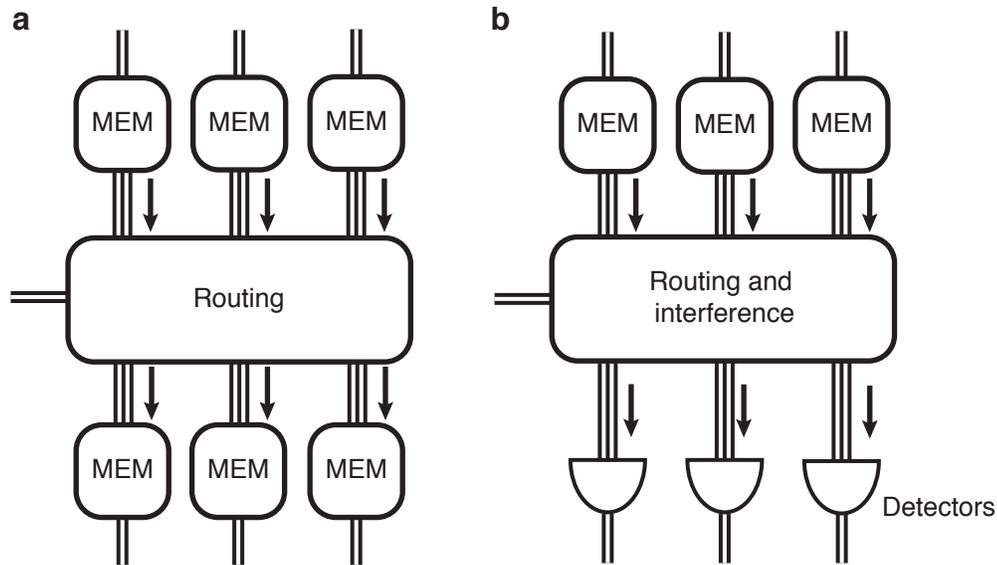
Figure 2.8: Entanglement-based Modular Architecture. a) Classically-controlled router allows uni-directional transfer of qubit states from top modules to bottom modules, and generation of resource entanglement. b) Router interferes emission from modules and directs resultant correlated light onto photon detectors. Detectors project qubits in modules into entangled states.

an entangled pair, which can then be used for teleportation.

This scheme is more general than I have described it. There are other ways of using flying light to effect a measurement which projects two qubits into an entanglement state which don't require interference with a beamsplitter, see for instance [45]. Alternatively, it is possible to interfere bosonic atoms instead of light, so this approach could be used with atomic qubits if the spatial control is of sufficient quality.

Entanglement purification

The entanglement prepared in this architecture will not be perfect, which would result in errors in the subsequent teleported operation. Fortunately, there are ways to improve the quality of entanglement by generating multiple entangled pairs between modules and performing local operations and measurements [46, 47, 48, 34]. This is known as *purification* or *distillation* of entanglement. This approach may be essential

to producing Bell pairs which enable teleportation below error-correction thresholds [49, 50]. Purification schemes require the ability to generate Bell pairs quickly enough that multiple pairs can be gathered together before they decohere, and has already been demonstrated in a platform where the generation is probabilistic [51].

Role of the router

In this design, the router is probably a passive element which directs and/or interferes light, but need not provide any measurements or operations itself. But the router/processor in the shuttling architecture (Subsection 2.5.2) could probably play the role of the router and detector bank, so it may be possible to use both approaches in a single platform.

Place in this thesis

The experiments in Chapter 6 and Chapter 7 both provide ways of preparing entangled states on demand, so either approach lends itself to this approach.

2.5.4   Gate-based architecture

We come now to the final architecture discussed here. This approach concerns itself not with the ability to send information between modules, but instead to perform gates directly between them. This is the least well-developed of the three, since direct gates between qubits in disparate modules have not been demonstrated yet. However, a two-qubit gate (e.g. CNOT, CPHASE, or SWAP-like) which can be implemented between modules with high fidelity might cut down drastically on the hardware overhead needed. It may not be necessary to devote resources to communication qubits like in the teleported scheme, or way-station qubits in the shuttling based scheme. This architecture also need not have a router if interactions are flexible or dense enough.
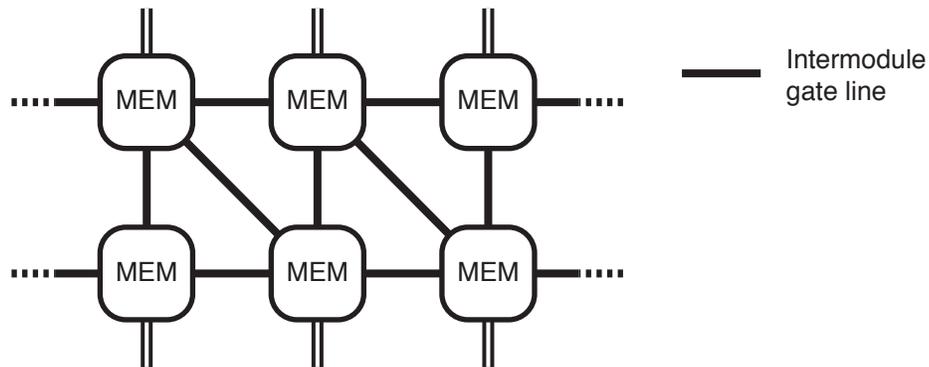
Figure 2.9: Gate-based Modular Architecture. Array of modules connected by lines which allow gates directly between qubits in disparate modules.

This approach, sketched in Figure 2.9, may lend itself well to topological error-correction codes, where qubits are usually arrayed in a two- or three-dimensional lattice with only nearest-neighbor interactions. The gates available between modules could allow for the lattice to extend between modules seamlessly[10].

Place in this thesis

Chapter 7 will demonstrate an approach for a link which should enable direct gates between modules in a few ways, though actually demonstrating this remains an outstanding goal.

2.6 Conclusion

In this chapter we described quantum bits and their essential features, then described some of the requirements we need to build a quantum information machine. We turned next to the discussion of what can go wrong, and talked a bit about how we can correct quantum errors. Finally, we considered a few prototypical architectures for

---

10. Since these inter-module gates might be of lower quality than the intra-module ones, the control apparatus would probably want to minimize the number of long-range gates.

building large quantum machines out of smaller modules, and how we might connect them.

But it is reasonable to expect that our modules might be of very high quality, and it is in braving the outside world, sending a qubit from one to another, that we might suffer the greatest chance of errors. This isn't necessarily going to be the case, but until we develop links which are as good as our memories and processors, we must be prepared to deal with errors in transit. Indeed this is one of the goals of this thesis.

So if we have communication errors, we probably need error correction here as much as we do anywhere else. But this presents a challenge. For instance, if we need to send a logical qubit to another module, we might think to decode it onto a single qubit and send that. But this would leave us completely exposed to errors in transmission with no recourse for correction. Instead, we could keep the qubits encoded in some repetition code, and send them one at a time through the link. This sounds decidedly inefficient. Another alternative would be to send them through multiple channels in parallel, but this requires a drastic increase in the connectivity between modules.

This brings us to a concept which will be a core topic in Chapter 3 — the "hardware-efficient logical qubit". This is an object which can be sent through a single link all in one go. If we are using photonic links, then we certainly need to send multiple photons together, because we need a large enough Hilbert space to satisfy the Knill-Laflamme conditions. So why don't we encode our information in a collection of photons to begin with? In the next chapter we'll describe exactly how to do that with bosonic qubits.

Chapter 3

Oscillators and Error Correction

In Chapter 2 we laid out some of the basic tenets of quantum information processing and error detection and correction, and discussed some ideas for modular architectures for quantum machines. We left that discussion on a final tension between the need to correct errors in quantum communication between modules, and the difficulties involved in sending a logical qubit through a communication channel. This motivated the idea of a "hardware-efficient" logical qubit for communication — some quantum resource which can hold redundant information in multiple excitations, but can be moved through a photonic link as a single object. In this chapter, we will describe such an object. In fact, it is the harmonic oscillator, one of the simplest textbook quantum systems.

In Section 3.1 we will describe the quantum LC resonator, the archetypal harmonic oscillator in cQED. We will cover useful bases, important classes of states, and the Wigner function, a phase-space visual representation that we will use throughout this thesis. Then in Section 3.2 we will describe the effect of dissipation on the oscillator in the Kraus formalism introduced in Chapter 2. Next Section 3.3 will turn to the discussion of error detection and correction in the harmonic oscillator. First we will consider some desirable qualities of an error-correction code; namely, the

ability to detect photon loss errors with a measurement of number parity. We will then introduce an encoding which has this quality, but does not make for a proper first-order error-detection code. Having understood what makes this code fail, we will move on to Section 3.4, where we will give the rest of our requirements for error-correction codes, and then introduce a few that we will use here, and compare their behavior and performance. Finally, we will give a quick summary of other bosonic codes for photon loss, and leave off with some thoughts about how we can use these encodings for robust quantum communication.

3.1 Harmonic oscillators in cQED

This section will serve as a very simple description of the quantum harmonic oscillator in the context of cQED. A much more detailed treatment can be found in [52], for example, which describes quantum circuits in general, including the Josephson junction element we will rely on in Chapter 5. Here we'll focus on introducing a useful basis in which to describe a few kinds of states of interest, as well as coupling to the environment through dissipation and drives. We'll make use of the Wigner function, which serves as a useful representation of the oscillator state in phase space, as it is one of the primary characterization tools used in experiments this thesis.

3.1.1 Ladder Operators and the Oscillator Hilbert Space

We begin with the simplest realization of a harmonic oscillator in cQED — the LC oscillator. The circuit, consisting of an inductance L and capacitance C in parallel, is shown in Figure 3.1. The quantum mechanical variables are the charge $\hat{q}$ on the capacitor and the flux $\hat{\phi}$ across the inductor. The flux is defined as the time integral of the voltage drop across the inductor, $\hat{\phi} = \int_{-\infty}^{t} \hat{v}(t')\mathrm{d}t'$. This is in analogy to the charge $\hat{q} = \int_{-\infty}^{t} \hat{i}(t')\mathrm{d}t'$, the integral of the current through the capacitor.
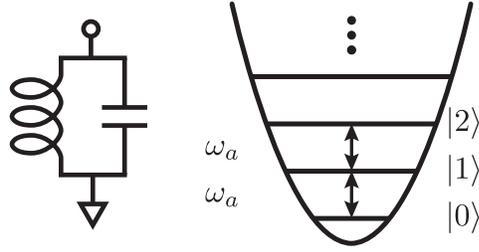
Figure 3.1: The LC Oscillator. The harmonic oscillator has an infinite number of levels, equally separated in energy by $\hbar\omega_a$.

These variables obey the canonical commutation relation $[\hat{\phi}, \hat{q}] = i\hbar$, and the dynamics are governed by the quadratic Hamiltonian

$$\mathcal{H} = \frac{\hat{q}^2}{2C} + \frac{\hat{\phi}^2}{2L}. \tag{3.1}$$

This is in direct analogy to a mechanical oscillator with mass m and spring constant k if one makes the replacement $\hat{\phi} \to \hat{x}$, $\hat{q} \to \hat{p}$, $C \to m$, $L \to k^{-1}$.

This is all well and good, but is not terribly well suited to the techniques available in cQED, as we rarely measure the charge or flux directly in this context. Instead, we adopt a basis for describing the state which will be naturally adapted to the kinds of experiments performed throughout this thesis. That will be the number, or Fock, basis, which uses as an index the number of excitations in the oscillator. This is made easy by defining the annihilation (a) and creation ($a^\dagger$) operators (collectively, ladder operators), which are so common throughout this thesis that we forgo the operator hats to reduce clutter. The ladder operators are defined such that[1]

$$
\begin{aligned}
\hat{\phi} &= \phi_a \left(a + a^\dagger\right) \\
\hat{q} &= \frac{1}{i} q_a \left(a - a^\dagger\right)
\end{aligned}
\qquad
\mathcal{H} = \frac{\hat{q}^2}{2C} + \frac{\hat{\phi}^2}{2L} = \omega_a \left(a^\dagger a + \frac{1}{2}\right),
\tag{3.2}
$$

---

1. As in Chapter 2, the Hamiltonian $\mathcal{H}$ has frequency units, which corresponds to $\hbar = 1$.

where the zero-point fluctuations $\phi_a = \sqrt{\hbar Z_a/2}$ and $q_a = \sqrt{\hbar/2Z_a}$, and the resonant frequency $\omega_a = \sqrt{1/LC}$ and characteristic impedance $Z_a = \sqrt{L/C}$ depend on the inductance and capacitance. In this light, it is straightforward to think of the operator a as proportional to the (integral of the) complex-valued voltage across the inductor and capacitor. The real part (in-phase, I) is related to the flux across the inductor, and the imaginary part (quadrature, Q) to the charge on the capacitor. This complex phase-space connection will be made more explicit in Subsection 3.1.2, and visualized in Subsection 3.1.3. As a final note, the ladder operators obey the commutation relation for bosons: $[a, a^\dagger] = 1$, as a direct consequence of the canonical commutation relation for $\hat{\phi}$ and $\hat{q}$.

3.1.2 Fock and coherent states

We said above that the ladder operators would make representing key kinds of states easy, so let's see what that looks like. The first class of states we will consider are the number, or Fock, states. They are eigenstates of the Hamiltonian in Equation 3.2, with eigenenergies $E_n = \omega_0(n + 1/2)$, for integer $n \geq 0$. We use the notation $|n\rangle$ for these states. They are not eigenstates of either ladder operator, but of the product: $a^\dagger a |n\rangle = \hat{n} |n\rangle = n |n\rangle$, where we have defined the number operator $\hat{n} = a^\dagger a$. The Fock states $\{|0\rangle, |1\rangle, |2\rangle, ...\}$ form an infinite orthonormal basis for the oscillator Hilbert space.

The second commonly encountered class of states is the coherent states, which are the eigenstates of the lowering operator: $a |\alpha\rangle = \alpha |\alpha\rangle$, for any complex number α. These are "classical" states, in that the expectation values of the charge and flux, defined in terms of the real and imaginary part of α, obey the classical equations of motion for an oscillator, and that the fluctuations on these quantities are the

minimum as allowed by the uncertainty principle. In the Fock basis,

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle . \tag{3.3}$$

The Fock state occupations of a coherent state (the square of the coefficients in Equation 3.3) are a Poisson distribution with mean $|\alpha|^2$, which is the expectation value of the photon number operator $a^\dagger a$. As such, the notation $|0\rangle$ is unambiguous, because the coherent state with zero amplitude is the number state 0, or vacuum state. For clarity, we use Greek symbols like α for coherent states and Roman letters like n to index Fock states. When there is ambiguity, we denote coherent states with a specific amplitude as $|\alpha = 1\rangle$, for example.

The coherent states cover the entire oscillator phase space, but are not exactly orthogonal, so they form an over-complete basis set, which isn't usually the most convenient representation except in special cases. The overlap of two coherent states is

$$|\langle \alpha | \beta \rangle|^2 = e^{-|\alpha - \beta|^2}, \tag{3.4}$$

which has the nice interpretation of being a function only of the distance between the two coherent states in complex phase space. This overlap rapidly drops as the distance increases, but it has important consequences for $|\alpha|, |\beta| \lesssim \sqrt{2}$.

Cat states

Before we move on to discussing creation and measurement of oscillator states, we describe one more class of states, which will be of interest for quantum error correction in Section 3.4 — superpositions of coherent states. The states are often called cat states, because they are superpositions of two "classical" states. An example is the

two-component cat state

$$\left| \mathcal{C}_\alpha^\pm \right\rangle = \frac{1}{\mathcal{N}_\alpha^\pm} \left(|\alpha\rangle \pm |-\alpha\rangle \right) \qquad \mathcal{N}_\alpha^\pm = \sqrt{2 \left(1 \pm e^{-2|\alpha|^2} \right)} \tag{3.5}$$

Because of the finite overlap of any two coherent states, the normalization constant $\mathcal{N}_\alpha^\pm \to \frac{1}{\sqrt{2}}$ as $\alpha \to \infty$ but $\mathcal{N}_\alpha^+ \neq \mathcal{N}_\alpha^-$ for any finite α. Due to the sign alternation in the amplitude of the Fock state components of $|-\alpha\rangle$ (Equation 3.3), the symmetric superposition $|\mathcal{C}_\alpha^+\rangle$ has only even photon number Fock components, and the antisymmetric superposition $|\mathcal{C}_\alpha^-\rangle$ has only odd. For this reason, we call these states "even cats" and "odd cats," respectively.

Since states with this property of even or odd number content will be central to our discussion of error correction of oscillators, we define the parity operator

$$\hat{P} = e^{i\pi a^\dagger a} = \sum_{n=0}^\infty (-1)^n |n\rangle\langle n| \tag{3.6}$$

Even states like $|\mathcal{C}_\alpha^+\rangle$ are eigenstates of $\hat{P}$ with eigenvalue $+1$, and odd states like $|\mathcal{C}_\alpha^+\rangle$ have eigenvalue -1. Likewise, Fock states $|n\rangle$ are also eigenstates of $\hat{P}$, with eigenvalue depending on the parity of n. Coherent states are in general not eigenstates of $\hat{P}$, except trivially for $\alpha = 0$.

3.1.3 Representing the oscillator state

It is useful to have a compact visual representation for oscillator states, akin to the Bloch sphere for qubits. For this we use the Wigner function, which will be a primary experimental characterization tool in Chapters 6, and 7, since it is a complete representation of the state, and is convenient to measure (see Chapter 5). The Wigner function for a state given by density matrix ρ is

$$W(\beta) = \frac{2}{\pi} \mathrm{Tr} \left(\hat{D}^\dagger(\beta) \rho \hat{D}(\beta) \hat{P} \right) \tag{3.7}$$
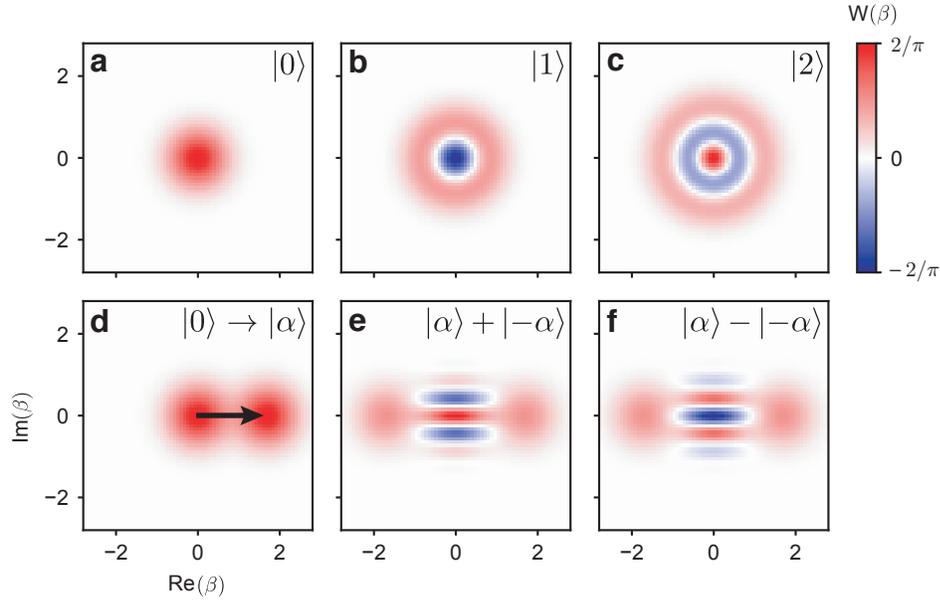
Figure 3.2:   The Wigner Function. a) The vacuum state is a symmetric Gaussian in phase space. b,c) Fock states $|1\rangle$, $|2\rangle$ are made of concentric rings. They carry no phase information, and so are rotationally symmetric. d) A coherent state with complex amplitude α is the vacuum state displaced by that amplitude. Here $\alpha = 1.7$. e,f) Two-legged cat states formed from symmetric (antisymmetric) superpositions of two coherent states of opposite phase have even (odd) parity.

This is understood as "the average parity of the state when displaced by β," and indeed it is measured by displacing the cavity and measuring $\langle \hat{P} \rangle$. The prefactor $2/\pi$ ensures that the Wigner function is normalized (its integral over all phase space is equal to one), even for mixed states. Knowledge of the Wigner function for all β is equivalent to knowledge of the entire density matrix, and an experimental measurement of $W(\beta)$ allows one to reconstruct ρ (see Chapter 5). Since the dimensionality of the oscillator density matrix is infinite, in practice one measures $W(\beta)$ for $|\beta|^2$ up to about $2\bar{n}$, where $\bar{n} = \langle a^\dagger a \rangle$ is the mean photon number of ρ.

The Wigner functions of a few common states are shown in Figure 3.2. Note in particular that $W(0)$ is simply the expectation value of the parity of the state (multiplied by $\frac{2}{\pi}$). For Fock states $|0\rangle$, $|2\rangle$ (generally, $|2n\rangle$) and the even cat, this center point is maximally positive (red), and for $|1\rangle$ (generally, $|2n + 1\rangle$) and the odd cat, it is maximally negative (blue). Figure 3.2d shows the coherent state $|\alpha = 1.7\rangle$,

which is just the vacuum state $|0\rangle$ shifted to a new center. The Wigner functions of the cat states show the two coherent states, but with characteristic fringes in between these two "legs," and opposite parity.

3.2 Dissipation as an error channel

So far we have considered some of the features of a harmonic oscillator in isolation, but any physical oscillator also includes some decay channel, where it can lose some energy into the environment. We describe dissipation in the same way we introduced bit flips and amplitude damping in Section 2.3 — with Kraus operators, and the Lindblad master equation.

3.2.1 The Amplitude Damping Kraus Map

In the continuous LME picture, the Lindbladian representing decay in an oscillator is the annihilation operator a, which removes one photon from the oscillator at a time. We define the rate at which this occurs as κ, and the LME in the oscillator frame reads

$$\dot{\rho} = \kappa a\rho a^\dagger - \frac{\kappa}{2}\rho a^\dagger a - \frac{\kappa}{2}a^\dagger a\rho \tag{3.8}$$

Consider the instantaneous differential evolution of this decay channel acting on a Fock state $\rho = |n\rangle\langle n|$. The first term reads $n\kappa \langle n|\rho|n\rangle |n{-}1\rangle\langle n{-}1|$ — it adds population to the Fock state $|n-1\rangle$, at a rate that grows linearly with n and is proportional to the occupation of $|n\rangle$. The last two terms, which maintain the normalization of the state, do not change the photon number, but reduce the amplitude of the n-th component. Together they read $-n\kappa \langle n|\rho|n\rangle |n\rangle\langle n|$.

The dual of the LME is the Kraus map, introduced in Subsection 2.3.2,

$$\rho_{\mathrm{f}} = \sum_k E_k \rho_{\mathrm{i}} E_k^\dagger \qquad \sum_k E_k^\dagger E_k = \hat{\mathbb{I}}, \tag{3.9}$$

which can help give a picture of what happens for some discrete time t, and makes clear the effect of loss of multiple photons. The Kraus operator E_k corresponding to loss of k photons is given in [53][2]:

$$E_k = \sqrt{\frac{(1-\eta)^k}{k!}} \eta^{\hat{n}/2} a^k \tag{3.10}$$

with $\eta = e^{-\kappa t}$, and $\hat{n} = a^\dagger a$ the number operator. As we will see, this operator has two effects: the part which reads a^k corresponds to applying the lowering operator k times, and $\eta^{\hat{n}/2} = \mathrm{diag}(\eta^{n/2})$ tends to shrink the occupation in higher number states, irrespective of which Kraus operator is applied. Note that this shrinkage term doesn't change which number states are occupied, just their relative weights.

These Kraus operators have a convenient action on a Fock state:

$$E_k \left| n \right\rangle = \sqrt{\binom{n}{k}} \sqrt{\eta^{n-k}(1-\eta)^k} \left| n - k \right\rangle$$

$$E_0 \left| n \right\rangle = \sqrt{\eta^n} \left| n \right\rangle \tag{3.11}$$

$$E_1 \left| n \right\rangle = \sqrt{n\eta^{n-1}(1-\eta)} \left| n - 1 \right\rangle$$

where we have given a few useful special cases. Finally, since the Kraus map outputs a mixture of such terms, the normalization of the above terms is the the probability of that Kraus operator being applied. For example, the probability of not losing any photons from $\left| n \right\rangle$ is $\langle n | E_0^\dagger E_0 | n \rangle = \eta^n$.

---

2. Alternative, equivalent definitions are found in [54, 55]

3.2.2   Damping on the Fock code

Let's unpack this with a simple example, considering the action on a logical qubit in the simplest encoding we might consider in an oscillator. This is the so-called Fock code, where $|\pm z_\mathrm{L}\rangle = \{|0\rangle, |1\rangle\}$. To understand the behavior of this code, we will need to know how the damping channel acts on various states in the code space. Once we have a handle on the basic behavior we will quantify the extent to which a logical qubit survives this channel, for which we will use the mean state fidelity introduced in Section 2.3.

Behavior of cardinal states

We begin with the local codeword $|-z_\mathrm{L}\rangle = |1\rangle$, but we need to work with the density matrix $\rho_\mathrm{i} = |1\rangle\langle 1|$. Since the code space occupies only the first two levels of the Hilbert space, we only need to consider the Kraus operators E_0 and E_1; all other error terms will be exactly zero on the logical qubit. The final state under the channel defined by Equation 3.10 is

$$
\begin{aligned}
\rho_\mathrm{f} &= E_0 \rho_\mathrm{i} E_0^\dagger + E_1 \rho_\mathrm{i} E_1^\dagger \\
&= e^{-\kappa t} |1\rangle\langle 1| + (1 - e^{-\kappa t}) |0\rangle\langle 0| \\
&= \eta |1\rangle\langle 1| + (1 - \eta) |0\rangle\langle 0|
\end{aligned}
\tag{3.12}
$$

This is sensible — the population in $|1\rangle$ decays exponentially in time, and the population in $|0\rangle$ grows. Moreover, we recognize η as the *energy efficiency* of this process. For an average number of photons $\bar{n}$ in the initial state, there will be on average $\eta\bar{n}$ photons in the final state. The efficiency is a useful notion, because we can just consider this action as a channel, which could be occurring in time or in space.

So the population will decay from a Fock state. Additionally, the action of the loss channel on the code word $|+z_\mathrm{L}\rangle = |0\rangle$ is identity, since we can't lose a photon from nothing. This story gets more interesting when we consider the action on the

superposition state $|\pm x_{\mathrm{L}}\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Now the Kraus map gives

$$
\begin{aligned}
\rho_{\mathrm{f}} &= E_0 \rho_{\mathrm{i}} E_0^\dagger + E_1 \rho_{\mathrm{i}} E_1^\dagger \\
&= \frac{1}{2} \big(|0\rangle + \sqrt{\eta}\,|1\rangle \big) \big(\langle 0| + \sqrt{\eta}\,\langle 1| \big) \\
&\quad + \frac{1-\eta}{2}\, |0\rangle\langle 0| \,.
\end{aligned}
\tag{3.13}
$$

The last line is the one-jump action, which always results in zero photons, with probability $(1-\eta)/2$. This is half the loss probability for one photon.

Less intuitively, the first line is the no-jump action, which results a pure state that is *not* the initial state. When no loss occurs, there is a continuous deformation of the superposition toward $|0\rangle$, even though the jump operator does not include a power of a. Strangely enough, it seems that the no-jump event is just as likely to remove a photon as the one-jump event is for this state. The qualitative explanation for this is that the environment is slowly measuring the state. If some time has gone on and no photon was absorbed by the environment, it becomes less likely that there was a photon in the oscillator to begin with. This has the effect of reducing the population in $|1\rangle$ as the state is slowly projected into $|0\rangle$. While the state remains pure, this is bad news for our logical qubit. For instance, $|\pm x_{\mathrm{L}}\rangle$ converge to the same state ($|0\rangle$), so the information is lost even when a photon is not absorbed.

Fidelity

We are now in a position to compute (and more importantly, explain) a performance metric for this encoding as a function of η. As in Subsection 2.4.2, we use the mean state fidelity to assess how well an arbitrary qubit state would make it through the channel. We apply the Kraus map to each of the six cardinal states $\{|\pm z\rangle, |\pm x\rangle, |\pm y\rangle\}$, compute their state fidelities to the ideal state individually, and then average. When looking at the results, it is best to plot the *infidelity* $1-\mathcal{F}$ versus the *inefficiency* $1-\eta$, since the scaling at low inefficiency will tell us what order the

code is sensitive to loss.

When we perform this analysis for the Fock encoding, we find that the mean infidelity scales linearly with the inefficiency; see Figure 3.3. In particular, we find $1 - \mathcal{F} \approx (1 - \eta)/3$ for small inefficiencies; this scaling is plotted as a guide to the eye. This factor of $1/3$ confused me for a long time. One might expect that the infidelity would be half the inefficiency, since the average photon number in this encoding is $\bar{n} = 0.5$. Indeed, the error *probability* is $(1 - \eta)/2$, but, just as in the bit flip code, we must also consider the resulting infidelity when an error occurs. In this case, since $E_1 \left| 1 \right\rangle \propto \left| 0 \right\rangle$, an error keeps us in the code space, and there is some overlap with the original state for most points on the Bloch sphere. The calculation requires calculating the probability of loss $p_{\text{loss}}^{(i)}$ for each state i, the fidelity $\mathcal{F}_{\text{loss}}^{(i)}$ when a photon is lost, and the fidelity $\mathcal{F}_{\text{noloss}}^{(i)}$ in the case the photon is not lost. To leading order, $\mathcal{F}_{\text{noloss}}^{(i)} = 1$ for all i. Averaging over the six cardinal states,

$$
\begin{aligned}
1 - \mathcal{F}_{\text{Fock}} &= \frac{1}{6} \sum_{i=1}^{6} \left[p_{\text{loss}}^{(i)} \times (1 - \mathcal{F}_{\text{loss}}^{(i)}) + (1 - p_{\text{loss}}^{(i)}) \times \underbrace{(1 - \mathcal{F}_{\text{noloss}}^{(i)})} \right] \\
&= \frac{1}{6} \sum_{i=1}^{6} p_{\text{loss}}^{(i)} \times (1 - \mathcal{F}_{\text{loss}}^{(i)}) \\
&= \frac{1}{6} \left[\underbrace{0}_{\left|0\right\rangle} + \underbrace{(1 - \eta) \times 1}_{\left|1\right\rangle} + 4 \underbrace{\left(\frac{1 - \eta}{2} \times \frac{1}{2} \right)}_{\left|\pm x\right\rangle, \left|\pm y\right\rangle} \right] \\
&= \frac{1}{6} \left[(1 - \eta) + \frac{4}{4}(1 - \eta) \right] \\
&= \frac{1}{3}(1 - \eta)
\end{aligned}
\tag{3.14}
$$

Roughly speaking, the factor of $1/3$ comes from the fact that two-thirds of the states have fidelity $1/2$ under loss. To actually predict this correctly, it is important to weight the infidelity by the loss probability for each state, as we have done.

Looking at Figure 3.3, we see that the infidelity deviates from this linear slope at
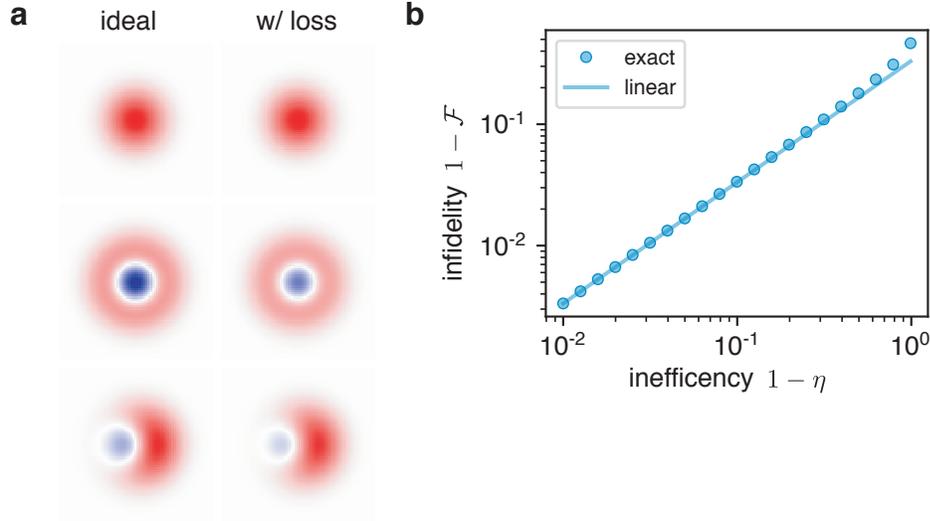
Figure 3.3: The Fock Code. a) Left: Wigner functions of $|+z_{\mathrm{L}}\rangle, |-z_{\mathrm{L}}\rangle, |+x_{\mathrm{L}}\rangle$ from top to bottom. Right: Same logical states with loss channel applied $\eta = 0.8$. b) Mean state infidelity versus inefficiency of channel, with linear approximation $(1 - \eta)/3$.

large losses. This is because I was sneaky in assuming $\mathcal{F}_{\mathrm{noloss}} = 1$ in Equation 3.14. This is decidedly not true; as we saw in Equation 3.13, the no-jump evolution does deform the equator states towards $|0\rangle$. This means there is an additional contribution to the infidelity from this effect, but it is second order in $1 - \eta$, so it only becomes apparent at large inefficiencies.

3.2.3 Damping on coherent states

The odd behavior of reduction of energy even in the case of no absorption of a photon by the environment (the action of E_0) is perhaps most pronounced for coherent states, which have the interesting property of being eigenstates of the annihilation operator: $a |\alpha\rangle = \alpha |\alpha\rangle$. As a result, the Kraus operator E_k has a fairly simple effect on a coherent state:

$$E_k |\alpha\rangle = \alpha^k \sqrt{\frac{(1 - \eta)^k}{k!}} |\sqrt{\eta}\alpha\rangle \tag{3.15}$$

Since the state vector in this expression is independent of k, the entire Kraus map acting on the coherent state remains pure; specifically, it is

$$|\alpha\rangle \rightarrow |\sqrt{\eta}\alpha\rangle \tag{3.16}$$

which decays continuously with η towards the vacuum state as η decreases towards zero. However, as we will see in Section 3.4, superpositions of coherent states become mixed.

3.3 Error detection with photons

Having defined the harmonic oscillator as it appears in a cQED context, we will begin to explore some of the ways it can be used as a hardware-efficient housing for a bosonic logical qubit.

Photons are one of the oldest candidates for qubits. In the optical domain, the idea of qubits stored in the polarization state of photons was the basis for thought experiments dating back to Einstein, Podolsky, and Rosen [56]. Moreover, proposals and realizations of entanglement and quantum information manipulation with photons have been around for more than three decades [57, 58]. These proposals often consider encoding information in multiple, spatially separate photonic modes (originally, optical fiber channels).

The story of *error-correction* with bosonic qubits began as an extension of those protocols. Chuang, Leung, and Yamamoto [54] presented the first multi-mode bosonic codes designed with amplitude damping in mind. Not long after, the first single-mode, multi-photon bosonic encoding, the Gottesman-Kitaev-Preskill (GKP) encoding [59], was envisioned as a way to protect information in an oscillator from noisy kicks from the environment. It turns out that many physical error channels can look like a small kick, so the GKP codes are an interesting candidate in a variety of platforms, only

recently realized in the motional mode of a trapped ion [60] and in a superconducting cQED cavity resonator [61].

In cQED in particular, much focus is on correcting the photon loss in quantum memories. While 3D cavities can have some of the lowest dissipation rates of any cQED qubit candidate [62, 63], photon loss is still the dominant intrinsic error. In fact, it was demonstrated that the inherent dephasing rate of a superconducting cavity resonator [64] can be more than an order a magnitude smaller than its dissipation rate, and probably even weaker than that. In the past few years, experimental demonstrations have reached the break-even point of error correction with codes designed for amplitude damping in cQED resonators [65, 66], with performance mainly limited by non-fault-tolerance. Namely, dephasing and other types of errors are introduced by the quantum degrees of freedom used to control and measure the resonators. Eliminating the propagation of errors from these ancillae into the memory is a recent topic of successful research in our lab [64, 67].

Apart from protecting quantum memories, bosonic error-correction has applications in quantum communication as well. The pioneering work of Chuang and co-authors [54] was motivated in part by the idea that using optical fiber for quantum communication would necessitate the ability to correct for losses over long distances. Development of multi-photon codes for this purpose has continued ever since [68, 69, 70].

This is part of the approach we are taking for developing ways to move qubits and generate entanglement in cQED modular architectures. Even over relatively short distances, it is reasonable to expect that links between modules might be a weak point where errors can creep in, as the information must leave the carefully sheltered environment inside the modules. So far, most demonstrations in this direction have been primarily limited by losses in transmission channels, routing elements, and interconnects [45, 71, 72, 16, 73, 74, 75]. As such, we seek to implement transfer of

encoded bosonic qubits between modules, with an error-detection step at the end to recover from losses in transmission. Even more powerfully, it has been pointed out that such schemes can be robust to thermal noise in the transmission channel as well [76, 77].

In this section, we describe the basic idea behind the types of bosonic codes we will use in this thesis, which involve code words with well-defined photon number parity. In Subsection 3.3.1 we explain how these kinds of codes automatically satisfy some of the Knill-Laflamme (K-L) conditions, at least approximately. Then in Subsection 3.3.2 we define a simple code with this structure, which seems at first blush to enable correction of losses, but instead only allows us to detect them. Motivated by this example, we will move on to Section 3.4, where we describe some codes which *do* allow for quantum error correction.

3.3.1 Detecting loss with parity measurements

Let's revisit the Kraus map for photon loss from Equation 3.10:

$$E_k = \sqrt{\frac{(1-\eta)^k}{k!}} \eta^{\hat{n}/2} a^k \tag{3.17}$$

If the efficiency of a channel is high $(1 - \eta \lesssim 0.1)$, the prefactor $\sqrt{(1-\eta)^k/k!}$ is suppressed as k increases. This means the dominant Kraus terms are E_0 and E_1 — no jump, and one photon lost. By considering only these two error terms, we will be developing a first-order error-correction code[3]. Up to second order in $1 - \eta$, all bets may be off.

In order to satisfy the first K-L error-correction criterion (Subsection 2.4.3), it is essential that the Kraus operators we are worried about take our two logical code

---

3. Of course the probabilities of the various Kraus operators being applied are state-dependent, so a more detailed analysis will be necessary to ensure that we have build a true first-order error-correction code, where the infidelity is proportional to $(1 - \eta)^2$.

words to orthogonal subspaces:

$$\langle +z_\mathrm{L}| \, E_0^\dagger E_1 \, |-z_\mathrm{L}\rangle = \langle +z_\mathrm{L}| \, E_1^\dagger E_0 \, |-z_\mathrm{L}\rangle = 0 \qquad (3.18)$$

One way to ensure this is to have the code states $|\pm z_\mathrm{L}\rangle$ both live in a subspace of well-defined photon number parity. Since the one-jump operator E_1 changes the photon number parity, and the no-jump operator E_0 does not, this automatically satisfies this half of the first K-L condition, and we have a fighting chance. This statement relies on the ability to detect single-photon loss errors, but using an encoding with well-defined parity makes this straightforward — we just need to measure the photon number parity in a QND way. If we find it has changed, we assume E_1 was applied. If not, we assume E_0. Of course, the next-order error E_2 will preserve the parity by removing two photons, which is an undetectable error, but we will ensure that this only happens to higher order in $1 - \eta$.

3.3.2   A bad encoding

Obviously the single-photon Fock encoding does not have this parity structure, and definitely does not satisfy Equation 3.18, since $E_1 |1\rangle = |0\rangle = E_0 |0\rangle$. What about using one more level of the oscillator: $|\pm z_\mathrm{L}\rangle = \{|0\rangle, |2\rangle\}$? We call this the "0–2" encoding. Both code words have even photon number parity, so Equation 3.18 is satisfied. However, consider applying the Kraus map and then measuring parity. If we find the parity has changed from even to odd, we know E_1 was applied. However, the probability of this operator occurring on the code word $|+z_\mathrm{L}\rangle = |0\rangle$ is exactly zero — we can't lose a photon if there were none to start with. This means that any time we measure odd parity, which happens with probability first-order in $1-\eta$, we will project the state onto $|1\rangle$, which completely erases the encoded information. This breakdown

is related to a failure to satisfy the second K-L condition: for this encoding

$$0 = \langle +z_\mathrm{L}| E_1^\dagger E_1 |+z_\mathrm{L}\rangle \neq \langle -z_\mathrm{L}| E_1^\dagger E_1 |-z_\mathrm{L}\rangle \propto 1 - \eta \tag{3.19}$$

Performance comparison

Indeed, a calculation of the performance of this code bears out this failure. As when we discussed the three-qubit bit flip code in Section 2.4, it is instructive to look at the performance of this code *without* an error syndrome measurement. In this case, one would naïvely assume the infidelity is equal to the inefficiency, since this encoding has on average one photon. Unlike the Fock encoding, this assumption is correct to first order, since loss of a photon takes us out of the code space, resulting in zero fidelity. This means the uncorrected 0–2 code, plotted in Figure 3.4, is three times worse than the 0–1 code.

This is a discussion of error correction, so what happens if we can perfectly detect and compensate for photon loss errors? It's useful to separate the results into the two syndrome outcomes — no jump (even parity), and one jump (odd). If there is no loss, we assess the fidelity to the original code space. In this case, we can to first order only consider the action of E_0, which in this space is

$$E_0 = |0\rangle\langle 0| + \eta |2\rangle\langle 2| \tag{3.20}$$

This keeps us in the code space, with a deformation on the equator states of the Bloch sphere. However, the infidelity induced by this deformation is only second-order in $1 - \eta$. This corresponds to exact satisfaction of the first K-L condition for the two dominant Kraus operators, and approximate satisfaction of the second

$$\langle +z_\mathrm{L}| E_1^\dagger E_1 |+z_\mathrm{L}\rangle \approx \langle -z_\mathrm{L}| E_1^\dagger E_1 |-z_\mathrm{L}\rangle, \tag{3.21}$$
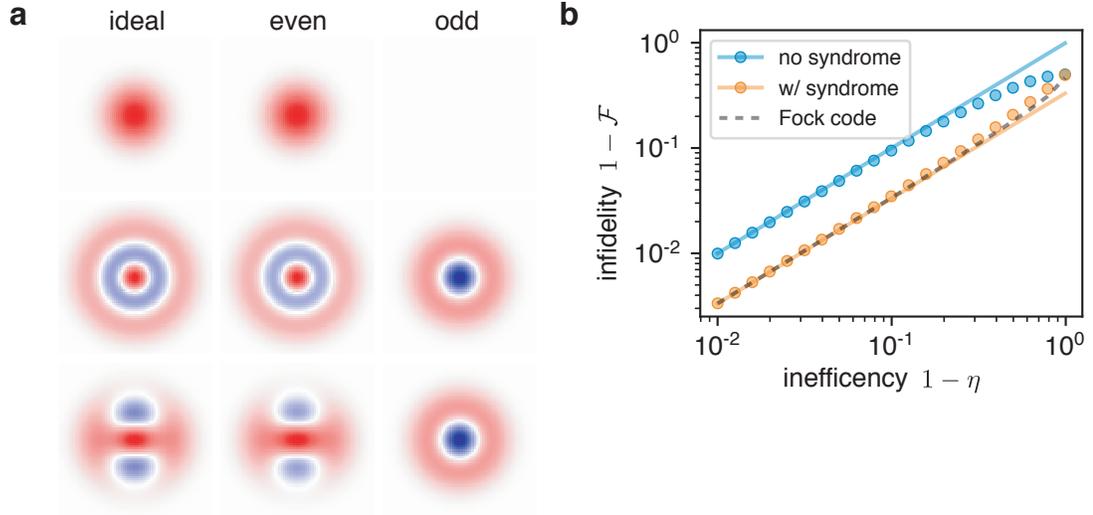
Figure 3.4: 0–2 Code. a) Left: Wigner functions of $|+z_{\mathrm{L}}\rangle, |-z_{\mathrm{L}}\rangle, |+x_{\mathrm{L}}\rangle$ from top to bottom. Center: Same logical states with loss channel applied $\eta = 0.8$, conditioned on even parity. Right: Same, conditioned on odd parity. b) Mean state infidelity versus inefficiency of channel, with and without parity conditioning. Linear approximation for no syndrome is $2(1-\eta)$, approximation with syndrome is $(1-\eta)/3$, same as Fock code.

at least to first order. So far, so good.

Things quickly get sad when we consider the one-jump error, signaled by a measurement of odd parity. In this case, there is only one possible state the oscillator can be in, $|1\rangle$. The best we can do is use the Fock code as our error space, to which we have a constant average infidelity of $1/2$. The resulting infidelity contribution is this constant factor times the error probability, which is on average $1 - \eta$. This is clearly a first-order error, so the scaling of the 0–2 code, even with perfect error detection, is no better than the Fock code. In fact, a full calculation of the error-detected mean fidelity in Figure 3.4 shows identical performance to the Fock code. This means that we can never expect to pass any break-even point with this encoding.

As a final note, another way to look at the failure of this encoding is to consider it from the point of view of the environment which is stealing photons from the oscillator. If the environment absorbs a photon, it effectively measures the state of the logical qubit to have been in $|-z_{\mathrm{L}}\rangle = |2\rangle$. The backaction of this measurement collapses the qubit into a known state. For this reason, this effect is sometimes

described as a dephasing error. Since the rate at which photons are lost from an oscillator is proportional to the average photon number $\bar{n}$, this effect will be present in any encoding where $\bar{n}$ is not constant as we move about the Bloch sphere. The dephasing rate is proportional to the difference in $\bar{n}$ between states in the encoding. This is precisely the K-L condition in Equation 3.21, which is not satisfied even to first order for E_1 in the 0–2 code.

3.4 Error correction with photons

The 0–2 code and its failure modes provide good instruction for building a first-order error-correcting code that uses number parity as the error syndrome. The requirements are:

1. A code space $|\pm z_{\mathrm{L}}\rangle$ with well-defined number parity

2. An set of error code words $|\pm z_{\mathrm{E}}\rangle \approx E_1 |\pm z_{\mathrm{L}}\rangle$ which are orthogonal to one another, and to $|\pm z_{\mathrm{L}}\rangle$

3. A uniform average photon number $\bar{n}$ throughout the encoding

As we will see, the second and third of these requirements may only be satisfied to first order in $1-\eta$ for most realistic encodings. This is acceptable as long as the second-order errors remain small. There will already be some second-order errors from the action of E_2, the two-jump error, so we're just looking to not make these too much worse.

3.4.1 The cat code

A primary code of interest in this thesis, and the first experimentally-demonstrated bosonic error-correction code, is the cat code, which has basis states formed from superpositions of coherent states [78]. As we saw in Subsection 3.1.2, a symmetric

superposition of two coherent states of opposite phase $|\alpha\rangle + |-\alpha\rangle$ has even parity. However, we need *two* orthogonal states of even parity to make up a logical encoding. For this, we use two distinct superpositions of *four* coherent states (legs) along the positive real, positive imaginary, negative real, and negative imaginary axes in phase space:

$$
\begin{aligned}
|+z_{\mathrm{L}}\rangle &= \frac{1}{\mathcal{N}_0(\alpha)} \left(|\alpha\rangle - |i\alpha\rangle + |-\alpha\rangle - |-i\alpha\rangle\right) \\
|-z_{\mathrm{L}}\rangle &= \frac{1}{\mathcal{N}_1(\alpha)} \left(|\alpha\rangle + |i\alpha\rangle + |-\alpha\rangle + |-i\alpha\rangle\right)
\end{aligned}
\tag{3.22}
$$

where $\mathcal{N}_{0,1}$ are normalization constants which differ from 2 to the extent that the legs are not orthogonal. This is actually a continuous family of codes, parametrized by the size α, assumed to be real and positive; the importance of this parameter will be discussed in Subsection 3.4.3.

Since each of the coherent states has the same weight in any given number state, but with different phases as per Equation 3.3, these superpositions interfere such that not only do the two code words have definite parity (number modulo 2), but they also have definite 4-parity (number modulo 4):

$$
\begin{aligned}
|+z_{\mathrm{L}}\rangle &= \frac{4}{\mathcal{N}_0(\alpha)} \sum_{n=2,6,10\ldots} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \equiv \left|\mathcal{C}_\alpha^{2\mathrm{mod}4}\right\rangle \\
|-z_{\mathrm{L}}\rangle &= \frac{4}{\mathcal{N}_1(\alpha)} \sum_{n=0,4,8\ldots} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \equiv \left|\mathcal{C}_\alpha^{0\mathrm{mod}4}\right\rangle,
\end{aligned}
\tag{3.23}
$$

where we have introduced notation similar to $|\mathcal{C}_\alpha^\pm\rangle$ for the two-legged cats. Due to this structure, the code words are manifestly orthogonal, and superpositions within the code space will have even parity, but undefined 4-parity. The 4-parity $\hat{P}_4$ is essentially the logical Z operator in this encoding:

$$
\hat{P}_4 = e^{i\pi a^\dagger a/2} = \sum_{n=0}^{\infty} i^n |n\rangle\langle n|
\tag{3.24}
$$

It is worth noting that in the limit of $\alpha \to 0$, the first term in the sums in Equation

3.23 dominates, and this encoding trends towards to $|\pm z_{\rm L}\rangle = |2\rangle, |0\rangle$, which is the

0–2 encoding up to an exchange of the logical code words.[4]

3.4.2 Cat code errors

Given the ability to measure parity, one can use this as an error syndrome for an

amplitude damping channel with the cat code. Under such a channel with some

finite efficiency, we observed in Subsection 3.2.3 that the amplitude of a coherent

state shrinks: $|\alpha\rangle \to |\alpha'\rangle = |\sqrt{\eta}\alpha\rangle$. This happens in the cat code as well, which

deforms the code space. Additionally, the parity can either remain even (no-jump)

or change to odd (one-jump). The code space in the no-jump case is the same but

with a smaller value of alpha:

$$
\begin{aligned}
|{+z'}_{\rm L}\rangle = \left|\mathcal{C}_{\alpha'}^{2\mathrm{mod}4}\right\rangle &= \frac{1}{\mathcal{N}_0(\alpha')} \left(|\alpha'\rangle - |i\alpha'\rangle + |-\alpha'\rangle - |-i\alpha'\rangle\right) \\
|{-z'}_{\rm L}\rangle = \left|\mathcal{C}_{\alpha'}^{0\mathrm{mod}4}\right\rangle &= \frac{1}{\mathcal{N}_1(\alpha')} \left(|\alpha'\rangle + |i\alpha'\rangle + |-\alpha'\rangle + |-i\alpha'\rangle\right)
\end{aligned}
\tag{3.25}
$$

For the case of a parity switch, the one-jump operator $E_1 \approx a$ has basically the same

effect, except that it also draws out a factor of the coherent state amplitude, since

the coherent state are eigenstates of the lowering operator. The resulting error states

are

$$
\begin{aligned}
|{+z}_{\rm E}\rangle = \left|\mathcal{C}_{\alpha'}^{1\mathrm{mod}4}\right\rangle &\propto \left(\alpha\,|\alpha'\rangle - i\alpha\,|i\alpha'\rangle - \alpha\,|-\alpha'\rangle + i\alpha\,|-i\alpha'\rangle\right) \\
|{-z}_{\rm E}\rangle = \left|\mathcal{C}_{\alpha'}^{3\mathrm{mod}4}\right\rangle &\propto \left(\alpha\,|\alpha'\rangle + i\alpha\,|i\alpha'\rangle - \alpha\,|-\alpha'\rangle - i\alpha\,|-i\alpha'\rangle\right)
\end{aligned}
\tag{3.26}
$$

In other words, one jump causes the four-parity to change, while also shrinking the

coherent state amplitude by the same amount as the no-jump case. By checking

for errors, we can assess the mean state fidelity in these new, shrunken encodings,

conditional on the parity signature. The resulting mean state fidelity, plotted in

---

4. This is an arbitrary choice of definition. In [79] the cat code basis states are swapped. In [78, 65] these are the x states, and the code words are two-legged cats.
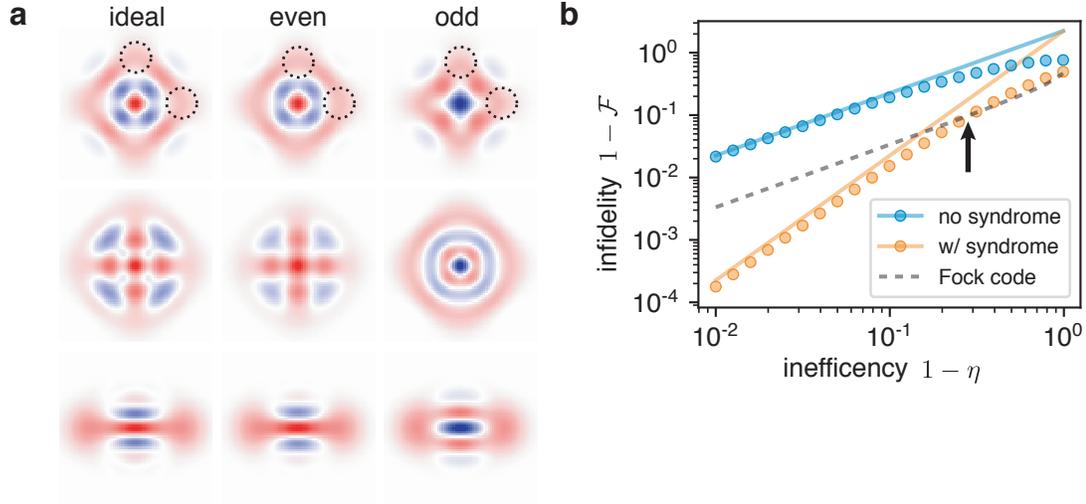
Figure 3.5: Cat Code. a) Cat code with $\alpha = 1.5$. Left: Wigner functions of $|+z_L\rangle, |-z_L\rangle, |+x_L\rangle$ from top to bottom. Dotted circles denote location of coherent state "legs." Center: Logical states with loss channel applied $\eta = 0.8$, conditioned on even parity. Circles are contracted towards the origin, centered at $\alpha' = \sqrt{\eta}\alpha$. Right: Same, conditioned on odd parity. b) Mean state infidelity versus inefficiency of channel, with and without parity conditioning. Linear approximation for no syndrome is $\alpha^2(1 - \eta)$. Approximation with syndrome is quadratic: $\alpha^2(1-\eta)^2$. Black arrow indicates break-even inefficiency, below which corrected infidelity is lower than Fock code, at $\eta = .72$.

Figure 3.5 for $\alpha = 1.5$, shows that this is a first-order error-correcting code, and it breaks even with the Fock code around $\eta = 0.72$.

3.4.3 Second order errors

The cat code is a fairly remarkable beast. Despite its seemingly complex structure, it has a very simple action under the no-jump and one-jump Kraus operators. So what can go wrong? The first, and most obvious, uncorrectable error is double-photon loss (E_2). Losing two photons causes the code words to change 4-parity by two, and they are mapped to one another. This is a logical *bit-flip* error, and it is missed by the parity check.

The probability of this double error depends both on the efficiency and the size of the cat. The larger the cat, the more photons it has, and the larger the prefactor of α^2 is in the Kraus operator (Equation 3.15). So to minimize this error probability,

we might think to make α smaller. The problem that arises is that for small coherent state amplitude ($\alpha \lesssim 2$), the form of the logical equator states becomes a strong function of alpha. When α is large, the legs of the cat are approximately orthogonal. In this limit, $\mathcal{N}_0(\alpha) \approx \mathcal{N}_1(\alpha) \approx 2$, and

$$
\begin{aligned}
|+x_{\mathrm{L}}\rangle &\approx \frac{1}{\sqrt{2}} (|\alpha\rangle + |-\alpha\rangle) = \left|\mathcal{C}_\alpha^+\right\rangle \\
|-x_{\mathrm{L}}\rangle &\approx \frac{1}{\sqrt{2}} (|i\alpha\rangle + |-i\alpha\rangle) = \left|\mathcal{C}_{i\alpha}^+\right\rangle
\end{aligned}
\tag{3.27}
$$

which are even parity two-legged cats, along the real and imaginary axes of phase space, respectively. However, when α is not that large, this approximation breaks down, and we actually have

$$
\begin{aligned}
|+x_{\mathrm{L}}\rangle &\approx \cos\theta_\alpha \left|\mathcal{C}_\alpha^+\right\rangle + \sin\theta_\alpha \left|\mathcal{C}_{i\alpha}^+\right\rangle \\
|-x_{\mathrm{L}}\rangle &\approx \cos\theta_\alpha \left|\mathcal{C}_{i\alpha}^+\right\rangle - \sin\theta_\alpha \left|\mathcal{C}_\alpha^+\right\rangle
\end{aligned}
\tag{3.28}
$$

for some small parameter θ_α which is a function of α. These two states are orthogonal (they must be, or this is not a good encoding), but they have overlapping support in phase space. If we apply the no-jump operator, which shrinks the coherent state amplitudes but doesn't change the relative weightings, we end up with

$$
\begin{aligned}
|+x'_{\mathrm{L}}\rangle &\approx \cos\theta_\alpha \left|\mathcal{C}_{\alpha'}^+\right\rangle + \sin\theta_\alpha \left|\mathcal{C}_{i\alpha'}^+\right\rangle \\
|-x'_{\mathrm{L}}\rangle &\approx \cos\theta_\alpha \left|\mathcal{C}_{i\alpha'}^+\right\rangle - \sin\theta_\alpha \left|\mathcal{C}_{\alpha'}^+\right\rangle
\end{aligned}
\tag{3.29}
$$

which are *not* orthogonal for $\alpha' \neq \alpha$, since the mixing angle θ_α is wrong for the new cat size α'. This means the equator states begin to overlap as the coherent states shrink, which is a logical *phase-flip* error.

It turns out this dephasing error, which is small as long as α is not too small, is exactly the same kind of effect we saw for the 0–2 encoding under the no-jump Kraus operator — it is related to the fact that $|\pm z_{\mathrm{L}}\rangle$ have different average photon number.
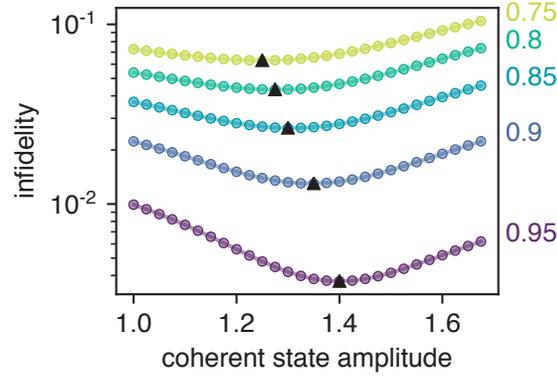
Figure 3.6: Fidelity vs Cat Size. For efficiencies ranging from 0.75 to 0.95 (text annotations at right), the ideal corrected infidelity versus cat size α. Optimum fidelity denoted by black triangles.

This leads to the measurement of parity gaining a little information about which code word the qubit is in, dephasing any superposition. This effect makes us prefer larger α, and so to optimize the performance we need to trade this against logical bit-flips from double jump errors. The performance versus α for a few different efficiencies is plotted in Figure 3.6. For a more complete discussion of this optimization, see [55, 69]. The takeaway is that lossier channels tend to prefer smaller α than less lossy ones. In a realistic implementation, we need to consider not just the efficiency of the channel, but also our ability to encode, decode, and perform syndrome measurements on the oscillator, which may also be α-dependent.

3.4.4 Other bosonic codes

There are a whole host of other bosonic error-correction encodings, and we give an incomplete list here.

General cat codes

We presented in Subsection 3.4.1 the four-legged cat code, but there are also versions of cat codes with six, eight, or more legs, which have larger separation in Fock space between code words. For example, in the six-legged cat code, all states have definite

photon number modulo 3 (0,3,6,9...), and so the code can tolerate up to two photon losses, provided that we can measure the three-valued syndrome three-parity. These codes have the potential advantage of correcting errors to higher order, or different kinds of errors, like photon gain. However, the legs are closer together as compared to the four-legged version, so to maintain the same protection from the shrinkage-induced non-orthogonality, we need to use more photons in our states. This requirement, along with the more complicated syndrome measurement, is part of the reason these codes have not been experimentally implemented yet.

Binomial codes

The so-called binomial codes proposed in [53] are constructed with the goal of exactly satisfying the K-L condition in Equation 3.21, which the cat code fails to do for small α. There is a very general construction given in the proposal, but codes in this class are constructed as superpositions of Fock states with binomial coefficients. This means the maximum photon number occupation is bounded, unlike in the cat code, where the occupation decays but is nonzero. The simplest version is

$$
\begin{aligned}
|+z_\mathrm{L}\rangle &= \left(|0\rangle + |4\rangle\right)/\sqrt{2} & |+z_\mathrm{E}\rangle &= |3\rangle \\
|-z_\mathrm{L}\rangle &= |2\rangle & |-z_\mathrm{E}\rangle &= |1\rangle
\end{aligned}
\tag{3.30}
$$

Indeed the mean photon number of any state in this encoding is $\bar{n} = 2$. However, unlike the cat code, the error space is a *bad* code in this respect, just like the 0–2 code. This means that, after measuring a parity jump, we should restore the state back to the original code space before carrying out another round of error-detection, as was done in [66].

This realization of the binomial code has very similar structure to the four-legged cat code for $\alpha = \sqrt{2}$, and the performance, shown in Figure 3.7, is very similar. We demonstrate the application of this code for communication in Chapter 6. Other
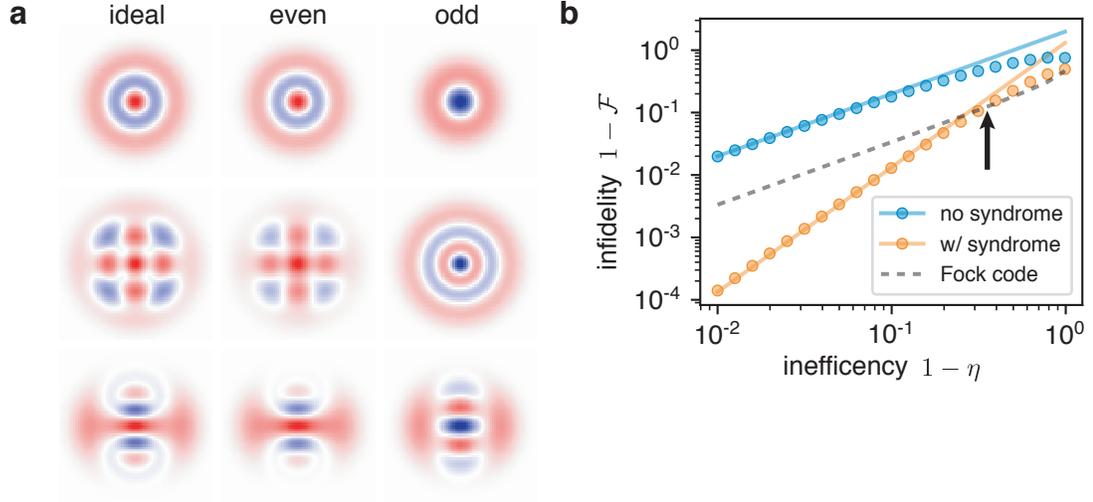
Figure 3.7: Binomial Code. a) Left: Wigner functions of $|+z_L\rangle , |-z_L\rangle , |+x_L\rangle$ from top to bottom. Center: Logical states with loss channel applied $\eta = 0.8$, conditioned on even parity. Right: Same, conditioned on odd parity. b) Mean state infidelity versus inefficiency of channel, with and without parity conditioning. Linear approximation for no syndrome is $2(1 - \eta)$. Approximation with syndrome is quadratic: $2(1 - \eta)^2$. Black arrow indicates break-even inefficiency, below which corrected infidelity is lower than Fock code, at $\eta = .68$.

versions of binomial codes are proposed in [53] to correct dephasing or photon gain as well as decay.

GKP codes

We mentioned earlier the grand-daddy of bosonic error correction, the GKP code [59], which has code words formed from superpositions of well-localized, Dirac-like peaks in phase space, arranged in a grid. As a result of the grid structure, these states are eigenstates of discrete displacements by the lattice constant, and so the syndrome involves measuring not parity, but a function of the displacement operator. Likewise, this means these codes can be protected from small random displacements by the environment. A recent surge of interest in GKP codes has come since the realization that these codes, somewhat remarkably, perform extremely well when exposed to amplitude damping errors. In fact, a comparison amongst the bosonic code zoo [55] showed that GKP codes are in some sense optimal, and perform extremely well even

for fairly modest efficiencies like the ones encountered in this thesis.

Other related encodings

Many of the codes discussed here were motivated by seemingly straightforward syndrome measurements. However, codes with less obvious syndrome measurements but other desirable properties exist. For example, the $\sqrt{17}$ code [53] corrects one decay error with $\bar{n} \approx 1.56$, and versions of cat and binomial codes with different structure to the signs of the Fock state amplitudes [80] allow for similar reduction of extent in Fock space. These codes sacrifice a sensible parity structure, so the syndrome measurement requires more complex control. Additionally, there are several proposals [53, 81] for encoding a single logical qubit in two or more oscillators, using some joint property as a syndrome.

Stabilized cat codes

It bears mentioning that there are alternatives to the scheme of encoding, allowing some free evolution of the oscillator, and then measuring a syndrome. There are several approaches to continuously stabilizing the code words which involve creation of two or more degenerate steady states in the oscillator into which a qubit can be encoded. Autonomous stabilization of cat states using exotic engineered dissipation was proposed [82] and realized in cQED [83], followed by manipulation of the state [84], and most recently, demonstration of exponential suppression of one type of error with cat size [85]. In parallel, a non-dissipative approach which engineers the Hamiltonian of a non-linear oscillator to create stable cat states was proposed [86] and implemented [28], demonstrating suppression of one error type. Both dissipative and Hamiltonian approaches have the potential to create noise-biased qubits with only one dominant error mechanism, which may provide efficient routes towards repetition codes with extremely forgiving thresholds [87, 88].

3.5 Conclusion

In this chapter we introduced the harmonic oscillator and its large Hilbert space as an object in which we may encode information in a hardware-efficient way. It is possible to build error-correction codes for the amplitude damping channel in a single harmonic oscillator, with a single error syndrome measurement. This is something which would have taken multiple physical two-level qubits and multiple syndrome measurements to do, as we discussed in Chapter 2.

We are going to use this ability to encode information in multi-photon quantum states like the cat code to move information around in very small modular quantum information machines. The primary error limiting these operation will be photon loss in the link between modules, and so we will leverage the error-correction strategy discussed here for the problem of quantum communication. In Chapter 4 we will discuss how to use photonic links as a means of sending information, and discuss a few approaches. Chapter 5 will get into the implementation of all of these tools in 3D cQED, and Chapters 6 and 7 will describe the experiments putting it all together.

Chapter 4

Building Bridges with Photons

In Chapter 2 we laid out three approaches towards modular quantum information machines: a shuttling architecture where qubits can be moved between modules; an entanglement-based approach where operations across the network are effected by teleported gates; and an architecture which implements gates directly between modules. In order to implement any of these architectures, we need a link or bridge between two or more modules which has the ability to support quantum information, at least transiently. As we explained in Chapter 2, we consider the links to carry microwave photons — in this chapter we will describe two approaches, one using propagating photons, and another which employs photons in a standing wave mode.

In Section 4.1 we describe some of the requirements we demand of such a link between modules, and what we expect to do with them. Section 4.2 gives the framework we use to describe the dynamical behavior when we couple two resonators together, and the useful unitary interaction which results. Next, Section 4.3 introduces the transmission lines which will make up the physical medium of the bridge. Here, we discuss coupling an oscillator to a transmission line in both the discrete mode (finite length) regime, as well as the continuum (infinite length) limit. The coupling can be physically realized in a few different ways, and this chapter will be mostly

agnostic to the implementation of the coupling. Our coupling of choice, which is frequency-converting parametric conversion, will be described in Chapter 5.

Next, we detail the two main approaches taken in this thesis. The first is the so-called "pitch and catch" scheme, or Cirac-Zoller-Kimble-Mabuchi (CZKM) protocol, named for the authors of the proposal [17], initially envisioned for atomic qubits in optical cavity, and first realized in [89]. In Section 4.4 we give the basic idea behind the in-principle lossless transfer of an arbitrary state from one resonator to another, using a uni-directional transmission line medium. We then describe a different and somewhat simpler coupling mechanism in Section 4.5. This approach uses a single mode of a finite-length section of transmission line, which we call a quantum bus. Here, we comment on some nice features enabled by this bi-directional link, especially when coupling resonators.

Finally, in Section 4.6 we describe how these different approaches perform when there is loss in the transmission channel, which is of great practical concern, as all implementation so far have been largely limited by such loss. We try to unify the ways one usually thinks about the loss in the two different regimes of interest, and describe how a few techniques for suppressing this loss can be implemented. Notably, we show that the simplest picture used usually breaks down in some limit, and the minimum amount of loss incurred is always that of a single pass through the transmission channel. We conclude with a few outgoing comments on how we hope to use bosonic error correction to surpass this hard limit.

4.1 What do we need from a bridge?

As we detailed in Chapter 2, we require links which enable some subset of the following operations:

– Moving qubit states between modules to perform local two-qubit gates

– Generating entanglement between qubits in separate modules to be used for state or gate teleportation

– Directly actuating entangling gates between qubits in separate modules

Since we have so far only realized the first two operations, we will mainly focus on these. The third approach will be discussed in Chapter 8, as we believe we have provided a promising step towards direct gates between modules.

We need not only to be able to perform one or more of these operations, but we need to do them *well*. By this I mean we need to not add too many errors, or at least introduce errors predominantly of only one type. To realize this goal, we require that the communication protocol is *rapid* compared to decoherence in the modules, so that the performance is not limited too much by qubit lifetimes. Additionally, we need to ensure we are not performing this operation when we do not intend to. For this we require at least some part of the scheme to be *switchable*. This can mean quite a few things, but here we mostly concern ourselves with a coupling at the interface between the module and link which can be turned on and off. Even more, we may require this control to be continuous; that is, that we can vary the coupling strength in time. This is an essential part of the CZKM protocol described in Section 4.4.

4.1.1 Photons as mediators

Interactions between qubits or resonators with transmission lines as demonstrated in cQED fulfill all of these requirements. A great body of work has demonstrated strong, switchable interactions of various kinds, including frequency-tunable resonators and qubits [90, 91, 92, 93], flux-tunable couplers [94, 95], and sideband or parametric modulation [96, 97, 71, 98], which is the tool used in this work. Importantly, the tunable couplings mentioned here enable the kind of continuous time-dependent control we need for certain approaches. These tools usually involve some sort of emission or

absorption of microwave light from or into the qubit.

Another feature of cQED is that interactions can be highly selective, meaning that the vast majority of photon emission from a qubit can be directed into a single channel. This is a great advantage over optical and atomic platforms, where emission into free space can be a major challenge in ensuring highly-efficient links. This selectivity comes from the ease with which microwave mode structures can be engineered on-chip, and from the effective one-dimensional confinement provided by microwave transmission lines and waveguides.

4.1.2   Hybrid quantum systems

Another advantage to using photons as mediators between modules is that many other platforms have shown strong interactions with microwave resonators or qubits, including but not limited to mechanical oscillators [99, 100], diamond color centers [101], quantum dots [102], and collective magnetic excitations of a bulk material [103]. This provides an appealing approach to building hybrid quantum networks, where modules might contain radically different physical hardware, all united by their ability to send and receive quantum information encoded in microwave photons. A compelling example of this is [104], where a quantum state was released from a superconducting resonator into a transmission line, then absorbed by a mechanical resonator in another module.

4.2   Coupling resonators

In this section, we discuss some couplings and operations of interest between two harmonic oscillators. Subsection 4.2.1 introduces a general type of unitary transformation between two resonators — the beamsplitter. This beamsplitter arises naturally in many contexts, and has a variety of uses, which we will discuss. Then in
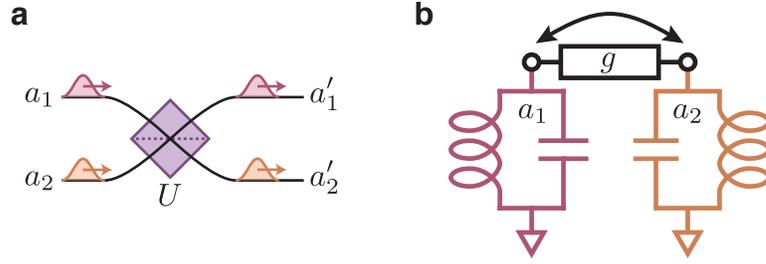
Figure 4.1: Coupled Oscillators. a) Schematic of beamsplitter transformation between two bosonic modes, here depicted as a physical beamsplitter with two propagating photonic inputs and outputs. b) Coupling configuration which effects the same transformation between stationary microwave modes. The element which coupled the oscillators with rate g can be capacitive, inductive, or a more exotic circuit.

Subsection 4.2.2 we turn to the typical cQED coupling Hamiltonian which generates a beamsplitter transformation, the conversion Hamiltonian. We'll discuss this conversion in a general setting, but then turn to the coupling between a resonator and a single mode of a transmission line. In Subsection 4.3.3 we introduce the effect of many modes, and when the mode spacing become dense (the line long), see the crossover to the continuum limit, where the response of the line becomes flat and dissipative. It is in this regime that we can convert between the standing oscillator and a wavepacket in the line by modulating the coupling strength in time.

4.2.1 Beamsplitter transformation

We consider two resonators, with field operators a_1 and a_2. In Figure 4.1a, these operators represent propagating photons incident on a beamsplitter. This is a central element in linear quantum optics [58, 105], and we like to use this picture because it is an easy way to to think about how the field operators behave. Any sequential combination of beamsplitters and single-mode phase shifters can be represented by a single two-dimensional unitary matrix U which acts on the vector $\vec{a} = [a_1, a_2]$, which is the input to the optical circuit. The output field operators are $\vec{a}' = [a_1', a_2'] = U\vec{a}$.

The most general two-mode unitary transformation is

$$\vec{a}' = U\vec{a} = e^{i\varphi} \begin{bmatrix} e^{i\zeta}\cos\theta & e^{i\phi}\sin\theta \\ -e^{-i\phi}\sin\theta & e^{-i\zeta}\cos\theta \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \tag{4.1}$$

which we generally call the *beamsplitter transformation*. The angle θ denotes how much the resonator modes mix as a result of the interaction, and is accordingly called the mixing, or beamsplitter, angle. The angles ζ and ϕ are relative phases between the arms of the beamsplitter inputs, and φ is a global phase imparted on the output.

This type of transformation is not on its own sufficient for quantum computation with bosonic modes. It belongs to a class know as *Gaussian operations*, because it effects a mode-mixing in phase space, which will transform Gaussian states[1] into Gaussian states. The beamsplitter does not create quantum states from classical ones, and Gaussian operations on Gaussian states can be efficiently computed classically, and therefore offers no quantum advantage. However, this does not mean Gaussian operations are not useful. On the contrary, when paired with some quantum resource like single photon inputs, or single-photon detectors on the output, linear operations are sufficient for universal quantum computation [58]. Since cQED offers well-understood methods of preparing and measuring single photons, Gaussian operations are of use for measurement and manipulation of bosonic qubits, which includes generation of multi-photon entangled states [106, 107].

4.2.2   Conversion coupling

Figure 4.1 shows the lumped-element cQED picture (b) corresponding to the linear optics one (a). The beamsplitter transformation in Equation 4.1 can be generated by

---

1. Gaussian states are those for which the Wigner function is a two-dimensional Gaussian. This includes symmetric Gaussians (coherent states) and asymmetric ones (squeezed states).

a bi-linear Hamiltonian of the form

$$\mathcal{H}_{\text{conv}} = g a_1^\dagger a_2 + g a_1 a_2^\dagger \tag{4.2}$$

For simplicity, we take g real. More generically, $\mathcal{H}_{\text{conv}} = g a_1^\dagger a_2 + g^* a_1 a_2^\dagger$ for any complex g. This Hamiltonian, which converts photons between the two modes, arises in many contexts in cQED. For example, coupling the two resonators with a capacitor or inductor and setting their frequencies to be equal results in this Hamiltonian in the rotating frame of the resonators. In Chapter 5 we will explain how it can arise even when the resonators are far-detuned using driven parametric conversion.

Throughout this chapter we will write the equations of motion for the field operators in the Heisenberg picture. This is convenient because it allows for easy semi-classical solutions when the Hamiltonian is bilinear like Equation 4.2. This will let us find the exact solution for coherent states by treating the operators $a_{1,2}$ as complex numbers with initial values $A_{1,2}$. This corresponds to taking the expectation value of the equations. Additionally, we can find the exact solution for the mean photon number $\langle a_k^\dagger a_k \rangle = |\langle a_k \rangle|^2$ in each mode. This will allow us to extract the state-independent efficiency η as defined in Section 3.2. Note that in most cases, this photon number calculation is only certain to be correct if one of the oscillators begins in vacuum (or both in coherent states). This is because the classical solution will capture the classical phase interference, but not the quantum-mechanical interference. It will not, for instance, accurately predict the statistics of Hong-Ou-Mandel interference [19], and we will need to use a quantum mechanical calculation for this.

The Heisenberg equations of motion resulting from Equation 4.2 are

$$\begin{aligned}
\dot{a}_1(t) &= -iga_2(t) & a_1(t=0) &= A_1 \\
\dot{a}_2(t) &= -iga_1(t) & a_2(t=0) &= A_2,
\end{aligned} \tag{4.3}$$

which admit solutions

$$a_1(t) = A_1 \cos(gt) - iA_2 \sin(gt)$$

$$a_2(t) = A_2 \cos(gt) - iA_1 \sin(gt).$$

(4.4)

We can write this evolution as a unitary in matrix form (a la Equation 4.1):

$$U(t) = \begin{bmatrix} \cos(gt) & -i\sin(gt) \\ -i\sin(gt) & \cos(gt) \end{bmatrix},$$

(4.5)

which is the general beamsplitter with $\varphi = \zeta = 0$, $\phi = -\pi/2$, and time-dependent mixing angle $\theta(t) = gt$. This means that an arbitrary mode mixing can be effected by tuning either g or t.

SWAP from beamsplitter

To give an example of how to work in this formalism and show the utility of this type of operation, we consider a simple example of how the beamsplitter can effect a SWAP operation at $gt = \pi/2$, where

$$U(t) = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} \qquad \text{or} \qquad \begin{matrix} a_1 \to -ia_2 \\ a_2 \to -ia_1. \end{matrix}$$

(4.6)

Since we're working with bosonic modes, there are many logical encodings we can consider, but we'll stick with the simplest one, the Fock encoding $\{|0_L\rangle\,|1_L\rangle = |0\rangle, |1\rangle\}$.

We can compute the matrix corresponding to U in the Fock basis, by calculating matrix elements like

$$\langle 01|U|10\rangle = \langle 00|a_2 U a_1^\dagger|00\rangle = \langle 00|a_2(ia_2^\dagger)U|00\rangle = i\langle 00|U|00\rangle = i,$$

(4.7)

where we use $\langle 00|U|00\rangle = 1$, since the Hamiltonian of Equation 4.2 has no action on

the vacuum state. By computing the rest of the matrix elements likewise, we find the logical unitary U_L in the Fock encoding is

$$
U_\mathrm{L} =
\begin{array}{cccc}
\ \ 00 & 01 & 10 & 11
\end{array}
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & i & 0 \\
0 & i & 0 & 0 \\
0 & 0 & 0 & -1
\end{bmatrix},
\tag{4.8}
$$

which is a SWAP operation between the two qubits, up to $Z_{\pi/2}$ gates on both qubits:
$U_\mathrm{L} = \left(Z_{\pi/2} \otimes Z_{\pi/2}\right) \times \mathrm{SWAP}$.

While we have shown this is true for the Fock encoding, the $\pi/2$ beamsplitter effects SWAP on any encoding. This can be seen by writing the logical states in the Fock basis as $\sum_n c_n |n\rangle = \sum_n \frac{c_n}{\sqrt{n!}} a^{\dagger n} |0\rangle$ and computing the matrix elements by passing the factors of $a^\dagger$ through U as above.

Entangling operations

Equation 4.8 looks like iSWAP, but it is not, due to the phase that appears on the $|11\rangle$ state. So the beamsplitter with $\theta = \pi/2$ is not an entangling operation. However, the 50:50 beamsplitter ($\theta = \pi/4$) can entangle the resonators, for certain input states. It is important to note, that although the 50:50 beamsplitter can be generated by simply halving the evolution time of the SWAP gate to $gt = \pi/4$, it is *not* $\sqrt{\mathrm{SWAP}}$. This is due to the interference properties of bosons. For example, even in the Fock encoding, the 50:50 beamsplitter on the separable input state $|11\rangle$ results in the state $|02\rangle + |20\rangle$ — this is the well-known Hong-Ou-Mandel interference [19]. While this is a maximally entangled state, it is outside of the original code space, and therefore is not a gate.

Although it is not a computational gate, the 50:50 beamsplitter is a useful entangler, even for the simple single-photon input state $|10\rangle$. In Chapter 7 we will exploit the properties of this beamsplitter to generate entanglement between modules in a simple network. Additionally, one can construct computational gates like $\sqrt{\text{SWAP}}$ [108, 107] by cascading 50:50 beamsplitters and single-resonator operations.

Other couplings

It is worth noting that Equation 4.1 is not the only operation which can result from bi-linear Hamiltonians. By including interaction terms which do not preserve photon number, like $a_1^\dagger a_2^\dagger + a_1 a_2$, we can generate squeezed light, which is quantum but still Gaussian. This kind of two-mode squeezing interaction is used in parametric amplifiers (see [109] for a review), and the single-mode version, along with a two-mode beamsplitter, has been used to simulate problems in quantum chemistry [110]. There are also non-linear couplings like self-Kerr ($a_1^\dagger a_1^\dagger a_1 a_1$) and cross-Kerr ($a_1^\dagger a_1 a_2^\dagger a_2$) interactions, to be expanded upon in Chapter 5, which can generate quantum non-Gaussian light.

4.3 Transmission line and distributed-element resonators

The physical link of choice in cQED will be microwave coaxial transmission line, which is compact, high-bandwidth, flexible, easy to couple to, commercially available, and can be very low loss when made from superconducting materials. Coaxial line of finite length can act as a multi-mode resonator. In this section, we'll give a quick description of transmission line of finite length, and consider coupling the modes of this distributed element object to an external resonator through one end. We'll then consider the continuum limit, where the line becomes long enough that the modes are

no longer discrete. In these regime we talk about propagation of photons wavepackets, which is a superposition of many standing modes. We'll then discuss the effects of dissipation in both regimes.

4.3.1 Transmission line

A length of transmission line is essentially a one-dimensional object, made of two conductors with mutual capacitance and inductance. We'll mostly be concerned with coaxial transmission line like the one pictured in Figure 4.2, which consists of an inner conductor and an outer ground shield, but most of this holds for other types as well. We first consider the lossless transmission line, with capacitance and inductance per unit length $\mathcal{C}$ and $\mathcal{L}$, respectively. The voltage $V(x,t)$ between the two conductors, and current $I(x,t)$ in the center conductor, as a function of time and space in the line, are related by the telegrapher's equations

$$
\begin{aligned}
\frac{\partial V(x,t)}{\partial x} &= -\mathcal{L}\frac{\partial I(x,t)}{\partial t} \\
\frac{\partial I(x,t)}{\partial x} &= -\mathcal{C}\frac{\partial V(x,t)}{\partial t},
\end{aligned}
\tag{4.9}
$$

We usually combine these into a wave equation for the current (or voltage)

$$
\frac{\partial^2 I(x,t)}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 I(x,t)}{\partial t^2},
\tag{4.10}
$$

with speed of light $c = 1/\sqrt{\mathcal{L}\mathcal{C}}$. Note that the speed of light in the line is usually slower than in vacuum, due to a capacitance which is increased by the presence of a dielectric.

If the length of the line is ℓ and it is open at each end, there are boundary conditions $I(x = 0, t) = I(x = \ell, t) = 0$. The one-dimensional wave equation with homogeneous boundary conditions has infinite discrete eigenmode solutions with complex
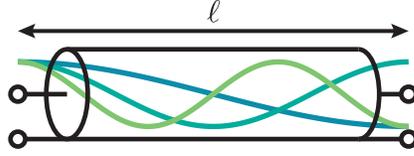
Figure 4.2: A Transmission Line Resonator. A length of open-ended transmission line supports standing modes. The voltage profiles of the first three modes are shown, with $\lambda_1 = 2\ell$, $\lambda_2 = \ell$, and $\lambda_3 = 2\ell/3$, from dark to light.

current and voltage profiles and eigenfrequencies

$$
\begin{aligned}
\mathcal{I}_n(x,t) &= I_n \sin\left(\frac{2\pi x}{\lambda_n}\right) e^{-i\omega_n t} \\
\mathcal{V}_n(x,t) &= V_n \cos\left(\frac{2\pi x}{\lambda_n}\right) e^{-i\omega_n t}
\end{aligned}
\qquad \lambda_n = \frac{2\ell}{n} \qquad \omega_n = \frac{n\pi c}{\ell} = 2\pi n \text{FSR}, \qquad (4.11)
$$

for positive integer n, where I_n and V_n are the maximum current and voltage for mode n, respectively. We have introduced the free spectral range, $\text{FSR} = c/2\ell$, which is the frequency spacing between modes in *angular* frequency units. The FSR is the inverse of the round-trip for light to travel from one end of the line to the other and back. This timescale will be relevant later on when we discuss the crossover from standing to propagating light. The first three voltage profiles are shown in Figure 4.2.

4.3.2 Coupling

We can couple the modes of a transmission line to an external, lumped- or distributed-element, resonator in any of the ways we discussed in Section 4.2. Typically, a lumped resonator can be connected to the end of the line by some coupling element, with rate g. This topology is shown in Figure 4.3a. When we do this, the lumped resonator will in fact couple to all of the modes of the line. The strength of these couplings may not be equal, but we'll take a simplified approach and assume they are[2]. The

---

2. Capacitive or inductive couplings have frequency-dependent rates, increasing with frequency for a capacitor and decreasing for an inductor. For the parametric couplings described in Chapter 5 and used in Chapter 7, the frequency dependence is nontrivial.

a **b**
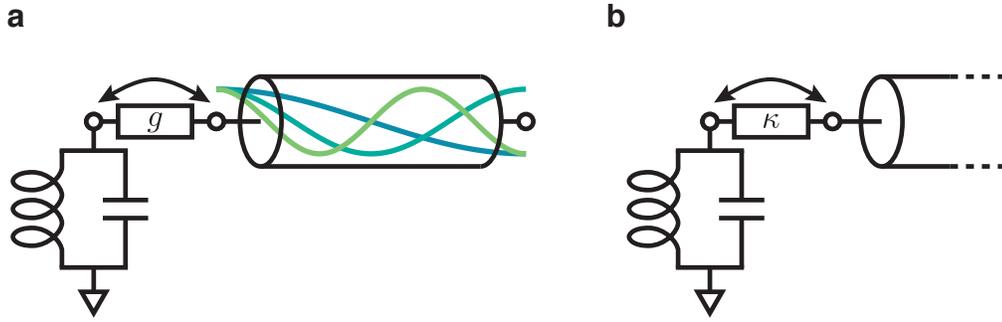
Figure 4.3: Coupling a Resonator to a Transmission Line. a) A lumped element resonator coupled to one end of a finite transmission line. The resonator can couple to any mode of the line. b) In the infinite length limit, where the mode frequency spacing is smaller than the coupling rate g, the resonator decays at a rate κ into the continuum of the line.

coupling to multiple modes is will mostly of concern when the FSR becomes small, so we may hope that at least for small changes in frequency the coupling can be assumed roughly constant.

4.3.3 The continuum limit

Here we'll describe the crossover from coupling to one mode (unitary dynamics) to the coupling to many modes (input-output theory and the Langevin equation) as the length of the transmission line is made long, as in Figure 4.3b. The discussion here closely follows the much more complete treatment in Chapter 12 of [111]. The goal is to eliminate the infinite number of transmission line modes to get an effective dynamics for the lumped mode. This will involve replacing the unitary coupling rate g with a decay constant κ which we will derive below. We'll then circle back to what happens in the line.

Let's call the lumped mode a and the modes of the line b_k, with frequencies $\omega_k = 2\pi k$FSR. The conversion Hamiltonian Equation 4.2 assumes the modes involved are resonant. To describe the coupling to multiple modes, we need to introduce detuning, which we will do with a time-dependent Hamiltonian, going into the rotating frame

for each of the line modes, but stay in the lab frame for mode a. The detuning of mode k from the lumped resonator is $\Delta_k = \omega_k - \omega_a$. The Hamiltonian reads

$$\mathcal{H}(t) = \mathcal{H}_a + \sum_k \left(g a(t) b_k^\dagger(t) e^{i\Delta_k t} + g a^\dagger(t) b_k(t) e^{-i\Delta_k t} \right), \tag{4.12}$$

where $\mathcal{H}_a$ is the rest of the Hamiltonian, involving only mode a, including the frequency term $\omega_a a^\dagger a$. We write the Heisenberg equations

$$\dot{a}(t) = i[\mathcal{H}_a, a] - ig \sum_k b_k(t) e^{-i\Delta_k t} \tag{4.13a}$$

$$\dot{b}(t)_k = -iga(t) e^{i\Delta_k t}, \tag{4.13b}$$

then integrate Equation 4.13b to get $b(t)$

$$b_k(t) = -ig \int_0^t a(\tau) e^{i\Delta_k \tau} d\tau \tag{4.14}$$

and substitute it into Equation 4.13a:

$$\dot{a}(t) = i[\mathcal{H}_a, a] - g^2 \sum_k \int_0^t a(\tau) e^{i\Delta_k(\tau - t)} d\tau. \tag{4.15}$$

This is generally true, and gives the dynamics for $a(t)$, which depends on all past times. But to get it into a more usable form, we now take the limit where the modes b_k of the line become very close together in frequency. This is the limit where $\ell \to \infty$ and FSR $\to 0$. We can replace the sum over k with an integral over Δ, using the

density of states, which is $1/(2\pi\mathrm{FSR})$:

$$
\begin{aligned}
\dot{a}(t) &= i[\mathcal{H}_a, a] - \frac{g^2}{2\pi\mathrm{FSR}} \int_{-\infty}^{\infty} \int_0^t a(\tau) e^{i\Delta(\tau-t)} \mathrm{d}\tau \mathrm{d}\Delta \\
&= i[\mathcal{H}_a, a] - \frac{g^2}{2\pi\mathrm{FSR}} \int_0^t a(\tau) 2\pi\delta(\tau-t) \mathrm{d}\tau \qquad\qquad (4.16) \\
&= i[\mathcal{H}_a, a] - \frac{g^2}{\mathrm{FSR}} \frac{1}{2} a(t).
\end{aligned}
$$

The integral over frequency yields a Dirac delta function, which when integrated over gives us half the integrand[3], resulting in this effective equation for $a(t)$. We make the identification $\kappa \equiv g^2/\mathrm{FSR}$, and have

$$
\dot{a}(t) = i[\mathcal{H}_a, a] - \frac{\kappa}{2} a(t). \qquad\qquad (4.17)
$$

This is the Langevin equation for an oscillator a coupled to a semi-infinite transmission line, though it can also be used to describe any other source of damping as well.

Frequency integration limits

We played a bit fast and loose with the conversion of the sum over modes of the line into an integral over frequency space. In particular, we integrated over all detunings, from $-\infty$ to ∞, when in reality there is a minimum frequency. This is justified because modes that are far away in frequency should not contribute much to the dynamics of a. However, there is an implicit assumption here that $g \ll \omega_a$, which is usually the case in cQED, though the so-called ultra-strong coupling regime has become of interest in recent years [112, 113, 114, 115, 116].

---

3. Half because the integrals stops at t, so we only integrate over half the delta function.

Input field

The derivation above assumes any amplitude in the modes b_k comes from a; that is, we assumed $b_k(t = 0) = 0$. But they could already have some occupation. This manifests as an additional term in Equation 4.15 which becomes an integral over all modes of the line at $t = 0$, and is proportional to

$$\int_{-\infty}^{\infty} b[\Delta]e^{-i\Delta t}\mathrm{d}\Delta, \tag{4.18}$$

and which we can call $a_{\text{in}}(t)$. It shows up in the final equation of motion as an additional term

$$\dot{a}(t) = i[\mathcal{H}_a, a] - \frac{\kappa}{2}a(t) - \sqrt{\kappa}a_{\text{in}}(t). \tag{4.19}$$

Since a_{in} is a superposition of eigenmodes in the line, it can be thought of as a wavepacket traveling towards the resonator. Any wavepacket whose spatial extent fits within the line can be Fourier expanded in this way. In other words, if the frequency bandwidth of the input field is larger than the FSR, we can think of it as an incoming wavepacket.

Output field

Likewise, the occupation of the transmission line modes also gives rise to an *outgoing* wavepacket, which we call $a_{\text{out}}(t)$. Again, it makes sense to think in this way as long as mode a is interacting with enough of the line modes that it generates a superposition in frequency space. This is the case if the coupling strength g is as large or larger than the FSR. In the time domain, we can say that if the interaction time $1/g$ is short compared to the round trip time $1/\text{FSR}$, then the state of the resonator a can be converted to propagating light in the line before that light has time to reflect off the far end, return, and interfere with itself. In this picture the concept of standing wave modes (ultimately an interference effect) breaks down, and we talk about wavepackets.

The interplay of the resonator field with the incoming and outgoing wavepackets is succinctly summarized in the *input-output relation*

$$a_{\text{out}}(t) = \sqrt{\kappa}a(t) + a_{\text{in}}(t). \tag{4.20}$$

This says that the outgoing field is a sum of the emitted resonator field and the reflected incoming field.

There is a phase convention we have used here between the definitions of a_{in} and a_{out}. This manifests here as a plus sign in Equation 4.20 — the incoming field is reflected with no phase flip. This is the case for a parallel LC resonator with a capacitor or inductor in series with the resonator and the line (as in Figure 4.1), since this looks like an open circuit off-resonance. For a *series* LC resonator with a parallel coupling capacitance or inductance, the reflected signal is inverted because the circuit looks like a short to ground off-resonance. This doesn't really change the discussion much, other than a sign convention on the output.

Relationship between coupling and dissipation rates

A final comment about the relation $\kappa = g^2/\text{FSR}$. This is a handy statement, but a little confusing. The left-hand side is a statement about the continuum limit, but the right side assumes discrete modes. Moreover, g and FSR are length-dependent, whereas κ is not. It turns out that the ratio g^2/FSR is constant as a function of length for a fixed coupling element, since $\text{FSR} \sim 1/\ell$ and $g \sim 1/\sqrt{\ell}$. The coupling strength decreases as the length increases because g^2 is the per-photon interaction energy. As the mode volume increases, the voltage and current at the ends of the line decrease, which lowers the interaction energy[4]. This also means that we may consider the continuum limit as 1) the regime where the length is longer than the

---

4. Put another way, the zero-point voltage and current fluctuations decrease as $1/\sqrt{\ell}$ for constant energy per photon (frequency).

wavepacket spatial extent, or 2) the limit where the round trip time is long compared to the interaction time. These are different ways of saying $g > \text{FSR}$.

4.3.4 Loss

While the transmission lines we use in this thesis are superconducting, they are are only lossless at zero frequency. The exact mechanisms which contribute to the loss at finite frequency are not entirely known, but it is likely some combination of dielectric loss, surface conductance loss, and resistive joint losses. These are all the same sorts of factors which dictate losses of superconducting resonators and qubits [117, 118, 119, 120]. Here we consider the effects of a few different kinds of losses on the modes of a transmission line resonator, and relate that to the effect on propagating wavepackets.

Single-pass loss

A simple assumption for the loss in a line is that the dissipation is equally distributed throughout the length of the line. This would be the case for a uniformly lossy dielectric, or a finite surface conductance. In this case, we can talk about an attenuation constant ζ, which has dimension of inverse length. For a wavepacket propagating down the line, the energy decays exponentially in the length x traveled, and is proportional to $e^{-\zeta x}$. Note that here ζ is the energy attenuation constant, which is twice the electric/magnetic field attenuation constant, usually referred to as α in the literature [121]. We use this somewhat nonstandard notation ζ for the attenuation constant to avoid confusion with α, which we have been using to denote coherent state amplitudes. Given ζ, the *single-pass loss* (SPL), the fraction of energy lost over a single transit from one end to the other of a length ℓ, is

$$\text{SPL} = 1 - e^{-\zeta \ell} \approx \zeta \ell, \tag{4.21}$$

where the approximation holds if the loss is fairly small. This approximation will be valid for everything considered here.

Damping rates and quality factors

We have spoken of wavepackets, but we know that a finite length of line has standing modes. Loss in the line means these modes have finite decay rates κ_n, where n indexes the mode number. The decay rate has dimension of inverse time, and the fraction of energy remaining in a time t is $e^{-\kappa t}$. Comparing this to the equivalent exponential involving the attenuation per length, and remembering that the speed of light c relates $x = ct$, we have

$$e^{-\kappa_n t} = e^{-\zeta x} = e^{-\zeta ct} \implies \kappa_n = c\zeta. \tag{4.22}$$

Another useful identity is to relate the SPL to the decay rate, which involves the FSR:

$$\text{SPL} \approx \zeta\ell = \zeta\frac{c}{2\text{FSR}} = \frac{\kappa_n}{2\text{FSR}}. \tag{4.23}$$

As an aside, recall that FSR is an angular frequency (in Hz), while κ is an inverse time, so there is a hidden factor of 2π in this expression.

We'll use κ_n below in some figures of merit, but an oft-quoted and related quantity is the dimensionless quality factor $Q_n = \omega_n/\kappa_n$, so it is useful to relate this to the above:

$$Q_n = \frac{\omega_n}{\kappa_n} = \frac{2\pi n\text{FSR}}{\kappa_n} = \frac{\pi n}{\text{SPL}} = \frac{\pi n}{\zeta\ell}. \tag{4.24}$$

There is an explicit dependence on the mode number in the quality factor, but there can also be an implicit dependence if ζ (and hence, the SPL) is frequency-dependent. If this is the case, then κ_n will be frequency-dependent as well. Whether this is the case will depend on the dissipative mechanism.

We should give some rough numbers for what we can expect for ζ is superconducting transmission line. For a typical on-chip stripline resonator (see Subsection

5.1.2), $l = 1\,\mathrm{cm}$ and the fundamental mode ($n = 1$) has $Q = 10^6$ or higher [122]. Then Equation 4.24 gives us $\zeta = \frac{\pi n}{Q\ell} = 1/(3\,\mathrm{km})$, which is roughly equal to that of optical fiber used for long-distance quantum communication. 3D coaxial cavity resonators, discussed in Subsection 5.1.3, store most of their energy in the lossless vacuum and can be even lower-loss, with $Q = 8 \times 10^7$ [79], which gives an attenuation constant of $\zeta = 1/(700\,\mathrm{km})$. While the transmission line links we employ for the experiments in this thesis are not yet at this quality (see Section 5.8), we should be optimistic about the prospects of superconducting microwave links.

Dielectric loss

If the loss is sourced by loss in a dielectric (like PTFE), which fills the entirety of the transmission line, then we often characterize the dissipation of this material with a dimensionless loss tangent $\tan\delta$ [121], which is the fraction of energy dissipated per cycle of oscillation, per the fraction of energy stored in the dielectric. For a full filling, all the electric energy lives in the dielectric, and

$$Q_n = \frac{1}{\tan\delta} \implies \kappa_n = \omega_n \tan\delta = 2\pi n \mathrm{FSR} \tan\delta. \tag{4.25}$$

If the loss tangent is frequency-independent, then the decay rate increases linearly with frequency.

Joint loss

Another possible source of loss arises because the transmission line must be connected to something at its ends. When we use a line to bridge two modules, there will be a connection between the ground shield of the line and the housing of the module at each end. This joint can have some resistance, which will dissipate energy if the modes of the line flow current across it. A simple model where the joint is envisioned
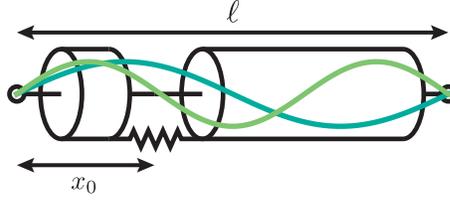
Figure 4.4: Joint Loss in a Transmission Line Resonator. A simple model for joint loss places a resistor in the round shield a distance x_0 from the end of the line. Current profiles for modes $n = 2, 3$ show strong dependence of current in the resistor on position This position is a current zero for mode $n = 3$ (light), while for $n = 2$ (dark) it is near a maximum.

as a interruption in the ground of the line with resistance R, a distance x_0 from one end, is pictured in Figure 4.4. In particular, different modes have different spatial current profiles, which can affect the amount of loss inherited by each.

To calculate the damping rate for mode n caused by the resistor, we need the ratio of power dissipated to total energy stored. The power dissipated is

$$P_n = \frac{1}{2} R |\mathcal{I}_n(x_0)|^2 = \frac{R I_n^2}{2} \sin^2 \left(\frac{2\pi x_0}{\lambda_n} \right) = \frac{R I_n^2 \beta_n}{2}, \tag{4.26}$$

which is half the maximum instantaneous power. We define $\beta_n = \sin^2 (2\pi x_0/\lambda_n)$, which is a mode-dependent geometric factor encoding the amount of current the mode puts across the resistor.

In analogy to a lumped element resonator, the total energy in the mode is the maximum energy stored in the inductor

$$E_n = \frac{1}{2} L I_n^2 = \frac{Z_n I_n^2}{2\omega_n}, \tag{4.27}$$

where Z_n is the characteristic impedance of mode n. For a transmission line with characteristic impedance Z_c, we have [121]:

$$Z_n = \frac{2}{n\pi} Z_c. \tag{4.28}$$

Z_c depends on geometry, and is usually $50\,\Omega$ or thereabouts.

The damping rate is then

$$\kappa_n = \frac{P_n}{E_n} = \frac{\omega_n R I_n^2}{Z_n I_n^2 \beta_n} = \frac{\pi n \omega_n \beta_n}{2}\frac{R}{Z_c} = \pi^2 n^2 \beta_n \mathrm{FSR}\frac{R}{Z_c}, \tag{4.29}$$

and the quality factor is

$$Q_n = \frac{\omega_n}{\kappa_n} = \frac{2}{\pi n \beta_n}\frac{Z_c}{R}. \tag{4.30}$$

We have written these expressions to make it clear that there is a dimensionless ratio of the resistance to the (mode-independent) characteristic impedance Z_c, as well as a mode-dependent geometric factor. The geometric factor β_n may not be monotonic in n, depending on where the joint is. For instance, if the joint is exactly in the middle $(x_0 = \ell/2)$, then $\beta_n = 1$ for odd n, but $\beta_n = 0$ for even n, and the loss vanishes.

This geometric dependence makes it hard to say anything about the scaling with n in general. However, an interesting regime is $x_0 \ll \lambda_n$ for all n of interest — this corresponds to the joint very near the end of the line. In this case $\beta_n \approx (2\pi x_0/\lambda_n)^2 = n^2\pi^2 x_0^2/L^2$. Then we have damping rate and quality factor

$$\begin{aligned}
\kappa_n &= \pi^4 n^4 \frac{x_0^2}{L^2}\mathrm{FSR}\frac{R}{Z_c} \\
Q_n &= \frac{1}{\pi^3 n^3}\frac{2L^2}{x_0^2}\frac{Z_c}{R},
\end{aligned} \tag{4.31}$$

which is a strong dependence on mode number (and frequency) indeed. This joint-induced scaling of $Q_n \sim 1/n^3$ could contribute to the stronger-than-expected behavior observed at low temperatures in [123].

We can estimate what sort of resistance R we might expect between a joint for two superconductors, which has been measured for cavity resonators [124]. In this case, the joint is not usually point-like, but instead is distributed over some seam with perimeter l_{seam}. We parametrize the quality of the seam with the conductance g_{seam},

which has units of $1/\Omega\,\mathrm{m}$, or conductance per length of seam. The total resistance is then $R = 1/(g_{\mathrm{seam}}l_{\mathrm{seam}})$. For the kind of coaxial cable used in Chapters 5 and 7, the perimeter is $l_{\mathrm{seam}} = 2\pi r \approx 3\,\mathrm{mm}$. For bulk aluminum alloy, $g_{\mathrm{seam}} \approx 10^4/\Omega\,\mathrm{m}$, which gives $R = 30\,\mathrm{m\Omega}$. From Equation 4.30, this would grant a worst-case $(\beta_n = 1)$ of $Q \approx 10^3$. However, there are two ways the seam-limited quality can be much higher than this. The use of indium films as a contact layer between metals can increase the seam conductance to $g_{\mathrm{seam}} \geq 10^8/\Omega\,\mathrm{m}$, which would improve this worst-case quality factor by four orders of magnitude to $Q \approx 10^7$. More recently, indium bump-bonding has increased this bound by an additional factor of 100 [125].

Moreover, we can improve the resonator quality by placing the joint in a better place, where the current is smaller. For instance, if the joint is within 2% of a wavelength from the end $(x_0 = 0.02\lambda_n)$, then $\beta)_n \approx .01$, and even with the worst seam quality from [124], we would have $Q \approx 10^5$, with plenty of room to improve with a more optimistic seam.

4.3.5 Waveguide

We pause for a moment to mention that transmission line is not the only possible photonic link. We could instead use waveguide, which is made of a single conductor with a hollow interior. Much of this discussion still applies, but there may be some advantages for waveguide links. For instance, without a center conductor, there is no need to use dielectric to support it, and so we can eliminate bulk dielectric loss. Waveguide also typically has larger lateral dimensions, which reduces the concentration of energy in the surface currents, and hence reduces conductor losses. Indeed, the superconducting waveguides measured in [123] were found to be lower loss than coaxial cable. Finally, waveguide has a cutoff frequency, below which it has no standing or propagating modes. The cutoff frequency is set by the lateral dimensions, and is independent of the length. This adds additional control to the mode structure,

which may be useful when tailoring interactions between the link and the modules. For instance, waveguide filters have been used to couple a resonator strongly to a transmission line without limiting the lifetime of a qubit below the cutoff frequency [71].

4.4 The CZKM protocol

Having described coupled resonators and resonators coupled to a transmission line continuum, we turn now to two approaches one can take to use transmission line as a coupling link between two resonators. These schemes can effect state transfer and generate entanglement between modules. The first uses a line of apparently infinite length, and so the coupling is mediated by a continuum. There is a "sender" resonator which emits a wavepacket into this continuum, and a "receiver" which perfectly absorbs this wavepacket. The other approach, described in Section 4.5, employs a single mode of a finite length of transmission line to enable unitary swapping dynamics between resonators coupled to either end.

4.4.1 The physical system

This first approach was proposed [17] as a way of transmitting the state of an atom through an optical fiber by exchanging the excitation from the atom to an optical cavity at a variable rate $g(t)$. The cavity couples to the fiber at a constant rate κ, and thus the excitation leaks out into the fiber at an effective rate which depends on both g and κ. The control of the effective rate allows for the emission of the cavity to be shaped into a smooth wavepacket. By a similar coupling on the other end of the fiber, this wavepacket can be perfectly absorbed into an atom in another cavity, over arbitrarily long distance. We will explain the protocol in the cQED context, replacing the atoms with resonators, but the essential details are the same.

a

b

Module 1   Module 2
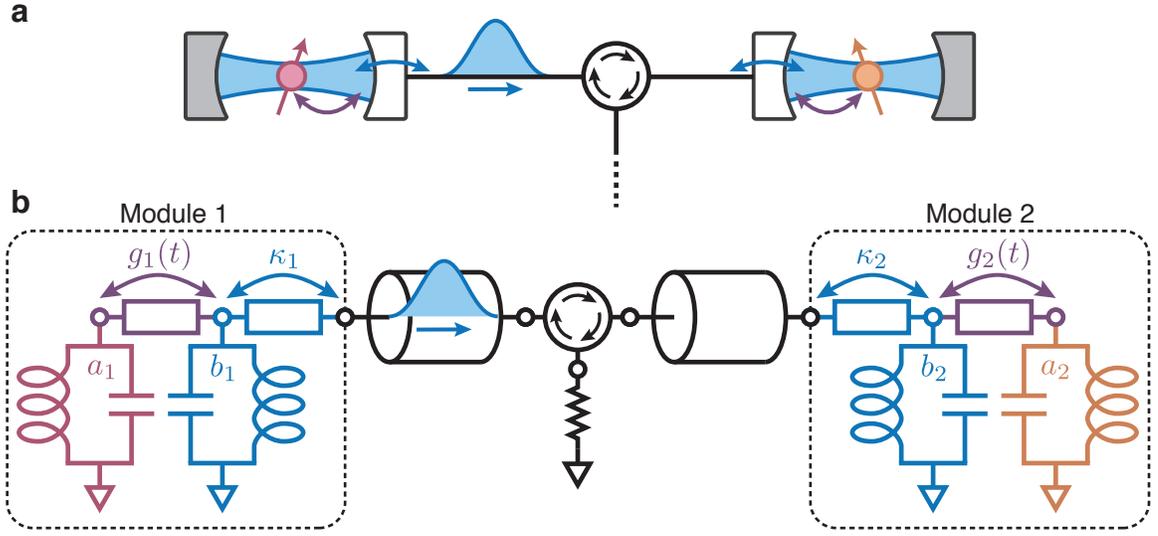
Figure 4.5: Physical layout for the CZKM Protocol. a) State transfer between atoms (red, orange) in cavities (blue) via emission through a directional optical fiber link. Modulated conversion coupling (purple arrows) and static decay (blue arrows, white mirrors) control emission profile. c) cQED realization. The modes a_1 and a_2 are made to couple through a directional channel and local output modes b_1 and b_2. Local couplings g_1 and g_2 are varied in time to release and capture a specified wavepacket shape. Output mode decay rates κ_1 and κ_2 are fixed.

The cQED layout of the CZKM protocol is shown in Figure 4.5. Two modules are connected by a transmission line. Each module, indexed by $k \in \{1, 2\}$, contains a storage resonator a_k and an output resonator b_k. The output resonator couples to the link at a fixed rate κ_k. The storage and output resonators couple with a time-variable rate $g_k(t)$.

The goal is to transfer an arbitrary resonator state from resonator a_1 to resonator a_2, initially assumed to be in the vacuum state. As we mentioned above, we will treat the transmission line as a continuum. For it to act exactly as a continuum, the line would need to be infinitely long[5]. Since there might be loss in the line, we don't actually want an infinite length, just a convenient one. We get around this by inserting a directional element, a circulator, into the line. This directs any emission from or reflections off module 2 into a matched resistor. In this way, from each

---

5. More precisely, long such that the round-trip time is long compared to times corresponding to the coupling rates g and κ, and the wavepacket bandwidth, which we will see below is $\sim 4g^2/\kappa$.

module, the line looks infinitely long, since emission never returns.

4.4.2 Effective coupling rate

We'll begin by considering the dynamics of module 1, since the circulators ensures they are independent of module 2. The coupling between modes a_1 and b_1 is the conversion Hamiltonian in Equation 4.2. The Langevin equations are

$$
\begin{aligned}
\dot{a}_1(t) &= ig_1^*(t)b_1(t) \\
\dot{b}_1(t) &= ig_1(t)a_1(t) - \frac{\kappa}{2}b_1(t).
\end{aligned}
\tag{4.32}
$$

Note that we allow g_1 to be complex. Since module 1 sees a matched input, $b_{\text{in},1} = 0$. The input-output relation then simply reads

$$
b_{\text{out},1}(t) = \sqrt{\kappa}\, b_1(t).
\tag{4.33}
$$

For a generic $g_1(t)$, these equations can be solved numerically. However, these equations admit an analytic solution for constant $g_1(t) = g$. Assuming the storage resonator has initial field amplitude $a_1(0)=a_0$, and the output resonator is empty ($b_1(0)=0$),

$$
\begin{aligned}
a_1(t) &= \frac{a_0}{2\sqrt{1-\frac{\kappa_{\text{eff}}}{\kappa}}}\left[\left(1+\sqrt{1-\frac{\kappa_{\text{eff}}}{\kappa}}\right) e^{-\frac{\kappa t}{4}\left(1-\sqrt{1-\frac{\kappa_{\text{eff}}}{\kappa}}\right)} \right. \\
&\qquad\qquad\left. - \left(1-\sqrt{1-\frac{\kappa_{\text{eff}}}{\kappa}}\right) e^{-\frac{\kappa t}{4}\left(1+\sqrt{1-\frac{\kappa_{\text{eff}}}{\kappa}}\right)}\right] \\
b_1(t) &= \frac{2iga_0}{\kappa\sqrt{1-\frac{\kappa_{\text{eff}}}{\kappa}}}\left[e^{-\frac{\kappa t}{4}\left(1-\sqrt{1-\frac{\kappa_{\text{eff}}}{\kappa}}\right)} - e^{-\frac{\kappa t}{4}\left(1+\sqrt{1-\frac{\kappa_{\text{eff}}}{\kappa}}\right)}\right],
\end{aligned}
\tag{4.34}
$$

where we define $\kappa_{\text{eff}} = 4|g|^2/\kappa$.

Under the approximation $|g| \ll \kappa$, the solution simplifies to

$$a_1(t) = a_0 e^{-\kappa_{\text{eff}} t/2}$$

$$b_1(t) = \frac{2iga_0}{\kappa} \left[e^{-\kappa_{\text{eff}} t/2} - e^{-\kappa t/2} \right], \tag{4.35}$$

which means that the field in storage resonator a_1 decays exponentially with a rate $\kappa_{\text{eff}}/2$, so the energy decays at κ_{eff}. Meanwhile, the field in the output resonator quickly rises, and then decays, also at $\kappa_{\text{eff}}/2$. The emitted field is then

$$b_{\text{out},1}(t) = \sqrt{\kappa} b_1(t) = 2ig\sqrt{\kappa} a_0 \left[e^{-\kappa_{\text{eff}} t/2} - e^{-\kappa t/2} \right]. \tag{4.36}$$

The initial energy content in the mode a_1 is $|a_0|^2$ (in units of photons), and integrating the output field reveals $\int_0^\infty |b_{\text{out},1}|^2 \mathrm{d}t = |a_0|^2$, which means that all of the energy is conveyed through the output resonator and emitted into the transmission line. The complete conversion means we have a chance of transmitting the entirety of the state in a_1 somewhere else, for any initial state, thanks to the linearity of the equations of motion. We just have to be able to absorb this emission somewhere else.

4.4.3 Absorbing reflections

The emission from module 1 will be incident on module 2 as $b_{\text{in},2}(t) = b_{\text{out},1}(t) = \sqrt{\kappa} b_1(t)$. The dynamics for this module are then dependent on what happens in module 1:

$$\dot{a}_2(t) = ig_2^*(t)b_2(t)$$

$$\dot{b}_2(t) = ig_2(t)a_2(t) - \frac{\kappa}{2}b_2(t) - \sqrt{\kappa}\, b_{\text{in},2}(t) \tag{4.37}$$

$$= ig_2(t)a_2(t) - \frac{\kappa}{2}b_2(t) - \kappa b_1(t),$$

with input-output relation

$$b_{\text{out},2}(t) = \sqrt{\kappa}\, b_2(t) + b_{\text{in},2}(t) = \sqrt{\kappa}\, (b_1(t) + b_2(t))\,. \tag{4.38}$$

Unfortunately, the constant conversion coupling considered above in Subsection 4.4.2 produces a wavepacket which is not efficiently absorbed with constant receiver coupling $g_2 = g_1$. This can be seen by considering the time-reverse of the emission process. The profile of the emitted field is a decaying exponential in time, with time constant $\kappa_{\text{eff}}/2$. In the time-reverse picture, we have perfect *absorption* of an incoming wavepacket with the reverse profile, which is a *rising* exponential. For this reason, constant couplings do not provide a good temporal-mode match between sender and receiver.

The observation made by Cirac and co-authors was that, by appropriately tuning $g_1(t)$ and $g_2(t)$, it is possible to enforce the condition $b_{\text{out},2}(t)=0$ for all time, which means that no power is reflected from module 2, and the incoming state is perfectly absorbed. This null reflection results from destructive interference between the incoming filed $b_{\text{in},2}$ with the intra-cavity field b_2:

$$b_2(t)=-b_1(t). \tag{4.39}$$

They solved this problem by choosing the wavepacket $b_{\text{out},1}(t)$ to be time symmetric: $b_{\text{out},1}(-t) = b_{\text{out},1}(t)$. This, taken with the interference condition, also forces b_1 and b_2 to be time symmetric. Then, by enforcing $g_2(t) = g_1(-t)$, they showed that there is a solution to the coupled dynamical equations.

The work of Cirac *et al.* was a critical step towards the first demonstration of communication with optical photons [89]. However, the solution presented therein leaves a little bit to be desired. For starters, this approach does not lend itself very well towards including certain additional Hamiltonian terms induced by the coupling

(see Chapter 5). Additionally, the conversion strength $g_1(t)$ is left at a maximum at the end of the protocol, and $g_2(t)$ is assumed to have been at a maximum for all time before hand. This is related to the fact that the structure of the problem is such that it admits no complete transfer of excitation in finite time. This can be seen in the constant-coupling solution in Equation 4.35, where the occupation of a_1 decays exponentially, but never reaches zero exactly. In practice one must choose a finite time for the protocol, such that the energy emitted is some fraction of the initial energy contained in a_1, which is as large as is practical. For these reasons, we present a related but more complete numerical approach to this problem.

4.4.4 Practical solutions

Guided by the results from [17], we break this problem into two separate pieces, treating the sender and receiver independently. The two subproblems are, given a wavepacket temporal envelope $b_{\mathrm{out},1}(t)$ defined on $t \in [0, T]$,

- find the sender coupling $g_1(t)$ which releases this wavepacket, containing the right amount of energy $\eta_1 = \int_0^T |b_{\mathrm{out},1}(t)|^2 \mathrm{d}t < 1$; and,

- find the receiver coupling $g_1(t)$ which *absorbs* this wavepacket, with efficiency $\eta_2 = 1 - \int_0^T |b_{\mathrm{out},2}(t)|^2 \mathrm{d}t < 1$.

Given the necessary imperfect release and capture efficiencies η_1 and η_2, the maximum transfer efficiency will be $\eta = \eta_1 \eta_2$. The envelope, and the release and capture efficiencies, are chosen by the user, but certain choices may be practical or not.

Release

The solution for the first subproblem involves some gymnastics with Equation 4.32, reproduced here to minimize page turning:

$$\dot{a}_1(t) = ig_1^*(t)b_1(t) \tag{4.40}$$

$$\dot{b}_1(t) = ig_1(t)a_1(t) - \frac{\kappa}{2}b_1(t). \tag{4.41}$$

Since $b_1(t)$ and its derivatives are determined by the outgoing wavepacket, we can regard these as equations for $a_1(t)$ and $g_1(t)$. We take the time derivative of the second equation, and substitute in the first:

$$\begin{aligned}\ddot{b}_1(t) &= i\dot{g}_1(t)a_1(t) + ig_1(t)\dot{a}_1(t) - \frac{\kappa}{2}\dot{b}_1(t)\\ &= i\dot{g}_1(t)a_1(t) - g_1(t)g_1^*(t)b_1(t) - \frac{\kappa}{2}\dot{b}_1(t).\end{aligned} \tag{4.42}$$

Next we multiply this equation by $g_1(t)$

$$g_1(t)\ddot{b}_1(t) = i\dot{g}_1(t)g_1(t)a_1(t) - g_1^2(t)g_1^*(t)b_1(t) - \frac{\kappa}{2}g_1(t)\dot{b}_1(t), \tag{4.43}$$

and solve Equation 4.40 for $ig_1(t)a_1(t)$ to substitute in:

$$\dot{g}_1(t)\left(\dot{b}_1(t) + \frac{\kappa}{2}b_1(t)\right) - g_1^2(t)g_1^*(t)b_1(t) - g_1(t)\left(\frac{\kappa}{2}\dot{b}_1(t) + \ddot{b}_1(t)\right) = 0. \tag{4.44}$$

We have succeeded in eliminating the unknown $a_1(t)$, and have just one equation for $g_1(t)$ in terms of $b_1(t)$ and its derivatives. This is a nonlinear equation for the complex-valued $g_1(t)$ which must in general be solved numerically, with the initial condition on $g_1(0)$ coming from Equation 4.41:

$$g_1(0) = \frac{\dot{b}_1(0) + \kappa b_1(0)/2}{ia_1(0)}. \tag{4.45}$$

Since the output resonator should be empty at the beginning of the protocol, it is wise to choose $b_{\text{out},1}(0) = 0$, or the solution cannot be exact. Additionally, by choosing $\dot{b}_1(0) = 0$, we have $g_1(0) = 0$. For these reasons, we usually use a wavepacket shape like $b_{\text{out},1}(t) \propto \sin^2(\pi t/T)$, which leads to smooth behavior of $g_1(t)$, without a step-like turn-on.

In practice we find that a solution for $g_1(t)$ can be found in a numerically stable way as long as the duration and shape of the wavepacket are chosen appropriately, and the released energy fraction $\eta_1 \lesssim 0.99$. The minimum duration is set both by the maximum attainable conversion rate g_{max}, as well as the bandwidth of the output mode κ. We typically find $1/T \ll g_{\text{max}} < \kappa$. This means the transfer time is slower than any of the dynamical timescales of the problem. This comes from the fact that we are utilizing damping of the communication resonator, which necessitates waiting several time constants to empty it fully.

Capture

Remarkably, the procedure for computing the capture coupling $g_2(t)$ is almost identical to the release. This is because, given the interference condition Equation 4.39, the equations of motion for module 2 are

$$
\begin{aligned}
\dot{a}_2(t) &= ig_2^*(t)b_2(t) \\
\dot{b}_2(t) &= ig_2(t)a_2(t) + \frac{\kappa}{2}b_2(t),
\end{aligned}
\tag{4.46}
$$

which look just like the equations for module 1, but with the sign of κ inverted. Since $b_2(t)$ is specified by the incoming wavepacket, we again can perform the same manipulation to eliminate $a_2(t)$. The difference is that we do not have a useful initial condition, but rather a final condition

$$
g_2(T) = \frac{\dot{b}_2(T) - \kappa b_2(T)/2}{ia_2(T)} \qquad a_2(T) = \sqrt{\eta_2 \int_0^T |b_{\text{in},2}(t)|^2 \mathrm{d}t}. \tag{4.47}
$$

For practical reasons, we then solve the equation for $g_2(t)$ in reverse, from $t = T$ to
0. If the wavepacket is time-symmetric and the decay constants κ are the same in
both modules, this will result in $g_2(t) = g_1(T - t)$, though these conditions are not
necessary.

4.4.5 Entanglement generation

The above procedure enables arbitrarily high fidelity transfer of unknown quantum
states. This is useful for moving qubits around in a modular architecture, and can
be used to generate entanglement between modules if we transmit half of a locally
entangled Bell pair. However, a very minor modification of this protocol enables direct
generation of an entangled state between sender and receiver. After preparing a single
photon in a_1, we can turn on the release process, but tune it so that only half of the
energy is released. This corresponds to specifying $\int_0^T |b_{\mathrm{out},1}(t)|^2 \mathrm{d}t = 0.5$, but can be
done with the same overall temporal envelope for the output field, which requires a
different release control profile $g_1(t)$. This generates an entangled state between a_1
and the propagating wavepacket. By *fully* absorbing the incoming state at module 2,
we can swap that entanglement into a_2, resulting in a single-photon Bell state between
the two modules. A nice feature of this approach is that the capture control $g_2(t)$ is
the same as for state transfer, since the problem is linear in the incoming wavepacket
$b_{\mathrm{in},2}(t)$. Indeed it is slightly easier than state transfer, because the maximum control
amplitude for partial release is smaller than for full state release.

It is worth noting that the general evolution of a state $|\psi\rangle_{\mathrm{s}}$ in the sender after
half-release is that of a 50:50 beamsplitter between the sender and the propagating
wavepacket, initially in $|0\rangle$. After absorption, the results is a 50:50 beamsplitter
between the sender and receiver, provided the initial state of the receiver is vacuum.
For example, with an initial coherent state $|\alpha\rangle_{\mathrm{s}} |0\rangle_{\mathrm{r}}$, the final state is $\left|\alpha/\sqrt{2}\right\rangle_{\mathrm{s}} \left|\alpha/\sqrt{2}\right\rangle_{\mathrm{r}}$,
which is not entangled.

4.5 The quantum bus

The CZKM protocol discussed above has some advantages; namely, it is independent of the length of the link. There are, however, some challenges. For instance, the output modes need to be matched in frequency, within their bandwidth κ, which is usually less than a part in a thousand precision on the frequency. This either presents the fabrication challenge of making identical resonators, or requires *in situ* frequency tunability. Additionally, as we will discuss below, the required circulator is usually lossy, which limits the fidelity of the transfer, and it enforces a directionality of communication, which may limit the flexibility of the interaction. These challenges will be discussed in more detail in the discussion of the experimental implementation in Chapter 6. In this section we present a somewhat simpler approach which uses a single mode of a finite section of line, with no directional element.

The physics discussed here bear a lot of resemblance to the original quantum bus experiments from Yale [90] and NIST [91], but we will consider coupling resonators instead of qubits. Additionally, we'll leave the exact coupling mechanism vague as we did in the last section. The analysis here applies to any coupling which realizes the conversion Hamiltonian in Equation 4.2. As we will see, this approach has some significant implications over the CZKM scheme.

4.5.1 The physical system

The hardware layout for coupling two resonators through a transmission line is sketched in Figure 4.6. The first simplification from the CZKM protocol is that there are no output resonators, reducing the hardware requirements. Instead, we envision modes a_1 and a_2 each directly coupled to a single mode of the line, b, at rate g. The coupling rate can be time dependent, and could differ on the two ends, but the simplest case is a gated coupling (simple on-off) with equal rates. We'll discuss
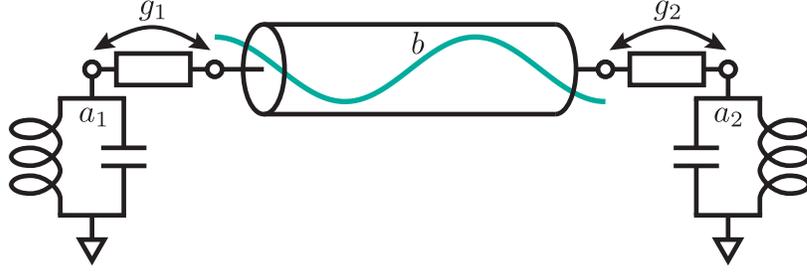
Figure 4.6: The Quantum Bus. Two resonators are coupled to either end of a length of transmission line. The resonators interact via a single mode of the line.

different approaches in Section 4.7. The lack of a need for numerical time-dependent control also simplifies the calibration requirements significantly.

The full Hamiltonian reads

$$\mathcal{H} = iga_1b^\dagger - iga_1^\dagger b - iga_2b^\dagger + iga_2^\dagger b. \tag{4.48}$$

Note the phase of the coupling g is different between the two ends. Here, this is just for convenience — it amounts to redefining the phase of a_2 with respect to a_1. However, in Section 4.7, when we consider multiple modes of the line, this relative phase will be important.

4.5.2 Swapping dynamics

The Heisenberg equations and initial conditions for the field operators under the Hamiltonian in Equation 4.48 are

$$\begin{aligned}
\dot{a}_1 &= -gb & a_1(0) &= A_1 \\
\dot{b} &= ga_1 - ga_2 & b(0) &= 0 \\
\dot{a}_2 &= gb & a_2(0) &= A_2,
\end{aligned} \tag{4.49}$$

where we assume b is in vacuum. The solution is

$$a_1(t) = A_1 \cos^2\left(gt/\sqrt{2}\right) + A_2 \sin^2\left(gt/\sqrt{2}\right)$$

$$a_2(t) = A_2 \cos^2\left(gt/\sqrt{2}\right) + A_1 \sin^2\left(gt/\sqrt{2}\right) \tag{4.50}$$

$$b(t) = \frac{1}{\sqrt{2}}\left(A_1 - A_2\right)\sin\left(\sqrt{2}gt\right).$$

These dynamics come with some interesting features, which we briefly explore here.

Population transfer

As we might expect from the Hamiltonian, population transfers between modes a_1 and a_2, through the bus mode b. If we are using the bus as an intermediary, we probably do not want our operation to leave population therein. Fortunately, for any initial state of the storage resonators, the bus periodically returns to the vacuum state. The first such emptying occurs at $t = \tau_{\text{SWAP}} = \pi/\sqrt{2}g$. At this time, we note that the argument of the oscillatory terms in $a_{1,2}(t)$ is π, and so

$$a_1(\tau_{\text{SWAP}}) = A_2$$

$$a_2(\tau_{\text{SWAP}}) = A_1 \tag{4.51}$$

$$b(\tau_{\text{SWAP}}) = 0,$$

which is exactly a SWAP operation, for any input states.

Dark states

Equation 4.50 has the interesting property of being constant in time if $A_1 = A_2$; that is, if the resonator modes begin in the same state, the bus will be empty at all times. This comes from a destructive interferences of the amplitudes being swapped into the bus from each end. This interference is a classical effect — the same situation happens with three masses coupled with springs in a chain. There is an eigenmode of

the system where both of the outer masses oscillate out of phase, and the forces they exert on the center mass (b) cancel. We call this a "dark state" of the system, and we will explore its properties more in Subsection 4.6.2 when considering dissipation in the bus. This still effects a SWAP at times τ_{SWAP}, since in this case SWAP is equivalent to identity.

We can see the dynamics will be trivial in this case by factoring Equation 4.48:

$$\mathcal{H} = ig(a_1 - a_2)b^\dagger - ig(a_1 - a_2)^\dagger b, \tag{4.52}$$

which is clearly zero on two-mode states $|\psi\rangle_{1,2}$ for which $a_1 |\psi\rangle_{1,2} = a_2 |\psi\rangle_{1,2}$. This is true of certain entangled states like $|10\rangle + |01\rangle$[6], as well as classical and separable ones like $|\alpha\rangle |\alpha\rangle$. This kind of dark state was used in [75]; we will comment more on this in Section 4.8.

4.5.3 Beamsplitters with the bus

As we mentioned in Section 4.2, the SWAP operation between resonators is a particular realization of the beamsplitter transformation, with $\theta = \pi/2$. With direct conversion coupling between two modes, we found we could generate any angle beamsplitter just by changing the interaction time: $\theta = gt$. When using the bus as an intermediary as described in Subsection 4.5.2, this is not the case, since the bus is populated at all times $0 < t < \tau_{\text{SWAP}}$. So to effect a proper two-mode beamsplitter between the resonators, we need an extra control parameter.

Since the beamsplitter rotation is a symmetric object with respect to its operation on the two modes a_1 and a_2 (except for the phase ϕ), we want to introduce a Hamiltonian term which treats a_1 and a_2 the same, and keeps the property that the Hamiltonian is invariant under $a_1 \leftrightarrow a_2$ (modulo a minus sign). We choose to

---

6. But not $|10\rangle - |01\rangle$, which is the one we will prepare later.

introduce a detuning Δ on the bus mode:

$$\mathcal{H} = iga_1 b^\dagger - iga_1^\dagger b - iga_2 b^\dagger + iga_2^\dagger b - \Delta b^\dagger b, \tag{4.53}$$

but otherwise the interaction is unchanged from Equation 4.48. The resulting equations of motion are updated to include

$$\dot{b} = ga_1 - ga_2 + i\Delta b. \tag{4.54}$$

Note that by this convention the detuning has opposite the usual sign. This is just for convenience; either positive or negative detuning will work, just changing some of the acquired phases.

The solution for the dynamics is a bit complicated, so we choose to write just the expression for the bus:

$$b(t) = \frac{g}{\sqrt{2}J}(A_1 - A_2)e^{i\Delta t/2}\sin\left(\sqrt{2}Jt\right) \qquad J \equiv g\sqrt{1 + \frac{\Delta^2}{8g^2}}, \tag{4.55}$$

where we have defined the effective coupling rate J in analogy to detuned vacuum Rabi oscillations. Now the amplitude of the population in the bus depends on the ratio of g and Δ, and is suppressed for $\Delta \neq 0$. However, the bus retains its periodic behavior, and there is still a time for which it is empty. We call this the beamsplitter time,

$$\tau_{\text{BS}}(\Delta) = \frac{\pi}{\sqrt{2}J} = \frac{2\pi}{\sqrt{8g^2 + \Delta^2}}. \tag{4.56}$$

At the times when the bus is eliminated, the dynamics must effect a two-mode beamsplitter on a_1 and a_2. Indeed, we find

$$a_1(\tau_{\text{BS}}) = e^{-i\theta}\left(A_1\cos\theta + iA_2\sin\theta\right)$$
$$a_2(\tau_{\text{BS}}) = e^{-i\theta}\left(A_2\cos\theta + iA_1\sin\theta\right), \tag{4.57}$$

which is the beamsplitter transformation given in Equation 4.1, with beamsplitter angle

$$\theta = \frac{\pi}{2}\left(1 - \frac{\Delta}{\sqrt{8}J}\right) = \frac{\pi}{2}\left(1 - \frac{\Delta}{\sqrt{8g^2 + \Delta^2}}\right) = \frac{\pi}{2}\left(1 - \frac{\Delta\,\tau_{\mathrm{BS}}}{2\pi}\right), \tag{4.58}$$

and phases $\phi = -\theta$, $\varphi = \pi/2$, and $\zeta = 0$. For instance, the useful 50:50 beamsplitter ($\theta = \pi/4$) is obtained at $\Delta = \sqrt{8/3}\,g$. At this detuning, $J = g\sqrt{4/3}$ and the interaction time is $\tau_{\mathrm{BS}} = \sqrt{3/8}\,\pi/g = \sqrt{3/4}\,\tau_{\mathrm{SWAP}}$.

This approach has a few nice features. For a given coupling strength g, we can enact an arbitrary beamsplitter angle by correctly choosing the detuning and time, with $\theta \in [0, \pi]$. Since the effective Rabi rate is faster for larger detuning, smaller angle beamsplitters take less time than the resonant SWAP. As we'll explore more in the next section, the population of the bus is suppressed by the detuning, which will make these operations a little less sensitive to loss in the bus. Finally, this beamsplitter can generate entangled states from single-photon input states. We'll demonstrate this for two different input states in Chapter 7. For instance, we will demonstrate Hong-Ou-Mandel interference. This kind of symmetric interference cannot occur in the CZKM protocol because of the directionality of the communication channel.

4.5.4 Large-detuning limit

We here consider an important limit of the quantum bus, which is for $\Delta \gg g$. We observed above that Equation 4.55 shows the field amplitude in the bus is suppressed by the detuning. As we will discuss in Section 4.6, this can be advantageous when the bus is lossy, and this the regime used in the original Yale experiment [90]. For now, we'll stick with the lossless case and just understand the dynamics.

Motivated by the observation that the occupation of the bus is suppressed as g/Δ when the detuning is large, we can approximate the equations of motion by

adiabatically eliminating the bus entirely. This is the approximation $\dot{b}(t) = 0$, which gives us

$$b = \frac{ig_1}{\Delta}a_1 + \frac{ig_2}{\Delta}a_2. \tag{4.59}$$

In Subsection 4.5.2, we set $g_1 = -g_2 = g$ (assumed real), but we will keep these as independent complex parameters for this discussion because it will be important later. Substituting this expression into the equations of motion Equation 4.51, we have

$$\begin{aligned}
\dot{a}_1 &= \frac{ig_1^* g_2}{\Delta}a_2 - \frac{i|g_1|^2}{\Delta}a_1 \\
\dot{a}_2 &= \frac{ig_1 g_2^*}{\Delta}a_1 - \frac{i|g_2|^2}{\Delta}a_2.
\end{aligned} \tag{4.60}$$

These are the same equations we had in Equation 4.3, with effective rate $g_{\text{eff}} = -g_1 g_2^*/\Delta$, which give rise to arbitrary-angle beamsplitter transformations with $\theta = g_{\text{eff}}t$.

There are additional detuning (Stark shifts) induced by the off-resonant coupling to the bus, but as long as $|g_1|^2 = |g_2|^2$, these are the same, and the interaction is still resonant in the right frame.[7] Assuming g_1 and g_2 are real, the time dynamics according to Equation 4.4 are

$$\begin{aligned}
a_1(t) &= A_1 \cos\left(\frac{g_1 g_2}{\Delta}t\right) + iA_2 \sin\left(\frac{g_1 g_2}{\Delta}t\right) \\
a_2(t) &= A_2 \cos\left(\frac{g_1 g_2}{\Delta}t\right) + iA_1 \sin\left(\frac{g_1 g_2}{\Delta}t\right).
\end{aligned} \tag{4.61}$$

In this limit, often dubbed the *virtual Raman* regime, we have an effective two-mode coupling, which, as we saw, generates arbitrary beamsplitter angles simply by tuning the time. The cost of this elimination is that the interaction is slowed down by a factor for g/Δ as compared to the resonant case. We will discuss this trade-off in Subsection 4.7.1.

---

7. Even if the couplings are not exactly matched, this coupling can still generate a 50:50 beam-splitter. The main effect is that the maximum angle is somewhat less than $\pi/2$, so we can't quite get to a full SWAP in this case.

4.6 Loss in the link

Having established a few ways in which we can use transmission line to couple distant resonators, we now turn to the issue of loss in the line. So far, all of the inter-module quantum communication experiments in cQED have been limited in large part by loss, either in the transmission line itself, of in the connection and directional elements [45, 71, 73, 16, 74, 75]. For the transfer of excitations, this has resulted in an energy efficiency (as defined in Chapter 3) of $\eta = 0.6 - 0.85$. While we hope to use bosonic error correction to improve the fidelity beyond that enabled by the efficiency, it is important to understand the contributions to this loss. We will examine the role of loss in the scenarios described above.

4.6.1 Loss in the CZKM protocol

Since the "pitch and catch" approach involves only a single pass of a wavepacket through the link, it is straightforward to model the inefficiency due to transmission loss. We simply modify the relationship between the outgoing wavepacket at module 1 and the incoming wavepacket at module 2: $b_{\text{in},2}(t) = \sqrt{\eta_t}\, b_{\text{out},1}(t)$. This way, the energy which leaves module 1 is reduced by a factor of η_t when it arrives.

What is the typical source of this loss? In the three experiments which realized this protocol [73, 16, 74], the length of superconducting coaxial cable between modules was about one meter[8]. Based on independent measurements of the attenuation constant in this commercial cable, which can be of order 1 km, the measured efficiency is likely not a result of attenuation in the cable itself. However, the required circulator *is* expected to be somewhat lossy. Though datasheets supplied by the vendor quote this loss to be of order 5% or less, it is difficult to calibrate insertion loss to this precision for cryogenic devices. Moreover, these measurements are typically performed at 4 K

---

8. The authors of [74] appear to have used non-superconducting cable, but it is not completely clear.

or above. It is known that loss of materials at microwave frequencies can worsen at lower temperatures or lower input powers. The ferrite core that gives these devices the required non-reciprocity is expected to be the main source of loss.

Additionally, the performance of a circulator can depend on the exact impedance seen from the third port, which is assumed to be perfectly matched. In a realistic setting, there will be some impedance mismatches on this port, potentially leading to incoming photons being directed the wrong way through the circulator and being lost. Even a circulator which is internally lossless may suffer from this problem. Finally, the connectors both at the circulator and at the modules can introduce some loss. The insertion loss of coaxial connectors is not well-known in this setting; in [74] this is estimated to be 2-3% each, including losses in the printed circuit board holding the sample.

It bears noting that the connections can cause loss without being dissipative themselves. Impedance mismatches at the modules, at the circulator, and even downstream after the third port of the circulator can harm the isolation of the circulator, and cause some amount of signal energy to go from port 1 to port 3, instead of to port 2 where it will make it to the receiver. While this isolation is indicated by the manufacturer to be of order 1%, these impedance mismatches may emerge as components contract at low temperatures, and may result in frequency-dependent isolation issues, which would be difficult to measure and resolve. Impedance mismatch is not a problem in the quantum bus, where it will serve to slightly modify the mode frequencies and couplings. This does not cause inefficiency as it does in the CZKM protocol, as there is no place for energy to escape as a result.

All told, these experiments find transmission efficiencies of around $\eta_\text{t} \approx 0.8$. In addition, as discussed in Subsection 4.4.4, there are inherent limitations on releasing and capturing the wavepacket, which result in energy left in the sender, and reflected from the receiver, respectively. These can also be a few percent each. Imperfections in

calibration of the control pulse shapes can further limit the transfer, as can dephasing of the storage and receiver resonators. The net result of all these factors results in an effective single-pass loss η which limits the fidelity of state transfer and entanglement generation.

State transfer

We can view this loss as the amplitude damping channel discussed in Section 2.3 and Section 3.2. For transfer of states encoded in single photon, we examined the mean state fidelity versus efficiency in Subsection 3.2.2. The result is that the mean infidelity goes as the inefficiency: $1-\mathcal{F} \approx (1-\eta)/3$. This means that for the losses found, the expected fidelity of state transfer is $\mathcal{F} \approx 0.9$ in the absence of other limitations.

Entanglement generation

As discussed in Subsection 4.4.5, entanglement can be generated between the two modules by half-release and capture of a single photon from module 1. The resulting ideal Bell state is $|O_+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. However, since the loss acts on the sent portion of the state, and not on what is left behind, the resulting state is asymmetric. We can model this simply as writing the state $|O_+\rangle$ and acting the loss channel with efficiency η on the second resonator. The resultant density matrix in the 0-1 subspace is

$$
\begin{aligned}
\rho_{\mathrm{f}} &= E_0 \rho_{\mathrm{i}} E_0^\dagger + E_1 \rho_{\mathrm{i}} E_1^\dagger \\
&= \frac{1}{2} \left(\eta |01\rangle\langle 01| + \sqrt{\eta} |01\rangle\langle 10| + \sqrt{\eta} |10\rangle\langle 01| + |10\rangle\langle 10| \right) \\
&\quad + \frac{(1-\eta)}{2} |00\rangle\langle 00| \\
&= \frac{(1+\eta)}{2} |O_{+,\eta}\rangle\langle O_{+,\eta}| + \frac{(1-\eta)}{2} |00\rangle\langle 00| ,
\end{aligned}
\tag{4.62}
$$

where we define

$$|O_{+,\eta}\rangle = \frac{1}{\sqrt{1+\eta}}\left(\sqrt{\eta}\,|01\rangle + |10\rangle\right), \tag{4.63}$$

which is a normalized pure state, equal to the Bell state $|O_+\rangle$ for $\eta = 1$.

The final density matrix in Equation 4.62 is a mixture of two pure states. The second component is the separable state $|00\rangle$ resulting from losing the photon, with probability $p_{\mathrm{jump}} = (1-\eta)/2$. This makes sense, as the probability of losing a photon is half the probability per photon (the inefficiency), since on average one half of the photon is sent. This vacuum state has zero overlap with the ideal Bell state.

The other component in Equation 4.62 is the no-jump case, $|O_{+,\eta}\rangle$, which is entangled, but not maximally so. This has fidelity to the ideal bell state of

$$\mathcal{F}_{\mathrm{nojump}} = |\langle O_+|O_{+,\eta}\rangle|^2 = \frac{(1+\sqrt{\eta})^2}{2(1+\eta)}, \tag{4.64}$$

with probability $p_{\mathrm{nojump}} = (1+\eta)/2$. However, the infidelity in this case is only second order in the inefficiency.

The resulting infidelity of the state after loss is

$$1 - \mathcal{F}_{\mathrm{Bell}} = 1 - p_{\mathrm{nojump}}\mathcal{F}_{\mathrm{nojump}} = 1 - \frac{1}{4}(1+\sqrt{\eta})^2 \approx \frac{1}{2}(1-\eta) = p_{\mathrm{jump}}, \tag{4.65}$$

which is a bit worse than the infidelity for state transfer.

4.6.2   Loss in the quantum bus

We now investigate the effect of loss in the quantum bus approach. Instead of talking about propagation lengths, we talk about decay rates, since we are using a standing-wave mode, and make the connection back later in this section.

We assume the bus resonator has an energy decay rate κ_{b}, which can be included

in the equations of motion by including the damping terms:

$$\dot{b} = ga_1 - ga_2 + i\Delta b - \frac{\kappa_b}{2}b. \tag{4.66}$$

This saps energy from the system at a rate proportional to $\kappa_b|b|^2$. Accordingly, the total amount of loss incurred will depend on the occupation of the bus, and the length of the protocol.

Resonant transfer

To understand the energy efficiency of the resonant quantum bus, we'll consider the case of one-way transfer. It's sufficient to consider the case of a coherent state $|\alpha\rangle$. This lets us work in the classical picture. In the language of Subsection 4.5.2, this is $A_1 = \alpha, A_2 = 0$. As long as the population oscillations are under-damped $(\kappa_b < 4\sqrt{2}g)$, the population in the bus will still periodically be eliminated, and the dynamics are modified from (4.50). In the limit of weak damping,

$$a_1(t) = \alpha e^{-\kappa_b t/4} \cos^2\left(\tilde{g}t/\sqrt{2}\right) + \frac{\alpha}{2}\left(1 - e^{-\kappa_b t/4}\right)$$
$$a_2(t) = \alpha e^{-\kappa_b t/4} \sin^2\left(\tilde{g}t/\sqrt{2}\right) + \frac{\alpha}{2}\left(1 - e^{-\kappa_b t/4}\right) \tag{4.67}$$
$$b(t) = \frac{\alpha}{\sqrt{2}} e^{-\kappa_b t/4} \sin\left(\sqrt{2}\tilde{g}t\right),$$

where the oscillation frequency is loaded by the damping: $\tilde{g} = g\sqrt{1 - (\kappa_b/4\sqrt{2}g)^2}$. Even with a fairly large amount of damping $(\kappa_b \approx g)$, this frequency loading is less than a 2% effect, so it has a very small impact on the duration τ_{SWAP}. As a result, we can neglect the distinction between g and $\tilde{g}$. In fact, the weak damping approximation is still good even when κ_b is as large as g. The field decay rate is $\kappa_b/4$ instead of the usual $\kappa_b/2$ because the bus mode is only occupied for half the transfer time on average.

At the first swap time we have

$$a_1(\tau_{\text{SWAP}}) = \frac{\alpha}{2}\left(1 - e^{-\kappa_b\tau_{\text{SWAP}}/4}\right)$$

$$a_2(\tau_{\text{SWAP}}) = \frac{\alpha}{2}\left(1 + e^{-\kappa_b\tau_{\text{SWAP}}/4}\right) \tag{4.68}$$

$$b(\tau_{\text{SWAP}}) = 0.$$

The energy efficiency for state transfer with the quantum bus is

$$
\begin{aligned}
\eta_{\text{bus}} &= \frac{|a_2(\tau_{\text{SWAP}})|^2}{|\alpha|^2} \\
&= \frac{1}{4}\left(1 + e^{-\kappa_b\tau_{\text{SWAP}}/4}\right)^2 \\
&\approx 1 - \frac{\kappa_b\tau_{\text{SWAP}}}{4}.
\end{aligned}
\tag{4.69}
$$

Since $\tau_{\text{SWAP}} = \pi/\sqrt{2}g$, we have the inefficiency as a function of g and κ_b:

$$1 - \eta_{\text{bus}} \approx \frac{\pi}{\sqrt{32}}\frac{\kappa_b}{g} \approx 0.55\frac{\kappa_b}{g} \tag{4.70}$$

This efficiency plays the same role as in the CZKM protocol, as this is also an amplitude damping channel. This gives us a way to compare the two approaches — in principle, whichever has a higher efficiency should yield higher fidelity state transfer. Of course, other sources of infidelity will appear, but we may hope these are less costly than the link loss.

An alternate way to compute the loss is to note that the rate at which energy decays from the three-mode system is $\kappa_b|b(t)|^2$. The total energy loss is then the integral thereof, and the inefficiency is the ratio of that energy loss to the initial

energy

$$
\begin{aligned}
1 - \eta_{\mathrm{bus}} &= \frac{1}{|\alpha|^2} \kappa_{\mathrm{b}} \int_0^{\tau_{\mathrm{SWAP}}} |b(t)|^2 \mathrm{d}t \\
&= \kappa_{\mathrm{b}} \int_0^{\tau_{\mathrm{SWAP}}} \frac{1}{2} e^{-\kappa_{\mathrm{b}} t/2} \sin^2\left(\sqrt{2}\tilde{g}t\right) \mathrm{d}t \\
&\approx \frac{\kappa_{\mathrm{b}}}{2} \frac{\tau_{\mathrm{SWAP}}}{2} \\
&= \frac{\pi \kappa_{\mathrm{b}}}{\sqrt{32}g},
\end{aligned}
\tag{4.71}
$$

where we assumed again that the oscillations are under-damped, ignoring the exponential in the integral. This is valid to leading order in κ_{b}/g.

Saturation behavior

The dynamics of Equation 4.67 are pretty interesting. We see that, at long time, the amplitude in the storage modes does not decay to zero, but rather saturates at one half. In terms of energy, there is one quarter of the original energy in each mode, the other half having dissipated in the bus. This is due to the presence of a dark mode, as discussed in Subsection 4.5.2. Symmetric superpositions of field in a_1 and a_2 are stationary and do not populate the bus, and hence do not decay. Antisymmetric superpositions will decay completely, and are called "bright." The initial state $|\alpha\rangle |0\rangle$ is halfway between bright ($|\alpha\rangle |\alpha\rangle$) and dark ($|\alpha\rangle |-\alpha\rangle$), and hence decays only halfway.[9]

Another result of this is that there is a little energy left behind in the original resonator after a full SWAP. The inefficiency contribution from this effect is pretty small. To leading order, the energy fraction left behind is $\frac{\pi^2}{128}\left(\frac{\kappa_{\mathrm{b}}}{g}\right)^2$, which is much smaller than the inefficiency due to loss, and can usually be neglected.

---

9. Here, "halfway between" means it is a superposition of bright and dark in the *classical* sense of the word — think wave mechanics or eigenmodes of coupled pendula.

Entanglement generation

As described in Subsection 4.5.3, by detuning the bus we can introduce an arbitrary angle beamsplitter. We will use a 50:50 beamsplitter ($\theta = \pi/4$) on a single photon input $|1\rangle|0\rangle$ to produce a maximally entangled state in Chapter 7. As a reminder, the conditions for the 50:50 beamsplitter are $\Delta = \sqrt{8/3}\,g$, $J = \sqrt{4/3}g$, and $\tau_{\mathrm{BS}} = \sqrt{3/8}\pi/g$. Here we compute the effect of dissipation on the state creation.

First, we can calculate the expected energy loss. We do this as we did above for resonant transfer, by integrating the dissipated power in the bus. There is a slight complication, though. Since the initial state is not classical, and has no well-defined phase, working with the field operators directly is a bit tricky. However, we can look at Equation 4.55, which gives the expectation value of $b(t)$ for initial coherent states A_1 and A_2. In this case, clearly $A_2 = 0$. We'll take the magnitude squared:

$$|b(t)|^2 = \frac{g^2}{2J^2}|A_1|^2 \sin^2\left(\sqrt{2}Jt\right). \tag{4.72}$$

If we interpret $|A_1|^2 = 1$ as the initial mean photon number in mode a_1, then the above equation gives us what we need to calculate the dissipated power[10]. In analogy to Equation 4.71, we have

$$
\begin{aligned}
1 - \eta_{\mathrm{BS}} &= \kappa_{\mathrm{b}}\frac{g^2}{2J^2}\int_0^{\tau_{\mathrm{BS}}} \sin^2\left(\sqrt{2}Jt\right)\mathrm{d}t \\
&= \frac{\kappa_{\mathrm{b}}g^2}{2J^2}\frac{\tau_{\mathrm{BS}}}{2} \\
&= \frac{\kappa_{\mathrm{b}}g^2}{2J^2}\frac{\pi}{g}\sqrt{\frac{3}{8}} \\
&= \frac{3\sqrt{3}\pi}{8\sqrt{32}}\frac{\kappa_{\mathrm{b}}}{g} \\
&\approx 0.36\frac{\kappa_{\mathrm{b}}}{g},
\end{aligned}
\tag{4.73}
$$

---

10. This is really an abuse of this framework. We said A_1 was the complex amplitude of $\langle a_1 \rangle$, which is zero for a Fock state. Here we are interpreting $|A_1|^2$ as $\langle a^\dagger a \rangle$

which is less loss than the resonant swap because the process is faster and populates the bus less.

Since there is only one excitation, the state which results from loss is $|00\rangle$. So the above inefficiency results in a mixture of the ideal Bell state and the vacuum state:

$$\rho_{\mathrm{f}} = \eta_{\mathrm{BS}} \, |O_+\rangle\langle O_+| + (1 - \eta_{\mathrm{BS}}) \, |00\rangle\langle 00| \qquad (4.74)$$

The second term has zero fidelity to the ideal state, so

$$1 - \mathcal{F}_{\mathrm{Bell,bus}} = 1 - \eta_{\mathrm{BS}} = \frac{3\sqrt{3}\pi}{8\sqrt{32}} \frac{\kappa_{\mathrm{b}}}{g} \approx 0.36 \frac{\kappa_{\mathrm{b}}}{g}, \qquad (4.75)$$

at least to first order in κ_{b}/g. The qualitative behavior is a bit different than in the CZKM protocol, because the loss results in a symmetric state, whereas when the loss happened in transit through a continuum, the final state was biased, with the receiver being less occupied than the sender

Comparison to CZKM

It is worth noting that, if the loss is dominated by dissipation in the transmission line itself, then the bus will perform significantly worse than the CZKM protocol. This is because, in this case, $\kappa_{\mathrm{b}} = 2\mathrm{SPL}\,\mathrm{FSR}$ (from Equation 4.23). The state transfer inefficiency is

$$1 - \eta \approx \frac{\kappa_{\mathrm{b}}}{2g} = \frac{2\mathrm{SPL}\,\mathrm{FSR}}{2g} = \mathrm{SPL}\frac{\mathrm{FSR}}{g} \gg \mathrm{SPL}. \qquad (4.76)$$

Since the use of a single-mode bus assumes $g \ll \mathrm{FSR}$, the inefficiency is always larger than the SPL. One qualitative way to see this is that the state takes multiple passes through the line, so it must incur the SPL many times. For example, in the experiment in Chapter 7, we have $g/2\pi = 560\,\mathrm{kHz}$ and $\mathrm{FSR} = 1.9\,\mathrm{GHz}$, so the loss-enhancement ratio is $\mathrm{FSR}/g \approx 500$.

However, as we have noted and will see in Chapter 6, the loss in the CZKM protocol is probably not dominated by the line loss, but rather the circulator and connectors. In this case, removing the circulator loss will provide a net gain, even though it increases the sensitivity to line loss. Fortunately, as discussed in Subsection 4.3.4, we have reason to believe the SPL through a meter of superconducting transmission line could be made be as low as 10^{-6}. Even with a loss-enhancement of 500, this could still allow an energy efficiency $\eta \approx 0.999$.

4.7 Suppressing loss

In Chapter 7 we will show an implementation of a quantum bus with SPL of around 2×10^{-3}, but the limited coupling strength g results in an inefficiency of around 0.1. In this situation, we might ask what we can do to improve the transfer for fixed parameters g and κ_{b}. Here we will discuss two approaches for suppressing the inefficiency directly, and one for improving the infidelity for fixed inefficiency.

4.7.1 Virtual Raman

As discussed in Subsection 4.5.4, the bus can be used in the far-off resonant regime, with detuning $\Delta \gg g$. The situation is diagrammed in Figure 4.7a. In this case, we have an effective direct interaction between modes a_1 and a_2. The adiabatic elimination of the bus allows for a nice compact representation of the effective equations of motion for the resonators, but we still need to compute the occupation of b to predict the loss rate. As per Equations 4.59 and 4.61, and assuming $g_1 = -g_2 = g$,

$$
\begin{aligned}
b(t) &= \frac{ig}{\Delta} \left(a_1(t) - a_2(t) \right) \\
&= \frac{ig}{\Delta} \left(A_1 - A_2 \right) \cos\left(\frac{g^2}{\Delta} t \right) - i \left(A_1 - A_2 \right) \sin\left(\frac{g^2}{\Delta} t \right) \\
|b(t)|^2 &= \frac{g^2}{\Delta^2} |A_1 - A_2|^2.
\end{aligned}
\tag{4.77}
$$
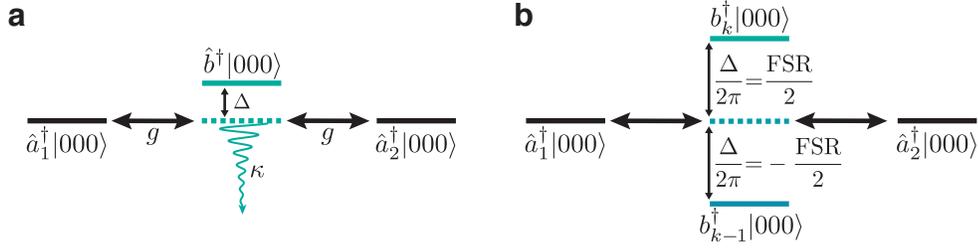
Figure 4.7: Virtual Raman via the Bus. a) Off-resonant coupling through a single bus mode. b) When the detuning is not small compared to the FSR, coupling to multiple modes must be considered.

The only thing that changes when we have $g_1 = +g_2 = g$ is that the relative phase between A_1 and A_2 becomes a sum instead of a difference. This is conveniently constant in time (by assumption).

Let's consider exchanging an excitation from a_1 to a_2; so, $A_1 = \alpha, A_2 = 0$. In this limit, the decay rate of energy from the system is $\kappa_{\text{eff}} = (g/\Delta)^2 \kappa_{\text{b}}$. This leads to oscillations of the excitation which decay at this rate. The effective swap rate $g_{\text{eff}} = g^2/\Delta$, so the SWAP time is $\pi/2g_{\text{eff}} = \pi/2/(g^2/\Delta)$, and the inefficiency after one swap is

$$1 - \eta_{\text{VR}} \approx \frac{\pi}{2g_{\text{eff}}}\kappa_{\text{eff}} = \left(\frac{\pi}{2}\frac{\Delta}{g^2}\right)\left(\frac{g^2}{\Delta^2}\kappa_{\text{b}}\right) = \frac{\pi\kappa_{\text{b}}}{2\Delta}. \tag{4.78}$$

Since we are assuming $\Delta \gg g$, this inefficiency is much smaller than that attained going resonantly through the bus:

$$1 - \eta_{\text{VR}} \approx \frac{\pi\kappa_{\text{b}}}{2\Delta} = \frac{\pi\kappa_{\text{b}}}{2g}\left(\frac{g}{\Delta}\right) \ll \frac{\kappa_{\text{b}}}{2g} \approx 1 - \eta_{\text{bus}}. \tag{4.79}$$

So, although it takes longer, there is a net gain by using the virtual transfer because the decay rate is suppressed by more than the swap time is increased.

Breakdown

Equation 4.78 makes it look like we can suppress the inefficiency as much as we want by making Δ larger and larger. This is not the case, since there is a maximum

detuning, $\Delta = 2\pi\text{FSR}/2 = \pi\text{FSR}$. This is because, by detuning more than this, we will be getting *closer* to the next mode. The best we can do is use this maximum detuning, at which we are equidistant from modes k and $k+1$, as in Figure 4.7b.

We might actually worry that in this case, the effective couplings mediated by each mode will cancel, since the detunings have opposite sign. However, due to the alternating sign of the eigenmodes (see Figure 4.3), the relative sign of g_1 and g_2 will alternate with k. For example, if mode k is even ($g_1^{(k)} = g_2^{(k)} = g$), then mode $k+1$ is odd ($g_1^{(k)} = -g_2^{(k)} = g$). Thus the effective coupling rates are

$$
\begin{aligned}
g_{\text{eff}}^{(k)} &= \frac{g^2}{\Delta_k} = \frac{g^2}{2\pi\text{FSR}/2} = \frac{g^2}{\pi\text{FSR}} \\
g_{\text{eff}}^{(k+1)} &= \frac{-g^2}{\Delta_{k+1}} = \frac{-g^2}{-2\pi\text{FSR}/2} = \frac{g^2}{\pi\text{FSR}},
\end{aligned}
\tag{4.80}
$$

which have the same magnitude *and* sign. The effective couplings thus add, and we should have $g_{\text{eff}} = 2g^2/\pi\text{FSR}$. However, since we populate both modes k and $k+1$, the effective decay rate is $\kappa_{\text{eff}} = 2(g/\Delta)^2\kappa_{\text{b}}$, and the inefficiency at max detuning is

$$
1 - \eta_{\text{VR,max}} \approx \frac{\pi}{2g_{\text{eff}}}\kappa_{\text{eff}} = \left(\frac{\pi}{2}\frac{\pi\text{FSR}}{2g^2}\right)\left(2\left(\frac{g}{\pi\text{FSR}}\right)^2\kappa_{\text{b}}\right) = \frac{\kappa_{\text{b}}}{2\text{FSR}},
\tag{4.81}
$$

which is *exactly* the single pass loss of the line, as shown in Subsection 4.3.4.

A more detailed calculation would consider all the modes of the line, since the coupling through the next nearest modes $k-1$ and $k+2$ are of the same order and *oppose* the coupling through k and $k+1$ considered here. Additionally, adding the decay rates is a bit of a slight of hand, because as we add more modes, the spatial pattern of field in the line becomes more complex (and wavepacket-like), and so we need to think about the interference effects here to really get this right. But the point is that we can't beat the single pass loss with the virtual Raman regime.

4.7.2 STIRAP

An alternate approach to suppress the loss also comes from atomic physics. In a simple three-mode (or three-level) system like the quantum bus, it is possible to transfer population from one resonator to the other continuously, while the population of the lossy intermediate mode is exactly zero. This is accomplished by modulating the couplings g_1 and g_2 slowly in time, and is called stimulated Raman by adiabatic passage, or STIRAP [126]. This allows for in-principle perfect population transfer, even if the intermediate mode is quite lossy.

One might think that STIRAP could be applied in this system to completely avoid the loss in the bus. Indeed it can be used to improve resonant transfer, as discussed in [75]. However, a detailed calculation including the effect of off-resonant coupling to the neighboring modes shows that, while the occupation of mode k can be eliminated, the coupling to $k \pm 1$ occupies those modes, contributing some loss. This calculation was done in [127], and numerics show that the minimum inefficiency is exactly the single pass loss.

4.7.3 Bosonic error correction

Virtual Raman and STIRAP are potentially useful approaches to improve the efficiency. However, both come at the cost of significant increase in the length of the protocol. In realistic scenarios where other sources of decoherence become more pronounced at longer times, this is not always practical way to enhance the fidelity.

A complementary approach, which is compatible with resonant transfer as well as STIRAP and the virtual Raman regime, is to encode a qubit in multi-photon states. By using any of the bosonic error-correcting codes discussed in Chapter 3, we can measure the photon number parity after transfer. If a single photon was lost in transit, we can detect and correct this loss. We will detail our efforts towards this in Chapter 6 and Chapter 7.

4.8   Related experiments

In addition to the experiments mentioned above, which realized the CZKM protocol around the same time, there were recently two other cQED experiments which are relevant to this discussion, both from Chicago.

The group of Andrew Cleland recently performed an experiment [128] which probed the crossover from standing modes to propagating photons. The experimental layout was similar to the one shown in Figure 4.6, but with two transmon qubits taking the place of the end-point resonators. Additionally, the entire experiment was located on a single chip, with a 78 cm coplanar waveguide transmission line connecting the qubits. The free spectral range of the line was 79 MHz. The coupling between the qubits and the line was a rapidly tunable inductor. When the coupling was kept much smaller than the FSR ($g/2\pi \approx 5$ MHz), the individual modes of the can be resolved, and they used one of them to transfer excitations and generate entanglement between the qubits. However, as the coupling was turned to its maximum rate of $g/2\pi = 47$ MHz, the dynamics lose their frequency structure, and the coupling is as if to a continuum. They show that by shaping the coupling in time, they are able to release and absorb something like a wavepacket, even in this regime where the coupling is comparable to the FSR. The results of this work make nicely clear the connection of these two seemingly different physical dynamics. As the coupling is increased, the superposition of multiple modes gives rise to a wavepacket in the line in precisely the same formalism used to describe this connection in Section 4.3.3.

An experiment performed in David Schuster's group used another approach to couple qubits on separate chips [75]. Again, the topology was similar to Figure 4.6, but with an important distinction. Since the transmission line (normal metal coaxial cable) was not particularly low-loss, the modes of the line were too low-Q to facilitate entanglement generation, and also would have induced too much dissipation on the qubits themselves. The authors added a resonator on each chip between the qubit

and the cable. In this sense, the layout resembles that in Figure 4.5, but with no circulator in the line. These communication resonators were made to be frequency-degenerate with one another, and hybridized strongly with a nearby mode of the cable. This three-mode hybrid system has one mode which has no participation in the lossy cable, and instead lives only in the on-chip resonators, and is significantly higher-Q than the bare cable modes. They used this "dark mode" as a quantum bus for entanglement generation between the qubits, avoiding the two lossy "bright modes". Since the rate of coupling from the qubit to the dark mode was much less than the FSR of the cable, or of the splitting of the dark and bright modes, the dynamics to good approximation are those discussed in Section 4.5.

The emergence of the dark mode is essentially the same physics pointed out in Subsection 4.5.2, which give rise to hybrid dark states of the three-mode system. While a simple analysis suggests the dark mode should be perfectly lossless, the authors explain that the presence of other nearby cable modes prevent this from being exactly true. A more careful analysis shows that the off-resonant coupling to other cable modes prevents the dark mode from being totally lossless, and the population transfer through it is limited by the single pass loss of the cable.

4.9   Conclusion

In this chapter we introduced most of the theoretical underpinnings of the communication and entanglement protocols used in Chapters 6 and 7. We began with the physics of coupled oscillators and the generation of beamsplitter transformations. We discussed the behavior of finite lengths of transmission line, and the crossover from standing modes to a continuum when coupling a resonator to the line. We also described some loss mechanisms in transmission line and how they imprint themselves on the damping rates of the modes therein.

We then described the CZKM protocol for communication and entanglement via propagating photons, and gave a framework for generating the time-dependent controls needed for release and capture of wavepackets. In Chapter 6 we will add some system-dependent implementation details critical to our realization of this protocol. We also introduced the quantum bus as a way to couple distant resonators, both for SWAP and entangling beamsplitter operations. Finally, we discussed the effects of loss in the communication link on these schemes, and some ways towards suppressing the resulting infidelity.

In the next chapter, we will provide a description of the hardware and couplings used to implement the schemes described here, before turning to the results in Chapters 6 and 7.

Chapter 5

3D cQED Hardware and
Techniques

This chapter will introduce the physical hardware for the experiments in Chapters 6 and 7, including the basic implementation of resonators and qubits in three-dimensional circuit quantum electrodynamics (3D cQED). Many more details of the hardware discussed here can be found in [15]. Section 5.1 will discuss the kinds of 3D resonators we use. In Section 5.2 we will discuss the problem of controlling a linear resonator, and describe the transmon, the source of nonlinearity and quantum control and measurement in our experiments. This section will describe the origin of the dispersive coupling between transmons and resonators. Then Section 5.3 we will describe the basic 3D cQED module, with a long-lived storage cavity, transmon, and readout resonator. Here we will also describe the coupling to the outside world.

Having discussed the physical hardware and couplings, we then turn to measurement and manipulation of these devices. For excellent discussions of calibration and tuneup of the operations here, see [13, 14]. We will mostly focus on details which have not been described in detail elsewhere. In Section 5.4 we discuss control and measurement of transmons with readout resonators strongly coupled to the environment.

Section 5.5 will describe measurement and control of 3D cavities as error-correctable quantum memories. Section 5.6 will turn to the final kind of control, used extensively here — parametric conversion. We'll give a general description of how new Hamiltonian terms can be selectively enabled through off-resonant driving. We'll then give the contexts in which we employ this kind of coupling, and some of the advantages of this kind of coupling, as well as some challenges.

Then in Section 5.7 we will discuss how we can measure a transmon which does not have a readout resonator, but is instead coupled to a memory cavity [129]. Finally, Section 5.8 will discuss the coupling of a module to a quantum bus resonator housed in coaxial cable.

5.1 Superconducting resonators

In this section, we will describe three kinds of 3D resonators used in this thesis.

5.1.1 Coaxial cable resonators

Perhaps the simplest realization of a superconducting resonator is a section of coaxial transmission line with open boundary conditions, which, as discussed in Chapter 4, hosts an infinite set of equally-spaced modes. The circuit model and fundamental mode voltage profile is shown in Figure 5.1a. Since all modes of this structure have voltage maxima at the ends, they are particularly easy to couple to capacitively.

One can easily construct this kind of resonator out of coaxial cable. There are a few commercial vendors of semi-rigid NbTi coax with PTFE dielectric, pictured in Figure 5.1b, including Coax Co. and Keycom, in the UT-85 form factor. This kind of cable has been used for low-loss transmission of signals from the $20\,\mathrm{mK}$ base plate of a dilution refrigerator up to the $4\,\mathrm{K}$ stage, where commercial amplifiers can be used. NbTi is a superconductor of choice for this application because it has a
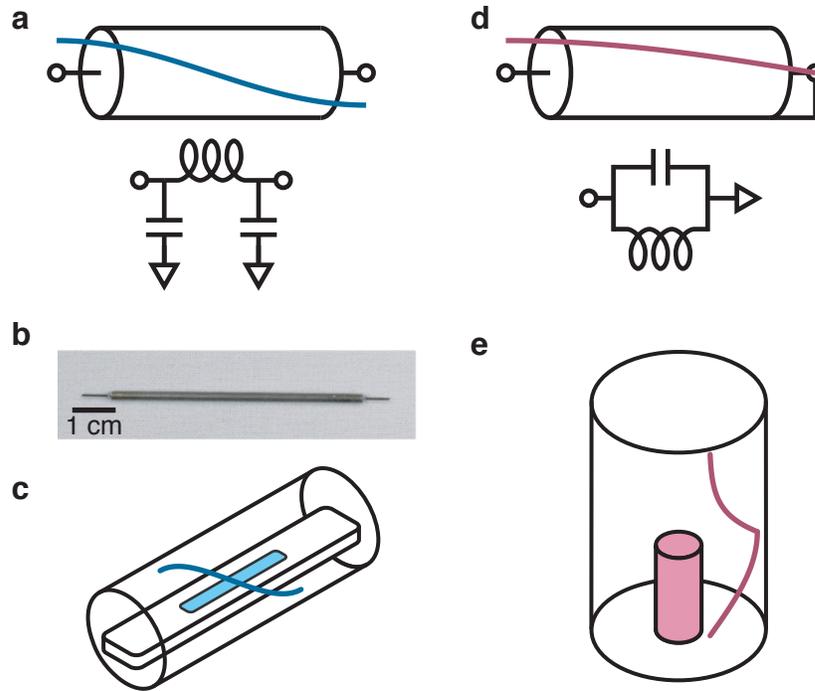
Figure 5.1: 3D Resonators. a) Schematic and circuit for fundamental mode of an open-ended transmission line resonator. Voltage profile in blue. b) NbTi coaxial cable resonator with outer conductor and dielectric removed at ends for coupling. c) Schematic of on-chip stripline resonator, made of aluminum trace (blue) on sapphire chip in aluminum tunnel. Blue line is electric field amplitude. d) Schematic and circuit for fundamental mode of coaxial transmission line resonator with one open end and one shorted end. Voltage profile in red. e) Schematic of 3D post cavity resonator. Red line is electric field amplitude profile, which decays exponentially in the waveguide section (top).

critical temperature around $10\,\mathrm{K}$. Indeed, it has been found that this kind of cable is extremely low-loss, and suitable for transmission of quantum signals of fairly long distances. In particular, this kind of cable can house modes with quality factors as high as 10^5 [123], which implies the attenuation length is as long as $1\,\mathrm{km}$. This method, using standing-wave modes of sub-meter lengths of cable, allows accurate measurement of small losses, since the quality factor measures the loss of many passes through the transmission line. See Section 5.8 for our characterization of this type of cable.

5.1.2 Stripline resonators

A ubiquitous player in 3D cQED experiments is the $\lambda/2$ stripline resonator, formed by a thin-film aluminum strip on a sapphire chip, in a bulk aluminum tunnel, as depicted in Figure 5.1c. This is just another open-ended transmission line, with the on-chip portion forming the center conductor and the tunnel playing the role of the ground. This is somewhat similar to the coplanar waveguide $\lambda/2$ resonators used extensively in planar cQED, but the large mode-volume and diffuse electromagnetic fields allow for internal quality factors 5-10 times higher in the 3D case. We typically use the fundamental mode in the 7 to 10 GHz range, which places the next harmonic well above 12 GHz. Due to the high dielectric constant of sapphire (ϵ_r =10–11), the speed of light is slowed, and a 9 mm stripline has a fundamental frequency around 9 GHz.

While the stripline is a fairly compact solution for an on-chip resonator, we often desire multiple elements on the same chip, and we wish to make the resonator take up less linear space. As a result, we sometimes further shorten the resonator by adding meanders, as will be shown in Subsection 5.3.3. The resulting increased inductance and capacitance lowers the frequency for fixed end-to-end length, or reduces the length of fixed frequency.

Straight striplines can have quality factors from 10^6 to 10^7 [122, 15]. Meandering striplines tend to store more energy in lossy surface dielectrics, and so may have somewhat worse internal quality factors. However, as we will discuss in Section 5.3, we typically intentionally limit the quality factor of these modes by coupling to an output transmission line to use them for transmon readout, so the internal quality factor is not critical.

5.1.3 3D post cavities

The final linear element in our arsenal is the $\lambda/4$ post cavity, sketched in Figure 5.1d,e. These resonators resemble a section of transmission line which is *shorted* at one end — the stub makes up the center conductor. Since there is no dielectric, the speed of light is that in vacuum, so we get a 6.5 GHz fundamental for a stub length of 11 mm. Since the next harmonic is at three times the fundamental, we can make post cavities with frequencies as low as ~3.5 GHz without causing frequency crowding issues.

Machined from a solid piece of aluminum, the post cavity stores the vast majority of its energy in vacuum, with no places of high confinement that we have with on-chip devices. Additionally, the seamless design has no points of large current losses. The result is a highly coherent resonator. Post cavities made from high-purity aluminum (99.99% pure or higher) and chemically treated [62] can have quality factors in excess of 10^8 [63]. This makes them ideal memory elements, and these cavities form the backbone of our experiments.

There must of course be a seam somewhere, since the cavity is machined and then closed up. However, we include a long continuation of the outer conductor above the post. With no center conductor, this acts as a length of waveguide. The lowest cutoff frequency in a circular waveguide of radius $r = 5$ mm belongs to the TE_{11} mode, and is at $f_{\text{cutoff}} \approx 0.3c/r \approx 18$ GHz [121]. As a result, the fields and currents of the fundamental mode, which is well below cutoff, fall off exponentially going up the waveguide. We design the housing to be long enough that current flowing across the seam at the end of the waveguide does not limit the quality of the mode at the level we need. For $r = 5$ mm, the field decay length is mostly frequency-independent below 10 GHz, and is about 3 mm. In practice we use waveguide lengths of around 3 cm. In fact, the coupling to the TE_{11} should be very small by symmetry. The attenuation length is closer to 2 mm, and is set by the TM_{01} mode, which has a cutoff frequency

about 30% higher than the TE_{11} mode [121, 130].

5.2 The transmon: introducing nonlinearity

Given the linear elements described above, how do we manipulate and measure them?
Here we will answer this question, introducing an additional key circuit element —
the transmon. In Subsection 5.2.1 we describe the oscillator control problem — with a
drive on a harmonic oscillator, one can only move the state around in phase space, and
cannot create quantum states or encode and manipulate a qubit. In Subsection 5.2.2,
we will describe the Josephson junction, a nonlinear inductor, which we can use to
make an *anharmonic* oscillator with an inherent nonlinearity that enables nontrivial
quantum control. Subsection 5.2.3 will describe some of the spectral features which
allow it to be easily manipulated like a qubit. Then in Subsection 5.2.5 we will give
a brief introduction to the formalism we use to talk about coupling of a transmon
to one or more resonators. A complete treatment of the transmon is found in the
original proposal [131], and the "black-box quantization" (BBQ) description of the
coupling of oscillators to one or more Josephson junctions which is favored here is in
[132]. We give the basic idea behind BBQ in Subsection 5.2.2.

5.2.1 Oscillator control with drives

We continue the discussion of oscillators in Chapter 3 to see why it is nontrivial
to produce the kinds of quantum states discussed there. We can introduce a time-
dependent control field to a resonator by irradiating it with microwave light near its
resonance frequency, in which case the Hamiltonian is[1]

$$\mathcal{H} = \omega_0 a^\dagger a + i\epsilon(t)a e^{-i\omega t} - i\epsilon^*(t)a^\dagger e^{i\omega t} \tag{5.1}$$

---

1. I have dropped the zero-point term $\frac{1}{2}\omega$ from the oscillator Hamiltonian, because it has no
consequences for the dynamics.

with complex drive amplitude $\epsilon(t)$. It is convenient to transform our Hamiltonian into the frame of this drive with the unitary transformation $U(t) = e^{i\omega a^\dagger a t}$, which cancels out the rapidly vary phase on the drive and yields

$$\mathcal{H} = \Delta a^\dagger a + \epsilon(t)a + \epsilon^*(t)a^\dagger \tag{5.2}$$

with detuning $\Delta = \omega_0 - \omega$.

The drive can clearly induce transitions between neighboring photon number states, moving up and down the Fock state ladder. However, since all of these transitions are at the same frequency, we have no control over the relative rates, and this Hamiltonian can only result in a very specific kind of evolution. In particular, for $\Delta = 0$, the propagator associated with this Hamiltonian is

$$U(t) = \hat{D}(\alpha(t)) = e^{-\alpha(t)a + \alpha^*(t)a^\dagger} \qquad \alpha(t) = \int_0^t \epsilon(t')\mathrm{d}t \tag{5.3}$$

This *displacement operator* has the property of simply shifting the state in phase space — it takes coherent state $|\beta\rangle \to |\beta + \alpha(t)\rangle$. This means that one cannot create quantum states like Fock or cat states with a simple drive. The consequence of this is that no two-dimensional subspace of a harmonic oscillator can easily serve as a qubit, since there is no natural way to perform an X gate by Rabi flopping, for instance.

5.2.2 The Josephson Junction

The key to creating nonlinear circuitry in cQED is the Josephson junction, an element made by separating two superconducting islands with a weak link (in our case, a few nanometer-thick insulating barrier) across which electrons tunnel in pairs. The zoology of Josephson junction qubits is a rich and historic field [133, 134, 135, 136, 131, 137], but in this work we'll mostly treat the junction as a special non-linear inductor which can replace the linear one supposed in our LC circuit in Figure 5.2a [132]. In
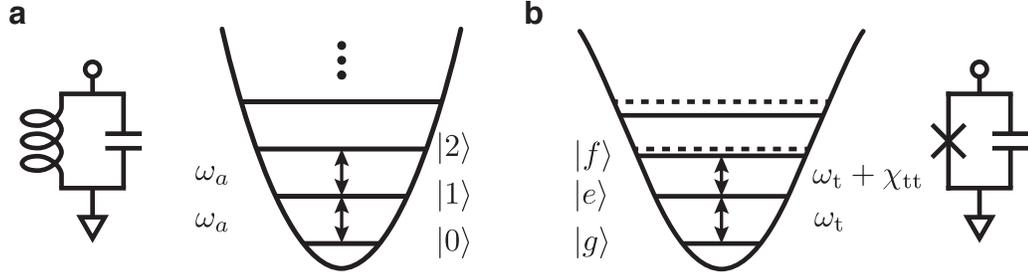
Figure 5.2: The Transmon as an Anharmonic Oscillator a) LC circuit and harmonic energy levels in quadratic potential. b) Nonlinear LC resonator with anharmonic spectrum in quadratic potential with negative quartic term. Energy levels are unequally spaced, with anharmonicity $\chi_{\mathrm{tt}} < 0$.

particular, the inductance of the Josephson is a function of the flux across it, which yields a nonlinear response. For convenience, we introduce the dimensionless phase $\varphi = \phi/\phi_0$, where $\phi_0 = \hbar/2e$ is the reduced flux quantum. The effective inductance of the junction is $L(\varphi) = \frac{L_{\mathrm{J}}}{\cos \varphi}$, where the Josephson inductance L_{J} is the "small signal" inductance. The Josephson inductance is a design parameter, set at time of device fabrication, though by combining two junctions in parallel, L_{J} can be tuned with a static magnetic field. We stick to fixed-inductance in this work for experimental convenience.[2]

In the language of Equation 3.1, replacing the linear inductor with a Josephson junction results in a Hamiltonian of the form[3]

$$\mathcal{H} = 4E_{\mathrm{C}}\hat{n}^2 - E_{\mathrm{J}} \cos \hat{\varphi}$$
$$= \omega_{\mathrm{t}} t^\dagger t - E_{\mathrm{J}} \cos_{\mathrm{nl}} \left(\varphi_{\mathrm{t}}(t + t^\dagger) \right) \tag{5.4}$$

with Josephson energy $E_{\mathrm{J}} = \phi_0^2/L_{\mathrm{J}}$, capacitive charging energy $E_{\mathrm{C}} = e^2/2C$, and

---

2. Introducing this kind of *in situ* tunability also comes with sensitivity to magnetic noise, which can adversely affect device coherence.

3. Here we use $\hat{n} = 2\hat{q}$, which is the number of Cooper pairs which have tunneled across the junction. In this simple parallel configuration, this is the same as the charge on the capacitor, but this is not generally the case. $\hat{n}$ is the correct variable to use in this context, as it obeys the canonical commutation relation with φ.

frequency $\omega_t = \sqrt{8E_J E_C}$. The non-linear cosine is the Taylor expansion of cosine starting at fourth order:

$$\cos_{nl}(x) = \cos(x) - \left(1 - \frac{1}{2!}x^2\right) = \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \dots \tag{5.5}$$

Finally, note the introduction of φ_t, the zero-point fluctuation of the phase. This is a measure of the per-photon phase drop across the junction, and it depends on the characteristic impedance of the resonance [132]; here $\varphi_t = \phi_0\sqrt{\hbar Z_t/2} = \phi_0\sqrt{\hbar/2\sqrt{L/C}}$, with $\phi_0 = \hbar/(2e)$ the reduced flux quantum.

5.2.3 The Transmon Spectrum

In the last line of Equation 5.4, we continue to use harmonic oscillator ladder operators, but call them $t, t^\dagger$, for "transmon." This is only a good approximation when the nonlinear terms are sufficiently small, and the oscillator is still largely harmonic. Quantitatively, this is the case when $E_J \gg E_C$ (typically at least 50 times larger) [131]. In this regime, the qualitative behavior is that of an oscillator with a small quartic term in the potential, as shown in Figure 5.2b. By ignoring higher order terms of the Taylor expansion, and working in the rotating wave approximation (RWA), we have

$$\begin{aligned}
\mathcal{H} &= \hbar\tilde{\omega}_t t^\dagger t - \frac{1}{4}E_J\varphi_t^4 t^{\dagger 2}t^2 \\
&= \hbar\tilde{\omega}_0 t^\dagger t - \frac{\hbar\chi_{tt}}{2}t^{\dagger 2}t^2
\end{aligned} \tag{5.6}$$

This fourth-order term introduces a change in the energy spectrum, and to leading order the eigenenergies of Equation 5.6 are $E_n = n\tilde{\omega}_t - n(n-1)\chi_{tt}/2$, with anharmonicity $\chi_{tt} \equiv E_J\varphi_t^4/2 \approx E_C \sim 2\pi \times 100 - 300$ MHz.

The deformation of the energy levels results in transition frequencies

$$\nu_{n-1,n} = (E_n - E_{n-1})/\hbar = \tilde{\omega}_0 - \chi_{tt}(n-1) \tag{5.7}$$

which are distinct for each n. This solves the control problem of the harmonic oscillator discussed in Subsection 5.2.1 — as long as the drive strength ϵ is weaker than the anharmonicity χ_{tt}, one can selectively address any transition between $|n\rangle$ and $|n+1\rangle$ with a single drive. This allows the operation of the first two levels $|0\rangle, |1\rangle$ of the transmon as a qubit. We often refer to the first three levels as $|g\rangle$, $|e\rangle$, and $|f\rangle$, for "ground," "excited," and "f is another letter near g and e." The fundamental frequency $\tilde{\omega}_0$ is also shifted from the "bare" harmonic frequency $\omega_0 = 1/\sqrt{LC}$ by the deformation of the potential (of order χ_{tt}).

A few comments before we move on. The energy spectrum of an isolated transmon can be predicted exactly in the model here. However, for more complicated circuits, which we employ shortly, it is much more convenient to solve the energy spectrum numerically. This makes it much easier to extract the effect of higher orders of non-linearity, non-RWA effects, and the spectrum in the presence of multiple oscillators. Finally, this discussion has completely ignored the offset charge (discussed at length in [131]). The entire point of the transmon is to suppress the effects of the offset charge, but it is important to check that one is operating in the "good transmon" regime, with $E_{\text{J}}/E_{\text{C}}$ sufficiently large. If not, noisy environmental charges cause the transmon frequency to fluctuate unacceptably. Some handy numerical results can be found in [13], but this dephasing mechanism is usually negligible as long as the relative anharmonicity $\chi_{\text{tt}}/\omega_0 \lesssim 5\%$.

5.2.4 3D transmons

The transmons used in this experiment are pattered on double-polished sapphire in a single electron-beam lithography step [138]. The capacitance consists of two large, roughly rectangular, co-planar pads, 100 to 500 µm wide and 1 to 2 mm long. The capacitor pads are connected by a single Josephson junction, fabricated with double-angle shadow evaporation. The devices in Chapter 6 were made with the bridge-free

technique for junction fabrication [139], while those in Chapter 7 were made using the Dolan bridge method [140].

5.2.5   Dispersive coupling for transmons and oscillators

In the previous subsection we introduced the transmon as a solution to the problem of control for a harmonic oscillator. Now we will explain how coupling between transmons and harmonic oscillators can be easily described using the BBQ formalism [132] already introduced, and how an oscillator coupled to the environment (as in Subsection 4.3.3) can be used to measure the transmon state. Then in the next two sections we will return to harmonic oscillators as storage elements for quantum information, and explain how a transmon can be used to manipulate and measure that information.

The treatment of the junction as a weakly non-linear inductor [132] allows for easy simulation of the transmon and any resonators to which it is coupled in a classical linear circuit modeler or finite-element package. We use ANSYS HFSS for this. The idea is to solve the classical linear circuit for the eigenfrequencies ω_k and the associated zero-point phases φ_k. Then each mode we wish to consider (including the transmon) appears in the Hamiltonian, extending Equation 5.4:

$$\mathcal{H} = \sum_k \omega_k a_k^\dagger a_k - E_{\mathrm{J}} \cos_{\mathrm{nl}} \left(\sum_k \varphi_k (a_k + a_k^\dagger) \right), \tag{5.8}$$

from which the non-linear properties can be computed. It is possible to write down the frequencies and zero-point phases for a lumped element circuit model with linear couplings, but the beauty of the BBQ method is that a 3D finite element eigensolver like the one in HFSS doesn't care about your circuit model — all it knows how to do is find the eigenmodes. This is because the quantities ω_k and φ_k are classical and linear. These quantities can be extracted from the modeler as in the original BBQ

method, which computes the impedance as seen by an imagined port in parallel with
junction, or from the electric and magnetic field distributions of the eigenmodes, as
in [141].

Either way, one extracts the parameters in Equation 5.8, and diagonalizes the
Hamiltonian. Typically, as long as the modes are sufficiently far detuned, it is then
sufficient to map the resulting numerical model onto the Hamiltonian

$$\mathcal{H} = \sum_k \omega_k a_k^\dagger a_k + \sum_k \chi_{kk} a_k^{\dagger\,2} a_k^2 + \sum_{k \neq l} \chi_{kl} a_k^\dagger a_k a_l^\dagger a_l \qquad (5.9)$$

This Hamiltonian, which treats transmons and oscillators as the same kind of
object, has a few obvious features. First, there is another round of dressing of the
frequencies due to the presence of nonlinear terms. We also have the anharmonicity,
or self-Kerr, for each mode ($\chi_{kk} = -\frac{1}{2} E_J \varphi_k^4$). As mentioned above, the self-Kerr of a
transmon-like mode is usually a few hundred MHz, and the self-Kerr of a resonator-
like mode will range from $1 - 50$ kHz in this thesis. Finally, we have the nonlinear
couplings, the cross-Kerrs, $\chi_{kl} = -E_J \varphi_k^2 \varphi_l^2$, between each pair of modes. These can
be read as "the frequency shift of mode k per photon in mode l." These cross-Kerrs,
or *dispersive shifts*, are the critical resource for measurement and operations, and are
usually on the order of 1 MHz between a resonator and a transmon. Cross-Kerrs
between two resonators, inherited from their mutual coupling to the junction of a
transmon, are usually small enough to be neglected, though they can be a resource
[142] or a nuisance [65]. The "typical" values given for these quantities are for the
work in this thesis, but different applications demand different dispersive shifts, and
they can range over several orders of magnitude throughout the literature.

5.3 3D cQED modules

Now that we have introduced the players, we turn to the question of putting them together into a simple quantum information processing module. In this section we'll describe the basic layout of a 3D cQED module. Subsection 5.3.1 will give a cartoon of the most basic such kind of module. Then in Subsection 5.3.2 we'll discuss how we selectively couple to one mode of the package to create a fast readout resonator. Finally, in 5.3.3, we'll diagram the actual hardware used in Chapter 7.

5.3.1 Integrating cavities and qubits

As we have discussed, we need a nonlinear element like a transmon coupled to our post cavity in order to perform quantum operations on it. Moreover, we need some kind of way to measure the state of that transmon, rapidly and with high fidelity.

In this thesis, we use extensively the integration scheme introduced in [122], diagrammed in Figure 5.3a. This architecture consists of a housing machined from a single piece of 99.99% pure aluminum which is chemically etched to remove surface damage and contaminants [62]. The housing contains a central post cavity with a circular waveguide topper, as described in Subsection 5.1.3. Intersecting this cavity near the top of the post is a tunnel, usually 3-4 mm in diameter. In the tunnel we place a sapphire chip, 1-3 mm wide and 20-30 mm long. The chip carries a 3D transmon and a stripline resonator for state readout.

The effective circuit diagram showing the couplings between the modes of the module is in Figure 5.3b. In particular, the size and spacing of the pads making up the transmon set its capacitance, a significant fraction of which is to the ground of the tunnel. Then, the placement of the chip with respect to the cavity post sets the coupling capacitance between the transmon and cavity, and hence the dispersive coupling χ_{at}. The top of the post is where the cavity electric field is largest, so this
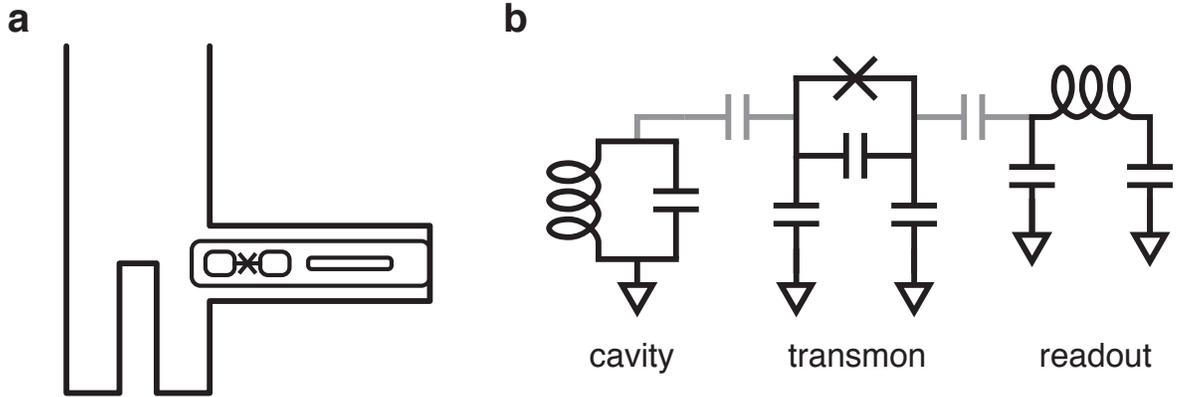
Figure 5.3: Cavity-Transmon-Readout Integration. a) Basic 3D cQED module with post cavity, transmon, and stripline readout resonator. b) Circuit model for a). Coupling capacitances in gray.

is an efficient place to couple. In practice, shifting the chip further in or out of the cavity is a good way to control this capacitance.

Next, the coupling of the transmon to its readout resonator is set by the geometry on chip. The coupling capacitance, and consequently the dispersive shift χ_{rt}, can be set quite precisely by adjusting the separation at the time of fabrication.

5.3.2 Coupling to the world

Now we discuss the coupling of the module to the outside world via input and output signal lines. We have been using more or less the same kind of coupling in the lab since the first 3D devices [138], which is an SMA connector to free-space launcher with an interchangeable pin. This mechanism is discussed at length in [122, 15]. The coupler pin extends into a small tunnel with diameter chosen to form a $50\,\Omega$ match with the pin diameter. The length of the pin can be chosen to vary the coupling quality factor Q_{c} to a given mode over three or four orders of magnitude. The ideal coupling rate depends on the application.

Drive pins

For signal delivery to any of the modes, we choose a coupling rate which does not limit the lifetime to the level of the intrinsic quality Q_i. In practice, Q_c is usually chosen to be at least a few times larger than the expected Q_i. For a transmon, this should be $Q_c \sim 10^7$. For post cavities with $Q_i \sim 10^8$, the coupling must be even weaker. These couplings must however be sufficient to deliver enough drive amplitude to effect fast transmon manipulation (see Subsection 5.4.2) and cavity displacements.

The readout resonator

In order to extract information from the system to perform transmon readout, we require a mode with an intentionally large coupling to the environment κ as defined in Subsection 4.3.3. If we hope to measure the transmon state much faster than its lifetime, we need a readout resonator which can emit a signal in a time $1/\kappa \lesssim 500$ ns. In the language of quality factors, this corresponds to $Q_c = \omega_r/\kappa \lesssim 3 \times 10^4$. So we see that the ability to vary the coupling over five orders of magnitude using a single technology is extremely useful.

5.3.3 The quantum bus

Here we have a look at the samples used in Chapter 7, shown in Figure 5.4. In addition to the features of the basic module discussed in Subsection 5.3.1, we point out a few new components.

The conversion transmon

One of the added components is a second transmon for coupling the cavity to a quantum bus, which we refer to as the conversion transmon (or converter), in the sense of parametric conversion to be discussed in Section 5.6. For differentiation, the transmon discussed so far is called the *ancilla*, and it is used to manipulate and measure
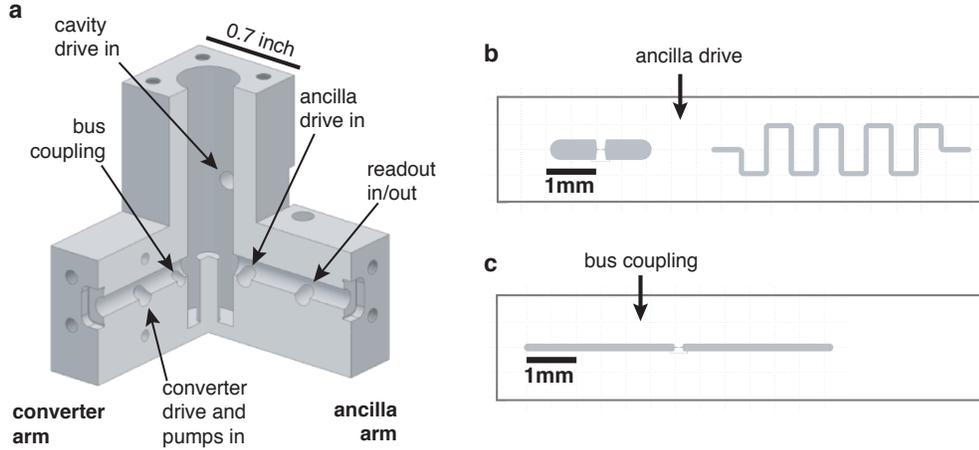
Figure 5.4: Quantum Bus Sample. a) Three-quarter cross section of module used in Chapter 7. Post cavity (center) intersected by two tunnels. Labels indicate coupling pin locations. b) Ancilla chip layout shows transmon pads (left) and meander stripline readout (right). Label indicates approximate location of ancilla drive port. Readout port and stripline Purcell filter, located further to the right, not shown. b) Conversion transmon chip layout. Label indicates approximate location of bus coupling pin. Drive port, located further to the right, not shown.

the cavity. The converter is housed on a separate chip in its own tunnel, orthogonal to the one housing the ancilla, and is "blind" in that it does not have a readout resonator. This sample is in this respect a reduced version of the one used in [129], which had three such blind transmons, and measurement of this transmon (discussed in Section 5.7) is a simplified version of a protocol presented in that reference.

The bandpass Purcell filter

The details of dispersive readout are found in Subsection 5.4.3, but the general rule is we also usually want a dispersive shift between the ancilla transmon and readout of $\chi_{\mathrm{rt}} \approx \kappa$. Unfortunately, strong coupling to a dissipative mode induces an unwanted dissipation on the ancilla. This channel, known as Purcell decay, induces in the simplest approximation a decay rate on the ancilla $\kappa_{\mathrm{P}} = \kappa \chi_{\mathrm{rt}}^2 / \chi_{\mathrm{tt}}^2$. Since we want κ and χ_{rt} to be large, but the maximum χ_{tt} is set by charge dispersion limitations, we are in a bind.

Fortunately, it is possible to engineer a filter which breaks this Purcell relationship. By placing an additional resonator between the readout mode and the coupling pin, we can effectively make the impedance seen by the sample strongly frequency dependent. The original "Purcell filter" [143] acted like a band-stop filter at the transmon frequency. This works, but it limits the filtered region to small bandwidth, and adds modes to the device near the transmon frequency, which may contribute to spectral crowding. An alternate approach, explored in [144, 145, 122] and employed here, is the band-pass Purcell filter. When the filter resonator is placed near in frequency to the readout resonator, it "passes" at the readout frequency, allowing large κ. However, the qubit is far-detuned from the filter, and so is less able to spontaneously emit out the coupling port. This leads to a more favorable relation between κ_P and κ. An alternate approach is the high-pass waveguide Purcell filter used in [71]. This works quite well thanks to the sharp cutoff of rectangular waveguide, though it adds quite a bit of volume to the sample.

5.4 Control and measurement of transmons

We will describe operation of these samples from the outside in, beginning with manipulation and measurement of the ancilla transmon before moving on to the storage cavity in Section 5.5, the conversion transmon in Section 5.7, and the quantum bus in Section 5.8.

5.4.1 Microwave drives

We introduced the Hamiltonian of a drive on a mode a in Subsection 5.2.1:

$$\mathcal{H}_\text{drive} = \omega_\text{a} a^\dagger a + i\epsilon(t)ae^{-i\omega t} - i\epsilon^*(t)a^\dagger e^{i\omega t} \tag{5.10}$$

In order to source such a drive, with arbitrary phase and amplitude control over $\epsilon(t)$, we generate an intermediate-frequency (IF) signal with the correct envelope $\epsilon(t)$ from a pair of digital-to-analog converters (DACs). The two DACs produce outputs $\text{Re}\left(\epsilon(t)e^{-i\omega_{\text{IF}}}\right)$ and $\text{Im}\left(\epsilon(t)e^{-i\omega_{\text{IF}}}\right)$. The DACs are housed on an Innovative Integration X6-1000M board, which generates the IF signals with a 2 ns sampling rate. The IF signals are typically centered around $\omega_{\text{IF}}/2\pi = -120$ to $120\,\text{MHz}$, with up to $250\,\text{MHz}$ bandwidth. The IF pair is up-converted with an IQ mixer (Marki IQ0307 or IQ0618) using a continuous-wave microwave local oscillator (LO, Vaunix LabBrick LMS-103 or Agilent/Keysight MXG) at $\omega_{\text{LO}}/2\pi = 3$ to $10\,\text{GHz}$. This results in a fully phase and amplitude controlled signal centered at $\omega_{\text{RF}} = \omega_{\text{LO}} + \omega_{\text{IF}}$, where, again, ω_{IF} can be chosen positive or negative.

Drive signals are typically filtered and attenuated both at room temperature and in the cryostat. The experimental wiring used in Chapter 7 is shown in Figure 5.5.

5.4.2   Transmon rotations

We effect rotations on the transmon with pulses sourced as described above. The difference between rotations about the X and Y axes is just the complex phase of $\epsilon(t)$. We do not typically have direct control about the Z axis, which can be introduced with serial combinations of X and Y rotations, though this is not needed in this thesis. For fast transmon rotations, we typically use Gaussian-shaped pulses $\epsilon(t) \propto e^{-t^2/2\sigma^2}$, truncated at 4σ in total width, with $\sigma \leq 10\,\text{ns}$. For manipulation in the $\{|e\rangle, |f\rangle\}$ manifold, we use the same drive chain to source pulses detuned by the qubit anharmonicity, which is within the bandwidth of our DACs. This requires a positive ω_{IF} for $\{|g\rangle, |e\rangle\}$ rotations, and a negative one for $\{|e\rangle, |f\rangle\}$ rotations.

Cavity displacements are sourced in the same manner, with similar pulse shapes. Readout drives were generated this way for Chapter 6 but with square envelopes. In Chapter 7 we instead gated a CW source at the readout frequency with a microwave
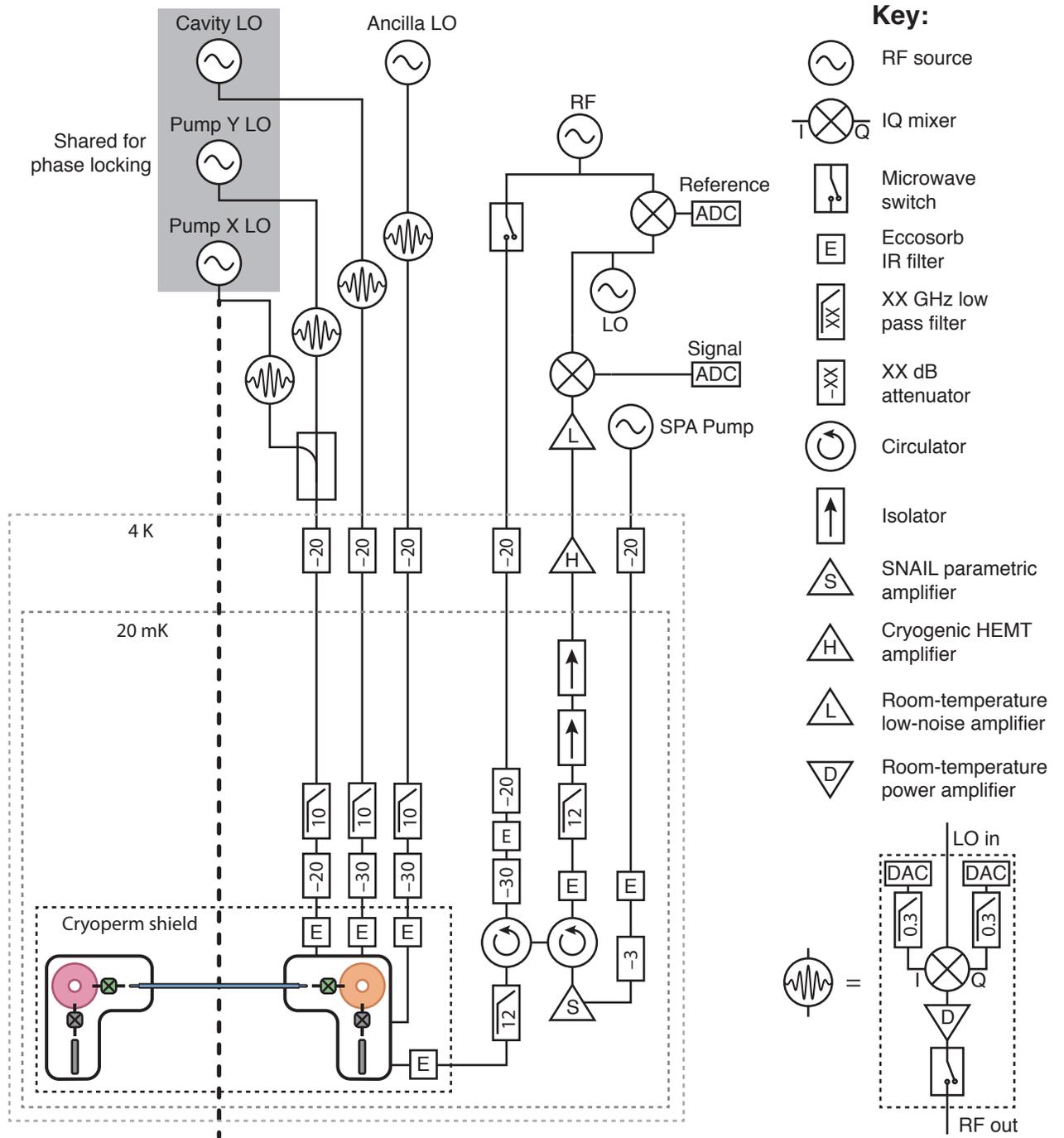
Figure 5.5: Quantum Network Bus Experimental Wiring Diagram. Input and output signal lines. Setup is duplicated across vertical dashed line. Three LOs in gray are shared between the two halves for phase-locking. Some attenuation and filtering at room temperature not shown.

switch and a digital output from the X6-1000M board, because we were using all available IF outputs for drives which required frequency and phase control.

5.4.3 Dispersive readout

We now turn to readout of the transmon state, giving just the essentials; more details can be found in [146, 147, 22, 144, 148], for instance. The task at hand is rapidly mapping the state of the transmon onto the state of some directly-measurable classical quantity. In this case, it will be the phase and amplitude of an outgoing microwave field.

The interaction

The transmon t coupled to the readout resonator r with dispersive shift χ_{rt}, which is in turn coupled to a transmission line at a rate κ_{r}. The dispersive Hamiltonian in this two-mode case is the simplified version of Equation 5.9:

$$\mathcal{H} = \omega_{\mathrm{t}} t^\dagger t + \chi_{\mathrm{tt}} t^{\dagger 2} t^2 + (\omega_{\mathrm{r}} + \chi_{\mathrm{rt}} t^\dagger t) r^\dagger r \tag{5.11}$$

where we have factored the dispersive interaction in a way that makes it clear that there is a frequency shift of χ_{rt} on the readout per excitation in the transmon. We're neglecting the self-Kerr of the readout, which makes it difficult to solve for the average behavior with classical methods, and is usually small enough to be neglected when predicting the general behavior.

The readout response

We irradiate the readout resonator with a square pulse of light r_{in} at a frequency ω near its resonance frequency, detuned by an amount $\Delta = \omega - \omega_{\mathrm{r}}$. The equation of motion for the readout resonator is given by the Langevin equation (Subsection

4.3.3):

$$\dot{r} = i(\omega_{\mathrm{r}} + \Delta + \chi_{\mathrm{rt}} t^{\dagger} t) - \frac{\kappa_{\mathrm{r}}}{2} r - \sqrt{\kappa} r_{\mathrm{in}}. \tag{5.12}$$

This makes clear the effective detuning depends on the transmon occupation number $n_t = \langle t^{\dagger} t \rangle = 0, 1, 2$ for the transmon in $|g\rangle, |e\rangle, |f\rangle$. The lovely fact is that the expectation value of this equation — the average trajectory $\langle r \rangle$ in phase space, and hence, the average output field $\langle r_{\mathrm{out}} \rangle = \langle r_{\mathrm{in}} \rangle + \sqrt{\kappa_{\mathrm{r}}} \langle r \rangle$ — is linear and can be solved in Fourier space. The response of the readout resonator to the incoming field is that of a Lorentzian filter of bandwidth κ_{r} and center frequency $\omega_{\mathrm{r}} + \Delta + \chi_{\mathrm{rt}} n_t$. The state-dependent center frequency of the readout leads to a state-dependent response of the average resonator field in the time domain. The average field is shown for the first three transmon states in Figure 5.6, for a square drive pulse placed near $\Delta = 0$.

Signal processing techniques

The quantum and classical fluctuations of the output field means that any single measured trajectory will have some (usually Gaussian) noise, so the trajectories need to be distinct enough to be reliably discriminated on a single-shot basis. In practice one needs a quantum-limited parametric amplifier on the output so the classical noise from subsequent amplification stages does not overwhelm the signal. We use a Josephson Parametric Converter [149, 45, 150] in Chapter 6, and two SNAIL Parametric Amplifiers [151] in Chapter 7. The next stage is a low-noise semiconductor HEMT amplifier located at the 4 K stage of the refrigerator, with more stages at room temperature.

As can be seen in the trajectories in Figure 5.6, the information content of the signal grows in time as the average trajectories become more distinct, and then decreases again after the pulse ends and the field rings down towards zero. However, the size of the noise is usually time-independent, so the signal-to-noise ratio is not constant. One can integrate the time trace to extract a single complex number, but
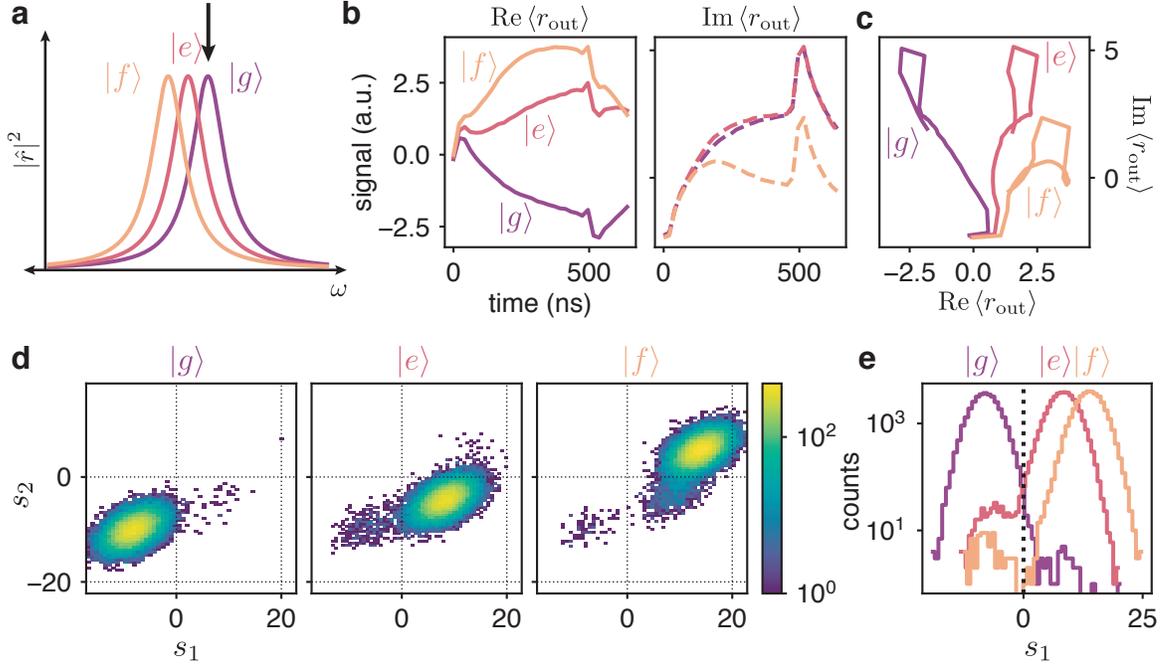
Figure 5.6: Dispersive Readout. a) Magnitude of steady-state field in readout resonator versus drive frequency for the first three transmon states. Parameters are $\chi_{rt} = 2\pi \times 0.8$ MHz, $\kappa_r = 2\pi \times 1.6$ MHz. Black arrow indicates $\omega = \omega_r$, where we place the drive in this example. b) Measured real (in-phase) and imaginary (quadrature) parts of average resonator field $\langle r \rangle$ for three transmon states for a constant pulse 500 ns in length. Sharp features at $t = 0$ and $t = 500$ ns correspond to turn on and turn off of pulse, which is visible in this reflection measurement. About 35 000 shots are included for each state. c) Trajectories from b) plotted in phase space. d) Histograms of weighted and integrated signals, sorted by prepared state. Horizontal (vertical) axis corresponds to s_1 (s_2) for distinguishing $|g\rangle$ and $|e\rangle$ ($|e\rangle$ and $|f\rangle$). e) Projection of histograms from d) onto s_1 axis. Vertical dotted line indicated threshold v_1.

this disproportionately weights short and long times. In practice, one can construct weighting envelopes $w_{1,2}(t)$ from the average responses for each transmon state, by which the trajectory is multiplied before integration. For example, when the noise is Gaussian and state-independent, the ideal weighting envelope to distinguish $|g\rangle$ and $|e\rangle$ is $w_1(t) = \left(\langle r_{out}(t) \rangle_e - \langle r_{out}(t) \rangle_g \right)^*$. This is the difference of the average signals for the two states, which weights most heavily the times when the difference is largest, and weights less the times which contain less information, like the early part of the signal. The complex conjugate serves to orient the signal in the real axis. The signal used to distinguish $|g\rangle$ and $|e\rangle$ is then $s_1 = \mathrm{Re} \int_0^T w_1(t) r(t) \mathrm{d}t$. Likewise,

we create a second envelope to distinguish $|e\rangle$ from $|f\rangle$. The optimal envelope is $w_2(t) = \left(\langle r_{\text{out}}(t)\rangle_f - \langle r_{\text{out}}(t)\rangle_e\right)^*$, and the signal is $s_2 = \text{Re} \int_0^T w_2(t)r(t)\mathrm{d}t$. Independent thresholds $v_{1,2}$ are then determined to assign a binary measurement outcome $m = 0$ for $s_1 < v_1$ and $m = 1$ for $s_1 > v_1$.

This discrimination problem can be cast in a framework in which the above weighting strategy corresponds to linear discriminant analysis. More sophisticated techniques like Bayesian filtering and quadratic discrimination [147, 152] can be applied to improve the fidelity of assignment when the noise is not Gaussian, or is state-dependent. Relaxation of the transmon during the measurement is a common example of this. This is often difficult to implement in a real-time quantum controller, and has not been widely adopted yet. Additionally, drive pulses which are not square, or are multi-chromatic, can improve the speed and fidelity with which readout is performed by separating the trajectories more rapidly, and by emptying the resonator faster [144, 153, 154, 61].

Chapter 6 makes use of a more complicated scheme where two transmon-readout pairs share a single measurement chain (see also [45, 71, 72]). This cascaded readout scheme is a bit more complicated, but the same semi-classical analysis is sufficient to describe the behavior.

5.4.4 Improving measurement contrast with higher levels

Typical readout discrimination fidelities for an appropriately designed sample can reliably be around 98%. The contrast is primarily dominated by decay events of the transmon during the readout acquisition, which cause an abrupt change in the readout trajectory, and can lead to mis-assignment of $|e\rangle$ as $|g\rangle$. For typical transmon lifetimes of $50\,\mu s$ and readout duration of 500 to 1000 ns, we expect this to happen around 1% of the time. However, this problem is exacerbated by the fact that population of the readout resonator can enhance the qubit decay rate [155, 156]. In many cases, this

measurement contrast can be calibrated out, but when we require true single-shot state assignment (for instance, for looking at entanglement correlations in Chapter 7), it is advantageous to have higher discrimination fidelity.

Exciting the transmon from $|e\rangle$ to $|f\rangle$ immediately prior to readout has two advantages that can increase the fidelity with which we can assign the initial state to $|g\rangle$ or $|e\rangle$. The readout response in the $|f\rangle$ state is included in Figure 5.6. Since the readout resonator is further detuned from the drive, the trajectory bends more severely. This effective larger dispersive shift ($\approx 2\chi_{\mathrm{rt}}$) provides a higher measurement rate for the same drive amplitude and system parameters, enhancing separation of the trajectories and distinguishability. Additionally, since the primary error channel for the $|f\rangle$ level is to decay to $|e\rangle$, the separation persists even during a decay event, to first order. This protection of the assignment fidelity from errors has been explored in depth recently in [157], though the idea is much older than that [158].

Example average trajectories and histograms which show the obvious improvement in distinguishability from using the $|f\rangle$ level are plotted in Figure 5.6, measured with sample 2 from Chapter 7. The ground state is prepared by measuring and post-selecting on $|g\rangle$, then measuring again to obtain the data shown. The first excited state is prepared by following this first measurement with a g–e π pulse. The second excited state is prepared as we prepare $|e\rangle$, but with an additional e–f π pulse. The assignment fidelity [147] is improved from 0.988 for g–e to 0.996 for g–f, both by separating the trajectories more rapidly and by reducing the effect of decay errors. A nice additional feature is that the measurement contrast becomes more symmetric between $|g\rangle$ and $|f\rangle$, due to the first-order protection from decay events.

The drawback of this approach is that the lifetime of the $|f\rangle$ state is usually shorter than that of the $|e\rangle$ state, so while decay errors do not impact the assignment fidelity, they may be more likely, and can degrade the fidelity of subsequent operations. For this reason, we only use this technique when the measurement is the final operation in

the experimental sequence and we only care about distinguishing power. Despite these challenges, the $|f\rangle$ state, or even higher levels, can in some cases be discriminated from both $|g\rangle$ and $|e\rangle$, and has been used for operation and measurement of the transmon as a multi-level system in several contexts [155, 159, 64, 67].

5.5 Controlling and measuring oscillators

Having discussed control and readout of the transmon, we can now turn to the task of controlling and measuring a long-lived storage cavity with a transmon. In this section, we will treat the dispersive measurement described in Subsection 5.4.3 as a subroutine, ignoring the actual dynamics of the readout resonator.

5.5.1 The dispersive spectrum

The coupling between the transmon and the cavity mode a can be factored as

$$\mathcal{H} = \omega_\mathrm{a} a^\dagger a + (\omega_\mathrm{t} + \chi_\mathrm{at} a^\dagger a)t^\dagger t + \chi_\mathrm{tt} t^{\dagger 2} t^2 \tag{5.13}$$

This makes clear that the transmon frequency is shifted by χ_at per photon in the cavity[4]. When the dispersive shift χ_at is large compared to the decay rate Γ_1 and dephasing rate Γ_2 of the transmon and decay rate of the cavity κ_a, we are in the so-called number-split regime. Since $\chi_\mathrm{at}/2\pi \sim 1\,\mathrm{MHz}$, $\Gamma_1/2\pi, \Gamma_2/2\pi \sim 5\,\mathrm{kHz}$, and $\kappa_\mathrm{a}/2\pi \sim 100\,\mathrm{Hz}$, we typically operate deep within this regime. In this limit, one can drive the transmon at frequency $\omega = \omega_\mathrm{t} + n\chi_\mathrm{at}$, and if the drive strength $\epsilon \ll \chi_\mathrm{at}$, excite it if and only if the cavity is in $|n\rangle$[160, 161]. The resulting spectrum as a function of drive frequency, with a coherent state in the cavity, is shown in Figure

---

4. Corrections to this Hamiltonian should be made when considering the $|f\rangle$ state and above, since higher orders of nonlinearity cause the dispersive shift to vary as we go up the transmon ladder. See for example [67].
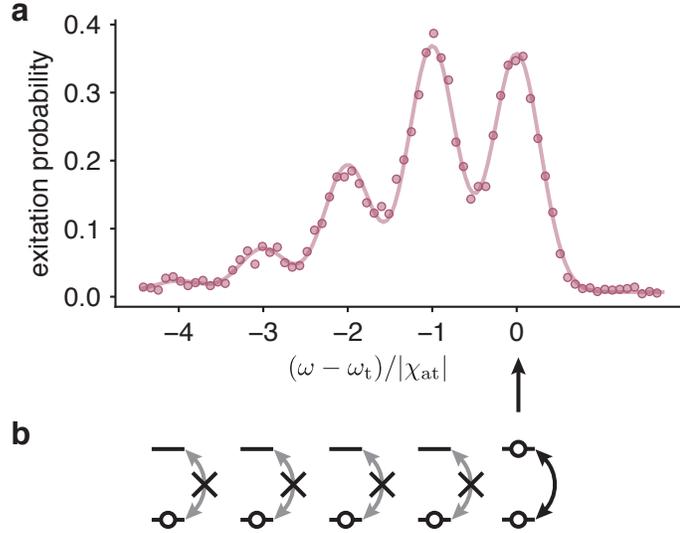
Figure 5.7: Number Splitting. a) Measured spectrum of transmon with a weak excitation tone, with a coherent state $|\alpha = 1\rangle$ in the cavity. Line-widths are determined by the spectral width of the excitation pulse, which is set so that the pulse time is as short as possible while still resolving the peaks. b) When driving at $\omega = \omega_t$ as indicated by the arrow in a), with $\epsilon \ll \chi_{at}$, a selective rotation is effected, which excites the transmon if and only if the cavity is in $|0\rangle$.

5.7. This type of experiment is used to measure the dispersive shift, which is the spacing between the peaks. Since the height of each number peak is proportional to the cavity occupation in that Fock state, we can fit this spectrum to extract the photon number distribution in the cavity.

5.5.2 Selective measurement and control

In the number-split regime, we have control over the transmon state conditioned on the cavity number state. In practice, we effect a *selective* π pulse with a Gaussian drive pulse with $\sigma \gtrsim 1/3\chi_{at}$. This type of control can be used for performing phase gates on the cavity [162, 163, 67, 164] via a 2π pulse, returning the transmon to its ground state but acquiring a number-state dependent phase. These phase gates, when combined with cavity displacements, offer universal control over the cavity state, as proposed in [162] and demonstrated in [163].

Selective transmon rotations are also an easy way to measure the state of the

cavity; specifically, it can be used to ask the binary question "is the cavity in $|n\rangle$?" in a QND way [161]. This is a basic form of tomography on the cavity, which is a useful characterization technique as well as the primary readout for certain bosonic simulation experiments [110]. In the specific case of $n = 0$, we use it extensively as a way of ensuring that the cavity is in the vacuum state before beginning an experiment, by mapping the answer "yes $n = 0$" onto the transmon excited state. Finally, in Chapter 6, by simultaneously applying drives on all of the even (or odd) number peaks, we can measure the photon number parity of the cavity, and hence the Wigner function (see also [163, 98, 16]).

These selective operations are not extremely fast compared to the transmon decoherence. For $\chi_{\mathrm{at}} = 2\pi \times 1\,\mathrm{MHz}$, the total π pulse duration is at least $2\,\mu\mathrm{s}$. This results in several percent probability of a decay or dephasing event during the pulse, both of which can result in the transmon ending in the ground state when it should have been excited. This results in the measurement "$n = 0$?" having a false negative probability of several percent. False positives are much less likely, so when we want to be sure the cavity is in vacuum, we map the result "yes $n = 0$" onto the excited state. As discussed above, readout errors also mostly give false negatives, so the outcome $|e\rangle$ gives us high confidence that the cavity is empty.

5.5.3 Ramsey-type Control and Measurement

While the selective operations discussed above offer straightforward control and measurement of coupled transmons and cavities, they have the disadvantage of necessarily being several times slower than $2\pi/\chi_{\mathrm{at}}$. Certain types of operations and measurements can be done in a time π/χ_{at}, or even faster.
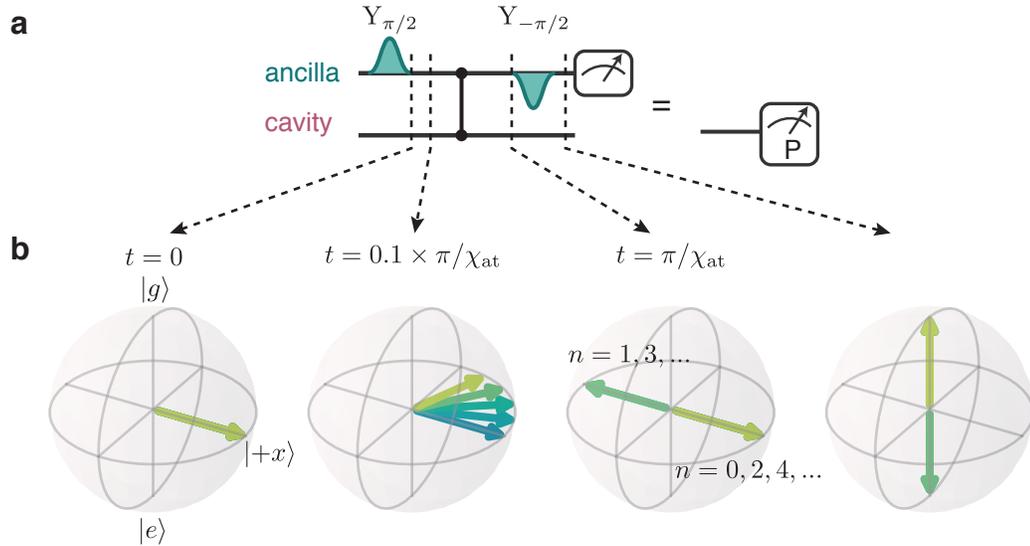
Figure 5.8: Ramsey-type Parity Measurement. a) Pulse sequence shows two trans-mon $\pi/2$ pulses with a delay between. During the delay, the dispersive interaction effects a parity-CPHASE. Measuring of the transmon state is equivalent to measuring the parity of the cavity state. b) Bloch sphere for $n = 0$ to $n = 4$ (dark to light) at different times in the sequence. First rotation brings all number states to $|+x\rangle$. Number states acquire phase at different rates, fanning out. At $t = \pi/\chi_{\mathrm{at}}$, even n components are anti-parallel to odd. A final rotation aligns even parity states with $|g\rangle$ and odd states with $|e\rangle$.

Parity-selective rotations

In Subsection 5.5.2 we considered transmon rotations with rates much weaker than χ_{at}. An alternate approach is to use *unselective* rotations much faster than the dispersive evolution, and delays. A common example is the parity map, which excites the transmon if the cavity has even parity, and leaves it in the ground state if the parity is odd. This sequence was used in [165] to generate entangled states, and then in [166, 65, 66] as an error syndrome measurement (see Section 3.4), and is used throughout the lab to measure Wigner functions. A more sophisticated version using three levels of the transmon has been used to measure the *joint* photon number parity of two cavities [167].

As discussed above, this can be managed with a comb of selective rotations at frequencies $\omega_{\mathrm{t}} + n\chi_{\mathrm{at}}$ for even n. However, the same task can be accomplished quite a bit faster faster with a Ramsey sequence, the dynamics of which are shown on

the $\{|g\rangle, |e\rangle\}$ Bloch sphere in Figure 5.8. First, an unselective $\pi/2$ rotation (length $\tau_{\pi/2} \ll 1/\chi_{\mathrm{at}}$) brings the transmon to the state $|+x\rangle$ for all cavity number states. Then, we wait. After a wait time t, the transmon acquires a phase $n\chi_{\mathrm{at}}t$ for each cavity number state $|n\rangle$. This phase, which amounts to a number-dependent Z rotation, appears as a fanning out on the Bloch sphere, since the number components rotate at different rates. After a time $t_{\mathrm{p}} = \pi/\chi_{\mathrm{at}}$, the phase is $n\pi$. Since the phase is only defined modulo 2π, the even components have return to $|+x\rangle$, while the odd components coalesce at $|-x\rangle$. This delay time has amounted to a parity-controlled $Z_{\pi/2}$ rotation, which is why we indicate it as a CPHASE gate in Figure 5.8. A final unselective $\pi/2$ with the opposite phase reorients the transmon Bloch vectors to Z eigenstates, turning the sequence into a parity-CNOT. Now, measuring the transmon state is tantamount to a measurement of the parity of the cavity state.

The use of the dispersive evolution during the delay to map the two parity manifolds onto orthogonal states of the transmon is what enables this protocol to work in minimal time, $\pi/\chi_{\mathrm{at}} + 2\tau_{\pi/2}$. Actually, it's slightly shorter than that, since some of the dispersive evolution occurs during the transmon pulses. We calibrate the time by performing this Ramsey experiment with a coherent state in the cavity, and fitting the measured transmon excitation probability as a function of wait time to a Poisson-distributed sum of cosines — see [168], Chapter 5 for more details.

5.5.4 Numerical universal control

The discussion above has so far only considered control sequences where selective or unselective transmon rotations and cavity displacements are interleaved, but we have not considered

- Operations which are neither fully selective nor fully unselective

- Driving both the cavity and the transmon simultaneously

This is because transmon rotations and cavity displacements do not commute with the dispersive Hamiltonian, nor even do displacements in different directions commute with one another, nor rotations about different axes. So it is rather difficult to qualitatively understand what happens when one drives the cavity and transmon simultaneously, even with constant amplitude and phase, let alone when these vary in time.

Fortunately, it is straightforward to simulate the unitary propagator of a single transmon and cavity under arbitrary time-varying simultaneous drives and an arbitrary static Hamiltonian, provided the drives can be approximated as piecewise constant. Moreover, it is possible to numerically compute the derivative of the final propagator with respect to the drive amplitudes at each discrete time point. This allows for convenient numerical optimization of a desired evolution over the drive parameters via gradient descent [169, 170], first used in nuclear magnetic resonance experiments [171, 172]. There are a few related approaches to this kind of optimization, like chopped random basis [173, 174] and the Krotov method [175], but gradient-descent pulse engineering (GRAPE) has become a workhorse of experiments in the Schoelkopf and Devoret labs over the last few years since the first demonstration of this method of control in cQED [79].

As used in this work, the problem statement is as follows. For a time-independent Hamiltonian $\mathcal{H}_0$, and time-dependent drive Hamiltonian

$$
\begin{aligned}
\mathcal{H}_\mathrm{d}(t) = {} & \mathrm{Re}\left(\epsilon_\mathrm{a}(t)\right)\left(a + a^\dagger\right) + i\mathrm{Im}\left(\epsilon_\mathrm{a}(t)\right)\left(a - a^\dagger\right) + \\
& + \mathrm{Re}\left(\epsilon_\mathrm{t}(t)\right)\left(t + t^\dagger\right) + i\mathrm{Im}\left(\epsilon_\mathrm{t}(t)\right)\left(t - t^\dagger\right),
\end{aligned}
\tag{5.14}
$$

the propagator after a time T is

$$
U(T) = \sum_{n=0}^{N-1} \exp\left[i\left(\mathcal{H}_0 + \mathcal{H}_\mathrm{d}(t_n)\right)\mathrm{d}t\right]
\tag{5.15}
$$

for time interval $\mathrm{d}t = T/N$. The optimization task is to map a set of M initial joint transmon-cavity states $\{|\psi_k^\mathrm{i}\rangle\}$ to final states $\{|\psi_k^\mathrm{f}\rangle\}$. The fidelity with which this task is accomplished for given set of drives is

$$\mathcal{F} = \frac{1}{M}\left|\sum_{k=0}^{M-1} \langle\psi_k^\mathrm{f}|U(T)|\psi_k^\mathrm{i}\rangle\right|^2 . \tag{5.16}$$

Due to the piecewise-constant approximation of $\mathcal{H}_\mathrm{d}$, the calculation of the $4N$ gradients of the fidelity can be computed in a time linear in N, which makes this optimization problem tractable on a desktop. The result is a numerically optimized control pulse (OCP).

We use the GRAPE algorithm for deterministically preparing quantum states in cavities, and for encoding and decoding arbitrary qubit states in a variety of logical encodings. It is easy to prepare an arbitrary superposition in the transmon $\{|g\rangle_\mathrm{t}, |e\rangle_\mathrm{t}\}$ manifold, and then we apply an OCP calculated to take $\{|\psi_k^\mathrm{i}\rangle\} = \{|g\rangle_\mathrm{t}|0\rangle_\mathrm{a}, |e\rangle_\mathrm{t}|0\rangle_\mathrm{a}\}$ into $\{|\psi_k^\mathrm{f}\rangle\} = \{|g\rangle_\mathrm{t}|0_\mathrm{L}\rangle_\mathrm{a}, |g\rangle_\mathrm{t}|1_\mathrm{L}\rangle_\mathrm{a}\}$. Now changing encodings is then simply a matter of using a different OCP.

The duration of an OCP is user-defined in the problem, and there is some flexibility in it. For performing nontrivial operations on a cavity, the pulse duration usually needs to be of order $T \sim 1/\chi_\mathrm{at}$. Making the pulse longer provides the algorithm with more control knobs, and more time in which to allow the dispersive evolution to work. Longer pulses usually require less amplitude and bandwidth of the drives, which is desirable. However, since the transmon is entangled with the cavity for most of the pulse, the actual infidelity is limited by the dephasing time T_2 of the transmon, at around T_2/T, which is usually of order 1%. Shorter pulses are less susceptible to the transmon decoherence, but usually require larger excursions in the oscillator phase space (akin to increasing the lever arm in the conditional displacement in Subsection 5.5.3). This enhances the sensitivity to cavity decoherence and higher-order terms in

the Hamiltonian, which can be included in $\mathcal{H}_0$ but may be known with less accuracy. The corresponding larger amplitude and bandwidth of the drives may also increase susceptibility to distortion and nonlinearities in the control hardware. These issues are explored in [79, 176] — the optimal balance is usually found empirically.

5.6 Parametric control

The control discussed so far, while made extremely general, flexible, and easy to implement by the GRAPE algorithm, is so far limited to drives resonant with the cavity and transmon, with rather narrow bandwidth. These numerical techniques don't generalize well to control of multiple cavities, since the complexity of the optimization grows rather quickly with Hilbert space size, which increases dramatically when adding a second cavity. Optimized control pulses of this form also require exciting the transmon, exposing the cavity to the typically much worse coherence properties of this element. While this approach is likely necessary in some contexts, there are a variety of operations which can be executed without entangling the cavity with the transmon, by applying *off-resonant* drives.

This section will explain how the nonlinearity of the Josephson junction can be used to perform nontrivial operations on one or more cavity modes. In particular, we will make use of the many rotating terms in the fourth-order Taylor expansion of the cosine in Equation 5.8 to enable new, resonant Hamiltonian terms. The explanation used here is related to the descriptions found in [83, 98, 106], but is often described as a type of sideband driving [96, 71] which traces its roots to atomic physics. This approach is an example of *parametric Hamiltonian engineering*, and bears similarities to the kind of control used in parametric amplifiers [149, 177] and flux-tunable superconducting circuits [97]. I like this description because it makes it very easy to estimate the relative strengths of these processes. An approach which allows numer-

ical calculation to higher order of nonlinearity can be found in [178].

5.6.1 From Off-resonant Drive to Parametric Pump

We begin with the multi-mode BBQ Hamiltonian in Equation 5.8, reproduced here with a drive applied to mode l at frequency ω_d:

$$\mathcal{H} = \sum_k \omega_k a_k^\dagger a_k - E_J \cos_{nl}\left(\sum_k \varphi_k(a_k + a_k^\dagger)\right) + \epsilon(t)e^{-i\omega_d t}a_l + \epsilon^*(t)e^{i\omega_d t}a_l^\dagger \quad (5.17)$$

The full details of the derivation will not be reproduced here (see [176], Appendices 3 and 4), but the result is that by making a few frame changes, we can work our way to a much simpler Hamiltonian from which we will quickly see new interactions arise. First, we move into the rotating frame of the each mode $k \neq l$, and the rotating frame of the *drive* for mode l, yielding

$$\mathcal{H} = \Delta a_l^\dagger a_l - E_J \cos_{nl}\left(\varphi_l(a_l e^{-i\omega_d t} + a_l^\dagger e^{i\omega_d t}) + \sum_{k \neq l} \varphi_k(a_k e^{-i\omega_k t} + a_k^\dagger e^{i\omega_k t})\right)$$
$$+ \epsilon(t)a_l + \epsilon^*(t)a_l^\dagger, \quad (5.18)$$

with drive detuning $\Delta = \omega_l - \omega_d$. This gets rid of the fast rotation on the drive term. We can now make a displacement transformation on mode l by the dimensionless amplitude $\xi(t) = \epsilon(t)/\Delta$ to eliminate the drive term, resulting in

$$\mathcal{H} = \Delta a_l^\dagger a_l - E_J \cos_{nl}\left(\varphi_l\left((a_l + \xi_l)e^{-i\omega_d t} + (a_l^\dagger + \xi_l^*)e^{i\omega_d t}\right) + \sum_{k \neq l} \varphi_k(a_k e^{-i\omega_k t} + a_k^\dagger e^{i\omega_k t})\right)$$
$$(5.19)$$

which moves the drive inside the non-linear cosine. Finally, we get rid of the first term by rotating mode l again into its proper rotating frame, ending with

$$\mathcal{H} = -E_J \cos_{nl}\left(\varphi_l \xi_l e^{-i\omega_d t} + \varphi_l \xi_l^* e^{i\omega_d t} + \sum_k \varphi_k\left(a_k e^{-i\omega_k t} + a_k^\dagger e^{i\omega_k t}\right)\right) \quad (5.20)$$

By making these transformations, we can say that instead of coupling to a particular mode a_l, we couple directly to the phase across the junction. In fact, Equation 5.20 doesn't really require us to prefer a particular mode — the strength of the effective phase bias depends only on the product $\varphi_l \xi$, and in general the same effects can be observed by driving any mode.

The nonlinear cosine term in the Hamiltonian in Equation 5.20 allows all products of any four of the terms in the sum in its argument. This results in two kinds of new Hamiltonian terms, which we discuss now.

5.6.2 Always-on terms

The first type of term which emerges from the driven Hamiltonian are time-independent for any combination of mode and drive frequencies. This gives rise to self- and cross-Kerr terms as discussed in Subsection 5.2.5. However, it now also yields low-order terms which are non-rotating if they include equal numbers of any term and its conjugate. An example is the so-called Stark shift:

$$
\begin{aligned}
& E_{\mathrm{J}} \left(\varphi_{\mathrm{a}} a e^{-i\omega_{\mathrm{a}}t} \right) \left(\varphi_{\mathrm{a}} a e^{-i\omega_{\mathrm{a}}t} \right)^{\dagger} \left(\varphi_l \xi_l e^{-i\omega_{\mathrm{d}}t} \right) \left(\varphi_l \xi_l e^{-i\omega_{\mathrm{d}}t} \right)^{\dagger} \\
& = E_{\mathrm{J}} \varphi_{\mathrm{a}}^2 \varphi_l^2 |\xi_l|^2 a^{\dagger} a e^{-i(\omega_{\mathrm{a}} - \omega_{\mathrm{a}} + \omega_{\mathrm{d}} - \omega_{\mathrm{d}})t}
\end{aligned}
\tag{5.21}
$$

which is a frequency shift on mode a dependent on the square of the pump amplitude. This sort of term is resonant for any mode and pump frequencies because the conjugates cancel the rotation. In fact, the drive induces a Stark shift on every mode which couples to the junction. The Stark shift can be useful for limited *in situ* frequency tuning [90, 179], but it is mostly an annoyance in this work.

5.6.3 Exotic Terms Made Resonant

The other class of emergent Hamiltonian term can enable new sorts of processes, and are made non-rotating for certain values of the pump frequency. An example which can be obtained with a single pump with two modes a and b is

$$E_{\mathrm{J}} \left(\varphi_{\mathrm{a}} a e^{-i\omega_{\mathrm{a}} t} \right)^2 \left(\varphi_{\mathrm{b}} b e^{-i\omega_{\mathrm{b}} t} \right)^\dagger \left(\varphi_l \xi_l e^{-i\omega_{\mathrm{d}} t} \right)^\dagger$$
$$= E_{\mathrm{J}} \varphi_{\mathrm{a}}^2 \varphi_{\mathrm{b}} \varphi_l \xi_l^* a^2 b^\dagger e^{-i(2\omega_{\mathrm{a}} - \omega_{\mathrm{b}} - \omega_{\mathrm{d}})t} \tag{5.22}$$

which is resonant for $\omega_{\mathrm{d}} = 2\omega_{\mathrm{a}} - \omega_{\mathrm{b}}$. This term (and its Hermitian conjugate) exchanges two photons in mode a for a photon in b and a photon in the pump, and is used in various forms where one of a and b is a transmon and the other is a resonator [96, 83, 64].

5.6.4 Parametric conversion

The process of choice for the experiments in this thesis involve two pumps at different frequencies. By applying the frame changes twice, we can include two pump terms inside the cosine, with strengths and frequencies $\xi_{1,2}$ and $\omega_{1,2}$. We will use this configuration to exploit the term

$$E_{\mathrm{J}} \left(\varphi_{\mathrm{a}} a e^{-i\omega_{\mathrm{a}} t} \right) \left(\varphi_{\mathrm{b}} b e^{-i\omega_{\mathrm{b}} t} \right)^\dagger \left(\varphi_{\mathrm{c}} \xi_1 e^{-i\omega_1 t} \right)^\dagger \left(\varphi_{\mathrm{c}} \xi_2 e^{-i\omega_2 t} \right)$$
$$= E_{\mathrm{J}} \varphi_{\mathrm{a}} \varphi_{\mathrm{b}} \varphi_{\mathrm{c}}^2 \xi_1^* \xi_2 a b^\dagger e^{-i(\omega_{\mathrm{a}} - \omega_{\mathrm{b}} - \omega_1 + \omega_2)t} \tag{5.23}$$

(and conjugate), which is resonant when $\omega_2 + \omega_{\mathrm{a}} = \omega_1 + \omega_{\mathrm{b}}$. In this case, we exchange a photon from mode a and pump 2 for one in mode b and pump 1. In other words, this processes turns on whenever the detuning between the two pumps is equal to the detuning between the modes, with rate $g_{\mathrm{ab}} = E_{\mathrm{J}} \varphi_{\mathrm{a}} \varphi_{\mathrm{b}} \varphi_{\mathrm{c}}^2 \xi_1^* \xi_2 = \sqrt{\chi_{\mathrm{ac}} \chi_{\mathrm{bc}}} \, \xi_1^* \xi_2$. The ability to quickly cast the mixing strength in terms of design parameters like cross-Kerrs is a nice feature of this formalism.

In this description we assume c is the transmon mode mutually coupled to two resonators a and b, though this process is quite general. This conversion coupling is exactly of the form we assumed in Section 4, and indeed we will use it in the two contexts described therein. In Chapter 6, a and b will be a storage cavity and stripline resonator, respectively, with b strongly coupled to a transmission line continuum, realizing the setup proposed for the Cirac protocol for communication with propagating photons (Section 4.4)[5]. Then in Chapter 7, a will again be a storage cavity, but b will be a quantum bus — a mode of a coaxial cable resonator, which will couple in this way to a module at each end, implementing the circuit in Section 4.5.

5.6.5   Unwanted effects and limitations

As we will discuss in the two experiments in the chapters to follow, this conversion process can be rapid and selective, but it is not perfect. Often by applying pumps to the system, we can excite other undesired processes, which may be detuned but much stronger than the process we desire. This usually manifests as excitations being created or destroyed or moved around, usually into the transmon mode which is most strongly coupled to the junction (large zero-point φ_c). Additionally, due to the large Stark shift experienced by the transmon, surprising processes can become resonant at particular pump strengths, leading to similar excitation. These spurious transmon excitations are poison, because the dispersive shift of modes a and b to the transmon are usually larger than g, so stochastic excitations of the converter will shift the modes away from the resonance condition, both deactivating the conversion and dephasing the cavities. These effects place maxima on the pump strengths we are willing and able to use, and hence on the conversion rate g. An additional concern is that the pumps, which are quite strong compared to most of our other microwave drives,

---

5. Note that in Chapter 6 we will assume the pumps couple to modes a and b, not to the transmon c, so the expression used there will differ slightly.

may locally heat attenuators at the base stage of the refrigerator, adding wide-band thermal noise which can excite modes of the sample, even after the pump is turned off.

Practical limits

The physics of the pump-induced transmon excitation is a subject of active research. A very thorough theoretical and experimental investigation was carried out in [178]. The full story is complicated and beyond the scope of this thesis. But there are a few things which can go wrong. One is that a transmon driven strongly and off-resonantly has dressed eigenstates[6] which are superpositions of the undriven eigenstates. This means that a decay channel in the undriven basis can look like an excitation in the driven basis. The same is true of dephasing. These new excitation mechanisms can lead to excitation rates which depend on the decay and dephasing rates of the transmon, the pump strength ξ, and in some cases the pump detuning Δ from the transmon frequency. The other kind of effect is a resonant one — since the non-linearity of the transmon allows all even-order mixing processes, we can accidentally drive strong processes which can excite the transmon. These processes are in principle predictable, but there are many of them, and their frequency is a strong function of the pump strength. The problem compounds when two pumps are applied, because the number of frequency collisions grows greatly.

The observations of [178] suggest a few practical considerations to soften these concerns. A generally useful approach is to place the pumps *blue-detuned* (positive detuning) from the transmon resonance. Since the transmon anharmonicity is negative, many more high-order transitions exist at lower frequencies, and these transitions also tend to shift down as the pumps strength is increased. Similarly, placing the cavity modes to be coupled likewise blue-detuned has a similar effect. We took

---

6. Strictly speaking, Floquet eigenstates.

this approach in Chapter 7. Other results of that work suggest that some of these concerns are mitigated by making the pump detuning Δ large compared to the transmon anharmonicity. This all favors pumps at rather high frequencies, which is challenging if they are far-detuned from all modes, since $\xi \approx \epsilon/\Delta$. Engineering microwave structures like waveguides which pass the pumps into the sample but do not induce spontaneous emission of the transmon and cavities is an active direction of research, and was for instance exploited in [180].

A general rule of thumb is that the phase across the junction $\varphi \xi$ should probably remain less than about 1, though, as discussed above, frequency collisions can cause problems before this if they are not avoided. For a transmon, the zero point phase $\varphi_c \approx (2E_C/E_J)^{1/4}$ is typically around 0.4, so we usually drive with $\xi < 1$. This limits the achievable conversion strength between cavity modes a and b to $g_{ab} < \sqrt{\chi_{ac}\chi_{bc}}$, which is usually on the MHz scale. Indeed, in Chapters 6 and 7, the maximum conversion strength is $g_{ab}/2\pi \approx 400$ to 550 kHz. This is improved by increasing the cross-Kerrs χ_{ac} and χ_{ac}, but this then induces additional self-Kerr, which can be problematic.

For the sample in Chapter 7, at the maximum conversion strength, we observed excitation rates of the driven transmon on the order of $1/(100\,\mu\text{s})$. Given a conversion strength of $500\,\text{kHz}$, the swap time is $500\,\text{ns}$, which yields a dephasing error rate of about 0.5%.[7] This is quite a bit larger than the error due to photon loss in a millisecond-lifetime 3D cavity, so further reducing these effects is a great practical concern. As far as we know, the conversion process itself is lossless, other than these ways in which it induces decoherence on the participating modes.

---

7. In this experiment, there are two transmons, and the swap time is a bit longer than this. All told, including improper initialization of these transmons to their ground state, the total error from uncontrolled excitations is about 4%.

5.7 Measuring a transmon through a high-Q cavity

Now we turn to the measurement of a transmon which does not have a readout resonator. This will be the conversion transmon that we use to drive parametric coupling between the cavity and the quantum bus in Chapter 7. For the purpose of experimental simplicity, and to reduce frequency crowding around the modes of interest, we do not include an additional resonator for measurement of this transmon. While we do not use this as a quantum degree of freedom for information processing, we do need to be able to measure it in order to characterize it, and to use it to measure the bus resonator. For this we map the ground and excited state of the conversion transmon onto pointer coherent states of the storage cavity, which we can measure quite well.

The mapping we use to measure the transmon is

$$
\begin{aligned}
|g\rangle |0\rangle &\to |g\rangle |0\rangle \\
|e\rangle |0\rangle &\to |e\rangle |\alpha'\rangle
\end{aligned}
\tag{5.24}
$$

followed by a measurement of whether the cavity is in vacuum, as described in Section 5.5. As long as α' is significantly larger than one, this allows exponentially perfect discrimination between the two transmon states, since the overlap between the two cavity pointer states is $|\langle 0|\alpha'\rangle|^2 = e^{-|\alpha'|^2}$. We present two approaches to this, which parallel the selective and Ramsey-style cavity-dependent transmon operations discussed above.

5.7.1 Cavity spectrum

As we explained in Subsection 5.2.1, drives on a harmonic oscillator cannot create quantum states. However, in the presence of the dispersive Hamiltonian, they can

have nontrivial action. By refactoring Equation 5.11

$$\mathcal{H} = (\omega_\mathrm{a} + \chi_\mathrm{at} t^\dagger t) a^\dagger a + \omega_\mathrm{t} t^\dagger t + \chi_\mathrm{tt} t^{\dagger 2} t^2 \tag{5.25}$$

we see the transmon-state-dependent dispersive shift of the cavity. This spectrum in the resolved-regime is illustrated in Figure 5.9a. By driving the cavity at frequency $\omega = \omega_\mathrm{a}$ with $\epsilon \ll \chi_\mathrm{at}$, we can selectively displace it if and only if the transmon is in $|g\rangle$. This can be used to create entangled states like $|g\rangle |\alpha\rangle + |e\rangle |0\rangle$. Then, if α is large, by measuring whether the cavity is in $|n{=}0\rangle$, we can effectively measure the state of the transmon.

Ramsey selective displacements

Much like the selective transmon rotations discussed in Subsection 5.5.2, this kind of selective displacement is fairly slow, so decay events of the transmon during the mapping sequence can limit the fidelity. Instead, we use the dual of the Ramsey-type parity-selective transmon rotation, which can effect an equivalent selective displacement, but in an arbitrarily short time (in principle). This is the simplest version of the protocol proposed in [181], and implemented in [129], where it is used to measure joint observables of multiple "blind" transmons.

A parallel of the transmon Ramsey sequence discussed in Subsection 5.5.3, this protocol consists of two cavity displacements of opposite amplitude, with a delay in between, as shown in Figure 5.9b. The evolution in cavity phase space is shown in Figure 5.9c, for the $|g\rangle$ and $|e\rangle$ states of the transmon. The displacements are assumed very fast compared to $1/\chi_\mathrm{at}$, though this is not a requirement [61]. The first displacement prepares a coherent state with amplitude α. During the delay, the $|g\rangle$ state does not evolve, but the coherent state rotates by an angle $\chi_\mathrm{at} t$ when the transmon is in $|e\rangle$. By now applying the opposite displacement, the cavity returns to
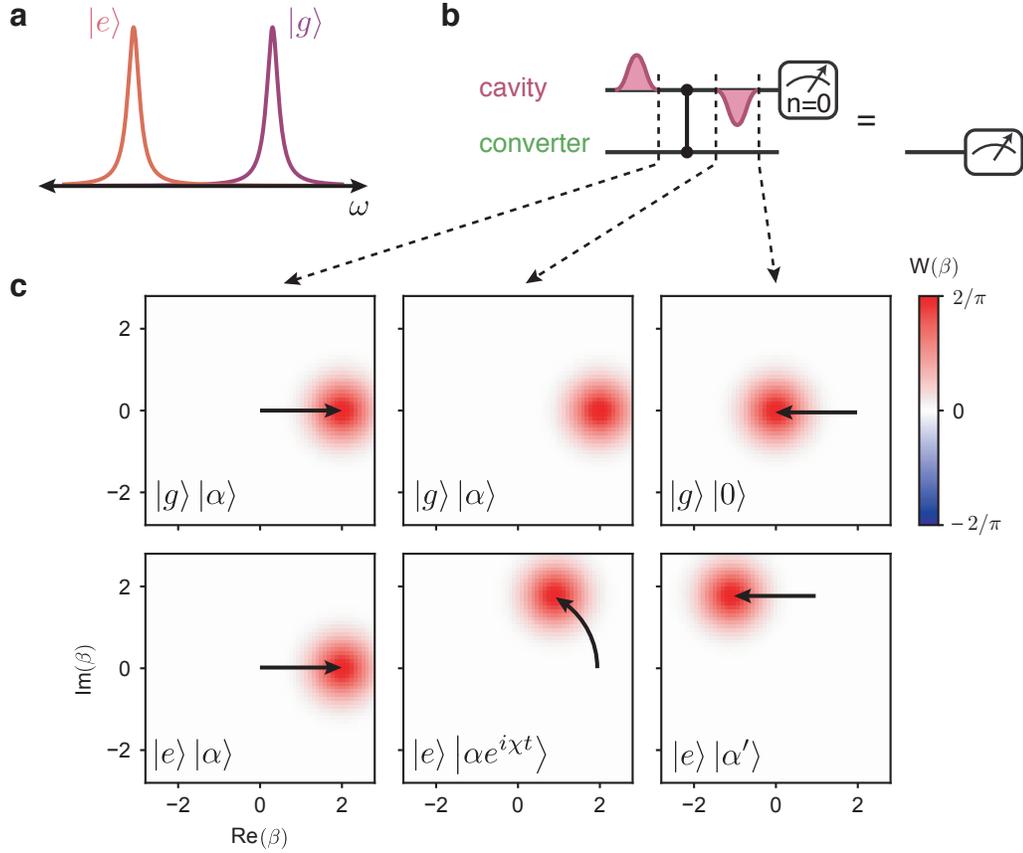
Figure 5.9: Transmon-Selective Cavity Displacement. a) Response of the cavity as a function of drive frequency ω, for first two transmon states. b) Pulse sequence for Ramsey-style selective cavity displacement. Displacement is applied, then a delay time during which the dispersive shift enacts a CPHASE-type operation, followed by the opposite displacement. Measuring the cavity to be in $|0\rangle$ is then a measurement that the transmon is in $|g\rangle$. c) Evolution of the cavity Wigner function in the frame rotating at ω_a during the sequence in b), for transmon in $|g\rangle$ (top) and $|e\rangle$ (bottom). Joint transmon-cavity state is indicated in each panel. Arrows indicate movement of the cavity state.

vacuum for the transmon in $|g\rangle$, but is displaced to $|\alpha'\rangle$ when the transmon is excited, where $\alpha' = \alpha(e^{i\chi_\mathrm{at} t} - 1)$. Now a measurement of the cavity in $n = 0$ constitutes a measurement of the blind transmon in $|g\rangle$.

Unlike the parity-selective rotation, this sequence does not have a minimum delay time. The separation between these two coherent states is maximal when $t = t_\mathrm{p} = \pi/\chi_\mathrm{at}$, but since the orthogonality is exponential in the separation, as long as $|\alpha'|^2 \gtrsim 4.5$, the overlap error is less than 1%. For instance, a modest displacement of $\alpha = 2$ and a delay time of $t_\mathrm{p}/2$ is sufficient, and the displacement size can be traded against

the time further. Higher order effects like the self-Kerr of the cavity, as well as pulse errors, can adversely affect the fidelity of this mapping sequence, but this idea of using the displacement size as a lever arm has broad applications, since the effect is to amplify the Hamiltonian, which semi-classically scales as $\bar{n}\chi$. For instance, some recent work from the Devoret group [61] embraces this approach, using a small dispersive shift and large cavity displacements to enact highly nontrivial control with conditional displacements. The authors demonstrate a few ways to suppress or cancel certain higher order effects like the self-Kerr, and provide an interesting and fruitful direction for future research.

5.8 Attaching a quantum bus

Now that we have described everything *inside* the module, we can talk about the physical implementation of the quantum bus prosed in Chapter 4. For this we use a coaxial cable resonator as described in Subsection 5.1.1. In this section we describe the characterization of these resonators.

5.8.1 Measurement *ex situ*

We can directly measure the mode structure and quality of the cable resonators in a simple standalone characterization rig, shown in Figure 5.10a. We capacitively couple one end of the cable to the continuum of a copper coax cable which is connected to our measurement apparatus. In this way we can measure the frequency response of the resonator modes with a vector network analyzer (VNA). This technique is similar to [123], though in that work the response was measured in transmission, with the external coupling assumed to be small. In reflection we can independently measure the internal and coupling rates, and there is only one coupling port, rather than two.
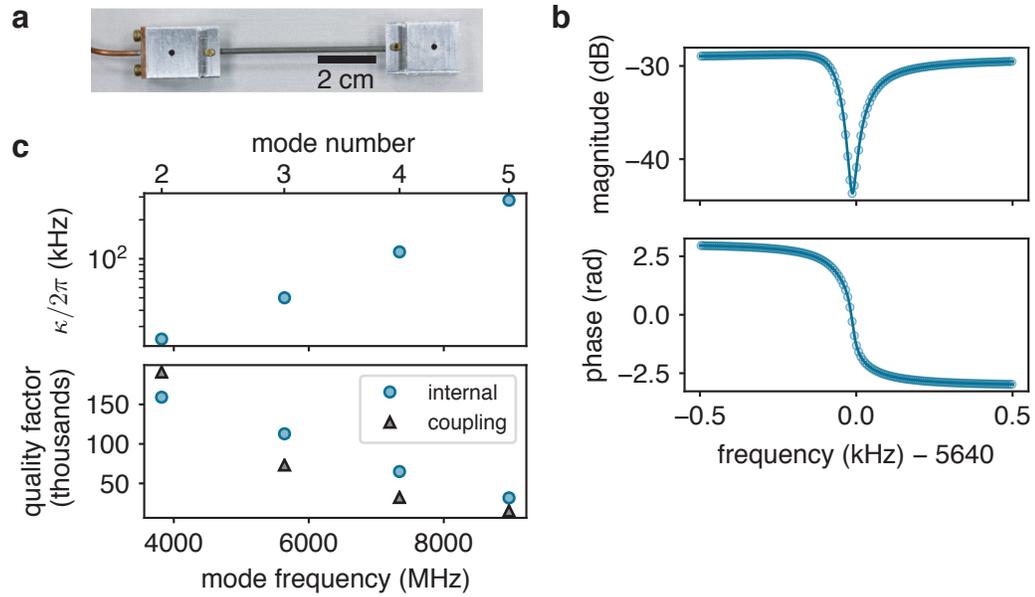
Figure 5.10: Bus Characterization *ex situ.* a) Setup for measuring cable resonators in reflection. Cable is embedded in aluminum tunnel, held in place with brass set screw. Copper cable (left) enters other end of tunnel, and is soldered into a flange. Both have outer conductor and dielectric removed in the tunnel (not shown). b) Reflection measurement of $n = 3$ mode, with fit. c) Extracted internal decay rate κ and internal and coupling quality factors for four modes of the same cable.

Connectorization

To minimize the complexity of the testing rig, and to make it as similar to the situation we will have when we integrate the cable into the modules, we use a simple connectorization and coupling scheme between the cable resonator and the signal line. We remove about 0.3 inches of outer conductor and dielectric to expose the inner conductor. The end of the cable is inserted into a loose-fitting tunnel in the body of an aluminum coupling block. A brass set screw pushes the outer conductor of the cable against the wall of the tunnel. This NbTi-Al joint is the only seam between the block and the cable. We tip the set screw with a small amount of indium, which deforms and provides a larger contact area, and allows us to apply more force on the joint before damaging the cable outer conductor. The copper cable has outer conductor and dielectric stripped similarly, soldered into a copper flange, and inserted into the other end of the block. The opposite end of the cable resonator

is embedded in another block, but with no copper cable. The small diameter of the tunnel confines the mode well within the tunnel, so the radiative losses from this end should be negligible. The flange is secured with screws and is visible in Figure 5.10a. The coupling is achieved by positioning the co-linear center conductors close together. To achieve coupling quality factors on the order of 10^5, the spacing between the two center conductors is typically 0.5 to 1 mm, though the coupling rate is frequency dependent.

Results

As an example, the VNA reflection trace of a 5.5 cm long cable resonator is shown in Figure 5.10b. This is the $n=3$ mode of this resonator, which has a free spectral range of just under 2 GHz. The extracted internal damping rates and internal and external quality factors of four modes of this cable, spanning 4 to 8 GHz, are shown in Figure 5.10c. In qualitative agreement with [123], the internal quality factor decreases at higher frequencies. While those authors attribute this to the increase of conductor dissipation with frequency, their data do not fit the expected power law dependence $Q \sim \omega^{-1}$ at low temperatures. As described in Chapter 4, losses in the seam between the outer conductor and the ground shell at the coupling end would also produce decrease in quality factor at higher frequency. Our data do not follow an obvious power law, so it is difficult to attribute the loss to one culprit.

We have measured cable manufactured by both Coax Co. and Keycom, and found no systematic difference in quality between the two, which have different density dielectric. The experimental results shown here and in Chapter 7 use cable from Coax Co. As suggested by the vendor, we have tried gently removing the NbTi oxide layer with fine-grit sandpaper before assembling, but the run-to-run variation so far seems to be larger than any benefit this conveys.

5.8.2 Integration

To use the quantum bus to couple our modules together, we need to integrate it into the modules with a capacitive coupling between the bus and the converter. The coupling is similar in spirit to the way we couple input and output signals (Subsection 5.3.2), with the outer conductor and dielectric removed as described in Subsection 5.8.1, and the center conductor protruding into the tunnel which holds the conversion transmon chip. The center conductor is orthogonal to the dipole moment of the conversion transmon, and crosses over it at the approximate location indicated in Figure 5.4. The pin is located 0.7 to 1 mm from the surface of the chip. If one wanted to increase this coupling rate in future, it should increase roughly linearly with the width of the transmon capacitor pad.

Spectroscopy

With the bus cable connected to a module at each end, there is no easy way to couple a diagnostic line to it directly, but we can measure it through the module without adding any complexity. Since the mode of interest has a dispersive shift to the conversion transmon, we can detect population in the bus by performing spectroscopy on the converter, much the way we measured storage mode population with the ancilla. The results of this spectroscopy with the cable driven to a coherent state is shown in Figure 5.11a. The cable and converter are mutually number split, with a dispersive shift of $\chi_{\mathrm{bc}} = -2\pi \times 4.3\,\mathrm{MHz}$.

Bus lifetime

Given the number-resolved spectrum in Figure 5.11a, the damping rate of the cable $\kappa_{\mathrm{b}} \ll \chi_{\mathrm{bc}}$. We can measure this rate directly by displacing the cable and driving the converter with a selective pulse at $\omega = \omega_{\mathrm{c}}$. The height of this spectroscopic peak corresponds to the occupation of the $|0\rangle$ state of the bus. This ring-down
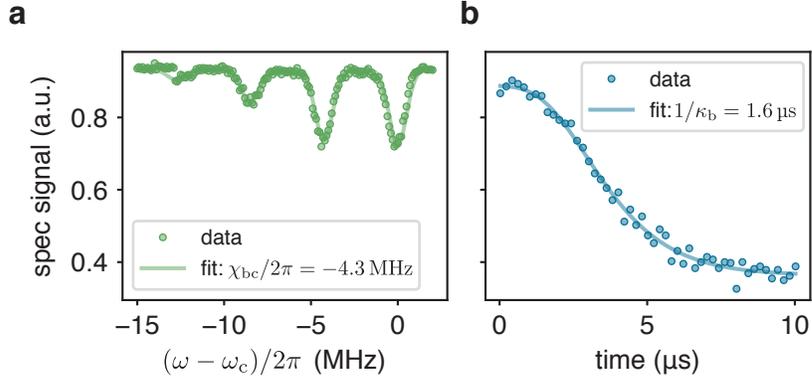
Figure 5.11: Bus Characterization *in situ*. a) Spectroscopy on conversion transmon for a coherent state $|\alpha = 1\rangle$ in the bus. b) Lifetime measurement of bus. Bus is displaced to $|\alpha = 2.5\rangle$, then after a variable delay, the conversion transmon is driven with a selective π pulse at $\omega = \omega_c$.

measurement, shown in Figure 5.11b, reveals a bus lifetime of $1.6\,\mu s$, or $\kappa_b/2\pi = 100\,kHz$.

While in the standalone configuration we measured the damping rate of the same mode of this cable to be as low as $\kappa_b/2\pi = 50\,kHz$, this is not yet consistently achieved from run to run. It could be that, given a different insertion into the package as compared to the testing housing in the standalone setup, the location of a lossy seam might be impacting the mode differently. It could also be variation in the quality of this seam given different mounting conditions.

In the language of Subsection 4.3.4, the measured quality factor would imply an effective resistance $\beta_n R = 200\,\mu\Omega$, where R is the actual resistance and $\beta_n \leq 1$ is a dimensionless geometric factor which depends on the location of the seam and the mode number n. Due to the distributed nature of the clamping method here, it is difficult to say with confidence what β_n is for these modes. Most likely, the large contact area reduces the effective resistance by providing multiple parallel channels through which the current can flow from the cable into the aluminum, since the seam conductance measured in [124] for aluminum alloy joints would predict $R \approx 30\,m\Omega$ (see Subsection 4.3.4). Alternatively, the soft aluminum might deform against the

harder NbTi, providing a seam of reasonable quality. More careful variation of the seam location and conditioning is required to fully understand this effect.

5.9 Conclusion

In this chapter we described the parts of the modules used in Chapter 6 and Chapter 7 to realize the protocols for entanglement and communication proposed in Chapter 4. We also gave an introduction to parametric conversion, which will be used as the coupling mechanism assumed in the two protocols. In the next two chapters we will describe the experimental results and conclusions taken from them.

Chapter 6

Pitch and Catch with Photons

6.1 Introduction

In this chapter we use all the tools and hardware we have developed in the previous chapters to describe our experiment achieving quantum state transfer and entanglement generation between 3D cavity memories using parametric conversion and propagating photons.

6.1.1 Communication with flying photons

We described the basic features of the approach taken here in Section 4.4, including the proposal by Cirac, Zoller, Kimble, and Mabuchi [17]. This protocol was originally envisioned as a way to perfectly transfer the state of one atom to another through an optical fiber. The idea is to map the state of the atom onto a propagating photonic wavepacket, and then capture that wavepacket at the other end. In this sense, the approach lends itself to implementation in the shuttling-type modular architecture presented in Subsection 2.5.2, provided we have a sufficiently flexible and non-reciprocal router.

However, by slight modification of the protocol, we can also generate entanglement

between the two atoms, thus making this approach applicable to the entanglement-based architecture described in Subsection 2.5.3. In this case, we begin with an excitation in one atom, and only release it halfway as a photon. This creates an entangled atom-photon pair. By completely absorbing the photon at the receiver, we swap the entanglement into the atom-atom pair. This entanglement can then be used with local operations and measurements for teleported operations [39].

6.1.2 Prior art

Versions of this protocol were first realized with atoms in optical cavities [89, 35]. Given the difficulties involved in efficiently coupling and transmitting optical photons, these demonstrations have been non-deterministic, with success probability below 1%. In the microwave domain, however, it is quite easy to direct emission from a resonator into a transmission line with efficiency as high as 99%. This has led to a rich body of research towards controllable interactions between superconducting qubits and microwave light.

Control of emission from a qubit or resonator into a continuum has been implemented in a variety of ways. The Santa Barbara group demonstrated tunability of the coupling of a resonator to a transmission line which enabled efficient emission of qubit states and absorption of classical light [95, 182]. The tunable coupling is a bit of a shortcut to the Cirac protocol, proposed in [183]. In this approach, one directly varies the decay rate κ, instead of a conversion rate into a communication resonator with fixed κ. This simplifies the dynamics a bit, but the spirit is very similar. A related approach using a qubit with tunable coupling to a fixed resonator, essentially tuning the Purcell decay, was used by the Princeton group to demonstrate temporal control over the emission profile [184]. One downside of these approaches is that they seem to suffer from poor on-off-ratio, requiring additional effort to protect the quantum memories from this tightly-coupled decay channel.

A complementary approach, more closely following the original Cirac proposal, involves microwave-drive-actuated frequency-conversion from a qubit to a communication mode with fixed decay rate into a transmission line. This sort of coupling can be understood as a parametric conversion processes in the language of Section 5.6, but is sometimes called a sideband transition in the literature. This sort of coupling is naturally present in any transmon-resonator system, and does not require explicit tunable coupler elements or flux-tunability, making it nicely suited for 3D cQED. This approach was first used to demonstrate fairly sophisticated temporal control over the emitted photon in the Zürich group [96]. A related driven coupling approach was used in the Devoret group to entangle distant qubits via flying photons [71], though not originally with the intent to demonstrate a state transfer scheme, but instead the kind of measurement-based entanglement scheme discussed in Subsection 2.5.3 [44].

Around this time, we began applying this type of conversion processes in our 3D cavity systems. By driving conversion between the long-lived storage cavity and the lossy readout mode, using the nonlinearity of the ancilla transmon junction, we found we could drastically reduce the lifetime of the cavity, using components which already needed to be present in the system. This tool became an easy and commonly-used way to increase experimental repetition rates [79], rather than wait many milliseconds for the cavity to spontaneously decay. Perhaps more important, this is a useful way to cool the cavity to its ground state, since we often measure appreciable equilibrium population above that which we would expect given the nominal temperature. However, this conversion was expected to be a coherent process, and indeed we demonstrated as much, including entanglement between the propagating field and the state left behind, by measuring the statistics of the emission [98]. Although the conversion made use of the transmon junction, the excitations converted thereby did not need to excite the transmon mode. Consequently, this results in the same kind of linear coupling between harmonic oscillators discussed in Section 4.2, and hence

the conversion process ought to be state-independent for multi-photon states. Indeed we demonstrated as much, seeing clear signatures of the initial quantum states in the emitted field. The ability to faithfully convert these kinds of states suggested we might be able to use this process to implement the Cirac protocol using the encodings discussed in Section 3.4 to correct for photon loss between modules.

6.1.3 Concurrent art

At the same time we were performing the experiment described in this chapter, two related efforts were carried out. Based on the work of [71], the Devoret group demonstrated deterministic entanglement generation between two transmons in separate modules with a directional channel and shaped photons [73]. We collaborated on this, and each experiment benefited from the other. Meanwhile, Andreas Wallraff's group in Zürich built upon their previous work coupling transmons to oscillators [96] to implement a similar experiment [74]. It bears noting that the Zürich experiment uses significantly stronger couplings, and the transfer is nearly an order of magnitude faster than the Yale. However, given the similar losses observed between the modules, the overall performance is comparable between all three experiments.

6.1.4 This chapter

Here we will lay out the details of our implementation, the spiritual successor to [98]. The report is published in [16], and many details can be found in Chris Axline's thesis [15]. I will focus on the essential components and the takeaways which motivated our subsequent work.

In Section 6.2 we discuss the hardware configuration and some design considerations. Then Section 6.3 demonstrates efficient release and capture of temporally-shaped microwave photons. This includes a description of the modifications we needed to make to the shaping protocol described in Subsection 4.4.4. Section 6.4 shows
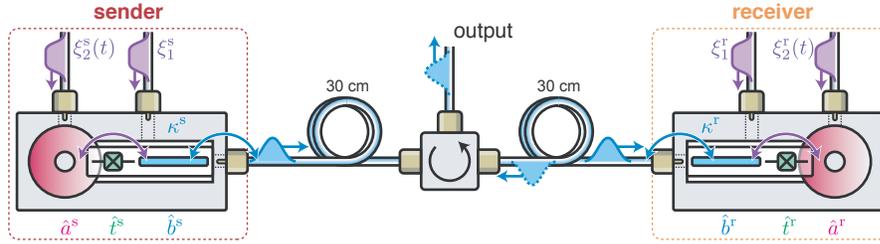
Figure 6.1: Cirac Protocol Hardware. Two nominally identical modules are connected through a uni-directional transmission channel. Each module ($i = $ s, r for "sender", "receiver") contains a 3D storage cavity (a^i, red), transmon (t^i, green), and stripline communication resonator (b^i, blue), also used for readout. Communication resonators couple to transmission line at rate κ^i. Conversion between storage and communication modes is actuated by two off-resonant pumps with strengths $\xi^i_{1,2}$, applied to storage and communication modes, respectively. Figure reproduced from [16] in accordance with Springer Nature copyright permissions.

transfer of arbitrary quantum states in the Fock encoding. Then in Section 6.6 we demonstrate the very minor modification to the protocol which allows for on-demand entanglement generation. Motivated by the dominant imperfection — photon loss — in Section 6.5 we demonstrate the compatibility of this approach with qubits encoded in an error-correctable manifold. Finally, in Section 6.8 we summarize the difficulties involved in this experiment and some of the lessons learned.

6.2 Cirac protocol implementation

The modules used in this experiment are more or less clones of the one used in [98], with some updates to make them more amenable to the scheme here.

6.2.1 Hardware configuration

The layout of the samples is shown in Figure 6.1. Each module contains the essentials described in Chapter 5: a 3D post cavity, a transmon, and a stripline readout resonator. The readout resonators also serve as the communication resonators for state transfer protocol. These resonators are each coupled to a superconducting trans-

mission line which terminates at a circulator, such that the emission from the first module is incident on the second, but the emission from the second is directed to the third port of the circulator, which carries the signal to a measurement chain. This circulation enforces the assumption that the line is a continuum, and that there are no standing wave modes. In the language of Chapter 4, this allows coupling via propagating modes without requiring the line to be long enough that the free spectral range is smaller than the coupling rate. This is important, because that length in this experiment would be longer than 100 meters. The circulator also breaks the symmetry between the roles of the modules, and so we call one the sender and the other the receiver. We label the storage, communication, and transmon modes in the sender (receiver) as a^s, b^s, and t^s (a^r, b^r, and t^r). The communication modes couple to the line at fixed rate κ^s and κ^r, which are similar but need not be exactly matched. The measured system parameters are summarized in Table 7.1.

6.2.2 Parametric conversion scheme

We use a four-wave mixing parametric conversion of the type introduced in Section 5.6. The layout in frequency space is shown in Figure 6.2. Since the detuning between the conversion pumps needs to equal the detuning between the storage and communication modes, it is convenient to locate the pumps symmetrically detuned from those modes by a small amount Δ. We choose $|\Delta|/2\pi$ between 30 and 50 MHz, which allows the pumps to be sourced by the same drive chain as the cavity and readout drives. The pump near the storage we call pump 1, and that near the readout pump 2.

By applying the pumps so close to the linear modes, we consider the coupling to the junction phase to be through these modes. As a result, the relevant Hamiltonian

Hamiltonian parameter (MHz)		Sender value	Receiver value
Frequency	$\omega_{\mathrm{a}}/2\pi$	4219.3	4269.6
	$\omega_{\mathrm{b}}/2\pi$	10031.5	10031.5*
	$\omega_{\mathrm{t}}/2\pi$	6156.1	6417.6
Cross-Kerr	$\chi_{\mathrm{ab}}/2\pi$	-16×10^{-3}	-12×10^{-3}
	$\chi_{\mathrm{ac}}/2\pi$	-2.86	-2.29
	$\chi_{\mathrm{bc}}/2\pi$	-2.4	-2.18
Self-Kerr	$\chi_{\mathrm{aa}}/2\pi$	-8×10^{-3}	-5×10^{-3}
	$\chi_{\mathrm{bb}}/2\pi$	-8×10^{-3}	-6×10^{-3}
	$\chi_{\mathrm{cc}}/2\pi$	-183.43	-196.17
Damping parameter (μs)			
Energy decay time	T_1^{a}	460 ± 10	770 ± 10
	T_1^{t}	26 ± 3	27 ± 3
	T_1^{b}	0.14 ± 0.01	0.11 ± 0.01
Ramsey decay time	$T_{2\mathrm{R}}^{\mathrm{a}}$	102 ± 3	130 ± 4
	$T_{2\mathrm{R}}^{\mathrm{t}}$	12 ± 2	12 ± 2
Hahn echo decay time	$T_{2\mathrm{E}}^{\mathrm{t}}$	15 ± 2	15 ± 2
Steady-state excitation		Sender value	Receiver value
Transmon	$1 - P(g)$	0.195	0.209
Cavity	$\bar{n}$	0.166	0.172

Table 6.1: Pitch and Catch Bus Sample Parameters. Uncertainties of measured Hamiltonian parameters are $< 0.1\%$ except when indicated by fewer significant digits. For the cavity and transmon decay times, the uncertainties given are the typical fluctuations observed over the course of one day. *This is the frequency set by tuning and is used during all phases of the experiment.
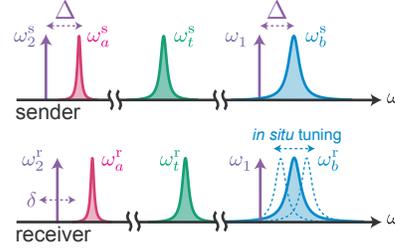
Figure 6.2: Frequency Layout. Off-resonant pumps (purple) are applied with approximately equal detuning Δ from storage and communication modes to enable conversion. Storage modes need not be frequency matched, but communication modes must be approximately resonant. Communication mode of receiver is mechanically tuned *in situ*. Figure reproduced from [16] in accordance with Springer Nature copyright permissions.

term for a single module is

$$\mathcal{H}_{\mathrm{FWM}} = E_{\mathrm{J}} \cos_{\mathrm{nl}} \left(\varphi_{\mathrm{a}}(ae^{-i\omega_{\mathrm{a}}t} + a^{\dagger}e^{i\omega_{\mathrm{a}}t}) + \varphi_{\mathrm{b}}(be^{-i\omega_{\mathrm{b}}t} + b^{\dagger}e^{i\omega_{\mathrm{b}}t}) + \varphi_{\mathrm{t}}(te^{-i\omega_{\mathrm{t}}t} + t^{\dagger}e^{i\omega_{\mathrm{t}}t}) \right.$$
$$\left. + \varphi_{\mathrm{a}}(\xi_1 e^{-i\omega_1 t} + \xi_1^* e^{i\omega_1 t}) + \varphi_{\mathrm{b}}(\xi_2 e^{-i\omega_2 t} + \xi_2^* e^{i\omega_2 t}) \right)$$
$$(6.1)$$

The desired conversion term is

$$\mathcal{H}_{\mathrm{conv}} = E_{\mathrm{J}}\varphi_{\mathrm{a}}^2\varphi_{\mathrm{b}}^2\xi_1^*\xi_2 ab^{\dagger}e^{-i(\omega_a-\omega_b-\omega_1+\omega_2)t} = \chi_{\mathrm{ab}}\xi_1^*\xi_2 ab^{\dagger}e^{-i\delta t} \qquad (6.2)$$

where $\chi_{\mathrm{ab}} = E_{\mathrm{J}}\varphi_{\mathrm{a}}^2\varphi_{\mathrm{b}}^2 = \chi_{\mathrm{at}}\chi_{\mathrm{bt}}/2\chi_{\mathrm{tt}}$ is the cross-Kerr between the storage and communication mode. We have subsumed all of the mode and pump frequencies into a single detuning δ, which denotes how far the pumps are from the nominal resonance condition $\omega_{\mathrm{a}} - \omega_1 = \omega_{\mathrm{b}} - \omega_2$. The prefactor $g \equiv \chi_{\mathrm{ab}}\xi_1^*\xi_2$ is the conversion rate, controlled by the pump amplitudes.

Stark shifts

As introduced in Subsection 5.6.2, there are frequency shifts induced by each pump on each mode. The Stark shifts relevant to the conversion process are

$$\delta_a = 2\chi_{aa}|\xi_1|^2 + \chi_{ab}|\xi_2|^2$$
$$\delta_b = 2\chi_{bb}|\xi_2|^2 + \chi_{ab}|\xi_1|^2 \qquad (6.3)$$
$$\delta_t = \chi_{at}|\xi_1|^2 + \chi_{bt}|\xi_2|^2$$

The last of these, the Stark shift on the transmon, is much larger than the first two, but since this mode does not directly participate in the conversion, the precise value is not important. These large shifts do lead to unwanted resonant transitions involving the transmon [178], which are dependent on the pump amplitudes and frequencies. We try to find pump frequencies and amplitudes which mitigate these transitions.

6.2.3 Frequency tunability

Since the Cirac protocol relies on the emission from the sender being compatible with the receiver, the communication modes of these module must be approximately matched in frequency. Fortunately, while the bandwidth of the emitted photon is on the order of $4g^2/\kappa$ (see Subsection 4.4.2), the tunability bandwidth of the conversion process (the frequency range over which the photon can efficiently be emitted) is κ, the line-width of the communication mode. This is somewhat larger, which makes the frequency matching of the modules more lenient. However, this is still a tighter tolerance than we can reliably make our stripline resonators, so in-situ tunability was needed in this experiment.

The frequency of the communication mode of the receiver module was enabled by introducing a superconducting aluminum pin in a tunnel which intersected the waveguide enclosure of the stripline. The pin acted as an additional capacitance to

ground, so bringing it closer to the device lowered the frequency, The pin was driven by a piezoelectric stage (Attocube ANPz101-A4), and by moving the stage over a range of a few millimeters, we could tune the frequency over a few hundred MHz, with sub-MHz precision. Once an acceptable frequency was attained, the rest of the experiment was performed for fixed pin position. Since the frequency could only be adjusted downward, we introduced a similar, stationary pin for the sender module to ensure it was within the tuning range.

Related experiments with similar frequency-matching requirements [71, 73] used an adjustable screw inserted into a 3D cavity used as a communication mode. The screw could only be adjusted at room temperature, but since these cavities can be measured while warm, it is possible to extrapolate the change in frequency from $300 \, \text{K}$ down to base temperature with a few cool-downs. Since the stripline resonators are too lossy at room temperature to measure accurately, this is not feasible with our samples.

It is difficult to know how much dissipation is induced by the tuning pin. Although it is superconducting, it may not be well-grounded, as contact with the walls of the tunnel can provide too much friction for the piezoelectric stage to overcome. As a result, the pin may be floating in its tunnel, forming a coaxial transmission line through which the communication mode can radiate into free space. We were not able to simply track the line-width of the mode with pin insertion, since the coupling bandwidth can change as the mode profile are perturbed. Additionally, it is quite hard to measure few-percent changes in the bandwidth. So it is possible this tuning mechanism adds some loss, which would result in transmission inefficiency. Flux tunability, or extreme diligence in device assembly, might provide a cleaner route towards the frequency matching requirement.

6.2.4   Joint readout

Since the communication resonators are also used for readout of the transmons, the measurement signal from the sender is incident on the receiver. This means we cannot measure the sender without also measuring the receiver[1]. This sort of configuration was used in similar experiments [45, 71]. The signal layout is shown in Figure 6.3a. To perform simultaneous joint readout, we applied a transmission tone on the sender resonator. The emission is also reflected off the receiver, and so gains information about both transmons. While for certain parameter regimes this might be sufficient, we found it necessary to apply an additional transmission tone on the receiver. The two measurement fields interfere, so the relative amplitude and phase are important. The sum of the two tones exits the third port of the circulator and is measured with a Josephson parametric converter (JPC) [149, 45, 150]. It is important to use a phase-preserving amplifier for this joint readout because we extract two bits of information, which is easiest done by using both quadratures of the signal.

While this joint readout is sufficient for most purposes, it is not ideal. In our realization, there is significant crosstalk between the state assignment of the two transmons. We characterize this in Figure 6.3b,c. We vary the amplitude of a transmon excitation pulse on both modules, and perform the joint readout and state assignment. As shown in the one-dimensional cuts in Figure 6.3c, there is a few percent of spurious signal which depends on the state of the opposite transmon. This crosstalk presents some challenges for the high-fidelity state discrimination required for quantum error correction, though it could be mitigated with more careful tuneup. The sequential style of readout also halves the bandwidth of the measurement signal, slowing down readout. Some of these issues will be addressed with the introduction of independent readout chains in Chapter 7.

---

1. The converse is allowed, and we did measure the receiver on its own for some calibration experiments. But in general we used the joint readout described here for most of our measurements.
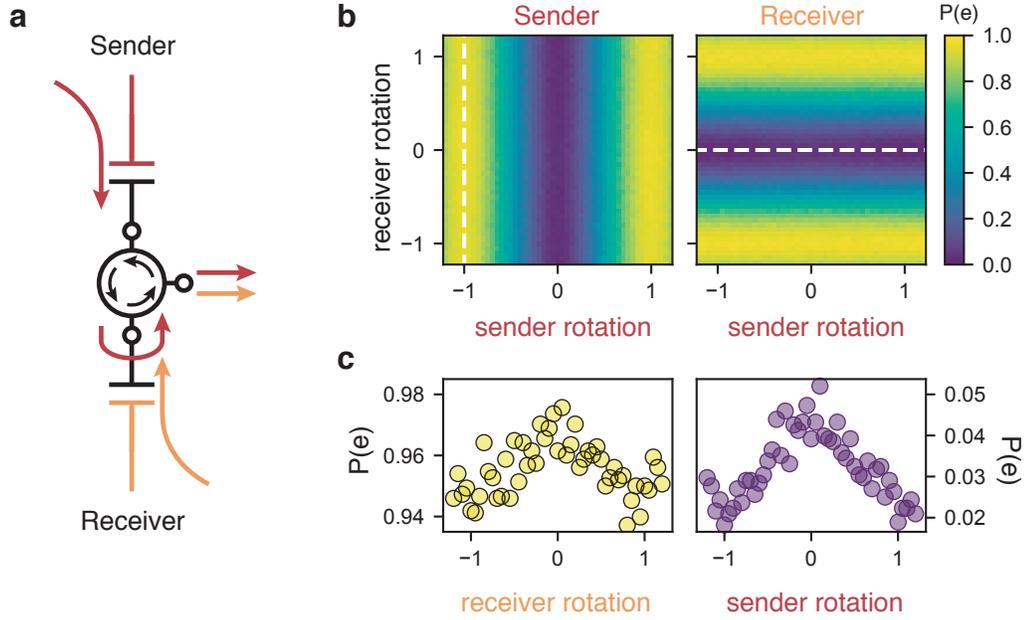
Figure 6.3: Joint Readout. a) Communication channel is also used for transmon measurement. Both communication resonators are driven in transmission at the same frequency. Emission from sender (red) also reflects off receiver, where the two signals combine. Both are routed out the third port of the circulator to a quantum-limited phase-preserving detection chain (not shown). b) Measured signal during simultaneous Rabi experiment. Both transmons are excited by different amounts before joint measurement. c) Cuts along dashed lines in b). Ideal behavior would show no signal versus receiver rotation in sender signal, and vice versa. Data show 2–3% unwanted crosstalk in assignment. Figure reproduced from [16] in accordance with Springer Nature copyright permissions.

6.3 Releasing and capturing shaped photons

Here we detail a few practical extensions of the Cirac wavepacket shaping protocol explained in Section 4.4. To recap, we pick a specified wavepacket shape $b_{\text{out},1}(t)$, and are tasked to find conversion couplings $g(t)$ for the sender which produce this wavepacket, and receiver couplings which absorb it. However, since the coupling g is dependent on the pump amplitudes $\xi_{1,2}$, the task is to find those controls. The additional complication is that the frequencies of the storage and communication modes depend on those amplitudes.

6.3.1 Control scheme

Since the conversion rate g (detuning δ) depend on the product of pump amplitudes (difference of pump frequencies), it is sufficient to vary the amplitude and frequency of only one of the pumps, leaving the other fixed. We choose to vary $\xi_1(t)$, since it is weaker than ξ_2, and hence the dynamic Stark shifts will be smaller if ξ_2 is fixed. We include 200 ns smooth turn-on and turn-off times for this constant pump so that its bandwidth is limited and it does not excite the cavity or communication modes.

The procedure for constructing a numerical solution for the pump is the same as presented in Subsection 4.4.4, but we cast the equations of motion in terms of $g(\xi_1(t), \xi_2)$, with the time dependence of g coming through $\xi_1(t)$. Hence the time derivatives of g in Equation 4.44 are replaced by derivatives of ξ_1. There are also additional terms resulting from the static and dynamical Stark shifts (and derivatives thereof). The full derivation can be found in [15] or the Methods section of [16]; here we just give the result, since the basic procedure is already written in Subsection 4.4.4. The equation to be solved for the sender is

$$
\begin{aligned}
0 = {} & (\dot{g}(t) - i\delta_{\mathrm{a}}(\xi_1)g(\xi_1))\left(\dot{b}(t) + \frac{\kappa}{2}b(t) + i\delta_{\mathrm{b}}(\xi_1)b(t)\right) \\
& - g^2(t)g^*(t)b(t) - g(t)\left(\frac{\kappa}{2}\dot{b}(t) + \ddot{b}(t) + i\dot{\delta}_{\mathrm{b}}(t) + i\delta_{\mathrm{b}}(\xi_1)\dot{b}(t)\right)
\end{aligned}
\tag{6.4}
$$

The equation for the receiver is similar, as explained in Subsection 4.4.4.

The functional form of the Stark shifts are given in Equation 6.3, and it is convenient to be able to to take the time derivative analytically, e.g. $\dot{\delta}_{\mathrm{b}} = 2\chi_{\mathrm{ab}}|\xi_1|\frac{\mathrm{d}|\xi_1|}{\mathrm{d}t}$. The same is true of the functional form of $g(\xi_1)$, though we found deviation from the simple form of Equation 6.2. At higher pump powers, we found from calibration experiments that we also needed to include a small dependence on higher powers of the pump strengths. This likely comes from higher order terms in the cosine, or the effect of the Stark shifts changing the detuning, and hence the effective pump strength ξ.
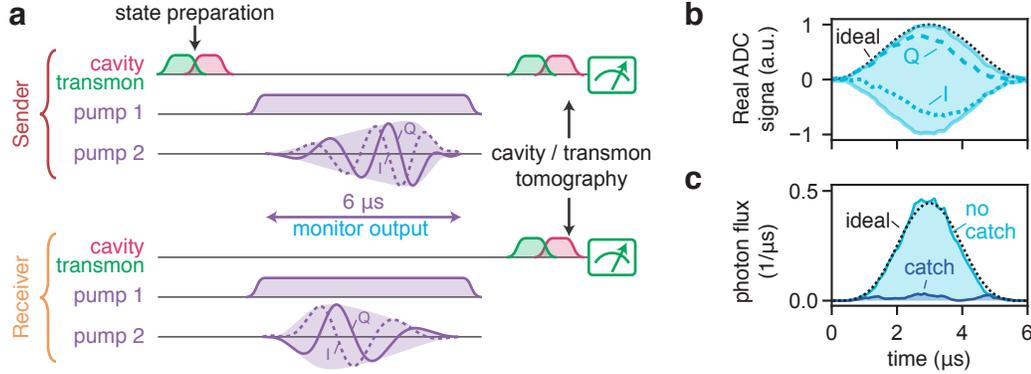
Figure 6.4:   Release and Capture of Shaped Wavepackets. a) Experimental sequence consists of state preparation in receiver, application of shaped conversion pumps, and measurement. Propagating signals are measured by quantum-limited heterodyne detection chain. b) Measured heterodyne signal for release of a coherent state $|\alpha = 1\rangle$, with no capture pulses applied. In-phase (I, dotted), quadrature (Q, dashed), and amplitude (solid, shaded), show agreement with ideal wavepacket shape (black, dotted). c) Measured photon flux ($I^2 + Q^2$, normalized) for coherent state. With capture pumps applied ("catch," dark), measured reflected signal is greatly suppressed. Figure reproduced from [16] in accordance with Springer Nature copyright permissions.

The maximum g used in this work was about $2\pi \times 400\,\text{kHz}$.

6.3.2   Wavepacket release and absorption

To verify the control of the emitted wavepacket, we prepare a coherent state in the sender cavity, and apply the shaped release pumps, as shown in Figure 6.4a. The pumps shown in this figure are the numerically-computed shapes. We measure the emitted field, which reflects uncaptured off the receiver and goes out the third port of the circulator to our quantum-limited detection chain. The average signal, measured in heterodyne [98] and shown in Figure 6.4b, shows very close match to the ideal shape.

To demonstrate absorption of the emitted signal, we perform the same experiment, while simultaneously applying the pumps on the receiver. Figure 6.4c shows the measured power (the absolute square of the average heterodyne signal) with and without the capture, showing a large reduction in the reflected power. The fractional change in reflected energy is a direct measure of the receiver efficiency, which is
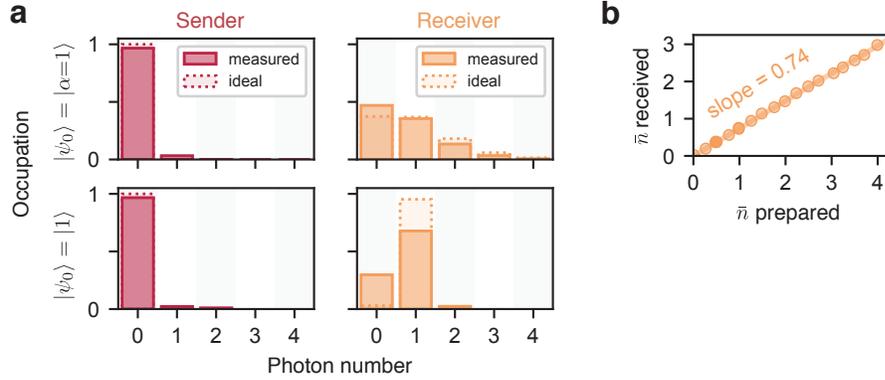
Figure 6.5: Single-photon Efficiency. a) Measured cavity occupations in Fock states $n = 0$–4 after release and capture for coherent state $|\alpha\rangle = 1$ (top) and Fock state $|1\rangle$ (bottom). b) Extracted efficiency for states of varying mean photon number $\bar{n}$. Slope of line is efficiency η. Figure reproduced from [16] in accordance with Springer Nature copyright permissions.

$\eta_{\mathrm{r}} = 0.93(1)$.

6.3.3 Total efficiency

Having established the ability to release and capture a coherent state, we can evaluate the end-to-end efficiency of the process by measuring the energy in both cavities after capture be measuring the number-split transmon spectrum (see Subsection 5.5.1). The results of this experiment, for a coherent state with $\bar{n} = 1$ and Fock state $|1\rangle$, are shown in Figure 6.5. For both states, we find the sender cavity is with high probability emptied — the sender efficiency is $\eta_{\mathrm{s}} = 0.97(1)$. The measured mean photon number in the receiver yields the total efficiency $\eta = 0.74(3)$ for both states. By preparing and transmitting a variety of coherent and Fock states we find that the efficiency is state-independent up to $\bar{n} = 4$ (Figure 6.5c). This is essential for the ability to transmit bosonic error-correction code words, and a result of the fact that the conversion process does not necessarily excite the transmon.

6.3.4 Contributions to inefficiency

While the measured efficiency is high enough to enable quantum communication, it is important to understand the origins of the loss. The major contributors are unwanted excitation of the transmons and transmission loss.

Transmon excitations

As stated above, the transmons which enable the conversion process ideally do not become excited during release and capture. However, we observe significant excitation rates of both. A large fraction of that comes from their rather high equilibrium occupation ($\sim$20%) and short lifetimes ($\sim 25\,\mu$s), resulting in an excitation rate of $\sim 1/(100\,\mu$s). Given the $6\,\mu$s transfer time, and imperfect initialization, we expect as much as 5-10% excitation of these transmons by the end of the transfer time. Additionally, the pumps add to this excitation rate [178]. The result is that the receiver transmon has an 87% probability of being found in its ground state at the end of the transfer, and 91% for the sender.

These excitation events have two effects. The first is that occupation of the transmon $|e\rangle$ state dispersively shifts the storage and communication modes by χ_{at} and χ_{bt} respectively, which are both larger than the conversion rate g. Since both modes are shifted, the conversion is still resonant to the extent that $\chi_{at} = \chi_{bt}$; they are not exactly matched but are fairly similar, such that the added detuning is of order g. In this limit, the conversion still mostly happens, but the photon is released at the wrong frequency if the sender transmon is excited, and not captured.

Fortunately, not all excitations events completely extinguish the transfer, since, for instance, an excitation which occurs in the last microsecond of time will have little effect. This effect is estimated to cause a loss of the photon about half of the time when the sender transmon is excited, which means the true sender efficiency $\eta_s = 0.93(1)$. Similarly, if the receiver transmon is excited, the capture process is

not resonant with the photon which arrives, and so it is not well-absorbed. This is a dominant contribution to the measured receiver efficiency.

The second effect of the transmon excitations are to confound our measurement of the cavity state in the receiver. Since the transmon spectrum is used to perform tomography on the cavity, stochastic excitation events primarily reduce measurement contrast. Since we always normalize the spectrum to give physical probabilities for the number state occupations (as in Figure 6.5a), we can say little about the occupation of the cavity when the receiver transmon is excited. For this reason we say that the measured efficiency is conditioned or heralded on the transmon being found in its ground state. This places a finite success probability on the protocol, which is not fundamental to the scheme but is important in this realization. That success probability is $p_\mathrm{s} = 0.87(3)$, the probability of finding the receiver transmon in $|g\rangle$ quoted above. For all numbers where this is relevant, we give the directly measured value, which is implicitly conditioned on this success, and an estimated deterministic number which conservatively assumes complete failure in the event that the transmon is excited. For the efficiency, this deterministic number is $\eta_\mathrm{d} \geq p_\mathrm{s} \times \eta = 0.87 \times 0.74 = 0.64(3)$. In what follows we present the deterministic (directly measured) value, with the conservative lower bound of the fully deterministic value in parentheses.

Other inefficiencies

The dominant additional contribution to the inefficiency is loss in the transmission path, including connectors, cabling and circulator. This is difficult to measure precisely, but we estimate it by performing a control experiment which looks at the transmission of the pump ξ_2 from sender to receiver, which is near the communication frequency. By comparing the relative Stark shifts of the two transmons, we can estimate the attenuation of the pump leaking from one system to the other. From this experiment we find a transmission efficiency $\eta_\mathrm{tx} = 0.80 \pm 0.15$. The large error comes

from the rather large uncertainty in the cross-Kerrs χ_{bt} which are used to convert the Stark shift into a drive amplitude. In the related experiment performed in the Devoret group [73], a similar calibration was done *at* the communication frequency by measuring the relative dephasing of the transmons due to emission from the sender; a similar value was found, though that experiment had several additional components in the transmission path.

6.4 Quantum state transfer

Given the measured transmission efficiency, we expect to be able to transfer quantum states from sender to receiver. In particular, we can encode an arbitrary quantum bit in the sender cavity, and should transmit it faithfully to the receiver. Since we showed the transfer process to be state-independent, we have a choice of encoding. To demonstrate this capability, we used the simplest, the Fock code, with $|\pm z_{\mathrm{L}}\rangle = \{|1\rangle, |0\rangle\}$. To keep consistent with the notation of previous chapters, we have permuted 0 and 1 in the above definition, because in this chapter we identify the top of the Bloch sphere $|+z\rangle$ with the Fock state $|1\rangle$.

6.4.1 State preparation

We prepare the six cardinal states of this encoding in the sender cavity using numerically optimized control pulses (OCP). Two of the prepared states, characterized by Wigner tomography, are shown in Figure 6.6a. Since one of the code words is the vacuum state, we do not need a control pulse to generate this state. Since the equator states (equal superpositions of $|0\rangle$ and $|1\rangle$) are the same up to the phase of the cavity, these states are generated with a single OCP. The phase of the cavity drive during the pulse dictates the phase of the final state. All told, only two distinct numerical OCPs are used to prepare these six states.
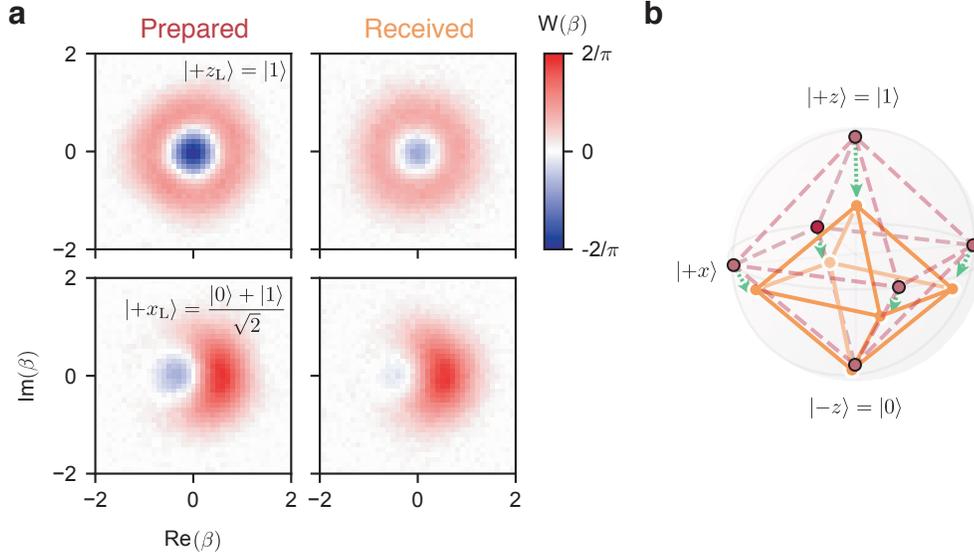
Figure 6.6: **Quantum State Transfer Between Modules.** a) Measured Wigner function of states as prepared before transfer (left) and received after transfer (right). Two states in Fock encoding are shown. b) Reconstructed logical Bloch sphere. Prepared states shown in red, received in orange. Green arrows denote expected movement due to amplitude damping with measured efficiency. Figure reproduced from [16] in accordance with Springer Nature copyright permissions.

6.4.2 Received fidelity

For each transferred state, we measure the resultant Wigner function. Each Wigner function is normalized and reconstructed using a convex optimization technique [14]; two of those received Wigner functions are shown in Figure 6.6a. The resulting mean state fidelity, averaged over all six states, is $\mathcal{F}_{\text{Fock}} = 0.87(4)$ (deterministic: $\mathcal{F}_{\text{Fock,d}} \geq p_{\text{s}} \times \mathcal{F}_{\text{Fock}} = 0.76(4)$. The implicitly conditioned and deterministic fidelities both exceed the classical bound of 2/3, which is the best fidelity with which one can reconstruct an unknown quantum state with only classical communication [185]. Thus we have demonstrated deterministic quantum communication between remote modules.

In addition to the average fidelity, it is instructive to see how each state evolves under the transmission channel. Figure 6.6b shows a representation of the Bloch sphere with all six reconstructed states as prepared and received. The states show

clear decay towards $|0\rangle$ as would be expected from a lossy channel with the measured efficiency. Moreover, the measured fidelity $\mathcal{F}_{\text{Fock}}$ is in good agreement with the fidelity expected from a pure amplitude damping channel with this efficiency, $\mathcal{F}_{\text{Fock,expected}} = 0.91(1)$, as estimated and calculated in Subsection 3.2.2 (see Figure 3.3).

6.5 Transferring error-correctable states

Given that the heralded transfer fidelity is primarily limited by photon loss, this is an opportunity to attempt to improve the communication with error-correctable bosonic codes. As a demonstration of this capability, we use the lowest-order binomial code, which, as defined in Subsection 3.4.4, has $|\pm z_{\text{L}}\rangle = \{(|0\rangle + |4\rangle)/\sqrt{2}, |2\rangle\}$ Since the transfer process is state-independent, this amounts to simply preparing new input states.

6.5.1 State preparation

The cardinal states in the binomial encoding are prepared using OCPs as in the Fock code. Since both code words are nontrivial states, we need a pulse for each of the north and south poles. Additionally, the equator states are not all related by a cavity phase shift, so the logical states $|+x_{\text{L}}\rangle$ and $|+y_{\text{L}}\rangle$ are made with different OCPs. However, due to the symmetry of the code, $|+x_{\text{L}}\rangle$ and $|-x_{\text{L}}\rangle$ are related by a $\pi/2$ cavity phase, and likewise for $|\pm y_{\text{L}}\rangle$, so we need four pulses to prepare the six cardinal states. Three of these logical states are shown in Figure 6.7a.

6.5.2 Transfer fidelity

Using the same sequence to transfer these states, we again measure the resultant Wigner functions for each of the six states. Three of these are shown in Figure 6.7a. The measured fidelity is $\mathcal{F}_{\text{bin}} = 0.54(4)$ (deterministic: $\mathcal{F}_{\text{bin,d}} \geq p_{\text{s}} \times \mathcal{F}_{\text{bin}} = 0.47(4)$).
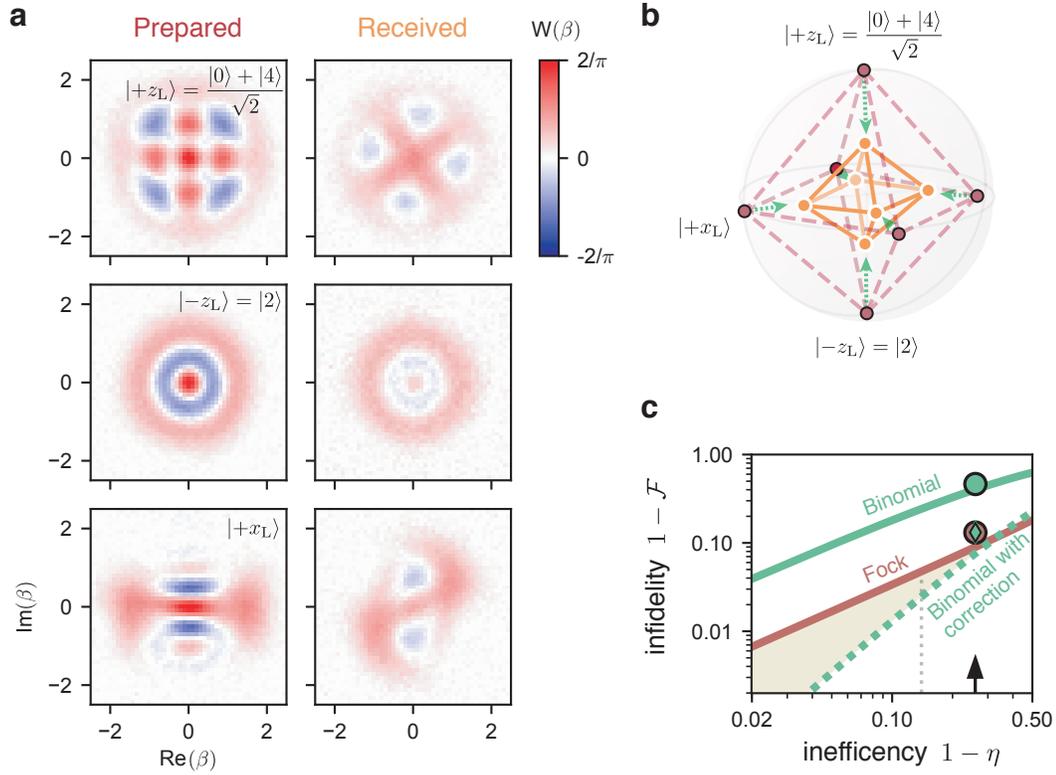
Figure 6.7: Error-Correctable State Transfer. a) Measured Wigner function of states in binomial encoding as prepared before transfer (left) and received after transfer (right). b) Reconstructed logical Bloch sphere. Prepared states shown in red, received in orange. Green arrows denote expected movement due to amplitude damping with measured efficiency. c) Calculated (lines) and measured (points) infidelity versus inefficiency for Fock and binomial codes. Dashed line and diamond denote projected performance of binomial code with ideal error detection. Shaded region is beyond break-even where corrected binomial code can out-perform that off uncorrectable Fock code. Figure reproduced from [16] in accordance with Springer Nature copyright permissions.

This again agrees with the expected fidelity from the loss channel, $\mathcal{F}_{\text{bin,expected}} = 0.60(3)$. The Bloch sphere representation in Figure 6.7b shows a characteristic symmetric shrinkage towards the origin, since the dominant error is leakage outside the code-space into the odd manifold. The error is roughly equal for all states, since they have equal mean photon number.

Due to the overhead of the larger photon number ($\bar{n} = 2$) of the binomial code, the measured fidelity is lower than in the Fock encoding. However, the parity structure of this code should allow for detection and correction of single-photon loss errors.

While we are not able to perform the parity measurement with high fidelity in this sample (see Section 6.7), we have access to the full density matrix for each state from the Wigner reconstruction. To asses the extent to which we could enhance the fidelity with parity measurement, we can perform error correction "in software." We project the density matrix onto the even (odd) subspaces and asses the fidelity within the ideal code space (error space). The even parity no-loss encoding includes the effect of deformation on the code word $|+z_\mathrm{L}\rangle = (|0\rangle + |4\rangle)\sqrt{2}$, which becomes $(\sqrt{1+\epsilon}\,|0\rangle + \sqrt{1-\epsilon}\,|4\rangle)\sqrt{2}$ after no photon loss, for some positive ϵ, which we determine in post-processing. The resulting fidelity assuming ideal parity measurement would be $\mathcal{F}_\mathrm{bin,corr} = 0.87$, which is equal to the measured fidelity in the Fock encoding.

The measured and extracted heralded infidelities, along with calculations of the infidelity versus inefficiency are plotted in Figure 6.7c. This plot, akin to the ones in Chapter 3, shows that our measured inefficiency is somewhat below the break even point for the binomial code, $\eta \approx 0.67$, which means that with no further improvements to the losses, it may be possible to beat the Fock code with error correction. Note that all measured values fall slightly above the expected calculated lines. This is likely due to errors such as imperfect state preparation, and the fact that stochastic transmon excitations may induce additional dephasing errors, even if the state is still partially captured. These errors are thus important to reduce in future implementations, to preserve the simple error structure needed to successfully apply this type of error correction.

6.5.3 Cavity self-Kerr

In addition to loss of contrast and blurring of fine features due to loss, the received Wigner functions shown in Figure 6.7 show distortion and rotations due to the self-Kerr of both sender and receiver cavities. The Kerr unitary $U_K = e^{iKa^\dagger a^\dagger aa/2}$ applies a number-dependent phase $\phi_n = Kn(n-1)/2$ to each Fock component. The codeword

$|-z_\mathrm{L}\rangle = |2\rangle$ is undistorted because it consists only of a single Fock state, and so no phase is visible. Since the effect of Kerr on a state with two Fock components is a relative phase, the effect on the state $|+z_\mathrm{L}\rangle = (|0\rangle + |4\rangle)\sqrt{2}$ is indistinguishable from a rotation. However, $|+x_\mathrm{L}\rangle$ has three Fock components in it, so the evolution is a rotation and shearing which distorts the state.

In this encoding, the Kerr evolution would cause an error were it not deterministic. Because the effect is unitary, it can be compensated for, and primarily results in a redefinition of the code space. After reconstructing the states from the Wigner functions, we compensate for this by applying the opposite unitary in software, resulting in the fidelities quoted above. This compensation could also be applied in real-time [163, 79, 66].

There is a loss in fidelity due to the Kerr effect. Although it is a unitary evolution, the Kerr Hamiltonian does not commute with the photon loss operator a. This results in some uncertainty in the amount of Kerr evolution undergone in the single-loss case (odd manifold), which is a kind of dephasing. We estimate this effect to contribute 1-2% infidelity in the in-software error-corrected performance $\mathcal{F}_{\mathrm{bin,corr}}$ by computing the uncertainty on the phase from the measured amount of Kerr evolution and loss.

6.5.4 Outlook for bosonic codes

The final conclusion to be gained from this projection of the error-correction performance is that modest improvements in the efficiency should allow drastic enhancement of the fidelity. This is because the infidelity of the corrected binomial code is *second order* in the inefficiency. This means that, further below the break-even point, there are large gains to be made. Possible improvements to this scheme which may allow this are discussed in Section 6.8.
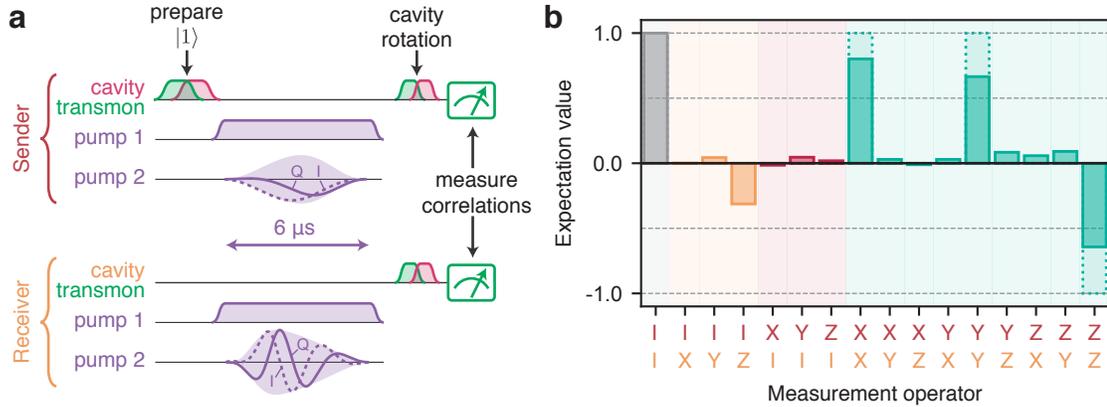
Figure 6.8: On-demand Entanglement. a) Preparation of single-photon Bell state by half-release. Capture pumps are unchanged from state transfer. b) Reconstructed logical two-qubit joint Pauli operator expectation values. Dashed bars denote ideal Bell state. Orange and red are single-qubit operators, green are two-qubit. Figure reproduced from [16] in accordance with Springer Nature copyright permissions.

6.6 On-demand entanglement generation

In addition to state transfer, we can use the tools here to generate entanglement between the two modules. This is of use in an entanglement-based modular architecture, where Bell states which span between modules can be used to teleport qubit state, or to perform teleported gates between logical qubits across the network [39].

6.6.1 Partial release

A Bell pair can be created between the sender and receiver cavities by half-release of a single photon, as described in Subsection 4.4.5. We prepare a single photon in the sender, then apply pumps computed to release only half the energy, but with the same temporal envelope. The pump profile used is shown in Figure 6.8a. Since the dynamics are linear in the field incident on the receiver, the same capture pulses used for state transfer are used here as well.

6.6.2 Entangled state tomography

In order to assess the fidelity of the resultant Bell state, it is necessary to measure correlations between the two cavities. After preparing the Bell state, we measure the cavity states in the z, x, and y bases.

The measurement in the z basis is performed by applying selective π pulses to the transmon, then measuring the state of the transmon. Since the readout and selective π pulse contrasts are both less than unity and asymmetric, we use the $n = 0, 1,$ and -1 number peaks to provide a symmetric and normalized measurement. The $n = -1$ measurement is effectively a measurement of the background transmon excitation, which must be subtracted out, or it will yield false correlations. This subtraction is equivalent to the implicit heralding discussed above for state transfer, but it conditions on both transmons, so the success probability is lower: $p_{\text{s,ent}} = 0.78(4)$.

To measure in another basis, we apply an OCP with maps the x or y eigenstates onto $|0\rangle$ and $|1\rangle$. This corresponds to a logical $Y_{\pi/2}$ or $X_{\pi/2}$ gate on the cavity. We measure in all $3 \times 3 = 9$ combinations of bases, and in each basis we measure the $3 \times 3 = 9$ number peak measurements described above. For each pair of bases, we assess the correlations to extract the occupation probabilities for the four diagonal elements of the density matrix in that joint basis. The combination of 9 basis pairs is sufficient to reconstruct the full density matrix using maximal likelihood estimation. The data described is fed into a reconstruction routine developed by Kevin Chou and described in Section 5.5 of his thesis [14]. The routine produces a physical, normalized density matrix, corresponding to the normalization of the data described above. As noted, this implicitly heralds on both transmons remaining in the ground state, and effectively normalizes on the cavities remaining in the logical subspace. In a sense, this method corrects for all measurement infidelity except for the infidelity of the OCP rotations for measuring in the x and y bases. See Subsection 7.6.2 for further discussion.

6.6.3 Results

The expectation values of the joint Pauli operators for the reconstructed state are shown in Figure 6.8b. The data show dominant two-qubit expectation values expected for the state $|O_+\rangle = |01\rangle + |10\rangle$, with mostly small non-ideal correlations. The only significant non-ideal Pauli expectation value is $\langle IZ \rangle$, which is polarization of the receiver cavity, and is negative. This is exactly the bias towards zero photons we expect from transmission loss, which indicates that the entangled state is primarily limited by loss. The fidelity to the ideal Bell state is $\mathcal{F}_{\mathrm{Bell}} = 0.77(2)$ (deterministic: $\mathcal{F}_{\mathrm{Bell,d}} = p_{\mathrm{s,ent}} \times \mathcal{F}_{\mathrm{Bell}} = 0.61(2)$. Both of these values are well above the threshold for verified entanglement, $\mathcal{F} = 0.5$, so we can cay we have demonstrated deterministic entanglement between modules.

6.7 Limitations and drawbacks

While this experiment successfully demonstrates on-demand entanglement and deterministic transfer of a quantum bit, as well as a route towards future improvement with bosonic codes, the scheme itself has some undesirable qualities, which we discuss briefly here.

6.7.1 Loss

As we discussed above, this and related experiments [73, 74] suffer from significant losses in the transmission channel. Fortunately, this was expected, and there are many ways of dealing with this limitation. As we showed here, there is potential to suppress the effect of losses with bosonic error correction. A complementary approach for enhancing the Bell state fidelity by detecting photon loss involves two rounds of pitch and catch using so-called *time-bin entanglement*. Such a protocol was recently demonstrated in cQED in the Zürich group [159]. In this experiment, the error

detection was performed in software much like we did here with the binomial code. However, recently a variant on this with a much simpler error-detection scheme was realized in the Devoret group, where it was shown that this time-bin approach can indeed improve the fidelity by heralding on no loss [186]. Additionally, generation of multiple entangled pairs and entanglement distillation provides a route towards further improvement of the fidelity.

However, all of these approaches benefit greatly from reduced loss. As shown in Chapter 5 and in [123], the attenuation expected from a single pass through the meter lengths of superconducting coaxial cable employed here is much less than 1%. The loss then likely comes from couplers, connectors, and circulators. There is a large body of work towards all-superconducting circulators [187, 188, 189] which used Josephson junctions and parametric modulation as their source of non-reciprocity. While it has not yet been demonstrated that a superconducting circulator can have lower insertion loss than a commercial ferrite one, there is reason to believe that this can be the case. One important aspect of this will be to ensure that the packaging is as low-loss as possible, lest we lose any gains from the superconducting internal circuitry.

6.7.2 Time

One of the drawbacks of the Cirac protocol (or the related proposal by Korotkov [183, 95, 182]) which use a coupling to a continuum is that they necessarily rely on an exponential decay to release a state into the transmission line. As such, if there is a maximum coupling rate κ, any protocol which releases the entirety of the state needs to take at least several time constants $1/\kappa$. The same is true of the capture process, which is the time-reverse process. So this time limit applies to entanglement by partial release as well.

For instance, in this experiment, the Rabi time corresponding to the maximum

conversion strength is $\pi/2g \approx 600\,\text{ns}$, and the time constant for the communication mode decay is $1/\kappa \approx 160\,\text{ns}$, but the transfer time is $6\,\text{µs}$, an order of magnitude slower. While increasing these rates can speed up the transfer (see [74], where the total transfer time is only $180\,\text{ns}$), the time is still much slower than the coupling rates might imply. This time can be substantially reduced for the same coupling rates using resonant swapping dynamics instead of exponential decays, as we will show in Chapter 7.

Further, and somewhat counter-intuitively, for fixed conversion rate g_{max}, increasing the communication resonator bandwidth κ actually slows down the protocol if $\kappa > g_{\text{max}}$. This is related to the physics of the instantaneous effective decay rate $4g^2/\kappa$, as described in Subsection 4.4.2. In this over-damped regime, which is where the Cirac protocol operates[2], a larger bandwidth reduces the occupation of the communication mode, which slows down the effective dynamics.

6.7.3   Frequency matching

Another consequence of the use of flying photons is that the sender and receiver must be able to communicate at the same frequency. While the channel itself is of extremely wide bandwidth, the communication modes are rather narrow-band. As a consequence, frequency tunability or matching in fabrication is essential, and adds complexity and constraints to the design of the modules. In Chapter 7 we will present an approach which requires no frequency matching, though it of course comes with its own design trade-offs.

---

2. For fixed κ, increasing $g > \kappa$ doesn't really help. This is because, in this case, the speed bottleneck is $\kappa/2$, the rate at which field decays into the transmission line.

6.7.4 Transmon excitation

As we detailed in Subsection 6.3.4, a major contributor to transfer inefficiency is unwanted excitation of the transmons. The thermal populations in these samples was unusually high, and it is possible to do much better, though there is still some work to be done on doing so in a reproducible way. Additionally, some progress has been made on understanding and predicting the pump-induced transitions [155, 178], so we hope to be able to systematically reduce these. However, using transmons or elements with appreciable dispersive shifts to the storage and communication modes will always present this problem to some extent. The use of Kerr-free three-wave mixing elements [177] might be a workaround, but the addition of DC flux bias presents engineering challenges in the 3D cQED architecture, as well as an additional dephasing channel which may add appreciable uncorrectable errors.

6.8 Future directions and lessons learned

What would be the next step towards better communication with flying photons? As discussed above, we would really like to be able to implement the error-correction protocol. There are some barriers to that in this module design. We'll go into some of those challenges, and then from there discuss some slightly more complicated structures which introduce a lot of new design avenues.

6.8.1 Effecting good parity measurements

The ability to perform rapid and high-fidelity measurements of the parity of the cavity is dependent on a few critical features:

– Fast parity mapping ($t_{\mathrm{p}} = \pi/\chi_{\mathrm{at}}$)

– Rapid QND measurement of the transmon state (requiring relatively large χ_{bt}

and κ)

– A transmon measurement which does not dephase the cavity

This last requirement is somewhat at odds with our conversion process require-ments. This is because photons in the readout resonator can "measure" the cavity state, due to the cross-Kerr $\chi_{ab} \approx \chi_{ab}\chi_{bt}/4\chi_{tt}$ between the two modes. This coupling is small compared to κ, so the dephasing rate [146] is approximately $\bar{n}_m\chi_{ab}^2/\kappa$, where $\bar{n}_m$ is the mean photon number in the readout resonator during measurement, and is typically of order 5 to 10. This means that if χ_{ab} is too large, the measurement of the transmon will somewhat dephase the cavity, and our parity measurement in-troduces uncorrectable errors. This effect can be minimized by appropriate choice of these parameters [65]. However, the strength of the conversion process g is di-rectly proportional to χ_{ab}, so it is not possible to optimize for both of these concerns at the same time. This added dephasing, combined with the complications of joint readout discussed in Section 6.2.4, make it challenging to implement fast, accurate, nondestructive parity measurements in the current hardware.

6.8.2 New module concepts

Given this tension between conversion and parity measurements, it is likely a good idea to separate this functionality. In Figure 6.9a, we present a design for a module with two transmons coupled to the storage cavity. One can be used for state prepara-tion, tomography, and syndrome measurements, and the other can be optimized with stronger couplings and a tunable-frequency communication resonator to allow rapid state release and capture. This topology also means the communication channel is not used for readout, and so the two modules can have independent measurement chains, simplifying the tuneup of measurement.

This idea of separation of concerns is a powerful and flexible one. In fact, we have
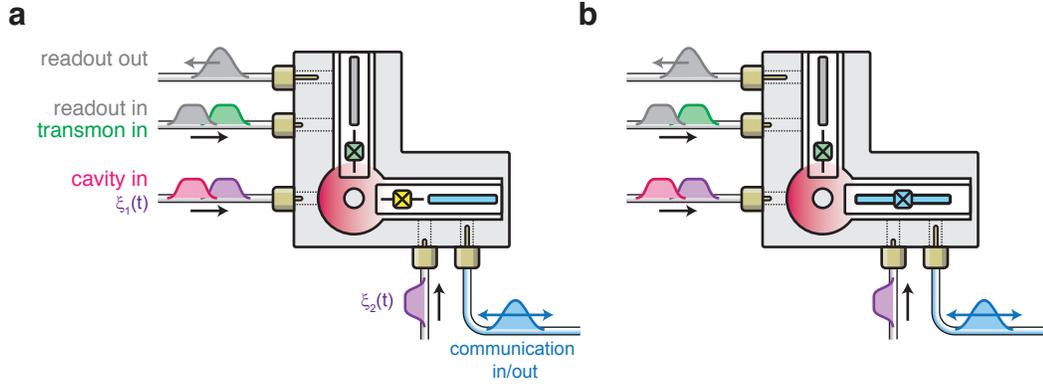
Figure 6.9: Possible Future Modules. a) Addition of a dedicated ancilla and readout transmon (top) separates error-detection circuitry from communication hardware. b) By embedding a Josephson junction in the communication resonator, a conversion transmon may not be needed.

a lot more freedom in this design. For instance, the communication arm of the sample now no longer needs a transmon *per se*. All it needs is a way to effect parametric conversion between the cavity and communication resonator. This can be implemented by embedding a Josephson junction directly into the stripline communication resonator, and using this as the source of the parametric nonlinearity. This reduces the number of modes coupled to the pump, easing some frequency-crowding concerns, and gets rid of this transmon which can become excited to ill effect. Additionally, by making the communication mode the most nonlinear part of the system, it changes the optimization landscape quite a bit. This kind of low-Q, mostly linear resonator[3] bears similarity to the kinds of elements found in parametric amplifiers, where much larger mixing strengths are regularly attained. So this may be a viable approach for increased conversion strength.

This approach is not without its challenges. One concern is that the low-Q communication mode is directly coupled to the storage mode, whereas before there was a transmon filtering the coupling. This may induce excess Purcell decay on the cavity, and necessitate careful engineering of the coupling, or adding a Purcell filter [143, 122].

---

3. Linear is meant in the sense that $\chi_{\rm bb} \ll \kappa$, which is required to maintain state-independence.

Moving the nonlinearity into the communication mode would weaken the participation of the storage mode in the junction, requiring stronger pumps. These pumps will induce larger Stark shifts on the communication resonator, which still needs to be frequency matched to the other module. This might be a feature or a bug, depending on if the relative Stark shift can be used for fine-tuning the frequency-matching.

Whether these alternate module designs could facilitate higher fidelity communication with the Cirac protocol is an open question. In the end, the added flexibility opened up by a dedicated readout ancilla led us to a different approach which forgoes propagating photons. This experiment is detailed in Chapter 7.

Chapter 7

The Quantum Network Bus

7.1   A low-loss, bi-directional link

In this chapter we present the final results of this thesis, in which we demonstrate a two-node bi-directional network with a quantum bus connection. A manuscript describing these results is currently in preparation [18]. Motivated by some of the challenges encountered in our implementation of the Cirac protocol for directional state transfer in Chapter 7, use this simpler communication scheme and find superior performance.

In Section 7.2 we describe the hardware for this experiment. Next, Section 7.3 will explore the control landscape of the bus. Following the theoretical description in Section 4.5, we demonstrate the tunable beamsplitter with the bus, and find two working points. Section 7.4 will use the first working point, which realizes a SWAP between modules. Then in Section 7.5, we use an error-correctable logical encoding and a single syndrome measurement to demonstrate error-tracked state transfer at the break-even point.

In Section 7.6, we turn to the problem of generating shared entanglement across the network, for which we use a 50:50 beamsplitter operation. Here we generate a

high-fidelity single-photon Bell state, the same one created in Chapter 6. However, the bi-directional nature of the bus allows us to generate a two-photon Bell state with Hong-Ou-Mandel interference, detailed in Section 7.7. Critically, thanks to the parity structure of the state, we are able to detect photon loss errors which occur during the interference. By post-selecting against these errors, we obtain probabilistic entanglement with higher fidelity that we can achieve with a single photon. Finally, in Section 7.8, we discuss some of the possibilities, opportunities, and hurdles for scaling, improving, and extending this flexible networking tool.

7.2 Physical implementation

In Chapter 6, we realized state transfer and entanglement generation with a uni-directional, nonreciprocal communication channel. We found the performance to be largely dominated by loss in the link, mostly due to the circulator used to enforce directionality, and possibly the connectors which join the transmission line to the modules and circulator. This provides some of the motivation for moving towards a link without a circulator, and with all-superconducting connections.

We were also inspired by the work by the Schuster group in Chicago [75], where a finite length of coaxial cable was used to form a bi-directional channel for rapid entanglement generation. In that experiment, a clever approach was taken to avoid the rather large loss of the standing-wave cable mode, but this scheme requires additional on-chip resonators which must be frequency-matched to one another. Additionally, while this trick is quite successful, the protocol still suffers somewhat from the loss of the cable mode, and has inherent speed limits due to the presence of lossy nearby resonances.

The approach taken in this chapter is a significant simplification of the Chicago experiment, enabled by the use of a low-loss bus resonator. The simplicity enables
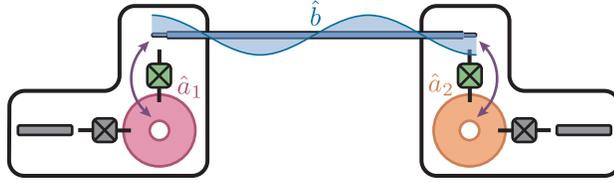
Figure 7.1: Quantum Bus Hardware. Two nominally identical modules $i = 1, 2$, with storage cavities a_i (red, orange), coupled via a quantum bus b. The $n = 3$ mode, voltage profile shown, is used in this experiment. Cavities couple to bus simultaneously via parametric conversion, mediated by the nonlinearity of a conversion transmon (green). Conversion is induced by applying two drives (not shown) directly to the conversion transmon. Each module has an ancilla transmon and readout resonator (gray) with Purcell filter (not shown).

a variety of straightforward mechanisms for entanglement generation, and the use of bosonic qubits provides us means for using error-detection to further suppress the effects of remaining losses.

7.2.1 Hardware

The hardware schematic for our simple network is shown in Figure 7.1. The modules $(i = 1, 2)$, described in more detail in Section 5.3.3, consist of a single 3D cavity memory (a_i). Each cavity has a dedicated ancilla transmon and readout resonator, used for state preparation, tomography, and error syndrome measurement. The ancillae have separate output chains, a simplification from some of the difficulties encountered in Chapter 6. Each cavity also has an additional conversion transmon (c_i) on a separate chip. This conversion transmon couples capacitively to the storage cavity, and to a $5.5\,\mathrm{cm}$ NbTi coaxial cable resonator. Details of the mounting scheme, and characterization of the bus resonator, can be found in Section 5.8.2. The measured system parameters are summarized in Table 7.1.

The conversion transmon is driven with two blue-detuned pumps to actuate parametric conversion between the storage cavity and a single mode of the cable b, which is the $n = 3$ resonance. This is the mode we use as the quantum bus. The bus has a quality factor of $50\,000$, corresponding to a damping rate $\kappa_\mathrm{b}/2\pi = 110\,\mathrm{kHz}$, or a

Hamiltonian parameter (MHz)		Module 1	Module 2
Frequency	$\omega_a/2\pi$	6514.3	6505.2
	$\omega_t/2\pi$	5838.5	5668.7
	$\omega_c/2\pi$	5081.6	5149.2
	$\omega_r/2\pi$	8970.7	9014.9
	$\omega_f/2\pi$	9077	9114
	$\omega_b/2\pi$	5643	
Cross-Kerr	$\chi_{at}/2\pi$	-1.138	-0.953
	$\chi_{ac}/2\pi$	-0.765	-1.077
	$\chi_{bc}/2\pi$	-4.3	-2.7
	$\chi_{rt}/2\pi$	-1	-1
Self-Kerr	$\chi_{aa}/2\pi$	-3.9×10^{-3}	-4.1×10^{-3}
	$\chi_{tt}/2\pi$	-213.5	-203.4
	$\chi_{cc}/2\pi$	†	-112
Decay parameter (µs)			
Energy decay time	T_1^a	300	450
	T_1^t	35	65
	T_1^c	10	20
	T_1^r	0.10	0.10
	T_1^f	0.004	0.005
	T_1^b	1.6	
Ramsey decay time	T_{2R}^a	100	140
	T_{2R}^t	15*	30
	T_{2R}^c	10	20
Hahn echo decay time	T_{2E}^t	35	80
	T_{2E}^c	20	40
Steady-state excitation			
Ancilla	$1 - P(g)$	0.10	0.12
Cavity	$\bar{n}$	0.01	0.01
Converter	$1 - P(g)$	0.1	0.1

Table 7.1: Quantum Bus Sample Parameters. Uncertainties of measured Hamiltonian parameters are $< 0.1\%$ except when indicated by fewer significant digits. Subscript f refers to Purcell filter. Decay parameters are observed to fluctuate around 10%; typical values are given.

† χ_{cc} of module 1 was not measured, but is expected to be similar to that of module 2.

∗ Ramsey decay of ancilla 1 was not a simple exponential, indicating a frequency instability. Reported number is decay scale at short times.

lifetime of $1.6\,\mu s$. While this bus is an order of magnitude longer-lived than the cable modes used in [75], the lifetime is moderate at best compared to other 3D resonators This suggests there is significant room to improve the bus quality in the future. As explained in Section 5.8.2, we believe the lifetime is limited in part by resistive losses at the joint with the modules, as the same resonator had been measured to have a quality factor three times larger in a different mounting configuration. Improvement in the joint quality or location could grant significant increases in κ_{b}. In addition, the choice of commercial coax is not fundamental, and we are presently exploring resonators which can serve the same purpose, but with lifetimes more comparable to state-of-the-art 3D cavities.

7.2.2 Parametric conversion

In addition to a reasonably long lifetime, we are able to couple rapidly to the bus. Figure 7.2 demonstrates parametric conversion from one cavity into the bus, with the coupling turned off at the other module. The clear and rapid chevron pattern shows strong coupling compared to the bus lifetime. From this, we extract a conversion rate of $g/2\pi = 550\,\mathrm{kHz}$. Furthermore, the decay rate found with the pumps on is $\kappa/2\pi = 110\,\mathrm{kHz}$, consistent with that measured without pumps in Subsection 5.8.2. Operating far in the under-damped regime ($g = 5\kappa$) allows efficiency SWAPs between modules.

7.2.3 Advantages

As mentioned above, this scheme provides several simplifying advantages for coupling the storage modes through the bus. Because all the interactions are actuated by parametric conversion, the precise frequencies of all of the modes are irrelevant. The frequencies of the conversion transmon and the bus mode simply need to be somewhere nearby one another to facilitate a large dispersive shift between the two,

a

b
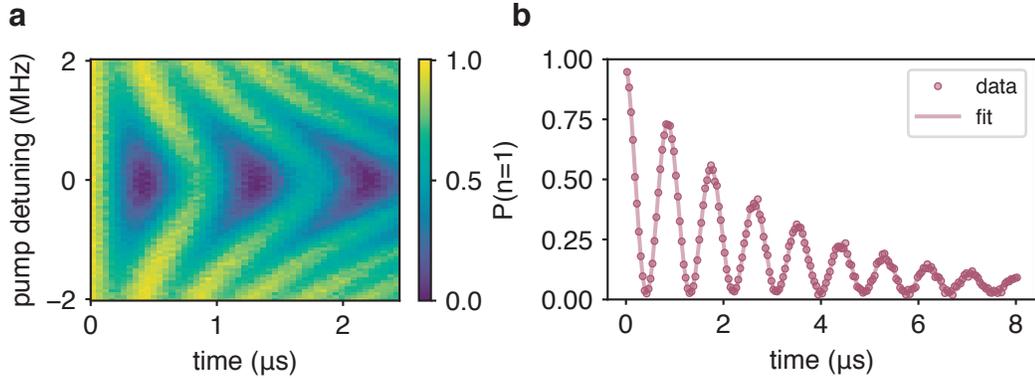
Figure 7.2: Cavity-Bus Swapping Dynamics. a) A single photon is prepared in cavity 1, then conversion to the bus is enable by turning on both pumps. Measured occupation of $|1\rangle$ in cavity 1 shown versus time and detuning of one pump from resonance. b) Line cut at zero detuning shows ring-down of population from decay in the bus. Value of conversion strength and decay rate in text is extracted from fit.

which enhances the strength of the parametric conversion. To our knowledge, this is the first scheme for coupling cQED modules which does not require any precise frequency matching or tunability. We chose the cavities to be nominally identical for convenience, but this is not a requirement. In fact, the near-degeneracy of these two modes, detuned by only 8 MHz, causes a small amount of crosstalk between control pulses. We find when we displace one cavity with a wide-band pulse, the other is displaced by 2–3% in amplitude. This may cause small problems in control and tomography, which can be easily mitigated by a small international detuning between the two by machining the lengths of the posts to be slightly different.

Another advantage to this approach compared to the Cirac protocol [17] (see Section 4.4) is that the interaction time is comparatively rapid. This speedup comes from the use of unitary swapping dynamics, rather than relying on the exponential decay of a communication resonator. For example, the swap time from the storage cavity to the bus mode is about 500 ns, compared to the release time of 6 µs in Chapter 6, while the maximum parametric coupling strengths is only slightly larger in this experiment. This more rapid interaction thus benefits from reduced decoherence from thermal or pump-induced excitation of the transmons.

Finally, we point out the general simplicity of the protocol. As described in Section 4.5, the state transfer and entanglement operations we will employ here do not require time-dependent control, or careful pre-calibration of Stark shifts and conversion strengths as in Chapter 6. As a result, the characterization needed to implement the experiment was rather easier. Another nice feature is that there is no need for direct measurement of the conversion transmons or the bus resonator itself. These are measured indirectly through the storage mode as detailed in Chapter 5. Indeed, no direct manipulation of the conversion transmon was necessary, save for initial characterization experiments like the ones shown in Section 5.8.2.

7.2.4 System initialization

The coherence and thermal occupations of the modes in these samples was a bit better than in Chapter 6, but the transmons still have a 5–10% chance to be found out of their ground state in equilibrium. To ensure the modules begins in a known state, we use an active feedback cooling sequence that makes use of the ability of our control hardware to perform simultaneous and independent branching when resetting the ancilla transmons.

The set of nested subroutines used at the beginning of every experimental sequence is shown in Figure 7.3. The cooling sequence relies on the measurement techniques introduced in chapter 5. The sequence begins by ensuring both ancillae are in their ground states, actively resetting as necessary. Then a π pulse, selective on $N = 0$ photons in the cavity, is applied to each ancilla, followed by measurement. Measurement of the ancilla in $|e\rangle$ heralds an empty cavity. If both cavities are empty, we continue (see next paragraph). If not, we reset the ancillae to $|g\rangle$, then actively empty the cavities by performing swaps with the bus (as in Figure 7.2), one at a time, with a $10\,\mu s$ delay after the swap to allow the state to decay in the bus. This is at least two orders of magnitude faster than waiting for the long-lived cavities to decay on their

own. We then start the sequence over, beginning with ancilla reset, and repeating as
necessary to ensure both cavities are in the vacuum state.

Once the cavities are confirmed empty, we use them to measure the state of the
conversion transmons. This uses the Ramsey-style selective displacement described
in Section 5.7, which displaces the cavity if the transmon is not in its ground state.
We then repeat the cavity measurement as described again. If the cavities are again
found in vacuum, we know the converters were in $|g\rangle$. We then reset the ancillae one
last time, and begin the experimental sequence. If either cavity is not in the vacuum
state, this means its converter was not in $|g\rangle$. We then empty both cavities and
begin the entire sequence from the beginning. Since the time to empty the cavities
is relatively long, we do not actively reset the converters, but simply allow them to
decay during this time.

After successful completion of this cooling routine, we find the ancilla transmons
and cavities with less than 1% probability each of being out of their respective ground
states. The conversion transmons are difficult to measure to this degree of accuracy,
since their measurement sequence is fairly long and involved. Given the length of
this sequence, it is likely that they have re-thermalized to about the 1% level (each)
by the time the cooling is successfully completed. A discussion of the effect of this
thermal population can be found in Subsection 7.3.1.

7.3 A tunable beamsplitter

There are a wide variety of ways to use the quantum bus for communication and
entanglement. The original experiments from Yale and NIST demonstrated two dif-
ferent approaches. The Yale experiment [90] operated with a qubit coupled to each
end of the bus at a rate g. The qubits were resonant with one another, but detuned
by an amount $\Delta \gg g$ from the bus. This is the virtual Raman regime discussed in
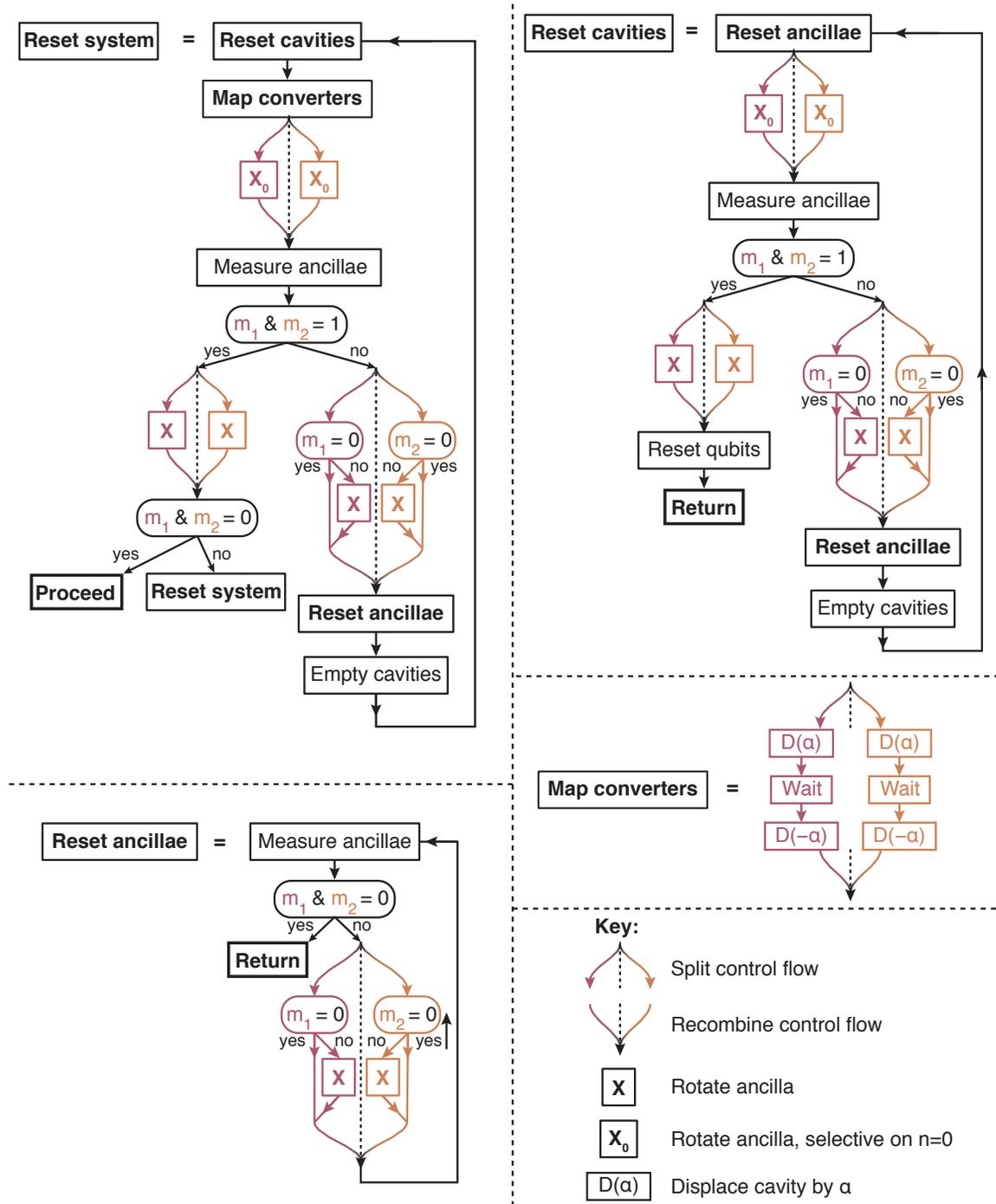
Figure 7.3: Quantum Bus System Initialization. Logical control flow for system reset and verification. Each experimental sequence begins with a call to "Reset System," which follow the control flow until the conditions have been met to reach "Proceed." Individual subroutines loop until reaching "Return," at which point they return to the previous sequence.

Subsection 4.7.1, where the bus is adiabatically eliminated, and the effective dynamics are described by a direct exchange between the endpoint qubits. This is especially useful when the bus is quite lossy, as was the case in that experiment. However, the effective exchange rate is suppressed by g/Δ, which is problematic when other sources of decoherence come into play. Alternately, the NIST experiment [91] involved sequential pairwise swaps between the qubits and bus — the first qubit is brought into resonance with the bus for a time, then moved away as the second qubit is moved into resonance. This sequential coupling can be used for full or partial swaps between qubits, but it is quite sensitive to the loss of the bus, because the qubit state must fully pass through the bus.

The approach we use here is a compromise between fast operation and reduced sensitivity to loss, while retaining full control over the degree of interaction. This is the tunable beamsplitter described in Section 4.5, where the storage cavities couple to the bus at equal rates g, with a common detuning Δ, as indicated in Figure 7.4a. In this manner we can generate an arbitrary beamsplitter by tuning the detuning and operation time. We review the basic idea here.

An important consideration when utilizing this coupling is to ensure that no information is left in the bus at the end of the protocol. For a fixed conversion rate and detuning, the bus is periodically emptied at time $\tau_{\text{BS}}(\Delta) = 2\pi/\sqrt{8g^2 + \Delta^2}$. At this time, the effective evolution involves only the two storage cavities. The resulting transformation (reproduced from Equation 4.57) is

$$
\begin{aligned}
a_1(\tau_{\text{BS}}) &= e^{-i\theta}\left(A_1 \cos\theta + iA_2 \sin\theta\right) \\
a_2(\tau_{\text{BS}}) &= e^{-i\theta}\left(A_2 \cos\theta + iA_1 \sin\theta\right),
\end{aligned}
\tag{7.1}
$$

which is a beamsplitter with mixing angle

$$
\theta = \frac{\pi}{2}\left(1 - \frac{\Delta}{\sqrt{8g^2 + \Delta^2}}\right) = \frac{\pi}{2}\left(1 - \frac{\Delta\,\tau_{\text{BS}}}{2\pi}\right).
\tag{7.2}
$$
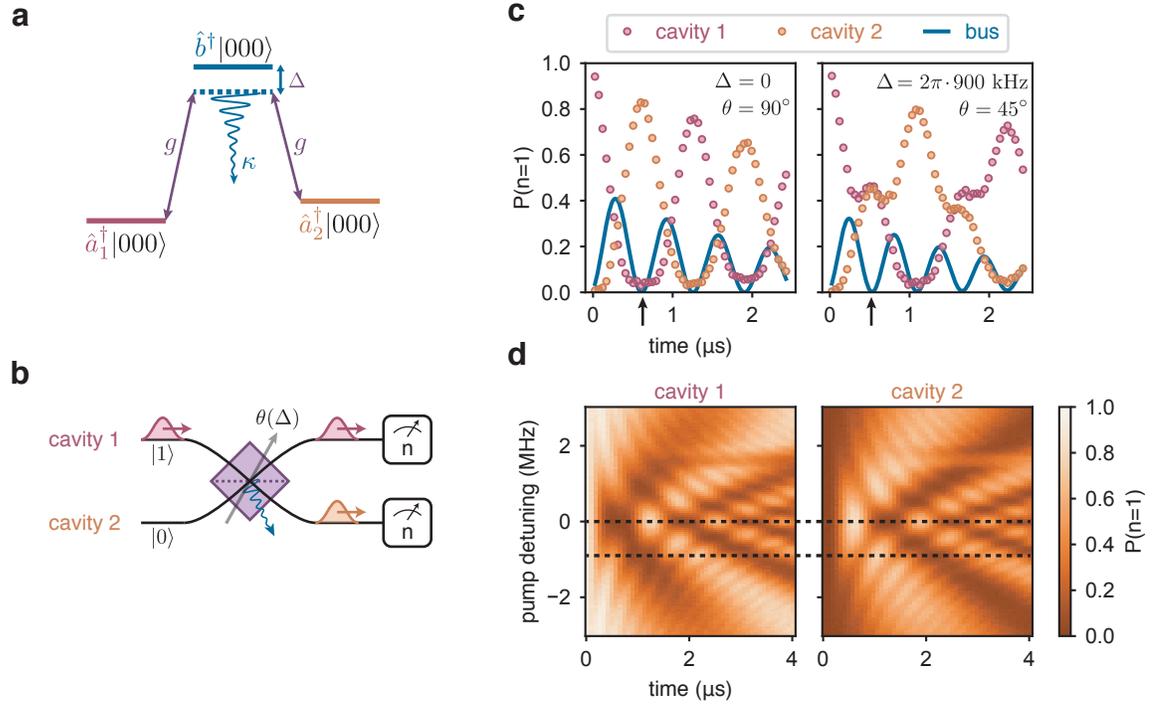
Figure 7.4: A Tunable Beamsplitter. a) Energy-level diagram denoting conversion (g) between cavity modes, mediated by the bus. Conversion can be detuned by Δ. Photons decay in the bus at rate κ_b. b) The conversion process acts as a tunable beamsplitter, with angle controlled by detuning. The beamsplitter mixes initial cavity states (left), producing final states (right), which can then be measured in the photon number basis. c) Measured cavity occupations as a function of conversion time, for a single photon prepared in cavity 1. Solid blue line is population of bus, inferred from fits to the measured cavity occupation. Operations are realized when the bus is empty, indicated by arrows on time axis. SWAP (left) inverts the initial cavity occupations, while 50:50 beamsplitter (right) results in equal occupation. d) Full dynamics as a function of detuning and time. Dashed lines indicate line cuts plotted in c). Faint splitting visible around detuning 1.5 MHz is an unknown pump-induced transition.

We indicate this beamsplitter transformation schematically with a physical beamsplitter as in Figure 7.4b, where the initial (final) cavity states are represented in inputs to (outputs from) the beamsplitter.

The beamsplitter angle θ is fully tunable from opaque to transparent, but in this chapter we use two particular working points, both shown in Figure 7.4c. The first is $\Delta = 0$, which realized a fully transparent beamsplitter ($\theta = \pi/2$), and swaps the state of the two cavities (up to local phases). This point is used for state transfer, with $\tau_{SWAP} = 624$ ns. From this data we extract the efficiency, $\eta = 0.84(1)$, which is

significantly higher than in Chapter 6 and related works.

The other useful operation is the 50:50 beamsplitter ($\theta = \pi/4$), obtained at $\Delta = \sqrt{8/3}\,g = 2\pi \times 900\,\mathrm{kHz}$, which we will use for entanglement generation. The 50:50 beamsplitter time is $\tau_{\mathrm{BS}} = 520\,\mathrm{ns}$. The full dynamics of the system for a single photon input state are seen in Figure 7.4d, demonstrating the continuous tunability of this operation.

7.3.1 Inefficiency contributions

Both SWAP and the 50:50 beamsplitter are several hundred times faster than the decay rate of either cavity ($\kappa_{\mathrm{a1},a2}^{-1} = 300, 450$ µs), so the excitation decay during the beamsplitter is dominated by dissipation in the bus ($\kappa_{\mathrm{b}}^{-1} = 1.5$ µs). As explained in Subsection 4.6.2, the inefficiency contribution due to loss is approximately $\frac{\pi \kappa_{\mathrm{b}}}{\sqrt{32}g} \approx 0.55\kappa_{\mathrm{b}}/g = 11\%$; thus, most of the measured inefficiency results from dissipation in the bus. Accordingly, these operations can be improved in future with increased conversion rate or bus quality.

The bulk of the remaining inefficiency is due to excitation of either conversion transmon. Both transmons have significant thermal excitation rates (see Table 7.1), and additional excitation events are caused by pump-induced transitions [178]. These events are responsible for 3–4% inefficiency. The value of g used here is a trade-off between speed, which reduces loss errors, and excitation errors induced by the pumps; an element which makes this trade-off more favorable can in future improve the beamsplitter (see Subsection 7.8.1 for discussion). The remaining 1–2% is due to single-photon preparation and measurement errors.

7.3.2 Tune-up procedure

The working points for SWAP and the 50:50 beamsplitter require both cavities to convert into the bus at the same detuning. Due to the Stark shifts on the cavities

and the bus by the pumps, this requires a little bit of fine-tuning. The procedure we used to find the working point shown in Figure 7.4 is as follows:

1. Find the bus mode frequency spectroscopically via the conversion transmon (as described in Section 5.8.2).

2. Find a working point for parametric conversion for each cavity, one at a time. This is done by setting the detuning between the two pumps to be approximately the detuning between a_i and b, and sweeping one pump frequency over a few MHz to produce the plot in Figure 7.2. This can be fit to extract the center frequency and coupling rate.

3. Still working one module at a time, adjust the pump amplitude (and frequency) in one of the modules to match the conversion rate found in the other module to within about 10%.

4. Now estimate the bus Stark shift induced by each pump in module 2. To do this, we measured the shift of the conversion chevron center frequency from module 1 with one of the pumps in module 2 turned on, then the other. The total added Stark shift is roughly the sum of those two shifts; this is how much one of the pumps on module 1 must be adjusted.

5. Repeat step 5 in reverse, measuring the bus Stark shift induced by the pumps in module 1 from module 2. Adjust one of the pump frequencies in module 2 accordingly.

6. In principle this would result in matched conversion rates and detunings in both modules when all pumps are on, and would determine the time for a SWAP, but some fine-tuning is still required. This is because the Stark shifts are large and do not necessarily add exactly (see for example [178]). To fine-tune, prepare a single photon in module 1, then turn on all four pumps, and

measure the occupation of each cavity as a function of four parameters: time, frequency and amplitude of one pump in module 1, and frequency of one pump in module 2. This provides enough knobs to match detunings and conversion rates. The optimum point in the four-dimensional sweep is the point at which the population of cavity 1 is minimized and that of cavity 2 is maximized. This point defines $\Delta = 0$ in Figure 7.4c,d.

Fortunately, by getting close to the right working point via the iterative steps 4 and 5, the large sweep in step 6 need not be over a very large range. In practice we swept over a time-span of about 40 ns, 100 to 150 kHz in detuning, at 10% in pump amplitude. The analysis of this 4D sweep is also iterative. Each time trace of population, corresponding to a particular set of detunings and amplitude, was fit to a 4th order polynomial, and the optimum time and occupation for those settings was extracted from the fit. Then, for each pump amplitude, the optimum occupation was fit to another 4th order polynomial, and so on, reducing the dimensionality by one each step. The final step was a fit over the last detuning, which extracted the optimum pump settings. We then looked back at the first fit to find the optimum time for those settings.

7.4   Moving a qubit with SWAP

Having established a working point for the SWAP operation, we now turn to the task of using it for transfer of a logical qubit from one module to the other. The operation is bi-directional, and so it supports simultaneous transfer in both directions — but for simplicity of characterization we focus on preparing a set of quantum states in module 1 and transferring them to module 2, as indicated in Figure 7.5a.

7.4.1 Encoding a qubit

As a first demonstration, we encode logical qubit states in the storage cavity of module 1 in the Fock code $|\pm z\rangle = \{|0\rangle\,|1\rangle\}$. To do this, we create an optimal control pulse (OCP) which maps the state of the ancilla transmon onto the cavity in this manifold; that is, an arbitrary superposition of transmon states is mapped onto the same superposition of cavity states:

$$(c_+\,|g\rangle + c_-\,|e\rangle) \otimes |0\rangle \rightarrow |g\rangle \otimes (c_+\,|0\rangle + c_-\,|1\rangle) \tag{7.3}$$

This same kind of encoding operation is employed in [65, 79], albeit in a different logical space.

The OCP encoding process is supposed to end with the transmon in its ground state to be re-used for measurement of the cavity later. Certain errors in the encoding (coherent or incoherent) can result in the transmon ending in $|e\rangle$ instead, which denotes a failure of the preparation of the logical state. To detect these cases, which will spoil all proceeding operations, we measure the transmon state immediately after encoding. If it is found in $|g\rangle$, we carry on with the experiment. If instead the transmon is in $|e\rangle$, we know that the encoding has failed, and we reset the entire system and try again. While this results in a probabilistic protocol, we regard this as part of the state initialization, not the communication. The success probability is about 98%, and the average fidelity of encoding is found to be 99% for the Fock code, as measured from Wigner tomography of the prepared states. Two of these states are shown in Figure 7.5b. The small rotation on the logical $|+x\rangle$ state is due to the backaction of the transmon measurement on the cavity; it is deterministic and so just applies an overall cavity phase, which can be compensated for in software.
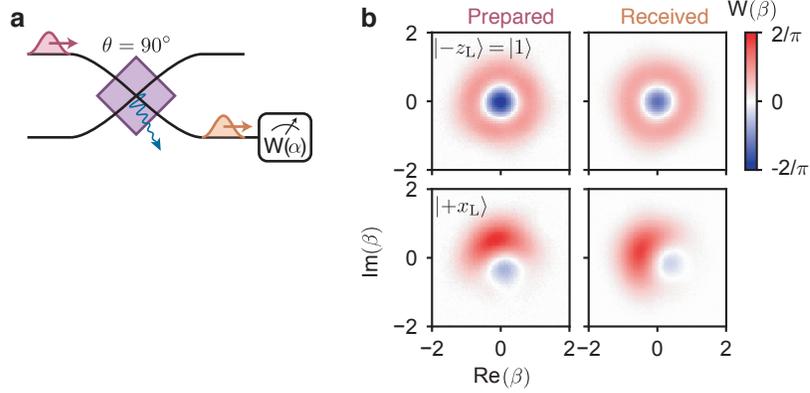
Figure 7.5: SWAP to Transfer a Qubit. a) A qubit is prepared in cavity 1, with the fully transparent beamsplitter applied to swap cavity states. The Wigner function of cavity 2 is measured after. b) Measured Wigner functions of two logical states in the Fock code, as prepared in cavity 1 (left) and received in cavity 2 (right).

7.4.2 Transfer fidelity

After encoding, we apply the SWAP pumps, and then perform Wigner tomography on cavity 2 to characterize the state. Two of these Wigner tomograms are shown in Figure 7.5b. As in Chapter 6, we do this for the six cardinal points on the Bloch sphere, reconstructing each resulting density matrix. The reconstructed mean state fidelity is $\mathcal{F}_{\text{Fock}} = 0.92(1)$. This is an improvement in infidelity of about a factor of two compared to the implementation in Chapter 6, due to the reduced loss of the link and the much more rapid protocol. The full dataset can be found in Figure 7.11.

As discussed in Subsection 3.2.2, the expected infidelity from the 11% loss is $.11/3 \approx 0.04$. There is an additional 1% infidelity from encoding. Finally, excitation of either conversion transmon primarily prevents the SWAP, and so result in $|0\rangle$ in cavity 2. In this encoding, the mean fidelity in this case is 0.5, so the 4% excitation probability contributes 0.02 infidelity. This simple error budget adds up to 0.07, consistent with the measured infidelity of 0.08(1).

7.5 Reaching break-even with error tracking

The attained state transfer fidelity in the Fock encoding is very encouraging, but as discussed, a large chunk of the infidelity comes from loss in the bus. A close look at Figure 3.5 suggests that we may be within the efficiency regime where we expect to be able to use error-detection to suppress the infidelity due to loss to second order. The slightly modified protocol is shown in Figure 7.5a. Our choice of error-correctable encoding is the cat code, as defined in Subsection 3.4.1. We encode the qubit in module 1, transfer, and then measure parity in cavity 2, and measure the Wigner function. It would instead be possible to apply one of two parity-dependent feedback correction pulses to restore to the original code space, or to conditionally decode the qubit onto the transmon state, as in [65]. However, since both the even and odd cat code are both viable logical encodings, it is sufficient to just record the error and proceed, possibly with fed-forward updates to future operations. For this reason, we simply sort the resulting tomographic data shot-by-shot, conditioned on the measured parity outcome, and present the data as error-tracked.

7.5.1 Choice of encoding

We chose the cat code to demonstrate the working of error correction in this encoding, as defined in Equation 3.23:

$$
\begin{aligned}
|+z_{\mathrm{L}}\rangle &= \frac{4}{\mathcal{N}_0(\alpha)} \sum_{n=2,6,10\ldots} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \equiv \left|\mathcal{C}_\alpha^{2\bmod 4}\right\rangle \\
|-z_{\mathrm{L}}\rangle &= \frac{4}{\mathcal{N}_1(\alpha)} \sum_{n=0,4,8\ldots} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \equiv \left|\mathcal{C}_\alpha^{0\bmod 4}\right\rangle,
\end{aligned}
\tag{7.4}
$$

Figure 3.6 tells us that for $\eta = 0.85$ we should use $\alpha = 1.3$, and so we did. More careful consideration might suggest using a slightly larger cat size, since the loss-limited efficiency is closer to 0.9, but the dependence of the fidelity on α is rather
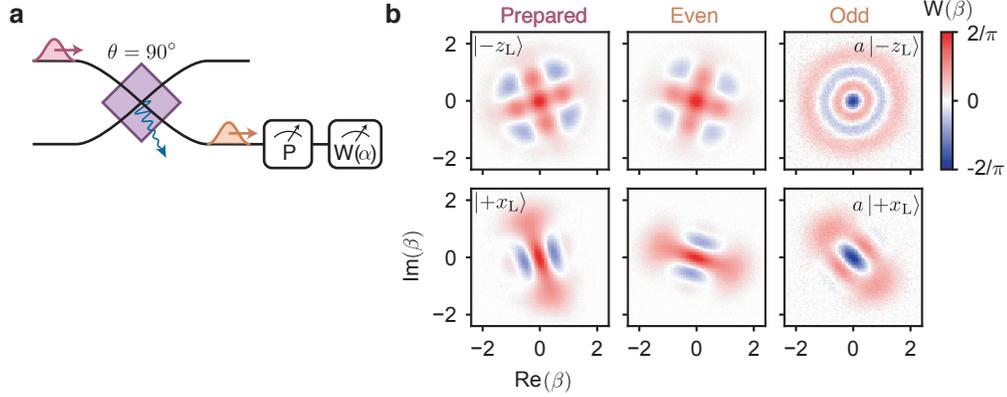
Figure 7.6: Error-tracked State Transfer. a) A qubit is encoded and swapped as before, but a parity measurement is performed before tomography. b) Two logical states in the cat code as prepared (left), and received, sorted by even (center) and odd (right) parity. Additional rotation for odd states are a deterministic result of the dispersive interaction from the ancilla being measured in $|e\rangle$.

weak around the optimum, so the exact value is not crucial. Reading off this plot, we expect a best-case infidelity from double errors and shrinkage of 0.01–0.02%, and a final cat size of $\alpha' = \sqrt{\eta}\,\alpha \approx 1.2$.

The qubit is encoded in the cavity in the same way as before, but with a new OCP computed for this encoding. Measured Wigner functions of two of the logical states as prepared are shown in Figure 7.6b. The average preparation fidelity is 0.98, slightly lower than in the Fock encoding due to the more complex nature of the state. Note again the same deterministic phase shift, a result of the verification measurement which checks for failures in state preparation.

7.5.2 Transfer fidelity

Due to the overhead of using a code with $\bar{n} = \alpha^2 = 1.7$ photons, the measured fidelity *without* the syndrome measurement is $\mathcal{F}_{\text{cat}} = 0.80(1) < \mathcal{F}_{\text{Fock}}$ (see Figure 7.11). However, the parity measurement can detect loss events, mitigating the impact on fidelity and resulting in a net gain.

When including the syndrome measurement before tomography, we overcome this

overhead to reach the break even point with respect to the Fock encoding. The measured Wigner functions sorted by parity outcome are shown in Figure 7.6b. The dominant outcome is that the parity remains even ($p_{\text{even}} = 84\%$), and the fidelity in this no-error case is $\mathcal{F}_{\text{even}} = 0.93(1)$. We detect a single loss error when the parity changes ($p_{\text{odd}} = 16\%$), and the coherence is preserved in the odd manifold, with fidelity $\mathcal{F}_{\text{odd}} = 0.86(3)$. By averaging over all results, the deterministic error-tracked fidelity is $\mathcal{F}_{\text{cat,tracked}} = 0.92(2)$. Even in the presence of code overhead, as well as errors in encoding and syndrome measurement, the cat code reaches the break-even point at which the transfer fidelity is as high as in the Fock encoding. The full dataset can be found in Figure 7.11.

As mentioned above, the expected infidelity due to second order errors not detected in this encoding are 0.01–0.02. The remaining errors in the cat encoding are largely the same as in the Fock code. State preparation accounts for 0.02, errors in assigning the parity an additional 0.01–0.02, and transmon excitation events 0.03, which gives a total expected error around 0.06–0.09, consistent with the measured 0.08.

In future, there are a few improvements which could improve the fidelity of the cat code in this application, enabling error-corrected state transfer beyond break-even. Use of a fault-tolerant parity measurement [64] can allow repeated measurement of the parity to improve confidence, without introducing more errors. Suppression of conversion transmon excitations will hopefully reduce this rather large, uncorrectable, effect. Finally, further reducing the loss of the bus resonator to improve the efficiency will push further past the break-even point, quickly suppressing the second-order errors.

7.6 Entanglement on demand

So far we have demonstrated quantum state transfer, which should be useful in a modular architecture which operates by shuttling qubits between modules. This approach can also be used to generate intermodule entanglement by preparing local Bell states and moving one qubit to another module. However, as in Chapter 6, there are a few readily-available, and more hardware-efficient, ways to generate entanglement in this system.

7.6.1 Entangling with a 50:50 beamsplitter

We demonstrated in Section 7.3 an operation which takes as input a single-photon and produces equal population in both cavities on the output. If this operation is in fact a 50:50 beamsplitter, the output state ought to be the odd Bell state $|O_+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. In order to confirm the creation of this state, we must measure joint correlations in three bases. We do this in a bit of a more visually nice way than in Chapter 6.

7.6.2 Entanglement verification

To clearly illustrate the correlations between the two cavities, we perform Wigner tomography on cavity 1, post-selected on a logical measurement in cavity 2 in the x, y, and z bases, as indicated in Figure 7.7a. Since the Wigner function is a complete description of the state, these conditional Wigner tomograms provides enough information to reconstruct the full two-qubit state. To execute the logical measurements, we use an OCP to decode cavity 2 onto its ancilla. This decoding is the inverse of the encoding in Equation 7.3 — it leaves the cavity in vacuum and maps the $\{|0\rangle, |1\rangle\}$ subspace onto the ancilla $\{|g\rangle, |e\rangle\}$ states. We then measure the ancilla to effect a z basis measurement, or rotate the ancilla into the appropriate basis with a $\pi/2$ pulse
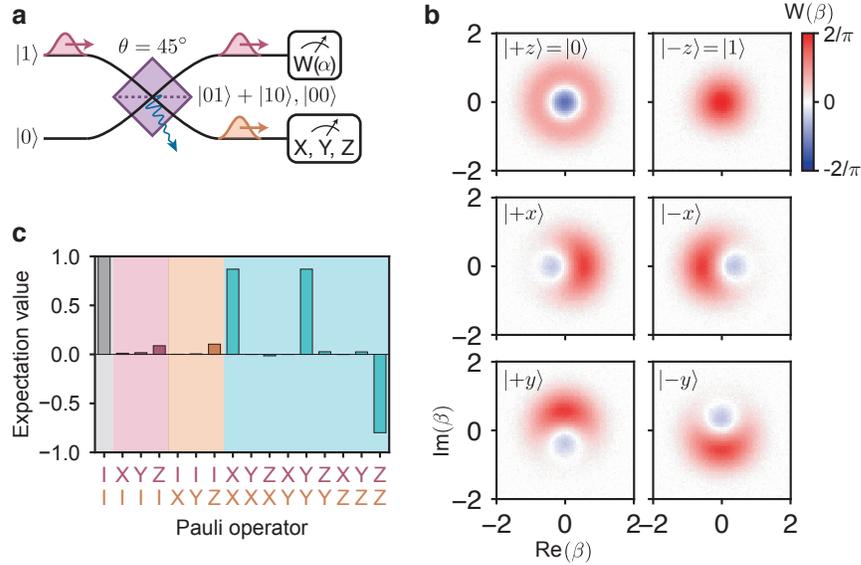
Figure 7.7: Entanglement by 50:50 Beamsplitter. a) A photon is prepared in cavity 1, and the 50:50 beamsplitter creates a Bell state. Correlations are assessed by measuring the Wigner function of cavity 1, conditioned on a measurement in the x, y or z basis in cavity 2. Ideal state and primarily error state appear at the outputs of beamsplitter. b) Wigner tomograms of cavity 1, conditioned on the measurement result in cavity 2 (inset label). States are anticorrelated in z basis and correlated in x and y. c) Reconstructed joint Pauli observables of the two-qubit entangled state.

around the Y (X) axis to measure in the x (y) basis.

The conditional Wigner functions are shown in Figure 7.7b. The Wigner function of cavity 1 shows strong correlations in all bases with the measurement outcome in cavity 2. We reconstruct each Wigner function individually to produce the density matrix of cavity 1, conditioned on the measurement outcome in cavity 2. As before, the measured Wigner function is normalized before reconstruction, to correct for measurement contrast in the parity mapping and ancilla readout. The result is two conditional density matrices for each of the three bases.

We use these conditional density matrices to reconstruct the logical two-qubit density matrix. This reconstruction uses a routine developed by Kevin Chou and described in Section 5.5 of his thesis [14], adapted for our tomography scheme. The reconstruction routine is the same as that used in Chapter 6, but the underlying data is different, and far fewer corrections are made for measurement contrast. This is

made possible by the improved reliability of the ancilla transmons in this experiment, and the independent readout chains which allow better measurement contrast. Additionally, the use of the $|f\rangle$ level of the ancilla for enhanced and roughly symmetric measurement contrast obviates the need for additional measurements to symmetrize the resultant data (see Subsection 5.4.4).

Each joint choice of bases for the two cavities is given by $\{k, l\} \in \{x, y, z\}^{\otimes 2}$, where k corresponds to the basis choice for cavity 1 and l for cavity 2. We refer to the measured probabilities of the logical measurements in cavity 2 as $p_{\pm l}$, and the conditional density matrices of cavity 1 as $\rho_{\pm l}$. The goal is to produce the expectation values $p_{\pm k, \pm l}$ of the four projectors $\Pi_{\pm k, \pm l} = |\pm k\rangle_1 |\pm l\rangle_2 \langle \pm k|_1 \langle \pm l|_2$, which correspond to the probability of measuring the joint state to be in $|\pm k\rangle_1 |\pm l\rangle_2$. This joint probability is $p_{\pm k, \pm l} = p_{\pm l} P(\pm k | \pm l)$, where $P(\pm k | \pm l)$ is the conditional probability of measuring $\pm k$ in cavity 1 given the result $\pm l$ in cavity 2. This conditional probability is the expectation value of the single-cavity projector $\Pi_{\pm k} = |\pm k\rangle_1 \langle \pm k|_1$, given the result $\pm l$ in cavity 2. There are $3 \times 3 = 9$ joint bases, and $2 \times 2 = 4$ joint probabilities per basis, for a total of 36 conditional probabilities.

To compute these conditional probabilities, we take the conditional density matrix $\rho_{\pm l}$ and evaluate $P(\pm k | \pm l) = \langle \Pi_{\pm k} \rangle_{\pm l} \equiv \mathrm{Tr}\,(\rho_{\pm l} \Pi_{\pm k})$, which is the squared overlap of the measured cavity state with the logical state $|\pm k\rangle_1$ given outcome $\pm l$. This is essentially the probability we would measure cavity 1 to be in $|\pm k\rangle_1$ with an ideal projective measurement. It is important to note here that, since the cavity density matrix is of dimension larger than 2, leakage out of the logical subspace (here, $\{|0\rangle, |1\rangle\}$) results in $P(+k | \pm l) + P(-k | \pm l) < 1$. We will see in a moment that this results in a reconstructed logical two-qubit state with trace slightly less than 1. Since we assign a binary outcome to the logical measurement of cavity 2, $p_{+l} + p_{-l} = 1$ by construction. This means that leakage out of the logical space on cavity 2 is not directly observed. However, such leakage will contribute to infidelity. Since the de-

code operation cannot account for this leakage, the result is some arbitrary outcome of the ancilla measurement, which is we assume to be uncorrelated with the result in cavity 1. Thus, while the decode-and-measure sequence will mask this leakage, it should convert it to infidelity in the form of a statistical mixture. Put another way, this local operation cannot increase the amount of entanglement, so it does not result in overestimation of the fidelity.

The computed joint probabilities $p_{\pm k, \pm l}$ are fed into the MLE reconstruction routine [14], which is a convex optimization over the space of all physical two-qubit (2^2 dimensional) density matrices. To ensure physicality, the resultant density matrix ρ_{L} is constrained to be Hermitian and positive semi-definite. In addition, $\mathrm{Tr}\,(\rho_{\mathrm{L}}) \leq 1$ to account for the possibility of leakage out of the logical space as discussed above. The reconstructed state is essentially the Bell state between cavity 1 and *ancilla* 2, since we do not correct for errors in the decoding procedure, which introduces about 1% of infidelity. In addition, the measurement errors on ancilla 2 contribute a small infidelity ($\sim 0.5\%$), also not corrected for.

The extracted fidelity to the ideal Bell state is $\mathcal{F}_{\mathrm{Bell,01}} = 0.88(1)$. The resulting joint Pauli expectation values, shown in Figure 7.7c, show clearly dominant two-qubit expectations. The largest non-ideal bars are IZ and ZI, which correspond to equal polarization of both qubits towards $|0\rangle$, due to loss in the beamsplitter. The trace of ρ_{L} (the value of the II bar) is found to be 0.999, consistent with very small ($< 10^{-3}$) occupation of Fock states above $n = 1$ for the reconstructed cavity density matrices.

7.7 Improved entanglement with error-detection

The Bell state generated using the 50:50 beamsplitter is a significant improvement over previous efforts. However, the fidelity is clearly limited by photon loss, consistent with a 9–10% mixture of $|00\rangle$. There is no way to detect the loss in this subspace

using local measurements, since no local observable can distinguish this dominant error state from $|01\rangle$ and $|10\rangle$ without collapsing the entangled state. However, by using multiple photons and an entangled state with well-defined *local* parity, we can detect photon loss errors.

7.7.1 Hong-Ou-Mandel interference

So far, all of the communication protocols demonstrated in this chapter can be executed with a directional link. However, we have not yet taken advantage of the bi-directionality of our quantum bus. In particular, a 50:50 beamsplitter has the interesting property that, when the input states are Fock states with the same photon number, the output ports, which will generically be entangled, both have *even* photon number. This is a property of the indistinguishability of photons, and has been used to measure the purity [190] and overlap [106] of quantum states.

The minimal and most canonical example of this property is Hong-Ou-Mandel (HOM) interference [19], where the input state $|11\rangle$ results in the output $(|02\rangle +$ $|20\rangle)/\sqrt{2}$. This corresponds to an odd logical Bell state in the encoding $|\pm x\rangle =$ $\{|0\rangle, |2\rangle\}$. While we noted in Chapter 3 that this is not a full error-correcting encoding, it has some nice properties, and has been used as the logical basis for the control qubit in a two-qubit logical CNOT gate [191]. For instance, we noted in Section 3.3 that this encoding allows error *detection*.

7.7.2 Error detection

Since the HOM state has more photons than the single-photon Bell state in Section 7.6, it is more sensitive to loss in the bus. However, the dominant error states resulting from loss are $|01\rangle$ and $|10\rangle$, which have exactly one photon, and can be distinguished from the desired entangled state by measuring *local* parity in each of the cavities before tomography, as in Figure 7.8a. We then perform tomography on
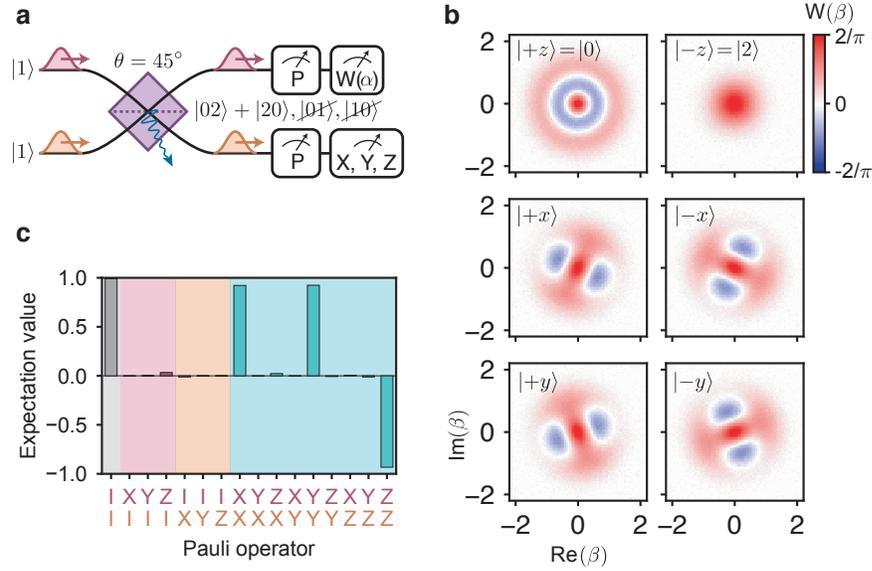
Figure 7.8: Error-detected Hong-Ou-Mandel Entanglement. a) A photon is pre-
pared in each cavity before application of 50:50 beamsplitter, resulting in two-photon Bell
state. Prior to tomography, the parity of each cavity is measured, and the data is post-
selected on even parity in both. This rejects dominant error states, shown in beamsplitter
output. b) Wigner tomograms of cavity 1, post-selected on success and conditioned on log-
ical measurement in cavity 2 (inset label indicates measurement outcome). Anticorrelation
observed in z, correlation in x and y. Rotations in x and y bases result from beamsplit-
ter phases and parity measurement, which are not corrected for in this measurement. c)
Reconstructed joint Pauli observables of the two-qubit entangled state.

the cavities as in Section 7.6, but with a decoding OCP for the 0-2 manifold. By

conditioning on even parity in *both* cavities ($p_s = 79\%$), we reject the dominant error

cases, resulting in improved performance at the cost of a small failure rate. The post-

selected conditional Wigner tomograms are show in Figure 7.8b. These data show

clear correlations as before, but in the two-photon manifold. There is a rotation of

the x and y states, which is a combination of phases acquired from the beamsplitter

pumps and the backaction of the parity measurement, but these are deterministic and

can be corrected for or calibrated out[1].

As in Subsection 7.6.2, we can reconstruct the logical two-qubit state ρ_L, but now

using the logical space $\{|0\rangle, |2\rangle\}$ to define the projection operators $\Pi_{\pm k}$ for cavity

---

1. The beamsplitter phase is calibrated out in Figure 7.7, which is why no rotation is observed
there.

1. Upon reconstruction and correction of single-qubit Z rotation (deterministic cavity phase rotation), we find fidelity to the ideal state $\mathcal{F}_{\mathrm{Bell,02}} = 0.94(1)$. This is a reduction of the *infidelity* by a factor of two compared to the single-photon state. The joint Pauli bars, shown in Figure 7.8c, show clear suppression of the IZ and ZI observables, which indicates successful suppression of photon loss. The value of the II bar (and hence the trace of ρ_{L}) is 0.991, consistent with a typical 1% occupation of the $|1\rangle$ state in the measured Wigner functions of cavity 1, due to imperfect syndrome measurement and cavity decay during syndrome measurement and tomography. In fact, the occupation of the error state $|1\rangle$ is found to be largest for states with large occupation of $|2\rangle$, suggesting errors during Wigner tomography are primarily responsible. Finally, we reiterate that this tomographic method only corrects for contrast in the Wigner function, but no other sources of error. The estimated error of the decode-and-measure sequence is 2%, which suggests the fidelity of the underlying cavity-cavity entanglement is in excess of 0.95.

7.7.3 Increasing the success rate

Generally speaking, sacrificing success rate for fidelity is acceptable when generating entanglement as a resource for later operations, as long as it can be generated rapidly. In particular, we highlight the high success probability of this approach, which is an advantage when compared to many other heralding approaches. This high success probability is due thanks to the low-loss beamsplitter and accurate parity measurements.

That said, it is useful to be able to increase the success probability when necessary, particularly if multiple Bell pairs need to be generated for entanglement distillation, for instance. This scheme readily admits an increase in the success rate, at a small cost in the fidelity. The state of the cavities upon failure (loss of a photon) is known — either $|10\rangle$ or $|01\rangle$, depending on which cavity was found to have odd parity. This
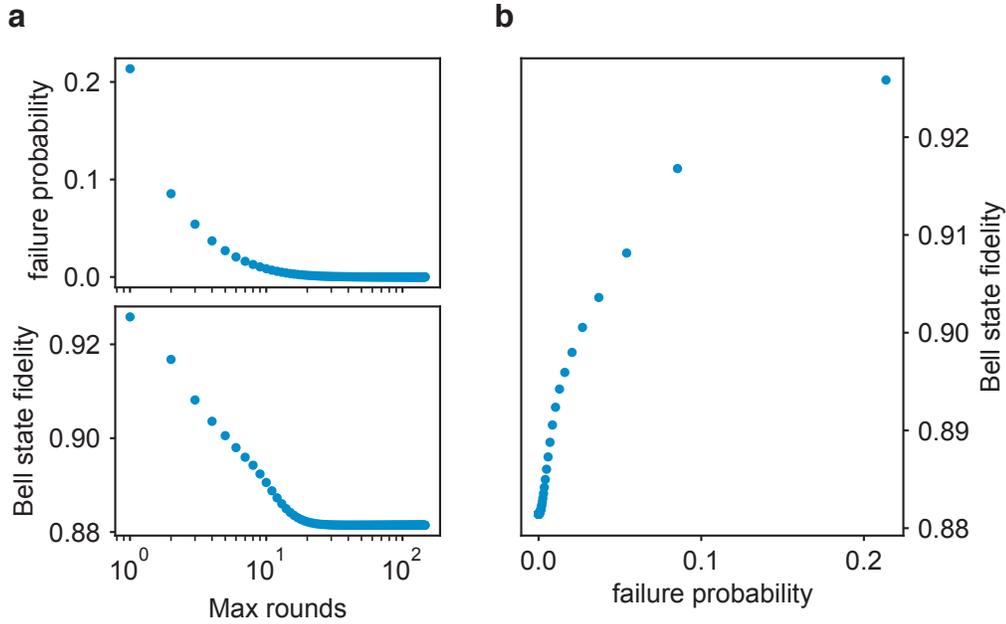
Figure 7.9: Repeated Hong-Ou-Mandel entanglement. a) Cumulative failure probability and fidelity of repeated Hong-Ou-Mandel interference and parity measurement, as a function of the number of rounds accepted. b) Fidelity versus failure probability.

involved resetting one transmon to its ground state while applying OCP to the other. We can rapidly reload a photon into the empty cavity and apply the beamsplitter again, and repeat this until we succeed, finding both cavities to have even parity. Because the success probability is high, the chance of succeeding on successive rounds rapidly increases.

We take all of the data from this "try and re-try" experiment, and plot the success probability and fidelity versus the maximum number of rounds kept. These are plotted in Figure 7.9. Indeed we see a rapid increase in the probability of success versus number of rounds. Within ten attempts, the success probability is above 99%.

By repeating until success, we can make this scheme effectively deterministic. However, the cumulative fidelity decreases monotonically with the number of rounds. This effect is primarily due to undetectable dephasing-like errors from conversion transmon excitation or ancilla decay during parity mapping or readout. The deterministic fidelity is $\mathcal{F}_{\text{Bell},02,\text{det}} = 0.88(1)$, which turns out to be equal to that obtained

in the deterministic single-photon Bell state in Section 7.6. However, this approach allows an extra degree of freedom which trades fidelity for success rate. In the first few rounds, this trade-off might be useful in a distillation context, since the success probability rises faster than the fidelity drops. Note that the fidelity including the first round is slightly lower than in the single-round experiment presented in Figure 7.8. This data was taken at a different time, and small drifts in state preparation or measurement fidelity may account for this roughly 1% change, which is within the reconstruction uncertainty.

7.8 Conclusions

This chapter represents a coming together of a few of the ideas throughout this thesis. We have presented a method of joining independent modules with a low-loss, rapid communication link, which grants several simplifications from previous work like Chapter 6; in particular, this approach does not require specific time-dependent controls or frequency-matching of samples. Here we give a few thoughts on challenges in this architecture, and possible opportunities moving forward.

7.8.1 Improving the link

Though we have presented several approaches for suppressing infidelity resulting from loss in the bus, all of the experiments demonstrated here would benefit from improved efficiency to begin with. For state transfer, the error-correctable approach benefits from reduced loss by rapidly suppressing second-order errors, as well as expanding the available choice of encodings. The uncorrectable Fock encoding also of course benefits from improvements. If the inefficiency of this link could approach the single-pass loss — which is more than three order of magnitude smaller than the loss incurred here — then the need for bosonic error correction in this situation might be obviated to

begin with. Likewise, both entanglement schemes would benefit from lower loss in fidelity and/or success probability.

Since the inefficiencies of these protocols depends on the ratio $\kappa_{\rm b}/g$, we have two avenues along which we can improve. Reducing the damping rate $\kappa_{\rm b}$ of the bus is an obvious target. We already know these coaxial cable resonators can have quality factors three times higher at this frequency, so learning how to improve the joint quality could make the bus reproducibly better. Materials improvements to the bus itself, perhaps by replacing or removing the PTFE dielectric, using high-purity aluminum conductors, or even increasing the diameter of the coax, could reduce contributions to loss from dielectric and surface resistance. There is little reason to doubt a connection could be made between modules with quality comparable to the 3D cavity resonators we use as quantum memories. Further discussion on this topic can be found in Chapter 8.

In parallel, improvements in the parametric conversion scheme would greatly benefit the protocols demonstrated here. The single-junction transmon is a simple converter, but it is far from optimal for this setting. More exotic elements, made with arrays of junctions, or flux loops, may give rise to cleaner and simpler mixing processes, with different trade-offs in terms of the Kerr and dissipation the induce on the cavities and bus [177, 178]. Some of these elements are challenging to implement in a 3D architecture, but we believe this is more a question of careful and clever engineering than a fundamental problem. An alternate approach would be to make the converter itself very low-Q, so that induced excitations decay rapidly, mitigating the dephasing effect. If this could be done without inducing too much damping on the cavity and bus (e.g. by Purcell filtering), it might improve the performance, especially by reducing uncorrectable errors.

Longer links

One obvious difference between the approach taken here and schemes utilizing propagating photons is the length of the link. When communicating via a continuum, the length of the transmission line is irrelevant, save for a small amount of added loss. Here, however, the mode structure of the bus changes as the length is increased. As discussed in Chapter 4 and demonstrated in [128], the presence of nearby modes is not really a problem, as it simply leads to an onset of the continuum regime. However, the presence of extra modes may crowd the frequency spectrum, making it more difficult to find bands in which to place the parametric pumps, conversion element, and storage cavity without accidental collisions. Further, these modes may contribute additional Purcell decay to the conversion transmon, and ultimately to the storage mode. Finally, with fixed capacitance, as the length of the link increases, the zero-point phase φ_{b} of the bus across the converter junction decrease with the square root of the length, weakening the strength of any parametric process. This may require increasing the physical coupling capacitance, which will worsen Purcell effects.

One approach which may simultaneously solve a few of these problems is to place a filter mode between the mixing element and the bus. Bus modes near this filter will couple more strongly to the transmon, while harmonics further away will couple more weakly, reducing Purcell effects and suppressing the effects of frequency collisions. This is similar to the on-chip modes used in [75], but in this context, the exact frequency of the filter modes would not be crucial; they would simply need to be nearby a particular mode of the link.

7.8.2   Inter-module gates

The quantum bus has uses beyond the communication and entanglement implemented here. The bus has long been used as a way to perform gates between qubits [93, 192, 97, 193]. The scheme used here is immediately extensible to operations between

transmons, such as RIP and cross-resonance gates, which could be used to realize the kind of architecture proposed in Subsection 2.5.4, with gates performed directly between modules without the need for teleportation or shuttling.

Using the bus for gates between bosonic qubits is slightly more challenging, but it is within reach. The beamsplitter itself is not a computational gate, since it takes the qubits outside their code spaces. However, this operation is a key component of linear optical quantum computing (LOQC) protocols [58, 105], and so this tool may enable new avenues for modular computing along these lines, with the added benefits of deterministic state preparation and QND measurement afforded by cQED.

A particularly interesting extension of traditional LOQC is the exponential-SWAP gate (eSWAP) [108], which has already been realized within a single 3D cQED module [107]. The eSWAP is a continuous family of encoding-independent gates which can be used to construct logical qubits out of two or four cavity modes, and to perform single- and two-qubit logical gates. This gate consists of two 50:50 beamsplitters, between which a local operation, similar in structure to the parity mapping used here, is applied to one cavity. As such, all of the requirements for this gate are already present in this system. Since the beamsplitter must be applied twice, a lower-loss bus will be extremely beneficial in demonstrating this type of control. This is an obvious next application of the quantum bus for interfacing bosonic qubits over a network, hopefully leading to new approaches for quantum networking in cQED.
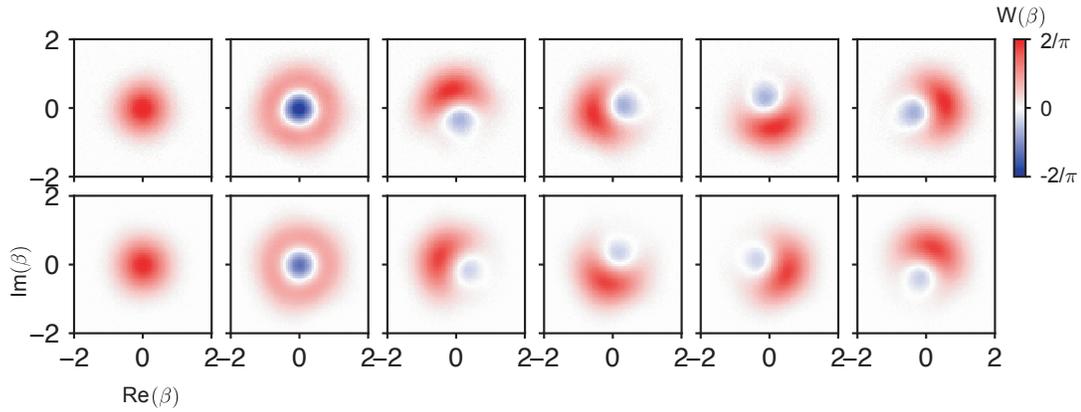
Figure 7.10: Fock Code Transfer Extended Data. Measured Wigner functions for all six cardinal states of Fock encoding, as prepared in module 1 (top) and received in module 2 (bottom).
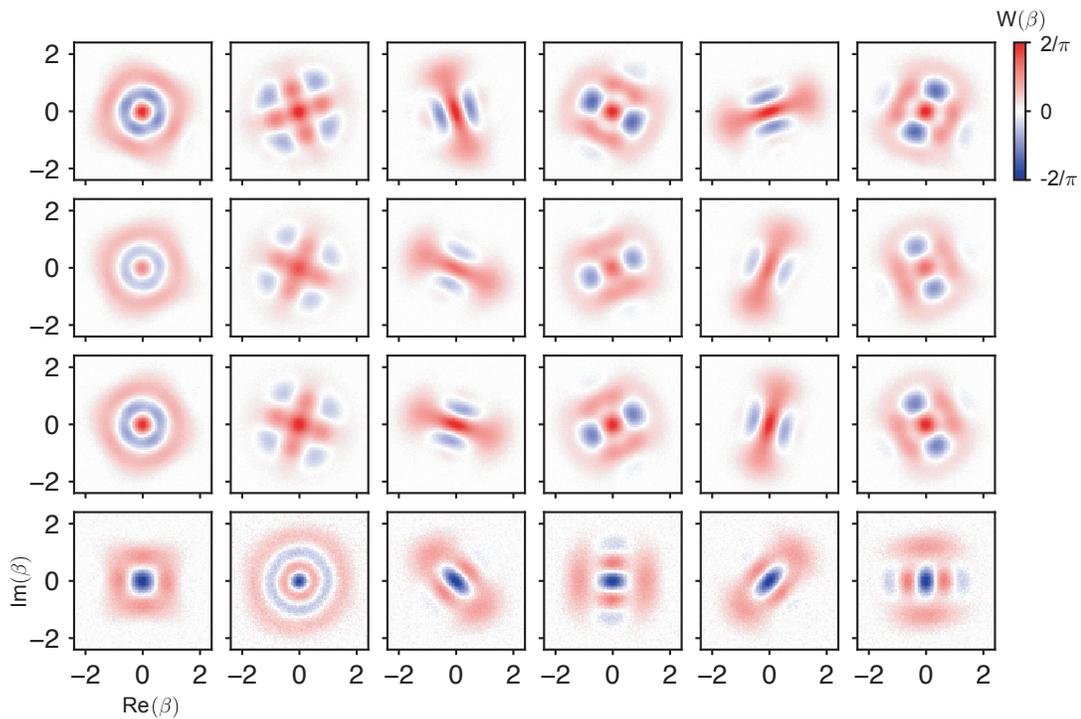
Figure 7.11: Fock Code Transfer Extended Data. Measured Wigner functions for all six cardinal states of cat encoding, as prepared in module 1 (first row), received in module 2 with no parity measurement, (second row), and received and conditioned on even (third row) and odd (fourth row) parity.

Chapter 8

Outlook

I'd like to conclude by taking some of the tools developed and lessons learned in this thesis, and suggesting some possible extensions and future directions.

Lower loss links

Obviously, all off the experiments presented here would have benefited from less loss in the communication links. Building better interfaces and links is an ongoing task. For the directional pitch and catch scheme, this requires a better circulator, and better connectors and couplers. The connectorization scheme used for the quantum bus resonator, which eschews commercial connectors, it quite low-loss, though some work is required to make it more robust and controlled. There is ongoing work on all-superconducting circulators based on interference in Josephson junction devices [187, 188, 189]. However, use of such an element would still require careful design and packaging to ensure the total transmission loss is not limited by connectorization and the required interference elements. Still, there is great promise in this, especially given that the large instantaneous bandwidth of commercial ferrite circulators is not a requirement for the comparatively narrow-band communication.

The path towards a lower-loss quantum bus is more straightforward, due to the

relative simplicity of the device. More careful treatment and placement of the joint between bus and module might improve the robustness of quality, but commercial coaxial cable is probably not the ideal link moving forward. 3D resonators which are sensitive to their seam resistance have been made with quality factors as high as 300 million [125], so there are many orders of magnitude over which we can improve the current iteration. To get there, we probably need to build our own links from materials of known quality, like high-purity aluminum [62] or indium platings or coatings [124, 125]. As mentioned in Subsection 4.3.5, aluminum waveguide might be a nice alternative which requires no dielectric, possesses a mode structure which can be tailored, and can be coupled to through apertures. As the materials quality is increased, the joint will become increasingly critical, so these advances must be made together.

If the communication process can be executed with sufficiently high efficiency, we may not even require dedicated error-correction, would would simplify certain aspects of the scheme. As a rule of thumb, if the communication errors are as low as the local gate errors, *and* they are of the same type, then they can probably be handled by whatever error-correction scheme is being used throughout the rest of the quantum machine. Loss can probably be made to fill these criteria, since amplitude damping is already an error which must be tacked in most superconducting qubit modalities.

Connecting more modules

While the last few years have seen an explosion of work on cQED networks, so far no demonstration exists for more than two modules. An obvious question to ask is how one might wire together networks with $N > 2$ modules. Either of the methods demonstrated here could be extended by adding more "communication arms" to each module and working with pairwise links. However, it would be more efficient to use a single channel to connect multiple modules.
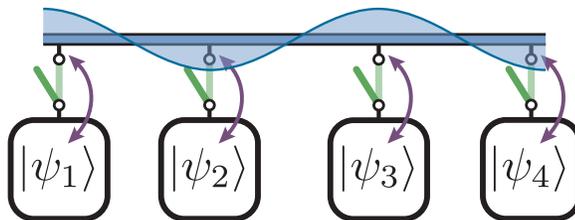
Figure 8.1:   Extending the Quantum Bus. Multiple modules connected by a single bus.

It is not clear to me how the pitch and catch protocol with a directional channel generalizes to more modules. Simply tiling more circulators to make a long unidirectional channel could work for small N, but may quickly become cumbersome with the addition of many circulators. A conceptually appealing approach is to devise a router like the ones described in Section 2.5. Switchable directionality in a low-loss router would enable many flexible approaches to scaling with propagating photons.

The quantum bus should provide a reasonable way to to interface more than two modules. In this work, we coupled to the ends of the bus for convenience. However, equal coupling strengths could be had by coupling to *each* voltage maximum, as indicated in Figure 8.1. It's not clear that this would be particularly easy to do with commercial coaxial cable, as it would require removing sections of the outer conductor from the interior. However, with a custom link as discussed above, this could be designed in from the start. A nice feature of this approach is that the voltage *maxima* are current *minima*, so resistive joint losses could be mitigated by appropriate design. This design also provides any-to-any communication, or multi-body interactions coupling through a single bus mode. It might even allow simultaneous pairwise communication by working in the virtual Raman regime (Subsection 4.7.1) with different detunings.

Toward entanglement purification and teleportation

As we stated in Subsection 2.5.3, high-fidelity entanglement can be created from multiple Bell pairs of lesser quality. So far, this kind of entanglement purification has only been demonstrated between quantum memories in separate network nodes with nitrogen-vacancy centers [51]. Using a highly probabilistic remote entanglement scheme, those authors required many attempts to successfully generate the raw pairs. Fortunately, that platform has a long-lived memory element (nuclear spin) and a communication element (electron spin) which couples strongly to light, so the first pair can be efficiently stored while the second pair is generated.

The two deterministic entanglement schemes presented here lend themselves to a distillation experiment, since in both cases the entanglement procedure is fast compared to the coherence times of the cavity memories. The entangled pairs could be created sequentially using a module where the conversion transmon couples *two* cavities [167, 106] to the bus. Alternatively, only one "communication" cavity needs to couple to the bus if its state can then be swapped with that of a "memory" cavity [106]. That same architecture has been used to demonstrate a local entangling gate [191] which could be used to perform the required Bell measurement for distillation. In this sense, all the tools are already in place to take this next step in 3D cQED. Furthermore, a teleported gate has been performed in a single module [40], and these tools can be readily applied to the remote version.

Heterogeneous modular networks

The two approaches for communication presented here have their own advantages and challenges. While the quantum bus allows rapid and faithful entanglement generation, and perhaps direct intermodule gates, we detailed some of the challenges for scaling to longer bus connections in Subsection 7.8.1. While these problems have potential solutions, a hybrid approach might be of use here. One can imagine grouping a few

modules into small "clusters" connected by a single bus for dense local interactions. But then we could bridge the clusters with longer-distance propagating photon channels, through which one plays pitch and catch. This long-haul communication role could also be filled by probabilistic entanglement schemes [71, 159] which are robust to loss, with part of each module working repeatedly to generate entanglement while deterministic local operations are being carried out.

There are a lot of options for the connectivity graph of this kind of inhomogeneous system. Each module in each cluster could have connections to other clusters, in which case it might look like cities connected by highways, with multiple on- and off-ramps per city. Alternatively, each cluster could have a module dedicated for long-distance communication, in which case the structure would be tree-like. A graph can look tree-like at multiple layers, so this might be the most efficient way to scale up while keeping the number of operations required to get between any two modules logarithmic in the system size.

I am optimistic that that the tools and methods explored in this thesis will be of use in the near future for modular superconducting quantum processors.

Bibliography

[1] Gordon E Moore. Cramming more components onto integrated circuits. *Electronics*, 38(8):114–117, 1965.

[2] Intel. The story of the Intel 4004. `https://www.intel.co.uk/content/www/uk/en/history/museum-story-of-intel-4004.html`.

[3] Institute of Electrical and Electronics Engineers. International roadmap for devices and systems. `https://irds.ieee.org/images/files/pdf/2016_MM.pdf`, 2016.

[4] Robert H. Dennard, Fritz H. Gaensslen, Hwa Nien Yu, V. Leo Rideout, Ernest Bassous, and Andre R. Leblanc. Design of Ion-Implanted MOSFET's With Very Small Physical Dimensions. *IEEE Journal of Solid-State Circuits*, 9(5):256–268, 1974.

[5] Mark Bohr. A 30 Year Retrospective on Dennard's MOSFET Scaling Paper. *IEEE Solid-State Circuits Newsletter*, 12(1):11–13, 2007.

[6] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982.

[7] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[8] Lov K. Grover. A fast quantum mechanical algorithm for database search. *Proceedings of the Annual ACM Symposium on Theory of Computing*, F1294:212–219, jul 1996, 9605043.

[9] Ashley Montanaro. Quantum algorithms: An overview. *npj Quantum Information*, 2:15023, 2016.

[10] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man Hong Yung, Xiao Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5:4213, jul 2014, arXiv:1304.3061 [quant-ph].

[11] Alexander Randall. Q&A: A lost interview with ENIAC co-inventor J. Presper Eckert. https://www.computerworld.com/article/2561813/q-a--a-lost-interview-with-eniac-co-inventor-j--presper-eckert.html, 2006.

[12] Eduardo Pinheiro, WD Weber, and LA Barroso. Failure trends in a large disk drive population. https://www.usenix.org/legacy/events/fast07/tech/pinheiro.html, 2007.

[13] Jacob Blumoff. *Multiqubit Experiments in 3D Circuit Quantum Electrodynamics.* PhD thesis, New Haven CT, 2017.

[14] Kevin Chou. *Teleported Operations between Logical Qubits in Circuit Quantum Electrodynamics.* PhD thesis, New Haven CT, 2018.

[15] Christopher J Axline. *Building Blocks for Modular Circuit QED Quantum Computing.* PhD thesis, New Haven CT, 2018.

[16] Christopher J. Axline, Luke D. Burkhart, Wolfgang Pfaff, Mengzhen Zhang, Kevin Chou, Philippe Campagne-Ibarcq, Philip Reinhold, Luigi Frunzio, S. M. Girvin, Liang Jiang, M. H. Devoret, and R. J. Schoelkopf. On-demand quantum state transfer and entanglement between remote microwave cavity memories. *Nature Physics*, 14(7):705–710, jul 2018.

[17] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi. Quantum State Transfer and Entanglement Distribution among Distant Nodes in a Quantum Network. *Physical Review Letters*, 78(16):3221–3224, apr 1997, 9611017.

[18] Luke D. Burkhart, James Teoh, Yaxing Zhang, Christopher J. Axline, Luigi Frunzio, M. H. Devoret, Liang Jiang, S. M. Girvin, and R. J. Schoelkopf. Error-detected state transfer and entanglement in a superconducting quantum network (in preparation). apr 2020, 2004.06168.

[19] C. K. Hong, Z. Y. Ou, and L. Mandel. Measurement of subpicosecond time intervals between two photons by interference. *Physical Review Letters*, 59(18):2044–2046, 1987.

[20] Wayne M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland. Quantum Zero effect. *Physical Review A*, 41(5):2295–2300, 1990.

[21] M. Brune, E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J. M. Raimond, and S. Haroche. Observing the progressive decoherence of the meter in a quantum measurement. *Physical Review Letters*, 77(24):4887–4890, 1996.

[22] M. Hatridge, S. Shankar, M. Mirrahimi, F. Schackert, K. Geerlings, T. Brecht, K. M. Sliwa, B. Abdo, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret. Quantum back-action of an individual variable-strength measurement. *Science*, 339(6116):178–181, jan 2013, arXiv:1903.11732 [quant-ph].

[23] D. H. Slichter, C. Müller, R. Vijay, S. J. Weber, A. Blais, and I. Siddiqi. Quantum Zeno effect in the strong measurement regime of circuit quantum electrodynamics. *New Journal of Physics*, 18(5), may 2016, arXiv:1512.04006 [quant-ph].

[24] Shay Hacohen-Gourgy, Leigh S. Martin, Emmanuel Flurin, Vinay V. Ramasesh, K. Birgitta Whaley, and Irfan Siddiqi. Quantum dynamics of simultaneously measured non-commuting observables. *Nature*, 538(7626):491–494, 2016, arXiv:1608.06652 [quant-ph].

[25] Z. K. Minev, S. O. Mundhada, S. Shankar, P. Reinhold, R. Gutiérrez-Jáuregui, R. J. Schoelkopf, M. Mirrahimi, H. J. Carmichael, and M. H. Devoret. To catch and reverse a quantum jump mid-flight. *Nature*, 570(7760):200–204, jun 2019, arXiv:1803.00545 [quant-ph].

[26] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, 2nd edition, 2010.

[27] U. Vool, S. Shankar, S. O. Mundhada, N. Ofek, A. Narla, K. Sliwa, E. Zalys-Geller, Y. Liu, L. Frunzio, R. J. Schoelkopf, S. M. Girvin, and M. H. Devoret. Continuous Quantum Nondemolition Measurement of the Transverse Component of a Qubit. *Physical Review Letters*, 117(13), sep 2016, arXiv:1605.08004 [quant-ph].

[28] Alexander Grimm, Nicholas E. Frattini, Shruti Puri, Shantanu O. Mundhada, Steven Touzard, Mazyar Mirrahimi, Steven M. Girvin, Shyam Shankar, and Michel H. Devoret. The Kerr-Cat Qubit: Stabilization, Readout, and Gates. jul 2019, arXiv:1907.12131 [quant-ph].

[29] Klaus Mølmer and Anders Sørensen. Multiparticle entanglement of hot trapped ions. *Physical Review Letters*, 82(9):1835–1838, 1999, 9810040.

[30] Norbert Schuch and Jens Siewert. Natural two-qubit gate for quantum computation using the XY interaction. *Physical Review A - Atomic, Molecular, and Optical Physics*, 67(3):8, 2003.

[31] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Physical Review A*, 55(2):900–911, 1997, 9604034.

[32] Richard Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics*, 41(12):2315–2323, 1994.

[33] H J Kimble. The quantum internet. *Nature*, 453(7198):1023–1030, jun 2008, arXiv:0806.4195 [quant-ph].

[34] Liang Jiang, Jacob M. Taylor, Anders S. Sørensen, and Mikhail D. Lukin. Distributed quantum computation based on small quantum registers. *Physical Review A*, 76(6):062323, dec 2007.

[35] T. E. Northup and R. Blatt. Quantum information transfer using photons. *Nature Photonics*, 8(5):356–363, apr 2014.

[36] C. Monroe, R. Raussendorf, A. Ruthven, K. R. Brown, P. Maunz, L. M. Duan, and J. Kim. Large-Scale Modular Quantum-Computer Architecture with Atomic Memory and Photonic Interconnects. *Physical Review A*, 89(2):1–16, feb 2014, arXiv:1208.0391 [quant-ph].

[37] M. A. Rowe, A. Ben-Kish, B. DeMarco, D. Leibfried, V. Meyer, J. Beall, J. Britton, J. Hughes, W. M. Itano, B. Jelenkovic, C. Langer, T. Rosenband, and D. J. Wineland. Transport of Quantum States and Separation of Ions in a Dual RF Ion Trap. may 2002, arXiv:0205094 [quant-ph].

[38] D. Kielpinski, C. Monroe, and D. J. Wineland. Architecture for a large-scale ion-trap quantum computer. *Nature*, 417(6890):709–711, jun 2002.

[39] Daniel Gottesman and Isaac L. Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402(6760):390–393, nov 1999.

[40] Kevin S. Chou, Jacob Z. Blumoff, Christopher S. Wang, Philip C. Reinhold, Christopher J. Axline, Yvonne Y. Gao, L. Frunzio, M. H. Devoret, Liang Jiang, and R. J. Schoelkopf. Deterministic teleportation of a quantum gate between two logical qubits. *Nature*, 561(7723):368–373, sep 2018.

[41] Yong Wan, Daniel Kienzler, Stephen D. Erickson, Karl H. Mayer, Ting Rei Tan, Jenny J. Wu, Hilma M. Vasconcelos, Scott Glancy, Emanuel Knill, David J. Wineland, Andrew C. Wilson, and Dietrich Leibfried. Quantum gate teleportation between separated qubits in a trapped-ion processor. *Science*, 364(6443):875–878, may 2019.

[42] C. Cabrillo, J. I. Cirac, P. García-Fernández, and P. Zoller. Creation of entangled states of distant atoms by interference. *Physical Review A - Atomic, Molecular, and Optical Physics*, 59(2):1025–1033, 1999, 9810013.

[43] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller. Long-distance quantum communication with atomic ensembles and linear optics. *Nature*, 414(6862):413–418, nov 2001.

[44] Sean D. Barrett and Pieter Kok. Efficient high-fidelity quantum computation using matter qubits and linear optics. *Physical Review A - Atomic, Molecular, and Optical Physics*, 71(6), jun 2005, 0408040.

[45] N. Roch, M. E. Schwartz, F. Motzoi, C. Macklin, R. Vijay, A. W. Eddins, A. N. Korotkov, K. B. Whaley, M. Sarovar, and I. Siddiqi. Observation of Measurement-Induced Entanglement and Quantum Trajectories of Remote Superconducting Qubits. *Physical Review Letters*, 112(17):170501, apr 2014.

[46] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Physical Review A - Atomic, Molecular, and Optical Physics*, 53(4):2046–2052, 1996, 9511030.

[47] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Physical Review Letters*, 76(5):722–725, 1996, 9511027.

[48] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A - Atomic, Molecular, and Optical Physics*, 54(5):3824–3851, 1996, 9604024.

[49] Naomi H. Nickerson, Ying Li, and Simon C. Benjamin. Topological quantum computing with a very noisy network and local error rates approaching one percent. *Nature Communications*, 4:1756, apr 2013.

[50] Naomi H. Nickerson, Joseph F. Fitzsimons, and Simon C. Benjamin. Freely scalable quantum technologies using cells of 5-to-50 qubits with very lossy and noisy photonic links. *Physical Review X*, 4(4), 2014, arXiv:1406.0880 [quant-ph].

[51] N. Kalb, A. A. Reiserer, P. C. Humphreys, J. J.W. Bakermans, S. J. Kamerling, N. H. Nickerson, S. C. Benjamin, D. J. Twitchen, M. Markham, and R. Hanson. Entanglement distillation between solid-state quantum network nodes. *Science*, 356(6341):928–932, jun 2017, arXiv:1703.03244 [quant-ph].

[52] Uri Vool and Michel Devoret. Introduction to quantum electromagnetic circuits. *International Journal of Circuit Theory and Applications*, 45(7):897–934, 2017.

[53] Marios H. Michael, Matti Silveri, R. T. Brierley, Victor V. Albert, Juha Salmilehto, Liang Jiang, and S. M. Girvin. New class of quantum error-correcting codes for a bosonic mode. *Physical Review X*, 6(3), 2016, arXiv:1602.00008 [quant-ph].

[54] Isaac L. Chuang, Debbie W. Leung, and Yoshihisa Yamamoto. Bosonic quantum codes for amplitude damping. *Physical Review A*, 56(2):1114–1125, 1997.

[55] Victor V. Albert, Kyungjoo Noh, Kasper Duivenvoorden, Dylan J. Young, R. T. Brierley, Philip Reinhold, Christophe Vuillot, Linshu Li, Chao Shen, S. M. Girvin, Barbara M. Terhal, and Liang Jiang. Performance and structure of single-mode bosonic codes. *Physical Review A*, 97(3), mar 2018, arXiv:1708.05010 [quant-ph].

[56] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47(10):777–780, 1935.

[57] G. J. Milburn. Quantum optical Fredkin gate. *Physical Review Letters*, 62(18):2124–2127, 1989.

[58] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, jan 2001.

[59] Daniel Gottesman, Alexei Kitaev, and John Preskill. Encoding a qubit in an oscillator. *Physical Review A. Atomic, Molecular, and Optical Physics*, 64(1):123101–1231021, 2001, 0008040.

[60] C. Flühmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. P. Home. Encoding a qubit in a trapped-ion mechanical oscillator. *Nature*, 566(7745):513–517, feb 2019.

[61] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret. A stabilized logical quantum bit encoded in grid states of a superconducting cavity. jul 2019, arXiv:1907.12487 [quant-ph].

[62] Matthew Reagor, Hanhee Paik, Gianluigi Catelani, Luyan Sun, Christopher Axline, Eric Holland, Ioan M. Pop, Nicholas A. Masluk, Teresa Brecht, Luigi Frunzio, Michel H. Devoret, Leonid Glazman, and Robert J. Schoelkopf. Reaching 10 ms single photon lifetimes for superconducting aluminum cavities. *Applied Physics Letters*, 102(19), may 2013, arXiv:1302.4408 [cond-mat.supr-con].

[63] Matthew Reagor, Wolfgang Pfaff, Christopher Axline, Reinier W. Heeres, Nissim Ofek, Katrina Sliwa, Eric Holland, Chen Wang, Jacob Blumoff, Kevin Chou, Michael J. Hatridge, Luigi Frunzio, Michel H. Devoret, Liang Jiang, and Robert J. Schoelkopf. Quantum memory with millisecond coherence in circuit QED. *Physical Review B*, 94(1), jul 2016, arXiv:1508.05882 [quant-ph].

[64] S. Rosenblum, P. Reinhold, M. Mirrahimi, Liang Jiang, L. Frunzio, and R. J. Schoelkopf. Fault-tolerant detection of a quantum error. *Science*, 361(6399):266–270, jul 2018, arXiv:1803.00102 [quant-ph].

[65] Nissim Ofek, Andrei Petrenko, Reinier Heeres, Philip Reinhold, Zaki Leghtas, Brian Vlastakis, Yehan Liu, Luigi Frunzio, S. M. Girvin, L. Jiang, Mazyar Mirrahimi, M. H. Devoret, and R. J. Schoelkopf. Extending the lifetime of a quantum bit with error correction in superconducting circuits. *Nature*, 536(7617):441–445, jul 2016.

[66] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. P. Song, C. L. Zou, S. M. Girvin, L. M. Duan, and L. Sun. Quantum error correction and universal gate set operation on a binomial bosonic logical qubit. *Nature Physics*, 15(5):503–508, may 2019, arXiv:1805.09072 [quant-ph].

[67] Philip Reinhold, Serge Rosenblum, Wen-Long Ma, Luigi Frunzio, Liang Jiang, and Robert J. Schoelkopf. Error-corrected gates on an encoded qubit. jul 2019, arXiv:1907.12327 [quant-ph].

[68] Marcel Bergmann and Peter Van Loock. Quantum error correction against photon loss using NOON states. *Physical Review A*, 94(1), jul 2016, arXiv:1512.07605 [quant-ph].

[69] Linshu Li, Chang Ling Zou, Victor V. Albert, Sreraman Muralidharan, S. M. Girvin, and Liang Jiang. Cat Codes with Optimal Decoherence Suppression for a Lossy Bosonic Channel. *Physical Review Letters*, 119(3), jul 2017, arXiv:1609.06386 [quant-ph].

[70] Sreraman Muralidharan, Chang Ling Zou, Linshu Li, Jianming Wen, and Liang Jiang. Overcoming erasure errors with multilevel systems. *New Journal of Physics*, 19(1), jan 2017, arXiv:1504.08054 [quant-ph].

[71] A. Narla, S. Shankar, M. Hatridge, Z. Leghtas, K.M. Sliwa, E. Zalys-Geller, S.O. Mundhada, W. Pfaff, L. Frunzio, R.J. Schoelkopf, and M.H. Devoret. Robust Concurrent Remote Entanglement Between Two Superconducting Qubits. *Physical Review X*, 6(3):031036, sep 2016.

[72] C. Dickel, J. J. Wesdorp, N. K. Langford, S. Peiter, R. Sagastizabal, A. Bruno, B. Criger, F. Motzoi, and L. DiCarlo. Chip-to-chip entanglement of transmon qubits using engineered measurement fields. *Physical Review B*, 97(6), feb 2018, arXiv:1712.06141 [quant-ph].

[73] P. Campagne-Ibarcq, E. Zalys-Geller, A. Narla, S. Shankar, P. Reinhold, L. Burkhart, C. Axline, W. Pfaff, L. Frunzio, R. J. J Schoelkopf, and M. H. H Devoret. Deterministic Remote Entanglement of Superconducting Circuits through Microwave Two-Photon Transitions. *Physical Review Letters*, 120(20):200501, may 2018.

[74] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J-c C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff. Deterministic quantum state transfer and remote entanglement using microwave photons. *Nature*, 558(7709):264–267, jun 2018.

[75] N. Leung, Y. Lu, S. Chakram, R. K. Naik, N. Earnest, R. Ma, K. Jacobs, A. N. Cleland, and D. I. Schuster. Deterministic bidirectional communication and remote entanglement generation between superconducting qubits. *npj Quantum Information*, 5:18, dec 2019.

[76] Ze Liang Xiang, Mengzhen Zhang, Liang Jiang, and Peter Rabl. Intracity quantum communication via thermal microwave networks. *Physical Review X*, 7(1), mar 2017, arXiv: 1611.10241.

[77] B. Vermersch, P. O. Guimond, H. Pichler, and P. Zoller. Quantum State Transfer via Noisy Photonic and Phononic Waveguides. *Physical Review Letters*, 118(13), mar 2017, arXiv: 1611.10240.

[78] Zaki Leghtas, Gerhard Kirchmair, Brian Vlastakis, Robert J. Schoelkopf, Michel H. Devoret, and Mazyar Mirrahimi. Hardware-Efficient Autonomous Quantum Memory Protection. *Physical Review Letters*, 111(12):120501, sep 2013.

[79] Reinier W. Heeres, Philip Reinhold, Nissim Ofek, Luigi Frunzio, Liang Jiang, Michel H. Devoret, and Robert J. Schoelkopf. Implementing a universal gate set on a logical qubit encoded in an oscillator. *Nature Communications*, 8(1), dec 2017, arXiv:1608.02430 [quant-ph].

[80] Linshu Li, Dylan J. Young, Victor V. Albert, Kyungjoo Noh, Chang-Ling Zou, and Liang Jiang. Designing good bosonic quantum codes via creating destructive interference. jan 2019, arXiv:1901.05358 [quant-ph].

[81] Victor V. Albert, Shantanu O. Mundhada, Alexander Grimm, Steven Touzard, Michel H. Devoret, and Liang Jiang. Pair-cat codes: Autonomous error-correction with low-order nonlinearity. *Quantum Science and Technology*, 4(3), 2019, arXiv:1801.05897 [quant-ph].

[82] Mazyar Mirrahimi, Zaki Leghtas, Victor V Albert, Steven Touzard, Robert J Schoelkopf, Liang Jiang, and Michel H Devoret. Dynamically protected cat-qubits: a new paradigm for universal quantum computation. *New J. Phys. New Journal of Physics New Journal of Physics*, 16(16), 2014.

[83] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, M. J. Hatridge, M. Reagor, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret. Confining the state of light to a quantum manifold by engineered two-photon loss. *Science*, 347(6224):853–857, feb 2015, arXiv:1412.4633 [quant-ph].

[84] S. Touzard, A. Grimm, Z. Leghtas, S. O. Mundhada, P. Reinhold, C. Axline, M. Reagor, K. Chou, J. Blumoff, K. M. Sliwa, S. Shankar, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret. Coherent Oscillations inside a Quantum Manifold Stabilized by Dissipation. *Physical Review X*, 8(2), apr 2018, arXiv:1705.02401 [quant-ph].

[85] Raphaël Lescanne, Marius Villiers, Théau Peronnin, Alain Sarlette, Matthieu Delbecq, Benjamin Huard, Takis Kontos, Mazyar Mirrahimi, and Zaki Leghtas. Exponential suppression of bit-flips in a qubit encoded in an oscillator. jul 2019, arXiv:1907.11729 [quant-ph].

[86] Shruti Puri, Samuel Boutin, and Alexandre Blais. Engineering the quantum states of light in a Kerr-nonlinear resonator by two-photon driving. *npj Quantum Information*, 3:18, dec 2017, arXiv:1605.09408 [quant-ph].

[87] Shruti Puri, Lucas St-Jean, Jonathan A. Gross, Alexander Grimm, N. E. Frattini, Pavithran S. Iyer, Anirudh Krishna, Steven Touzard, Liang Jiang, Alexandre Blais, Steven T. Flammia, and S. M. Girvin. Bias-preserving gates with stabilized cat qubits. may 2019, arXiv:1905.00450 [quant-ph].

[88] Jérémie Guillaud and Mazyar Mirrahimi. Repetition Cat Qubits for Fault-Tolerant Quantum Computation. *Physical Review X*, 9(4):041053, dec 2019.

[89] Stephan Ritter, Christian Nölleke, Carolin Hahn, Andreas Reiserer, Andreas Neuzner, Manuel Uphoff, Martin Mücke, Eden Figueroa, Joerg Bochmann, and Gerhard Rempe. An elementary quantum network of single atoms in optical cavities. *Nature*, 484(7393):195–200, apr 2012.

[90] J. Majer, J. M. Chow, J. M. Gambetta, Jens Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Coupling superconducting qubits via a cavity bus. *Nature*, 449(7161):443–447, sep 2007.

[91] Mika A. Sillanpää, Jae I. Park, and Raymond W. Simmonds. Coherent quantum state storage and transfer between two phase qubits via a resonant cavity. *Nature*, 449(7161):438–442, sep 2007.

[92] A. Palacios-Laloy, F. Nguyen, F. Mallet, Patrice Bertet, D. Vion, and D. Esteve. Tunable resonators for quantum circuits. *Journal of Low Temperature Physics*, 151(3-4 PART 2):1034–1042, may 2008, arXiv:0712.0221 [quant-ph].

[93] L. Dicarlo, J. M. Chow, J. M. Gambetta, Lev 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, jul 2009, arXiv:0903.2030 [cond-mat.mes-hall].

[94] R. C. Bialczak, M. Ansmann, M. Hofheinz, M. Lenander, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, T. Yamamoto, A. N. Cleland, and J. M. Martinis. Fast tunable coupler for superconducting qubits. *Physical Review Letters*, 106(6), feb 2011.

[95] Yi Yin, Yu Chen, Daniel Sank, P. J. J. O'Malley, T. C. White, R. Barends, J. Kelly, Erik Lucero, Matteo Mariantoni, A. Megrant, C. Neill, A. Vainsencher, J. Wenner, Alexander N. Korotkov, A. N. Cleland, and John M. Martinis. Catch and Release of Microwave Photon States. *Physical Review Letters*, 110(10):107001, mar 2013.

[96] M. Pechal, L. Huthmacher, C. Eichler, S. Zeytinolu, A.A. Abdumalikov, S. Berger, A. Wallraff, and S. Filipp. Microwave-Controlled Generation of Shaped Single Photons in Circuit Quantum Electrodynamics. *Physical Review X*, 4(4):041010, oct 2014.

[97] David C. McKay, Stefan Filipp, Antonio Mezzacapo, Easwar Magesan, Jerry M. Chow, and Jay M. Gambetta. Universal Gate for Fixed-Frequency Qubits via a Tunable Bus. *Physical Review Applied*, 6(6), dec 2016, arXiv:1604.03076 [quant-ph].

[98] Wolfgang Pfaff, Christopher J Axline, Luke D Burkhart, Uri Vool, Philip Rein-
hold, Luigi Frunzio, Liang Jiang, Michel H Devoret, and Robert J Schoelkopf.
Controlled release of multiphoton quantum states from a microwave cavity
memory. *Nature Physics*, 13(9):882–887, jun 2017.

[99] A. D. O'Connell, M. Hofheinz, M. Ansmann, Radoslaw C. Bialczak, M. Lenan-
der, Erik Lucero, M. Neeley, D. Sank, H. Wang, M. Weides, J. Wenner, John M.
Martinis, and A. N. Cleland. Quantum ground state and single-phonon control
of a mechanical resonator. *Nature*, 464(7289):697–703, apr 2010.

[100] J. D. Teufel, T. Donner, Dale Li, J. W. Harlow, M. S. Allman, K. Cicak,
A. J. Sirois, J. D. Whittaker, K. W. Lehnert, and R. W. Simmonds. Side-
band cooling of micromechanical motion to the quantum ground state. *Nature*,
475(7356):359–363, jul 2011, arXiv:1103.2144 [quant-ph].

[101] Y. Kubo, F. R. Ong, P. Bertet, D. Vion, V. Jacques, D. Zheng, A. Dréau, J. F.
Roch, A. Auffeves, F. Jelezko, J. Wrachtrup, M. F. Barthe, P. Bergonzo, and
D. Esteve. Strong coupling of a spin ensemble to a superconducting resonator.
Physical Review Letters, 105(14), sep 2010, arXiv: 1006.0251.

[102] J. J. Viennot, M. C. Dartiailh, A. Cottet, and T. Kontos. Coherent coupling
of a single spin to microwave cavity photons. *Science*, 349(6246):408–411, jul
2015.

[103] Yutaka Tabuchi, Seiichiro Ishino, Atsushi Noguchi, Toyofumi Ishikawa, Rekishu
Yamazaki, Koji Usami, and Yasunobu Nakamura. Coherent coupling between
a ferromagnetic magnon and a superconducting qubit. *Science*, 349(6246):405–
408, jul 2015, arXiv:1410.3781 [quant-ph].

[104] A. P. Reed, K. H. Mayer, J. D. Teufel, L. D. Burkhart, W. Pfaff, M. Reagor,
L. Sletten, X. Ma, R. J. Schoelkopf, E. Knill, and K. W. Lehnert. Faithful
conversion of propagating quantum information to mechanical motion. *Nature
Physics*, 13(12):1163–1167, sep 2017, arXiv:1703.02548 [quant-ph].

[105] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling,
and G. J. Milburn. Linear optical quantum computing with photonic qubits.
Reviews of Modern Physics, 79(1):135–174, 2007.

[106] Yvonne Y. Gao, Brian J. Lester, Yaxing Zhang, Chen Wang, Serge Rosen-
blum, Luigi Frunzio, Liang Jiang, S. M. Girvin, and Robert J. Schoelkopf. Pro-
grammable Interference between Two Microwave Quantum Memories. *Physical
Review X*, 8(2):021073, 2018, arXiv:1802.08510 [quant-ph].

[107] Yvonne Y. Gao, Brian J. Lester, Kevin S. Chou, Luigi Frunzio, Michel H.
Devoret, Liang Jiang, S. M. Girvin, and Robert J. Schoelkopf. Entangle-
ment of bosonic modes through an engineered exchange interaction. *Nature*,
566(7745):509–512, 2019.

[108] Hoi-Kwan Lau and Martin B. Plenio. Universal Quantum Computing with Arbitrary Continuous-Variable Encoding. *Physical Review Letters*, 117(10):100501, aug 2016.

[109] Ananda Roy and Michel Devoret. Introduction to parametric amplification of quantum signals with Josephson circuits. *Comptes Rendus Physique*, 17(7):740–755, aug 2016.

[110] Christopher S. Wang, Jacob C. Curtis, Brian J. Lester, Yaxing Zhang, Yvonne Y. Gao, Jessica Freeze, Victor S. Batista, Patrick H. Vaccaro, Isaac L. Chuang, Luigi Frunzio, Liang Jiang, S. M. Girvin, and Robert J. Schoelkopf. Efficient multiphoton sampling of molecular vibronic spectra on a superconducting bosonic processor. aug 2019, arXiv:1908.03598 [quant-ph].

[111] Daniel Adam Steck. Quantum and Atom Optics. http://atomoptics-nas.uoregon.edu/~dsteck/teaching/quantum-optics/quantum-optics-notes.pdf, 2007.

[112] Neereja M. Sundaresan, Yanbing Liu, Darius Sadri, László J. Szocs, Devin L. Underwood, Moein Malekakhlagh, Hakan E. Türeci, and Andrew A. Houck. Beyond strong coupling in a multimode cavity. *Physical Review X*, 5(2), 2015, arXiv:1502.02660 [quant-ph].

[113] P. Forn-Díaz, J. J. García-Ripoll, B. Peropadre, J. L. Orgiazzi, M. A. Yurtalan, R. Belyansky, C. M. Wilson, and A. Lupascu. Ultrastrong coupling of a single artificial atom to an electromagnetic continuum in the nonperturbative regime. *Nature Physics*, 13(1):39–43, jan 2017, arXiv:1602.00416 [quant-ph].

[114] Fumiki Yoshihara, Tomoko Fuse, Sahel Ashhab, Kosuke Kakuyanagi, Shiro Saito, and Kouichi Semba. Superconducting qubit-oscillator circuit beyond the ultrastrong-coupling regime. *Nature Physics*, 13(1):44–47, jan 2017, arXiv:1602.00415 [quant-ph].

[115] Sal J. Bosman, Mario F. Gely, Vibhor Singh, Alessandro Bruno, Daniel Bothner, and Gary A. Steele. Multi-mode ultra-strong coupling in circuit quantum electrodynamics. *npj Quantum Information*, 3(1), dec 2017, arXiv:1704.06208 [quant-ph].

[116] Roman Kuzmin, Nitish Mehta, Nicholas Grabon, Raymond Mencia, and Vladimir E. Manucharyan. Superstrong coupling in circuit quantum electrodynamics. *npj Quantum Information*, 5(1), dec 2019, arXiv:1809.10739 [cond-mat.supr-con].

[117] John M. Martinis and A. Megrant. UCSB final report for the CSQ program: Review of decoherence and materials physics for superconducting qubits. oct 2014, arXiv:1410.5793 [quant-ph].

[118] C. Wang, C. Axline, Y. Y. Gao, T. Brecht, Y. Chu, L. Frunzio, M. H. Devoret, and R. J. Schoelkopf. Surface participation and dielectric loss in superconducting qubits. *Applied Physics Letters*, 107(16), oct 2015, arXiv:1509.01854 [quant-ph].

[119] G. Calusine, A. Melville, W. Woods, R. Das, C. Stull, V. Bolkhovsky, D. Braje, D. Hover, D. K. Kim, X. Miloshi, D. Rosenberg, A. Sevi, J. L. Yoder, E. Dauler, and W. D. Oliver. Analysis and mitigation of interface losses in trenched superconducting coplanar waveguide resonators. *Applied Physics Letters*, 112(6), feb 2018, arXiv:1709.10015 [quant-ph].

[120] W. Woods, G. Calusine, A. Melville, A. Sevi, E. Golden, D. K. Kim, D. Rosenberg, J. L. Yoder, and W. D. Oliver. Determining Interface Dielectric Losses in Superconducting Coplanar-Waveguide Resonators. *Physical Review Applied*, 12(1), jul 2019, arXiv:1808.10347 [quant-ph].

[121] David A Pozar. *Microwave Engineering*. Wiley, 4 edition, 2011.

[122] C. Axline, M. Reagor, R. Heeres, P. Reinhold, C. Wang, K. Shain, W. Pfaff, Y. Chu, L. Frunzio, and R. J. Schoelkopf. An architecture for integrating planar and 3D cQED devices. *Applied Physics Letters*, 109(4):042601, jul 2016.

[123] Philipp Kurpiers, Theodore Walter, Paul Magnard, Yves Salathe, and Andreas Wallraff. Characterizing the attenuation of coaxial and rectangular microwave-frequency waveguides at cryogenic temperatures. *EPJ Quantum Technology*, 4(1):8, dec 2017.

[124] T. Brecht, M. Reagor, Y. Chu, W. Pfaff, C. Wang, L. Frunzio, M. H. Devoret, and R. J. Schoelkopf. Demonstration of superconducting micromachined cavities. *Applied Physics Letters*, 107(19), nov 2015, arXiv:1509.01119 [cond-mat.supr-con].

[125] Chan U Lei, Lev Krayzman, Suhas Ganjam, Luigi Frunzio, and Robert J. Schoelkopf. High coherence superconducting microwave cavities with indium bump bonding. jan 2020, arXiv:2001.09216 [quant-ph].

[126] K. Bergmann, H. Theuer, and B. W. Shore. Coherent population transfer among quantum states of atoms and molecules. *Reviews of Modern Physics*, 70(3):1003–1025, jul 1998.

[127] B. Vogell, B. Vermersch, T. E. Northup, B. P. Lanyon, and C. A. Muschik. Deterministic quantum state transfer between remote qubits in cavities. *Quantum Science and Technology*, 2(4), dec 2017, arXiv:1704.06233 [quant-ph].

[128] Y. P. Zhong, H. S. Chang, K. J. Satzinger, M. H. Chou, A. Bienfait, C. R. Conner, Dumur, J. Grebel, G. A. Peairs, R. G. Povey, D. I. Schuster, and A. N. Cleland. Violating Bell's inequality with remotely connected superconducting qubits, aug 2019, arXiv:1808.03000 [quant-ph].

[129] J. Z. Blumoff, K. Chou, C. Shen, M. Reagor, C. Axline, R. T. Brierley, M. P. Silveri, C. Wang, B. Vlastakis, S. E. Nigg, L. Frunzio, M. H. Devoret, L. Jiang, S. M. Girvin, and R. J. Schoelkopf. Implementing and characterizing precise multiqubit measurements. *Physical Review X*, 6(3), 2016, arXiv:1606.00817 [quant-ph].

[130] Matthew Reagor. *Superconducting Cavities for Circuit Quantum Electrodynamics.* PhD thesis, New Haven CT, 2015.

[131] Jens Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Charge-insensitive qubit design derived from the Cooper pair box. *Physical Review A - Atomic, Molecular, and Optical Physics*, 76(4), 2007, 0703002 [cond-mat.mes-hall].

[132] Simon E. Nigg, Hanhee Paik, Brian Vlastakis, Gerhard Kirchmair, S. Shankar, Luigi Frunzio, M. H. Devoret, R. J. Schoelkopf, and S. M. Girvin. Black-box superconducting circuit quantization. *Physical Review Letters*, 108(24), jun 2012, arXiv:1204.0587 [cond-mat.mes-hall].

[133] John M. Martinis, Michel H. Devoret, and John Clarke. Energy-level quantization in the zero-voltage state of a current-biased Josephson junction. *Physical Review Letters*, 55(15):1543–1546, 1985.

[134] V. Bouchiat, D. Vion, P. Joyez, D. Esteve, and M. H. Devoret. Quantum coherence with a single cooper pair. *Physica Scripta T*, 76:165–170, 1998.

[135] T. P. Orlando, J. E. Mooij, Lin Tian, Caspar H. Van Der Wal, L. S. Levitov, Seth Lloyd, and J. J. Mazo. Superconducting persistent-current qubit. *Physical Review B - Condensed Matter and Materials Physics*, 60(22):15398–15413, 1999, 9908283.

[136] Y. Nakamura, Yu A. Pashkin, and J. S. Tsai. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature*, 398(6730):786–788, apr 1999, 9904003 [cond-mat.mes-hall].

[137] Vladimir E. Manucharyan, Jens Koch, Leonid I. Glazman, and Michel H. Devoret. Fluxonium: Single cooper-pair circuit free of charge offsets. *Science*, 326(5949):113–116, 2009, arXiv:0906.0831 [cond-mat.mes-hall].

[138] Hanhee Paik, D. I. Schuster, Lev S. Bishop, G. Kirchmair, G. Catelani, A. P. Sears, B. R. Johnson, M. J. Reagor, L. Frunzio, L. I. Glazman, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf. Observation of high coherence in Josephson junction qubits measured in a three-dimensional circuit QED architecture. *Physical Review Letters*, 107(24), dec 2011, arXiv:1105.4652 [quant-ph].

[139] Florent Lecocq, Ioan M. Pop, Zhihui Peng, Iulian Matei, Thierry Crozes, Thierry Fournier, Cécile Naud, Wiebke Guichard, and Olivier Buisson. Junction

fabrication by shadow evaporation without a suspended bridge. *Nanotechnology*, 22(31), aug 2011.

[140] G. J. Dolan. Offset masks for lift-off photoprocessing. *Applied Physics Letters*, 31(5):337–339, 1977.

[141] Zlatko Minev. *Catching and Reversing a Quantum Jump Mid-Flight.* PhD thesis, New Haven CT, 2018.

[142] E. T. Holland, B. Vlastakis, R. W. Heeres, M. J. Reagor, U. Vool, Z. Leghtas, L. Frunzio, G. Kirchmair, M. H. Devoret, M. Mirrahimi, and R. J. Schoelkopf. Single-Photon-Resolved Cross-Kerr Interaction for Autonomous Stabilization of Photon-Number States. *Physical Review Letters*, 115(18), oct 2015, arXiv:1504.03382 [quant-ph].

[143] M. D. Reed, B. R. Johnson, A. A. Houck, L. Dicarlo, J. M. Chow, D. I. Schuster, L. Frunzio, and R. J. Schoelkopf. Fast reset and suppressing spontaneous emission of a superconducting qubit. *Applied Physics Letters*, 96(20), may 2010, arXiv:1003.0142 [cond-mat.mes-hall].

[144] Evan Jeffrey, Daniel Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Megrant, P. J. O'Malley, C. Neill, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and John M. Martinis. Fast accurate state measurement with superconducting qubits. *Physical Review Letters*, 112(19), may 2014, arXiv:1401.0257 [quant-ph].

[145] Eyob A. Sete, John M. Martinis, and Alexander N. Korotkov. Quantum theory of a bandpass Purcell filter for qubit readout. *Physical Review A - Atomic, Molecular, and Optical Physics*, 92(1), jul 2015, arXiv:1504.06030 [quant-ph].

[146] Jay Gambetta, Alexandre Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting. *Physical Review A - Atomic, Molecular, and Optical Physics*, 74(4), 2006.

[147] Jay Gambetta, W. A. Braff, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf. Protocols for optimal readout of qubits using a continuous quantum nondemolition measurement. *Physical Review A - Atomic, Molecular, and Optical Physics*, 76(1), jul 2007, arXiv: 0701078.

[148] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff. Rapid High-Fidelity Single-Shot Dispersive Readout of Superconducting Qubits. *Physical Review Applied*, 7(5), may 2017, arXiv: 1701.06933.

[149] N. Bergeal, F. Schackert, M. Metcalfe, R. Vijay, V. E. Manucharyan, L. Frunzio, D. E. Prober, R. J. Schoelkopf, S. M. Girvin, and M. H. Devoret. Phase-preserving amplification near the quantum limit with a Josephson ring modulator. *Nature*, 465(7294):64–68, may 2010, arXiv:0912.3407 [cond-mat.mes-hall].

[150] Baleegh Abdo, Flavius Schackert, Michael Hatridge, Chad Rigetti, and Michel Devoret. Josephson amplifier for qubit readout. *Applied Physics Letters*, 99(16), oct 2011, arXiv:1103.1405 [cond-mat.supr-con].

[151] N. E. Frattini, V. V. Sivak, A. Lingenfelter, S. Shankar, and M. H. Devoret. Optimizing the Nonlinearity and Dissipation of a SNAIL Parametric Amplifier for Dynamic Range. *Physical Review Applied*, 10(5):054020, nov 2018, arXiv: 1806.06093 [quant-ph].

[152] Easwar Magesan, Jay M. Gambetta, A. D. Córcoles, and Jerry M. Chow. Machine Learning for Discriminating Quantum Measurement Trajectories and Improving Readout. *Physical Review Letters*, 114(20), may 2015, arXiv:1411.4994 [quant-ph].

[153] Yanbing Liu, Srikanth J. Srinivasan, D. Hover, Shaojiang Zhu, R. McDermott, and A. A. Houck. High fidelity readout of a transmon qubit using a superconducting low-inductance undulatory galvanometer microwave amplifier. *New Journal of Physics*, 16, nov 2014.

[154] D. T. McClure, Hanhee Paik, L. S. Bishop, M. Steffen, Jerry M. Chow, and Jay M. Gambetta. Rapid Driven Reset of a Qubit Readout Resonator. *Physical Review Applied*, 5(1), jan 2016, arXiv:1503.01456 [quant-ph].

[155] Daniel Sank, Zijun Chen, Mostafa Khezri, J. Kelly, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Mutus, M. Neeley, C. Neill, P. J.J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, T. White, J. Wenner, Alexander N. Korotkov, John M. Martinis, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, T. White, J. Wenner, Alexander N. Korotkov, and John M. Martinis. Measurement-Induced State Transitions in a Superconducting Qubit: Beyond the Rotating Wave Approximation. *Physical Review Letters*, 117(19):190503, nov 2016, arXiv:1606.05721 [quant-ph].

[156] Th. K. Mavrogordatos, G. Tancredi, M. Elliott, M.J. Peterer, A. Patterson, J. Rahamim, P.J. Leek, E. Ginossar, and M.H. Szymańska. Simultaneous Bistability of a Qubit and Resonator in Circuit Quantum Electrodynamics. *Physical Review Letters*, 118(4):040402, jan 2017.

[157] Salvatore S. Elder, Christopher S. Wang, Philip Reinhold, Connor T. Hann, Kevin S. Chou, Brian J. Lester, Serge Rosenblum, Luigi Frunzio, Liang Jiang, and Robert J. Schoelkopf. High-Fidelity Measurement of Qubits Encoded in Multilevel Superconducting Circuits. *Physical Review X*, 10(1):011001, jan 2020.

[158] François Mallet, Florian R. Ong, Agustin Palacios-Laloy, François Nguyen, Patrice Bertet, Denis Vion, and Daniel Esteve. Single-shot qubit readout in circuit quantum electrodynamics. *Nature Physics*, 5(11):791–795, 2009, arXiv:1005.5615 [quant-ph].

[159] Philipp Kurpiers, Marek Pechal, Baptiste Royer, Paul Magnard, Theo Walter, Johannes Heinsoo, Yves Salathé, Abdulkadir Akin, Simon Storz, Jean-Claude C. Besse, Simone Gasparinetti, Alexandre Blais, and Andreas Wallraff. Quantum Communication with Time-Bin Encoded Microwave Photons. *Physical Review Applied*, 12(4):044067, nov 2019, arXiv:1811.07604 [quant-ph].

[160] D. I. Schuster, A. A. Houck, J. A. Schreier, A. Wallraff, J. M. Gambetta, A. Blais, L. Frunzio, J. Majer, B. Johnson, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Resolving photon number states in a superconducting circuit. *Nature*, 445(7127):515–518, feb 2007, 0608693 [cond-mat.mes-hall].

[161] B. R. Johnson, M. D. Reed, A. A. Houck, D. I. Schuster, Lev S. Bishop, E. Ginossar, J. M. Gambetta, L. Dicarlo, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Quantum non-demolition detection of single microwave photons in a circuit. *Nature Physics*, 6(9):663–667, 2010, arXiv:1003.2734 [cond-mat.mes-hall].

[162] Stefan Krastanov, Victor V. Albert, Chao Shen, Chang Ling Zou, Reinier W. Heeres, Brian Vlastakis, Robert J. Schoelkopf, and Liang Jiang. Universal control of an oscillator with dispersive coupling to a qubit. *Physical Review A - Atomic, Molecular, and Optical Physics*, 92(4), oct 2015, arXiv:1502.08015 [quant-ph].

[163] Reinier W. Heeres, Brian Vlastakis, Eric Holland, Stefan Krastanov, Victor V. Albert, Luigi Frunzio, Liang Jiang, and Robert J. Schoelkopf. Cavity State Manipulation Using Photon-Number Selective Phase Gates. *Physical Review Letters*, 115(3), sep 2015, arXiv:1503.01496 [quant-ph].

[164] Yuan Xu, Yuwei Ma, Weizhou Cai, Xianghao Mu, Wei Dai, Weiting Wang, Ling Hu, Xuegang Li, Jiaxiu Han, Haiyan Wang, Yipu Song, Zhen-Biao Yang, Shi-Biao Zheng, and Luyan Sun. Geometrically manipulating photonic Schrödinger cat states and realizing two-logical-qubit gates. oct 2018, arXiv:1810.04690 [quant-ph].

[165] Brian Vlastakis, Gerhard Kirchmair, Zaki Leghtas, Simon E. Nigg, Luigi Frunzio, S. M. Girvin, Mazyar Mirrahimi, M. H. Devoret, and R. J. Schoelkopf. Deterministically encoding quantum information using 100-photon Schrödinger cat states. *Science*, 342(6158):607–610, 2013.

[166] L. Sun, A. Petrenko, Z. Leghtas, B. Vlastakis, G. Kirchmair, K. M. Sliwa, A. Narla, M. Hatridge, S. Shankar, J. Blumoff, L. Frunzio, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf. Tracking photon jumps with repeated quantum non-demolition parity measurements. *Nature*, 511(7510):444–448, 2014, arXiv:1311.2534 [quant-ph].

[167] Chen Wang, Yvonne Y. Gao, Philip Reinhold, R. W. Heeres, Nissim Ofek, Kevin Chou, Christopher Axline, Matthew Reagor, Jacob Blumoff, K. M. Sliwa,

L. Frunzio, S. M. Girvin, Liang Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf. A Schrödinger cat living in two boxes. *Science*, 352(6289):1087–1091, may 2016, arXiv:1601.05505 [quant-ph].

[168] Brian Vlastakis. *Controlling Coherent State Superpositions with Superconducting Circuits*. PhD thesis, New Haven CT, 2015.

[169] Navin Khaneja, Timo Reiss, Cindie Kehlet, Thomas Schulte-Herbrüggen, and Steffen J. Glaser. Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms. *Journal of Magnetic Resonance*, 172(2):296–305, feb 2005.

[170] P. De Fouquieres, S. G. Schirmer, S. J. Glaser, and Ilya Kuprov. Second order gradient ascent pulse engineering. *Journal of Magnetic Resonance*, 212(2):412–417, oct 2011.

[171] Florian Dolde, Ville Bergholm, Ya Wang, Ingmar Jakobi, Boris Naydenov, Sébastien Pezzagna, Jan Meijer, Fedor Jelezko, Philipp Neumann, Thomas Schulte-Herbrüggen, Jacob Biamonte, and Jörg Wrachtrup. High-fidelity spin entanglement using optimal control. *Nature Communications*, 5, feb 2014, arXiv:1309.4430 [quant-ph].

[172] B. E. Anderson, H. Sosa-Martinez, C. A. Riofrío, Ivan H. Deutsch, and Poul S. Jessen. Accurate and Robust Unitary Transformations of a High-Dimensional Quantum System. *Physical Review Letters*, 114(24), jun 2015.

[173] Tommaso Caneva, Tommaso Calarco, and Simone Montangero. Chopped random-basis quantum optimization. *Physical Review A - Atomic, Molecular, and Optical Physics*, 84(2), aug 2011, arXiv:1103.0855 [quant-ph].

[174] N. Rach, M. M. Müller, T. Calarco, and S. Montangero. Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape. *Physical Review A - Atomic, Molecular, and Optical Physics*, 92(6), dec 2015, arXiv:1506.04601 [quant-ph].

[175] O. V. Morzhin and A. N. Pechen. Krotov method for optimal control of closed quantum systems. *Russian Mathematical Surveys*, 74(5):851–908, oct 2019.

[176] Philip Reinhold. *Controlling Error-Correctable Bosonic Qubits*. PhD thesis, New Haven CT, 2019.

[177] N. E. Frattini, U. Vool, S. Shankar, A. Narla, K. M. Sliwa, and M. H. Devoret. 3-wave mixing Josephson dipole element. *Applied Physics Letters*, 110(22):222603, may 2017.

[178] Yaxing Zhang, Brian J. Lester, Yvonne Y. Gao, Liang Jiang, R. J. Schoelkopf, and S. M. Girvin. Engineering bilinear mode coupling in circuit QED: Theory and experiment. *Physical Review A*, 99(1):012314, 2019.

[179] Yiwen Chu, Prashanta Kharel, William H. Renninger, Luke D. Burkhart, Luigi Frunzio, Peter T. Rakich, and Robert J. Schoelkopf. Quantum acoustics with superconducting qubits. *Science*, 358(6360):199–202, oct 2017, arXiv:1703.00342 [quant-ph].

[180] S. O. Mundhada, A. Grimm, J. Venkatraman, Z. K. Minev, S. Touzard, N. E. Frattini, V. V. Sivak, K. Sliwa, P. Reinhold, S. Shankar, M. Mirrahimi, and M. H. Devoret. Experimental Implementation of a Raman-Assisted Eight-Wave Mixing Process. *Physical Review Applied*, 12(5):054051, nov 2019, arXiv: 1811.06589.

[181] Simon E. Nigg and S. M. Girvin. Stabilizer quantum error correction toolbox for superconducting qubits. *Physical Review Letters*, 110(24), jun 2013, arXiv:1212.4000 [quant-ph].

[182] J. Wenner, Yi Yin, Yu Chen, R. Barends, B. Chiaro, E. Jeffrey, J. Kelly, A. Megrant, J. Y. Mutus, C. Neill, P. J.J. O'Malley, P. Roushan, D. Sank, A. Vainsencher, T. C. White, Alexander N. Korotkov, A. N. Cleland, and John M. Martinis. Catching time-reversed microwave coherent state photons with 99.4% absorption efficiency. *Physical Review Letters*, 112(21), may 2014.

[183] Alexander N. Korotkov. Flying microwave qubits with nearly perfect transfer efficiency. *Physical Review B - Condensed Matter and Materials Physics*, 84(1), jul 2011, arXiv:1103.0592 [cond-mat.mes-hall].

[184] Srikanth J. Srinivasan, Neereja M. Sundaresan, Darius Sadri, Yanbing Liu, Jay M. Gambetta, Terri Yu, S. M. Girvin, and Andrew A. Houck. Time-reversal symmetrization of spontaneous emission for quantum state transfer. *Physical Review A - Atomic, Molecular, and Optical Physics*, 89(3), mar 2014, arXiv:1308.3471 [quant-ph].

[185] S. Massar and S. Popescu. Optimal Extraction of Information from Finite Quantum Ensembles. *Physical Review Letters*, 74(8):1259–1263, feb 1995.

[186] Evan Zalys-Geller. .. PhD thesis, New Haven CT, 2020.

[187] K. M. Sliwa, M. Hatridge, A. Narla, S. Shankar, L. Frunzio, R. J. Schoelkopf, and M. H. Devoret. Reconfigurable josephson circulator/directional amplifier. *Physical Review X*, 5(4), 2015, arXiv:1503.00209 [quant-ph].

[188] F. Lecocq, L. Ranzani, G. A. Peterson, K. Cicak, R. W. Simmonds, J. D. Teufel, and J. Aumentado. Nonreciprocal Microwave Signal Processing with a Field-Programmable Josephson Amplifier. *Physical Review Applied*, 7(2), feb 2017, arXiv:1612.01438 [quant-ph].

[189] Baleegh Abdo, Nicholas T. Bronn, Oblesh Jinka, Salvatore Olivadese, Antonio D. Córcoles, Vivekananda P. Adiga, Markus Brink, Russell E. Lake, Xian

Wu, David P. Pappas, and Jerry M. Chow. Active protection of a superconducting qubit with an interferometric Josephson isolator. *Nature Communications*, 10(1), dec 2019, arXiv:1810.07234 [quant-ph].

[190] Rajibul Islam, Ruichao Ma, Philipp M. Preiss, M. Eric Tai, Alexander Lukin, Matthew Rispoli, and Markus Greiner. Measuring entanglement entropy in a quantum many-body system. *Nature*, 528(7580):77–83, dec 2015.

[191] S. Rosenblum, Y. Y. Gao, P. Reinhold, C. Wang, C. J. Axline, L. Frunzio, S. M. Girvin, Liang Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf. A CNOT gate between multiphoton qubits encoded in two cavities. *Nature Communications*, 9(1), dec 2018, arXiv:1709.05425 [quant-ph].

[192] Jerry M. Chow, A. D. Córcoles, Jay M. Gambetta, Chad Rigetti, B. R. Johnson, John A. Smolin, J. R. Rozen, George A. Keefe, Mary B. Rothwell, Mark B. Ketchen, and M. Steffen. Simple all-microwave entangling gate for fixed-frequency superconducting qubits. *Physical Review Letters*, 107(8), aug 2011, arXiv:1106.0553 [quant-ph].

[193] Hanhee Paik, A. Mezzacapo, Martin Sandberg, D. T. McClure, B. Abdo, A. D. Córcoles, O. Dial, D. F. Bogorin, B. L.T. Plourde, M. Steffen, A. W. Cross, J. M. Gambetta, and Jerry M. Chow. Experimental Demonstration of a Resonator-Induced Phase Gate in a Multiqubit Circuit-QED System. *Physical Review Letters*, 117(25), dec 2016, arXiv:1606.00685 [quant-ph].