THE UNIVERSITY OF CHICAGO

QUANTUM COMPUTING WITH DISTRIBUTED MODULES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY
NELSON LEUNG

CHICAGO, ILLINOIS
MARCH 2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would first like to thank my advisor Professor Dave Schuster for setting up a great environment for performing excellent scientific research and assembling an exceptional group of people to work together. I am very fortunate to join the lab at a time that I had lots of opportunities working with Dave side by side, where I have learned from him a tremendous amount of scientific knowledge and problem-solving skills. I have been greatly benefited from his ideas, guidance, and discussion throughout my graduate study.

This work has been possible through collaborations with my lab mates, in particular, Yao Lu, Srivatsan Chakram and Ravi Naik, whom I have the greatest gratitude and respect. It has been an extraordinary experience to be in a lab with such heterogeneous combination of research projects, and be surrounded by colleagues with diverse and very brilliant set of skills and expertise. I would like to thank everyone in the Schuster lab for their help, camaraderie, and kindness.

Thank you to all my friends for their companionship and support. Most importantly, I would like to thank my family - my parents and my wife Ziwei for their unconditional love, encouragement, and support.

ABSTRACT

This thesis describes our work on optimizing quantum control and scaling up quantum computers. We demonstrate the application of automatic differentiation in quantum optimal control, which allows us to specify advanced optimization criteria and incorporate them in the optimization process of quantum control with ease. These criteria enable more intricate control on the evolution path, suppression of departures from the truncated model subspace, as well as minimization of the physical time needed to perform high-fidelity state preparation and unitary gates. We subsequently present our experimental efforts in efficiently scaling up a quantum computer. The innovation of quantum random access memory (qRAM) architecture with multimode circuit QED establishes a promising path towards expanding the information processing power of a quantum device, without introducing control-resources overhead. Using a single Josephson junction transmon circuit serving as the central processor, we demonstrate universal operations on a nine-qubit random access memory. We further present our experimental results in establishing bidirectional and deterministic photonic communication between two remote superconducting multimode processors, connected through a one-meter long coaxial cable. This device demonstrates a prototype of setting up a scalable and distributed quantum computing cluster that could solve a harder computational problem.

PREFACE

With a finite amount of control resources, how can we improve information processing capabilities? With multiple computational units in hand, how can they be combined to solve a larger problem? How can we improve the control efficiency of a computational device? Classical computing seems to have answers to all these problems. Rarely a classical computing architecture would exclusively use expensive transistors for computation. There are usually components of random access memory (RAM) and storage device. A more powerful supercomputer is not just formed by a more advanced individual processing chip, it is also formed by a larger and more connected network of chips. The optimization of software compilation has been crucial in pushing the limit of computational performance. How could these classical ideas be transferred to the realm of quantum computing technologies is not straightforward, since the physics of quantum computing are governed by the laws of quantum mechanics, and its nature is very different from the classical computing world. This thesis presents our effort in tackling these challenging problems for quantum computers, with the goal of building a scalable universal quantum computer.

CHAPTER 1

QUANTUM COMPUTING

1.1 Introduction to Quantum Computing

For an n-qubit system, the quantum state $|\psi\rangle$ can be mathematically described by

$$|\psi\rangle = \sum_{i_1,i_2\ldots i_n=0}^{1} a_{i_1 i_2 \ldots i_n} |i_1 i_2 \ldots i_n\rangle \tag{1.1}$$

$$\sum_{i_1,i_2\ldots i_n=0}^{1} |a_{i_1 i_2 \ldots i_n}|^2 = 1 \tag{1.2}$$

$$\langle i'_1 i'_2 \ldots i'_n | i_1 i_2 \ldots i_n \rangle = \delta_{i_1}^{i'_1} \delta_{i_2}^{i'_2} \ldots \delta_{i_n}^{i'_n} \tag{1.3}$$

The space of the quantum state is span by 2^n orthogonal basis vectors $|i_1 i_2 \ldots i_n\rangle$. The $a_{i_1 i_2 \ldots i_n}$ are complex numbers, with the condition that the sum of L^2-norm of $a_{i_1 i_2 \ldots i_n}$ is normalized to 1. A quantum gate can be described by a unitary transformation acting on the state $|\psi\rangle$, which would update the coefficients $a_{i_1 i_2 \ldots i_n}$ but preserve the normalization condition. The fact that the state size of an n-qubit system scales exponentially with n is the reason that a quantum computer is very difficult to simulate with classical resources.

That does not imply a quantum computer necessarily have exponentially large computational power. A caveat of the former description is the coefficients $a_{i_1 i_2 \ldots i_n}$ are not directly observable. After a measurement, an n-qubit system would collapse to one of the basis state $|i_1 i_2 \ldots i_n\rangle$, where the quantity of $|a_{i_1 i_2 \ldots i_n}|^2$ is the probability of obtaining the basis state $|i_1 i_2 \ldots i_n\rangle$ in a single measurement. Thus, the coefficients can only be interpreted through an exponentially large number of repeated state preparations and measurements. Only specially designed quantum algorithms could provide a speedup over classical algorithms, and we will describe a few examples in the next section. For a comprehensive introduction of quantum computing, the reader should consult the

book written by Isaac and Nielsen [137].

1.2 High-Level Description of Some Quantum Algorithms

1.2.1 Grover's algorithm

Given a function with a domain of N discrete elements each represented by a $log_2 N$ bit binary string, with an output

$$f_\omega(x) = \begin{cases} -1, & \text{if } x = \omega \\ 1, & \text{else} \end{cases} \tag{1.4}$$

The ω is a member of the N discrete elements, and its identity is unknown to us a priori. The Grover's algorithm can be used for inverting this given function (i.e., providing the identity of ω) with the number of queries scaling as $\sqrt{N}$ [69]. It works through amplitude amplification, which iteratively amplifies the magnitude of the coefficient corresponding to $|\omega\rangle$. A classical algorithm can perform no better than brute force search with repeated trial-and-error, which has a scaling of N. The quadratic speedup of Grover's algorithm can be used to accelerate classical brute force search for hard problems, which could have a wide range of applications.

1.2.2 Shor's algorithm

Shor's algorithm is an integer factorization algorithm, which finds the prime factors of a given integer N [168]. Shor's algorithm runs with a time-complexity polynomial of $logN$, which is almost exponentially faster than the fastest known classical algorithm, the general number field sieve. Factoring large integers is believed to be classically intractable and forms the theoretical basis of security for the widely used public-key cryptography scheme RSA. This became a strong motivator for the design and construction of quantum computers. It has also induced research on new cryptography schemes that are secure from powerful quantum computers, known as post-

quantum cryptography.

1.2.3 Variational Quantum Eigensolver (VQE)

The Variational Quantum Eigensolver (VQE) is a classical-quantum hybrid algorithm that can be used as an eigensolver for the ground state of a Hamiltonian H [116]. Recent progress in the application of chemistry with near-term quantum hardware was largely driven by this algorithm. The algorithm consists of two major steps:

1. Prepare a quantum state $\left|\psi(\vec{\theta})\right\rangle$ in a quantum computer. The parameter $\vec{\theta}$ determines the variables of qubit controls. For example, the rotation angles of a set of single qubit gates or C-phase gates.

2. Measure the expectation value of $\left\langle\psi(\vec{\theta})\right| H \left|\psi(\vec{\theta})\right\rangle$ with the quantum computer.

A classical optimization algorithm can thus interact with the quantum computer by providing the parameter $\vec{\theta}$ and obtain a numerical value $\left\langle\psi(\vec{\theta})\right| H \left|\psi(\vec{\theta})\right\rangle$, which can be treated as the cost function for the optimization. Mathematically, this can be expressed as

$$\min_{\vec{\theta}} \left\langle\psi(\vec{\theta})\right| H \left|\psi(\vec{\theta})\right\rangle \tag{1.5}$$

After the optimization converges, we obtain the ground eigenstate $\left|\psi(\vec{\theta'})\right\rangle$, which we can subsequently calculate its properties.

1.3 Quantum Architectures

1.3.1 Trapped Ions

Trapped ion quantum computer consists of charged atomic particles (ions) confined and suspended in free space with electromagnetic fields [33]. Qubits are defined with stable electronic states of each ion, with relatively long coherence time. The trapped ion quantum computer is one of the

earliest experimental platforms for quantum computing research and has remained to be one of the most promising candidate architectures for a scalable universal quantum computer [90].

1.3.2 Silicon

One of the most promising silicon quantum computing architectures is based on an array of individual phosphorus atoms embedded in pure silicon lattice, which each phosphorus atoms can be used as a qubit [85]. The process to fabricate such architecture involves using a scanning tunneling microscope (STM) for atomically-precise fabrication for the placement of phosphorus atoms [165]. The silicon quantum computer is readily scalable to a large array of qubits and maintaining a clean environment with the pure silicon environment.

1.3.3 Superconducting qubits

Superconducting quantum computing is one of the most popular and promising architectures for realizing a scalable universal quantum computer and is currently pursued by industrial efforts including Google, IBM, Intel, Raytheon, Rigetti Computing, etc. Several qualities of superconducting architecture made it an attractive candidate. While most other architectures use natural atoms, superconducting qubits are made of human-designed electrical circuits, which is the reason sometimes they are called artificial-atoms [40]. This allows physicists to easily design and engineer atomic properties they desire. The macroscopic nature of these quantum circuits made it relatively easy to control and scale up. Moreover, the fabrication techniques required for superconducting qubits is very similar to the existing technology developed by the semiconductor industry. Our lab, and thus this thesis, focuses on the development of superconducting quantum computing technologies.

1.4 Thesis Overview

This chapter summarizes the basic theories, algorithms and experimental architectures of quantum computing. Chapter 2 describes the physical theory of how qubits can be formed with superconducting circuits. Chapter 3 discusses the theory and experimental setup for controlling superconducting qubits. The subsequent three Chapters report my published work during my doctoral research in Schuster Lab. They involve numerical optimization of qubit control (Chapter 4), design and implementation of quantum random access memory (Chapter 5), and establishing photonic quantum communications between remote devices (Chapter 6).

CHAPTER 2

CIRCUIT QUANTUM ELECTRODYNAMICS

This chapter describes the basic theories of superconducting qubits. We begin our discussion with the theory of quantizing a quantum harmonic oscillator, and how the collection of electrons in a LC resonator would behave like a single particle. We then introduce the circuit element Josephson junction, which serves as a crucial component for realizing superconducting qubits. We subsequently describe the quantization theory of a Transmon qubit and coupled resonators. The dispersive shift between Transmon and resonator allows one to conveniently measure the transmon state.

2.1 Quantization of Quantum Harmonic Oscillator

One of the most elementary and fundamental systems in quantum mechanics is the quantum harmonic oscillator. The theory used for solving a simple quantum harmonic oscillator is essential for understanding the mechanisms of quantum bits (qubits). Here, we begin our discussion with a simple Harmonic oscillator in the position (x) and momentum (p) space, with the Hamiltonian

$$H = \frac{kx^2}{2} + \frac{p^2}{2m} = \frac{1}{2}m\omega^2 x^2 + \frac{p^2}{2m}, \tag{2.1}$$

where $\omega = \sqrt{\frac{k}{m}}$ is the angular frequency of the oscillator. One obtains the following relationships

$$\begin{aligned} \frac{\partial H}{\partial x} &= m\omega^2 x = kx = -\dot{p} \\ \frac{\partial H}{\partial p} &= \frac{p}{m} = \dot{x} \end{aligned} \tag{2.2}$$

Thus the variables x and p are canonical conjugate variables. A fundamental relation in quantum mechanics is the canonical commutation relation between canonical conjugate quantities. For position and momentum, we have the relation

$$[x, p] = i\hbar. \tag{2.3}$$

This is the only proposition we assume to be true a priori. Every derivation in this section can be derived from this proposition. We can thus decompose the Hamiltonian to be

$$\begin{aligned} H &= \frac{m\omega^2}{2}(x - i\frac{1}{m\omega}p)(x + i\frac{1}{m\omega}p) - i\frac{1}{2}\omega[x, p] \\ &= \hbar\omega[\sqrt{\frac{m\omega}{2\hbar}}(x - i\frac{1}{m\omega}p)][\sqrt{\frac{m\omega}{2\hbar}}(x + i\frac{1}{m\omega}p)] + \frac{1}{2}\hbar\omega. \end{aligned} \tag{2.4}$$

We can defined the operators a and its adjoint $a^\dagger$ as

$$\begin{aligned} a &= \sqrt{\frac{m\omega}{2\hbar}}(x + i\frac{1}{m\omega}p) \\ a^\dagger &= \sqrt{\frac{m\omega}{2\hbar}}(x - i\frac{1}{m\omega}p) \end{aligned} \tag{2.5}$$

Using equation 2.3, we obtain

$$\begin{aligned} [a, a^\dagger] &= 1 \\ [a^\dagger a, a^\dagger] &= a^\dagger \\ [a^\dagger a, a] &= -a \end{aligned} \tag{2.6}$$

$$\tag{2.7}$$

We may rewrite the Hamiltonian at equation 2.4 to be

$$H = \hbar\omega(a^\dagger a + \frac{1}{2}) \tag{2.8}$$

From this equation, it is obvious that the eigenstate $a^\dagger a$ is the eigenstate of H. Let $|n\rangle$ be the

eigenstate of $a^\dagger a$ with eigenvalue n. $|n\rangle$ is just a notation of an eigenstate, and n is a real number associated with the eigenstate. At this point, we make no assumption of the nature of n (such as whether it is an integer). Now consider the relation

$$\begin{aligned}(a^\dagger a)a^\dagger |n\rangle &= (a^\dagger a^\dagger a + [a^\dagger a, a^\dagger]) |n\rangle \\ &= (a^\dagger a^\dagger a + a^\dagger) |n\rangle \\ &= (n+1)a^\dagger |n\rangle\end{aligned} \tag{2.9}$$

$$\tag{2.10}$$

Similarly,

$$(a^\dagger a)a |n\rangle = (n-1)a |n\rangle \tag{2.11}$$

Therefore, $a^\dagger |n\rangle$ and $a |n\rangle$ are eigenstates of $a^\dagger a$. For any given eigenstate $|n\rangle$, one can apply the operators a and $a^\dagger$ to generate new eigenstates of $a^\dagger a$.

One can obtain the norm of the eigenstate $a |n\rangle$ by the following equations:

$$n = \langle n| (a^\dagger a |n\rangle) \tag{2.12}$$

$$= (\langle n| a^\dagger)(a |n\rangle) \tag{2.13}$$

$$= (a |n\rangle)^\dagger (a |n\rangle) \tag{2.14}$$

Since $a |n\rangle$ satisfies $(a^\dagger a)a |n\rangle = (n-1)a |n\rangle$ and $(a |n\rangle)^\dagger (a |n\rangle) = n$, we can express $a |n\rangle$ as

$$a |n\rangle = \sqrt{n} |n-1\rangle , \tag{2.15}$$

which automatically satisfies the above relations. Similarly, one can obtain the relation

$$a^\dagger |n\rangle = \sqrt{n+1}\,|n+1\rangle\,. \tag{2.16}$$

Now, assume n is non-integer and positive valued λ, we can apply the a operator indefinitely to generate infinite number of eigenstates with infinitely negative eigenvalues. For instance, for an integer k,

$$a^k |\lambda\rangle = \sqrt{(\lambda)(\lambda-1)(\lambda-2)\dots(\lambda-k)}\,|\lambda-k\rangle\,. \tag{2.17}$$

The coefficient $\sqrt{(\lambda)(\lambda-1)(\lambda-2)\dots(\lambda-k)}$ is non-zero if and only if λ is a non-integer. if λ is a non-integer, when k tends to ∞, the eigenvalues tend to $-\infty$ without bound, which is unphysical as every realistic physical system has a stable ground state. Therefore, we conclude that λ must be an integer, with the ground state equal to $|0\rangle$. Thus, a harmonic oscillator has eigenvalues $\hbar\omega(n+\frac{1}{2})$ with n being integers. The harmonic oscillator is considered to be a linear system as it has an equal spacing of energy differences.

2.2 Circuit Quantum Harmonic Oscillator

The aforementioned harmonic oscillator is common to most Physicists, which the wave functions can be understood as describing a single particle confined in a harmonic field. On the other hand, one rarely encounters the quantization of another popular harmonic oscillator we learned in Physics: the LC oscillator formed by an inductor (L) and a capacitor (C). The LC oscillator has a Hamiltonian very similar to the aforementioned harmonic oscillator:

$$H = \frac{1}{2}CV^2 + \frac{1}{2}LI^2 = \frac{q^2}{2C} + \frac{\phi^2}{2L}, \tag{2.18}$$

with $q = VC$ and $\phi = LI$. The capacitor and inductor are assumed to be non-dissipative elements, which can be realized by using superconductors.

A distinct feature of the LC oscillator is the variables q and ϕ are formed by a larger collection

of electrons. Remarkably, the collective motion of electrons forms a single set of conjugate variables! This has been experimentally demonstrated numerous times. The following shows how this can be understood theoretically.

The inductor and capacitor of an LC oscillator share the same voltage, therefore

$$V = -L\dot{I} = \frac{q}{C}. \tag{2.19}$$

We obtain

$$\begin{aligned} \frac{\partial H}{\partial q} &= \frac{q}{C} = -L\dot{I} = -\dot{\phi} \\ \frac{\partial H}{\partial \phi} &= \frac{\phi}{L} = I = \dot{q}. \end{aligned} \tag{2.20}$$

Thus the variables q and ϕ are canonical conjugate variables, which they must satisfy the commutative relation

$$[\phi, q] = i\hbar \tag{2.21}$$

Using a similar procedure resulting equation 2.5, the annihilation and creation operators of a LC oscillator are given by

$$\begin{aligned} a &= \frac{1}{\sqrt{2\hbar Z_c}}(\phi + iZ_c q) \\ a^\dagger &= \frac{1}{\sqrt{2\hbar Z_c}}(\phi - iZ_c q), \end{aligned} \tag{2.22}$$

where $Z_c = \sqrt{L/C}$.

By rearranging the equations of 2.22, we can obtain a set of useful expressions of the conjugate variables (charge and flux) in terms of creation and annihilation operators as

$$
\begin{aligned}
q &= i\sqrt{\frac{\hbar}{2Z_c}}(a^\dagger - a) \\
\phi &= \sqrt{\frac{\hbar Z_c}{2}}(a^\dagger + a),
\end{aligned}
\tag{2.23}
$$

These equations would be used for some following derivations.

Everything we derived from the last section can be applied here in the same fashion. We obtain the quantized Hamiltonian with angular frequency $\omega = 1/\sqrt{LC}$. The inductor and capacitor are macroscopic objects, where piecing them together remarkably makes them behave like a single particle in a harmonic field. This forms the basis of creating an artificial atom, which has atomic properties designed by Physicists and quantum engineers. To make an artificial atom useful for quantum computation, we require an additional circuit element: the Josephson junction.

2.3 Josephson Junction

It is probably not an overstatement that non-linear interaction is what makes the world interesting, and what makes computation and information processing possible. To be more specific in the case of quantum computation, it is impossible to construct a universal quantum computer with only the LC oscillator, since it does not permit conditional operations, a crucial part of realizing universal operations. In quantum circuits, the non-linear elements must also obey the additional requirement of being non-dissipative in order to preserve the coherence of quantum states.

Josephson junction is the only known circuit element that is non-dissipative and strongly non-linear at low temperature [41]. It consists of two superconductors coupled by a weak link, typically consists of a thin insulating barrier. Contrasting to the linear inductor current-flux relation $I = \phi/L$, the Josephson junction has a relation expressed as

$$I = I_0 \sin \frac{2\pi\phi}{\Phi_0} \tag{2.24}$$
$$E = -E_J \cos \frac{2\pi\phi}{\Phi_0},$$

where I_0 is the Josephson critical current, E_J is the Josephson energy, and Φ_0 is the the magnetic flux quantum constant $h/2e$. These relations resemble linear inductor at the first order, seen from performing a Taylor expansion on the $\sin$ and $\cos$ term

$$I = I_0(\frac{2\pi\phi}{\Phi_0} + O(\phi^3)) \tag{2.25}$$
$$E = -E_J(1 - \frac{1}{2}(\frac{2\pi\phi}{\Phi_0})^2 + O(\phi^4)).$$

These equations closely resemble the behavior of an inductor. In fact, the element can be thought as a non-linear inductive element, which its inductance is highly dependent on the excitation present in the circuit. For the next section, we replace the linear inductor with the Josephson junction in our LC circuit, and discuss how this can form a qubit.

2.4 Quantization of Transmon Qubit

The Hamiltonian of a system formed by a capacitor and a Josephson junction can be expressed as

$$H = \frac{q^2}{2C} - E_J \cos \frac{2\pi\phi}{\Phi_0} = 4E_C n^2 - E_J \cos \frac{2\pi\phi}{\Phi_0}, \tag{2.26}$$

where $E_C = e^2/2C$, and n is the number of cooper pairs.

We can re-express the equations in 2.23 to be

$$
\begin{aligned}
n &= i\frac{1}{2e}\sqrt{\frac{\hbar}{2Z_c}}(a^\dagger - a) \\
\phi &= \sqrt{\frac{\hbar Z_c}{2}}(a^\dagger + a).
\end{aligned}
\tag{2.27}
$$

Given $E_c = e^2/2C$, we can express $C = e^2/2E_c$. We begin our analysis by approximating the Josephson junction as a linear inductor. By comparing the second order term of E_J in equation 2.25 with the energy of a linear inductor, one can obtain the relation

$$
E_J\frac{1}{2}(\frac{2\pi\phi}{\Phi_0})^2 = \frac{\phi^2}{2L} \tag{2.28}
$$

Thus,

$$
L = (\frac{\Phi_0}{2\pi})^2\frac{1}{E_J} \tag{2.29}
$$

By substituting $\Phi_0 = h/2e$, we obtain

$$
L = (\frac{h}{2\pi 2e})^2\frac{1}{E_J} = (\frac{\hbar}{2e})^2\frac{1}{E_J} \tag{2.30}
$$

Then,

$$
\begin{aligned}
\frac{L}{C} &= (\frac{\hbar}{2e})^2\frac{1}{E_J}\frac{2E_C}{e^2} \\
&= (\frac{\hbar}{2e^2})^2\frac{2E_C}{E_J} \\
Z_c = \sqrt{\frac{L}{C}} &= \frac{\hbar}{2e^2}\sqrt{\frac{2E_C}{E_J}}
\end{aligned}
\tag{2.31}
$$

Thus we can express the n and ϕ operators as

$$
\begin{aligned}
n &= i\frac{1}{2e}\sqrt{\frac{\hbar}{2Z_c}}(a^\dagger - a) = i\frac{1}{2}(\frac{E_J}{2E_C})^{1/4}(a^\dagger - a) \\
\phi &= \sqrt{\frac{\hbar Z_c}{2}}(a^\dagger + a) = \frac{\hbar}{2e}(\frac{2E_C}{E_J})^{1/4}(a^\dagger + a),
\end{aligned} \tag{2.32}
$$

which directly implies

$$
\frac{2\pi\phi}{\Phi_0} = (\frac{2E_C}{E_J})^{1/4}(a^\dagger + a). \tag{2.33}
$$

Keeping terms up to fourth order for the Josephson energy of 2.26, we obtain the Hamiltonian

$$
H = 4E_C n^2 - E_J + \frac{E_J}{2}(\frac{2\pi\phi}{\Phi_0})^2 - \frac{E_J}{24}(\frac{2\pi\phi}{\Phi_0})^4. \tag{2.34}
$$

This Hamiltonian can be viewed as a harmonic oscillator with a fourth order perturbation inducing anharmonicity. Within this harmonic oscillator approximation, we can substitute n and ϕ by annihilation and creation operators defined in equation 2.32.

$$
\begin{aligned}
H &= 4E_C(i\frac{1}{2}(\frac{E_J}{2E_C})^{1/4}(a^\dagger - a))^2 - E_J + \frac{E_J}{2}((\frac{2E_C}{E_J})^{1/4}(a^\dagger + a))^2 \\
&\quad -\frac{E_J}{24}((\frac{2E_C}{E_J})^{1/4}(a^\dagger + a))^4 \\
&= \sqrt{8E_J E_C}(a^\dagger a + \frac{1}{2}) - E_J - \frac{E_C}{12}(a^\dagger + a)^4
\end{aligned} \tag{2.35}
$$

The first term is recognized as a linear harmonic oscillator with angular frequency $\sqrt{8E_J E_C}$. The second term is simply a constant. The third term gives rise to a fourth order correction to the eigenenergies, given by

$$
E_j^1 = -\frac{E_C}{12}\langle j|\,(a^\dagger + a)^4\,|j\rangle = -\frac{E_C}{12}(6j^2 + 6j + 3). \tag{2.36}
$$

We thus obtain an anharmonicity $\alpha = (E_2 - E_1) - (E_1 - E_0) = -E_C$. In general, $(E_{m+1} - E_m) - (E_m - E_{m-1}) = -E_C$. This forms the basic theoretical description of the transmon qubit.

2.5 Quantization of Split Junction Transmon Qubit with External Flux

From equation 2.35 we learn that the transmon qubit can be understood as a linear harmonic oscillator with angular frequency $\sqrt{8E_J E_C}$, with anharmonicity equal to $-E_C$. The ability to tune E_J would allow one to tune the transmon frequency, which is important to many applications in this thesis. The ability to tune E_J is enabled by splitting the junction (making it as a SQUID), and threading external flux to the inner loop of the SQUID.

Each Josephson junction has a characteristic tunneling energy (E_{J1}, E_{J2}) and superconducting phase (ϕ_1, ϕ_2) across it. The Josephson part of the Hamiltonian is thus

$$H_J = E_{J1} \cos \phi_1 + E_{J2} \cos \phi_2. \tag{2.37}$$

With magnetic flux Φ threading inside the center of SQUID, the phases ϕ_1 and ϕ_2 are related by

$$\frac{2\pi\Phi}{\Phi_0} = \phi_1 - \phi_2 \tag{2.38}$$

Therefore H_J can be expressed as

$$H_J = E_{J1} \cos \phi_1 + E_{J2} \cos(\phi_1 - \frac{2\pi\Phi}{\Phi_0}), \tag{2.39}$$

and thus the full Hamiltonian is

$$H = 4E_C n^2 + E_{J1} \cos \phi_1 + E_{J2} \cos(\phi_1 - \frac{2\pi\Phi}{\Phi_0}). \tag{2.40}$$

The next natural step is to quantize this Hamiltonian with the creation and annihilation operators (a and $a^\dagger$). Recall that the physical origin of the quantization representation of transmon is the

approximation of a harmonic oscillator. In that case, $\phi = 0$ represents the point of minimal energy of the harmonic oscillator, which we subsequently represent ϕ by the creation and annihilation operators. Therefore, we need to find an expression of ϕ which $\phi = 0$ represents the minimal of H_J. To find the minimal point of ϕ_1 of H_J, we can simply differentiate H_J against ϕ_1:

$$\frac{\partial H_J}{\partial \phi_1} = -E_{J1} \sin \phi_1 - E_{J2} \sin(\phi_1 - \frac{2\pi\Phi}{\Phi_0}) \tag{2.41}$$

Setting this derivative to be zero, we obtain

$$\begin{aligned} 0 &= -E_{J1} \sin \bar{\phi} - E_{J2} \sin(\bar{\phi} - \frac{2\pi\Phi}{\Phi_0}) \\ E_{J1} \sin \bar{\phi} &= -E_{J2} \sin \bar{\phi} \cos(\frac{2\pi\Phi}{\Phi_0}) + E_{J2} \cos \bar{\phi} \sin(\frac{2\pi\Phi}{\Phi_0}) \\ [E_{J1} + E_{J2} \cos(\frac{2\pi\Phi}{\Phi_0})] \sin \bar{\phi} &= E_{J2} \sin(\frac{2\pi\Phi}{\Phi_0}) \cos \bar{\phi} \\ \tan \bar{\phi} &= \frac{E_{J2} \sin(\frac{2\pi\Phi}{\Phi_0})}{E_{J1} + E_{J2} \cos(\frac{2\pi\Phi}{\Phi_0})} \end{aligned} \tag{2.42}$$

Thus,

$$\begin{aligned} \sin \bar{\phi} &= \frac{E_{J2} \sin(\frac{2\pi\Phi}{\Phi_0})}{\sqrt{E_{J1}^2 + E_{J2}^2 + 2E_{J1}E_{J2} \cos(\frac{2\pi\Phi}{\Phi_0})}} \\ \cos \bar{\phi} &= \frac{E_{J1} + E_{J2} \cos(\frac{2\pi\Phi}{\Phi_0})}{\sqrt{E_{J1}^2 + E_{J2}^2 + 2E_{J1}E_{J2} \cos(\frac{2\pi\Phi}{\Phi_0})}} \end{aligned} \tag{2.43}$$

We can thus express $\phi_1 = \phi + \bar{\phi}$, and obtain

$$
\begin{aligned}
E_J &= E_{J1}\cos(\phi+\bar{\phi}) + E_{J2}\cos(\phi+\bar{\phi}-\frac{2\pi\Phi}{\Phi_0}) \\
&= E_{J1}[\cos\phi\cos\bar{\phi}-\sin\phi\sin\bar{\phi}] \\
&\quad +E_{J2}[\cos\phi(\cos\bar{\phi}\cos(\frac{2\pi\Phi}{\Phi_0})+\sin\bar{\phi}\sin(\frac{2\pi\Phi}{\Phi_0})) \\
&\quad -E_{J2}[\sin\phi(sin\bar{\phi}\cos(\frac{2\pi\Phi}{\Phi_0})-\cos\bar{\phi}\sin(\frac{2\pi\Phi}{\Phi_0})) \\
&= \sqrt{E_{J1}^2+E_{J2}^2+2E_{J1}E_{J2}\cos(\frac{2\pi\Phi}{\Phi_0})}\cos\phi
\end{aligned}
\tag{2.44}
$$

Therefore, we obtain the Hamiltonian

$$
H = 4E_C n^2 + \sqrt{E_{J1}^2+E_{J2}^2+2E_{J1}E_{J2}\cos(\frac{2\pi\Phi}{\Phi_0})}\cos\phi. \tag{2.45}
$$

We obtain an effective $E_J = \sqrt{E_{J1}^2+E_{J2}^2+2E_{J1}E_{J2}\cos(\frac{2\pi\Phi}{\Phi_0})}$. This results in a transmon frequency $\sqrt{8E_C\sqrt{E_{J1}^2+E_{J2}^2+2E_{J1}E_{J2}\cos(\frac{2\pi\Phi}{\Phi_0})}}$, which is tunable with the external magnetic flux. The anharmonicity remains to be $-E_C$, independent to the external magnetic flux.

2.6 Coupling between Resonators

For two capacitively coupled LC oscillators. We have the Hamiltonian

$$
H = \frac{1}{2}C_1V_1^2 + \frac{1}{2}C_2V_2^2 + \frac{1}{2}C_g(V_2-V_1)^2 + \frac{1}{2}L_1I_1^2 + \frac{1}{2}L_2I_2^2. \tag{2.46}
$$

Substituting the relation $LI=\phi$ and $V=-L\dot{I}=-\phi$, we obtain

$$
H = \frac{1}{2}C_1\dot{\phi_1}^2 + \frac{1}{2}C_2\dot{\phi_2}^2 + \frac{1}{2}C_g(\dot{\phi_2}-\dot{\phi_1})^2 + \frac{\phi_1}{2L_1} + \frac{\phi_2}{2L_2}. \tag{2.47}
$$

Thus, the Lagrangian can be expressed as

$$\mathcal{L} = \frac{1}{2}C_1\dot{\phi_1}^2 + \frac{1}{2}C_2\dot{\phi_2}^2 + \frac{1}{2}C_g(\dot{\phi_2} - \dot{\phi_1})^2 - \frac{\phi_1}{2L_1} - \frac{\phi_2}{2L_2}. \tag{2.48}$$

The kinetic part of the Hamiltonian can be written as

$$T = \frac{1}{2}(C_1 + C_g)\dot{\phi_1}^2 + \frac{1}{2}(C_2 + C_g)\dot{\phi_2}^2 - C_g\dot{\phi_1}\dot{\phi_2}, \tag{2.49}$$

or in matrix form

$$T = \frac{1}{2}\dot{\phi}^T C\dot{\phi} \tag{2.50}$$

$$\phi = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \tag{2.51}$$

$$C = \begin{bmatrix} C_1 + C_g & -C_g \\ -C_g & C_2 + C_g \end{bmatrix} \tag{2.52}$$

The conjugate variable (charge) of ϕ is

$$q = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial \dot{\phi_1}} \\ \frac{\partial \mathcal{L}}{\partial \dot{\phi_2}} \end{bmatrix} = \begin{bmatrix} (C_1 + C_g)\dot{\phi_1} - C_g\dot{\phi_2} \\ -C_g\dot{\phi_1} + (C_2 + C_g)\dot{\phi_2} \end{bmatrix} = C\dot{\phi} \tag{2.53}$$

$$C^{-1}q = \dot{\phi}. \tag{2.54}$$

$$C^{-1} = \frac{1}{C_1C_2 + C_1C_g + C_2C_g}\begin{bmatrix} C_2 + C_g & C_g \\ C_g & C_1 + C_g \end{bmatrix} \tag{2.55}$$

We can thus express the total Hamiltonian as

$$
\begin{aligned}
H &= q\dot{\phi} - \mathcal{L} \\
&= \frac{1}{2}q^T C^{-1} q + \frac{\phi_1}{2L_1} + \frac{\phi_2}{2L_2} \\
&= \frac{q_1^2}{2\bar{C_1}} + \frac{q_2^2}{2\bar{C_2}} + \frac{q_1 q_2}{2\bar{C_g}} + \frac{\phi_1}{2L_1} + \frac{\phi_2}{2L_2}
\end{aligned} \tag{2.56}
$$

$$
\begin{aligned}
\bar{C_1} &= C_1 + \frac{C_2 C_g}{C_2 + C_g} \\
\bar{C_2} &= C_2 + \frac{C_1 C_g}{C_1 + C_g} \\
\bar{C_g} &= \frac{C_1 C_2 + C_1 C_g + C_2 C_g}{C_g}
\end{aligned} \tag{2.57}
$$

By substituting the creation and annihilation operator defined in equation 2.23, we can express the coupling Hamiltonian as

$$
H_g = -\frac{\hbar}{2\sqrt{Z_1 Z_2}\bar{C_g}}(a_1^\dagger - a_1)(a_2^\dagger - a_2), \tag{2.58}
$$

where $Z_i = \sqrt{L_i/\bar{C_i}} = 1/\omega\bar{C_i}$

Thus,

$$
\begin{aligned}
g &= -\frac{1}{2\sqrt{Z_1 Z_2}\bar{C_g}} \\
&= -\sqrt{\omega_1\omega_2}\frac{\sqrt{\bar{C_1}\bar{C_2}}}{2\bar{C_g}} \\
&= -\sqrt{\omega_1\omega_2}\frac{C_g}{2\sqrt{(C_1 + C_g)(C_2 + C_g)}}
\end{aligned} \tag{2.59}
$$

The total Hamiltonian can be written as

$$\hbar\omega_1 a_1^\dagger a_1 + \hbar\omega_2 a_2^\dagger a_2 + \hbar g(a_1^\dagger - a_1)(a_2^\dagger - a_2) \tag{2.60}$$

2.7 Dispersive Shift between Transmon and Resonator

Recall that a Transmon and a resonator has an identical form of the capacitance part of the Hamiltonian. This implies the derivation of coupling between Transmon and oscillator is essentially identical to the last section. Thus, the Hamiltonian of a Transmon and a resonator can be written as

$$H = \hbar\omega_t a_t^\dagger a_t + \frac{\hbar}{2}\alpha_t a_t^\dagger a_t(a_t^\dagger a_t - 1) + \hbar\omega_r a_r^\dagger a_r + \hbar g(a_t^\dagger - a_t)(a_r^\dagger - a_r), \tag{2.61}$$

where $\omega_{[t,r]}$ is the [Transmon, resonator] frequency and α_t is the transmon anharmonicity. One can easily verify this form of Hamiltonian has a frequency ω_t with anharmonicity α. With rotating wave approximation, we omit the coupling terms $a_t^\dagger a_r^\dagger$ and $a_t a_r$.

$$H = \hbar\omega_t a_t^\dagger a_t + \frac{\hbar}{2}\alpha_t a_t^\dagger a_t(a_t^\dagger a_t - 1) + \hbar\omega_r a_r^\dagger a_r - \hbar g(a_t^\dagger a_r + a_t a_r^\dagger), \tag{2.62}$$

Let $|m, n\rangle$ be the state with m transmon excitations and n resonator excitations. The first-order perturbation of eigenenergies from $g(a_t^\dagger a_r + a_t a_r^\dagger)$ is zero, because $\langle m, n|\, g(a_t^\dagger a_r + a_t a_r^\dagger)\, |m, n\rangle = 0$. The second-order perturbation is

$$\begin{aligned} & \frac{|\,\langle m-1, n+1|\, g(a_t a_r^\dagger)\, |m, n\rangle\,|^2}{E_{m,n} - E_{m-1,n+1}} + \frac{|\,\langle m+1, n-1|\, g(a_t^\dagger a_r)\, |m, n\rangle\,|^2}{E_{m,n} - E_{m+1,n-1}} \\ = \;& \frac{g^2(m(n+1))}{\omega_t + (m-1)\alpha - \omega_r} - \frac{g^2((m+1)n)}{\omega_t + m\alpha - \omega_r} \end{aligned} \tag{2.63}$$

Thus, the eigenenergies of $|0, n\rangle$ are

$$n\omega_r - \frac{g^2 n}{\omega_t - \omega_r} = n(\omega_r - \frac{g^2}{\Delta}), \tag{2.64}$$

where $\Delta = \omega_t - \omega_r$, and the eigenenergies of $|1, n\rangle$ are

$$\begin{aligned} & \omega_t + n\omega_r + \frac{g^2(n+1)}{\omega_t - \omega_r} - \frac{g^2 2n}{\omega_t + \alpha - \omega_r} \\ = \; & (\omega_t + \frac{g^2}{\omega_t - \omega_r}) + n\omega_r + \frac{g^2 n}{\omega_t - \omega_r} - \frac{g^2 2n}{\omega_t + \alpha - \omega_r} \\ = \; & (\omega_t + \frac{g^2}{\Delta}) + n\omega_r + \frac{g^2 n}{\Delta} - \frac{g^2 2n}{\Delta + \alpha} \\ = \; & (\omega_t + \frac{g^2}{\Delta}) + n(\omega_r - \frac{g^2}{\Delta}(\frac{2}{1 + \alpha/\Delta} - 1)), \end{aligned} \tag{2.65}$$

where $(\omega_t + g^2/\Delta)$ is the eigenfrequency of $|1, 0\rangle$. Comparing the two above equations, we obtain

$$\begin{aligned} |1, n\rangle - |1, 0\rangle - |0, n\rangle &= n(-\frac{g^2}{\Delta}(\frac{2}{1 + \alpha/\Delta} - 2)) \\ &= n(\frac{g^2}{\Delta}\frac{2\alpha/\Delta}{1 + \alpha/\Delta}) \\ &= n(\frac{g^2}{\Delta^2}\frac{2\alpha}{1 + \alpha/\Delta}). \end{aligned} \tag{2.66}$$

The dispersive shift is thus

$$\chi = \frac{g^2}{\Delta^2}\frac{\alpha}{1 + \alpha/\Delta}. \tag{2.67}$$

As expected, the dispersive shift tends to zero as α tends to zero, matching our expectation for two coupled linear oscillators. For the ground state and first excited state of the transmon, we can write the Hamiltonian as

$$H = \hbar(\omega_r + \chi\sigma_z)a_r^\dagger a + \frac{1}{2}\hbar\omega_t\sigma_z \tag{2.68}$$

CHAPTER 3

CONTROLLING SUPERCONDUCTING QUBITS

This chapter describes how the quantum systems described earlier can be controlled and measured by quantum computing experimentalists or engineers. Next, we describe the importance of a cryogenic environment in implementing superconducting qubits. We begin with the discussion on the theory of charge and flux controls and their applications. We then describe the necessary microwave components (filters, amplifiers, etc) that allow us to operate of the quantum system. We next talk about the room temperature electronics we used for the control operations. The final chapter puts all these components together and shows a complete setup for operating a quantum computer.

3.1 Charge Control

The charge control is responsible for driving qubit excitations. This is achieved by coupling a voltage source $V_d(t)$ with the Transmon qubit by a capacitor C_d. The transmon Hamiltonian with the voltage source can be written as

$$H = \frac{1}{2}CV^2 + \frac{1}{2}C_d(V_d(t) - V)^2 + \frac{LI^2}{2}. \tag{3.1}$$

Substituting the relation $LI = \phi$ and $V = -L\dot{I} = -\dot{\phi}$, the transmon Lagrangian can be written as

$$\mathcal{L} = \frac{1}{2}C\dot{\phi}^2 + \frac{1}{2}C_d(V_d(t) - \dot{\phi})^2 - \frac{\phi^2}{2L} \tag{3.2}$$

The conjugate variable of ϕ can be obtained by

$$q = \frac{\mathcal{L}}{\partial\dot{\phi}} = C_\Sigma\dot{\phi} - C_dV_d(t), \tag{3.3}$$

where $C_\Sigma = C + C_d$. Thus,

$$\dot{\phi} = \frac{1}{C_\Sigma}(q + C_d V_d(t)) \tag{3.4}$$

By expressing the Hamiltonian and a function of q and ϕ, we get

$$\begin{aligned} H &= q\dot{\phi} - \mathcal{L} \\ &= \frac{q^2}{2C_\Sigma} + \frac{C_d V_d(t) q}{C_\Sigma} + \frac{C_d^2 V_d(t)^2}{2C_\Sigma} + \frac{\phi^2}{2L} \end{aligned} \tag{3.5}$$

We can then quantize the conjugate operators with creation and annihilation operators according to equation 2.23. We discard the last term since it does not contain any conjugate variable. For a Transmon, we have

$$H = \hbar\omega a^\dagger a + \frac{\hbar}{2}\alpha a^\dagger a(a^\dagger a - 1) + i\frac{C_d V_d(t)}{C_\Sigma}\sqrt{\frac{\hbar}{2Z_c}}(a^\dagger - a) \tag{3.6}$$

The prefactor of the time-dependent driving term can be represented by $\Omega(t)$, thus the full Hamiltonian is

$$H = \hbar\omega a^\dagger a + \frac{\hbar}{2}\alpha a^\dagger a(a^\dagger a - 1) + i\Omega(t)(a^\dagger - a) \tag{3.7}$$

In the number basis of transmon state, one can understood the term $a - a^\dagger$ to be coupling the states $|n\rangle$ with $|n-1\rangle$ and $|n+1\rangle$. By targeting the frequency of $\Omega(t)$ with the corresponding transition energy between E_n and E_{n-1} or E_{n+1}, the coupling term $i\Omega(t)(a^\dagger - a)$ can induce Rabi oscillations between these states. By controlling the strength and coupling time of $\Omega(t)$, one can control the degree of Rabi oscillations. This forms the basis of single qubit gate operations.

3.2 Flux Control

From section 2.5 we found that for a split junction Transmon, E_J can be tunable with external time-dependent magnetic flux $\Phi(t)$ in the SQUID loop, with the effective E_J equals to

$$\sqrt{E_{J1}^2 + E_{J2}^2 + 2E_{J1}E_{J2}\cos(\frac{2\pi\Phi(t)}{\Phi_0})} \tag{3.8}$$

The mechanism for applying magnetic flux is rather straightforward. We apply current to an inductive element that has finite mutual inductance coupling with the SQUID loop. A time-dependent magnetic flux in the SQUID loop can be obtained by varying the applied current. The Transmon Hamiltonian can thus be written as

$$H = \sqrt{8E_C\sqrt{E_{J1}^2 + E_{J2}^2 + 2E_{J1}E_{J2}\cos(\frac{2\pi\Phi(t)}{\Phi_0})}}(a^\dagger a + \frac{1}{2}) + \frac{\hbar}{2}\alpha a^\dagger a(a^\dagger a - 1) \tag{3.9}$$

For simplicity, one can write the prefractor of $a^\dagger a + \frac{1}{2}$ to be $\hbar\omega(t)$. The Hamiltonian is thus

$$H = \hbar\omega(t)(a^\dagger a + \frac{1}{2}) + \frac{\hbar}{2}\alpha a^\dagger a(a^\dagger a - 1) \tag{3.10}$$

This forms the basis of flux control, which allows us to tune the Transmon frequency at will. This ability enables one to bring qubits on- and off-resonance at will, which we leverage to form two-qubit gates. We use the flux control in this thesis extensively for two-qubit gate operations.

3.3 Transmon State Measurement

From the section 2.7 we conclude that the resonator frequency is dependent on the transmon state. This is basically the key to transmon state measurement. One can measure the transmon state by coupling the transmon with a linear LC resonator. A change of transmon state results in a shift of resonator frequency, which then results in a shift of transmission frequency and phase of a

microwave pulse transmitting through the resonator.

3.4 Cryogenic Environment

Superconducting qubits require an operating environment with cryogenic temperature. This is mainly due to two reasons. First, most superconductors have a very low critical temperature. For example, aluminum and niobium have critical temperature of 1.2K and 9.3K respectively. The superconductivity (zero-resistance) property is critical for a highly coherent circuit. Any introduction of unwanted resistive elements would damage the qubit coherence. Second, the suppression of thermal noise is crucial to make the qubits well behaved. Thermal noise randomly excites the qubit in a stochastic fashion, which highly impedes the execution of a target control sequence.

As an estimation, a typical superconducting qubit has a frequency of $8\,\text{GHz}$. This corresponds to a thermal temperature of

$$\frac{h \times 8\,\text{GHz}}{k_B} \approx 384\,\text{mK}, \tag{3.11}$$

where h is the Planck constant and k_B is the Boltzmann constant. Ideally, one would place the qubit in an environment that is an order of magnitude colder than 384mK. A dilution fridge is able to cooldown to a temperature down to $10 \sim 20$ mK, which is perfect for our application.

3.5 Microwave Control Electronics

3.5.1 Oscillator

An oscillator is used for outputting a set frequency and amplitude microwave signal. The switching times of these settings are typically rather long (much longer than qubit coherence time). Therefore, the oscillators are used as a static source of microwave signal. Most generators support pulsing, meaning the on/off of signal output can be controlled by an external control signal. The control signal can be provided by an arbitrary waveform generator.

3.5.2 Arbitrary waveform generator

The arbitrary waveform generator is a generator that allows the output of programmable voltage or current waveforms. These waveforms are not totally arbitrary. Typically, there are constraints in amplitude, bandwidth, output resolution and time precision of the waveforms.

Some of the most expensive commercial arbitrary waveform generators have high enough clock rate which allows direct synthesis of microwave pulses at the qubit frequencies (4 - 8 GHz). The Schuster lab has resources to purchase a few of these, which I am fortunate to have first-hand experience using these state-of-the-art electronics. For instance, the Keysight M8195A is able to generate 16GSa/s waveform signals for 4 channels, which was perfect for some of our multi-qubit experiments.

For a lower clock rate arbitrary waveform generator, it would require frequency up-conversion to generate pulses at qubit frequencies. To be specific, one would use IQ mixers for the frequency up-conversion, which we will describe in the next section.

3.5.3 IQ mixer

The IQ mixer is a 4-port device (L,R,I,Q). The L-port typically requires an input of microwave signal from an oscillator at a fixed power ($\sim 13\,\mathrm{dBm}$). Given an input frequency ω_l of the L-port. The signal of (R,I,Q) ports are related by

$$R(t) = I(t)\cos(\omega_l t) + Q(t)\sin(\omega_l t) \tag{3.12}$$

Remarkably, the R and I,Q ports can be used as input or output interchangeably. To use the R-port as output, the I- and Q-ports are used as inputs, which the signals are generated by arbitrary waveform generators. The $R(t)$ have an envelope defined by $I(t)$ and $Q(t)$, with oscillation frequency equals to ω_l. In this case, the IQ mixer is used as a frequency up-converter. On the other hand, the R-port can be used as input and the I- and Q-ports become outputs. The signals of $I(t)$ and $Q(t)$ are defined by the same equation, which $[I(t), Q(t)]$ is the $[\sin, \cos]$ quadrature of $R(t)$.

In this case, the IQ mixer is used as a frequency down-converter, which is useful for qubit state measurement.

3.6 Other Electronics

An analog-to-digital converter (ADC) is a system that converts an analog signal into a digital signal. It is used for qubit state measurement, where we take the down-converted measurement signal to ADC and use a computer to analyze the digital signals.

We use a DC-voltage source to tune the frequency of Transmon qubits, discussed in 2.5. The stability of the DC-voltage source is crucial for high qubit coherence since any voltage noise would result in qubit frequency noise. We also require the voltage source to be high precision and finely tunable, such that we can tune qubit frequencies with high preciseness.

3.7 Microwave Components

3.7.1 Attenuator

While the quantum device operates in cryogenic temperature, most of our control electronics operate in room temperature. It is thus unavoidable to bring in room temperature signals, which contains a large amount of thermal noise. The attenuator, usually made of resistive (dissipative) components, is thus employed to attenuate the thermal noise. A dilution fridge is multi-stage, and typically each stage operating in temperatures 50K , 4K, 700mK, and 10mK. Attenuators are typically placed in stages 4K, 700mK and 10mK to attenuate the thermal noise successively. Of course, the signals would also be attenuated. Thus, the input signals at room temperature are set to high amplitudes, and both signals and noises are attenuated as they propagate to the colder stage of the fridge.

3.7.2 Filter

Filters are usually made of inductive and capacitive (non-dissipative) components, resulting in different transmission amplitudes for a different range of frequencies. In most experiments, we have a well-expected frequency range of signals for each control line. We can thus use the filters to filter out any transmission from unwanted frequencies, thus lowering the total thermal noise reaching the quantum device.

3.7.3 Isolator

A circulator is a passive non-reciprocal three-port device, in which a microwave or radio frequency signal entering any port is transmitted to the next port in rotation. When one port of a three-port circulator is terminated in a matched load, it can be used as an isolator, since a signal can travel in only one direction between the remaining ports. The isolator is especially useful for the measurement output channel, where signals are traveling from the quantum device towards room temperature electronics. The isolator ensures that a minimal thermal noise from room temperature can be conducted through the measurement channel.

3.7.4 Amplifier

The measurement signals at the quantum device are very weak and would be overwhelmed by thermal noise if the signals are not amplified. Therefore it is critical to amplify the measurement signals within the cryogenic environment, before reaching room temperatures. Special amplifiers such as parametric amplifier, and High Electron Mobility Transistor (HEMT) are typically employed for amplification within the fridge. A proper amplification is crucial for obtaining high measurement fidelity.

3.8 Software For Quantum Operations and Measurements

From our experience, having a well-structured software is extremely important for high productivity in performing physical experiments. I spent a good portion of time on building lab software framework for executing experiments (which I really enjoyed). That involves writing drivers for individual hardware, integration of various hardware components, measurement data analysis and code instruction framework for generating physical waveforms. This chapter introduces the basic components of software programming for controlling a quantum computer.

3.8.1 Hardware Drivers

Most programmable electronic hardware has a simple programming interface called The Standard Commands for Programmable Instruments (SCPI), which can commonly be communicated through LANs or PCIs. Most operations can be set with a single line command, such as setting the amplitude or frequency of an oscillator or setting the output voltage of a DC-voltage source. Defining the output of an arbitrary waveform generator involves more sophisticated programming, which most of the following sections in this chapter will be dedicated to this topic.

3.8.2 Essential properties

As seen from section 3.1, the voltage $V(t)$ for controlling a qubit typical involves an oscillation microwave pulse at some given frequency, amplitude and duration. Controlling a quantum computer is essentially sending streams of microwave pulses into the system. We require a software framework that has the following properties:

- Discrete gate abstraction: When performing gate sequence programming, one would rather think and program in discrete gate set, and not prefer to think in terms of continuous streams of microwave pulses.

- Flexibility in defining individual gate: Most of the numerically optimized pulses have a non-analytic waveform. Thus the software is required to support the generation or arbitrarily

defined waveforms.

- Incorporating generators with different requirements on waveforms: Most generator has specifications on the sample size of waveforms, such as a minimum sample size and the sample size must be an integer multiple of a certain value (e.g. $2^4 = 16$).

- Incorporating generators at different clock rate: We typically have to group various models of generators to perform an experiment. To synchronize the output of each generator, it would be convenient to only think in terms of time, instead of dealing with the different clock cycle of each generator.

- The software programming of each experiment should be agnostic to the underlying hardware: This is closely related to the two above items. It is common to use different waveform generators for a similar set of experiments. For example, almost every qubit related experiment would require implementing the Rabi, Ramsey and T_1 experiments. It would be highly inefficient if one has to reprogram their experiments when one is given a different set of hardware.

These are the guiding principles I followed when programming our lab's quantum control software, and so far it has been successful in applying the same software to various experimental projects in the lab. Looking forward, the programming of a system with feedback operation would introduce the next level of complexity, and would likely need some careful design on how one should write software for that.

3.9 Putting all together: Cryogenic setup and control instrumentation

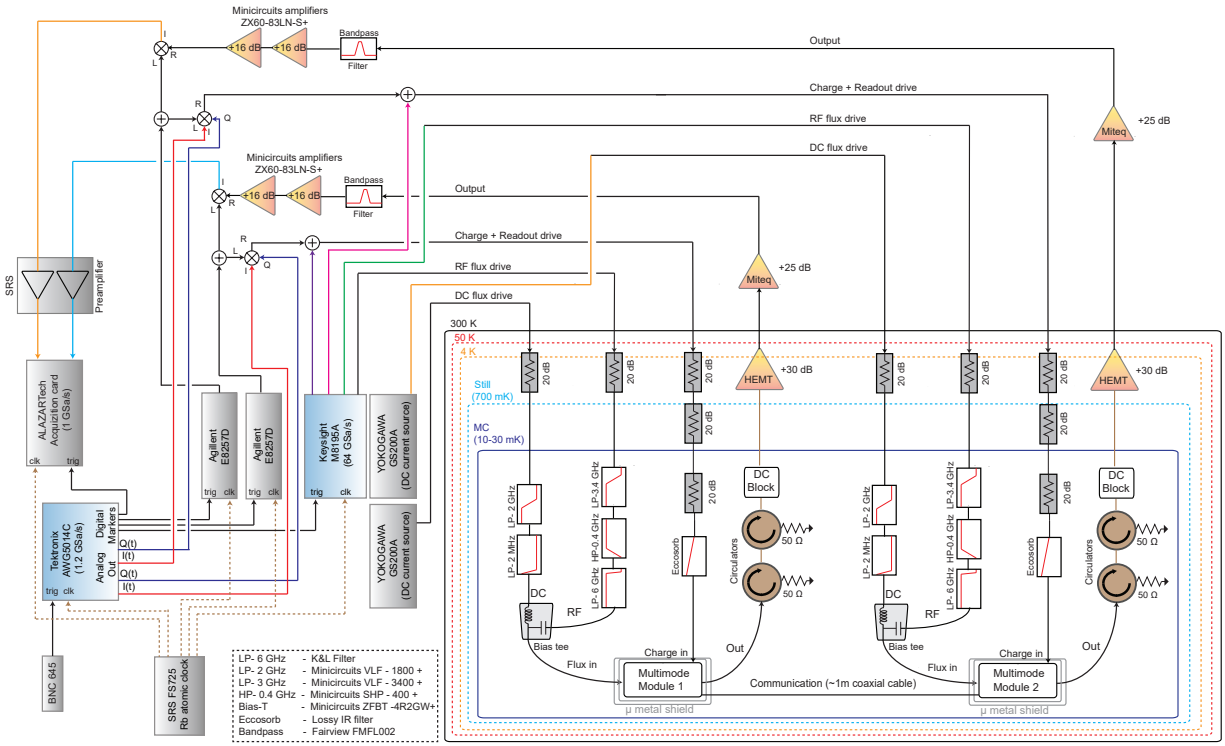

Figure 3.1: Detailed schematic of the cryogenic setup, control instrumentation, and the wiring of microwave and DC connections to the device.

Here we put all the components together and describes an experimental setup we used for conducting quantum computations. The quantum device is heat sunk via an OFHC copper post to the base stage of a Bluefors dilution refrigerator (10-30 mK). The sample is surrounded by a can containing two layers of μ-metal shielding and a layer of lead shielding, thermally anchored using an inner close fit copper shim sheet, attached to the copper can lid. The schematic of the cryogenic setup, control instrumentation, and the wiring of the device is shown in Figure 3.1. Each device is connected to the rest of the setup through three ports: a charge port that applies qubit and readout drive tones, a flux port for shifting the qubit frequency using a DC-flux bias current and for applying RF sideband flux pulses, and an output port for measuring the transmission from the readout resonator. The readout pulses are generated by mixing a local oscillator tone (generated

from an Agilent 8257D RF signal generator), with pulses generated by a Tektronix AWG5014C arbitrary waveform generator (TEK) with a sampling rate of 1.2 GSa/s, using an IQ-Mixer (MARQI MLIQ0218). The charge drive pulses are generated with Keysight M8195A arbitrary waveform generator by direct synthesis, and subsequently combined with the readout drive pulse. The combined signals are sent to the device, after being attenuated a total of 60 dB in the dilution fridge, using attenuators thermalized to the 4K (20 dB), still (20 dB) and base stages (20 dB). The charge drive line also includes a lossy ECCOSORB CR-117 filter to block IR radiation, and a low-pass filter with a sharp roll-off at 6 GHz, both thermalized to the base stage. The flux-modulation pulses are also directly synthesized by the Keysight M8195A arbitrary waveform generator and attenuated by 20 dB at the 4 K stage, and bandpass filtered to within a band of 400 MHz - 3.4 GHz at the base stage, using the filters indicated in the schematic. The DC flux bias current is generated by a YOKOGAWA GS200 low-noise current source, attenuated by 20 dB at the 4 K stage, and low-pass filtered down to a bandwidth of 1.9 MHz. The DC flux bias current is combined with the flux-modulation pulses at a bias tee thermalized at the base stage. The state of the transmon is measured using the transmission of the readout resonator, through the dispersive circuit QED readout scheme [195]. The transmitted signal from the readout resonator is passed through a set of cryogenic circulators (thermalized at the base stage) and amplified using a HEMT amplifier (thermalized at the 4 K stage). Once out of the fridge, the signal is filtered (narrow bandpass filter around the readout frequency) and further amplified. The amplitude and phase of the resonator transmission signal are obtained through a heterodyne measurement, with the transmitted signal demodulated using an IQ mixer and a local oscillator at the readout resonator frequency. The heterodyne signal is amplified (SRS preamplifier) and recorded using a fast ADC card (ALAZARtech).

CHAPTER 4

QUANTUM OPTIMAL CONTROL

This chapter describes how numerical techniques can be employed to optimize the control efficiency of a quantum computer. We implement a quantum optimal control algorithm based on automatic differentiation and harness the acceleration afforded by graphics processing units (GPUs). Automatic differentiation allows us to specify advanced optimization criteria and incorporate them in the optimization process with ease. We show that the use of GPUs can speedup calculations by more than an order of magnitude. Our strategy facilitates efficient numerical simulations on affordable desktop computers, and exploration of a host of optimization constraints and system parameters relevant to real-life experiments. We demonstrate optimization of quantum evolution based on fine-grained evaluation of performance at each intermediate time step, thus enabling more intricate control on the evolution path, suppression of departures from the truncated model subspace, as well as minimization of the physical time needed to perform high-fidelity state preparation and unitary gates. The content of this chapter is based on our work "Speedup for quantum optimal control from GPU-based automatic differentiation" published in *Physical Review A* [105].

4.1 Introduction

The techniques and algorithms used to optimize the control of quantum systems [63, 89, 38, 174, 53, 66, 138, 159, 146, 184, 205, 206, 144, 114, 143, 17, 46] and those underlying the field of deep neural networks [72, 73] share a number of common elements. Both areas heavily use linear algebra operations combined with gradient descent optimization. Thus, advanced hardware and software technology recently emerging from the rapid development of machine learning also paves the way for a significant boost of optimal quantum control techniques.

A crucial factor for recent impressive progress in machine learning has been the leveraging of massive parallelism native to graphics processing units (GPUs) [142, 23, 167, 151, 179]. Similarly GPUs have been used to accelerate computations in many areas of quantum physics and chem-

istry [14, 34, 202, 188, 194, 145, 76]. Specifically, GPUs are extremely efficient in multiplying very large matrices [37, 56]. Such multiplications also form a central step in the simulation and optimal control of quantum systems. Exploiting this advantageous feature of GPUs, we achieve significant speed improvements in optimizing control schemes for systems at the current frontiers of experimental quantum computation. As the number of qubits in these experiments is increasing [36, 39, 87], it becomes increasingly important to take advantage of optimal control techniques. Moreover, recent advances in commercially available electronics – e.g., arbitrary waveform generators enabling base-band synthesis of the entire microwave spectrum [88] – afford new capabilities which quantum optimal control is uniquely well-suited to harness.

There have been numerous theoretical developments of numerical and analytical methods for quantum optimal control (see Ref. [63] for a recent review). The algorithms involved are predominantly based on gradient methods, such as realized in gradient ascent pulse engineering (GRAPE) [89, 38], Krotov algorithms [174, 53, 66, 138, 159, 146, 184] or rapid monotonically convergent algorithms [205, 206, 144, 114, 143, 17, 46], and are available in several open-source packages, including QuTiP [83, 84], DYNAMO [112], Spinach [75], and SIMPSON [187]. Quantum optimal control has been remarkably successful in determining optimized pulse sequences [16], designing high-fidelity quantum gates [176, 157, 138, 135, 86, 113, 51, 106, 156, 50, 98], and preparing entangled states [47, 150, 198, 64, 65].

Optimal control is a versatile concept which can be applied to a vast variety of quantum systems. Typically there is a primary goal (e.g. maximizing fidelity to a target state/unitary), as well as additional constraints/costs associated with specific experimental systems. Examples of such constraints include fixed maximum amplitudes of control pulses [173, 94], maximum overall power of control signals [95], and limited time resolution of arbitrary waveform generators [132]. Further, finite coherence of quantum systems motivates minimizing the overall time needed for reaching the intended state or unitary (time-optimal control) [24]. In certain cases, steering the quantum system among an optimal path (time-dependent target) may be desired [166]. Incorporating new constraints in the optimization process often requires the analytical derivation and implementa-

tion of additional contributions to the gradient calculation, and may necessitate significant effort to deploy on large computer clusters. This issue can greatly impede the ability to quickly develop control strategies for new problems.

To overcome these obstacles, we have implemented a quantum optimal control scheme that incorporates constraints via automatic differentiation [7, 199] and utilizes GPUs for boosting computational efficiency. Specifically, automatic differentiation handles the updating of gradient calculations in the backward propagation algorithm [73], and thus eliminates the need to hard-code additional gradient contributions from constraints. For the actual optimal control applications we present in this paper, we find that the computational speed-up from utilizing GPUs becomes significant for Hilbert space sizes exceeding dimensions of the order of one hundred, see Fig. 4.3. Together, these features allow a quick turnaround for varying optimization constraints and system parameters, rendering this approach invaluable for the study of quantum optimal control. In this paper, we describe the implementation of automatic differentiation, demonstrate its application to quantum optimal control of example systems relevant to quantum computing and quantum optics, and discuss the performance gains achieved by utilizing GPUs.

4.2 Theory

We briefly review the essential idea of quantum optimal control and introduce the notation used throughout our paper. We consider the general setting of a quantum system with intrinsic Hamiltonian $\mathcal{H}_0$ and a set of external control fields $\{u_1(t), \ldots, u_M(t)\}$ acting on the system via control operators $\{\mathcal{H}_1, \ldots, \mathcal{H}_M\}$. The resulting system Hamiltonian is given by $H(t) = \mathcal{H}_0 + \sum_{k=1}^{M} u_k(t)\mathcal{H}_k$. Optimal control theory aims to minimize deviations from a target state or target unitary by appropriate adjustments of the control fields $u_k(t)$. To implement this optimization, the time interval of interest is discretized into a large number N of sufficiently small time steps δt.

Denoting intermediate times by $t_j = t_0 + j\,\delta t$, the Hamiltonian at time t_j takes on the form

$$H_j = \mathcal{H}_0 + \sum_{k=1}^{M} u_{k,j}\mathcal{H}_k. \tag{4.1}$$

The control fields subject to optimization now form a set $\{u_{k,j}\}$ of $d = M \cdot N$ real numbers.

The quantum evolution from the initial time $t = t_0$ to time t_j is described by a propagator K_j, decomposed according to

$$K_j = U_j U_{j-1} U_{j-2} \ldots U_1 U_0 \tag{4.2}$$

where

$$U_j = \exp(-iH_j\delta t) \tag{4.3}$$

is the propagator for the short time interval $[t_j, t_j + \delta t]$. (Here and in the following, we set $\hbar = 1$.) Evolution of a select initial state $|\Psi_0\rangle$ from $t = t_0$ to $t = t_j$ then takes the usual form

$$|\Psi_j\rangle = K_j|\Psi_0\rangle. \tag{4.4}$$

In the decomposition of K_j, each short-time propagator U_i can be evaluated exactly by matrix exponentiation or approximated by an appropriate series expansion. Propagation methods which go beyond the piecewise-constant approximation for the propagation, can further improve speed and accuracy [113].

Optimization of the discretized control fields $\mathbf{u} \in \mathbb{R}^d$ can be formulated as the minimization of a cost function $C(\mathbf{u})$ where $C : \mathbb{R}^d \to \mathbb{R}^+$. Table 4.1 shows some of the most important cost function contributions used for quantum optimal control. The total cost function is a linear combination of these cost functions, $C = \sum_\mu \alpha_\mu C_\mu$. The weight factors α_μ must be determined empirically, and depend on the specific problem and experimental realization at hand. In the following, we discuss these relevant cost function contributions.

4.2.1 *Important types of cost function contributions*

The first cost contribution, $C_1(\mathbf{u})$, is the primary tool for realizing a target unitary K_T, such as a single or multi-qubit gate. Cost is incurred for deviations between the target unitary and the realized unitary K_N at a given final time t_N. For a system with Hilbert space dimension D, its expression $1 - |\operatorname{tr}(K_T^\dagger K_N)/D|^2$ [89] represents the infidelity obtained from the trace distance between the target unitary and the realized unitary. Minimizing this cost function is the principle goal of the quantum control problem.

The second cost function, $C_2(\mathbf{u}) = 1 - |\langle\Psi_T|\Psi_N\rangle|^2$ measures the distance between a desired target state $|\Psi_T\rangle$ and the state $|\Psi_N\rangle$ realized at the final time t_N, as obtained from evolution of a given initial state $|\Psi_0\rangle$. In addition, generalizing C_2 to multiple initial and target states is useful for performing a unitary K_T which is only defined on some subspace $H_\mathcal{S}$ of the modeled Hilbert space. Such restriction to a selected subspace is of practical importance whenever a desired unitary is to be implemented within some computational subspace only, as is common for quantum computation applications. There, evolution of higher excited states or auxiliary systems outside the computational subspace is immaterial. Optimal control, then, can be achieved by simultaneous evolution of a set of initial states $\{|\Psi_0^s\rangle\}$ ($s = 1, 2, \ldots, S$) that forms a basis of $H_\mathcal{S}$. Optimal control fields are obtained from minimizing the composite state infidelity $C_{2\Sigma}(\mathbf{u}) = 1 - |\frac{1}{S}\sum_s \langle\Psi_T^s|P_\mathcal{S}|\Psi_N^s\rangle|^2$ relative to the desired target states $|\Psi_T^s\rangle = K_T|\Psi_0^s\rangle$. (Here, $P_\mathcal{S}$ is the projector onto subspace $H_\mathcal{S}$.)

This composite state-transfer cost function when used over a complete basis is equivalent to the gate fidelity, but has several advantages. Most importantly it is more memory efficient requiring only the current state to be stored rather than the whole unitary. In addition, it is very amenable to distributed computing approaches. However, when the unitary transfer matrix can be stored in memory, propagating the full unitary can take advantage of the parallelism of the GPU for smaller problems (see Fig. 4.3).

Like many optimization problems, quantum optimal control is typically underconstrained. In order to obtain control fields that are consistent with specific experimental capabilities and limitations, it is often crucial to add further constraints on the optimization. Control fields must be

μ	**Cost function contribution**	$C_\mu(\mathbf{u})$
1	Target gate infidelity	$1 - \lvert \operatorname{tr}(K_T^\dagger K_N)/D \rvert^2$
2	Target state infidelity	$1 - \lvert \langle \Psi_T \vert \Psi_N \rangle \rvert^2$
3	Control amplitudes	$\lvert \mathbf{u} \rvert^2$
4	Control variations	$\sum_{j,k} \lvert u_{k,j} - u_{k,j-1} \rvert^2$
5	Occupation of forbidden state	$\sum_j \lvert \langle \Psi_F \vert \Psi_j \rangle \rvert^2$
6	Evolution time (target gate)	$1 - \frac{1}{N} \sum_j \lvert \operatorname{tr}(K_T^\dagger K_j)/D \rvert^2$
7	Evolution time (target state)	$1 - \frac{1}{N} \sum_j \lvert \langle \Psi_T \vert \Psi_j \rangle \rvert^2$

Table 4.1: Relevant contributions to cost functions for quantum optimal control. Names of contributions indicate the quantity to be *minimized*.

realizable in the lab, should be robust to noise, and avoid large control amplitudes and rapid variations based on signal output specifications of instruments employed in experiments. Exceedingly strong control fields may also be problematic due to heat dissipation which may, for instance, raise the temperature inside a dilution refrigerator. These points motivate the consideration of additional cost function contributions in the following.

One such contribution, $C_3(\mathbf{u}) = |\mathbf{u}|^2$ suppresses large control-field amplitudes globally, and is commonly employed in quantum optimal control studies [89, 173, 94, 74]. (The generalization to more fine-grained suppression of individual control fields is straightforward to implement as well.) Penalizing the L^2 norm of the control fields favors solutions with low amplitudes. It also tends to spread relevant control fields over the entire allowed time window. While C_3 constitutes a "soft" penalty on control-field amplitudes, one may also apply a trigonometric mapping to the amplitudes to effect a hard constraint strictly enforcing fixed maximum amplitudes [55].

The fourth type of contribution to the cost function, $C_4(\mathbf{u}) = \sum_{j,k} |u_{k,j} - u_{k,j-1}|^2$, penalizes rapid variations of control fields by suppressing their (discretized) time derivatives [74]. The resulting smoothening of signals is of paramount practical importance, since any instrument generating a control field has a finite impulse response. If needed, contributions analogous to C_4 which suppress higher derivatives or other aspects of the time dependence of fields can be constructed. Together, limiting the control amplitudes and their time variation filters out high-frequency "noise" from control fields, which is an otherwise common result of less-constrained optimization. Smoother

control fields also have the advantage that essential control patterns can potentially be recognized and given a meaningful interpretation.

The contribution $C_5(\mathbf{u}) = \sum_j |\langle \Psi_F | \Psi_j \rangle|^2$ to the cost function has the effect of suppressing occupation of a select "forbidden" state $|\Psi_F\rangle$ (or a set of such states, upon summation) throughout the evolution. The inclusion of this contribution addresses an important issue ubiquitous for systems with Hilbert spaces of large or infinite dimension. In this situation, truncation of Hilbert space is needed or inevitable due to computer memory limitations. (Note that this need even arises for a single harmonic oscillator.) Whenever the evolution generated by optimal control algorithms explores highly excited states, truncation introduces a false non-linearity which can misguide the optimization. Including additional states can, in principle, mitigate this problem, but is generally computationally very expensive. An independent physics motivation for avoiding occupation of highly-excited states consists of spontaneous relaxation in realistic systems: high-energy states are often more lossy (as is usually the case, e.g., for superconducting qubits), and possibly more difficult to model. Active penalization of such states therefore has the two-fold benefit of keeping Hilbert space size at bay, and reducing unwanted fidelity loss from increased relaxation. To address these challenges, we employ an intermediate-time cost function [166, 147]: the cost function C_5 limits leakage to higher states during the entire evolution, and at the same time prevents optimization to be misinformed by artificial non-linearity due to truncation. We note that the efficacy of this strategy is system dependent: it works well, for example, for harmonic oscillators or transmon qubits [96] which have strong selection rules against direct transitions to more distant states, but may be less effective in systems such as the fluxonium circuit [115] where low-lying states have direct matrix elements to many higher states.

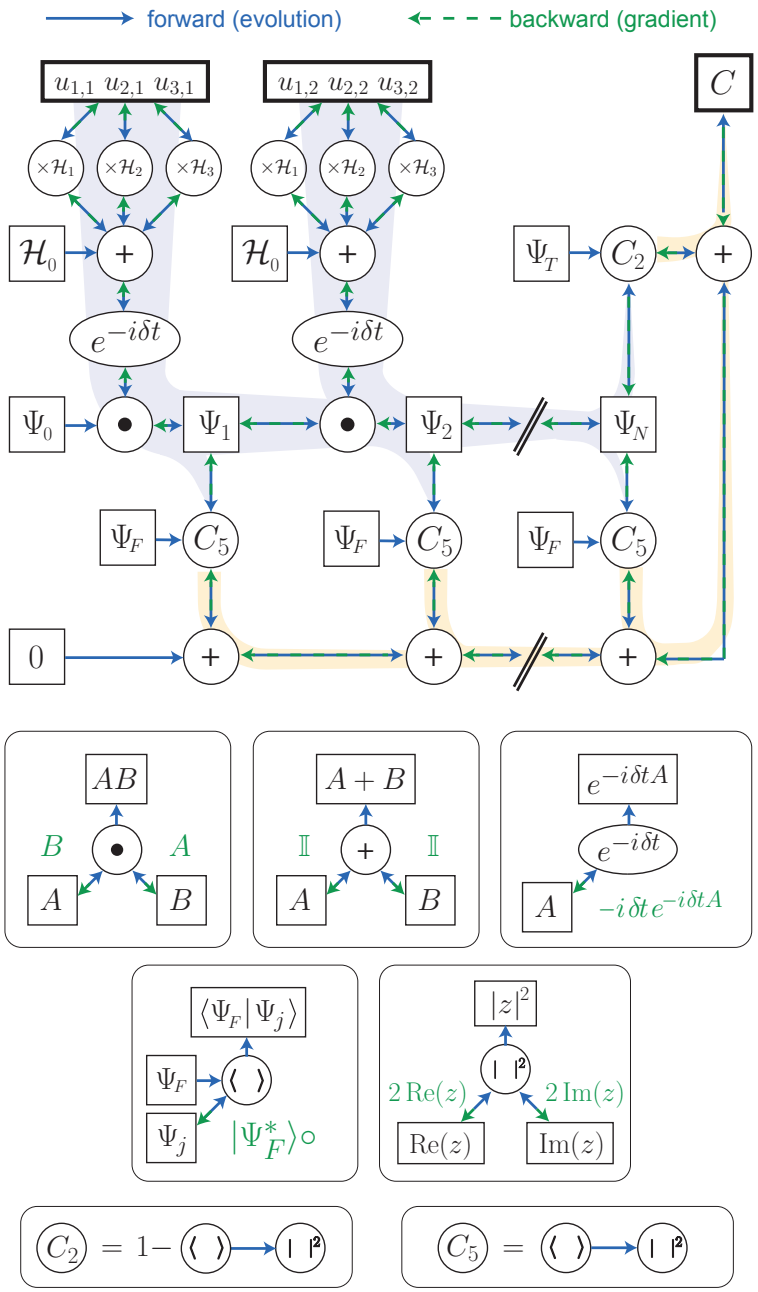

Figure 4.1: Computational network graph for quantum optimal control. Circular nodes in the graph depict elementary operations with known derivatives (matrix multiplication, addition, matrix exponential, trace, inner product, and squared absolute value). Backward propagation for matrices proceeds by matrix multiplication, or where specified, by the Hadamard product $\circ$. In the forward direction, starting from a set of control parameters $u_{k,j}$, the computational graph effects time evolution of a quantum state or unitary, and the simultaneous computation of the cost function C. The subsequent "backward propagation" extracts the gradient $\nabla_{\mathbf{u}} C(\mathbf{u})$ with respect to all control fields by reverse-mode automatic differentiation. This algorithm is directly supported by TensorFlow [1], once such a computational network is specified.

Customarily, algorithms minimizing the cost function $C = \sum_\mu \alpha_\mu C_\mu$ for a given evolution time interval $[t_0, t_N]$ aim to match the desired target unitary or target state at the very end of this time interval. To avoid detrimental effects from decoherence processes during the evolution, it is often beneficial to additionally minimize the gate duration (or state preparation) time $\Delta t = t_N - t_0$ itself. Instead of running the algorithms multiple times for a set of different Δt, we employ cost function contributions of the form $C_6(\mathbf{u}) = 1 - \frac{1}{N}\sum_j |\operatorname{tr}(K_T^\dagger K_j)/D|^2$ for a target unitary, or $C_7(\mathbf{u}) = 1 - \frac{1}{N}\sum_j |\langle\Psi_T|\Psi_j\rangle|^2$ for a target state, respectively. These expressions penalize deviations from the target gate or target state not only at the final time t_N, but *at every time step*. This contribution to the overall cost function therefore guides the evolution towards a desired unitary or state in as short a time as possible under the conditions set by the other constraints, and thus results in a time-optimal gate.

We will demonstrate the utility of these cost function contributions in the context of quantum information processing in Section 4.5. The versatility of automatic differentiation allows straightforward extension to other contexts such as optimization of quantum observables.

4.2.2 Gradient evaluation

The weighted sum of cost functions, $C = \sum_\mu \alpha_\mu C_\mu$, can be minimized through a variety of gradient-based algorithms. Such algorithms are a very popular means of optimization thanks to their good performance and effectiveness in finding optimized solutions for a wide range of problems. At the most basic level, gradient-based algorithms minimize the cost function $C(\mathbf{u})$ by the method of steepest descent, updating the controls $\mathbf{u}$ in the opposite direction of the local cost-function gradient $\nabla_{\mathbf{u}} C(\mathbf{u})$:

$$\mathbf{u}' = \mathbf{u} - \eta \nabla_{\mathbf{u}} C(\mathbf{u}). \tag{4.5}$$

The choice of the update step size η for the control field parameters $\mathbf{u}$, plays an important role for the convergence properties of the algorithm. A number of schemes exist which adaptively determine an appropriate step size η in each iteration of the minimization algorithm. Our im-

plementation supports second order methods such as L-BFGS-B [20] as well as gradient descent methods developed for machine learning such as ADAM [92].

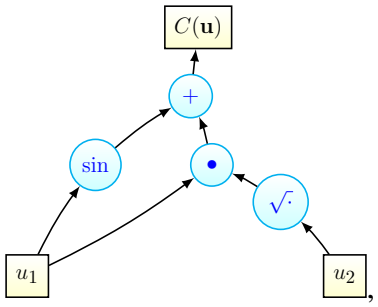

Figure 4.2: Sample computational graph for automatic differentiation. Automatic differentiation utilizes the decomposition of the multivariable cost function $C(\mathbf{u})$ into its computational graph of elementary operations, each of which has a known derivative. In reverse-accumulation mode, all partial derivatives of C are evaluated in a recursion from the top level (C) back towards the outermost branches (variables $\mathbf{u}$).

For the evaluation of the gradient $\nabla_{\mathbf{u}} C$ we make use of automatic differentiation [199, 7] in reverse-accumulation mode. In brief, this algorithm utilizes the decomposition of the multivariable cost function $C(\mathbf{u})$ into its computational graph of elementary operations (addition, matrix multiplications, trace, etc.), each of which has a known derivative. In reverse-accumulation mode, all partial derivatives of C are evaluated in a recursion from the top level (C) back towards the outermost branches (variables $\mathbf{u}$) – rather similar to the procedure of obtaining a derivative with pencil and paper. For instance, for the simple function

$$C(\mathbf{u}) = \sin(u_1) + u_1 \cdot \sqrt{u_2} = f_+ \Big[\sin(u_1),\ f_\bullet(u_1, \sqrt{u_2}) \Big]$$

one obtains all partial derivatives by a recursion starting with the evaluation of

$$\frac{\partial}{\partial u_j} C = D_1 f_+ \, [\frac{\partial}{\partial u_j} \sin] \ + \ D_2 f_+ \, [\frac{\partial}{\partial u_j} f_\bullet] = \cdots .$$

Here, $D_j f$ stands for the derivative of a multivariable function f with respect to its j-th argument; square brackets denote subsequent numerical evaluation of the enclosed term. (Function arguments

are suppressed for brevity.)

Automatic differentiation has become a central tool in machine learning [8], and equally applies to the problem of optimal control of quantum systems. In this approach, the gradient of a set of elementary operations is defined and more complex functions are built as a graph of these operations. The value of the function is computed by traversing the graph from inputs to the output, while the gradient is computed by traversing the graph in reverse via the gradients. This methodology gives the same numerical accuracy and stability of analytic gradients without requiring one to derive and implement analytical gradients specific to each new trial cost function.

All cost functions summarized in table 4.1 can be conveniently expressed in terms of common linear-algebra operations. Figure 4.1 shows the network graph of operations in our software implementation, realizing quantum optimal control with reverse-mode automatic differentiation. For simplicity, the graph only shows the calculation of the cost functions C_2 and C_5. The cost function contributions C_1, C_6, and C_7 are treated in a similar manner. The suppression of large control amplitudes or rapid variations, achieved by C_3 and C_4, is simple to include, since the calculation of these cost function contributions is based on the control signals themselves and does not involve the time-evolved state or unitary. The host of steps for gradient evaluation is based on basic matrix operations like summation and multiplication.

Reverse-mode automatic differentiation [73] provides an efficient way to carry out time evolution and cost function evaluation by one forward sweep through the computational graph, and calculation of the full gradient by one backward sweep. In contrast to forward accumulation, each derivative is evaluated only once, thus enhancing computational efficiency. The idea of backward propagation is directly related to the GRAPE algorithm for quantum optimal control pioneered by Khaneja and co-workers [89], see Chapter 4.6. While the original GRAPE algorithm bases minimization exclusively on the fidelity of the final evolved unitary or state, advanced cost functions (such as C_5 through C_7) require the summation of cost contributions from each intermediate step during time evolution of the system. Such cost functions go beyond the usual GRAPE algorithm, but can be included in the more general backward propagation scheme described above. [Chap-

ter 4.6 shows analytical forms for gradients for cost functions that are based on time evolution ($\{C_1, C_2, C_5\}$).]

4.3 Implementation

Our quantum optimal control implementation utilizes the TensorFlow library developed by Google's machine intelligence research group [1]. This library is open source, and is being extended and improved upon by an active development community. TensorFlow supports GPU and large-scale parallel learning, critical for high-performance optimization. The simple interface to Python allows non-software professionals to implement high-performance machine learning and optimization applications without excessive overhead.

Typical machine-learning applications require most of the same building blocks needed for quantum optimal control. Predefined operations, along with corresponding gradients, include matrix addition and multiplication; matrix traces; and vector dot products. In addition, we have implemented an efficient kernel for approximate evaluation of the matrix exponential and its gradient. Using these building blocks, we have developed a fast and accurate implementation of quantum optimal control, well-suited to deal with a broad range of engineered quantum systems and realistic treatment of capabilities and limitations of control fields.

In common applications of quantum optimal control, time evolving the system under the Schrödinger equation – more specifically, approximating the matrix exponential for the propagators U_j at each time step t_j – requires the biggest chunk of computational time. Within our matrix-exponentiation kernel, we approximate $e^{-iH_j\delta t}$ by series expansion, taking into account that the order of the expansion plays a crucial role in maintaining accuracy and unitarity. The required order of the matrix-exponential expansion generally depends on the magnitude of the matrix eigenvalues relative to the size of the time step. General-purpose algorithms such as expm() in Python's SciPy framework accept arbitrary matrices M as input, so that the estimation of the spectral radius or matrix norm of M, needed for choosing the appropriate order in the expansion, often costs more computational time than the final evaluation of the series approximation itself.

Direct series expansion with only a few terms is sufficient for $H_j\delta$ with spectral radius smaller than 1. In the presence of large eigenvalues, series convergence is slow, and it is more efficient to employ an appropriate form of the "scaling and squaring" strategy, based on the identity

$$\exp M = \left[\exp\left(\frac{M}{2^n}\right)\right]^{2^n}, \tag{4.6}$$

which reduces the spectral range by a factor of 2^n at the cost of recursively squaring the matrix n times [128]. Overall, this strategy leads to an approximation of the short-time propagator of the form

$$U_j \approx \left[\sum_{k=0}^{p} \frac{(-iH_j\delta t/2^n)^k}{k!}\right]^{2^n}, \tag{4.7}$$

based on a Taylor expansion truncated at order p. Computational performance could be further improved by employing more sophisticated series expansions [3, 35] and integration methods [81].

As opposed to the challenges of general-purpose matrix exponentiation, matrices involved in a specific quantum control application with bounded control field strength $(iH_j\delta t)$, will typically exhibit similar spectral radii. Thus, rather than attempting to determine individual truncation levels p_j, and performing scaling-and-squaring at level n_j in each time step t_j, we make a conservative choice for global p and n at the beginning and employ them throughout. This simple heuristic speeds up matrix exponentiation over the default SciPy implementation significantly, primarily due to leaving out the step of spectral radius estimation.

By default, automatic differentiation would compute the gradient of the approximated matrix exponential via backpropagation through the series expansion. However, for sufficiently small spectral radius of M, we may approximate [89]

$$\frac{d}{dx}e^{M(x)} \approx M'(x)\, e^{M(x)}, \tag{4.8}$$

neglecting higher-order corrections reflecting that $M'(x)$ and $M(x)$ may not commute. (Higher-order schemes taking into account such additional corrections are discussed in Ref. [38].) Equation

(4.8) simplifies automatic differentiation: within this approximation, only the same matrix exponential is needed for the evaluation of the the gradient. We make use of this in a custom routine for matrix exponentiation and gradient-operator evaluation, further improving the speed and memory performance.

The TensorFlow library currently has one limitation relevant to our implementation of a quantum optimal control algorithm. Operators and states in Hilbert space have natural respresentations as matrices and vectors which are generically complex-valued. TensorFlow, designed primarily for neural network problems, has currently only limited support for complex matrices. For now, we circumvent this obstacle by mapping complex-valued matrices to real matrices via the isomorphism $H \overset{\cong}{\longmapsto} \mathbb{I} \otimes H_{\mathrm{re}} - i\sigma_y \otimes H_{\mathrm{im}}$, and state vectors $\vec{\Psi} \overset{\cong}{\longmapsto} (\vec{\Psi}_{\mathrm{re}}, \vec{\Psi}_{\mathrm{im}})^t$. Here, $\mathbb{I}$ is the 2×2 unit matrix and σ_y one of the Pauli matrices. Real and imaginary part of the matrix H are denoted by $H_{\mathrm{re}} = \mathrm{Re}\, H$ and $H_{\mathrm{im}} = \mathrm{Im}\, H$, respectively; similarly, real and imaginary parts of state vectors are $\vec{\Psi}_{\mathrm{re}} = \mathrm{Re}\, \vec{\Psi}$ and $\vec{\Psi}_{\mathrm{im}} = \mathrm{Im}\, \vec{\Psi}$. Written out in explicit block matrix form, this isomorphism results in

$$H\vec{\Psi} \overset{\cong}{\longmapsto} \begin{pmatrix} H_{\mathrm{re}} & -H_{\mathrm{im}} \\ H_{\mathrm{im}} & H_{\mathrm{re}} \end{pmatrix} \begin{pmatrix} \vec{\Psi}_{\mathrm{re}} \\ \vec{\Psi}_{\mathrm{im}} \end{pmatrix}, \tag{4.9}$$

rendering all matrices and vectors real-valued. For the Hamiltonian matrix, this currently implies a factor two in memory cost (due to redundancy of real and imaginary part entries). There are promising indications that future TensorFlow releases may improve complex-number support and eliminate the need for a mapping to real-valued matrices and vectors.

4.4 Performance Benchmarking

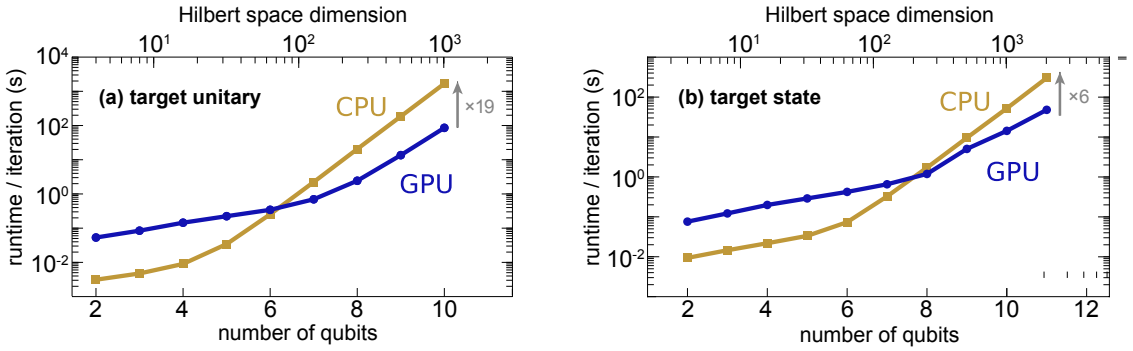

Figure 4.3: Benchmarking comparison between GPU and CPU for (a) a unitary gate (Hadamard transform), and (b) state transfer (GHZ state preparation). Total runtime per iteration scales linearly with the number of time steps. For unitary-gate optimization, the GPU outperforms the CPU for Hilbert space dimensions above ~ 100. For state transer, GPU benefits set in slightly later, outperforming the CPU-based implementation for Hilbert space dimensions above ~ 300. The physical system we consider, in this case, is an open chain of N spin-1/2 systems with nearest neighbor $\sigma_z\sigma_z$ coupling, and each qubit is controlled via fields Ω_x and Ω_y.

Obtaining a fair comparison between CPU-based and GPU-based computational performance is notoriously difficult [101]. We attempt to provide a specific comparison under a unified computation framework. TensorFlow allows for straightforward switching from running code on a CPU to a GPU. For each operation (matrix multiplication, trace, etc.), we use the default CPU/GPU kernel offered by TensorFlow. Note that properly configured, TensorFlow automatically utilizes all threads available for a given CPU, and GPU utilization is found to be near 100%. Not surprisingly, we observe that the intrinsic parallelism of GPU-based matrix operations allows much more efficient computation beyond a certain Hilbert space size, see Fig. 4.3.

In this example, we specifically inspect how the computational speed scales with the Hilbert space dimension when optimizing an n-spin Hadamard transform gate and n-spin GHZ state preparation for a coupled chain of spin-1/2 systems presented in Section 4.5.4. (Details of system parameters are described in the same section.) We benchmark the average runtime for a single iteration for various spin-chain sizes and, hence, Hilbert space dimensions. We find that the GPU quickly outperforms the CPU in the unitary gate problem, even for a moderate system size of ~ 100 basis states. For optimization of state transfer, we observe that speedup from GPU usage, relative CPU

performance, sets in for slightly larger system sizes of approximately ~ 300 basis states.

The distinct thresholds for the GPU/CPU performance gain stem from the different computational complexities of gate vs. state-transfer optimization. Namely, optimizing unitary gates requires the propagation of a unitary operator (a matrix), involving matrix-matrix multiplications, while optimizing state transfer only requires the propagation of a state (a vector), involving only matrix-vector multiplications:

$$U_j|\Psi\rangle \approx \sum_{k=0}^{p} \frac{(-i\delta t)^k}{k!}(H_j \dots (H_j(H_j|\Psi\rangle))), \tag{4.10}$$

Computing the matrix-vector multiplication is generally much faster than computing the matrix exponential itself [169]. For an n-dimensional matrix, the computation of the matrix exponential involves matrix-matrix multiplication, which scales as $O(n^3)$. The computation of state transfer only involves matrix-vector multiplication, which scales as $O(n^2)$ [or even $O(n)$ for sufficiently sparse matrices].

For optimization of the Hadamard transform as well as the GHZ state preparation, we observe a 19-fold GPU speedup for a 10-qubit system (Hilbert space dimension of 1,024) in the former case, and a 6-fold GPU speedup for an 11-qubit system (Hilbert space dimension of 2,048) in the latter case. Since matrix operations are the most computationally intensive task in our software, this speedup is comparable to other GPU application studies that heavily use matrix operation [142, 23, 167, 151, 179, 185, 101]. We emphasize that these numbers are indicative of overall performance trends, but detailed numbers will certainly differ according to the specific system architecture in place. The CPU model we used was an Intel® Core™ i7-6700K CPU @ 4.00 GHz, and the GPU model was an NVIDIA® Tesla® K40c. In this study, all computations are based on dense matrices. Since most physically relevant Hamiltonians are sparse (evolution generally affects sparsity, though), future incorporation of sparse matrices may further improve computation speed for both CPU and GPU [13, 108].

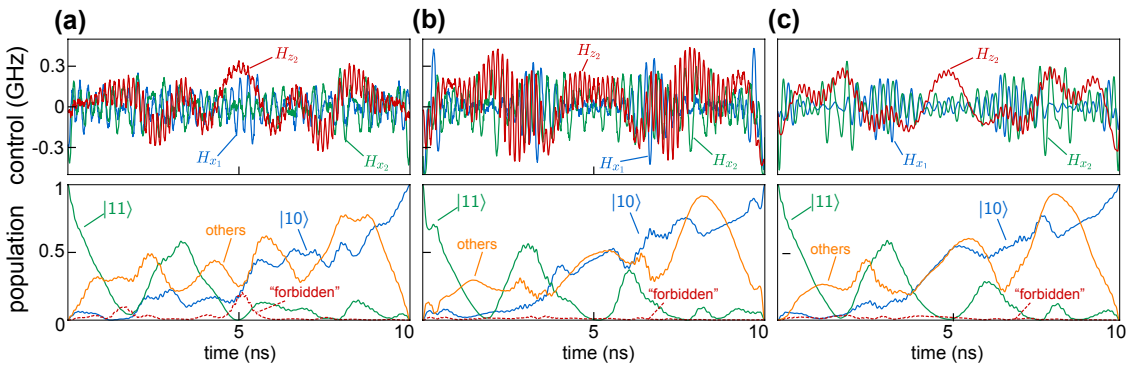

Figure 4.4: Control pulses and evolution of quantum state population for a CNOT gate acting on two transmon qubits, (a) only targeting the desired final unitary, (b) employing an additional cost function suppressing occupation of higher-lying states (C_5), and (c) including additional pulse-shape cost functions (C_3, C_4). Here, only the evolution of state $|11\rangle$ is shown, as the evolution of state $|11\rangle$ is most susceptible to the occupation of higher level states. In all three cases, the CNOT gate converged to a fidelity of 99.9%. The results differ in important details: in (a), both high-frequency "noise" on the control signals and significant occupation of "forbidden" states (3rd and 4th excited transmon level), shown as dashed red line, are visibile throughout the evolution; in (b), forbidden-state occupation is suppressed at each time step during evolution; in (c), this suppression is maintained and all control signals are smoothened. The maximum occupation of forbidden states is reduced from $\sim 20\%$ in (a) to $\sim 3\%$ in (b) and (c). The population of "others" states (non- $|11\rangle$, $|10\rangle$ or "forbidden") is also shown for completeness. For demonstration purposes, all three examples use the same gate duration of 10 ns, despite being subject to different constraints. In practice, one would typically increase the gate time for a more constrained problem to achieve the best result in maximizing gate fidelity, minimizing forbidden state occupation, and achieving a realistic control signal.

4.5 Showcase Applications

In this last section, we present a set of example applications of experimental relevance. The first application demonstrates the importance of cost functions suppressing intermediate occupation of higher-lying states during time evolution, as well as cost functions accounting for realistic pulse shaping capabilities. In a second application, we show how the cost function C_6 can yield high-fidelity state transfer within a reduced time interval. Third, we discuss the application of Schrödinger-cat state preparation – an example from the context of quantum optics and of significant interest in recent schemes aiming at quantum information processing based on such states [74, 125, 193]. This application combines considerable system size with a large number of time

steps, and utilizes most of the cost functions discussed in Section 4.2.1. In the fourth application, we demonstrate the algorithm performance in finding optimal solutions for GHZ state preparation and implementation of a Hadamard transform gate in a chain of qubits with a variable number of qubits. We use either the Adam [92] or L-BFGS-B optimization algorithm [20] for pulse optimization, and achieve a minimum fidelity of 99.9% in all of our following examples.

4.5.1 CNOT gate for two transmon qubits

In the first example, we study realization of a 2-qubit CNOT gate in a system of two coupled, weakly anharmonic transmon qubits. For each transmon qubit ($j = 1, 2$) [96], we take into account the lowest two states spanning the qubit computational space, as well as the next three higher levels. The system Hamiltonian, including the control fields $\{\Omega_{x_1}(t),\, \Omega_{x_2}(t),\, \Omega_{z_2}(t)\}$, then reads

$$\begin{aligned} H(t) = & \sum_{j=1,2} \left[\omega_j b_j^\dagger b_j + \tfrac{1}{2}\alpha_j\, b_j^\dagger b_j (b_j^\dagger b_j - 1)\right] \\ & + J(b_1 + b_1^\dagger)(b_2 + b_2^\dagger) \\ & + \Omega_{x_1}(t)(b_1 + b_1^\dagger) + \Omega_{x_2}(t)(b_2 + b_2^\dagger) + \Omega_{z_2}(t) b_2^\dagger b_2. \end{aligned} \tag{4.11}$$

Here, the ladder operators b_j, and $b_j^\dagger$ are truncated at the appropriate level. (The qubit frequencies $\omega_j/2\pi$ are chosen as 3.5 and 3.9 GHz, respectively; both transmons have an anharmonicity of $\alpha/2\pi = -225$ MHz; and the qubit-qubit coupling strength used in the simulation is $J/2\pi = 100$ MHz.) Consistent with recent circuit QED experiments utilizing classical drives as well as parametric modulation, we investigate control fields acting on $H_{x_1} = b_1 + b_1^\dagger$, $H_{x_2} = b_2 + b_2^\dagger$, and $H_{z_2} = b_2^\dagger b_2$.

We next optimize control fields for the realization of a CNOT gate, with transmon qubit $j = 1$ acting as the control qubit. Our control-field optimization reaches a prescribed fidelity of 99.9% for a 10 ns gate duration in all cases, as seen in Fig. 4.4. Results shown in Fig. 4.4(a) are obtained with the standard target-gate infidelity cost function (C_1) only. It is evident that the solution encounters

two issues: the occupation of the 3rd and 4th excited transmon level ("forbidden") is significant, and control fields are polluted by high-frequency components. Including a cost function contribution of type C_5 succeeds in strongly suppressing occupation of higher levels, see Fig. 4.4(b). This both reduces exposure to increased relaxation rates and ensures that the evolution is minimally influenced by our numerical truncation of Hilbert space. In the final improvement step, shown in Fig. 4.4(c), our optimization additionally suppresses excessive control amplitudes and derivatives via cost contributions of type C_3 and C_4. The inclusion of these terms in the overall cost lessens superfluous "noise" in the control signals, and also helps improve convergence of the algorithm – without reducing the achieved target-gate fidelity.

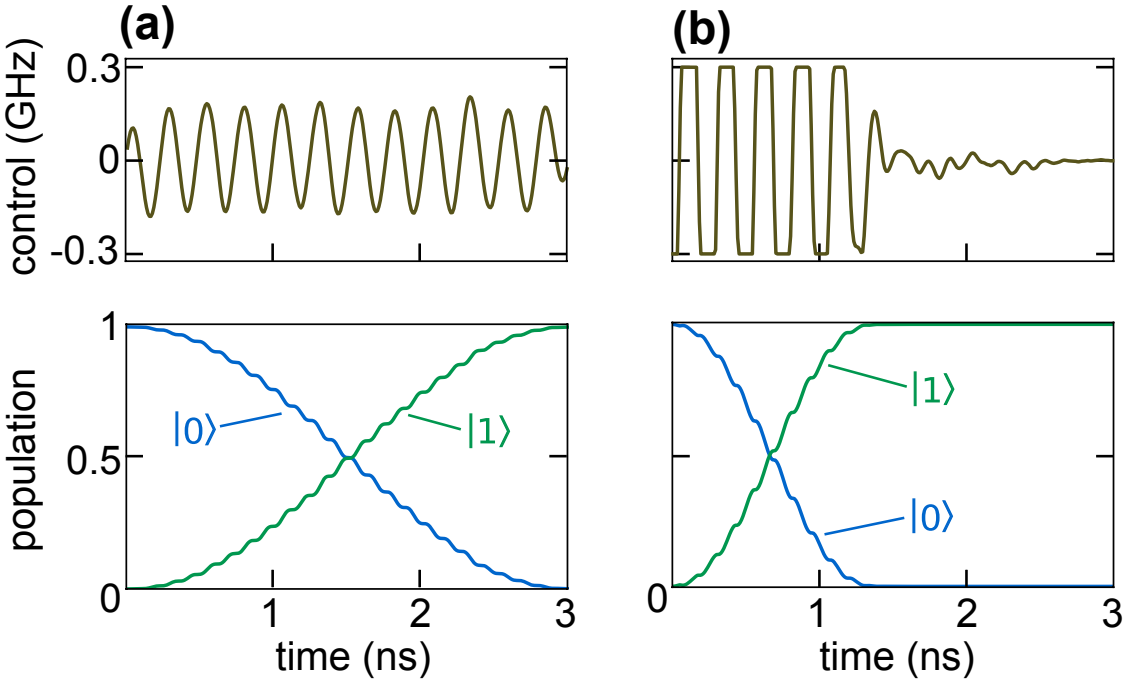

Figure 4.5: Minimizing evolution time needed for a high-fidelity state transfer. (a) No time-optimal award function. (b) With time-optimal award function. (a) Without penalty for the time required for the gate, the control field spreads across the entire given time interval. (b) Once evolution over a longer time duration is penalized with a contribution of type C_6 or C_7 (see table I), the optimizer achieves target state preparation in a shorter time, without loss of fidelity.

4.5.2 Reducing duration of state transfer

In this second example, we illustrate the use of cost function contributions (types C_6, C_7) in minimizing the time needed to perform a specific gate or prepare a desired state. To this end, we consider a two-level spin qubit ($\omega/2\pi$: $3.9\,$GHz). The system and control Hamiltonians combined are taken to be

$$H = \frac{\omega}{2}\sigma_z + \Omega(t)\,\sigma_x. \tag{4.12}$$

We allow for a control field acting on the qubit σ_x degree of freedom, and constrain the maximum control-field strength $\Omega_{\text{max}}/2\pi$ to $300\,$MHz. When the evolution time needed to perform the state transfer is fixed (rather than subject to optimization itself), we observe that control fields generically spread across the prescribed gate duration time. The desired target state is realized only at the very end of the allowed gate duration. When we incorporate a C_6 or C_7-type cost contribution, the optimal control algorithm also aims to minimize the overall gate duration, so as to realize the target unitary or state in as short a time as possible, given other active constraints. In our example, this reduces the time for a state transfer from $3\,$ns to less than $1.5\,$ns, see Fig. 4.5. We note that it is further possible to adaptively change the overall simulation time during optimization. For instance, if further optimization was desired in the case of Fig. 4.5(b), then the simulation time interval could be adaptively reduced to $\sim 1.5\,$ns – resulting in a significant cutback in overall computation time.

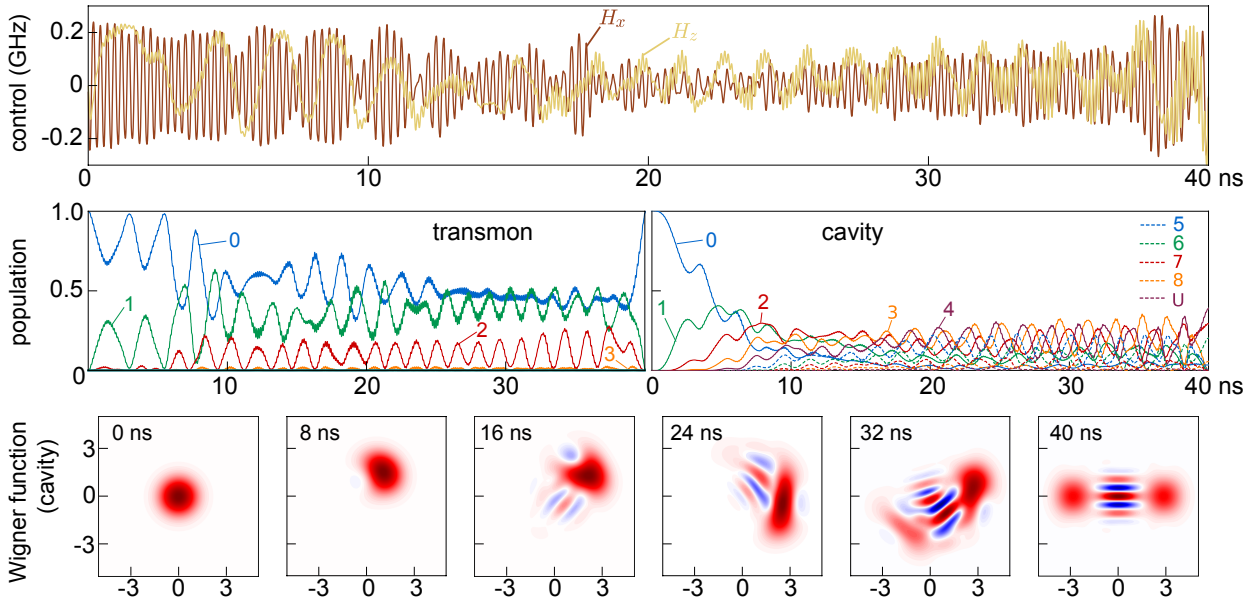

Figure 4.6: Cat state generation. Control pulse, state evolution in Fock basis, and Wigner function tomography of the cavity evolution. Photonic cat state generation is shown as a test of state transfer, challenging the quantum control algorithm with a system of considerable size, large number of required time steps, and inclusion of multiple types of cost function. The desired Schrödinger cat state in the resonator is created indirectly, by applying control fields to a transmon qubit coupled to the resonator, and reached within a prescribed evolution time of 40 ns with a fidelity of 99.9%. (Note that occupation of transmon level 4, 5, 6 remains too small to be visible in the graph.)

4.5.3 Generating photonic Schrödinger cat states

As an example of quantum state transfer, we employ our optimal control algorithm to the task of generating a photonic Schrödinger-cat state. The system we consider to this end is a realistic, and recently studied [125, 193] circuit QED setup, consisting of a transmon qubit capacitively coupled to a three-dimensional microwave cavity. External control fields are restricted to the qubit. Working in a truncated subspace for the transmon (limiting ourselves to levels with energies well below the maximum of the cosine potential), the full Hamiltonian describing the system is

$$H(t) = \omega_q b^\dagger b + \tfrac{1}{2}\alpha\, b^\dagger b(b^\dagger b - 1) + \omega_r a^\dagger a \\ + g(a + a^\dagger)(b + b^\dagger) + \Omega_x(t)(b + b^\dagger) + \Omega_z(t) b^\dagger b \tag{4.13}$$

Here, a and b are the usual lowering operators for photon number and transmon excitation number, respectively. The frequencies $\omega_q/2\pi = 3.5\,\text{GHz}$ and $\alpha/2\pi = -225\,\text{MHz}$ denote the transmon 0-1 splitting and its anharmonicity. The frequency of the relevant cavity mode is taken to be $\omega_r/2\pi = 3.9\,\text{GHz}$. Qubit and cavity are coupled, with a strength parameterized by $g/2\pi = 100\,\text{MHz}$. In our simulation, the overall dimension is $154 = (7 \text{ transmon levels}) \times (22 \text{ resonator levels})$. Note that the rotating wave approximation is not applied in order to reflect the capabilities of modern arbitrary waveform generation.

The state-transfer task at hand, now, is to drive the joint system from the zero-excitation state $|0\rangle_q \otimes |0\rangle_r$ (the ground state if counter-rotating terms in the coupling are neglected) to the photonic cat state $|0\rangle_q \otimes |\text{cat}\rangle_r$. Here, the cat state in the resonator corresponds to a superposition of two diametrically displaced coherent states: $|\text{cat}\rangle_r = \frac{1}{\sqrt{2}}(|\lambda\rangle + |-\lambda\rangle)$. Coherent states are defined in the usual way as normalized eigenstates of the photon annihilation operator a, and correspond to displaced vacuum states $|\lambda\rangle = e^{-|\lambda|^2/2} e^{\lambda a^\dagger}|0\rangle$. The cat state $|\text{cat}\rangle_r$ is approximately normalized for sufficiently large λ. As our concrete target state, we choose a cat state with amplitude $\lambda = 2$ (normalization error of $\sim 0.03\%$). The state transfer is to be implemented by control fields $\Omega_x(t)$ and $\Omega_z(t)$ acting on the transverse and longitudinal qubit degrees of freedom, $H_x = (b + b^\dagger)$ and $H_z = b^\dagger b$, respectively. Matching experimental realizations of multi-mode cavity QED systems[120], we do not allow for any direct control of the cavity degrees of freedom.

This state-transfer problem provides an excellent test for an optimal control algorithm. It incorporates the simultaneous challenges of a large number of time steps (8,000), a considerable evolution time ($40\,\text{ns}$), and the application of most of the cost functions we discussed in Sect. 4.2.1 and summarized in Table 4.1. Specifically, in addition to minimizing the target state infidelity (C_2), we penalize occupation of transmon levels 3 to 6 and cavity levels 20 and 21 (C_5) to avoid artifacts from truncation, and penalize control variations (C_4) [1]. Results from the optimization are presented in 4.6, which shows the control-field sequence, as well as the induced state evolution. At the end of the $40\,\text{ns}$ time interval, the control fields generate the desired cat state with a fidelity of

1. A cost function for reducing evolution time (C_7) was not included in this example.

99.9%. The maximum populations at the truncation levels of transmon and cavity are $\sim 6 \times 10^{-6}$ and $\sim 7 \times 10^{-10}$, respectively. We independently confirm convergence with respect to truncation by simulating the obtained optimized pulse for enlarged Hilbert space (8 transmon and 23 cavity levels), and find that the evolution continues to reach the target state with 99.9% fidelity.

4.5.4 Hadamard transform and GHZ state preparation

We present a final set of examples illustrating the algorithm performance for increasing system size. To that end, we consider a coupled chain of N qubits, or spin-1/2 systems. We assume that all spins are on-resonance in the multiple-rotating frame. This system is described by the Hamiltonian

$$H(t) = \sum_{n=1}^{N} \left[\Omega_x^{(n)}(t)\sigma_x^{(n)} + \Omega_y^{(n)}(t)\sigma_y^{(n)} + J\,\sigma_z^{(n)}\sigma_z^{(n+1)} \right], \tag{4.14}$$

where the coupling term is understood to be dropped for the final summand ($n = N$). The qubit-qubit coupling strength is fixed to $J/2\pi = 100\,\text{MHz}$. Each qubit (n) is controlled via fields $\Omega_x^{(n)}$ and $\Omega_y^{(n)}$, with a maximum allowed drive strength of $\Omega_{x,y}^{(n)}/2\pi = 500\,\text{MHz}$.

As a first optimization task, we search for control fields to implement the unitary realizing a Hadamard transform, commonly used in various quantum algorithms. The gate time we allow for the Hadamard transform is $(2N)$ ns, simulated with $10N$ time steps. Figure 4.7(a) shows the number of iterations and wall-clock time required to converge to the desired 99.9% process fidelity. For the same spin-chain system, we have also employed our code to optimize control fields for transferring the system ground state to a maximally entangled GHZ state. The overall time and time steps we allow for the GHZ state preparation is identical to that used for the Hadamard transform gate. Figure 4.7(b) shows the number of iterations necessary and the total wall-clock time spent for reaching convergence to a result with 99.9% state fidelity. For both examples, we employed computation on either CPU or GPU, depending which one is faster. (This performance benchmarking data was shown in Section 4.4). We note that, when using a modest desktop PC with graphics card, optimal control problems for small Hilbert space size converge within seconds. For a

10-qubit Hadamard gate (Hilbert space dimension of 1,024) or 11-qubit GHZ state (Hilbert space dimension of 2048), it takes $\sim$1 day to obtain a solution meeting the 99.9% fidelity threshold. The total wall-clock time could likely have been reduced significantly by appropriate choice of optimizer, hyperparameters, and/or initial control fields. In the case of spin-chain system, like many quantum information systems, as the number of elements increase, not only does the Hilbert space grow exponentially, the number of control fields and the required number of time steps also get larger. This further increases the complexity of the problem.

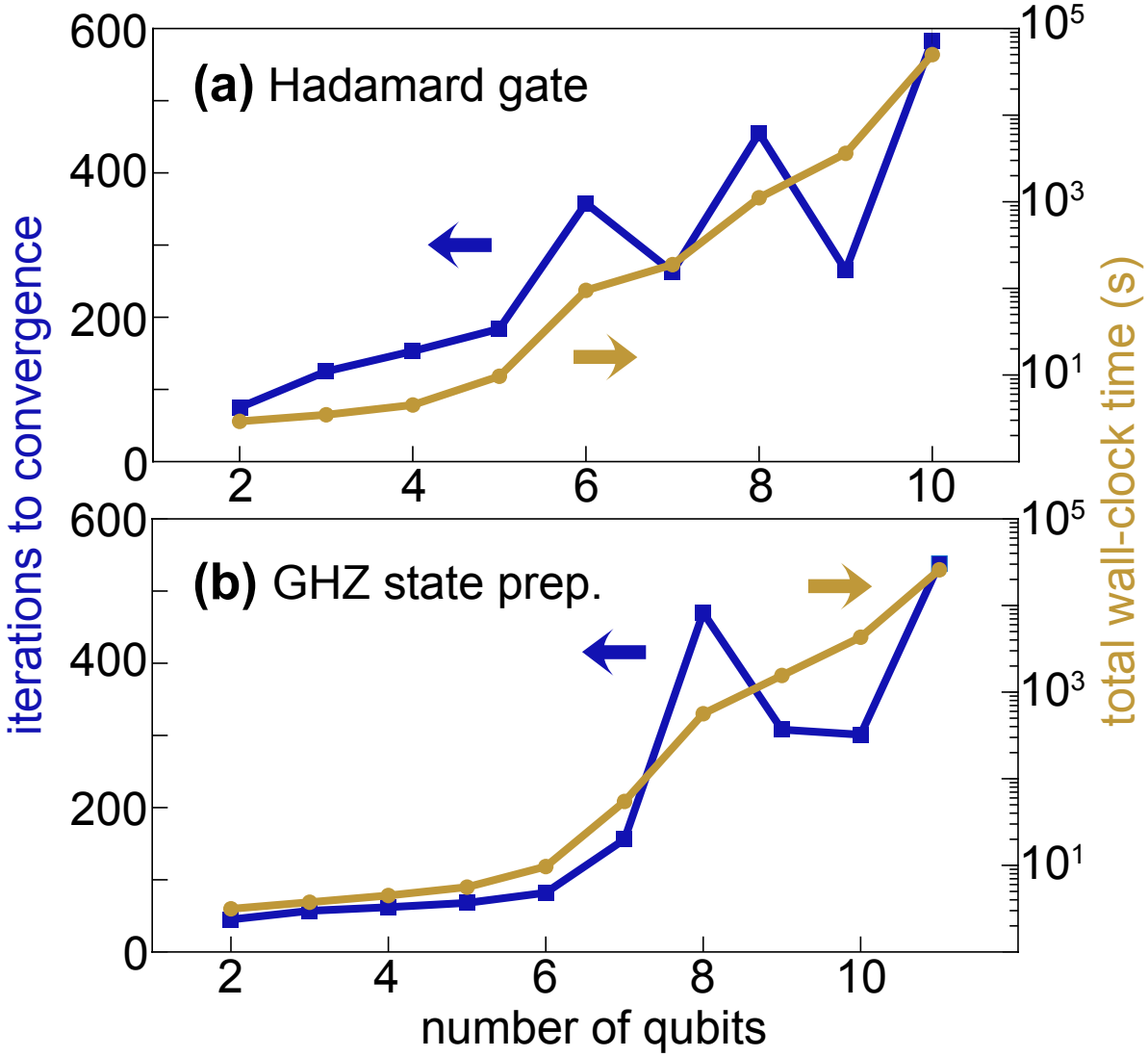

Figure 4.7: Performance of optimal control algorithm as a function of qubit number for (a) a Hadamard transform gate, and (b) GHZ state preparation. As system size increases, total time and number of iterations for the algorithm grow rapidly. The larger number of control parameters and complexity of the target state add to the challenge of quantum optimal control for systems with many degrees of freedom.

4.6 Analytical gradients and algorithms

In the following, we outline the analytical calculation of gradients for cost functions such as those summarized in Table 4.1. We stress that our automatic-differentiation implementation evaluates these gradients autonomously, without the need of these analytical derivations or hard-coding any

new gradients. The following derivations are thus merely intended as illustrations for a better mathematical understanding (and appreciation) of the gradients calculated without user input by means of automatic differentiation.

For a systematic treatment of the different types of cost functions, we note that most cost functions involve an absolute value squared of an inner product between target and final states or target and final unitaries (Hilbert-Schmidt inner product). To obtain the gradients of expressions such as $C_1(\mathbf{u}) = 1 - |\operatorname{tr}(K_T^\dagger K_N)|^2$ with respect to the control parameters, we note that control parameters enter via the final states or unitaries through the evolution operators, $K_N = U_N(\mathbf{u})U_{N-1}(\mathbf{u})\cdots U_1(\mathbf{u})U_0$. To streamline our exposition, we first summarize multiple matrix-calculus relations of relevance.

Consider two complex-valued matrices A and B, compatible in row/column format such that the matrix product AB is defined. Then, one finds

$$\frac{\partial \operatorname{tr}(AB)}{\partial B_{ji}} = \frac{\partial (A_{nm}B_{mn})}{\partial B_{ji}} = A_{ij}. \tag{4.15}$$

Throughout this Chapter, we use Einstein convention for summation, and follow the Jacobian formulation (also known as numerator layout) for derivatives with respect to matrices. We will further encounter expressions of the following form, involving a third matrix C of the same dimensions as B^t:

$$\begin{aligned}
\operatorname{tr}\left[\frac{\partial[|\operatorname{tr}(AB)|^2]}{\partial B}C\right] &= \frac{\partial[\operatorname{tr}(AB)\operatorname{tr}(AB)^*]}{\partial B_{ji}}C_{ji} \\
&= \frac{\partial \operatorname{tr}(AB)}{\partial B_{ji}}\operatorname{tr}(AB)^* C_{ji} = A_{ij}\operatorname{tr}(AB)^* C_{ji} \\
&= \operatorname{tr}(AC)\operatorname{tr}(AB)^*.
\end{aligned} \tag{4.16}$$

In the framework of Wirtinger derivatives in complex analysis, derivatives treat quantities X and X^* as independent variables, and Eq. (4.15) is used in the step from line 1 to line 2.

The evaluation of cost-function gradients requires the application of the chain rule to expres-

sions of the type $\frac{\partial}{\partial u_i} c(M(\mathbf{u}))$. Here, c maps a complex-valued $\ell \times \ell$ matrix M (e.g., the propagator K_N with ℓ denoting the Hilbert space dimension) to a real number (the cost). The matrix $M = (M_{mn})$ itself depends on the real-valued control parameters $\mathbf{u} \in \mathbb{R}^d$. The subscript in u_i is understood as a multi-index $i = (k, j)$ encoding the control-field label k and discretized-time index j. The matrix-calculus result

$$\begin{aligned} \frac{\partial}{\partial u_i} c(M(\mathbf{u})) &= \frac{\partial c}{\partial M_{mn}} \frac{\partial M_{mn}}{\partial u_i} + \frac{\partial c}{\partial M^*_{mn}} \frac{\partial M^*_{mn}}{\partial u_i} \\ &= \operatorname{tr}\left(\frac{\partial c}{\partial M} \frac{\partial M}{\partial u_i} \right) + \text{c.c.} \end{aligned} \tag{4.17}$$

is straightforward to derive with the "regular" chain rule by re-interpreting the functions involved as $c\colon \mathbb{C}^{\ell^2} \to \mathbb{R}$ and $M\colon \mathbb{R}^d \to \mathbb{C}^{\ell^2}$. In the following, Eqs. (4.16) and (4.17) are used to obtain the analytical expressions for several examples of cost-function gradients.

4.6.1 Gradient for C_1: target-gate infidelity

The cost function $C_1 = 1 - |\operatorname{tr}[K_T^\dagger K_N(\mathbf{u})]/D|^2$, penalizes the infidelity of the realized unitary $K_N = U_N U_{N-1} \ldots U_1 U_0$ with respect to the target propagator K_T. In the following, we omit the constant factor D since it affects all the gradients only by a constant factor. The cost function then has the gradient

$$\begin{aligned} \frac{\partial C_1}{\partial u_{k,j}} &\overset{(4.17)}{=} \operatorname{tr} \frac{\partial C_1}{\partial K_N} \frac{\partial K_N}{\partial u_{k,j}} + \text{c.c.} \\ &= -\operatorname{tr}\left[\frac{\partial [|\operatorname{tr}(K_T^\dagger K_N)|^2]}{\partial K_N} \frac{\partial K_N}{\partial u_{k,j}} \right] + \text{c.c.} \\ &\overset{(4.16)}{=} -\operatorname{tr}\left(K_T^\dagger \frac{\partial K_N}{\partial u_{k,j}} \right) \operatorname{tr}(K_T^\dagger K_N)^* + \text{c.c.} \\ &= \operatorname{tr}\left(K_T^\dagger \Big[\prod_{j'>j} U_{j'} \Big] i\,\delta t\, \mathcal{H}_k K_j \right) \operatorname{tr}(K_T^\dagger K_N)^* + \text{c.c.} \\ &= -2\,\delta t \operatorname{Im}\left\{ \operatorname{tr}\left(K_T^\dagger \Big[\prod_{j'>j} U_{j'} \Big] \mathcal{H}_k K_j \right) \operatorname{tr}(K_T^\dagger K_N)^* \right\} \end{aligned} \tag{4.18}$$

where $\prod$ is understood to produce a time-ordered product.

This expression shows that automatic reverse-mode differentiation requires the propagators K_j from every time step. Within TensorFlow, the set of intermediate propagators $\{K_j\}$ is stored in memory during the forward evolution. The resulting memory demand therefore scales as $O(\ell^2 \times N)$.

Memory-efficient algorithm.— We note that storage of $\{K_j\}$ can be avoided by applying the strategy introduced in the original GRAPE paper [89]: since the evolution is unitary, one may time-reverse the evolution step by step, and re-calculate the intermediate propagator via $K_j = U_{j+1}^\dagger K_{j+1}$. Here, each short-time propagator U_j is re-generated locally in time, using only the control fields at time t_j. Such a backwards-propagation algorithm leads to an increase in computation time by roughly a factor of 2 (each U_j is then calculated twice), but has a memory demand of only $O(\ell^2)$ – which does not scale with N, the number of time steps. Thus for large problems the memory efficient algorithm is superior. This memory-efficient algorithm, currently not realized in this implementation, is given by

Algorithm 1 C_1 gradient via backwards propagation

$P = \mathrm{tr}(K_T^\dagger K_N)^* K_T^\dagger$
$X = K_N$
for $j = N$ to 0 **do**
  for all k **do**
    $\partial C_1/\partial u_{k,j} = -2\delta t\, \mathrm{Im}[\mathrm{tr}(P\, \mathcal{H}_k\, X)]$
  end for
  $X = U_j^\dagger X$
  $P = P U_j$
end for
return $\partial C_1/\partial \mathbf{u}$

4.6.2 Gradient for C_2: target-state infidelity

For state preparation or unitaries specified only in a subspace, it is sufficient to optimize the evolution for only a few initial states, rather than for the complete basis. This is achieved by minimizing a cost function based on $C_2(\mathbf{u}) = 1 - |\langle \Psi_T | \Psi_N \rangle|^2$, where the realized final state $|\Psi_N\rangle$ depends

on the control parameters $\mathbf{u}$. Again applying equations (4.17) followed by (4.16) (and using that the trace of a number results in that number), we obtain

$$\frac{\partial C_2}{\partial u_{k,j}} = -2\,\delta t\,\mathrm{Im}\left[\langle\Psi_T|\Big[\prod_{j'>j} U_{j'}\Big]\mathcal{H}_k|\Psi_j\rangle\langle\Psi_T|\Psi_N\rangle^*\right]$$

Memory-efficient algorithm.— In TensorFlow-based automatic differentiation algorithm here, the intermediate states $\{|\Psi_j\rangle\}$ are stored, leading to a memory requirement of $O(\ell\times N)$, rather than $O(\ell^2 \times N)$ for the full propagators. By using the same backward propagation strategy as above, a more memory-efficient algorithm with memory requirement $O(\ell)$ independent of the time-step number is possible:

Algorithm 2 C_2 gradient via backwards propagation

$P = \langle\Psi_T|\Psi_N\rangle^*\langle\Psi_T|$
$X = |\Psi_N\rangle$
for $j = N$ to 0 **do**
    for all k **do**
        $\partial C_2/\partial u_{k,j} = -2\delta t\,\mathrm{Im}[P\,\mathcal{H}_k\,X]$
    end for
    $X = U_j^\dagger X$
    $P = P U_j$
end for
return $\partial C_2/\partial\mathbf{u}$

4.6.3 Gradient for C_5: occupation of forbidden state

Occupation of a "forbidden" state is discouraged by the cost function $C_5 = \sum_j |\,\mathrm{tr}(\Psi_F^\dagger\Psi_j)|^2$. This cost function differs qualitatively from the gate and state infidelity cost functions: the latter are evaluated based on the result at the final time, while forbidden-state occupation involves intermediate states at every time step. Accordingly, the corresponding gradient takes a different form.

First, Eq. (4.17) is replaced by

$$\frac{\partial}{\partial u_i} c\big(\Psi_0(\mathbf{u}), \Psi_1(\mathbf{u}), \ldots, \Psi_N(\mathbf{u})\big) = \operatorname{tr} \frac{\partial c}{\partial \Psi_j} \frac{\partial \Psi_j}{\partial u_i} + \text{c.c.} \tag{4.19}$$

where introduction of the trace of a c-number is convenient for direct application of Eq. (4.16). We then obtain

$$\begin{aligned}
\frac{\partial C_5}{\partial u_{k,j}} &\overset{(4.19)}{=} \sum_J \operatorname{tr} \frac{\partial C_5}{\partial \Psi_J} \frac{\partial \Psi_J}{\partial u_{k,j}} + \text{c.c.} \\
&= \sum_{J \geq j} \sum_{j'} \operatorname{tr} \left[\frac{\partial [|\operatorname{tr}(\Psi_F^\dagger \Psi_{j'})|^2]}{\partial \Psi_J} \frac{\partial \Psi_J}{\partial u_{k,j}} \right] + \text{c.c.} \\
&\overset{(4.16)}{=} \sum_{J \geq j} \operatorname{tr} \left(\Psi_F^\dagger \frac{\partial \Psi_J}{\partial u_{k,j}} \right) \operatorname{tr}(\Psi_F^\dagger \Psi_J)^* + \text{c.c.} \\
&= 2\,\delta t \sum_{J \geq j} \operatorname{Im} \left[\big\langle \Psi_F \big| \big[\textstyle\prod_{j'=j+1}^{J} U_{j'} \big] \mathcal{H}_k \big| \Psi_j \big\rangle \big\langle \Psi_J \big| \Psi_F \big\rangle \right]
\end{aligned} \tag{4.20}$$

The double sum of eq. (4.20) makes it appear as though the computation of this gradient would take $O(N^2)$, however after simplification, the relationship between the limits of the sum and product allow it to be calculated in $O(N)$ time. The corresponding backward propagation algorithm then takes the following form:

Algorithm 3 C_5 gradient via backwards propagation

$P = \langle \Psi_N | \Psi_F \rangle \langle \Psi_F |$
$X = |\Psi_N\rangle$
for $j = N$ to 0 **do**
    for all k **do**
        $\partial C_5 / \partial u_{k,j} = 2\,\delta t \operatorname{Im}[P \mathcal{H}_k X]$
    end for
    $X = U_j^\dagger X$
    $P = P U_j + \langle X | \Psi_F \rangle \langle \Psi_F |$
end for
return $\partial C_5 / \partial \mathbf{u}$

This cost function and gradient is also used as the time-optimal award function, using a negative cost to reward rather than penalize the target state at every time step (rather than just at the end). The gradients of cost functions involving only control fields do not involve the time propagation, so we also omit their derivation.

4.6.4 Discussion of the algorithms

Algorithms for each cost function along with their computation and memory costs have been presented. The computation time of the algorithms all scale linearly with the number N of time steps. Automatic gradient calculation which requires caching of each step causes memory to scale like N, while reducing the run time by a constant factor of 2. By contrast, algorithms which directly exploit the unitary structure of quantum evolution can have memory requirements do not scale with the number of time steps. Hence, it may be worth implementing analytic gradients for very long computations which otherwise would not fit in memory.

Computing the fidelity and gradient for the whole unitary evolution as in algorithm 1, requires $O(\ell^2)$, whereas state transfer requires $O(\ell)$ memory. It should be noted that full unitary evolution fidelity can also be calculated as ℓ^2 state transfer computations over a complete basis. This has the memory requirements of state transfer, and the same computation requirements as algorithm 1, though is less efficient by a constant factor. In principle, each state transfer can be performed in parallel and assembled to compute the total cost and gradient. In addition, the Hamiltonians of many physical problems can be represented sparsely allowing a significant speedup in computation as well. For practical problems, the number time steps required may scale with the size of the problem, as more complex quantum gates/algorithms require more time than simple ones.

4.7 Conclusion

In conclusion, we have presented a quantum optimal control algorithm harnessing two key technologies that enable fast and low-overhead numerical exploration of control signal optimization.

First, we have demonstrated that automatic differentiation can be leveraged to facilitate effortless inclusion of diverse optimization constraints, needed to obtain realistic control signals tailored for the specific experimental capabilities at hand. Automatic differentiation dramatically lowers the overhead for adding new cost functions, as it renders analytical derivations of gradients unnecessary. For illustration, we have presented concrete examples of optimized unitary gates and state transfer, using cost functions relevant for applications in superconducting circuits and circuit QED. We emphasize that this is but one instance within a much larger class of quantum systems for which optimal control is instrumental, and the methods described here are not limited to the specific examples shown in this paper.

The second key technology we have incorporated is the implementation of GPU-based numerical computations, which offers a significant speedup relative to conventional CPU-based code. The use of the TensorFlow library [1] hides the low-level details of GPU acceleration, allowing implementation of new cost functions at a high level. The reduction in computational time will generally depend on a number of factors including system type, Hilbert space size, and the specific hardware employed by the user. We observe that runtime speedup by an order of magnitude is not unusual when using a standard desktop PC, enabling the development of sophisticated quantum control without enormous investments into powerful computing equipment. The underlying libraries also have support for high-performance distributed computing systems for larger optimizations. Our software implementation is open source and can be downloaded at: *github.com/SchusterLab/quantum-optimal-control*.

The increased efficiency and ease of optimal quantum control due to the employment of GPUs and automatic differentiation makes our work valuable to a broad range of research. Future work will address sparse-matrix implementations, as well as the deployment of adaptive step size and Runge-Kutta methods for time evolution.

CHAPTER 5

MULTIMODE CIRCUIT QED

This chapter will focus on our development of a new component of quantum computing architecture, the quantum random access memory. Qubit connectivity is an important property of a quantum processor, with an ideal processor having random access – the ability of arbitrary qubit pairs to interact directly. This is a challenge with superconducting circuits, as state-of-the-art architectures rely on only nearest-neighbor coupling. Here, we implement a random access superconducting quantum information processor, demonstrating universal operations on a nine-qubit memory, with a Josephson junction transmon circuit serving as the central processor. The quantum memory uses the eigenmodes of a linear array of coupled superconducting resonators. We selectively stimulate vacuum Rabi oscillations between the transmon and individual eigenmodes through parametric flux modulation of the transmon frequency. Utilizing these oscillations, we perform a universal set of quantum gates on 38 arbitrary pairs of modes and prepare multimode entangled states, all using only two control lines. We thus achieve hardware-efficient random access multi-qubit control in an architecture compatible with long-lived microwave cavity-based quantum memories. The content of this chapter is based on our work "Random access quantum information processors using multimode circuit quantum electrodynamics" published in *Nature Communications* [133].

5.1 Introduction

Superconducting circuit quantum electrodynamics (cQED) is rapidly progressing towards small and medium-scale quantum computation [40]. Superconducting circuits consisting of lattices of Josephson junction qubits [87, 30] have been used to realize quantum information processors relying on nearest-neighbor interactions for entanglement. An outstanding challenge in cQED is the realization of architectures with high qubit connectivity, the advantages of which have been demonstrated in ion trap quantum computers [78, 39, 107]. Classical computation architectures typically address this challenge by using a central processor which can randomly access a large

memory, with the two elements often comprising distinct physical systems. We implement a quantum analog of this architecture, realizing a random access quantum information processor using cQED.

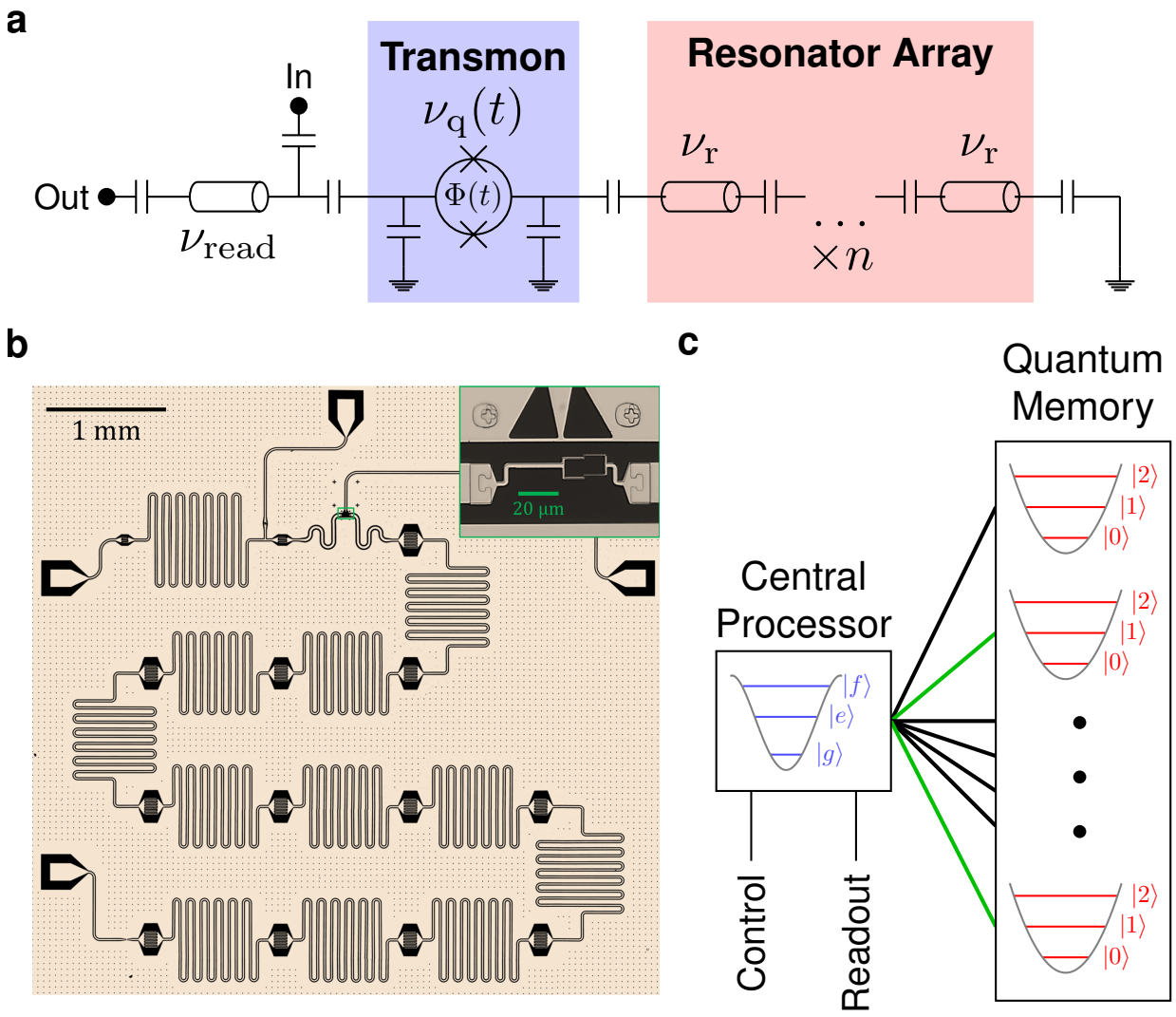

Figure 5.1: **Random access superconducting quantum information processor. a and b,** Schematic and optical image, respectively, of the superconducting microwave circuit. The circuit comprises an array of 11 identically designed, co-planar waveguide (CPW) half-wave resonators, capacitively coupled strongly to each other. The top end of the array is capacitively coupled to a tunable transmon qubit. The transmon is measured with a separate resonator, whose input line doubles as a charge bias for the transmon. The inset shows the tunable SQuID of the transmon, as well as its flux bias above it. **c,** Random access with multiplexed control. The quantum memory consists of the eigenmodes of the array, with each mode accessible to the transmon. This allows for quantum operations between two arbitrary memory modes (such as those highlighted in green) via the central processing transmon and its control lines.

As in the classical case, quantum logic elements, such as superconducting qubits, are expensive in terms of control resources and have limited coherence times. Quantum memories based on harmonic oscillators, instead, can have coherence times two orders of magnitude longer than the best qubits [154, 155, 141], but are incapable of logic operations on their own. This observation suggests supporting each logic-capable processor qubit with many memory qubits. In the near term, this architecture provides a means of controlling tens of highly coherent qubits with minimal cryogenic and electronic-control overhead. To build larger systems compatible with existing quantum error correction architectures [57, 60, 91, 130], one can connect individual modules consisting of a single processor qubit and a number of bits of memory while still accessing each module in parallel.

Here, we describe and experimentally demonstrate the use of a single non-linear element to enable universal quantum logic with random access on a collection of harmonic oscillators. We store information in distributed, readily accessible, and spectrally distinct resonator modes. We show how to perform single qubit gates on arbitrary modes by using frequency-selective parametric control [9, 181, 172, 117, 204, 136] to exchange information between a superconducting transmon qubit [97] and individual resonator modes. Next, using higher levels of the transmon, we realize controlled-phase (CZ) and controlled-NOT (CX) gates on arbitrary pairs of modes. Therefore, we demonstrate all the ingredients necessary for universal quantum computation with harmonic modes. Finally, we use these tools to prepare multi-mode entangled states as an important step towards quantum error correction.

5.2 Multimode quantum memory

To build a multimode quantum memory we use the eigenmodes of a linear array of $n = 11$ identical, strongly coupled superconducting resonators [121] (see Figure 5.1). For a linear array, the eigenmodes correspond to distributed "momentum" states (See Chapter 5.6). Importantly, every mode has non-zero amplitude at the edge, allowing the transmon to couple to each mode. The

Hamiltonian of the combined system is:

$$\begin{aligned}\hat{H} = h\nu_{\mathrm{q}}(t)\hat{a}^\dagger\hat{a} \quad &+ \quad \frac{1}{2}h\alpha\,\hat{a}^\dagger\hat{a}(\hat{a}^\dagger\hat{a}-1) + \sum_{k=1}^{n} h\nu_k\hat{b}_k^\dagger\hat{b}_k \\ &+ \quad \sum_{k=1}^{n} hg_k(\hat{b}_k + \hat{b}_k^\dagger)(\hat{a}+\hat{a}^\dagger), \end{aligned} \tag{5.1}$$

where the transmon is treated as a Duffing oscillator [97] with anharmonicity α, coupled to the modes with frequency ν_k ($6-7$ GHz) and coupling strength g_k ($50-200$ MHz, Chapter 5.12). The operators $\hat{a}^\dagger$ ($\hat{a}$) and $\hat{b}_k^\dagger$ ($\hat{b}_k$) create (annihilate) photons in the transmon and in eigenmode k, respectively. While this implementation is straightforward, the idea of a multimode memory also applies to related systems with many harmonic degrees of freedom, including long transmission-line [182] or 3D waveguide cavities. We limit ourselves to the zero- and one-photon Fock states of the eigenmodes. It is also possible to use more of the oscillator Hilbert space, allowing logical encoding in terms of cat [126] and binomial code [124] states.

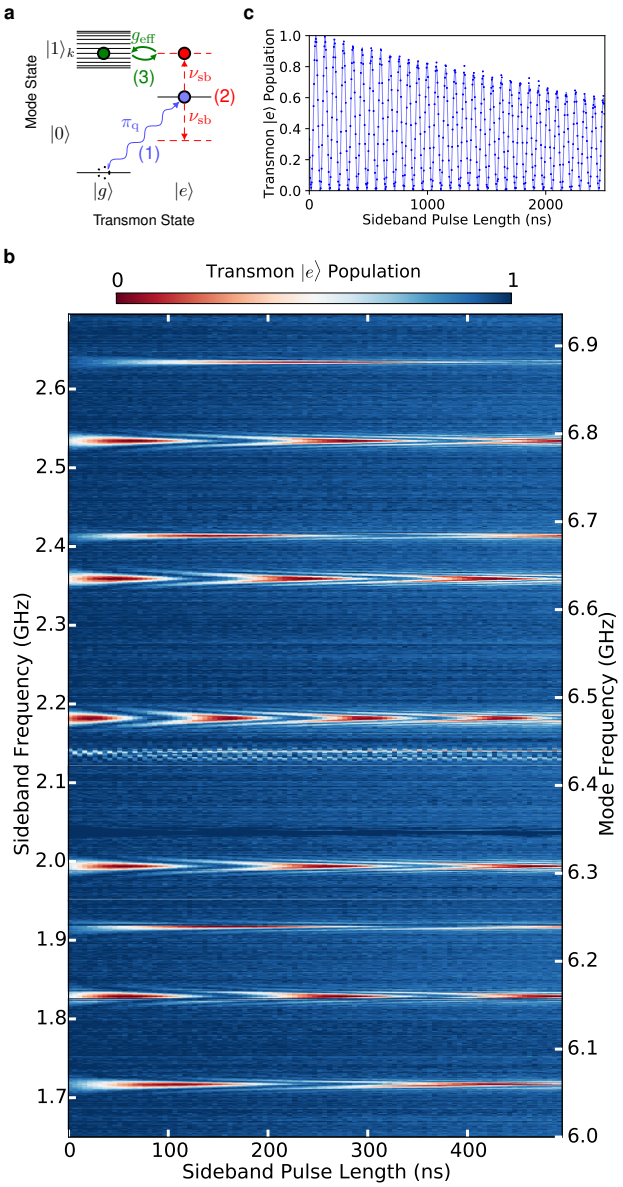

Figure 5.2: **Stimulated vacuum Rabi oscillations.** **a,** Generation of stimulated vacuum Rabi oscillations. $|1\rangle_k$ is the state with a single photon in mode k; all other modes are in the ground state. (1) An excitation is loaded into the transmon via its charge bias. (2) The transmon frequency is flux-modulated to create sidebands. (3) When a sideband is resonant with a mode, single-photon vacuum Rabi oscillations occur between transmon and the mode. **b,** Experimental results obtained from this protocol for a range of sideband modulation frequencies, with the transmon biased at $\nu_q = 4.28$ GHz. The length of the flux modulation pulse is swept for each frequency and the excited state population of the transmon is measured after the pulse ends. Chevron patterns indicate parametrically induced resonant oscillations with each of the memory modes. Two of the eleven modes are weakly coupled to the transmon and are not visible at these flux modulation amplitudes. The distribution of the modes can be understood through Hamiltonian tomography [111] (Chapter 5.13). **c,** Resonant oscillations between transmon and mode 6.

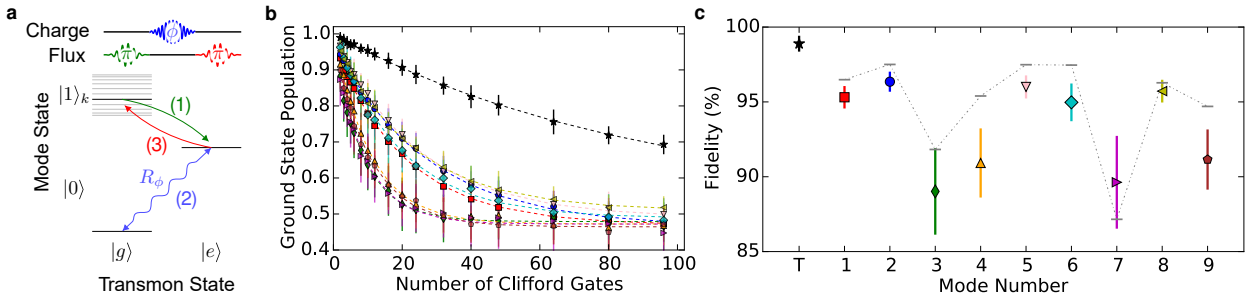

Figure 5.3: **Single-mode gate protocol and benchmarking. a,** The sequence for generating arbitrary single-qubit gates of a memory mode: (1) The mode's initial state, consisting of a superposition of 0 and 1 photon Fock states, is swapped to the transmon (initially in its ground state), using a transmon-mode iSWAP (see text). (2) The transmon is rotated by the desired amount (R_ϕ) via its charge control line. (3) The rotated state is swapped back to the mode, by reversing the iSWAP gate in (1). Segments of this sequence are used to achieve state preparation [steps (2) and (3)] and measurement [steps (1) and (2)] of each mode. **b,** Single-mode randomized benchmarking. We apply sequences of varying numbers of consecutive Clifford gates, then invert each sequence with a unique Clifford gate. We measure the transmon ground-state population after inversion and average over 32 different random sequences, with the standard deviation (s.d.) plotted as error bars for each sequence length. **c,** From fitting the resulting data, we find single-mode gate fidelities from 89.0±2.9% to 96.3±0.7% and a transmon (T in the figure) gate fidelity of 98.9±1.3%. These are consistent with the expected coherence-limited fidelities, plotted as gray bars (s.d. from fit plotted as error bars.)

Given access to the multimode memory via the transmon, we demonstrate methods to address each mode individually. In many circuit QED schemes, excitations are loaded into modes by adiabatically tuning the qubit frequency through or near a mode resonance [42]. This works well for single modes, but for a multimode manifold, one must carefully manage Landau-Zener transitions through several modes [121], to avoid leaving residual excitations elsewhere in the manifold. Also, the qubit must be returned to the far-dispersive regime to minimize spurious unwanted interactions, requiring longer gate durations.

5.3 Stimulated vacuum Rabi oscillations

We induce resonant interactions between the transmon and an individual mode by modulating the transmon excitation energy via its flux bias. The modulation creates sidebands of the transmon excited state, detuned from the original resonance by the frequency of the applied flux tone. When

one of these sidebands is resonant with a mode of the memory, the system experiences stimulated vacuum Rabi oscillations: parametrically induced exchange of a single photon between the transmon and the selected mode. These are similar to resonant vacuum Rabi oscillations [160], but occur at a rate that is controlled by the modulation amplitude [181, 9] $g_{\mathrm{eff},k} = g_k J_1(\epsilon/2\nu_{\mathrm{sb}})$, where J_1 is the first Bessel function, ϵ and ν_{sb} are the amplitude and frequency of the modulation, respectively, and g_k is the bare coupling rate to eigenmode k (Chapter 5.7). The rate of photon exchange is linear to lowest order in ϵ and can be as large as $g_k/2$.

To illustrate the application of parametric control for addressing the multimode memory, we employ the experimental sequence shown in Figure 5.2a. First, the transmon is excited via its charge bias. Subsequently, we modulate the flux to create sidebands of the transmon excited state at the modulation frequency. This is repeated for different flux pulse durations and frequencies, with the population of the transmon excited state measured at the end of each sequence. When the frequency matches the detuning between the transmon and a given eigenmode, we observe full-contrast stimulated vacuum Rabi oscillations. In Figure 5.2b, we see the resulting characteristic chevron patterns [181] as the modulation frequency approaches the detuning between the transmon and each of the modes. For long modulation times, the excited state population approaches zero. This is evident in the stimulated vacuum Rabi oscillation between the transmon and mode 6 shown in Figure 5.2c. This indicates that the original photon is being exchanged between the transmon and the mode and no other photons are being pumped into the system. We achieve photon exchange between the transmon and individual modes in 20-100 ns, depending on the mode. This rate is limited by spectral crowding arising from neighboring modes and sideband transitions involving the transmon $|f\rangle$ level. This operation is coherent and can be used to transfer arbitrary qubit states between the transmon and the memory mode, corresponding to a transmon-mode iSWAP [137] in the single-excitation subspace.

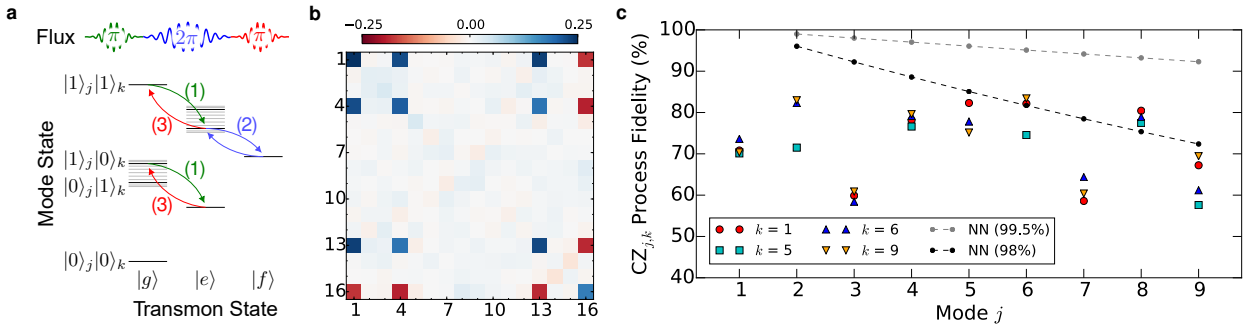

Figure 5.4: **Controlled-phase gate between two arbitrary modes. a,** Protocol for controlled-phase (CZ) gate between an arbitrary pair of modes, with j indicating the control mode and k indicating the target mode of the gate: (1) The state of mode j is swapped to the transmon via a transmon-mode iSWAP pulse at the frequency difference between the transmon $|g\rangle$-$|e\rangle$ transition and mode k. (2) A CZ gate is performed between mode k and the transmon, by applying two frequency-selective iSWAPs from energy level $|e1\rangle$ to level $|f0\rangle$ and back, mapping the state $|e1\rangle$ to $-|e1\rangle$. (3) The state of the transmon is swapped back to mode j, reversing the iSWAP in (1). **b,** Process matrix for the CZ gate between modes $j = 6$ and $k = 8$, corresponding to a process fidelity of 82% (see Chapter 5.21 for details on state preparation and measurement.) **c,** Fidelities from process tomography for 38 pairs of memory modes with $k = 2, 5, 6, 8$. The process fidelities are extracted from sequences that include SPAM errors, and are conservative estimates of the gate fidelities. For comparison, the dashed black and gray lines show the decay in fidelity for a two-qubit gate between qubit 1 and qubit j in a corresponding linear array comprising only nearest-neighbor gates with fidelities of 99.5 [6] and 98%, respectively.

5.4 Universal quantum control

The transmon-mode iSWAP and arbitrary rotations of the transmon state via its charge bias provide a toolbox for universal state preparation, manipulation, and measurement of each mode of the quantum memory. In Figure 5.3, we illustrate how to perform these operations. To characterize the quality of our single-mode operations, we perform randomized benchmarking (RB) [93, 31]. We generate single-mode Clifford operations by sandwiching single-qubit Clifford rotations of the transmon with transmon-mode iSWAPs (Chapter 5.16). We achieve RB fidelities ranging from 89.0±2.9% to 96.3±0.7%. These fidelities approach the expected coherence limit, indicated by the gray bars in the figure. The coherence limits are estimated based on the qubit RB fidelity, the iSWAP times ($20 - 100$ ns) and the coherence times ($T_1 = 1 - 5$ μs, $T_2^* = 1 - 8.5$ μs) of individual modes (Chapter 5.14). Each single-mode gate consists of two transmon-mode iSWAPs,

and a single transmon gate. From the error in the single-mode and transmon RB, we estimate the fidelities of the individual transmon-mode iSWAP operations to range from 95 to 98.6%.

To achieve universal control of the quantum memory, we extend our parametric protocols to realize two-mode gates. We perform conditional operations between the transmon and individual modes by utilizing the $|e\rangle$-$|f\rangle$ transition of the transmon. A controlled-phase (CZ) gate between the transmon and an individual mode consists of two sideband iSWAPs resonant to the $|e1\rangle$-$|f0\rangle$ transition, selectively mapping the state $|e1\rangle$ to -$|e1\rangle$, leaving all other states unchanged due to the anharmonicity of the transmon (Chapter 5.10). To enact a CZ gate between two arbitrary modes, the control mode is swapped into the transmon, a transmon-mode CZ is performed, and the mode is swapped back as illustrated in Figure 5.4a. In our device, gate speeds (250-400 ns) are primarily limited by crosstalk between iSWAP operations on the $|g\rangle$-$|e\rangle$ and $|e\rangle$-$|f\rangle$ transitions of modes with difference frequencies approaching the anharmonicity of the transmon. This crosstalk can be reduced by tailoring the frequency spacing of the memory modes and the anharmonicity of the transmon. In addition to the CZ gate, we obtain controlled-X and Y gates (CX, CY) between modes by swapping $|e\rangle$ and $|f\rangle$ transmon state populations in the middle of the pulse sequence for the CZ gate. These gate protocols can be extended to realize two-mode SWAP gates (Chapter 5.17), as well as multi-qubit gates such as Toffoli and controlled-controlled-phase (CCZ) gates [77] between arbitrary modes.

To perform high-fidelity gates between modes, several issues must be considered. These include: (1) DC shifts of the transmon frequency during iSWAP pulses (~ 10 MHz), (2) dispersive shift of the $|e1\rangle$ state (~ 1 MHz), and (3) stimulated dispersive shifts of non-targeted modes during iSWAP pulses ($\sim 10-100$ kHz). We fully compensate effect (1) and correct the phase error arising from (2) by calibrating the phase errors and suitably adjusting the relative phases of the iSWAP pulses (Chapter 5.18). The error from (3) is relatively small and currently adds to the gate error.

Our multimode architecture allows for straightforward measurements of arbitrary multi-bit correlators, forming a basis for tomography, and for the stabilizer measurements required for error correction. An arbitrary correlator comprises products of Pauli operators applied to each of the

memory bits, and corresponds to a generalized parity measurement. This is exactly the back-action on the phase measurement of a transmon while serving as the control of a CZ (CX) gate targeting a memory mode [54]. The value of an arbitrary stabilizer can thus be measured by performing Ramsey interferometry of the transmon with a series of CZ (CX) gates applied to the desired memory modes.

We use correlator measurements to characterize a CZ gate between a given pair of modes via process tomography. We perform process tomography by applying the gate on 16 linearly independent input states that form a basis for an arbitrary two-qubit density matrix [140]. The resulting density matrices are reconstructed through state tomography. For two-qubit state tomography, we map all correlators to individual measurements of the transmon, using combinations of single- and two-mode gates.

In order to obtain a fair estimate of the gate fidelity, each of the process tomography sequences has a single two-mode gate. Additional gates required for tomography are combined with the characterized CZ gate (Chapter 5.21). The process matrix obtained using this protocol for a CZ gate between modes 6 and 8 is shown in Figure 5.4b. We use this protocol to characterize the fidelities for gates between 38 mode pairs, as shown in Figure 5.4c. The fidelities from full process tomography range approximately from $60 - 80\%$ for the CZ gates between the mode pairs indicated. These fidelities incorporate state preparation and measurement (SPAM) errors, with the SPAM sequences being of similar duration as the gates. Conservative estimates from single-mode and transmon RB (see Figure 5.3c) give SPAM errors of $5 - 10\%$, depending on the modes addressed. The gate fidelities are largely limited by the coherence times of the modes ($\sim 5 - 15\%$ error). Future devices based on 3D superconducting cavities [154] may have up to three orders of magnitude enhancement in memory mode coherence times. The process fidelities are additionally limited by dephasing of the transmon ($\sim 5\%$ error), and residual coherent errors arising from bare and stimulated dispersive shifts. The error from the dephasing can be reduced by coupling a fixed-frequency transmon to the multimode memory using a tunable coupler [26, 117, 109]. Additionally, biasing the tunable coupler at a point with small static coupling also reduces coherent errors from the the

bare dispersive shift.

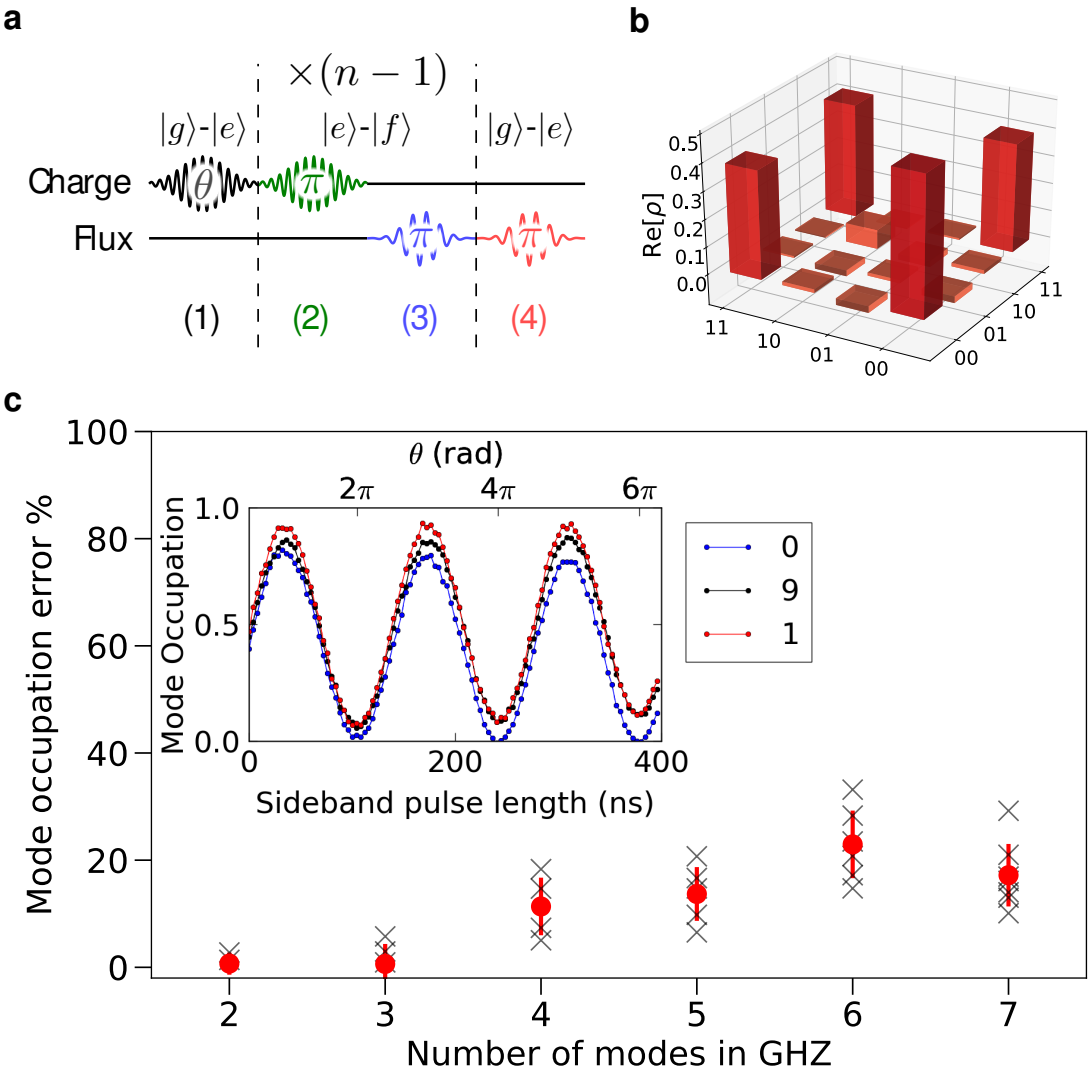

Figure 5.5: **Multimode entanglement.** **a,** Pulse sequence for generating n-mode maximally-entangled states. Step (1) creates a superposition of the transmon $|g\rangle$ and $|e\rangle$ states, with the relative amplitudes of the superposition controlled by the rotation angle θ. Steps (2) and (3) load photons into modes of the memory, conditioned on the transmon state by utilizing the transmon $|f\rangle$ state. These steps are repeated $n-1$ times to entangle additional modes. Step (4) performs a $|g\rangle$-$|e\rangle$ iSWAP to the last mode, disentangling the transmon from the modes. **b,** The real part of the density matrix (ρ) of the $|\Phi^+\rangle$ Bell state of mode 6 and 8, obtained from tomography. **c,(inset)** Correlated oscillations resulting from sweeping θ and measuring each mode individually. This demonstrates control of the relative amplitudes of the entangled state superposition. **c,** Deviation from expected mean populations of each of the modes, upon preparation of the GHZ state ($\theta = \frac{\pi}{2}$). The red filled circles and error bars indicate the average and s.d., respectively, over individual mode measurements (black crosses).

Figure 5.4c highlights the advantages of random access in a quantum computing architecture. An entangling gate between the first and the jth qubit of an array with only nearest-neighbor coupling would require $2j-1$ gates (such as CXs or iSWAPs). This results in an exponential decay of the fidelity with increasing distance between the corresponding qubits. Conversely, in a random access quantum information processor, there is no additional computational cost to perform gates between arbitrary pairs of qubits. Even without considering potential improvements in the coherence times, we see (Figure 5.4c) that the processor performs competitively with state-of-the-art gates [6] between distant qubits in a nearest-neighbor architecture. While we have highlighted the advantages of this processor in terms of random access and minimal control hardware, a resulting requirement is the need to perform sequential operations. The number of modes which can be multiplexed to a single qubit without loss of fidelity, is given by the ratio of the loss from idling in a cavity mode to the loss in performing qubit operations, which for modern 3D cavities can be up to 100 [154].

5.5 Multimode entanglement

We use universal control of the quantum memory to build maximally entangled states spanning several modes, using the protocol described in Figure 5.5a. First, we create a superposition of the transmon ground and excited states. Next, we add a photon to the desired mode, conditioned on the transmon state. This is repeated for each mode in the entangled state. Finally, we disentangle the transmon from the memory modes, transferring the remaining population into the final mode. In Figure 5.5b, we show full state tomography for a Bell state [11] with state fidelity $F = 0.75$, including errors from tomography (Chapter 5.20). In the inset of Figure 5.5c, we apply the protocol to three modes and show populations of each of the modes as a function of the initial qubit rotation angle, θ. Finally, in Figure 5.5c, we show the population error from the target state at $\theta = \pi/2$, corresponding to a photonic Greenberger-Horne-Zeilinger (GHZ) state [68] of up to seven modes. While the three mode GHZ state can be demonstrated to be tripartite entangled through a measurement of the Mermin witness [123, 42] (Chapter 5.23), full characterization of entangled states

of more than two modes is hampered by the additional gates required for tomography and the gate fidelities of the current device. This protocol however illustrates the ease with which a random access quantum information processor can be used to generate multimode entangled states of arbitrary modes. Variants of this sequence can be used to create other classes of multimode entangled states, including W states, Dicke states [44] and cluster states [153]. Such states are valuable resources in several quantum error correction schemes and are useful for quantum optics and sensing [15].

5.6 Multimode quantum memory Hamiltonian

The following sections describe in detail on the characterization and calibrations of the multimode quantum random access memory. We begin with an in-depth discussion of our system Hamiltonian. Our multimode quantum memory implementation uses the eigenmodes of a linear array of $n = 11$ identical, strongly coupled superconducting resonators [121], as shown in Figure 5.1. The array is described by the Hamiltonian:

$$\hat{H}_{\mathrm{mm}} = \sum_{j=1}^{n} h\nu_{\mathrm{r}} \hat{c}_j^{\dagger}\hat{c}_j + \sum_{j=1}^{n-1} hg_{\mathrm{r}}(\hat{c}_j^{\dagger}\hat{c}_{j+1} + \hat{c}_j\hat{c}_{j+1}^{\dagger}), \tag{5.2}$$

where ν_r is the resonance frequency of the identical resonators, g_r is the tunnel-coupling between neighboring resonators, and $\hat{c}_j^{\dagger}$ ($\hat{c}_j$) is the operator that creates (annihilates) photons in the resonator at spatial index j. The coupling ($g_r \sim 250$ MHz) between neighboring resonators is larger than the disorder in the resonator frequencies ($\nu_r = 6.45 \pm 0.1$ GHz). Thus, the single-photon eigenmodes of this circuit are 11 distributed "momentum" states of the array, with eigenfrequencies in a band from 6 to 7 GHz and mode spacing varying from 50 to 150 MHz. The eigenmodes

and eigenfrequencies of the tight-binding Hamiltonian of Equation (5.2) are [121]:

$$|\psi\rangle_k = \sum_{j=1}^{n} \sqrt{\frac{2}{n+1}} \sin\left(\frac{jk\pi}{n+1}\right) |1\rangle_j, \tag{5.3}$$

$$\nu_k = \nu_\mathrm{r} - 2g_\mathrm{r} \cos\left(\frac{k\pi}{n+1}\right) \quad k \in \{1, n\}, \tag{5.4}$$

where $|\psi\rangle_k$ and ν_k are the kth eigenstate and eigenfrequency, respectively, and $|1\rangle_j$ is the state with a single photon in the jth resonator of the array and with all other resonators in the ground state. The coupling between the transmon and a given eigenmode is given by:

$$g_k = g_\mathrm{q} \sqrt{\frac{2}{n+1}} \sin\left(\frac{k\pi}{n+1}\right). \tag{5.5}$$

5.7 Parametric control Hamiltonian

The frequency of the transmon is tunable using the magnetic flux threading the SQuID loop of the transmon, controlled by passing a current through a nearby flux line. For a sinusoidally modulated flux, the transmon $|g\rangle - |e\rangle$ transition frequency is:

$$\nu_\mathrm{ge}\left(\Phi\left(t\right)\right) = \nu_\mathrm{ge}\left(\Phi_b + \epsilon_\Phi \sin\left(2\pi\nu_\mathrm{sb}t + \phi\right)\right) \sim \bar{\nu}_\mathrm{ge} + \epsilon \sin\left(2\pi\nu_\mathrm{sb}t + \phi_\mathrm{m}\right) \tag{5.6}$$

where $\bar{\nu}_\mathrm{ge} = \nu_\mathrm{ge}\left(\Phi_b\right) + \delta\nu_\mathrm{DC}(\Phi_b, \epsilon_\Phi)$ is the mean qubit frequency during the flux modulation. The relation between the frequency (ϵ) and flux (ϵ_Φ) modulation amplitudes and the DC-shift of the transmon frequency during flux-modulation are:

$$\epsilon = \epsilon_\Phi \left.\frac{d\nu_\mathrm{ge}}{d\Phi}\right|_{\Phi_b} \quad \text{and} \quad \delta\nu_\mathrm{DC} = \frac{\epsilon_\Phi^2}{4} \left.\frac{d^2\nu_\mathrm{ge}}{d\Phi^2}\right|_{\Phi_b}, \tag{5.7}$$

respectively. The frequency is shifted from its bare value due to the non-linear flux dependence of the transmon frequencies, and is quadratic in the modulation amplitude.

We obtain a simple description for parametric control of this system by considering the Hamil-

tonian of Equation (1) in the main text and restricting to the lowest two transmon levels. The Hamiltonian then reduces to:

$$\hat{H} = \sum_{k=1}^{n} h\nu_k \hat{b}_k^\dagger \hat{b}_k + \frac{1}{2} h\nu_q(t)\hat{\sigma}_z + \sum_{k=1}^{n} hg_k(\hat{b}_k + \hat{b}_k^\dagger)(\hat{\sigma}_- + \hat{\sigma}_+), \tag{5.8}$$

The lowest two levels form a qubit, whose frequency is modulated over time by using the flux bias. In this work, we focus on iSWAP interactions between the transmon and the resonator mode. As a result, we modulate the transmon frequency near the the difference frequencies of the memory modes and transmon. The $(\hat{b}_k\hat{\sigma}_- + \text{c.c.})$ terms in Equation (5.8) can be therefore be dropped in the rotating-wave approximation. When the modulation frequency is resonant with the detuning between the eigenmode and the transmon, this realizes effectively resonant interactions and stimulated vacuum Rabi oscillations of a single photon between the transmon and the mode. These sidebands manifest in a rotating frame defined by the transformation [9, 181]:

$$U(t) = \exp\left[-2\pi i\left(\bar{\nu}_{ge}t - \frac{\epsilon}{2\nu_{sb}}\cos\left(2\pi\nu_{sb}t\right)\right)\hat{\sigma}_z - 2\pi i\nu_k\hat{b}_k^\dagger\hat{b}_k t\right]. \tag{5.9}$$

In this rotating frame, the Hamiltonian is transformed to:

$$\begin{aligned}\hat{H}' = U\hat{H}U^\dagger - iU\partial_t U^\dagger &= \sum_{j=1}^{n} hg_k J_0\left(\frac{\epsilon}{2\nu_{\text{sb}}}\right)\left(e^{-2\pi i\Delta_k t}\hat{b}_k^\dagger\hat{\sigma}_- + e^{2\pi i\Delta_k t}\hat{b}_k\hat{\sigma}_+\right) \\ &+ \sum_{j=1}^{n} hg_k\hat{b}_k^\dagger\hat{\sigma}_-\left(\sum_{m=1}^{\infty}(-1)^m J_m\left(\frac{\epsilon}{2\nu_{\text{sb}}}\right)e^{2\pi i(m\nu_{\text{sb}}-\Delta_k)t}\right) + \text{c.c.},\end{aligned} \tag{5.10}$$

where $\Delta_k = \nu_k - \bar{\nu}_{ge}$, is the detuning between the qubit and the k^{th} eigenmode. When $\nu_{sb} = \Delta_k$, we obtain resonant first-order sideband transitions between the transmon and mode k, described by:

$$H'_{\text{sb},k} = hg_k J_1\left(\frac{\epsilon}{2\nu_{\text{sb}}}\right)\left(\hat{b}_k^\dagger\hat{\sigma}_- + \hat{b}_k\hat{\sigma}_+\right) \Rightarrow \tau_{\text{iSWAP}} = \frac{1}{2g_k J_1\left(\frac{\epsilon}{2\nu_{\text{sb}}}\right)}. \tag{5.11}$$

We perform universal operations on the multimode memory using iSWAP operations between a

given mode and both the $|g\rangle - |e\rangle$ and $|e\rangle - |f\rangle$ transmon transitions, with the latter allowing the realization of entangling gates between arbitrary eigenmodes. The minimal description of our gate operations on the multimode memory therefore involves three transmon levels, with the parametric control of the eigenmodes described by an extension of the Hamiltonian of Equation (5.10) to a single qutrit coupled to the harmonic memory modes. In addition to sideband transitions, the Hamiltonian also includes dispersive shifts arising from photons in the memory modes, due to the bare coupling between the eigenmodes and the transmon. We ignore the correction to the dispersive shift due to the modulation amplitude dependence of the bare term ($\propto J_0(\epsilon_{\rm m}/2\nu_{\rm sb})$), whose lowest order contribution is quadratic in $\epsilon/2\omega_{\rm m}$. These effects are described by the following simplified Hamiltonian:

$$\begin{aligned}
\tilde{H}(t) &= \tilde{H}_{\rm sb} + \tilde{H}_{\rm q} + \tilde{H}_{\chi}, && (5.12)\\
\tilde{H}_{\rm sb} &= \sum_k \sum_{\alpha \in ge,ef} \left(g^{\alpha}_{eff,k}(t) \hat{b}^{\dagger}_k \hat{\sigma}^{-}_{\alpha} e^{2\pi i(\nu_{sb} - \Delta^{\alpha}_k)t} + \text{c.c.} \right), && (5.13)\\
\tilde{H}_{\rm q} &= \sum_{\alpha \in ge,ef} \left(\Omega^{\alpha}(t)\, \hat{\sigma}^{-}_{\alpha} + \text{c.c.} \right), && (5.14)\\
\tilde{H}_{\chi} &= \sum_k \left(\chi^{e}_k |e\rangle\langle e| + \chi^{f}_k |f\rangle\langle f| \right) \hat{b}^{\dagger}_k \hat{b}_k. && (5.15)
\end{aligned}$$

In the above, $\hat{\sigma}^{-}_{ge} = |g\rangle\langle e|$, $\hat{\sigma}^{-}_{ef} = |e\rangle\langle f|$ and $g^{\alpha}_{{\rm eff},k}(t) = g^{\alpha}_k J_1\left(\frac{\epsilon^{\alpha}_k(t)}{2\nu_{\rm sb}}\right)$. $\epsilon^{\alpha}_k(t)$ for $\alpha \in \{ge, ef\}$ are the strengths of time-dependent parametric frequency modulation tones addressing mode k and $\Delta^{\alpha}_k = \nu_k - \bar{\nu}_{\alpha}$ is the detuning between mode k and the frequency of the corresponding transmon transition frequency $\alpha \in \{ge, ef\}$. $\Omega^{\alpha}(t)$ are the strengths of the transmon charge drives and $\chi^{e,f}_k$ are the dispersive shifts of the $|e\rangle$ and $|f\rangle$ levels resulting from the addition of a photon in mode k. In addition to the dispersive shift, there are second-order tunneling terms of the form $b^{\dagger}_l b_k$ for $l \neq k$ arising from the virtual hopping of photons between different eigenmodes via the transmon. These terms are of the same order as the dispersive shift, but their effect can be ignored since they correspond to off-resonant tunneling ($\sim$ 1 MHz) between non-degenerate levels (spaced by $\sim$ 100 MHz). We note that there is also a shift (DC-offset) of the qubit frequency during the flux

modulation, arising from the non-linear flux-frequency relation of the transmon. Given that the flux pulses used in the experimental sequences in this work are sequential, we include their effect as:

$$\tilde{H}_{\mathrm{DC}} = \sum_{j;\beta\in\{e,f\}} \delta\nu_{\mathrm{DC}}^{j,\beta}\left(\epsilon_j^{\beta}(t)\right)|\beta\rangle\langle\beta|\,, \tag{5.16}$$

with additional cross terms being present if different flux tones were turned on simultaneously. We can further simplify the Hamiltonian above by ignoring off-resonant terms. If the transmon charge and flux-modulation tones are of the form $\epsilon_{q,\mathrm{sb}}\cos\left(\omega_{q,\mathrm{sb}}t+\phi_{q,\mathrm{sb}}\right)$, and we consider near resonant operations with a single eigenmode k, the drive phases ($\phi_{q,\mathrm{sb}}$) enter the effective Hamiltonian as:

$$\tilde{H}_{\mathrm{sb},\alpha}(t) = g_{\mathrm{sb},\alpha}(t)\hat{a}_k^{\dagger}\hat{\sigma}_{\alpha}^{-}e^{-i\phi_{\mathrm{sb}}} + \mathrm{c.c.} \qquad \tilde{H}_{\mathrm{q}} = \Omega_{\alpha}(t)\hat{\sigma}_{\alpha}^{+}e^{-i\phi_q} + \mathrm{c.c.} \quad \alpha\in\{ge,ef\} \tag{5.17}$$

Reducing to 2×2 subspaces over which each of these terms act, and taking the top row to be the state with the higher transmon level, and with $\theta(t)=2\Omega_{\alpha}t$ and $2g_{\mathrm{sb},\alpha}t$ for the sideband and qubit drives, we obtain:

$$\hat{H}_{\mathrm{sb},ge}(t) = \begin{pmatrix} 0 & g_{\mathrm{sb},ge}e^{i\phi_{\mathrm{sb}}} \\ g_{\mathrm{sb},ge}e^{-i\phi_{\mathrm{sb}}} & 0 \end{pmatrix} \tag{5.18}$$

$$\Rightarrow \hat{U}_{\mathrm{sb},ge}^{\theta} = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -i\sin\left(\frac{\theta}{2}\right)e^{i\phi_{\mathrm{sb}}} \\ -i\sin\left(\frac{\theta}{2}\right)e^{-i\phi_{\mathrm{sb}}} & \cos\left(\frac{\theta}{2}\right) \end{pmatrix},$$

$$\hat{H}_{\mathrm{q},ge}(t) = \begin{pmatrix} 0 & \Omega_{ge}e^{-i\phi_{\mathrm{q}}} \\ \Omega_{ge}e^{i\phi_{\mathrm{q}}} & 0 \end{pmatrix} \tag{5.19}$$

$$\Rightarrow \hat{U}_{\mathrm{q},ge}^{\theta} = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -i\sin\left(\frac{\theta}{2}\right)e^{-i\phi_{\mathrm{q}}} \\ -i\sin\left(\frac{\theta}{2}\right)e^{i\phi_{\mathrm{q}}} & \cos\left(\frac{\theta}{2}\right) \end{pmatrix}.$$

5.8 Device design and fabrication

The CPW resonators in the array have a center pin width of 12 μm and a gap width of 6 μm. They are coupled to each other via interdigitated capacitors, where each side of the capacitor has 6 digits that are 107 μm long, 6 μm wide, and spaced by 6 μm. The capacitor coupling the array to the qubit is identical to the intra-array capacitors to minimize disorder of the resonators in the array. The transmon is capacitively coupled to ground via CPW capacitors on either side of the SQuID, with a center pin width of 20 μm and gap of 10 μm. The SQuID is a 20 μm by 10 μm loop, with two square junctions that are 170 nm and 125 nm wide. The flux bias line 6 μm from the SQuID is dipolar, with 25 μm long and 2 μm wide wires on each end. The ground plane of the chip has an array of 5 μm wide square holes spaced by 50 μm for flux vortex pinning [175]. The linear elements of the full circuit design have been simulated with a commercial 3D finite element analysis software (ANSYS HFSS).

The device (shown in Figure 5.1) was fabricated on a 430 μm thick C-plane sapphire substrate. The base layer of the device, which includes the majority of the circuit (excluding the Josephson junctions of the transmon), consists of 100 nm of aluminum deposited via electron-beam evaporation at 1 Å/s, with features defined via optical lithography and reactive ion etch (RIE) at wafer-scale. The wafer was then diced into 7x7 mm chips. The junction mask was defined via electron-beam lithography with a bi-layer resist (MMA-PMMA) in the Manhattan pattern, with overlap pads for direct galvanic contact to the optically defined capacitors. Before deposition, the overlap regions on the pre-deposited capacitors were milled *in-situ* with an argon ion mill to remove the native oxide. The junctions were then deposited with a three step electron-beam evaporation and oxidation process. First, an initial 35 nm layer of aluminum was deposited at 1 Å/s at an angle of 29° relative to the normal of the substrate, parallel azimuthally to one of the fingers in the Manhattan pattern [62] for each of the junctions. Next, the junctions were exposed to 20 mBar of high-purity O_2 for 12 minutes for the first layer to grow a native oxide. Finally, a second 120 nm layer of aluminum was deposited at 1 Å/s at the same angle relative to the normal of the substrate, but orthogonal azimuthally to the first layer of aluminum. After evaporation, the remaining resist

was removed via liftoff in N-Methyl-2-pyrrolidone (NMP) at 80°C for 3 hours, leaving only the junctions directly connected to the base layer, as seen in the inset of Figure 1 of the main text. After both the evaporation and liftoff, the device was exposed to an ion-producing fan for 15 minutes, in order to avoid electrostatic discharge of the junctions.

The device is mounted and wirebonded to a multilayer copper PCB microwave-launcher board. Additional wirebonds connect separated portions of the ground plane to eliminate spurious differential modes. The device chip rests in a pocketed OFHC copper fixture that presses the chip against the launcher board. Notably, the fixture contains an additional air pocket below the chip to alter 3D cavity modes resulting from the chip and enclosure, shifting their resonance frequencies well above the relevant band by reducing the effective dielectric constant of the cavity volume.

5.9 Transmon and readout resonator properties

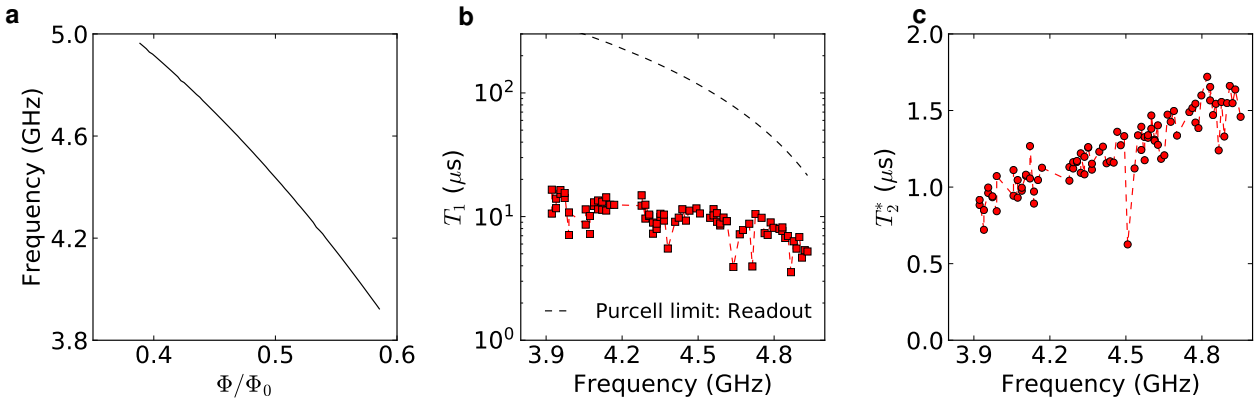

Figure 5.6: **a,** Qubit frequency as a function of applied DC flux bias in units of flux quanta. **b,** Energy relaxation time (T_1) as a function of transmon frequency, with the Purcell limit from the readout resonator shown for comparison. **c,** Ramsey (T_2^*) coherence times as a function of transmon frequency.

The parameters of the transmon are obtained by fitting the spectrum obtained as a function of the applied DC flux. The Josephson and electrostatic charging energies extracted from these fits are $E_{\mathrm{J,max}} = 22.2$ GHz and $E_{\mathrm{C}} = 192$ MHz, while the SQuID loop junction asymmetry, $(E_{\mathrm{J1}} - E_{\mathrm{J2}})/(E_{\mathrm{J1}} + E_{\mathrm{J2}}) = 0.1$. These parameters correspond to maximum and minimum qubit

frequencies of 5.84 GHz and ~ 2 GHz, respectively. The experiments in this work were typically performed with the transmon biased between $3.9-4.7$ GHz (see Figure 5.6a). This frequency band is ~ 2 GHz away from the eigenmodes of the resonator array. As a result, photons in the multi-mode manifold cause relatively small dispersive shifts of the transmon frequency. Additionally, the slope of the flux-frequency curve in this regime allows for sufficiently large frequency modulation amplitudes, while limiting sensitivity to flux noise to maintain transmon coherence. The transmon qubit state is probed using a capacitively coupled CPW readout resonator. The frequency and the quality factor of the readout resonator are $\nu_{\text{read}} = 5.255$ GHz and $Q = 15000$, and the coupling to the qubit is $g_{\text{read}} = 47$ MHz. For the typical transmon frequency range, we obtain single-shot readout fidelities between $0.3 - 0.85$ using dispersive [196] and high-power [158] circuit QED readout schemes. The readout signal is calibrated by appending a sequence with no pulse, and one with a transmon $|g\rangle - |e\rangle$ π pulse at the of each set of experimental sequences. Upon averaging over 1000-2000 experiments, the readout signal results in a visibility of $\sim 99\%$, limited by the fidelity of the single qubit gates, and consistent with the RB data.

The coherence of the transmon is characterized by standard lifetime (T_1) and Ramsey (T_2^*) experiments. The measured T_1 of the $|e\rangle$ state and T_2^* of the $|g\rangle - |e\rangle$ transition are show as a function of the $|g\rangle - |e\rangle$ transition frequency in Figure 5.6b and 5.6c. The T_1 is found to show a slight decrease with increasing frequency in this regime, explained partially by increased Purcell loss from coupling to the readout resonator (see Figure 5.6b). The T_1 at a given flux bias slowly varies with time (over the course of weeks) by $\sim 25 - 30\%$. The T_2^* is found to increase with frequency, consistent with reduced sensitivity to flux noise due to the decreasing slope of the frequency-flux curve transform. We note that T_2^* showed no improvement from reducing the cutoff frequency of the external cryogenic low-pass filter on the DC flux bias from 100 to 3 kHz. At $\nu_{\text{q}} = 4.36$ GHz, the T_2 obtained from a spin-echo experiment with a single π pulse is 3.7 μs. This time could be increased to ~ 14 μs using a dynamical decoupling sequence (Carr-Purcell-Meiboom-Gill with 61 pulses) [19]. We note that the measured T_2^* jumped from 400 ns to 1.2 μs, 2-3 weeks following cooling the fridge to the base temperature (20 mK), coincident with a shift

and stabilization of the applied current corresponding to a flux quantum. The coherence of the $|f\rangle$ level is characterized by analogous lifetime and Ramsey experiments. The lifetime of the $|f\rangle$ level at $\nu_{\mathrm{q}}^{ge} = 4.3325$ GHz is $T_{1,ef} = 3.7$ μs while the phase coherence time is $T_{2,ef}^{*} = 1.2$ μs.

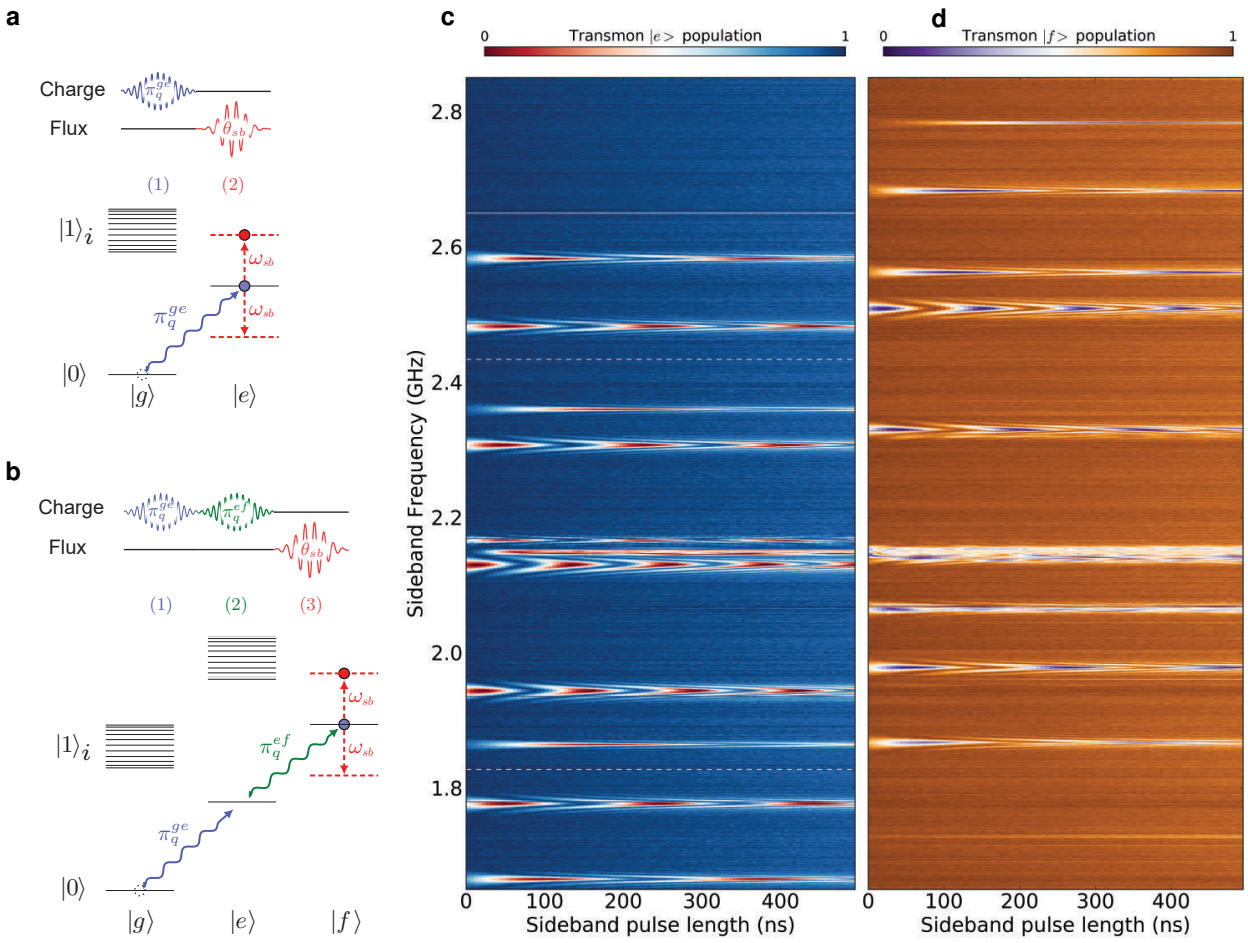

Figure 5.7: **a and b,** Level diagrams and experimental protocols for probing sideband transitions across $|g\rangle - |e\rangle$ and $|e\rangle - |f\rangle$, respectively. **c and d,** The corresponding stimulated vaccum Rabi spectra with the transmon biased at $\nu_{ge} = 4.3325$ GHz. These spectra are taken at a fixed flux-modulation amplitude, after correcting the distortion of the flux pulses due to the transmission profile of the flux bias (see chapter 5.11). At this drive amplitude, only nine of the 11 modes can be resolved, with the two remaining modes being weakly coupled. The locations of the two weakly coupled modes, found by driving at higher flux-modulation amplitudes, are indicated by the white dashed lines in **c**.

5.10 Stimulated vacuum Rabi oscillations with multiple transmon levels

The eigenmodes of the resonator array are probed by sideband spectroscopy using the RF flux bias. The flux-modulation pulses are directly synthesized using a Tektronix AWG70000A arbitrary waveform generator (AWG) with a sampling rate of 50 GSa/s. The target pulse waveform envelopes are typically square, with Gaussian edges ($\sigma = 10 - 20$ ns, 2σ cutoff) added to reduce the pulse bandwidth and thus minimize crosstalk between modes. The protocol used to the measure the $|g\rangle - |e\rangle$ sideband spectrum is shown in Figure 5.2 and reiterated in Figure 5.7a, with the corresponding measured spectrum shown in Figure 5.7c . The spectrum for $|e\rangle - |f\rangle$ sideband transitions are obtained using a similar protocol (see Figure 5.7b), beginning instead by first loading the transmon in the $|f\rangle$ state. Following a sideband flux pulse with frequency ν_{sb} and duration τ, we measure the population in the transmon $|f\rangle$ state. The $|f\rangle$ state population is obtained by mapping $|e\rangle - |f\rangle$ to $|g\rangle - |e\rangle$, and subsequently measuring the $|e\rangle$ population as before. The spectrum thus obtained as a function of ν_{sb} and τ is plotted in Figure 5.7d. These spectra are taken at a fixed frequency-modulation amplitude, after calibrating and correcting for the transfer function of the RF flux bias as described in the following section.

5.11 Correcting for the transmission profile of the flux bias line

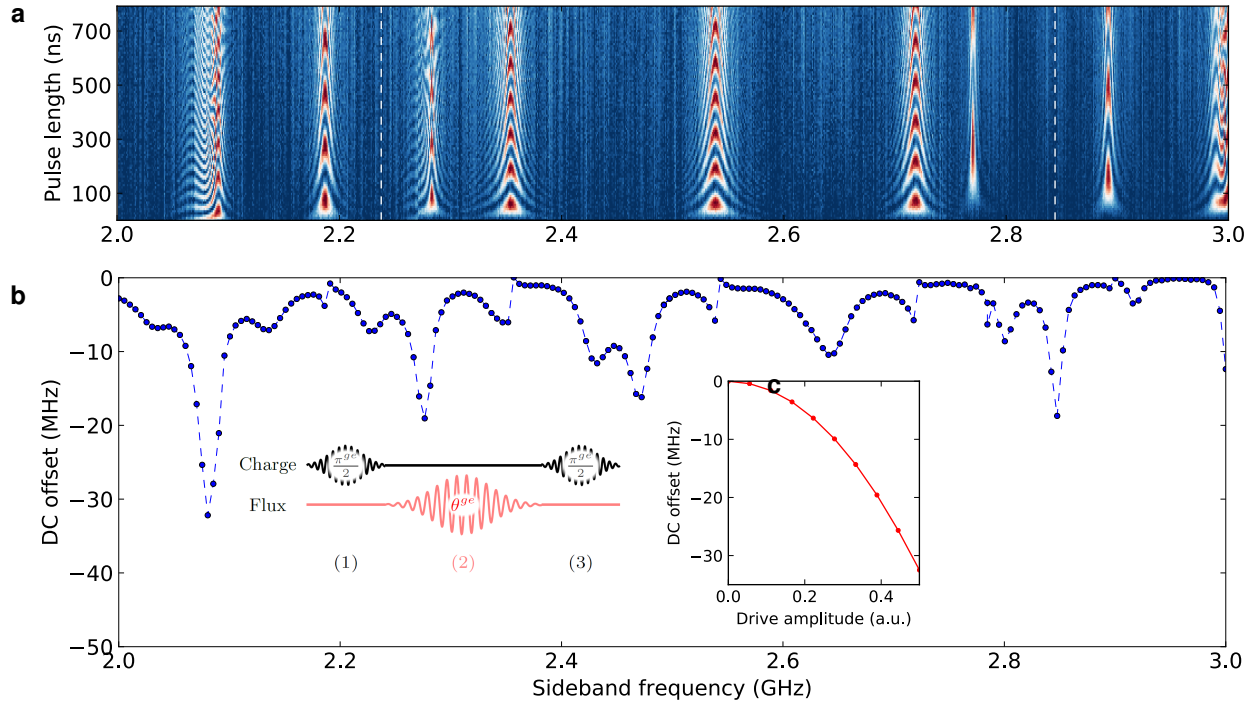

Figure 5.8: **a,** Sideband spectrum on the $|g\rangle - |e\rangle$ transition with the transmon biased at $\nu^{ge} = 3.9225$ GHz. **b,** Qubit DC-offset at a fixed drive amplitude as a function of sideband frequency measured using a transmon Ramsey experiment on the $|g\rangle - |e\rangle$ transition with a flux pulse inserted during the idle time. **c,** DC-offset as a function of drive amplitude for $\nu_{\mathrm{sb}} = 2.63$ GHz, showing the expected quadratic dependence with drive amplitude.

We directly calibrate the frequency-modulation amplitude using the DC-shift of the qubit frequency during flux modulation. This frequency shift is measured using a transmon Ramsey interferometry experiment, with a flux pulse inserted during the idle time (Figure 5.8b inset). For a fixed external RF voltage amplitude, the measured DC-offset as a function of the flux-modulation frequency is shown in Figure 5.8b, along with the corresponding $|g\rangle - |e\rangle$ sideband spectrum is shown in Figure 5.8a. The stimulated vacuum Rabi chevrons for some modes are found to be distorted from the expected shape [181] (see mode 0 and 10). These distortions are due to resonances in the transmission profile of the flux bias line, as seen in the DC-offset spectroscopy. The memory modes appear as avoided crossings in a Ramsey experiment on $|g\rangle - |e\rangle$ transition, due to

interference of Ramsey fringes arising from the DC-offset and resonant stimulated vacuum Rabi oscillations. At a given frequency, the DC-offset shows a quadratic dependence on the modulation amplitude seen in Figure 5.8(c), as expected from equation (5.7). The amplitude profile of the transfer function $T(\nu)$ of the flux bias line is obtained from the DC-offset at fixed AWG drive voltage, with $|T(\nu)| \propto \sqrt{|\delta\nu_{\mathrm{DC}}|}$.

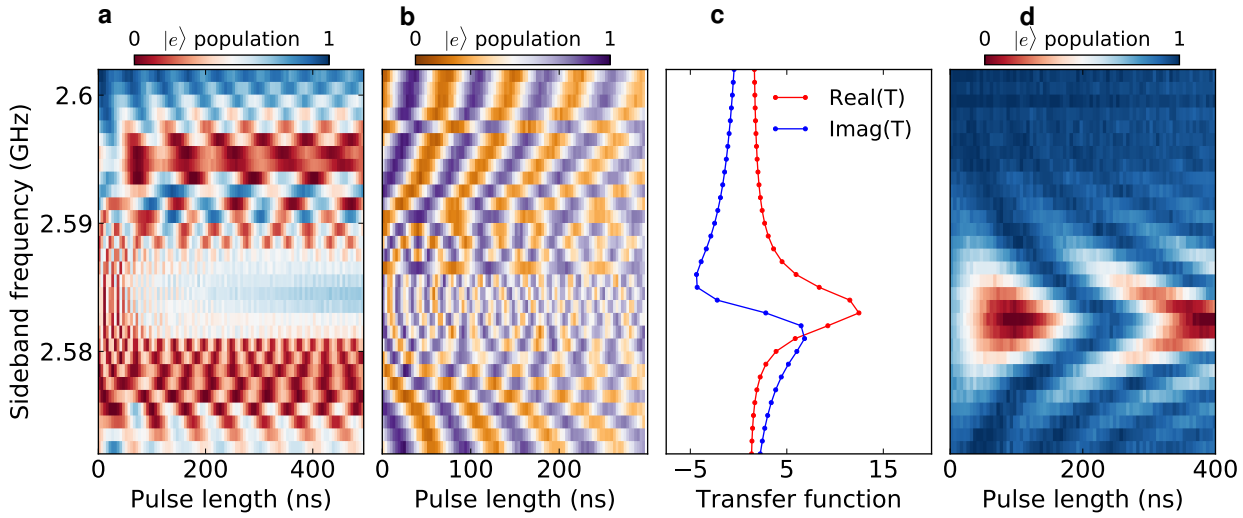

Figure 5.9: **a,** Stimulated vacuum Rabi spectrum on the $|g\rangle - |e\rangle$ transition near mode 10 with the transmon biased at 4.3325 GHz. **b,** Ramsey experiment on the $|g\rangle - |f\rangle$ transition with a flux-pulse inserted during the idle time. This allows for a precise measurement of the line attenuation near the modes. **c,** Real and Imaginary parts of the Transfer function obtained by fitting the experimentally measured transfer function amplitude to Equation (5.20). This form is automatically constrained to be causal, with the real and imaginary parts satisfying the Kramers-Kronig relations. **d,** Corrected spectrum, revealing a chevron pattern.

For short pulse durations and large stimulated vacuum Rabi rates, the bandwidth of the pulse becomes commensurate with the frequency scale over which the transfer function of the flux bias varies significantly, causing distortion of the flux pulses. However, this effect is corrected using the knowledge of the transfer function of the flux bias. The complete complex transfer function (characterizing the amplitude and the phase of the flux bias distortion) is extracted only from the amplitude of the transfer function, by assuming that the response of the line is causal and enforcing the Kramers-Kronig relations [186]. We enforce these relations here by fitting the amplitude of the

measured transfer function to the functional form:

$$|T(\nu)| = \left| y_0 + \sum_i \frac{A_i}{(\nu^2 - \nu_{0,i}^2) - i\gamma\nu} \right| . \tag{5.20}$$

We account for flux-pulse distortion by modifying the pulses generated by the AWG to account for the transfer function of the flux bias line. The AWG waveform used to generate a given pulse $f(t) = \mathrm{Re}[f_c(t)]$ at the location of the qubit is:

$$f_{\mathrm{AWG}}(t) = \mathrm{Re}\left[\mathrm{IFFT}\left(\frac{\mathrm{FFT}(f_c(t))}{T(\nu)}\right)\right] . \tag{5.21}$$

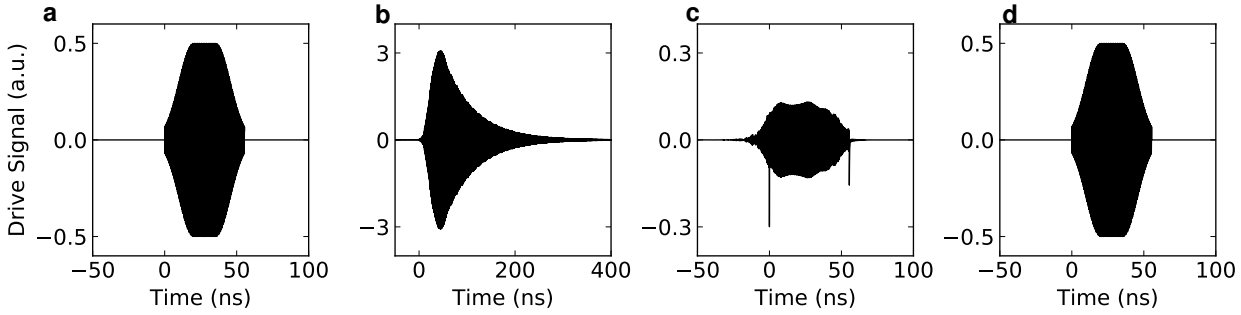

Figure 5.10: **a,** Ideal sideband pulse used perform a iSWAP between the transmon and mode 6. **b,** Expected distorted pulse at the location of the sample based on the measured transfer function of the flux line. **c,** Corrected AWG waveform calculated using Equation (5.21). **d,** Expected pulse at the location of the sample due to distortion of the pulse in **c** through the flux line.

5.12 Coupling between the transmon and the multimode memory

The rate of stimulated vacuum Rabi oscillations are related to the modulation strength and the bare coupling according to Equation (5.11) of the main text. We can therefore extract the bare couplings from the measured sideband Rabi oscillation rates, particularly since the strength of the modulation can be independently calibrated from spectroscopy of the DC-offset of the transmon frequency (see Chapter 5.11).

Instead, we measure the eigenmode-state dependent dispersive shift of the transmon frequency.

The shift for each mode k is measured with a transmon Ramsey interference experiment conducted after loading a photon into mode k, according to the protocol shown in Figure 5.11(a).

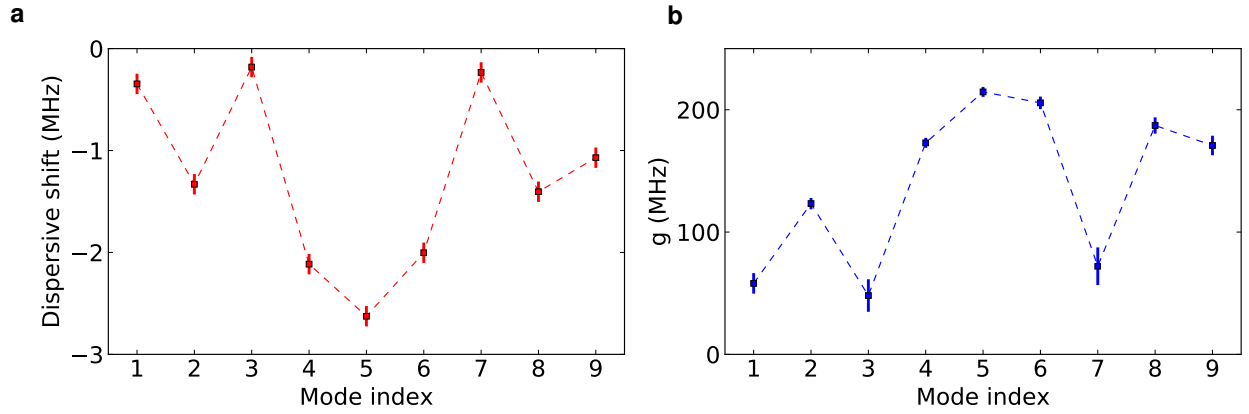

Figure 5.11: **a,** Dispersive shift χ_k as a function of mode index with the transmon biased at $\nu^{ge} = 4.3325$ GHz. The qubit frequency is stabilized to 50 kHz prior to each measurement and corresponds to the indicated error bar. **b,** The transmon-mode couplings g_k, extracted from the measured dispersive shift using equation (5.22), with the errors inferred from those in χ.

The dispersive shift is related to the measured oscillation frequency ν_{osc} and the Ramsey frequency ν_{Ram} according to $\delta_i = \nu_{\mathrm{osc}} - \nu_{\mathrm{Ram}}$, and is plotted as a function of mode number in Figure 5.11(b). We extract the coupling rate g_k from the measured dispersive shift χ_k, which are related by [97]:

$$\chi_k = \frac{g_k^2 \alpha}{\Delta_k(\Delta_k + \alpha)}, \tag{5.22}$$

where α is the transmon anharmonicity and $\Delta_k = \nu_q - \nu_k$ is the detuning between the transmon and mode k. The g_k's extracted from this expression are shown in Figure 5.11(c). The bare coupling rates extracted from the stimulated vacuum Rabi rate and from the dispersive shift are found to be consistent.

5.13 Hamiltonian tomography

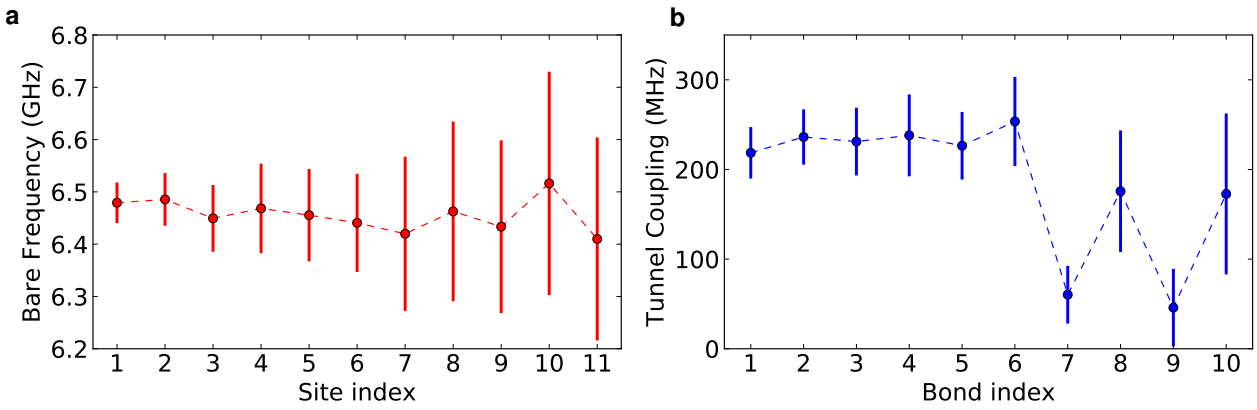

Figure 5.12: **a,** Bare resonator frequencies $\nu_{\mathrm{r},i}$ and **b,** nearest-neighbor tunnel couplings $g_{\mathrm{r},i}$ obtained from Hamiltonian tomography. The error bars in the bare frequencies and the couplings are obtained assuming a 25% uncertainty in the measured g_k's and a 0.5% uncertainty in the measured eigenfrequencies.

We use Hamiltonian tomography [111] to extract the $2N-1$ parameters of a chain of N nearest-neighbour coupled resonators. We assume Hamiltonian for this chain is given by equation (1) of the main text, but allowing disorder of the individual resonator frequencies ($\nu_{\mathrm{r},i}$) and tunnel couplings ($g_{\mathrm{r},i}$):

$$\hat{H} = \sum_{i=1}^{N} \nu_{\mathrm{r},i} \hat{a}_i^\dagger \hat{a}_i + \sum_{i=1}^{N-1} g_{\mathrm{r},i} \left(\hat{a}_{i+1}^\dagger \hat{a}_i + \mathrm{c.c.} \right) \tag{5.23}$$

We extract these parameters using the frequencies (ν_k) and couplings to the transmon (g_k) of the eigenmodes of the array ($2N$ numbers). The coupling to the transmon is proportional to the amplitude of the memory-mode wavefunction at the edge resonator ($g_{\mathrm{eff},k} \propto |\phi_1^k|$), where the creation operator for eigenmode k is $\hat{b}_k^\dagger = \sum_i \phi_i^k \hat{a}_i^\dagger$. The bare frequencies and tunnel-couplings of the resonator are then extracted by iteratively solving the Schrödinger equation starting from the transmon end of the chain, while imposing the constraints from wavefunction normalization

($\sum_k \left(\phi_i^k\right)\left(\phi_j^k\right)^* = \delta_{ij}, \sum_i \left(\phi_i^k\right)\left(\phi_i^q\right)^* = \delta_{kq}$), as shown below:

$$\nu_{r,i} = \sum_k \nu_k \left|\phi_i^k\right|^2 \quad , \quad \phi_i^k = \frac{\left(\nu_k - \nu_{\mathrm{r},i-1}\right)\phi_{i-1}^k - g_{\mathrm{r},i-2}\phi_{i-2}^k}{g_{\mathrm{r},i-1}} \quad , \tag{5.24}$$

$$g_{r,i}^2 = \sum_k \left(\nu_k - \nu_{\mathrm{r},i}\right)^2 \left|\phi_i^k\right|^2 - g_{\mathrm{r},i-1}^2.$$

The bare frequencies and tunnel couplings thus extracted are shown in Figure 5.12a and 5.12b. We infer that two of the normal modes (see Figure 5.7) being extremely weakly coupled is likely a result of coupler 7 and 9 being defective.

5.14 Coherence of the multimode memory

The coherence times of the memory modes are characterized through protocols analogous to those for the transmon, with the qubit pulses sandwiched with a pair of transmon-mode iSWAP pulses to transfer the quantum state between the transmon and the mode. The results of T_1 and T_2^* measurements for mode 1 of the memory are shown in Figure 5.13a and b, with the coherence times for all the modes, summarized in Figure 5.13c. The T_2^* is not found to be T_1 limited for many of the modes.

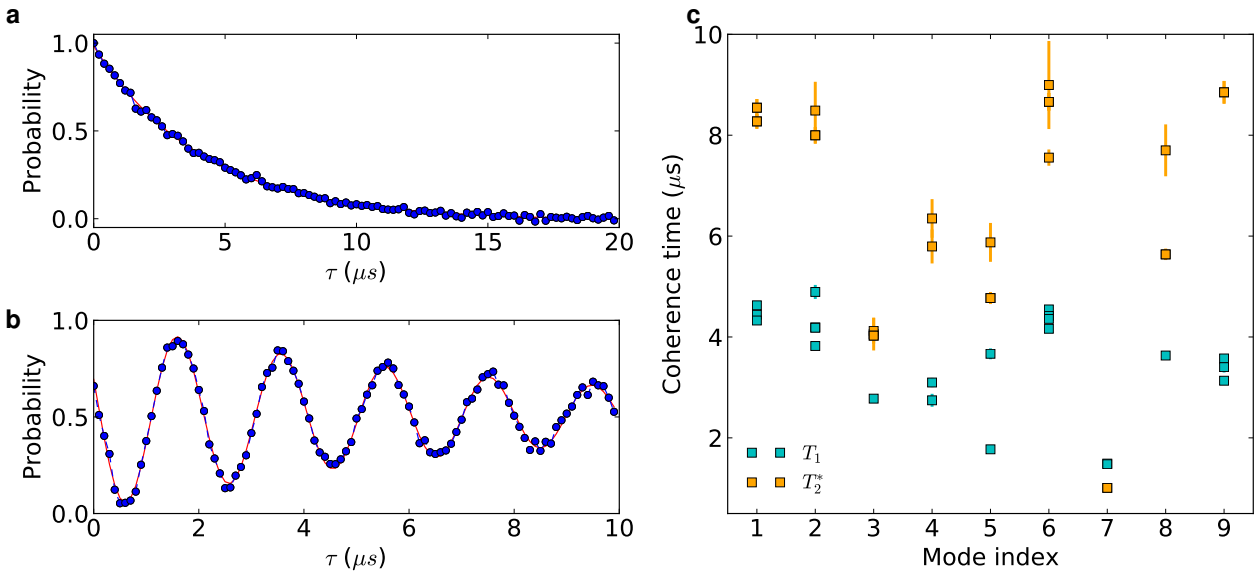

Figure 5.13: **a,** Single-photon lifetime measurement on mode 1. **b,** Ramsey (T_2^*) experiment on mode 1. The oscillation frequency of the Ramsey fringe can be used to infer the DC-shift of the transmon frequency during the iSWAP pulse. **c,** T_1 and T_2^* for the modes in the multimode memory. The error bars for individual points are as extracted from the fits. Because of insufficient statistics, the figure includes all the coherence data obtained with the transmon biased at $\nu_q \sim 4.59$ GHz.

5.15 Single-mode gate calibration

The frequency of the iSWAP pulse acting on a particular mode is obtained by choosing the frequency corresponding to maximum contrast of the stimulated vacuum Rabi chevrons, such as those in Figure 5.7. The amplitude and pulse bandwidth of the flux-modulation pulses are optimized to maximize the oscillation rate, while minimizing cross talk with neighboring modes and sideband transitions across other transmon levels. The length of an iSWAP pulse is obtained using fits of stimulated vacuum Rabi oscillations such as in Figure 5.2c. To achieve high fidelity gate operations, we also calibrate and correct phase errors arising during the sideband pulses.

The main phase error in the flux-modulation pulse is due to the DC-shift of the transmon frequency during flux-modulation ($\delta\nu_{\rm DC}$ in equation (5.7)). The frequency of the center of the stimulated vacuum Rabi chevron is detuned from the difference frequency between the mode and

the relevant transmon transition by $-\delta\nu_{\mathrm{DC}}$. Since the flux-pulse frequency is set to the center of the chevron, the clock (rotating frame) of the drive generator is shifted from the the frame of the Hamiltonian of equation 5.1. If the drive generator idles on resonance, there is an additional phase that accrues during that time. In the Ramsey experiment measuring the coherence time (T_2^*) of the modes (see Figure 5.13a (top)), the accrued phase shifts the frequency of the Ramsey fringes by $\delta\nu_{\mathrm{DC}}$. We can then account for the misalignment of clocks by advancing the phase of the subsequent pulse by $2\pi\delta\nu_{\mathrm{DC}}\tau$. This correction can be easily implemented by keeping the drive clock aligned with the bare qubit-resonator system when the flux pulse is off, and incrementing the drive frequency by $\delta\nu_{\mathrm{DC}}$ during the iSWAP pulse to bring it into resonance with the DC-shifted frame.

Fixing the drive clock to be in sync with the Hamiltonian of Equation (5.1) results in the absence of idle-time dependent phase errors. We additionally need to calibrate an additional dynamical phase (σ_z error) that occurs due to the change in the qubit frequency during the ramp up of the flux pulse. This phase is manifest in a rotating frame corresponding to the instantaneous qubit frequency $\bar{\nu}_{ge}(t)$ in equation (5.9). Repeating the transformation of equation (5.10) with a time-dependent qubit frequency results in an additional term:

$$\delta\hat{H} = -\frac{1}{2}h\frac{\partial\bar{\nu}_{ge}\left(t\right)}{\partial t}t\hat{\sigma}_z. \tag{5.25}$$

If we consider a square flux pulse with modulation amplitude corresponding to a DC-offset of ν_{DC} and pulse duration of t_π, the additional term in the Hamiltonian results in a dynamical phase of $\pi\nu_{\mathrm{DC}}t_\pi$. This error is calibrated using the sequence shown in Figure 5.14a and corrected by adjusting the relative iSWAP pulse phases. The result of this calibration for one of the memory modes is shown in Figure 5.14b. After calibrating the iSWAP phase ($\phi^{\pi-\pi}$), we add (subtract) $\frac{\phi^{\pi-\pi}}{2}$ to every iSWAP pulse for loading (unloading) an excitation into each of the memory modes. Subsequent $|g\rangle - |e\rangle$ iSWAP pulses in all circuit diagrams include this phase correction, and are represented by $\tilde{\pi}$ when represented in an equation.

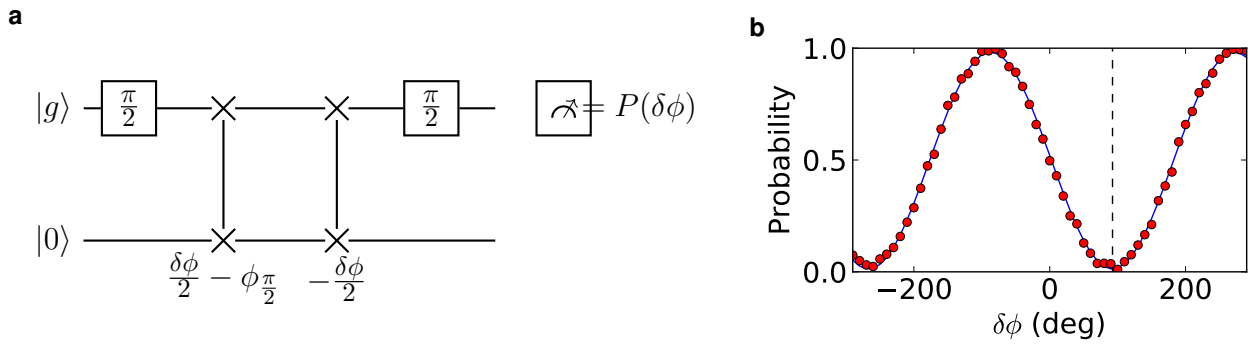

Figure 5.14: **a,** Circuit for calibrating phase of the iSWAP gate, where we sweep the phase $\frac{\delta\phi}{2}$ added and subtracted from the iSWAP pulses used to load and unload the state to the memory mode. We introduce an additional offset phase ($\phi_{\frac{\pi}{2}}$) which corrects σ_z errors occurring in the qubit pulse, ensuring that the mode state at the end of the first sideband pulse is free from all σ_z errors. The optimal phase for the iSWAP pulses is obtained by minimizing $P(\delta\phi)$. **b,** iSWAP phase calibration for mode 9 of the multimode memory, optimal phase ($92.5 \pm 0.1^\circ$) indicated by the black dashed line.

5.16 Randomized benchmarking

We characterize the fidelity of single-mode operations using randomized benchmarking (RB) [93]. As described in the main text (see Figure 5.3), single-mode Clifford gates are realized by sandwiching single-qubit Clifford gates (C_i) with transmon-mode iSWAP pulses.

$$\tilde{C}_i = U_{\tilde{\pi}_{\mathrm{sb}}} C_i U_{\tilde{\pi}_{\mathrm{sb}}(\phi=\pi)} = C_i$$

To load the excitation to and from the transmon, we use σ_z error corrected sideband iSWAP pulses that are 180° out of phase with each other, so that the mode Cliffords are mapped directly from their transmon qubit counterparts. The Cliffords are generated by concatenating an operator each from $\{0, \frac{\pi}{2}_y, \pi_y, -\frac{\pi}{2}_y\}$ and $\{0, \frac{\pi}{2}_x, \pi_x, -\frac{\pi}{2}_x, \frac{\pi}{2}_z, -\frac{\pi}{2}_z\}$, to generate all 24 elements of the single qubit Clifford group. The circuit showing the sequence used for RB of the modes is shown in Figure 5.15.

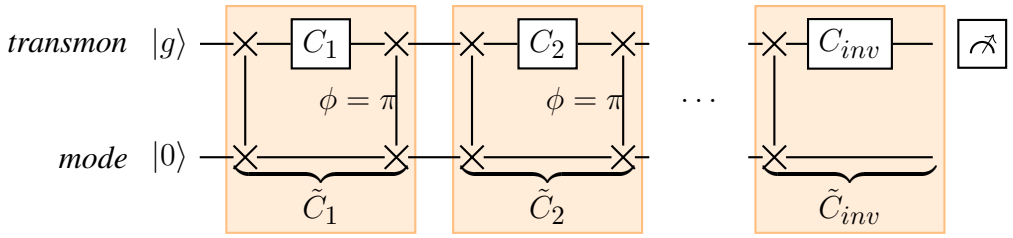

Figure 5.15: Randomized benchmarking (RB) characterizes the average random gate fidelity by acting random sequences of single-mode Clifford gates of increasing length, inverting the sequence, and then measuring the qubit state as a function of length. The sequence is acted on the multimode system initialized in the ground state, with the mode occupation error (ϵ) measured at the end of each sequence. The RB fidelity is extracted from the decay of the occupation fidelity ($1-\epsilon$) as a function of the length of the benchmarking sequence. For the data shown in Figure 5.3, we use sequence lengths corresponding to those in [93, 31] and average over 32 random sequences.

The RB fidelity (p) is extracted by fitting the decay curves to the form $Ap^m + B$, where m is the sequence length. We estimate the coherence limit to the RB fidelity to be:

$$p_i = p_q - 2\left(1 - \exp\left[-\frac{t^{\pi}_{\mathrm{sb},k}}{T_{1,k}}\right]\right), \tag{5.26}$$

where $t^{\pi}_{sb,k}$ and $T_{1,k}$ are the iSWAP durations and lifetimes, respectively, of mode k, and p_q is the RB fidelity of the transmon. The experimentally measured fidelities approach these coherence limits (Figure 5.3).

5.17 Two mode gates using sideband transitions

The level diagram describing the relevant multimode states and transitions involved in the CZ gate are shown in Figure 5.16a. We break the 2π $|e\rangle - |f\rangle$ sideband pulse used to provide a conditional phase between the transmon and the second mode ($|e1\rangle \rightarrow -|e1\rangle$, see Figure 5.4a of main text) into two π pulses. Control of the relative phases between these pulses allows for the correction of additional phase errors arising from the dispersive shift and the realization of an arbitrary controlled phase gate.

We obtain a CNOT gate by inserting an $|e\rangle - |f\rangle$ transmon charge π pulse (π^{ef}_{q}) between the two $|e\rangle - |f\rangle$ sideband iSWAP pulses. This allows for mapping $|e0\rangle$ to $|e1\rangle$ (and vice versa) via

the $|f0\rangle$ (qubit-mode CNOT), which again becomes a mode-mode CNOT gate when sandwiched with two $|g\rangle - |e\rangle$ sideband iSWAP pulses. The pulse sequence, energy level diagram, and relevant transitions for the CNOT gate are shown in Figure 5.16b.

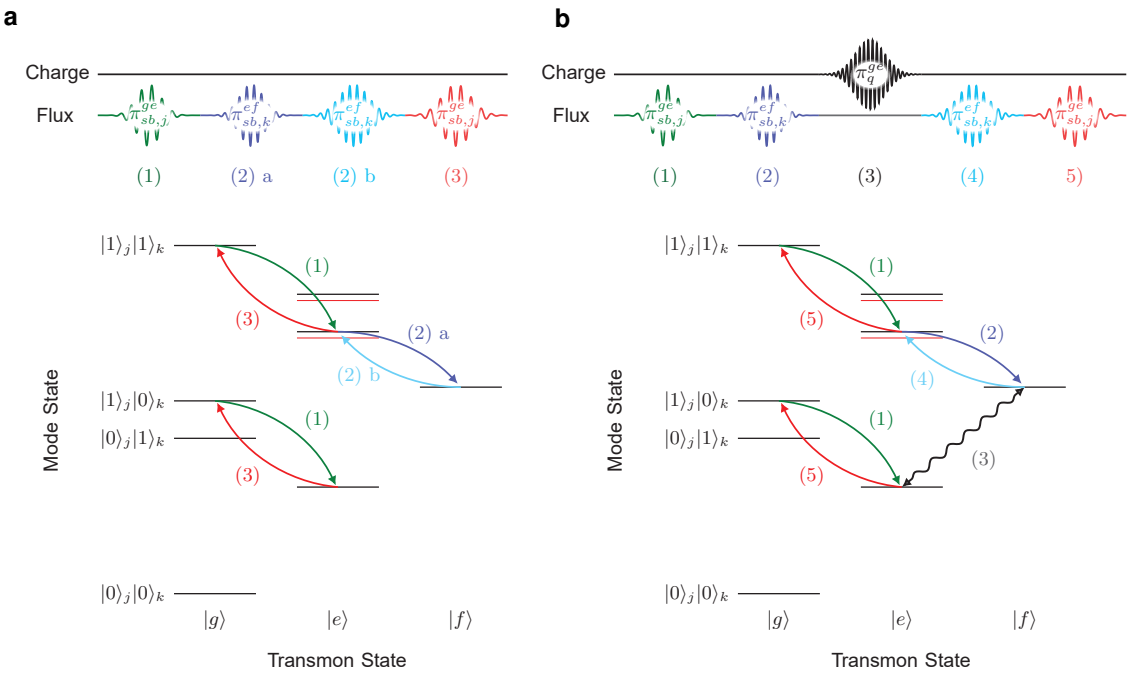

Figure 5.16: **a and b,** Energy level diagrams showing the multimode states and pulses involved in CZ and CNOT gates, respectively. The $e^{i\pi}$ phase factor arising from sideband SWAP operations between $|e01\rangle \longleftrightarrow |f00\rangle$ is modified by the dispersive shift (χ) of $|e01\rangle$ level (red). The additional phase arising from the dispersive shift is corrected by adjusting the phase of an ef sideband pulse on mode 1. CY gates are realized by adjusting the phase of the π_{q}^{ef} pulse.

Slight modifications of these pulse sequences allow the realization of other two-mode gates such as the mode-mode CY and SWAP gates. The pulse sequences (without corrections from the dispersive shift) for realizing these two-mode gates are summarized in Table 5.1.

5.18 Phase error corrections to two-mode gates

The previous discussion of two-mode gates only involved resonant first-order sideband transitions and ideal transmon charge drive pulses. This idealized description is corrected by additional terms in the Hamiltonian of equation (5.12). The dominant additional effects are from: (1) dispersive

Two-mode gate	Pulse Sequence
$\mathrm{CZ_{j,k}}$	$\pi^{ge}_{\mathrm{sb},j} + \pi^{ef}_{\mathrm{sb},k} + \pi^{ef}_{\mathrm{sb},k} + \pi^{ge}_{\mathrm{sb},j}(\phi = \pi)$
$\mathrm{CX_{j,k}}$	$\pi^{ge}_{\mathrm{sb},j} + \pi^{ef}_{\mathrm{sb},k} + \pi^{ef}_{\mathrm{q},y} + \pi^{ef}_{\mathrm{sb},k} + \pi^{ge}_{\mathrm{sb},j}(\phi = \pi)$
$\mathrm{CY_{j,k}}$	$\pi^{ge}_{\mathrm{sb},j} + \pi^{ef}_{\mathrm{sb},k} + \pi^{ef}_{\mathrm{q},x} + \pi^{ef}_{\mathrm{sb},k} + \pi^{ge}_{\mathrm{sb},j}(\phi = \pi)$
$\mathrm{SWAP_{j,k}}$	$\pi^{ge}_{\mathrm{sb},j} + \pi^{ef}_{\mathrm{sb},k} + \pi^{ge}_{\mathrm{sb},k} + \pi^{ef}_{\mathrm{sb},k} + \pi^{ge}_{\mathrm{sb},j}$

Table 5.1: Pulse sequences used for realizing various two-mode gates. j and k are indices corresponding to the control and target mode, respectively.

shifts arising from photons in the multimode memory, (2) the qubit DC-offset due to flux modulation, and (3) phases from AC Stark shifts due to off-resonant first-order sidebands. These shifts result in corrections to the transmon rotation and transmon-mode iSWAP unitaries.

For the case of the dispersive shift, the corrections to the target unitaries depend on the quantum state of the multimode memory and result in a transmon-mode ZZ error. If we ignore photons in the rest of the memory, the effect of the dispersive shift on ther modes involved in a two-mode entangling gate can be inferred from the energy level diagram of Figure 5.16. The dispersive shift causes the $|e10\rangle$ and $|e01\rangle$ levels to be shifted (red) from their bare values (black). As a result, sidebands resonant with $|e00\rangle \longleftrightarrow |g10\rangle$ are off-resonant from $|e01\rangle \longleftrightarrow |g11\rangle$.

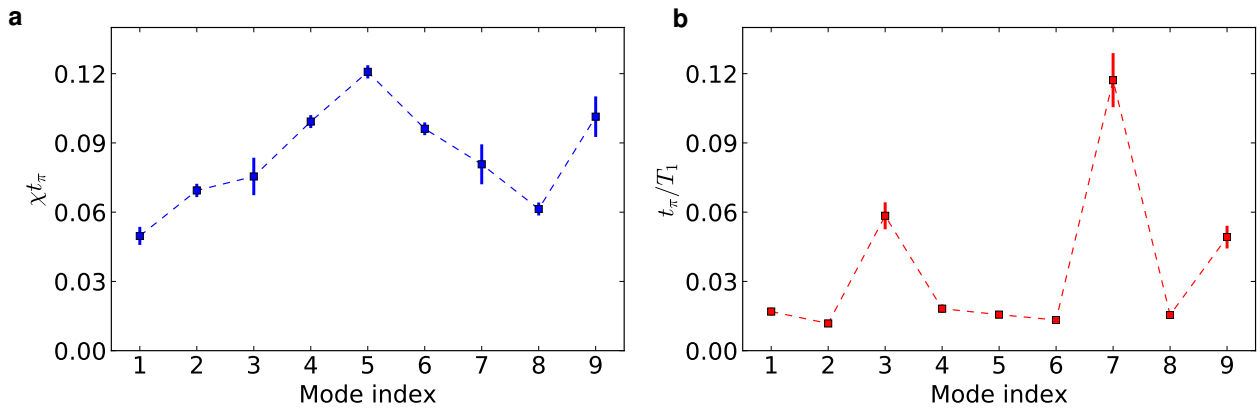

Figure 5.17: **a,** Error arising from the dispersive shift χ during an iSWAP pulse of duration t_π. The linear order phase errors ($\sim 25 - 60^\circ$ over the course of a CZ gate) are corrected during the gate, leaving residual amplitude error ($\propto (\chi t_\pi)^2$). The error bar in χt_π is dominated by the error in the measurement of χ. **b,** Error arising from T_1 loss during a sideband, plotted for comparison to the dispersive shift error. The $T_1's$ are the mean of the data in Supplemetary Figure 5.13, with the error in $\frac{t_\pi}{T_1}$ extracted from a 10% uncertainty in the T_1 of each mode.

The dispersive shift results in a phase and population error in the $|e01\rangle$ state. For the dispersive shifts χ and iSWAP times t_π used in this work, this population error is $\propto (\chi t_\pi)^2$ and, at worst, $\sim 5\%$ over the course of a CZ gate. This error is uncorrected and factors into the total gate error. The phase error on other hand is $\propto (\chi t_\pi)$ and results in a more significant effect (see Figure 5.17a). Given that the $|e01\rangle$ state affected by the dispersive shift is selectively addressed by the $|e\rangle - |f\rangle$ sideband pulses used in the gate (see Figure 5.16), this phase error is calibrated and corrected by adjusting the relative phase between these pulses. The experimental protocols used for the gate calibration and phase error correction for the CZ gate are described in Note 5.19. The state dependent phases arising in the gate can be calculated by considering the effective Hamiltonian of

equation (5.15) in the 8×8 subspace of levels relevant for the gates and shown in Figure 5.16:

$$\tilde{H}(t) = \begin{pmatrix} 0 & 0 & 0 & 0 & \Omega_{ge}^* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & g_{k,ge}^* & \Omega_{ge}^* & 0 & 0 \\ 0 & 0 & 0 & 0 & g_{j,ge}^* & 0 & \Omega_{ge}^* & 0 \\ 0 & 0 & 0 & 0 & 0 & g_{j,ge}^* & g_{k,ge}^* & 0 \\ \Omega_{ge} & g_{k,ge} & g_{j,ge} & 0 & 0 & 0 & 0 & \Omega_{ef}^* \\ 0 & \Omega_{ge} & 0 & g_{j,ge} & 0 & \delta_k & 0 & g_{k,ef}^* \\ 0 & 0 & \Omega_{ge} & g_{k,ge} & 0 & 0 & \delta_j & g_{j,ef}^* \\ 0 & 0 & 0 & 0 & \Omega_{ef} & g_{k,ef} & g_{j,ef} & 0 \end{pmatrix} \begin{matrix} |g\rangle \otimes \left|0_j 0_k\right\rangle \\ |g\rangle \otimes \left|0_j 1_k\right\rangle \\ |g\rangle \otimes \left|1_j 0_k\right\rangle \\ |g\rangle \otimes \left|1_j 1_k\right\rangle \\ |e\rangle \otimes \left|0_j 0_k\right\rangle \\ |e\rangle \otimes \left|0_j 1_k\right\rangle \\ |e\rangle \otimes \left|1_j 0_k\right\rangle \\ |f\rangle \otimes \left|0_j 0_k\right\rangle \end{matrix} \tag{5.27}$$

Here, the multimode state is labeled $\left|n_j, n_k\right\rangle$ and the phases of equation (5.17) have been absorbed into the g's and Ω's (which are time dependent), i.e.,

$$g_{i,\alpha} \to g_{i,\alpha} e^{i\phi_{\mathrm{sb},\alpha}} \quad , \quad \Omega_\alpha \to \Omega_\alpha e^{-i\phi_{\mathrm{q},\alpha}} \quad , \quad \alpha \in \{ge, ef\} \quad , \quad i \in \{j, k\}\,. \tag{5.28}$$

The $|e\rangle - |f\rangle$ sideband pulses act only on one transition and are unaffected by the state dependent shift when considering only two modes. We chose the $|e\rangle - |f\rangle$ sideband freqeuncy to be resonant with the $|f00\rangle$ and the dispersively shifted $|e01\rangle$ level. In the rotating frame of equation (5.27), this corresponds to $|e\rangle - |f\rangle$ first-order sidebands acquiring the following time-dependence:

$$g_{j,ef}(t) = \tilde{g}_{j,ef} e^{-2\pi i \delta_k t} \quad , \quad g_{k,ef}(t) = \tilde{g}_{k,ef} e^{-2\pi i \delta_j t}. \tag{5.29}$$

$\tilde{g}_{i,ef}$ is proportional to the envelope of the $|e\rangle - |f\rangle$ sideband pulse, and δ_k and δ_j are the dispersive shifts of $\left|e0_j 1_k\right\rangle$ and $\left|e1_j 0_k\right\rangle$ respectively.

We compute the action of the CZ and CNOT gate sequences by evolving the Hamiltonian above, with time dependent coefficients and phases as per Table 5.2.

In these pulse sequences, only one of the drive terms is on at any given time and the corresponding unitaries obtained upon integration of the Schrödinger equation are generalizations of

Gate	Pulse Sequence
CZ	$\tilde{\pi}^{ge}_{\mathrm{sb},j}(\phi_\mathrm{a}) + \tilde{\pi}^{ef}_{\mathrm{sb},k}(\phi_\mathrm{b}) + \tilde{\pi}^{ef}_{sb,k}(\phi_\mathrm{c}) + \tilde{\pi}^{ge}_{\mathrm{sb},j}(\phi_\mathrm{d})$
CNOT	$\tilde{\pi}^{ge}_{\mathrm{sb},j}(\phi_\mathrm{a}) + \tilde{\pi}^{ef}_{\mathrm{sb},k}(\phi_\mathrm{b}) + \pi^{ef}_{q}(\phi_\mathrm{e}) + \tilde{\pi}^{ef}_{\mathrm{sb},k}(\phi_\mathrm{c}) + \tilde{\pi}^{ge}_{\mathrm{sb},j}(\phi_\mathrm{d})$

Table 5.2: Nomenclature for the pulse phases used in the CZ and CNOT gates resulting in the unitary operators in equation (5.31).

those in Equation (5.20), with corrections arising from the dispersive shift. The effective unitary thus realized for the CZ and CNOT gates, to lowest order in χ/g_{sb} and χ/Ω are;

$$\mathrm{U_{CZ}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -e^{i(\phi_\mathrm{a}-\phi_\mathrm{d})} & 0 \\ 0 & 0 & 0 & e^{i\left(\phi_\mathrm{a}-\phi_\mathrm{d}+\phi_\mathrm{b}-\phi_\mathrm{c}-2\pi t^{\pi}_{\mathrm{sb},j,ge}\delta_k\right)} \end{pmatrix} \tag{5.30}$$

$$\mathrm{U_{CNOT}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & e^{i\left(\phi_\mathrm{a}-\phi_\mathrm{d}+\phi_\mathrm{b}+\phi_\mathrm{e}-\pi t^{\pi}_{\mathrm{sb},j,ge}\delta_k\right)} \\ 0 & 0 & e^{i\left(\phi_\mathrm{a}-\phi_\mathrm{d}-\phi_\mathrm{c}-\phi_\mathrm{e}-\pi t^{\pi}_{\mathrm{sb},j,ge}\delta_k\right)} & 0 \end{pmatrix}$$

We see that one can choose sideband pulse phases that cancel phases arising from the dispersive shift and thereby realize the target unitaries for the CZ and CNOT gates.

5.19 CZ gate phase calibration sequences

Here, we describe protocols used to calibrate and correct each of the additional phases arising in the CZ gate. The phases of the iSWAP pulses used in the CZ gate are defined below, where $\phi_{1,2}$ are the controlled phases:

$$CZ_{j,k}(\phi_1,\phi_2) = \tilde{\pi}^{ge}_{\mathrm{sb},j} + \pi^{ef}_{\mathrm{sb},k}(\phi_1) + \pi^{ef}_{\mathrm{sb},k}(0) + \tilde{\pi}^{ge}_{\mathrm{sb},j}(\phi_2). \tag{5.31}$$

Sequence	Phase error	Initial state	# CZ's	Measured mode
a	Dispersive shift (SPAM)	$\|10\rangle + \|11\rangle$	1	Target (k)
b	Dispersive shift (Gate)	$\|10\rangle + \|11\rangle$	2	Target (k)
c	DC-offset during gate	$\|01\rangle + \|11\rangle$	1	Control (j)
d	Off-resonant sidebands	$\|00\rangle + \|01\rangle$	1	Target (k)

Table 5.3: Summary of CZ gate phase calibration experiments that correct for each of the sources of the phase errors. The initial states used to calibrate each of the phase errors are indicated.

j is the control mode and k is the target mode of the CZ gate, with the states labeled $\left|n_j, n_k\right\rangle$, and $\tilde{\pi}$ indicating iSWAP pulses for which the DC-offset σ_z error is corrected within the pulse. From equation (5.30), we see that only two relative phase adjustments ($\phi_{\mathrm{a}} - \phi_{\mathrm{d}}$ and $\phi_{\mathrm{b}} - \phi_{\mathrm{c}}$) are required to correct the dispersive shift error. Here, we adjust these relative phases by controlling $\phi_1 = \phi_{\mathrm{b}}$ and $\phi_1 = \phi_{\mathrm{d}}$, while leaving ϕ_{a} and ϕ_{c} fixed. We measure each phase error through Ramsey experiments with initial states that are appropriate superpositions of the basis states, as indicated in Table 5.3.

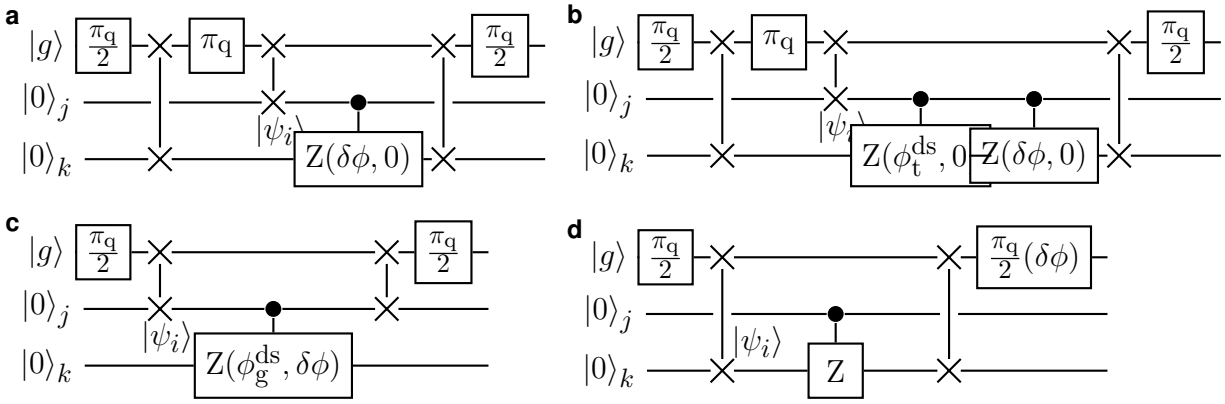

Figure 5.18: CZ gate calibration sequences. **a,** Measures phase error from the dispersive shift arising from the entire sequence. **b,** Isolates the error from the dispersive shift error occuring only during the gate. **c,** Measures the phase error arising from the qubit dc-offset occuring during the gate. **d,** Phase error arising from the AC Stark shift due to off-resonant first order sidebands.

The phase error from the dispersive shift is obtained by preparing the system in the state $\left|\psi_p\right\rangle = |10\rangle + |11\rangle$. We measure the relative phase between the basis states after applying the CZ gate, using the sequence in Figure 5.18a. The dispersive shift results in the $|11\rangle$ acquiring an additional phase

in the preparation (ϕ_{p}), gate (ϕ_{g}) and measurement segments (ϕ_{m}). Similar additional phases also accrue during the gate (ϕ_{g}) and the measurement (ϕ_{m}) segments. We sweep the phase ($\delta\phi$) of the first $|e\rangle - |f\rangle$ sideband pulse of the CZ gate (see Equation (5.31)). The phase that *maximizes* the final measured transmon population provides the total added phase, $\phi_{\mathrm{t}}^{\mathrm{ds}} = \phi_{\mathrm{p}} + \phi_{\mathrm{g}} + \phi_{\mathrm{m}}$. $CZ_{j,k}(\phi_{\mathrm{t}}^{\mathrm{ds}}, 0)$ is a combination of an ideal CZ gate and the $C\phi$ gate that cancels additional phases arising from the dispersive shift over the entire sequence. We isolate the state-dependent phase error arising only during the CZ gate by adding a second $\mathrm{C}\phi$ gate to the previous sequence, as shown in Figure 5.18b. We sweep the phase ($\delta\phi$) of the (first) $|e\rangle - |f\rangle$ sideband pulse of the second $\mathrm{C}\phi$ gate, with $\phi_1 = \phi_{\mathrm{ds}}^{\mathrm{t}}$ for the first $\mathrm{C}\phi$ gate. Given the same preparation and measurement sequences, the SPAM phases are corrected by construction by the first $\mathrm{C}\phi$ gate. We find the phase $\delta\phi$ that *minimizes* the population of the transmon, thus realizing a CZ gate that flips the sign of the $|11\rangle$ state. The $CZ_{k,j}(\phi_{\mathrm{g}}, 0)$ gate therefore is corrected for phases from dispersive shifts occurring during the gate sequence.

We obtain a fully corrected CZ gate by correcting the relative phase between the $\{|00\rangle, |01\rangle\}$ and $\{|10\rangle, |11\rangle\}$ manifolds. These state manifolds have a relative phase resulting from the transmon frequency DC-offset occurring during the $|e\rangle - |f\rangle$ sidebands of the CZ gate. We correct this additional phase by adjusting the phase of the final $|g\rangle - |e\rangle$ sideband pulse of the CZ gate (ϕ_2 in equation (5.31)), using the experimental sequence of Figure 5.18c. The resulting $\mathrm{CZ}_{\mathrm{j,k}}(\phi_{\mathrm{g}}^{\mathrm{ds}}, \phi^{\mathrm{DC}})$ gate is therefore corrected of errors from dispersive shifts and qubit flux-modulation DC-offsets.

The phase error resulting from dispersive shifts due to off-resonant first-order sidebands are measured by acting the CZ gate on the $|00\rangle + |01\rangle$ state. The CZ gate nominally does not change this state, and we correct this phase error using subsequent qubit pulses as shown in Figure 5.18d. This phase is significant only for gates between modes with spectral spacing near the anharmonicity of the transmon.

5.20 Two-mode quantum state tomography

Reconstructing the density matrix of an arbitrary two-qubit state requires the measurement of all possible two-qubit correlations $\{\langle \mathrm{XI}\rangle, \langle \mathrm{XX}\rangle \ldots \langle \mathrm{ZZ}\rangle\}$, i.e.;

$$C_{i,j} = \langle B_i \otimes B_j \rangle \;\mid\; B_i \in \{\mathrm{I}, \mathrm{X}, \mathrm{Y}, \mathrm{Z}\} \tag{5.32}$$

These correlators can be measured through Ramsey interferometry, as described in the main text [54]. We equivalently measure all the necessary correlations with the aid of the single and two-mode gate operations prior to measuring the state of the transmon. A sideband iSWAP pulse (π_{sb}) on the $|g\rangle - |e\rangle$ transition, along with single qubit rotations alone can be used to measure all single-mode correlators $C_{ij} \in \{\langle B_i \otimes I\rangle \text{ or } \langle I \otimes B_i\rangle\} \mid B_i \in \{\mathrm{X}, \mathrm{Y}, \mathrm{Z}\}$.

The entanglement information is present in two-mode correlators,

$$C_{i,j} = \langle \mathrm{B_i} \otimes \mathrm{B_j}\rangle \mid \mathrm{B_i} \in \{\mathrm{X}, \mathrm{Y}, \mathrm{Z}\}.$$

We measure these correlators by acting two-mode gates before measuring a single-mode correlators. For instance, the $\langle XX\rangle$ correlator of a given state ($|\psi_i\rangle$) is measured by acting CX gate prior to the measurement of $\langle \mathrm{XI}\rangle$. In the Heisenberg picture [67], the transformation is shown below:

$$C = \langle \psi_f | \, X \otimes I \, | \psi_f \rangle = \langle \psi_i | \, U_{\mathrm{CX}}^{\dagger} \, (X \otimes I) \, U_{\mathrm{CX}} \, | \psi_i \rangle = \langle \psi_i | \, X \otimes X \, | \psi_i \rangle \tag{5.33}$$

Here $|\psi_i\rangle$ is the two-mode state to be measured and $|\psi_f\rangle = U_{\mathrm{CX}} \, |\psi_f\rangle$ is the state obtained following action of the CX gate. A summary of pulse sequences used for the measurement of each of the correlations required for two-mode tomography are shown in Table 5.4.

#	Measured Correlation	Pulse Sequence
0	$-\langle \mathrm{IX} \rangle$	$\pi_{\mathrm{sb},k} + \frac{\pi}{2}_y$
1	$\langle \mathrm{IY} \rangle$	$\pi_{\mathrm{sb},k} + \frac{\pi}{2}_x$
2	$\langle \mathrm{IZ} \rangle$	$\pi_{\mathrm{sb},k}$
3	$-\langle \mathrm{XI} \rangle$	$\pi_{\mathrm{sb},j} + \frac{\pi}{2}_y$
4	$-\langle \mathrm{XX} \rangle$	$\mathrm{CX} + \pi_{\mathrm{sb},j} + \frac{\pi}{2}_y$
5	$-\langle \mathrm{XY} \rangle$	$\mathrm{CY} + \pi_{\mathrm{sb},j} + \frac{\pi}{2}_y$
6	$\langle \mathrm{XZ} \rangle$	$\mathrm{CZ} + \pi_{\mathrm{sb},j} + \frac{\pi}{2}_y$
7	$\langle \mathrm{YI} \rangle$	$\pi_{\mathrm{sb},j} + \frac{\pi}{2}_y$
8	$\langle \mathrm{YX} \rangle$	$\mathrm{CX} + \pi_{\mathrm{sb},j} + \frac{\pi}{2}_x$
9	$\langle \mathrm{YY} \rangle$	$\mathrm{CY} + \pi_{\mathrm{sb},j} + \frac{\pi}{2}_x$
10	$-\langle \mathrm{YZ} \rangle$	$\mathrm{CZ} + \pi_{\mathrm{sb},j} + \frac{\pi}{2}_x$
11	$\langle \mathrm{ZI} \rangle$	$\pi_{\mathrm{sb},1}$
12	$-\langle \mathrm{ZX} \rangle$	$\mathrm{CZ} + \pi_{\mathrm{sb},k} + \frac{\pi}{2}_y$
13	$\langle \mathrm{ZY} \rangle$	$\mathrm{CZ} + \pi_{\mathrm{sb},k} + \frac{\pi}{2}_x$
14	$\langle \mathrm{ZZ} \rangle$	$\mathrm{CX} + \pi_{\mathrm{sb},k}$

Table 5.4: Pulse sequences used for the measurement of all two-qubit correlations between mode pairs. These correlations are used to reconstruct a two-qubit density matrix using equation 5.34.

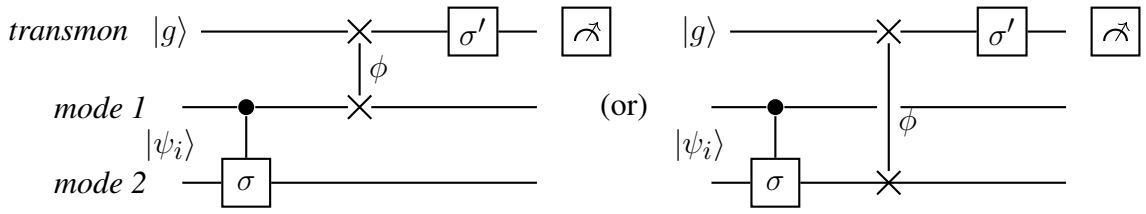

Figure 5.19: General two-mode correlator measurement sequence, where $\sigma, \sigma' \in \{\mathrm{I}, \mathrm{X}, \mathrm{Y}, \mathrm{Z}\}$. To measure and correct additional phase shifts (as described in Note 5.19) arising in the tomography sequence, we sweep the phase, ϕ, of the final sideband pulse of the sequence used to measure each correlator.

We extract the correlators and construct the density matrix of the two-mode state from the measured transmon population P_{ij} at the end of the sequence for each correlator C_{ij} using;

$$\begin{aligned} C_{i,j} &= \langle B_i \otimes B_j \rangle = 2P_{ij} - 1, \\ \rho &= \sum_{ij} \frac{C_{ij} B_i \otimes B_j}{4}. \end{aligned} \tag{5.34}$$

In general, fast measurement and reset [61] of the transmon would allow us to perform sequential measurements of two-mode correlations using the transmon without requiring mode-entangling gate operations. For each mode, we would map the mode state to the transmon with an iSWAP, measure the transmon, and reset it to the ground state. The transmon state could be reset with an iSWAP back to the measured mode or to an auxiliary mode. The transmon can subsequently be used to measure the next mode. Additionally, we can perform Wigner tomography [71] of the multimode chain through direct measurements of the multimode fields and parametric amplification. These techniques pose more stringent conditions on the measurement fidelity and speed and are beyond the scope of this work.

We account for spurious phases arising in the Bell state tomography sequence by varying the phase of the final sideband pulse used in each correlator measurement. The results of such phase sweeps for the $|\Phi^+\rangle$ and $|\Psi^+\rangle$ Bell states are shown in Figure 5.24a and b. We note that for these states, the only correlators that are functions of the final sideband phase are $\mathrm{XX}, \mathrm{XY}, \mathrm{YX}, \mathrm{YY}$ and the $|\Phi^+\rangle$ and $|\Psi^+\rangle$ Bell states give opposite answers for measurements of the two-mode parity

ZZ. The optimal sideband phase that accounts for the additional phases shifts are indicated by the dashed lines. We extract the density matrices from the results of these measurements using equation (5.34). The state fidelities for the two states are calculated from the overlap of the ideal ρ_{id} and measured ρ_{m} density matrices: $F_p = \mathrm{Tr}(\rho_{\mathrm{id}}\rho_{\mathrm{m}})$.

5.21 Process tomography of two-mode gates

Process tomography of a two-qubit gate consists of quantum state tomography after acting the gate on a set of 16 linearly independent input states that form a basis for representing an arbitrary two-qubit density matrix [140]. Process tomography of two-mode gates therefore consists of a set of 240 measurement sequences of the form shown in Figure 5.20.

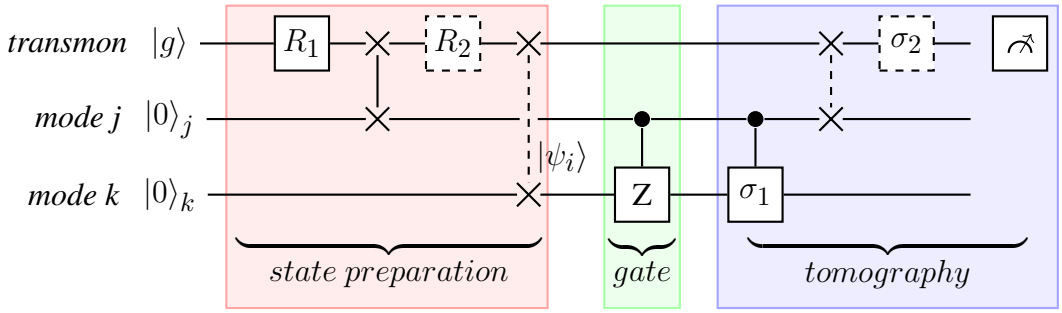

Figure 5.20: Process tomography sequence for two-mode gates, broken down into preparation (red), gate (green), and tomography (blue) segments. For the preparation sequence, we use qubit rotations $R_{1,2} = \left\{I, R_y\left(\frac{\pi}{2}\right), R_x\left(\frac{\pi}{2}\right), R_x\left(\pi\right)\right\}$ and DC-offset corrected sideband iSWAP pulses with an additional $-\frac{\pi}{2}$ phase, such that the target multimode state at the end of the preparation sequence is $|\psi_i\rangle = R_1 \otimes R_2 \left|0_j 0_k\right\rangle$. We measure the density matrix of the gate outputs for given input density matrices $\rho_i = |\psi_i\rangle\langle\psi_i|$ using the state tomography protocols of Chapter 5.20 and the sequences in Table 5.4, corresponding to $\sigma_{1,2} = \{I, X, Y, Z\}$ in the sequence shown above. We note that the iSWAP gate acts on mode k for some of the correlators, and that the tomography sequences that measure single-mode correlators have no additional two-mode gate, corresponding to $\sigma_1 = \nVdash$. The qubit and iSWAP operations that are indicated by the dashed lines have errors arising from the dispersive shift.

The gate calibration protocols described in Chapter 5.19 for the CZ gate, and analogous protocols for the CX gate, correct phase errors due to dispersive shifts and the transmon DC-offset from flux modulation during the gate. We additionally correct errors arising from the dispersive shift during the state preparation and tomography (SPAM) segments of the various process tomography

sequences. These errors occur in the qubit and iSWAP operations indicated by dashed lines in Figure 5.20. The dispersive shift causes amplitude and phase errors in the transmon and sideband pulses. We again correct only phase errors to first-order in χ/Ω_{sb}. These controlled-phase errors can be formally incorporated as $\mathrm{C}\Phi$ gates at the end of the preparation sequence ($\mathrm{C}\Phi_\mathrm{p}$) and prior the to measurement sequence ($\mathrm{C}\Phi_\mathrm{m}$). These additional gates are concatenated into the gate and tomography sequences as shown in Figure 5.21.

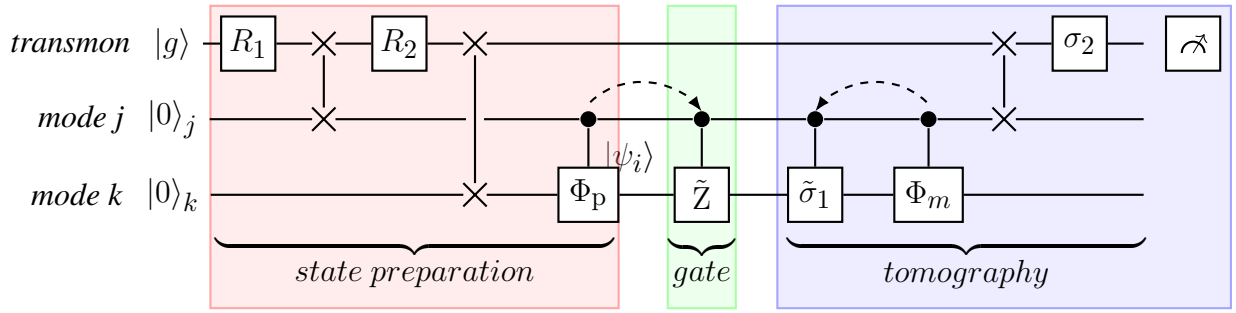

Figure 5.21: Protocol for correcting errors from the dispersive shift, in state preparation and measurement during process tomography of multimode gates.

$\mathrm{C}\tilde{\mathrm{Z}}$ and $\mathrm{C}\tilde{\sigma}_1$ are chosen to give phase-corrected CZ and $\mathrm{C}\sigma_1$ gates when concatenated with $\mathrm{C}\Phi_\mathrm{p}$ and $\mathrm{C}\Phi_\mathrm{m}$, respectively. The preparation error is corrected through an added phase (ϕ_p) in the first $|e\rangle - |f\rangle$ sideband of the first gate, while the tomography error is corrected through an added phase (ϕ_m) in the second $|e\rangle - |f\rangle$ sideband of the last gate of the sequence. We thereby correct the sequence to first-order in the dispersive-shift error. The sideband phases are chosen in this manner in order to correct errors in both the $|10\rangle$ and $|11\rangle$ states (see equation (5.30), (5.31)). The phase errors depend on the duration of the qubit and sideband pulses used in the sequence. In the absence of loss, they can be calculated based on the dispersive shift and pulse shapes.

We calibrate the additional phase errors through process tomography of the Identity ($\mathbb{1}$) gate (idling for 10 ns). We find the optimal phases by sweeping the added controlled-phases of the $\mathrm{C}\phi$ and $\mathrm{C}\sigma$ gates, and comparing results for corresponding correlators with and without CX/CY gates (such as XX and XI, respectively) as shown in Figure 5.22.

This scheme allows us to isolate state preparation and measurement errors (ϕ_p and ϕ_m). In the Heisenberg picture, working backward from the transmon measurement, we consider how

correlators are modified by the dispersive shift and the correcting two-mode gates.

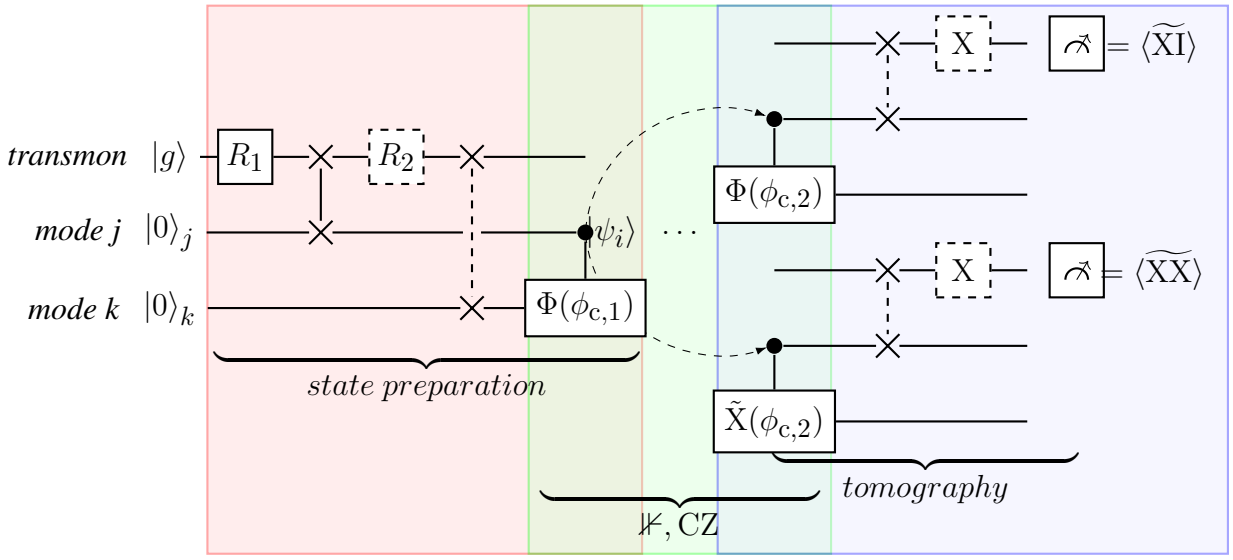

Figure 5.22: Process tomography of the $\mathbb{I}$ gate, used to calibrate the additional SPAM phase errors. We measuring the added phases to the $|10\rangle$ and $|11\rangle$ states by comparing the results of correlators with and without CNOT gates, thereby isolating state preparation and measurement errors.

As an example, for the prepared state $|\hat{x}\hat{x}\rangle$ ($R_{1,2} = Y_{\frac{\pi}{2}}$), the expected values of the correlators XI and XX are:

$$\langle\widetilde{\mathrm{XI}}\rangle = \cos^2\left(\frac{\phi_{\mathrm{c},1} + \phi_{\mathrm{c},2} - \phi_{\mathrm{p}} - \phi_{\mathrm{m}}}{2}\right), \tag{5.35}$$

$$\langle\widetilde{\mathrm{XX}}\rangle = \cos\left(\frac{\phi_{\mathrm{c},2} - \phi_{\mathrm{c},1} - \phi_{\mathrm{p}} - \phi_{\mathrm{m}}}{2}\right)\cos\left(\frac{\phi_{\mathrm{c},2} + \phi_{\mathrm{c},1} + \phi_{\mathrm{p}} - \phi_{\mathrm{m}}}{2}\right), \tag{5.36}$$

where ϕ_{p} and ϕ_{m} are the phase errors of the $|11\rangle$ state (relative to the other computational basis states) in state preparation and measurement, respectively. Finding and correcting ϕ_{p} and ϕ_{m} amounts to choosing $\phi_{\mathrm{c},1}$ and $\phi_{\mathrm{c},2}$ such that $\langle\widetilde{\mathrm{XI}}\rangle = \langle\widetilde{\mathrm{XX}}\rangle = 1$.

The additional phases only depend on the shape of the qubit and sideband pulse waveforms. As a result, we can calibrate ϕ_{p} and ϕ_{m} for all 240 sequences using a total of 13 unique experiments. We can then extract the full process matrix by measuring at the optimal angles obtained from the calibration experiments. We check that the validity of the calibrations by also additionally sweeping the phase of the final sideband pulse. In order to reduce SPAM error from decoherence,

we combine the state preparation and measurement correction gates (as indicated by the arrows in Figure 5.22) during process tomography of the $\nVdash$ and CZ gates, noting that $\mathrm{C}\Phi_{\phi_{\mathrm{c},1}}$ commutes with both of them.

We perform process tomography of the CZ gate by inserting it in place of the $\nVdash$ in Figure 5.22, after calibrating the tomography axes. A two-mode gate is fully characterized by the completely positive map $\mathcal{E}$;

$$\mathcal{E}(\rho) = \sum_{m,n=0}^{d^2-1} \chi_{mn} \hat{A}_m \rho \hat{A}_n^\dagger. \tag{5.37}$$

$\hat{A}_m = \hat{B}_i \otimes \hat{B}_j$, with $\hat{B}_i \in \{I, X, Y, Z\}$, forms a basis of operators acting on a two-mode state ρ. χ_{mn} is the process matrix characterizing the two-mode gate, and is extracted from the measured output density matrices ρ_j^{out} for 16 linearly independent input density matrices (ρ^j) as shown below:

$$\rho_j^{\mathrm{out}} = \mathcal{E}(\rho_j) = \sum_k \lambda_{jk} \rho_k \;=\; \sum_{m,n} \chi_{mn} \hat{A}_m \rho \hat{A}_n^\dagger = \sum_{m,n,k} \chi_{mn} \beta_{jk}^{mn} \rho_k, \tag{5.38}$$

$$\Rightarrow \lambda_{jk} = \mathrm{Tr}[\rho_k \rho_j^{\mathrm{out}}] \;=\; \sum_{mn} \beta_{jk}^{mn} \alpha_{mn}. \tag{5.39}$$

Equation (5.39) is directly inverted to obtain the process matrix α_{mn}. We do not impose the completeness condition, $\sum_{mn} \chi_{mn} \hat{A}_m \hat{A}_n^\dagger = \nVdash$ as a constraint. This constraint arises from the probabilities of states in the relevant two-mode space summing to 1. This is not necessarily the case when there are several memory modes. The process fidelities are extracted from the measured (χ^{m}) and ideal process matrices (χ^{t}) using $F_p = \mathrm{Tr}\left[\chi^{\mathrm{m}} \chi^{\mathrm{t}}\right]$.

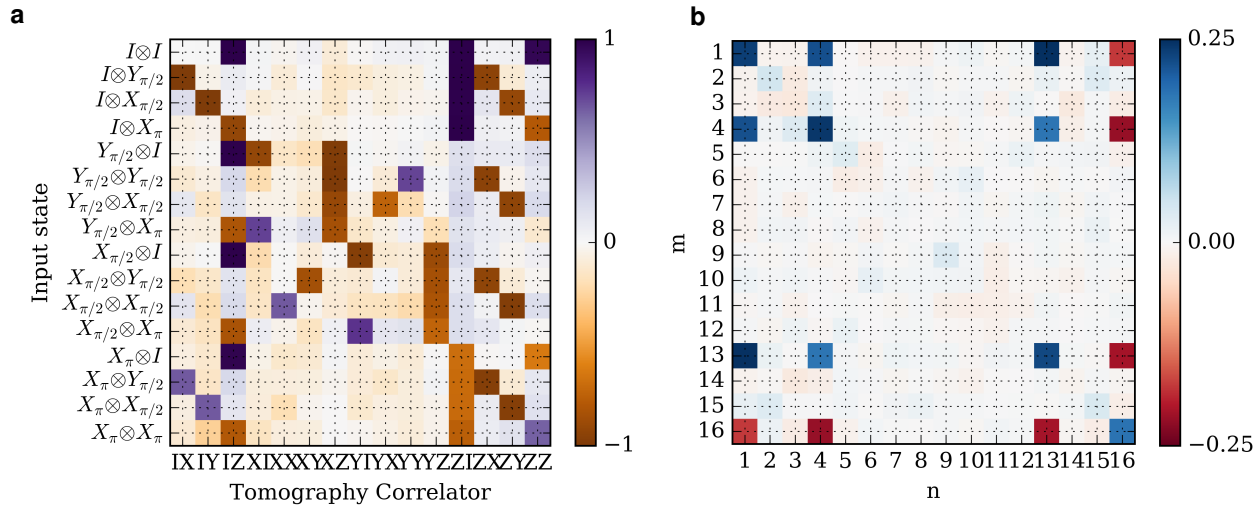

Figure 5.23: **a,** Experimentally measured correlators after correcting for phase errors arising during state preparation and measurement for process tomography of the CZ gate between mode $j = 2$ and $k = 6$. **b,** Process matrix extracted from the resulting measurements by inverting equation (5.39).

5.22 Preparation of entangled states

We use a slight variant of the scheme described in the main text to prepare multimode entangled states. We prepare Bell states [11] between two modes with the following protocol: starting with the transmon in its excited state, we swap half of the excitation via a sideband pulse ($\sqrt{\text{iSWAP}}$) to the first mode. This creates a $|\Psi^+\rangle$ Bell state between the first mode and the transmon:

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}\left(|g100...0\rangle + |e000...0\rangle\right). \tag{5.40}$$

We can rotate this state into the $|\Phi^+\rangle$ Bell state by flipping the transmon state via its charge bias control:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}\left(|g000...0\rangle + |e100...0\rangle\right). \tag{5.41}$$

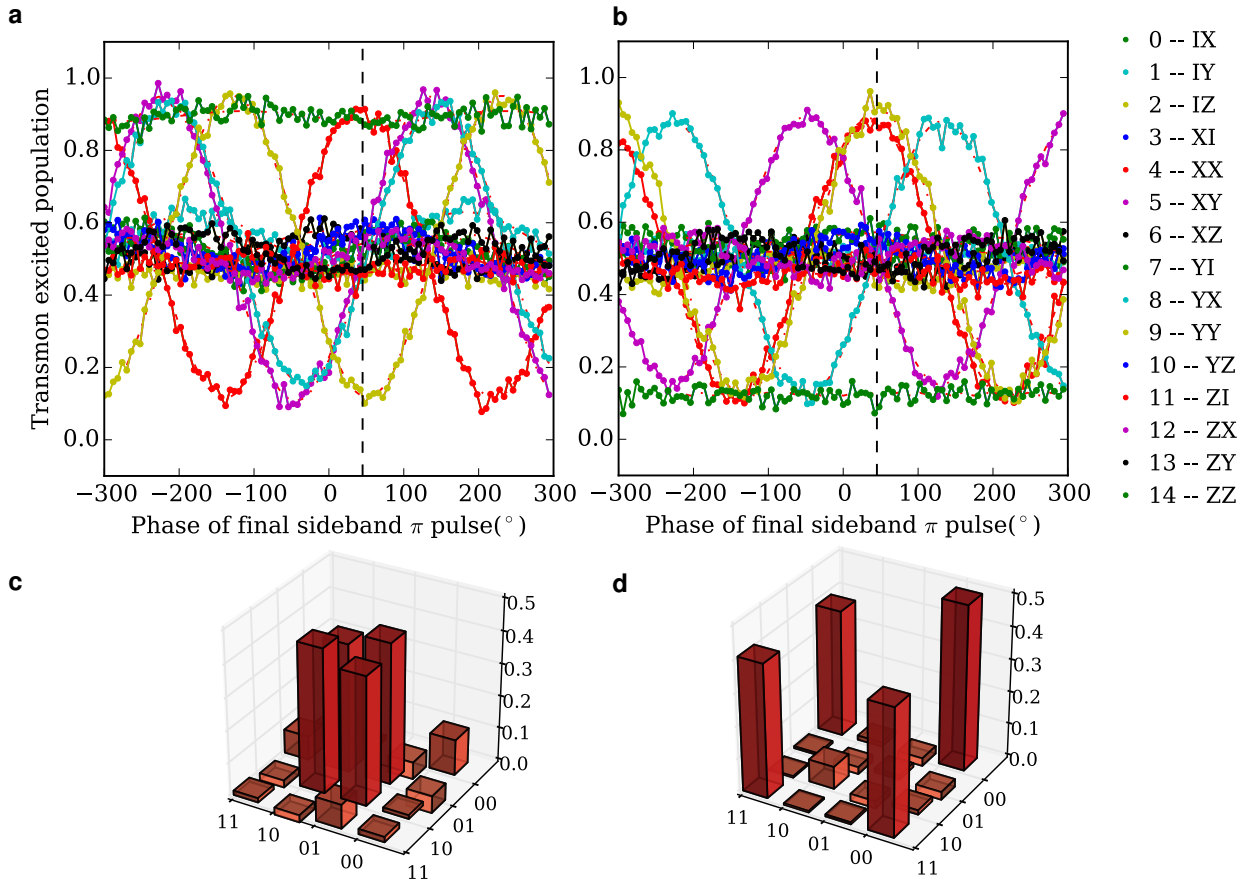

Figure 5.24: **a and b,** Transmon populations at the end of each of the two-mode correlator sequences for quantum state tomography of Ψ_+ and Φ_+ Bell states, respectively, as a function of the phase of the final sideband pulse. For these states, the only correlators that depend on the final sideband phase are $\mathrm{XX}, \mathrm{XY}, \mathrm{YX}, \mathrm{YY}$. Φ and Ψ Bell states are easily distinguished by a measurements of the two-mode parity operator ZZ. We account for additional phases by performing quantum state tomography using values measured at the sideband phase indicated by the dashed lines. **c and d,** The real parts of the density matrices. The corresponding state fidelities are found to be $F_p = 0.74$ and 0.75, respectively.

For either of the states in equations (5.40) and (5.41), we can create the corresponding Bell states between two arbitrary modes by simply swapping the transmon state to another mode.

To extend this protocol and create a n-mode Greenberger-Horne-Zeilinger (GHZ) state [68], we again utilize the anharmonicity of the transmon and map the population of $|e\rangle$ to $|f\rangle$. This allows us to transfer this excitation to the second mode via a sideband of the $|e\rangle - |f\rangle$ transition

without disturbing the population in the ground state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|g000...0\rangle + |e110...0\rangle\right). \tag{5.42}$$

These last two pulses can be repeated for each of the remaining eigenmodes before finally swapping the transmon back to the nth mode to complete the GHZ state:

$$|\psi_{\text{GHZ}}\rangle = \frac{1}{\sqrt{2}}\left(|g000...0\rangle + |g111...1\rangle\right). \tag{5.43}$$

5.23 GHZ entanglement witness

We witness the entanglement of the 3-qubit GHZ states through a measurement of the Mermin operator [123, 43],

$$\mathcal{W} = \mathrm{XXX} - \mathrm{YYX} - \mathrm{YXY} - \mathrm{XYY}. \tag{5.44}$$

For an ideal n-qubit GHZ state, the Einstein-Podolsky-Rosen (EPR) [52] bound for $\langle\mathcal{W}\rangle$ is $2^{(n-1)/2}$, while quantum mechanics predicts $\langle\mathcal{W}\rangle = 2^{n-1}$. For the 3-qubit GHZ state, $\langle\mathcal{W}\rangle > 2$ demonstrates tripartite entanglement.

We measure 3-qubit correlators by including an additional two-mode gate to the pulse sequences in Table 5.4 for measuring 2-qubit correlations. The XXX correlator is measured through the experimental sequence shown in Figure 5.25a. We reduce tomography error by compiling consecutive $|g\rangle - |e\rangle$ iSWAP pulses between successive CNOT gates, as well as the CNOT gate and the final iSWAP pulse. Additionally, we mitigate loss arising from transmon T_2^* by inserting a π pulse prior to the final $\pi/2$ qubit pulse, analogous to a spin-echo. The remaining correlators comprising $\mathcal{W}$ are measured by changing the phases of the CX gates and the final $\pi/2$ pulse. The results of the measurement of $\mathcal{W}$, as a function of the time τ between the π and $\pi/2$ pulses are shown in Figure 5.25b. The echo results in a revival of $\langle\mathcal{W}\rangle$ at a time $\tau \sim 530\,\text{ns}$. At the point indicated by the red star, $\langle\mathcal{W}\rangle = 2.35 \pm 0.04$, with the indicated error that arising from readout visibility

($4\times$visibility error). While the echo mitigates errors from transmon dephasing, the measured value of $\langle \mathcal{W} \rangle$ is a lower bound that includes loss from tomography. The tomography sequence prior to the echo includes 5 sideband and 2 qubit π pulses. Accounting for the loss in these pulses based on the RB data for the $|g\rangle - |e\rangle$ iSWAPs and qubit pulses gives $\langle \mathcal{W} \rangle > 2.65$. This is still a conservative estimate for the witness, since the $|e\rangle - |f\rangle$ iSWAPs have a lower fidelity than the corresponding $|g\rangle - |e\rangle$ iSWAPs.

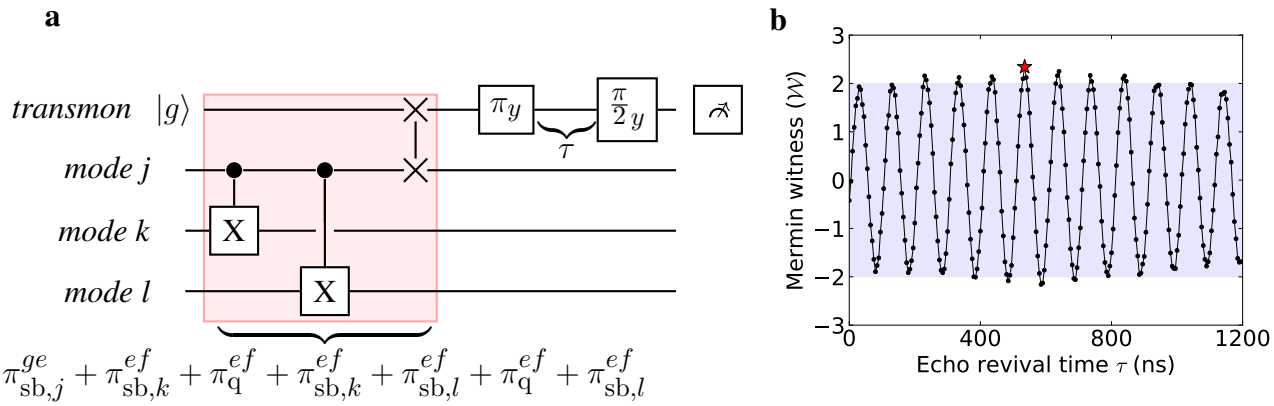

Figure 5.25: (a) Gate sequence for measurement of the XXX correlator with a transmon echo. As indicated by the pulse sequence, consecutive $|g\rangle - |e\rangle$ iSWAP pulses on the same mode in the red shaded region have been compiled and removed. (b) GHZ entanglement witness measured as a functon of time τ following the echo π pulse. The blue shaded region indicates the EPR bounds for $\mathcal{W}$, corresponding to the states being biseperable. The memory modes used in the measurement are $j = 6, k = 8$ and $l = 2$.

5.24 Conclusion

With minimal control-hardware overhead, we perform universal quantum operations between arbitrary modes of a nine-qubit memory using a single transmon as the central processor. The methods described in this work extend beyond this particular implementation of a multimode memory and in particular are compatible with the use of 3D superconducting cavities, which are naturally multimodal and have demonstrated the longest coherence times currently available in cQED [154], with the potential for even further improvements [163]. This architecture is compatible with the error-correcting codes that use higher Fock states of a single oscillator, such as the cat [103, 141]

and binomial [124] codes, as well as distributed qubit codes [21, 178], and is ideally suited to explore the potentially rich space of multi-qudit error-correcting codes that lie in between the two regimes [32, 148]. This makes cQED-based random access quantum information processors a promising new module for quantum computation and simulation.

CHAPTER 6

QUANTUM COMMUNICATIONS BETWEEN MULTIMODE MODULES

We further develop a more efficient and powerful quantum computer by connecting multiple quantum random access processors introduced in the last chapter. This establishes a readily scalable local quantum computing network. We propose and experimentally demonstrate an efficient scheme for bidirectional and deterministic photonic communication between two remote superconducting modules. The two chips, each consists of a transmon, are connected through a one-meter long coaxial cable that is coupled to a dedicated "communication" resonator on each chip. The two communication resonators hybridize with a mode of the cable to form a dark "communication mode" that is highly immune to decay in the coaxial cable. We overcome the various restrictions of quantum communication channels established by other recent approaches in deterministic communication for superconducting qubits. Our approach enables bidirectional communication, and eliminates the high insertion loss and large volume footprint of circulators. We modulate the transmon frequency via a parametric drive to generate sideband interactions between the transmon and the communication mode. We demonstrate bidirectional single-photon transfer with a success probability exceeding 60%, and generate an entangled Bell pair with a fidelity of $79.3 \pm 0.3\%$. The content of this chapter is based on our work "Deterministic bidirectional communication and remote entanglement generation between superconducting qubits" published in *npj Quantum Information* [104].

6.1 Introduction

A practical quantum computer requires a large number of qubits working in cooperation [58], a challenging task for any quantum hardware platform. For superconducting qubits, there is an ongoing effort to integrate increasing numbers of qubits on a single chip [139, 200, 27, 191, 59, 49, 164]. A promising approach to scaling up superconducting quantum computing hardware is to adopt a modular architecture [129, 18, 28] in which modules are connected together via communication

channels to form a quantum network. This reduces the number of qubits required on a single chip, and allows greater flexibility in reconfiguring and extending the resulting information processing system. In such an architecture, each module is capable of performing universal operations on multiple-bits, and neighboring modules are connected through photonic channels, allowing communication and entanglement generation between remote modules.

Remote entanglement between superconducting qubits has been realized probabilistically [162, 134, 45]. A deterministic quantum communication channel is advantageous over a probabilistic one because it lowers the threshold requirement for fault-tolerant quantum computation and can achieve higher entanglement rates [82]. Realizing deterministic photonic communication requires releasing a single photon from one qubit and catching it with the remote qubit. In the long-distance limit, the photon emission and absorption are from a continuum density of states. In this limit, static coupling limits the maximum transfer fidelity to only 54% [180, 197]. This limit is exceeded by dynamically tailoring the emission and absorption profiles [203, 177, 149, 201]. These capabilities are presently being used to perform photonic communication between superconducting qubits connected by a transmission line within a cryostat [4, 22, 45, 99]. In these experiments, the use of a circulator enables the finite-length transmission line to be modeled as a long line with a continuum density of states, at the cost of added transmission loss.

Here, we establish bidirectional photonic communication between two superconducting qubits through a multimodal communication channel. Rather than inserting a circulator, the multimode nature of the finite length transmission line is made manifest and exploited [79]. For intra-cryostat communication, the required connection coaxial cable length of 1 m or less results in a free spectral range on the order of hundreds of MHz. In this setting, the resonances of the coaxial cable form hybridized normal modes with on-chip communication resonators, and photons are transferred coherently through the discrete modes of the channel in contrast to emission/absorption through a continuum. We use parametric flux modulation of the qubit frequency to generate resonant sideband interactions between the qubit and the communication channel [102, 10, 181, 170, 118]. This approach avoids the loss due to the circulator that significantly limits the communication fidelity,

and enables bidirectional quantum communication. Moreover, avoiding large volume footprint circulators facilitates the establishment of communication channels between multiple devices within a cryostat.

6.2 Photonic communication by parametric interaction

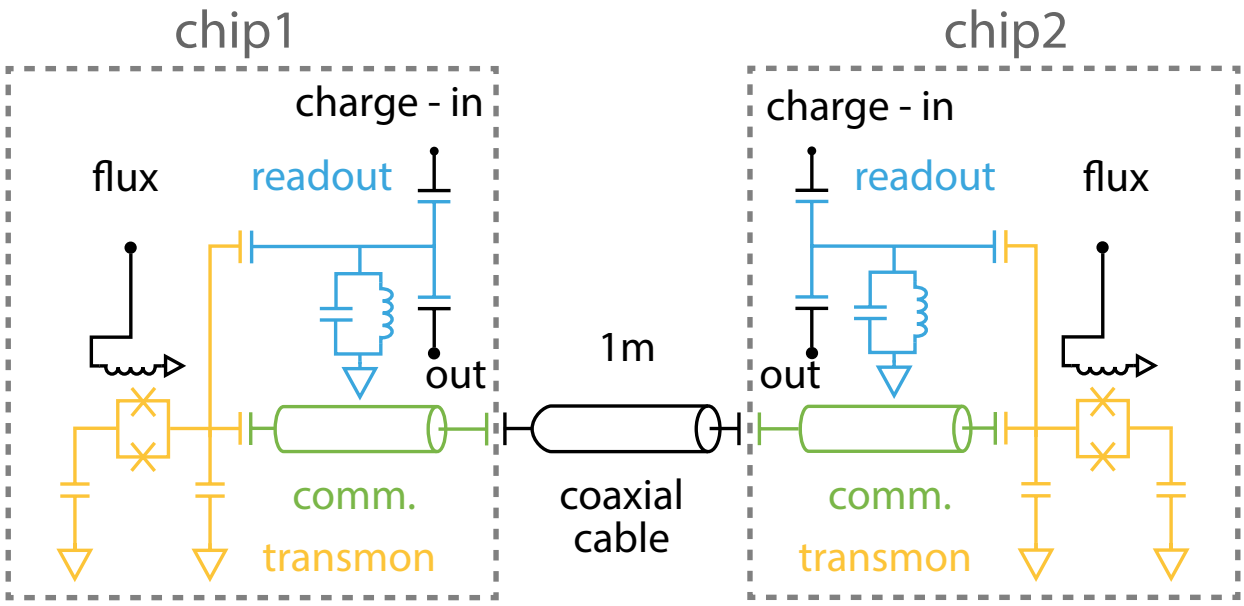

Figure 6.1: **Device circuit diagram.** Each chip consists of a frequency-tunable transmon, a resonator included for readout, and a second resonator coupled to the coaxial cable ($\sim$ 1 m) that provides the communication link between the chips. The cable we used is a Tin-dipped Hand-formable Microwave Cable (UT-085C-Form) with the silver-plated copper wire. The two communication resonators are designed to have identical frequencies. They are chosen to be coplanar waveguide resonators with a large center pin and gap width to make the frequency insensitive to fabrication variations [189]. The simple circuit diagram shows the circuit model of each module. We induce resonant interactions between the transmon and the communication mode by modulating the transmon frequency via its flux bias at the frequency difference (detuning) between the mode and the transmon.

Each chip consists of a single flux-tunable transmon and two additional resonators [122]. The first of these resonators is used for readout, and the second is coupled to the coaxial cable to enable the inter-module communication. The transmon can resonantly couple to the communication resonator through parametric flux modulation to realize inter-module photonic communications. Figure 6.1 shows a schematic of our two modules. The readout resonators have frequencies [module

1: 5.7463 GHz; module 2: 5.7405 GHz], and the communication resonators have frequencies [$\approx$ 7.88 GHz, see the Chapter 6.6 for detailed analysis of parameters]. We operate the transmons at the static frequency of [1: 4.7685 GHz; 2: 4.7420 GHz] with an anharmonicity of [1: 109.8 MHz; 2: 109.9 MHz]. Each chip also has an eight-bit quantum memory [133] with eigenmode frequencies spanning from 5.8 GHz to 7.7 GHz. These quantum memory modes can be utilized for universal quantum computation on each module locally but are not used in this work.

We induce resonant interactions between the transmon and the communication channel by modulating the transmon frequency via its flux bias. The modulation creates sidebands of the transmon excited state, detuned from the original resonance by the frequency of the applied flux tone. When one of these sidebands is resonant with a mode, the system experiences stimulated vacuum Rabi oscillations [133]. This process is similar to resonant vacuum Rabi oscillations [161], but occur at a rate that is controlled by the modulation amplitude [181, 10]. To illustrate the application of parametric control, we employ the following experimental sequence. First, the transmon is excited via its charge bias. Subsequently, we modulate the flux bias to create sidebands of the transmon excited state at the modulation frequency. This is repeated for different flux pulse durations and frequencies, with the population of the transmon excited state measured at the end of each sequence. When the frequency matches the detuning between the transmon and a given eigenmode, we observe full-contrast stimulated vacuum Rabi oscillations. Figure 6.2 shows that the transmon can selectively interact with each of the eigenmodes of the communication channel by choosing the appropriate modulation frequency. The photon transfer process between two remote qubits is initiated by switching on the sideband interactions targeting the communication resonator on each chip. As the bare frequencies of the transmon and the communication resonator are far detuned ($\Delta \approx 3$ GHz, $g \approx 50$ MHz), the sideband coupling scheme for photonic communication achieves a high on/off ratio. The following section explain the details of the multimode nature of our communication channel.

6.3 Multimode communication channel

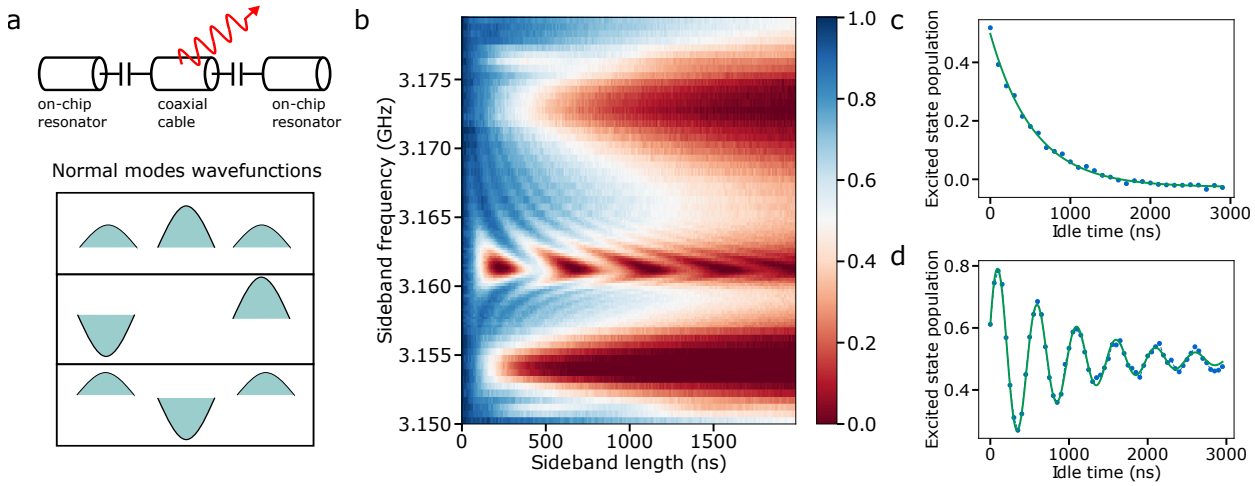

Figure 6.2: **Hybridized normal modes.** **a.** The schematic showing the wavefunctions of the coupled system involving the communication resonators and the coaxial cable. The three degenerate modes hybridize and form three normal modes with distinct signatures. The center normal mode has minimal participation in the lossy cable mode and has high quality factor. **b.** Stimulated vacuum Rabi oscillations around the communication modes. The near-degeneracy of the coaxial cable with the two communication resonators give rise to this almost equally-spaced three-mode structure. Being the two bright modes that include the lossy cable mode, and the dark "communication" mode of the two resonators. The latter couples more strongly to both qubits, and has a lifetime that is ideally only limited by the internal quality factors of the communication resonators. By fitting the simulation to experimental data, we found that the coaxial cable has a slightly higher frequency than the on-chip communication resonators [see Chapter 6.6]. **c., d.** T_1 and T_2^* data of the communication mode. The coherence time of the communication mode is characterized through protocols analogous to those for the transmon. We characterize both T_1 and T_2^* of the communication mode, and find $T_1 = 550ns$ and $T_2^* = 1\mu s$.

The two communication resonators are designed to have identical frequencies. They are chosen to be coplanar waveguide resonators with a large center pin and gap width to make the frequency insensitive to fabrication variations [189]. These resonators are coupled via the one-meter long coaxial cable, where the cable can be thought of as a multimode resonator with a free spectral range of around 100 MHz. The coupling strength between the cable and the communication resonators is $g_l \approx 7$ MHz. The cable mode that we use for communication has a frequency that is within g_l of the frequencies of the communication modes. Since the free spectral range of the coaxial cable is an order of magnitude larger than g_l, we consider the cable as a single mode nearly resonant with

the communication resonators. The cable and the communication resonators thus together produce three hybridized normal modes which are depicted in Figure 6.2. The near-degeneracy of the coaxial cable with the two communication resonators give rise to this almost equally-spaced three-mode structure, which can be seen from the three stimulated vacuum Rabi chevrons in Fig. 6.2 b. The center normal mode used for communication ideally has no participation in the cable mode, and as a result, its loss rate is limited by the internal quality factors of the communication resonators and small Purcell losses from neighboring cable modes. The transfer fidelity is ultimately limited by the single pass loss of the cable (see 6.9). In comparison to the neighboring modes, the center normal mode couples more strongly to both qubits due to higher wavefunction participation at the communication resonators. Thus, this communication mode has both the advantages of high quality factor and high coupling rate. For any practical device, the center normal mode does have a non-zero participation in the lossy coaxial cable due to a frequency mismatch between the two on-chip communication resonators. For our device, we found this frequency mismatch to be 5 MHz (detailed in Chapter 6.6), resulting in a less than 7% of cable mode participation in the communication mode.

The coherence time of the communication mode can be characterized using protocols analogous to those for the transmon. For instance, for mode T_1 one would first excite the transmon, then apply an iSWAP to transfer the excitation to the mode. After some variable time, another iSWAP is applied to transfer the excitation back to the transmon and we measure the transmon state subsequently [133]. We find $T_1 = 550\,\text{ns}$ and $T_2^* = 1\,\mu\text{s}$, corresponding to a quality factor of about 27000. This quality factor is reasonably high, considering that it involves losses from the long lossy cable, wirebonds, solder of the SMA connector, and the copper leads of the sample holder. The two neighboring normal modes have much lower coherence times due to the higher participation of the lossy cable mode. From fitting to fig. 6.2b we estimate the T_1 for the bare communication resonator and the bare cable mode to be ~ 5 us and ~ 40 ns, respectively (see Chapter 6.6). The dark state communication mode explored in this work can thus be utilized to improve transfer efficiency. In the following we demonstrate the use of the dark state communication mode

for bidirectional photon transfer and Bell state generation.

6.4 Bidirectional communication

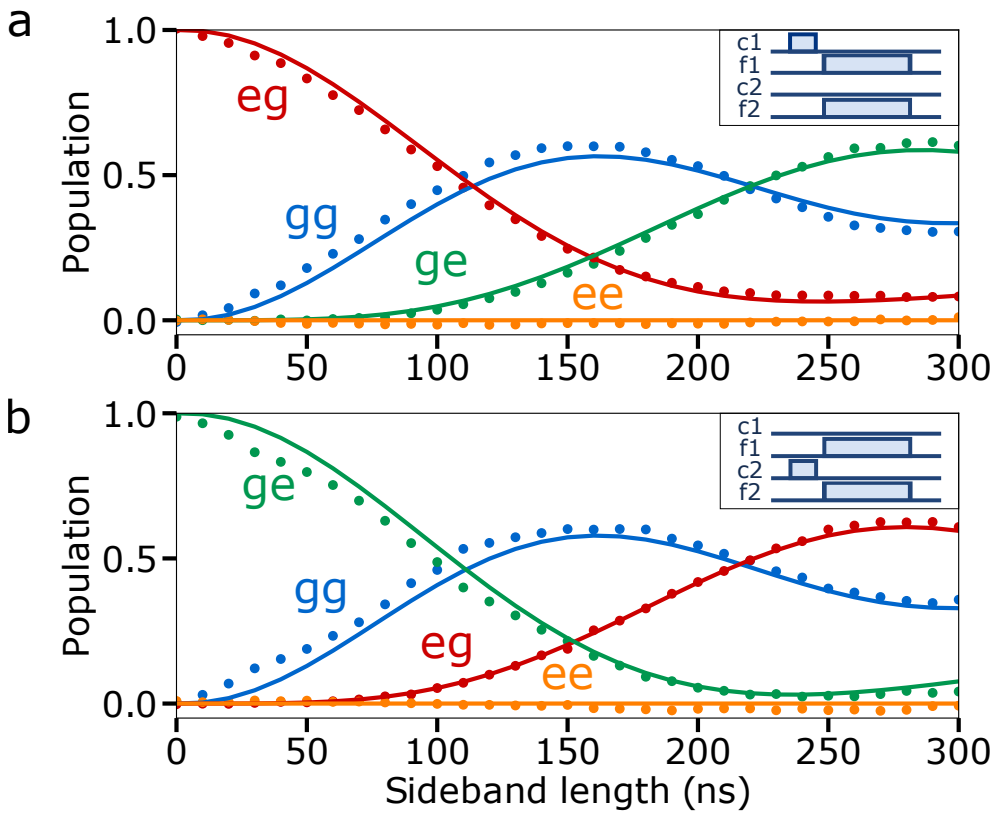

Figure 6.3: **Bidirectional excitation transfer.** The inset at top right shows the pulse sequence used to implement excitation transfer. The labels $c1$, $c2$ denote the charge drives on qubits 1 and 2, respectively, and $f1$, $f2$ the respective flux drives. We first apply a π pulse to excite one of the qubits, then simultaneously switch on the sideband flux pulse to drive the transfer process. Using the same sideband sequence, but instead applying the π pulse to the other qubit, we can send a single photon in the opposite direction. The transfer fidelity is limited by qubit dephasing and photon decay in the communication mode. Described in the following, the transfer process in different directions have slightly different loss mechanisms. **a.** Excitation transfer from qubit 1 to qubit 2. Notice that in this transfer process the sender qubit is not able to fully receive its excitation (population of $|eg\rangle$ does not reach zero). As confirmed by the master equation simulation, this is due to the dephasing of qubit 1. The remaining errors arise from communication cavity loss and dephasing of qubit 2, which is less than that of qubit 1. **b.** Excitation transfer from qubit 2 to qubit 1. In this process, while qubit 2 releases most of its excitations (population of $|ge\rangle$ comes close to zero), the dephasing of qubit 1 prevents it from capturing all the excitations in the communication mode, resulting in a slightly higher final population in $|gg\rangle$. The resulting fidelities for the transfer in the two directions are similar: $\{P_{|ge\rangle}, P_{|eg\rangle}\} \approx 61\%$, confirming the results from our numerical simulation.

To demonstrate photonic communication between the two chips, we send a single photon from one chip to the other. First, we excite the sender qubit, then we switch on sideband interactions simultaneously on both qubits, targeting the communication channel. This effectively creates a $\Lambda-$system between single photon states of the qubits and the communication mode. We send a photon in the reverse direction using the same sideband sequence but instead exciting the other qubit, thus demonstrating bidirectional photon transfer. Figure 6.3 shows the transmon population plotted as a function of the sideband pulse length. The master equation simulation results (solid lines) are shown along with the experimental data (dots). We are able to obtain photon transfer with a success rate of $\{P_{|ge\rangle}, P_{|eg\rangle}\} \approx 61\%$. We use simultaneous square pulses for the time-envelopes of the sideband interactions. From the simulations detailed in Chapter 6.8, we found that square pulses gave superior performance for our current circuit parameters. Note that the achieved transfer fidelity exceeds 54%, the maximum fidelity for absorbing a naturally shaped emission into a continuum [180, 197]. This demonstrates a qualitative difference in transferring via a multimode cable compared to that of releasing and catching flying photonic qubits through a continuum.

The transfer fidelity is limited by qubit dephasing and photon decay in the communication mode. Qubit 1 has a higher dephasing rate ($T_2^* \approx 700\,\text{ns}$) than qubit 2 ($T_2^* \approx 1.4\,\mu\text{s}$). The dephasing rate of qubit 1 is comparable to the sideband coupling rate, with the result that this qubit is not able to fully release its excitation during the transfer process. Conversely, for transfer in the other direction qubit 1 is not able to receive all of the excitations. This transfer infidelity can be largely mitigated by using a fixed-frequency qubit less susceptible to the flux noise, with its coupling strength to the communication mode parametrically controlled via a tunable coupler circuit [2, 25, 171, 119, 110]. The remaining loss of transfer fidelity comes from the loss in the communication mode. From our numerical simulations detailed in the Chapter 6.8, we estimate that the overall photon loss in both the qubits and the communication mode contribute to an infidelity of 24%, while the dephasing error of the two qubits accounts for an infidelity of 15%. The sideband coupling rate of the transmon is limited by the range over which its frequency can be parametri-

cally tuned, resulting in a maximum effective sideband coupling to the communication resonator of ≈ 2 MHz. With improved qubit coherence time, our simulation shows that more sophisticated transfer protocols such as STIRAP [70, 190] can be employed to boost transfer efficiency.

6.5 Bell state entanglement

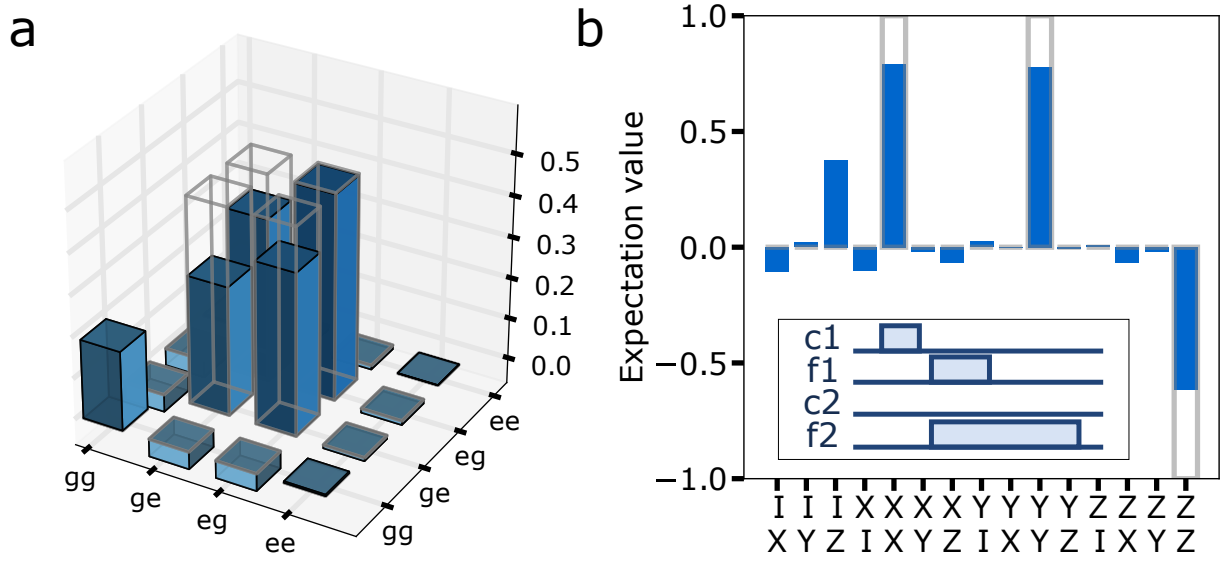

Figure 6.4: **Bell pair creation. a.** Real component of the density matrix. **b.** Expectation values of two-qubit Pauli operators. We create a Bell state between two remote qubits, one on each module. This is achieved by first applying the $\sqrt{i\text{SWAP}}$ gate between the excited qubit 1 and the communication mode, which is implemented by a sideband modulation pulse to qubit 1 to perform a $\pi/2$ rotation. A similar pulse, this time a π rotation, applied to the second qubit performs an iSWAP that transfers the entanglement from the communimation mode to the second qubit. As shown in the inset, we implement the two pulses simultaneously to reduce decoherence. We obtain the resulting Bell state with fidelity $\langle\Psi^+|\rho_{\text{exp}}|\Psi^+\rangle = 79.3 \pm 0.3$ %.

We now entangle two qubits by creating a Bell state between the transmons on the respective chips [12]. We can create such a state by first applying the $\sqrt{i\text{SWAP}}$ gate between the excited qubit 1 and the communication mode, which generates the Bell state $(|g1\rangle + |e0\rangle)/\sqrt{2}$ between them. We implement the $\sqrt{i\text{SWAP}}$ by applying a sideband modulation pulse to qubit 1 to perform a $\pi/2$ rotation. Subsequently, we transfer the state of the communication mode to qubit 2 through the iSWAP gate by applying a sideband modulation pulse to the latter to perform a π rotation. Ideally

this sequence prepares the Bell state $|\Psi^+\rangle = (|ge\rangle + |eg\rangle)/\sqrt{2}$ shared between the two remote qubits. To minimize decoherence the sender and receiver pulses can be applied simultaneously, so long as the lengths and amplitudes of the pulses are adjusted appropriately. Choosing qubit 1 as the sender and using square pulses, we found — both in our simulation and in the experiment — that maximal fidelity was obtained by setting both pulses at the same coupling rate and the length of the receiver pulse to be slightly longer than twice that of the sender,demonstrated by the pulse sequence diagram in Figure. 6.4b. The resulting Bell state has a fidelity of $\langle\Psi^+|\rho_{\text{exp}}|\Psi^+\rangle = 79.3 \pm 0.3\%$. We obtained the density matrix ρ_{exp} using quantum state tomography with an over-complete set of measurements complemented with the maximum likelihood method [80], and we corrected the measurement error by constructing a confusion matrix (detailed in Chapter 6.10). It can be inferred from the data that the fidelity is almost equally limited by photon decay in the cable and the qubit dephasing errors. We also note that the Bell state fidelity is significantly higher than the success probability we achieved for photon transfer. This is because the superposition state that is transferred has less participation in the communication mode which is the primary source of loss. The dephasing losses on both qubits are also reduced due to either shorter interaction time or lower average excitations.

6.6 Device Hamiltonian

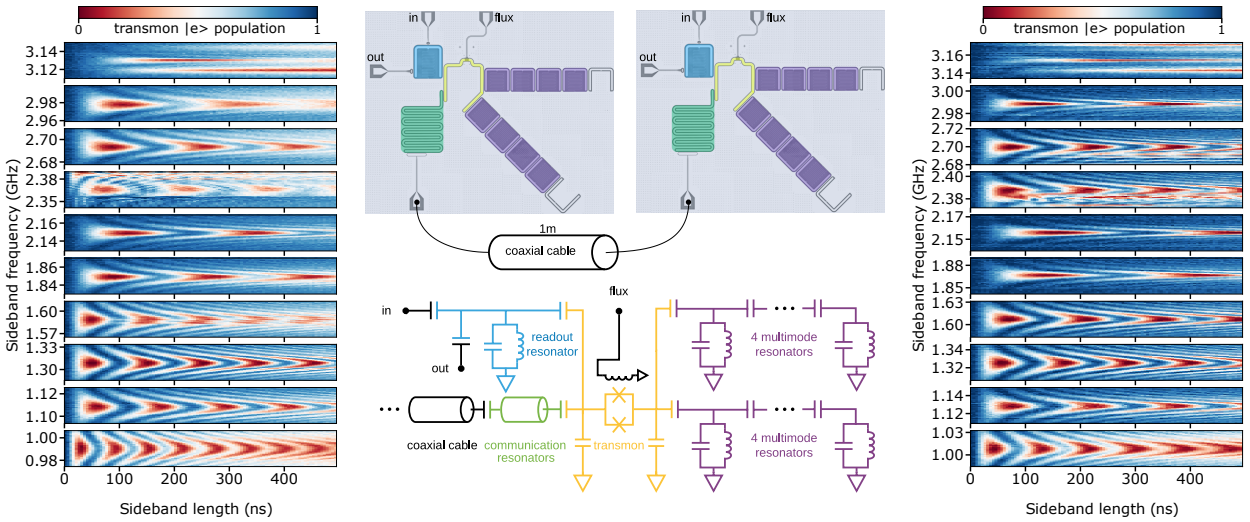

Figure 6.5: **Full Device schematic and whole spectrum stimulated vacuum Rabi oscillations.** Each chip consists of a frequency-tunable transmon and two chains of four identically designed, lumped-element resonators. In addition, a resonator is included for readout, and a second resonator is coupled to the coaxial cable ($\sim$ 1 m) that provides the communication link between the chips. The cable we used is a Tin-dipped Hand-formable Microwave Cable (UT-085C-Form). The wire is made of silver-plated copper. The dielectric is made of PTFE and the outer conductor is made of Tin dipped braid. The simple circuit diagram shows the circuit model of each module. We induce resonant interactions between the transmon and an individual mode by modulating the transmon frequency via its flux bias at the frequency difference (detuning) between the mode and the transmon. The chevron patterns indicate parametrically induced resonant oscillations with each of the modes. These patterns are generated by sweeping the length of the flux modulation pulse at each frequency and measuring the excited state population of the transmon after the pulse ends. The sideband modulation of the frequency-tunable transmon can access resonators in each chip by targeting their corresponding frequency detunings. On each chip, the lowest frequency corresponds to the readout resonator, and the highest frequency corresponds to the communication resonator. The eight memory mode frequencies are intermediate to these and spaced by $\approx$ 200 MHz. The local memory on each processor can be used for local processing and entanglement distillation in future work.

The following sections describe in detail on the characterization and calibrations of the two connected multimode modules. We begin with an in-depth discussion of our system Hamiltonian. Without connecting to the coaxial cable, the Hamiltonian of the i-th (i=1,2) circuit can be modeled by

$$\hat{H} = h\nu_{i,q}(t)\hat{a}_i^\dagger\hat{a}_i + \frac{1}{2}\alpha_i\hat{a}_i^\dagger\hat{a}_i(\hat{a}_i^\dagger\hat{a}_i - 1) + h\nu_{i,r}\hat{b}_{i,r}^\dagger\hat{b}_{i,r} + h\nu_{i,c}\hat{b}_{i,c}^\dagger\hat{b}_{i,c} + \sum_{m=1}^{8} h\nu_{i,m}\hat{b}_{i,m}^\dagger\hat{b}_{i,m}$$
$$+ hg_{i,r}(\hat{b}_{i,r} + \hat{b}_{i,r}^\dagger)(\hat{a}_i + \hat{a}_i^\dagger) + hg_c(\hat{b}_{i,c} + \hat{b}_{i,c}^\dagger)(\hat{a}_i + \hat{a}_i^\dagger) + \sum_{m=1}^{8} hg_{i,m}(\hat{b}_{i,m} + \hat{b}_{i,m}^\dagger)(\hat{a}_i + \hat{a}_i^\dagger) \tag{6.1}$$

where $\hat{a}_i$, $\hat{b}_{i,r}$, $\hat{b}_{i,c}$ and $\hat{b}_{i,m}$ stand for the annihilation operators of the flux-tunable qubit, the readout resonator, the communication cavity and the m-th multimode on the i-th chip. The communication cavities of the two chips are of identical coplanar waveguide resonator design with large center pin and gap width, leading to approximately the same resonant frequency $\nu_{1,c} \approx \nu_{2,c} = \nu_c$ and the same coupling strength g_l to the coaxial cable mode $\hat{b}_l$,

$$\hat{H}_{int} = \sum_{i=1}^{2} h\nu_c\hat{b}_{i,c}^\dagger\hat{b}_{i,c} + h\nu_l\hat{b}_l^\dagger\hat{b}_l + \sum_{i=1}^{2} hg_l(\hat{b}_l\hat{b}_{i,c}^\dagger + \hat{b}_l^\dagger\hat{b}_{i,c}). \tag{6.2}$$

Eq. 6.2 can be directly diagonalized, yielding three normal modes $\tilde{\hat{b}}_1$, $\tilde{\hat{b}}_2$ and $\tilde{\hat{b}}_c$,

$$\tilde{\hat{H}}_{int} = h\nu_c\tilde{\hat{b}}_c^\dagger\tilde{\hat{b}}_c + h\nu_1\tilde{\hat{b}}_1^\dagger\tilde{\hat{b}}_1 + h\nu_2\tilde{\hat{b}}_2^\dagger\tilde{\hat{b}}_2, \tag{6.3}$$

where

$$\begin{aligned}\nu_1 &= \nu_c + \frac{\delta}{2} + \sqrt{8g_l^2 + \delta^2},\\ \nu_2 &= \nu_c + \frac{\delta}{2} - \sqrt{8g_l^2 + \delta^2},\end{aligned} \tag{6.4}$$

and

$$
\begin{aligned}
\tilde{\hat{b}}_c &= \frac{1}{\sqrt{2}}(\hat{b}_{1,c} - \hat{b}_{2,c}), \\
\tilde{\hat{b}}_1 &= \frac{1}{\sqrt{2 + (r + \sqrt{2 + r^2})^2}}(\hat{b}_{1,c} + \hat{b}_{2,c} + (r + \sqrt{2 + r^2})\hat{b}_l), \\
\tilde{\hat{b}}_2 &= \frac{1}{\sqrt{2 + (r - \sqrt{2 + r^2})^2}}(\hat{b}_{1,c} + \hat{b}_{2,c} + (r - \sqrt{2 + r^2})\hat{b}_l).
\end{aligned} \tag{6.5}
$$

Here δ stands for the deviation of the cable mode frequency from the communication resonator frequency, i.e., $\delta = \nu_l - \nu_c$, and $r = \delta/2g_l$. The normal mode frequencies relative to the qubit frequency can be readily obtained from Fig. 6.2.b, so that δ and g_l can be calculated from Eqs. 6.3 and 6.4. Eqs. 6.1 and 6.5 together give the renormalized coupling strengths between the qubit and these normal modes,

$$
\begin{aligned}
\tilde{\hat{g}}_c &= \frac{g_c}{\sqrt{2}}, \\
\tilde{\hat{g}}_1 &= \frac{g_c}{\sqrt{2 + (h + \sqrt{2 + h^2})^2}}, \\
\tilde{\hat{g}}_2 &= \frac{g_c}{\sqrt{2 + (h - \sqrt{2 + h^2})^2}}.
\end{aligned} \tag{6.6}
$$

It is worth noting that the center normal mode, $\tilde{\hat{b}}_c$, is selected to be our communication channel mode in the experiment, for two obvious reasons: it contains only the two resonator modes with no convolution with the cable mode, as seen in Eq. 6.5, thus highly immune to the photon loss of the cable. Eq. 6.6 shows that it also couples more strongly to the qubit comparing to the other two normal modes, which also agrees well with Fig. 6.2.b where the center chevron has the fastest oscillation.

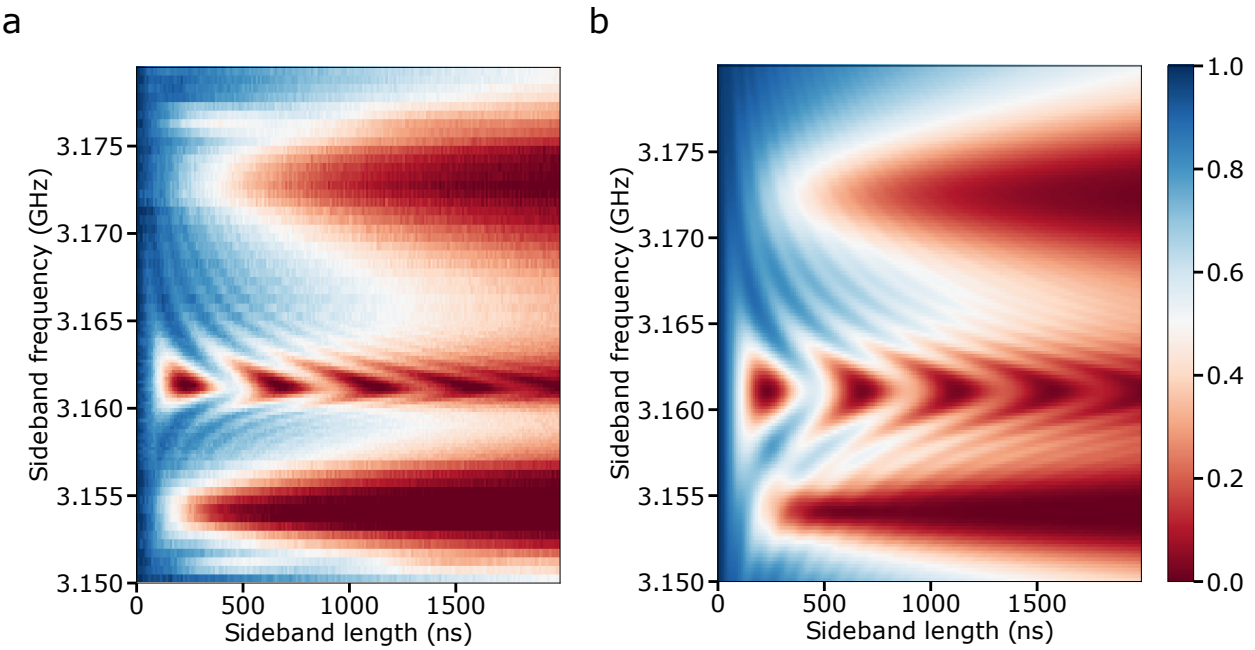

Figure 6.6: **Stimulated vacuum Rabi oscillation between the qubit and the communication mode.** By fitting to the experimental data (a) using our analytical model, we extracted the deviation of the cable mode frequency from the two communication resonators to be 4.25 MHz, while the coupling between the cable mode and the communication resonator is 6.46 MHz. Plugging these along with other circuit parameters obtained from the experiment into a master equation, we can simulate the experimental result with decent agreement (b).

Fitting Eq. 6.4 to Fig. 6.2b, we obtain $\delta = 4.25$MHz and $g_l = 6.46$MHz, from which we can numerically reproduce the chevron patterns observed in the experiment (Fig.6.6).

The above derivation is based on the ideal scenario where the two communication resonators have the same resonant frequency, resulting in a communication mode with no bare cable mode participation. However, in reality the imperfection of the circuit fabrication inevitably gives rise to a small detuning between the two resonators, with a typical value around a few MHz as observed from sample iterations. Also couplings between the two resonators and the rest of the cable modes need to be accounted for a more accurate calculation, as they lead to the an additional Purcell loss of the communication mode. Measure through our best fit of data with simulation, we find the detuning between the two communication resonators to be $\Delta = 5$ MHz, quality factors of the bare communication resonator and bare cable mode to be $Q_{i,c} = 2.65 \times 10^5$ and $Q_l = 2.1 \times 10^3$, and a free spectrum range of the cable of $\Delta\nu = 100$ MHz, then the Hamiltonian describing the

communication resonators and the cable modes can be written as

$$\hat{H}_{int} = \sum_{i=0}^{1} h(\nu_l - \delta + i\Delta)\hat{b}_{i,c}^{\dagger}\hat{b}_{i,c} + \sum_{j=-1}^{1} h(\nu_l + j\Delta\nu)\hat{b}_{j,l}^{\dagger}\hat{b}_{j,l} + \sum_{i,j} hg_{i,j,l}(\hat{b}_{j,l}\hat{b}_{i,c}^{\dagger} + \hat{b}_{j,l}^{\dagger}\hat{b}_{i,c}). \tag{6.7}$$

Plugging the values for each parameter in the above Hamiltonian, and being cognizant of the alternating phase of neighboring cable mode wave functions, we numerically diagonalize the Hamiltonian and obtain a participation of the lossy cable in the communication mode to be 6.5%, and we find the quality factors of the "dark" communication mode and the "bright" normal modes to be $Q_c = 2.67 \times 10^4$, $Q_1 = 5.5 \times 10^3$ and $Q_2 = 3.8 \times 10^3$, which quantitatively agrees with our measurement results in the main text.

Here we list the relevant circuit parameters in the following tables:

	qubit 1	qubit 2
frequency (GHz)	4.7685	4.7420
tunable range (GHz)	$\approx$ 3.0 - 5.9	$\approx$ 3.5 - 5.5
nonlinearity (MHz)	109.8	109.9
readout cavity frequency (GHz)	5.7463	5.7405
T_1 (μs)	10.1	7.9
T_2^* (μs)	0.7	1.4
coupling to comm resonator (MHz)	$\approx$ 40	$\approx$ 40

	comm resonator 1	cable mode	comm resonator 2	normal mode 1	comm mode	normal mode 2
frequency (GHz)	7.878	7.883	7.883	7.872	7.88	7.891
T_1 (μs)	5.35	0.043	5.35	0.15	0.55	0.1

6.7 Sideband interaction and calibrations

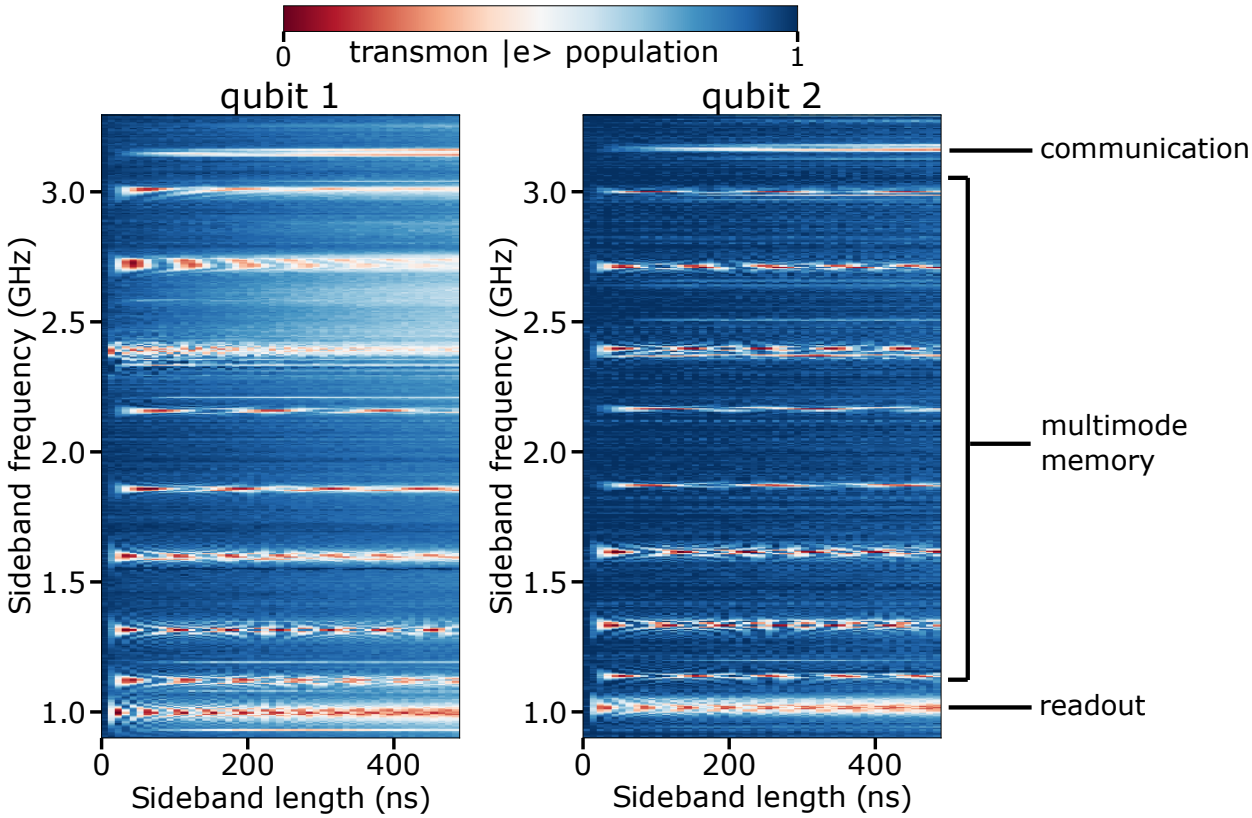

Figure 6.7: **Full sideband Rabi spectrum of each qubit.** Stimulated vacuum Rabi oscillation with sideband frequency scan covering the band of all resonance frequencies of the resonators. The clean chevron patterns indicate that our transmons are free from spurious crosstalks. We can clearly identify ten chevron patterns corresponding to one readout resonator (lowest frequency), eight multimode memory resonators and one communication resonator (highest frequency).

In figure 6.7, we can clearly identify ten chevron patterns corresponding to one readout resonator, eight multimode memory resonators, and one communication resonator. The crosstalk at sideband frequency ≈ 2.4 GHz corresponds to directly driving the g-e transition at the half of this frequency. The clean chevron patterns indicate that our transmons are free from spurious crosstalks. Compared to the segmented scans in the main text, these sideband scans are taken at higher amplitude

to broaden the chevron patterns for better visualization. The chevron patterns also show with faster oscillations and slightly higher frequency due to DC-offset, described in the next section.

The essential ingredient of photonic communication for our devices is the flux sideband interaction. It is therefore important to calibrate the sideband interactions well on both devices for obtaining high fidelity photonic communication. For our devices, this involves using the correct amplitude, frequency and timing of the sideband interactions. This section describes our calibration protocols for these parameters.

First, we run a 2D sweep (sideband amplitude and sideband frequency) of stimulated vacuum rabi around the communication mode frequencies. The main feature in figure 6.8 shows a clear pattern of three resonances, corresponds to the three hybridized normal modes of the communication channel. From the data, it is obvious that the resonance frequencies are dependent on the sideband amplitudes. The effect originated from the non-linear flux-frequency relation of the transmon, causing a shift (DC-offset) of the qubit frequency during the flux modulation. By doing a finer scan around the resonance frequency, we calibrated the on-resonance frequency of each sideband amplitude with an accuracy of 100 kHz.

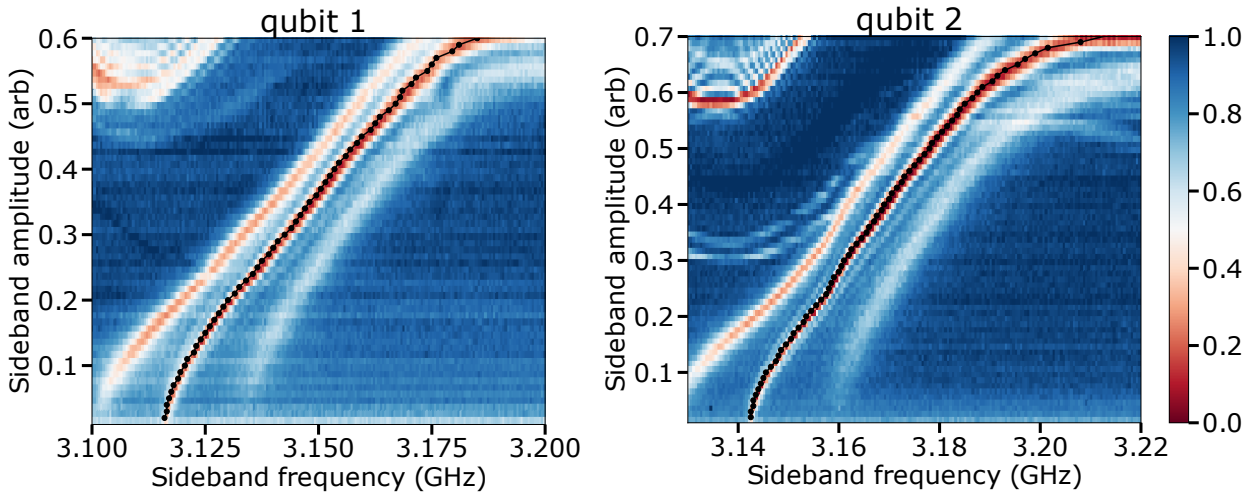

Figure 6.8: **DC offset scan.** There is a shift (DC-offset) of the qubit frequency during the flux modulation, arising from the non-linear flux-frequency relation of the transmon. To calibrate this effect, we sweep sideband transition frequency at different flux amplitudes and obtain the calibration with linear interpolation. The black dots on the figures show the tracked resonance sideband frequency for the considered range of amplitude. The pattern of three normal modes persisted for the considered range of sideband amplitudes. In this experiment, we set the sideband length to be inversely proportional to the sideband amplitude. This ensures high contrast features even for small sideband amplitude which the coupling is weak.

With the calibrated frequencies, we sweep the sideband length with a range of sideband amplitudes and obtain stimulated vacuum Rabi oscillation. The experimental data is displayed in figure 6.9. As expected, a higher sideband amplitude implies a higher effective coupling rate. Using this data, we obtained the effective qubit dissipation parameters during the sideband coupling. These dissipation parameters are subsequently being applied in master equation simulation of photon transfer and Bell entanglement generation.

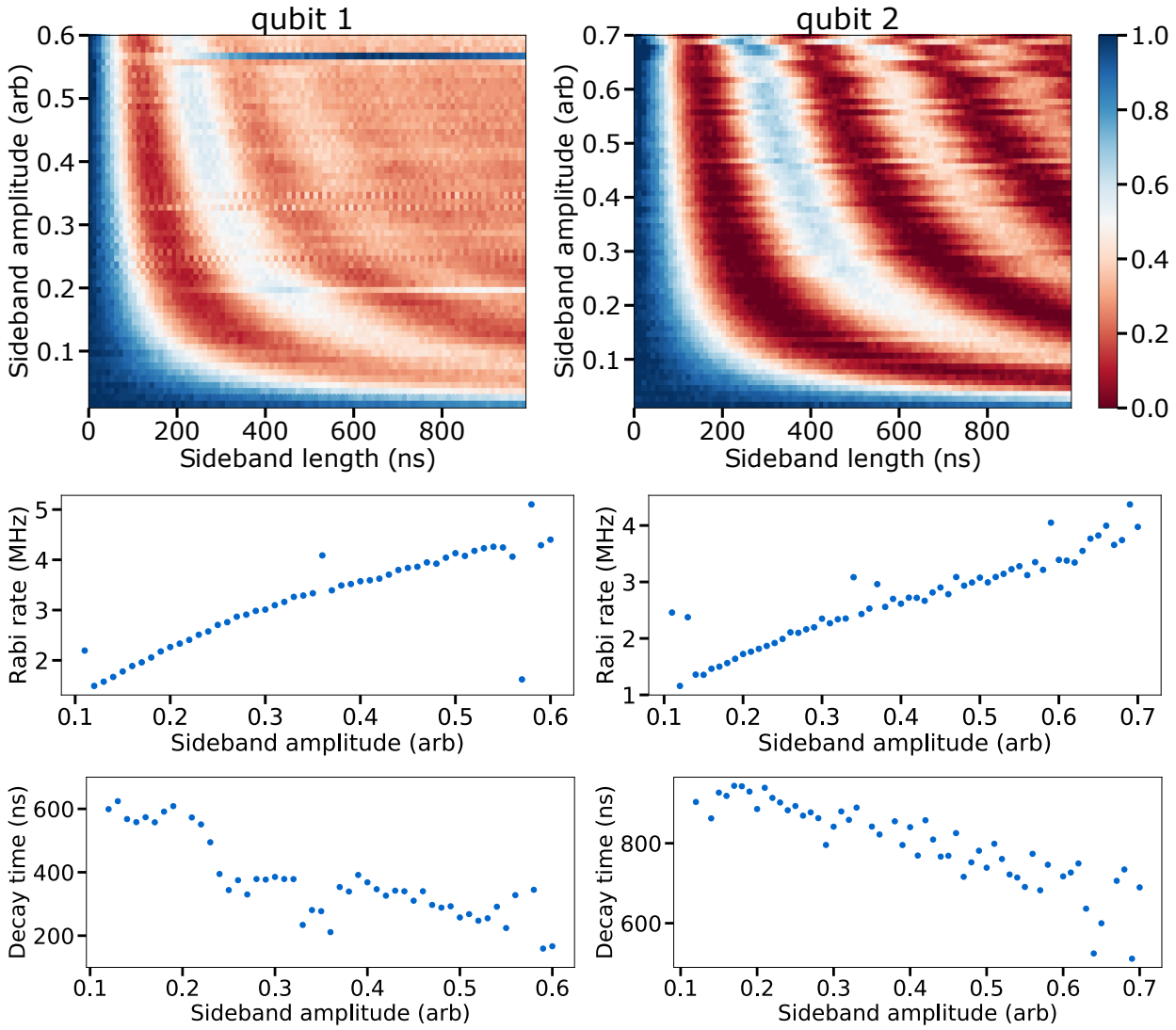

Figure 6.9: **Sideband Rabi sweep.** Using the calibrated DC offset, we obtained sideband rabi data between transmon and communication resonator for different sideband amplitude. Notice that the contrast of qubit 1 is much smaller than qubit 2. This is because qubit 2 has a higher coherence time. The trajectories of these scans are used for fitting the effective qubit decay parameters during the sideband coupling. The rabi rate and decay time plotted below are obtained through fitting the raw data shown in the color plot. These decay parameters are subsequently being applied in master equation simulation of photon transfer and Bell entanglement generation.

Lastly, we calibrated the timing of the two flux sideband pulses. Due to slightly different travel path length of flux line control from AWG to sample, we expect a slightly different timing between the two flux sideband pulses. Since the simultaneity of two flux sideband pulses is essential for high fidelity transfer, it is important to calibrate this systematic error. The experiment was conducted

with two equal length sideband pulses but sweeping the software delay between two pulses. Here, a negative receiver delay means the sender qubit (qubit 1) sideband pulse starts before the receiver qubit (qubit 2) sideband pulse. Figure 6.10 shows the population of the sender qubit with sweeping parameters of two sideband length and receiver delay. The center of the "K" pattern corresponds to the scenario where the photon is maximally captured by the receiver qubit. We obtained the "K" pattern as symmetric around receiver delay time of $\approx$ -10 ns, indicating the flux sideband pulse of the receiver qubit (qubit 2) lags the flux sideband pulse of the sender qubit (qubit 1). As a sanity check, we switched the role of sender and receiver qubit, such that sender is qubit 2 and receiver is qubit 1. In such case, we found that the pattern is symmetric around receiver delay time of $\approx$ +10 ns. This confirms our conclusion that indeed the qubit 2 lags the flux sideband pulse of qubit 1 due to a delay in the lines.

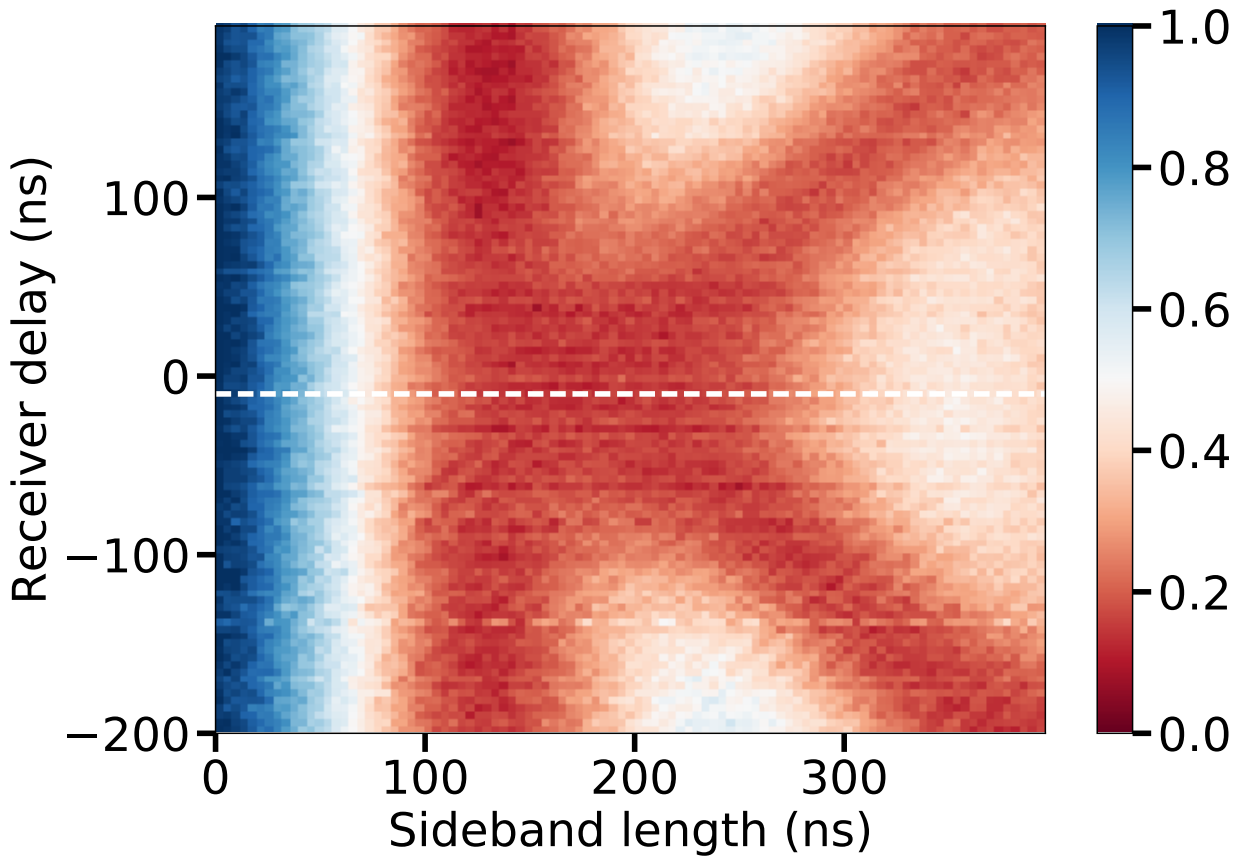

Figure 6.10: **Delay calibration.** This figure shows the population of the sender qubit with sweeping parameters of two sideband length and receiver delay. Ideally, the two flux sidebands during photon transfer should start simultaneously. However due to experimental conditions (e.g. different travel path length of flux line control from AWG to sample) causes the sideband pulses start at a different time on the devices, even the AWG is programmed to initiate two pulses simultaneously. To calibrate this effect, we sweep the delay between two sideband pulses and found that the flux control of qubit 2 is delayed by 10 ns. Throughout the experiment we time-advanced the control of qubit 2 flux by 10 ns in our pulse generation software.

6.8 Master equation simulation

In order to calculate the communication processes between the remote qubits using master equation simulations, we first write out the circuit Hamiltonian under flux modulations, based on Eq. 6.1∼6.6, as

$$\hat{H} = \sum_{i=1}^{2}\sum_{j=1}^{3} h(\nu_{i,q} + \epsilon_i \sin 2\pi f_i t)\hat{a}_i^\dagger \hat{a}_i + h\alpha_i \hat{a}_i^{\dagger 2}\hat{a}_i^2 + h\nu_j \hat{b}_j^\dagger \hat{b}_j + hg_j(\hat{b}_j + \hat{b}_j^\dagger)(\hat{a}_i + \hat{a}_i^\dagger), \quad (6.8)$$

where $\hat{b}_j$ stand for the three normal mode and g_j their coupling strengths to the two transmon qubits. Assuming weak flux modulation with $f_i \approx \nu_c - \nu_{i,q}$, and under the rotating frame transformation U =exp$[-i\sum_{i=1}^{2}\sum_{j=1}^{3}((2\pi\nu_{i,q}t - \frac{\epsilon_i}{2f_i}\cos 2\pi f_i t)\hat{a}_i^\dagger \hat{a}_i + 2\pi\nu_c \hat{b}_j^\dagger \hat{b}_j t)]$, Eq. 6.8 can be rewritten as

$$\hat{H} = \sum_{i=1}^{2}\sum_{j=1}^{3}\Bigg\{ h\alpha_i \hat{a}_i^{\dagger 2}\hat{a}_i^2 + h(\nu_j - \nu_c)\hat{b}_j^\dagger \hat{b}_j \quad (6.9)$$
$$-ihg_j J_1\left(\frac{\epsilon_i}{2f_i}\right)\left[\hat{b}_j \hat{a}_i^\dagger e^{2\pi i(f_i - \nu_j - \nu_c)t} - \hat{b}_j^\dagger \hat{a}_i e^{-2\pi i(f_i - \nu_j - \nu_c)t}\right]\Bigg\}.$$

Here $J_1(x)$ stands for the Bessel function of the first kind of the first order, and all the fast-oscillating terms have been abandoned. With the flux-modulation frequencies being $f_i = \nu_c - \nu_{i,q}$, and applying the two-level-approximation for the qubits, we find the "transfer Hamiltonian" as

$$\hat{H} = \sum_{i=1}^{2}\sum_{j=1}^{2} h(\nu_{l,j} - \nu_c)\hat{b}_{l,j}^\dagger \hat{b}_{l,j} - ihJ_1\left(\frac{\epsilon_i}{2f_i}\right)\left[g_{l,j}\left(\hat{b}_{l,j}\hat{\sigma}_i^+ - \hat{b}_{l,j}^\dagger \hat{\sigma}_i^-\right) + g_c\left(\hat{b}_c \hat{\sigma}_i^+ - \hat{b}_c^\dagger \hat{\sigma}_i^-\right)\right], \quad (6.10)$$

where $\hat{b}_{l,1}$ and $\hat{b}_{l,2}$ are the two lossy "bright" normal mode, and b_c is the "dark" communication channel mode. Plugging this into the master equation,

$$\dot{\rho} = -\frac{i}{\hbar}[\hat{H}, \rho] + \sum_{j=1}^{2}\kappa_{l,j}\mathcal{D}[\hat{b}_{l,j}]\rho + \kappa_c \mathcal{D}[\hat{b}_c]\rho + \sum_{i=1}^{2}\gamma_i \mathcal{D}[\hat{\sigma}_i^-]\rho + \sum_{i=1}^{2}\gamma_{i,\phi}\mathcal{D}[\hat{\sigma}_{i,z}]\rho, \quad (6.11)$$

with (in radians) $\kappa_{l,1} = 10\mu s^{-1}$, $\kappa_{l,2} = 25\mu s^{-1}$, $\kappa_c = 2\mu s^{-1}$, $\gamma_1 = 0.1\mu s^{-1}$, $\gamma_2 = 0.125\mu s^{-1}$, $\gamma_{1,\phi} = 1.43\mu s^{-1}$ and $\gamma_{2,\phi} = 0.71\mu s^{-1}$, we are able to simulate the bidirectional photon transfer experiment (Fig. 6.3) and the remote entanglement experiment (Fig. 6.4). The dephasing of qubit 1 is a major limiting factor of our Bell state fidelity. We found a very similar Bell state fidelity between using qubit 1 and qubit 2 as the sender. This is understood by realizing qubit 1 has high average photon occupation but reduced participation in the dynamics when qubit 1 is the sender, while qubit 1 has low average photon occupation but lengthened participation in the dynamics when qubit 1 is the receiver. It is therefore reasonable to expect the resulting total infidelity to be similar but better than the photon transfer case.

Simultaneous square sideband pulses are adopted in both the photon transfer and Bell state creation experiment to achieve the shortest pulse time possible. However, there is a possibility that better fidelities could be acquired through further minimizing the photon loss in the communication mode, by making use of adiabatic protocols in a manner akin to the stimulated Raman adiabatic passage (STIRAP). A typical STIRAP protocol has a pulse sequence shown in fig. 6.11a, where after the excitation of the sender qubit, the receiving pulse turns on first, and slowly ramps down together with the ramping up of the sending pulse. When the ramping of the pulses are done adiabatically with respect to the gap between the communication mode and the qubit modes, the transfer could be completed without inducing the communication mode population, and is therefore immune to the photon loss in the communication mode. However, this comes at the cost of much longer transfer time, which introduces more loss from the qubits.

For simplicity we model the sender and receiver pulses as two Gaussian pulses with the same maximum amplitude as the square pulse scheme used in our experiment. In the time domain, the two pulses are set to be

$$f_s(t) = \begin{cases} Ae^{-\frac{(t-t_0)^2}{2\sigma^2}}, & |t-t_0| \leqslant 5\sigma \\ 0, & |t-t_0| > 5\sigma \end{cases}, \quad f_r(t) = \begin{cases} Ae^{-\frac{(t-t_0-\Delta t)^2}{2\sigma^2}}, & |t-t_0-\Delta t| \leqslant 5\sigma \\ 0, & |t-t_0-\Delta t| > 5\sigma \end{cases}. \tag{6.12}$$

The fidelity yielded by this protocol is calculated as a function of both the pulse width σ and the delay time Δt, via master equation simulation with real circuit parameters. Fig. 6.11b shows that a maximum fidelity of 56% is achieved when two Gaussian pulses with $\sigma = 120$ ns overlap each other, which indicates that non-adiabatic transfer with shortest time is favorable in our current parameter regime. This also justifies our choice of the simultaneous square pulse scheme which is the fastest in all non-adiabatic schemes. In contrast, if the coherence of the qubit is improved to $T_1 = 20$ us and $T_2 = 20$ us, the same simulation results in a maximum fidelity of 85% at delay time $\Delta t =$ (fig. 6.11c) that is higher than the simultaneous square pulse fidelity of 82%, proving the usefulness of the adiabatic protocol for future improvements. Here we have only considered the simple Gaussian pulse scheme as a proof of principle demonstration. There are more sophisticated protocols that can further improve the transfer fidelity [192]. With better quality of qubits, the fidelity can potentially be improved through protocols with a small detuning between the parametric drive frequency and the frequency difference between the qubit and the communication mode. This virtual process mitigates the loss in the cables by avoiding excitation in the communication mode.

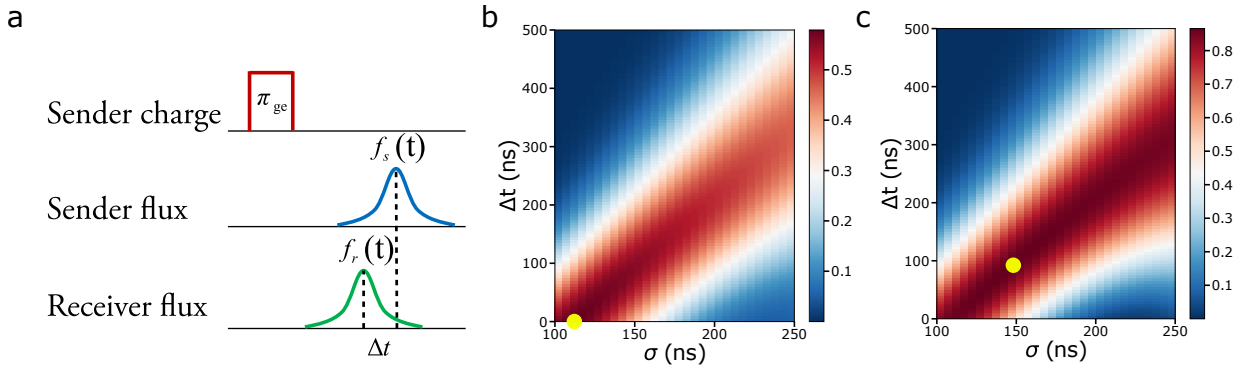

Figure 6.11: **STIRAP-like protocol.** (a) Pulse sequence of the STIRAP protocol for photon transfer. After initializing of the sender qubit state in the excited state, two Gaussian pulses with same duration and amplitude (set to be the maximum amplitude achievable in the experiment) are applied to the flux channels of the two qubits, with the receiver pulse turned on ahead of the sending pulse by a time of Δt. (b) Calculation of the transfer fidelity as a function of the Gaussian RMS width, σ, as well as the delay time Δt. A maximum fidelity of 56% occurs at $\{\sigma = 120\ ns, \Delta t = 0\ ns\}$ (labeled by the yellow dot), which is worse than the 60% fidelity achieved by the simultaneous square pulse scheme. This indicates that, in our current parameter regime, the fidelity is optimal with simultaneous square pulse scheme which has the shortest pulse length. (c) With better qubit coherence properties of $T1, T2 = 20$ us, the STIRAP protocol promises 85% maximum fidelity at $\{\sigma = 145\ ns, \Delta t = 95\ ns\}$ (labeled by the yellow dot), which is higher than the maximum fidelity of 82% yielded by the simultaneous square pulse scheme under the same parameters.

6.9 Single pass loss limitation of transfer fidelity

The transfer speed of photonic information (both quantum and classical) is fundamentally limited by the speed of light. Intuitively, this induces a limit towards the transfer fidelity, where the photon is exposed to cable loss with a duration of at least the single passage time. This section presents a formal argument of this claim from the perspective of quantum state transfer with a multimode cable.

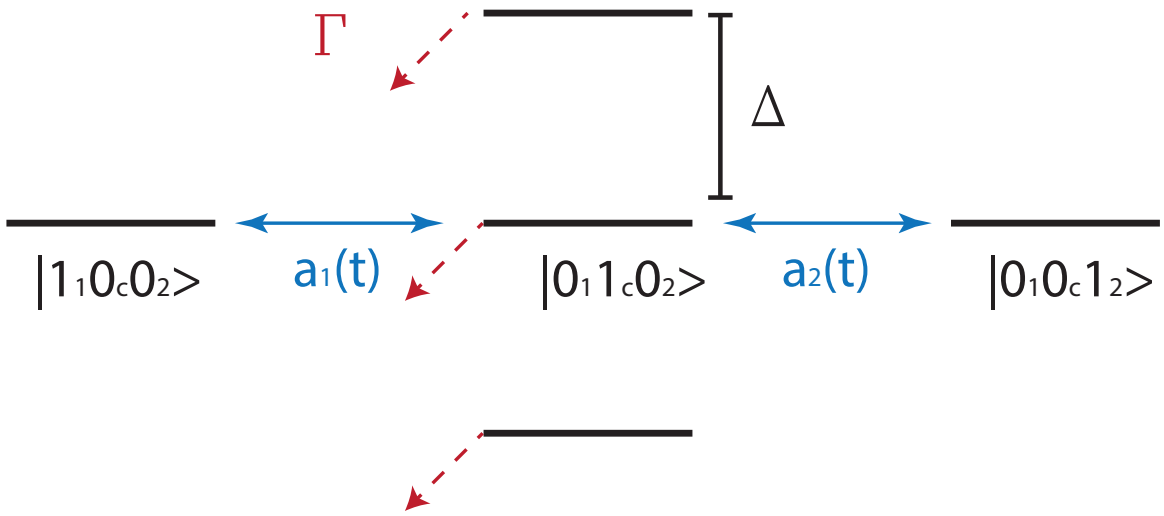

Figure 6.12: **Lambda system with extra lossy cable states.** Without the two extra lossy cable states, an adiabatic dark state protocol could conduct state transfer with almost perfect fidelity. The introduction of two extra lossy states in the model induces extra error which penalizes long transition times and makes adiabatic transfer non-optimal. The resulted optimal transfer under these conditions result in a total loss equivalent to the single pass loss through the cable.

Here we consider a $\Lambda-$system with two extra lossy cable states. Figure 6.12 shows the system with associated couplings and decays. The detuning Δ of these extra states is determined by the length of the cable L, with the relation $\Delta \sim \frac{v}{L}$ where v is the speed of light in cable. For simplicity of this analysis, we assume the cable is the only lossy component with decay rate Γ, and we induce a dark state protocol through STIRAP. We set the couplings of a STIRAP protocol to be $(a_1(t), a_2(t)) = \bar{a}(\sin\theta, \cos\theta)$, starting with $\theta(t=0) = 0$ and ending with $\theta(t=t_f) = \frac{\pi}{2}$. This results in a dark state $|dk\rangle = \cos\theta|1_1 0_c 0_2\rangle - sin\theta|0_1 0_c 1_2\rangle$ for each θ.

The error from non-adiabatic transition from $|dk\rangle$ to $|0_1 1_c 0_2\rangle$ is $err_{NA} \sim (\frac{\dot{\theta}}{\bar{a}})^2 \Gamma t_f \sim \frac{\Gamma}{\bar{a}^2 t_f}$. The non-resonant driving of neighboring modes induce a error $err_{NR} \sim (\frac{\bar{a}}{\Delta})^2 \Gamma t_f$. Note that the non-adiabatic error is higher with faster transition, while the non-resonant error from neighboring modes is higher with slower transition. We optimize t_f for minimal total error $err_{NA} + err_{NR}$, and obtain $t_{f,opt} = \frac{\Delta}{\bar{a}^2}$. Therefore, the minimal total error is $\sim \frac{\Gamma}{\Delta} \sim \frac{\Gamma}{v/L} \sim \Gamma\tau_{prop}$, where τ_{prop} is the single passage time. This error is equivalent to the single pass loss through the cable, thus

completing our argument.

6.10 Readout and state tomography

To measure the two-qubit state, we record the homodyne voltage for each qubit from every run. For example, run i of the experiment would result in a 4D heterodyne voltage values ($V_{I1,i}$, $V_{Q1,i}$, $V_{I2,i}$, $V_{Q2,i}$). These voltages are random numbers generated from a specific distribution corresponding to state projection and experimental noise. To measure the population in the four two-qubit basis states: $|gg\rangle, |ge\rangle, |eg\rangle, |ee\rangle$ we construct the histograms for these states by applying π pulses to the qubits. These histograms approximate the probability distribution for measuring a given voltage pair when the system is in a given basis state.

We employed logistic regression for classification of the two-qubit states. By setting decision thresholds for maximizing the classification accuracy for the two-qubit basis states according to the voltage distribution, we obtain a confusion matrix representing the correct and incorrect identification of basis state. For an unknown density matrix ρ we construct the classification distribution for ρ from N measurements, and project onto the basis states by applying the inverse of the calculated confusion matrix M (row: prepared state, column: measured state) for the states $\{|gg\rangle, |ge\rangle, |eg\rangle, |ee\rangle\}$, where

$$M = \begin{bmatrix} 0.8293 & 0.1053 & 0.0572 & 0.0082 \\ 0.1841 & 0.7514 & 0.0122 & 0.0523 \\ 0.101 & 0.0117 & 0.7923 & 0.095 \\ 0.0236 & 0.0979 & 0.1686 & 0.7099 \end{bmatrix}$$

We perform state tomography using the standard method by calculating the linear estimator,

$$\rho_{est} = \sum_{i,j} \frac{Tr[(\sigma_i \otimes \sigma_j)\rho](\sigma_i \otimes \sigma_j)}{4} \tag{6.13}$$

To calculate the term $Tr[(\sigma_i \otimes \sigma_j)\rho$ we apply a unitary operator U to ρ prior to measurement. For two-qubits, there are nine required measurements corresponding to the following unitary operators, $(I, R_Y(\pi/2), R_X(\pi/2)) \otimes (I, R_Y(\pi/2), R_X(\pi/2))$.

This simple linear estimator method can return unphysical results because it projects onto the space of all Hermitian matrices with Trace 1. However a physical density matrix must also be positive semi-definite. Following the maximum likelihood protocol outlined in [122, 80], we estimate the most likely physical density matrix by minimizing the function,

$$F[\rho_{est}] = \sum_{i=1,j=1}^{N,4} (\langle j|U_i^\dagger \rho_{est} U_i|j\rangle - P_{i,j})^2 \tag{6.14}$$

, where U_i are the set of N applied tomography pulses, $|j\rangle$ is the j^{th} basis state, $P_{i,j}$ is the measured probability, and ρ_{est} is a physical density matrix satisfying the physical constraints. The starting guess for the minimization is the density matrix estimated from the linear estimator with all negative eigenvalues set to zero. To form a over-complete set for a total of 17 tomography measurements, we also measure the negative pulse set [29] $(I, R_Y(-\pi/2), R_X(-\pi/2)) \otimes (I, R_Y(-\pi/2), R_X(-\pi/2))$.

6.11 Online Gaussian process for Bell state optimization

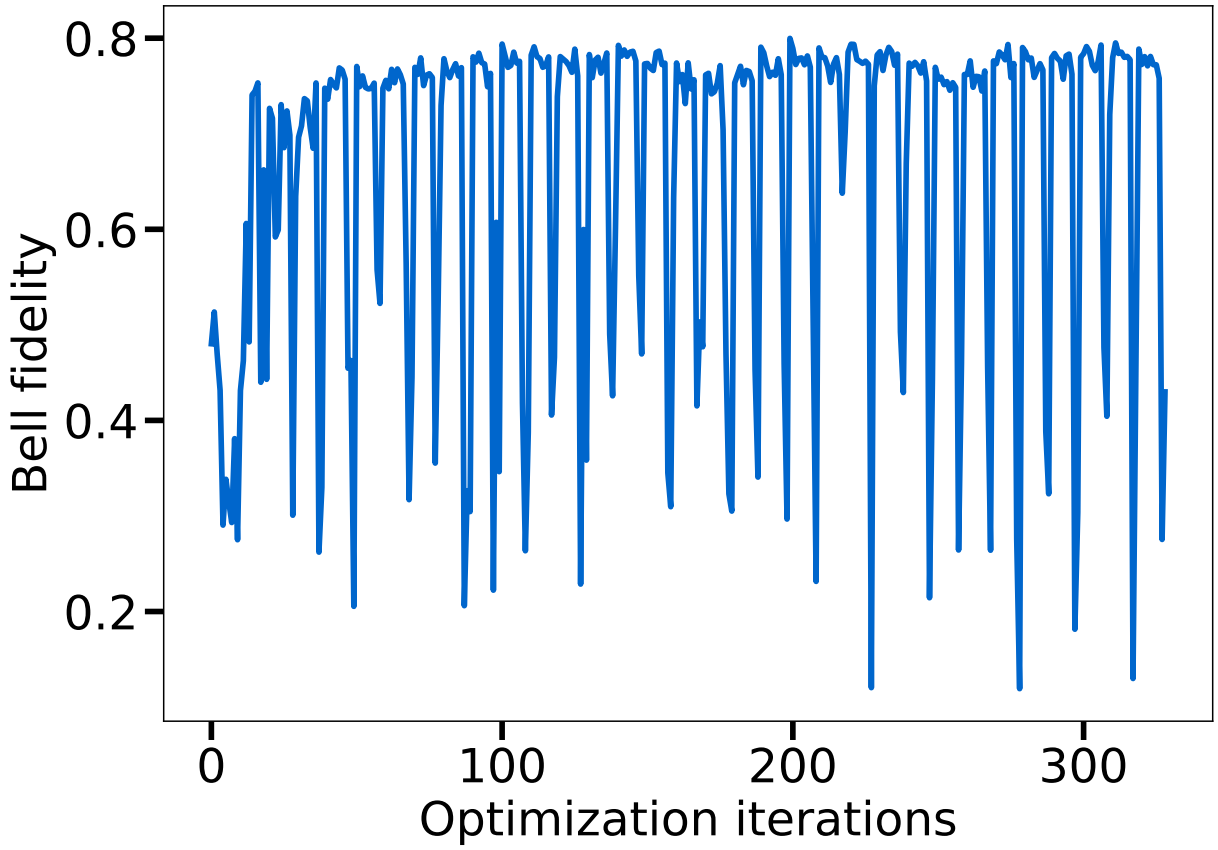

Figure 6.13: **Optimization of Bell state creation with an online Gaussian process.** We employed an online optimization directly applying on the experimental device. In each iteration, the Gaussian process model proposes 8 candidate solution (1 obtained from L-BFGS-B optimization on the Gaussian model, and 7 obtained from random sampling filtered with the best model prediction), and we also test two candidate solutions from pure random sampling to improve parameter space exploration. The random samplings lead to the apparent spikes of low fidelity Bell state during the optimization iterations. The model quickly starts to converge, and after some time we obtained Bell state with a fidelity close to 80%.

For two square pulses, there are in total 6 parameters (amplitude, frequency, and duration of each square pulse). The linear interpolation calibration of the DC offsets relates the amplitude and frequency parameters, thus resulting in 4 parameters to be optimized. All the parameters are fairly dependent on each other in the process of simultaneous transfer, meaning all 4 parameters have to be optimized together. Exhaustive search is quite forbidden even with just 4 parameters. There-

fore, we employed optimization techniques, in particular, the Gaussian process [152] to assist in optimizing Bell state creation. We employed an online optimization directly applied to the experimental device. In each iteration, the Gaussian process model proposes 8 candidate solution (1 obtained from L-BFGS-B optimization on the Gaussian model, and 7 obtained from random sampling filtered with the best model prediction), and we also test 2 candidate solution from pure random sampling to improve parameter space exploration. Figure 6.13 shows the optimization trajectory of Bell state creation. The model quickly starts to converge, and after some time we obtained Bell state with a fidelity close to 80%. Since only half of the excitation is being transmitted in the process, the transmission is less likely to be lost. We are able to obtain bell state creation with a fidelity higher than single photon transfer. During the optimization, we clipped the value of density matrix to a maximum of 0.5 for the calculation of fidelity. Without doing so, we found our numerical optimization results bias towards a higher excited population (> 0.5) of the sender qubit, where ideally one would expect the excited population to be 0.5. This artifact is likely due to the inner product definition of the fidelity, where > 0.5 excited population actually increases part of the inner products. We also took the absolute value of the resulting density matrix. This process optimized the Bell state up to a local qubit phase. To recover the target Bell state, we used the transfer parameters obtained from the optimizer and applied local phase advancement on one of the qubits. We repeated the Bell state creation experiment for 10+ times to obtain a statistics on the error of the Bell state fidelity. The resulted Bell state fidelity with this procedure was 79.3% $\pm$ 0.3%. The online optimization with Gaussian process works reasonably well even when we started with random initial parameters for the two square pulses. For arbitrarily shaped pulses, the high-dimensionality would necessarily require one to employ a model-based offline quantum optimal control [105] to facilitate the optimization process.

6.12 Heralding protocol for state transfer

It is possible to use heralding protocols to establish high fidelity quantum communications even the photons are transmitting through a lossy communication channel. The protocol we describe in

the following makes use of one qutrit and one communication resonator on each chip. The notation of the full state is

$$|\text{qubit, com}\rangle^1 \otimes |\text{qubit, com}\rangle^2$$

We initialize the system by creating an arbitrary superposition qubit state on chip1, using its $|g\rangle$ and $|f\rangle$ level:

$$\theta_{ge}^1 + \pi_{ef}^1 \Rightarrow (\alpha\,|g0\rangle^1 + \beta\,|f0\rangle^1) \otimes |g0\rangle^2 \tag{6.15}$$

Our goal is again to entangle the two remote qubits. However, starting from this step there are two schemes heralding photon loss errors in different ways, by using the $|g\rangle$ or the $|f\rangle$ level of second qubit respectively. The first scheme has the following steps:

$$\pi_{ge}^2 \Rightarrow \alpha\,|g0\rangle^1 \otimes |e0\rangle^2 + \beta\,|f0\rangle^1 \otimes |e0\rangle^2$$

$$\pi_{ef}^{sb1} + \text{1st photon flying} + \pi_{ef}^{sb2} \Rightarrow \alpha\,|g0\rangle^1 \otimes |e0\rangle^2 + \beta\,|e0\rangle^1 \otimes |f0\rangle^2$$

$$\pi_{ge}^1 + \pi_{ef}^1 + \pi_{ge}^2 \Rightarrow \alpha\,|f0\rangle^1 \otimes |g0\rangle^2 + \beta\,|g0\rangle^1 \otimes |f0\rangle^2$$

$$\pi_{ef}^{sb1} + \text{2nd photon flying} + \pi_{ge}^{sb2} \Rightarrow \alpha\,|e0\rangle^1 \otimes |e0\rangle^2 + \beta\,|g0\rangle^1 \otimes |f0\rangle^2$$

Now, it can be easily seen that if the 1st photon is lost, then the final state would be $|g0\rangle^1 \otimes |g0\rangle^2$, and if the 2nd photon is lost it would become $|e0\rangle^1 \otimes |g0\rangle^2$, hence the error is heralded by $|g\rangle^2$.

We now take a look at the other scheme which uses $|f\rangle^2$ for the heralding, which is arguably better because $|g\rangle$ and $|e\rangle$ are relatively more stable for storaging quantum information (Though, on the other hand, local decay can confuse $|f\rangle$ with $|e\rangle$ and hurt the fidelity, while the first scheme

is immune to this error). Still following line. (6.15),

$$\pi_{ef}^{sb1} + \text{1st photon flying} + \pi_{ge}^{sb2} \Rightarrow \alpha \left|g0\right\rangle^1 \otimes \left|g0\right\rangle^2 + \beta \left|e0\right\rangle^1 \otimes \left|e0\right\rangle^2$$

$$\pi_{ge}^{1} + \pi_{ef}^{1} + \pi_{ef}^{2} \Rightarrow \alpha \left|f0\right\rangle^1 \otimes \left|g0\right\rangle^2 + \beta \left|g0\right\rangle^1 \otimes \left|f0\right\rangle^2$$

$$\pi_{ef}^{sb1} + \text{2nd photon flying} + \pi_{ge}^{sb2} \Rightarrow \alpha \left|e0\right\rangle^1 \otimes \left|e0\right\rangle^2 + \beta \left|g0\right\rangle^1 \otimes \left|f0\right\rangle^2$$

$$\pi_{ge}^{2} + \pi_{ef}^{2} \Rightarrow \alpha \left|e0\right\rangle^1 \otimes \left|g0\right\rangle^2 + \beta \left|g0\right\rangle^1 \otimes \left|e0\right\rangle^2$$

Again the final state corresponding to the first and the second photon loss can be found to be $\left|g0\right\rangle^1 \otimes \left|f0\right\rangle^2$ and $\left|e0\right\rangle^1 \otimes \left|f0\right\rangle^2$, respectively, so measuring $\left|f\right\rangle$ state for the second qubit would herald the error and should be abandoned.

6.13 Conclusion

We have realized photonic communication between two remote modules, a first step in realizing a modular network. The sideband modulation of the transmon qubit in each module can be applied to implement local operations on the multimode resonators and to perform photon transfer between the two modules. The multimode characteristic of the communication channel (a coaxial cable) is enabled by the absence of a circulator. This mode structure results in normal modes that are superpositions of a mode of the inter-module communication cable and the on-chip resonators. One of the these normal modes is "dark" to the coaxial cable mode, thus avoiding much of the cable loss and allowing for high fidelity photon transfer. We characterized our system by performing single photon transfer with 61% fidelity and Bell-state preparation with 79.3% fidelity. These fidelities can be increased by improving the qubit coherence time and the strength of the coupling to the communication channel. This work sets the stage for future exploration in high

fidelity and scalable quantum communications. Fidelity can be further improved by implementing more sophisticated photon transfer protocols (e.g. STIRAP). Another exciting direction is scaling to multiple communication channels on the same cable, and exploring the crossover between distributed modes and the continuum as the cable gets longer. Finally, one can employ heralding [5, 131], and exploit the ability to perform high fidelity local gates in conjunction with photonic communication to distill entanglement [100, 183, 48] or perform forward error correction [127].

CHAPTER 7

REFERENCES

[1] Martín Abadi et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, March 2016.

[2] M.S. Allman, J.D. Whittaker, M. Castellanos-Beltran, K. Cicak, F. da Silva, M.P. DeFeo, F. Lecocq, A. Sirois, J.D. Teufel, J. Aumentado, and R.W. Simmonds. Tunable Resonant and Nonresonant Interactions between a Phase Qubit and L C Resonator. *Physical Review Letters*, 112(12):123601, 3 2014.

[3] M. Arioli, B. Codenotti, and C. Fassino. The Padé method for computing the matrix exponential. *Linear Algebra and its Applications*, 240:111–130, June 1996.

[4] 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, 7 2018.

[5] Koji Azuma, Kiyoshi Tamaki, and Hoi-Kwong Lo. All-photonic quantum repeaters. *Nature communications*, 6:6787, 2015.

[6] R Barends, J Kelly, A Megrant, A Veitia, D Sank, E Jeffrey, TC White, J Mutus, AG Fowler, B Campbell, et al. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature*, 508(7497):500–503, 2014.

[7] Michael Bartholomew-Biggs, Steven Brown, Bruce Christianson, and Laurence Dixon. Automatic differentiation of algorithms. *Journal of Computational and Applied Mathematics*, 124(1-2):171–190, December 2000.

[8] Atilim G. Baydin, Barak A. Pearlmutter, Alexey A. Radul, and Jeffrey M. Siskind. Automatic differentiation in machine learning: a survey, April 2015.

[9] Félix Beaudoin, Marcus P. da Silva, Zachary Dutton, and Alexandre Blais. First-order sidebands in circuit QED using qubit frequency modulation. *Phys. Rev. A*, 86:022305, Aug 2012.

[10] Flix Beaudoin, Marcus P. da Silva, Zachary Dutton, and Alexandre Blais. First-order sidebands in circuit QED using qubit frequency modulation. *Physical Review A*, 86(2):022305, 8 2012.

[11] John Stewart Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, 1964.

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

[13] Nathan Bell and Michael Garland. Efficient sparse matrix-vector multiplication on CUDA. NVIDIA Technical Report NVR-2008-004, NVIDIA Corporation, December 2008.

[14] Benjamin Block, Peter Virnau, and Tobias Preis. Multi-GPU accelerated multi-spin Monte Carlo simulations of the 2D Ising model. *Computer Physics Communications*, 181(9):1549–1556, September 2010.

[15] J. J. Bollinger, Wayne M. Itano, D. J. Wineland, and D. J. Heinzen. Optimal frequency measurements with maximally correlated states. *Phys. Rev. A*, 54:R4649–R4652, Dec 1996.

[16] Troy W. Borneman, Martin D. Hürlimann, and David G. Cory. Application of optimal control to CPMG refocusing pulse design. *Journal of Magnetic Resonance*, 207(2):220–233, December 2010.

[17] A. Borzı‘, J. Salomon, and S. Volkwein. Formulation and numerical solution of finite-level quantum optimal control problems. *Journal of Computational and Applied Mathematics*, 216(1):170–197, June 2008.

[18] Teresa Brecht, Wolfgang Pfaff, Chen Wang, Yiwen Chu, Luigi Frunzio, Michel H Devoret, and Robert J Schoelkopf. Multilayer microwave integrated quantum circuits for scalable quantum computing. *npj Quantum Information*, 2(1):16002, 11 2016.

[19] Jonas Bylander, Simon Gustavsson, Fei Yan, Fumiki Yoshihara, Khalil Harrabi, George Fitch, David G Cory, Yasunobu Nakamura, Jaw-Shen Tsai, and William D Oliver. Noise spectroscopy through dynamical decoupling with a superconducting flux qubit. *Nat. Phys.*, 7(7):565–570, 2011.

[20] Richard H. Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM J. Sci. Comput.*, 16(5):1190–1208, September 1995.

[21] A Robert Calderbank and Peter W Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54(2):1098, 1996.

[22] P. Campagne-Ibarcq, E. Zalys-Geller, A. Narla, S. Shankar, P. Reinhold, L. Burkhart, C. Axline, W. Pfaff, L. Frunzio, R.J. Schoelkopf, and M.H. Devoret. Deterministic Remote Entanglement of Superconducting Circuits through Microwave Two-Photon Transitions. *Physical Review Letters*, 120(20):200501, 5 2018.

[23] Bryan Catanzaro, Narayanan Sundaram, and Kurt Keutzer. Fast Support Vector Machine Training and Classification on Graphics Processors. In *Proceedings of the 25th International Conference on Machine Learning*, ICML '08, pages 104–111, New York, NY, USA, 2008. ACM.

[24] Qi M. Chen, Re B. Wu, Tian M. Zhang, and Herschel Rabitz. Near-time-optimal control for quantum systems. *Phys. Rev. A*, 92:063415, December 2015.

[25] Yu Chen, C. Neill, P. Roushan, N. Leung, M. Fang, R. Barends, J. Kelly, B. Campbell, Z. Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, A. Megrant, J.Y. Mutus, P.J.J. OMalley, C.M. Quintana, D. Sank, A. Vainsencher, J. Wenner, T.C. White, Michael R. Geller, A.N. Cleland, and John M. Martinis. Qubit Architecture with High Coherence and Fast Tunable Coupling. *Physical Review Letters*, 113(22):220502, 11 2014.

[26] Yu Chen, C Neill, P Roushan, N Leung, M Fang, R Barends, J Kelly, B Campbell, Z Chen, B Chiaro, et al. Qubit architecture with high coherence and fast tunable coupling. *Phys. Rev. Lett.*, 113(22):220502, 2014.

[27] Zijun Chen, A. Megrant, J. Kelly, R. Barends, J. Bochmann, Yu Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, J. Y. Mutus, P. J. J. O'Malley, C. Neill, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and John M. Martinis. Fabrication and characterization of aluminum airbridges for superconducting microwave circuits. *Applied Physics Letters*, 104(5):052602, 2 2014.

[28] Kevin S Chou, Jacob Z Blumoff, Christopher S Wang, Philip C Reinhold, Christopher J Axline, Yvonne Y Gao, Luigi Frunzio, MH Devoret, Liang Jiang, and RJ Schoelkopf. Deterministic teleportation of a quantum gate between two logical qubits. *Nature*, 561(7723):368, 2018.

[29] J. M. Chow, L. DiCarlo, J. M. Gambetta, A. Nunnenkamp, Lev S. Bishop, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Detecting highly entangled states with a joint qubit readout. *Physical Review A*, 81(6):062325, 6 2010.

[30] Jerry M Chow, Jay M Gambetta, Easwar Magesan, David W Abraham, Andrew W Cross, BR Johnson, Nicholas A Masluk, Colm A Ryan, John A Smolin, Srikanth J Srinivasan, et al. Implementing a strand of a scalable fault-tolerant quantum computing fabric. *Nat. Comm.*, 5:4015, 2014.

[31] JM Chow, JM Gambetta, Lars Tornberg, Jens Koch, Lev S Bishop, AA Houck, BR Johnson, L Frunzio, SM Girvin, and RJ Schoelkopf. Randomized benchmarking and process tomography for gate errors in a solid-state qubit. *Phys. Rev. Lett.*, 102(9):90502, 2009.

[32] Isaac L Chuang, Debbie W Leung, and Yoshihisa Yamamoto. Bosonic quantum codes for amplitude damping. *Phys. Rev. A*, 56(2):1114, 1997.

[33] Juan I Cirac and Peter Zoller. Quantum computations with cold trapped ions. *Physical review letters*, 74(20):4091, 1995.

[34] M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi. Solving lattice QCD systems of equations using mixed precision solvers on GPUs. *Computer Physics Communications*, 181(9):1517–1528, September 2010.

[35] W. J. Cody, G. Meinardus, and R. S. Varga. Chebyshev rational approximations to ex in [0, +) and applications to heat-conduction problems. *Journal of Approximation Theory*, 2(1):50–65, March 1969.

[36] A. D. Córcoles, Easwar Magesan, Srikanth J. Srinivasan, Andrew W. Cross, M. Steffen, Jay M. Gambetta, and Jerry M. Chow. Demonstration of a quantum error detection code using a square lattice of four superconducting qubits. *Nature Communications*, 6:6979, April 2015.

[37] Xiang Cui, Yifeng Chen, and Hong Mei. Improving Performance of Matrix Multiplication and FFT on GPU. In *Parallel and Distributed Systems (ICPADS), 2009 15th International Conference on*, pages 42–48, Washington, DC, USA, December 2009. IEEE.

[38] 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, October 2011.

[39] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe. Demonstration of a small programmable quantum computer with atomic qubits. *Nature*, 536(7614):63–66, August 2016.

[40] Michel H Devoret and Robert J Schoelkopf. Superconducting circuits for quantum information: an outlook. *Science*, 339(6124):1169–1174, 2013.

[41] Michel H Devoret, Andreas Wallraff, and John M Martinis. Superconducting qubits: A short review. *arXiv preprint cond-mat/0411174*, 2004.

[42] L DiCarlo, JM Chow, JM Gambetta, Lev S Bishop, BR Johnson, DI Schuster, J Majer, A Blais, L Frunzio, SM Girvin, et al. Demonstration of two-qubit algorithms with a superconducting quantum processor. *Nature*, 460(7252):240–244, 2009.

[43] L DiCarlo, MD Reed, L Sun, BR Johnson, JM Chow, JM Gambetta, L Frunzio, SM Girvin, MH Devoret, and RJ Schoelkopf. Preparation and measurement of three-qubit entanglement in a superconducting circuit. *Nature*, 467(7315):574–578, 2010.

[44] R. H. Dicke. Coherence in spontaneous radiation processes. *Phys. Rev.*, 93:99–110, Jan 1954.

[45] 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):064508, 2 2018.

[46] P. Ditz and A. Borzi‘. A cascadic monotonic time-discretized algorithm for finite-level quantum control computation. *Computer Physics Communications*, 178(5):393–399, March 2008.

[47] Florianín Dolde et al. High-fidelity spin entanglement using optimal control. *Nature Communications*, 5, February 2014.

[48] Ruifang Dong, Mikael Lassen, Joel Heersink, Christoph Marquardt, Radim Filip, Gerd Leuchs, and Ulrik L Andersen. Experimental entanglement distillation of mesoscopic quantum states. *Nature Physics*, 4(12):919, 2008.

[49] A. Dunsworth, R. Barends, Yu Chen, Zijun Chen, B. Chiaro, A. Fowler, B. Foxen, E. Jeffrey, J. Kelly, P. V. Klimov, E. Lucero, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, H. Neven, John M. Martinis, and A. Megrant. A method for building low loss multi-layer wiring for superconducting microwave devices. *Applied Physics Letters*, 112(6):063502, 2 2018.

[50] D. J. Egger and F. K. Wilhelm. Adaptive Hybrid Optimal Quantum Control for Imprecisely Characterized Systems. *Phys. Rev. Lett.*, 112:240503+, June 2014.

[51] D. J. Egger and F. K. Wilhelm. Optimized controlled-Z gates for two superconducting qubits coupled through a resonator. *Superconductor Science and Technology*, 27(1):014001+, January 2014.

[52] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47(10):777, 1935.

[53] Reuven Eitan, Michael Mundt, and David J. Tannor. Optimal control with accelerated convergence: Combining the Krotov and quasi-Newton methods. *Phys. Rev. A*, 83:053426, May 2011.

[54] Artur K Ekert, Carolina Moura Alves, Daniel KL Oi, Michał Horodecki, Paweł Horodecki, and Leong Chuan Kwek. Direct estimations of linear and nonlinear functionals of a quantum state. *Phys. Rev. Lett.*, 88(21):217901, 2002.

[55] John D. Farnum and David A. Mazziotti. Trigonometric mapping for the electric field strength in molecular optimal control theory. *Chemical Physics Letters*, 416(1-3):142–146, November 2005.

[56] K. Fatahalian, J. Sugerman, and P. Hanrahan. Understanding the Efficiency of GPU Algorithms for Matrix-matrix Multiplication. In *Proceedings of the ACM SIGGRAPH/EUROGRAPHICS Conference on Graphics Hardware*, HWWS '04, pages 133–137, New York, NY, USA, 2004. ACM.

[57] Austin G Fowler, Matteo Mariantoni, John M Martinis, and Andrew N Cleland. Surface codes: Towards practical large-scale quantum computation. *Phys. Rev. A*, 86(3):032324, 2012.

[58] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86(3):032324, 9 2012.

[59] B Foxen, JY Mutus, E Lucero, R Graff, A Megrant, Yu Chen, C Quintana, B Burkett, J Kelly, E Jeffrey, et al. Qubit compatible superconducting interconnects. *Quantum Science and Technology*, 3(1):014005, 2017.

[60] Jay M. Gambetta, Jerry M. Chow, and Matthias Steffen. Building logical qubits in a superconducting quantum computing system. *npj Quantum Information*, 3(1):2, 2017.

[61] K. Geerlings, Z. Leghtas, I. M. Pop, S. Shankar, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret. Demonstrating a driven reset protocol for a superconducting qubit. *Phys. Rev. Lett.*, 110:120501, Mar 2013.

[62] Sergey Gladchenko, David Olaya, Eva Dupont-Ferrier, Benoit Douçot, Lev B Ioffe, and Michael E Gershenson. Superconducting nanocircuits for topologically protected qubits. *Nat. Phys.*, 5(1):48–53, 2009.

[63] Steffen J. Glaser, Ugo Boscain, Tommaso Calarco, Christiane P. Koch, Walter Köckenberger, Ronnie Kosloff, Ilya Kuprov, Burkhard Luy, Sophie Schirmer, Thomas Schulte-Herbrüggen, Dominique Sugny, and Frank K. Wilhelm. Training Schrödinger's cat: quantum optimal control. *Eur. Phys. J. D*, 69(12):1–24, 2015.

[64] Michael H. Goerz, Giulia Gualdi, Daniel M. Reich, Christiane P. Koch, Felix Motzoi, K. Birgitta Whaley, J. Vala, Matthias M. Müller, Simone Montangero, and Tommaso Calarco. Optimizing for an arbitrary perfect entangler. II. Application. *Phys. Rev. A*, 91:062307+, June 2015.

[65] Michael H. Goerz, Felix Motzoi, K. Birgitta Whaley, and Christiane P. Koch. Charting the circuit QED design landscape using optimal control theory, August 2016.

[66] Caroline Gollub, Markus Kowalewski, and Regina de Vivie-Riedle. Monotonic Convergent Optimal Control Theory with Strict Limitations on the Spectrum of Optimized Laser Fields. *Phys. Rev. Lett.*, 101:073002, August 2008.

[67] Daniel Gottesman. The Heisenberg representation of quantum computers. *arXiv preprint quant-ph/9807006*, 1998.

[68] Daniel M Greenberger, Michael A Horne, and Anton Zeilinger. Going beyond Bell's theorem. In *Bell's theorem, quantum theory and conceptions of the universe*, pages 69–72. Springer, 1989.

[69] Lov K Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 212–219. ACM, 1996.

[70] T. Halfmann and K. Bergmann. Coherent population transfer and dark resonances in SO2. *The Journal of Chemical Physics*, 104(18):7068, 8 1998.

[71] Serge Haroche and Jean-Michel Raimond. *Exploring the Quantum: Atoms, Cavities, and Photons*. Oxford University Press, 2006.

[72] Simon Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1st edition, 1994.

[73] R. Hecht-Nielsen. Theory of the backpropagation neural network. In *Neural Networks, 1989. IJCNN., International Joint Conference on*, pages 593–605 vol.1. IEEE, June 1989.

[74] 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, August 2016.

[75] H. J. Hogben, M. Krzystyniak, G. T. P. Charnock, P. J. Hore, and Ilya Kuprov. Spinach A

software library for simulation of spin dynamics in large spin systems. *Journal of Magnetic Resonance*, 208(2):179–194, February 2011.

[76] Zhibo Hou, Han-Sen Zhong, Ye Tian, Daoyi Dong, Bo Qi, Li Li, Yuanlong Wang, Franco Nori, Guo-Yong Xiang, Chuan-Feng Li, and Guang-Can Guo. Full reconstruction of a 14-qubit state within four hours. *New Journal of Physics*, 18(8):083036, 2016.

[77] Ming Hua, Ming-Jie Tao, and Fu-Guo Deng. Fast universal quantum gates on microwave photons with all-resonance operations in circuit QED. *Scientific Reports*, 5(9274):9274, Mar 2015.

[78] D. Hucul, I. V. Inlek, G. Vittorini, C. Crocker, S. Debnath, S. M. Clark, and C. Monroe. Modular entanglement of atomic qubits using photons and phonons. *Nat. Phys.*, 11(1):37–42, Jan 2015.

[79] Kurt Jacobs, Rebing Wu, Xiaoting Wang, Sahel Ashhab, Qi-Ming Chen, and Herschel Rabitz. Fast quantum communication in linear networks. *EPL (Europhysics Letters)*, 114(4):40007, 5 2016.

[80] Daniel F. V. James, Paul G. Kwiat, William J. Munro, and Andrew G. White. Measurement of qubits. *Physical Review A*, 64(5):052312, 10 2001.

[81] Antony Jameson, Wolfgang Schmidt, and Eli Turkel. Numerical solution of the euler equations by finite volume methods using runge kutta time stepping schemes. In *14th fluid and plasma dynamics conference*, page 1259, 1981.

[82] 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, 12 2007.

[83] J. R. Johansson, P. D. Nation, and Franco Nori. QuTiP: An open-source Python frame-

work for the dynamics of open quantum systems. *Computer Physics Communications*, 183(8):1760–1772, August 2012.

[84] J. R. Johansson, P. D. Nation, and Franco Nori. QuTiP 2: A Python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 184(4):1234–1240, April 2013.

[85] Bruce E Kane. A silicon-based nuclear spin quantum computer. *nature*, 393(6681):133, 1998.

[86] J.ín Kelly et al. Optimal Quantum Control Using Randomized Benchmarking. *Phys. Rev. Lett.*, 112:240504, June 2014.

[87] J.ín Kelly et al. State preservation by repetitive error detection in a superconducting quantum circuit. *Nature*, 519(7541):66–69, March 2015.

[88] Keysight Technologies. *M8196A 92 GSa/s Arbitrary Waveform Generators — Keysight (formerly Agilent's Electronic Measurement)*, 2016.

[89] 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, February 2005.

[90] David Kielpinski, Chris Monroe, and David J Wineland. Architecture for a large-scale ion-trap quantum computer. *Nature*, 417(6890):709, 2002.

[91] H Jeff Kimble. The quantum internet. *Nature*, 453(7198):1023–1030, 2008.

[92] Diederik Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization, July 2015.

[93] Emanuel Knill, D Leibfried, R Reichle, J Britton, RB Blakestad, JD Jost, C Langer, R Ozeri, S Seidelin, and DJ Wineland. Randomized benchmarking of quantum gates. *Phys. Rev. A*, 77(1):012307, 2008.

[94] Kyryl Kobzar, Thomas E. Skinner, Navin Khaneja, Steffen J. Glaser, and Burkhard Luy. Exploring the limits of broadband excitation and inversion pulses. *Journal of Magnetic Resonance*, 170(2):236–243, October 2004.

[95] Kyryl Kobzar, Thomas E. Skinner, Navin Khaneja, Steffen J. Glaser, and Burkhard Luy. Exploring the limits of broadband excitation and inversion: II. Rf-power optimized pulses. *Journal of Magnetic Resonance*, 194(1):58–66, September 2008.

[96] 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. *Phys. Rev. A*, 76(4):042319, October 2007.

[97] 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. *Phys. Rev. A*, 76:042319, Oct 2007.

[98] Robert L. Kosut, Matthew D. Grace, and Constantin Brif. Robust control of quantum gates via sequential convex programming. *Phys. Rev. A*, 88:052326, November 2013.

[99] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff. Deterministic quantum state transfer and remote entanglement using microwave photons. *Nature*, 558(7709):264–267, 6 2018.

[100] Paul G Kwiat, Salvador Barraza-Lopez, Andre Stefanov, and Nicolas Gisin. Experimental entanglement distillation and hiddennon-locality. *Nature*, 409(6823):1014, 2001.

[101] Victor W. Lee, Changkyu Kim, Jatin Chhugani, Michael Deisher, Daehyun Kim, Anthony D. Nguyen, Nadathur Satish, Mikhail Smelyanskiy, Srinivas Chennupaty, Per Hammarlund, Ronak Singhal, and Pradeep Dubey. Debunking the 100X GPU vs. CPU Myth: An Evaluation of Throughput Computing on CPU and GPU. *SIGARCH Comput. Archit. News*, 38(3):451–460, June 2010.

[102] PJ Leek, S Filipp, P Maurer, M Baur, R Bianchetti, JM Fink, M Göppl, L Steffen, and A Wallraff. Using sideband transitions for two-qubit operations in superconducting circuits. *Physical Review B*, 79(18):180511, 2009.

[103] Zaki Leghtas, Gerhard Kirchmair, Brian Vlastakis, Robert J Schoelkopf, Michel H Devoret, and Mazyar Mirrahimi. Hardware-efficient autonomous quantum memory protection. *Phys. Rev. Lett.*, 111(12):120501, 2013.

[104] 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(1):18, 2019.

[105] Nelson Leung, Mohamed Abdelhafez, Jens Koch, and David Schuster. Speedup for quantum optimal control from automatic differentiation based on graphics processing units. *Physical Review A*, 95(4):042318, 4 2017.

[106] Per J. Liebermann and Frank K. Wilhelm. Optimal Qubit Control Using Single-Flux Quantum Pulses. *Phys. Rev. Applied*, 6:024022+, August 2016.

[107] Norbert M Linke, Dmitri Maslov, Martin Roetteler, Shantanu Debnath, Caroline Figgatt, Kevin A Landsman, Kenneth Wright, and Christopher Monroe. Experimental comparison of two quantum computing architectures. *Proceedings of the National Academy of Sciences*, 114:3305–3310, 2017.

[108] Weifeng Liu and Brian Vinter. An Efficient GPU General Sparse Matrix-Matrix Multiplication for Irregular Data. In *2014 IEEE 28th International Parallel and Distributed Processing Symposium*, pages 370–381. IEEE, May 2014.

[109] Yao Lu, S. Chakram, N. Leung, N. Earnest, R. K. Naik, Ziwen Huang, Peter Groszkowski, Eliot Kapit, Jens Koch, and David I. Schuster. Universal stabilization of a parametrically coupled qubit. *Phys. Rev. Lett.*, 119:150502, Oct 2017.

[110] Yao Lu, S. Chakram, N. Leung, N. Earnest, R.K. Naik, Ziwen Huang, Peter Groszkowski, Eliot Kapit, Jens Koch, and David I. Schuster. Universal Stabilization of a Parametrically Coupled Qubit. *Physical Review Letters*, 119(15):150502, 10 2017.

[111] Ruichao Ma, Clai Owens, Aman LaChapelle, David I Schuster, and Jonathan Simon. Hamiltonian tomography of photonic lattices. *Phys. Rev. A*, 95(6):062120, 2017.

[112] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Gruslys, S. Schirmer, and Thomas Schulte-Herbrüggen. Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework. *Phys. Rev. A*, 84:022305, August 2011.

[113] Shai Machnes, David J. Tannor, Frank K. Wilhelm, and Elie Assémat. Gradient optimization of analytic controls: the route to high accuracy quantum optimal control, July 2015.

[114] Yvon Maday and Gabriel Turinici. New formulations of monotonically convergent quantum control algorithms. *The Journal of Chemical Physics*, 118(18):8191–8196, May 2003.

[115] 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, October 2009.

[116] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, 2016.

[117] 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. *Phys. Rev. Appl.*, 6(6):064007, 2016.

[118] 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):064007, 12 2016.

[119] 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):064007, 12 2016.

[120] David C. McKay, Ravi Naik, Philip Reinhold, Lev S. Bishop, and David I. Schuster. High-Contrast Qubit Interactions Using Multimode Cavity QED. *Phys. Rev. Lett.*, 114:080501+, February 2015.

[121] David C. McKay, Ravi Naik, Philip Reinhold, Lev S. Bishop, and David I. Schuster. High-contrast qubit interactions using multimode cavity QED. *Phys. Rev. Lett.*, 114:080501, Feb 2015.

[122] David C. McKay, Ravi Naik, Philip Reinhold, Lev S. Bishop, and David I. Schuster. High-Contrast Qubit Interactions Using Multimode Cavity QED. *Physical Review Letters*, 114(8):080501, 2 2015.

[123] N David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65(15):1838, 1990.

[124] Marios H Michael, Matti Silveri, RT Brierley, Victor V Albert, Juha Salmilehto, Liang Jiang, and Steven M Girvin. New class of quantum error-correcting codes for a bosonic mode. *Phys. Rev. X*, 6(3):031006, 2016.

[125] 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 Journal of Physics*, 16(4):045014+, December 2013.

[126] 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 Journal of Physics*, 16(4):045014, 2014.

[127] Alexander E Mohr, Eve A Riskin, and Richard E Ladner. Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction. *IEEE journal on selected areas in communications*, 18(6):819–828, 2000.

[128] Cleve Moler and Charles Van Loan. Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later : SIAM Review: Vol. 45, No. 1 (Society for Industrial and Applied Mathematics). *SIAM Review*, 45(1):3–49, August 2006.

[129] 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):022317, 2 2014.

[130] Christopher R Monroe, Robert J Schoelkopf, and Mikhail D Lukin. Quantum connections. *Scientific American*, 314(5):50–57, 2016.

[131] Peter J Mosley, Jeff S Lundeen, Brian J Smith, Piotr Wasylczyk, Alfred B URen, Christine Silberhorn, and Ian A Walmsley. Heralded generation of ultrafast single photons in pure quantum states. *Physical Review Letters*, 100(13):133601, 2008.

[132] F. Motzoi, J. M. Gambetta, S. T. Merkel, and F. K. Wilhelm. Optimal control methods for rapidly time-varying Hamiltonians. *Phys. Rev. A*, 84(2):022307, August 2011.

[133] R. K. Naik, N. Leung, S. Chakram, Peter Groszkowski, Y. Lu, N. Earnest, D. C. McKay, Jens Koch, and D. I. Schuster. Random access quantum information processors using multimode circuit quantum electrodynamics. *Nature Communications*, 8(1):1904, 12 2017.

[134] 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, 9 2016.

[135] V. Nebendahl, H. Häffner, and C. F. Roos. Optimal control of entangling operations for trapped-ion quantum computing. *Phys. Rev. A*, 79(1):012312, January 2009.

[136] Francois Nguyen, Eva Zakka-Bajjani, Raymond W. Simmonds, and José Aumentado. Quantum interference between two single photons of different microwave frequencies. *Phys. Rev. Lett.*, 108:163602, Apr 2012.

[137] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, Cambridge University Press, 2000.

[138] R. Nigmatullin and S. G. Schirmer. Implementation of fault-tolerant quantum logic gates via optimal control. *New Journal of Physics*, 11(10):105032, October 2009.

[139] Omid Noroozian, Peter K. Day, Byeong Ho Eom, Henry G. Leduc, and Jonas Zmuidzinas. Crosstalk Reduction for Superconducting Microwave Resonator Arrays. *IEEE Transactions on Microwave Theory and Techniques*, 60(5):1235–1243, 5 2012.

[140] Jeremy L O'Brien, GJ Pryde, Alexei Gilchrist, DFV James, NK Langford, TC Ralph, and AG White. Quantum process tomography of a controlled-not gate. *Phys. Rev. Lett.*, 93(8):080502, 2004.

[141] Nissim Ofek, Andrei Petrenko, Reinier Heeres, Philip Reinhold, Zaki Leghtas, Brian Vlastakis, Yehan Liu, Luigi Frunzio, SM Girvin, L Jiang, et al. Extending the lifetime of a quantum bit with error correction in superconducting circuits. *Nature*, 536:441–445, 2016.

[142] Kyoung-Su Oh and Keechul Jung. GPU implementation of neural networks. *Pattern Recognition*, 37(6):1311–1314, June 2004.

[143] Yukiyoshi Ohtsuki, Yoshiaki Teranishi, Peter Saalfrank, Gabriel Turinici, and Herschel Rabitz. Monotonically convergent algorithms for solving quantum optimal control problems described by an integrodifferential equation of motion. *Phys. Rev. A*, 75:033407, March 2007.

[144] Yukiyoshi Ohtsuki, Gabriel Turinici, and Herschel Rabitz. Generalized monotonically convergent algorithms for solving quantum optimal control problems. *The Journal of Chemical Physics*, 120(12):5509–5517, March 2004.

[145] Roberto Olivares-Amaya, Mark A Watson, Richard G Edgar, Leslie Vogt, Yihan Shao, and Alán Aspuru-Guzik. Accelerating correlated quantum chemistry calculations using graphical processing units and a mixed precision matrix multiplication library. *Journal of chemical theory and computation*, 6(1):135–144, 2009.

[146] José P. Palao and Ronnie Kosloff. Optimal control theory for unitary transformations. *Phys. Rev. A*, 68(6):062308+, December 2003.

[147] José P. Palao, Ronnie Kosloff, and Christiane P. Koch. Protecting coherence in optimal control theory: State-dependent constraint approach. *Phys. Rev. A*, 77(6):063412, June 2008.

[148] Fernando Pastawski, Beni Yoshida, Daniel Harlow, and John Preskill. Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence. *Journal of High Energy Physics*, 2015(6):1–55, 2015.

[149] M. Pechal, L. Huthmacher, C. Eichler, S. Zeytinoğlu, 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, 10 2014.

[150] Felix Platzer, Florian Mintert, and Andreas Buchleitner. Optimal Dynamical Control of Many-Body Entanglement. *Phys. Rev. Lett.*, 105:020501, July 2010.

[151] Rajat Raina, Anand Madhavan, and Andrew Y. Ng. Large-scale Deep Unsupervised Learning Using Graphics Processors. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 873–880, New York, NY, USA, 2009. ACM.

[152] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Advanced lectures on machine learning*, pages 63–71. Springer, 2004.

[153] Robert Raussendorf, Daniel E Browne, and Hans J Briegel. Measurement-based quantum computation on cluster states. *Phys. Rev. A*, 68(2):022312, 2003.

[154] 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. *App. Phys. Lett.*, 102(19):192604, 2013.

[155] Matthew Reagor, Wolfgang Pfaff, Christopher Axline, Reinier W Heeres, Nissim Ofek, Katrina Sliwa, Eric Holland, Chen Wang, Jacob Blumoff, Kevin Chou, et al. Quantum memory with millisecond coherence in circuit qed. *Phys. Rev. B*, 94(1):014506, 2016.

[156] P. Rebentrost, I. Serban, Thomas Schulte-Herbrüggen, and F. K. Wilhelm. Optimal Control of a Qubit Coupled to a Non-Markovian Environment. *Phys. Rev. Lett.*, 102(9):090401, March 2009.

[157] P. Rebentrost and F. K. Wilhelm. Optimal control of a leaking qubit. *Phys. Rev. B*, 79(6):060507, February 2009.

[158] M. D. Reed, L. DiCarlo, B. R. Johnson, L. Sun, D. I. Schuster, L. Frunzio, and R. J. Schoelkopf. High-fidelity readout in circuit quantum electrodynamics using the Jaynes-Cummings nonlinearity. *Phys. Rev. Lett.*, 105:173601, Oct 2010.

[159] Daniel M. Reich, Mamadou Ndong, and Christiane P. Koch. Monotonically convergent optimization in quantum control using Krotov's method. *The Journal of Chemical Physics*, 136(10):104103, March 2012.

[160] Gerhard Rempe, Herbert Walther, and Norbert Klein. Observation of quantum collapse and revival in a one-atom maser. *Phys. Rev. Lett.*, 58(4):353, 1987.

[161] Gerhard Rempe, Herbert Walther, and Norbert Klein. Observation of quantum collapse and revival in a one-atom maser. *Physical Review Letters*, 58(4):353–356, 1 1987.

[162] 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, 4 2014.

[163] A Romanenko, A Grassellino, AC Crawford, DA Sergatskov, and O Melnychuk. Ultra-high quality factors in superconducting niobium cavities in ambient magnetic fields up to 190 mG. *App. Phys. Lett.*, 105(23):234103, 2014.

[164] D. Rosenberg, D. Kim, R. Das, D. Yost, S. Gustavsson, D. Hover, P. Krantz, A. Melville, L. Racz, G. O. Samach, S. J. Weber, F. Yan, J. L. Yoder, A. J. Kerman, and W. D. Oliver. 3D integrated superconducting qubits. *npj Quantum Information*, 3(1):42, 12 2017.

[165] SR Schofield, NJ Curson, MY Simmons, FJ Rueß, T Hallam, L Oberbeck, and RG Clark. Atomically precise placement of single dopants in si. *Physical review letters*, 91(13):136104, 2003.

[166] I. Serban, J. Werschnik, and E. K. U. Gross. Optimal control of time-dependent targets. *Phys. Rev. A*, 71:053810, May 2005.

[167] Toby Sharp. Implementing Decision Trees and Forests on a GPU. In David Forsyth, Philip Torr, and Andrew Zisserman, editors, *Computer Vision ECCV 2008*, volume 5305 of *Lecture Notes in Computer Science*, chapter 44, pages 595–608. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[168] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.

[169] Roger B. Sidje. Expokit: A Software Package for Computing Matrix Exponentials. *ACM Trans. Math. Softw.*, 24(1):130–156, March 1998.

[170] A. J. Sirois, M. A. Castellanos-Beltran, M. P. DeFeo, L. Ranzani, F. Lecocq, R. W. Simmonds, J. D. Teufel, and J. Aumentado. Coherent-state storage and retrieval between superconducting cavities using parametric frequency conversion. *Applied Physics Letters*, 106(17):172603, 4 2015.

[171] A. J. Sirois, M. A. Castellanos-Beltran, M. P. DeFeo, L. Ranzani, F. Lecocq, R. W. Simmonds, J. D. Teufel, and J. Aumentado. Coherent-state storage and retrieval between superconducting cavities using parametric frequency conversion. *Applied Physics Letters*, 106(17):172603, 4 2015.

[172] AJ Sirois, MA Castellanos-Beltran, MP DeFeo, L Ranzani, F Lecocq, RW Simmonds, JD Teufel, and J Aumentado. Coherent-state storage and retrieval between superconducting cavities using parametric frequency conversion. *App. Phys. Lett.*, 106(17):172603, 2015.

[173] Thomas E. Skinner, Timo O. Reiss, Burkhard Luy, Navin Khaneja, and Steffen J. Glaser. Reducing the duration of broadband excitation pulses using optimal control with limited RF amplitude. *Journal of magnetic resonance (San Diego, Calif. : 1997)*, 167(1):68–74, March 2004.

[174] Shlomo E. Sklarz and David J. Tannor. Loading a Bose-Einstein condensate onto an optical lattice: An application of optimal control theory to the nonlinear Schrödinger equation. *Phys. Rev. A*, 66(5):053619, November 2002.

[175] Chunhua Song, Michael P DeFeo, Kang Yu, and Britton LT Plourde. Reducing microwave loss in superconducting resonators due to trapped vortices. *App. Phys. Lett.*, 95(23):232501, 2009.

[176] A. Spörl, Thomas Schulte-Herbrüggen, S. J. Glaser, V. Bergholm, M. J. Storcz, J. Fer-

ber, and F. K. Wilhelm. Optimal control of coupled Josephson qubits. *Phys. Rev. A*, 75(1):012302, January 2007.

[177] 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*, 89(3):033857, 3 2014.

[178] Andrew Steane. Multiple-particle interference and quantum error correction. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 452, pages 2551–2577. The Royal Society, 1996.

[179] D. Steinkraus, I. Buck, and P. Y. Simard. Using GPUs for machine learning algorithms. In *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*, ICDAR '05, pages 1115–1120 Vol. 2, Washington, DC, USA, August 2005. IEEE.

[180] M. Stobińska, G. Alber, and G. Leuchs. Perfect excitation of a matter qubit by a single photon in free space. *EPL (Europhysics Letters)*, 86(1):14007, 4 2009.

[181] J. D. Strand, Matthew Ware, Félix Beaudoin, T. A. Ohki, B. R. Johnson, Alexandre Blais, and B. L. T. Plourde. First-order sideband transitions with flux-driven asymmetric transmon qubits. *Phys. Rev. B*, 87:220505, Jun 2013.

[182] Neereja M. Sundaresan, Yanbing Liu, Darius Sadri, László J. Szőcs, Devin L. Underwood, Moein Malekakhlagh, Hakan E. Türeci, and Andrew A. Houck. Beyond strong coupling in a multimode cavity. *Phys. Rev. X*, 5:021035, Jun 2015.

[183] Hiroki Takahashi, Jonas S Neergaard-Nielsen, Makoto Takeuchi, Masahiro Takeoka, Kazuhiro Hayasaka, Akira Furusawa, and Masahide Sasaki. Entanglement distillation from gaussian input states. *Nature photonics*, 4(3):178, 2010.

[184] DavidJ Tannor, Vladimir Kazakov, and Vladimir Orlov. Control of Photochemical Branching: Novel Procedures for Finding Optimal Pulses and Global Upper Bounds. In J. Broeckhove and L. Lathouwers, editors, *Time-Dependent Quantum Molecular Dynamics*, volume 299 of *Nato ASI Series*, pages 347–360. Springer US, 1992.

[185] Theano Development Team. Theano: A Python framework for fast computation of mathematical expressions. *arXiv e-prints*, abs/1605.02688, May 2016.

[186] John S Toll. Causality and the dispersion relation: logical foundations. *Phys. Rev.*, 104(6):1760, 1956.

[187] Zdeněk Tošner, Thomas Vosegaard, Cindie Kehlet, Navin Khaneja, Steffen J. Glaser, and Niels C. Nielsen. Optimal control in NMR spectroscopy: Numerical implementation in SIMPSON. *Journal of Magnetic Resonance*, 197(2):120–134, April 2009.

[188] Ivan S Ufimtsev and Todd J Martinez. Graphical processing units for quantum chemistry. *Computing in Science & Engineering*, 10(6):26–34, 2008.

[189] D L Underwood, W E Shanks, Jens Koch, and A A Houck. Low-disorder microwave cavity lattices for quantum simulation with photons. *Physical Review A*, 86, 2012.

[190] G. S. Vasilev, A. Kuhn, and N. V. Vitanov. Optimum pulse shapes for stimulated Raman adiabatic passage. *Physical Review A*, 80(1):013417, 7 2009.

[191] V Vesterinen, O-P Saira, A Bruno, and L DiCarlo. Mitigating information leakage in a crowded spectrum of weakly anharmonic qubits. *arXiv preprint arXiv:1405.0450*, 2014.

[192] NV Vitanov, KA Suominen, and BW Shore. Creation of coherent atomic superpositions by fractional stimulated raman adiabatic passage. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 32(18):4535, 1999.

[193] 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, November 2013.

[194] Leslie Vogt, Roberto Olivares-Amaya, Sean Kermes, Yihan Shao, Carlos Amador-Bedolla, and Alán Aspuru-Guzik. Accelerating resolution-of-the-identity second-order møller- plesset quantum chemistry calculations with graphical processing units. *The Journal of Physical Chemistry A*, 112(10):2049–2057, 2008.

[195] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Approaching Unit Visibility for Control of a Superconducting Qubit with Dispersive Readout. *Physical Review Letters*, 95(6):060501, 8 2005.

[196] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Approaching unit visibility for control of a superconducting qubit with dispersive readout. *Phys. Rev. Lett.*, 95:060501, Aug 2005.

[197] Yimin Wang, Ji Minář, Lana Sheridan, and Valerio Scarani. Efficient excitation of a two-level atom by a single photon in a propagating mode. *Physical Review A*, 83(6):063842, 6 2011.

[198] Paul Watts, J. Vala, Matthias M. Müller, Tommaso Calarco, K. Birgitta Whaley, Daniel M. Reich, Michael H. Goerz, and Christiane P. Koch. Optimizing for an arbitrary perfect entangler. I. Functionals. *Phys. Rev. A*, 91:062306+, June 2015.

[199] R. E. Wengert. A Simple Automatic Derivative Evaluation Program. *Commun. ACM*, 7(8):463–464, August 1964.

[200] J Wenner, M Neeley, Radoslaw C Bialczak, M Lenander, Erik Lucero, A D OConnell, D Sank, H Wang, M Weides, A N Cleland, and John M Martinis. Wirebond crosstalk and cavity modes in large chip mounts for superconducting qubits. *Superconductor Science and Technology*, 24(6):065001, 6 2011.

[201] J. Wenner, Yi Yin, Yu Chen, R. Barends, B. Chiaro, E. Jeffrey, J. Kelly, A. Megrant, J.Y. Mutus, C. Neill, P.J.J. OMalley, 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):210501, 5 2014.

[202] Karl A. Wilkinson, Paul Sherwood, Martyn F. Guest, and Kevin J. Naidoo. Acceleration of the GAMESS-UK electronic structure package on graphical processing units. *Journal of computational chemistry*, 32(10):2313–2318, July 2011.

[203] Yi Yin, Yu Chen, Daniel Sank, P. J. J. OMalley, 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, 3 2013.

[204] Eva Zakka-Bajjani, Francois Nguyen, Minhyea Lee, Leila R. Vale, Raymond W. Simmonds, and Jose Aumentado. Quantum superposition of a single microwave photon in two different 'colour' states. *Nat. Phys.*, 7(8):599–603, Aug 2011.

[205] Wusheng Zhu, Jair Botina, and Herschel Rabitz. Rapidly convergent iteration methods for quantum optimal control of population. *The Journal of Chemical Physics*, 108(5):1953–1963, February 1998.

[206] Wusheng Zhu and Herschel Rabitz. A rapid monotonically convergent iteration algorithm for quantum optimal control over the expectation value of a positive definite operator. *The Journal of Chemical Physics*, 109(2):385–391, July 1998.