Theory and Mitigation of Crosstalk on Quantum Information Processors

by

Adam Winick

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics

Waterloo, Ontario, Canada, 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Successfully implementing large-scale quantum computation has proven to be an exceptionally arduous task. Decoherence and imperfect control limit the coherent manipulation of large ensembles of particles. While quantum error correction provides robust schemes for executing quantum algorithms on error-prone systems, the methods usually assume that the errors are well-behaved and lie below some threshold. The burden of QEC can be substantial, and reaching error rates well below these thresholds can dramatically improve the processing capabilities of a device.

A hierarchy of error processes exists with increasingly desirable properties at the cost of realism and generality. For example, quantum circuits subject to Markovian errors typically have higher error thresholds than those under general errors. We can further divide Markovian errors into coherent and incoherent processes, with the former having much lower thresholds.

This thesis examines crosstalk, a type of coherent error process, and mainly studies its role in superconducting quantum computing devices.

The first part of our work details a systematic framework for modeling crosstalk that occurs during the operation of a quantum computer, i.e., what happens on the device while performing gates. We break this crosstalk down into local and nonlocal effects. We show how to model local crosstalk on a digital computer without approximations efficiently. Unlike local crosstalk, nonlocal crosstalk cannot be modeled efficiently on a digital computer without approximations. Thus, we develop a framework for approximating the effect of nonlocal crosstalk. We observed a negligible difference between our approximation and the exact system dynamics in typical systems.

The second part of this thesis details our attempts to characterize and efficiently mitigate crosstalk on fixed-frequency superconducting qubits experimentally. The first obstacle we encountered was learning the crosstalk affecting a system. When the crosstalk is weak, existing methods prove difficult, so we developed a new approach to measure crosstalk. The second problem we needed to address was verifying that our model was correct. Using the results from our first measurements, we compare predicted evolutions with experimental data in a setting much different than the measurement procedure. We see excellent agreement between experiment and theory, indicating the model is reasonable. The last outstanding puzzle piece in this investigation is using this model to mitigate crosstalk, and our research is ongoing.

Acknowledgements

I want to thank my supervisor Joseph Emerson for his guidance and support over the years I have worked with him in academia and at Quantum Benchmark. I also want to thank my thesis committee members, Joel J. Wallman and David Cory. This thesis was only made possible by the assistance and dedication of my experimental collaborators Jan Balewski, Gang Huang, and Yilun Xu at Lawrence Berkeley National Laboratory. Finally, I would like to thank my parents and my partner, Sara, without whom none of this would have been possible.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CR cross-resonance viii, 21, 22

DRAG derivative removal by adiabatic gate viii, 19, 20

IQ in-phase and quadrature 17

NISQ noisy intermediate-scale quantum 2, 7

QEC Quantum Error Correction 1

QIP Quantum Information Processor 1

RWA rotating wave approximation 17, 21

SPAM state preparation and measurement errors 9, 27

Chapter 1

Introduction

Over the past two decades, Quantum Information Processors (QIPs) have progressed from two-qubit novelties [9] to finely-tuned complex systems consisting of more than one hundred qubits [7]. Progress in processor fabrication and experimental techniques, as well as theoretical developments improving both the control and understanding of quantum systems, has made this remarkable growth possible. Nevertheless, engineering such systems have proven exceedingly difficult. Decoherence and imperfect control limit the coherent manipulation of large ensembles of particles. While Quantum Error Correction (QEC) [42, 16, 20] provides robust schemes for executing quantum algorithms on error-prone systems, the methods usually require that errors are well-behaved and lie below some threshold [2, 43]. The burden of QEC can be substantial, and reaching error rates well below these thresholds can dramatically improve the processing capabilities of a device.

A central part of the problem is the inherent dichotomy between the implementation of simultaneous one-qubit and two-qubit gates. Fast, high-fidelity two-qubit gates require qubits that are spatially and spectrally near one another. Such closeness reduces the addressability of the constituent subsystems because a control pulse intended for one part can interact with the other [15, 33]. Researchers have typically dealt with this crosstalk by either maximizing the gap between qubits or executing local operations asynchronously [40, 32, 4]. The former solution requires the ability to tune couplings or extra engineering. However, the added complexity can dramatically impact coherence times and requires additional control wires. As QIPs grow, the 2-3 times more wires needed for hundreds or thousands of qubits will become an immense hurdle. In the latter approach, depending on the degree to which the control fields affect neighboring subsystems, the time overhead can be significant and is directly proportional to the size of the neighborhood.

Crosstalk describes a broad range of effects that violate one of two assumptions: spatial locality and independence of operations [37, 38, 1]. Gates and other operations are supposed to act on disjoint subsets of qubits. However, unintended interactions can couple the qubits, producing nonlocal correlated noise. Even if an operation has a well-defined action on a particular subset of qubits, the effective noise might depend on its context—what operations affect other qubits.

1.1 Modeling crosstalk

In the first part of this thesis, we introduce a scalable framework for accurately modeling idle and operation crosstalk on experimental devices. Our technique exploits the tensor product structure of local (classical) crosstalk to express its impact on gates efficiently. Through a perturbative expansion, we extend our ideas to nonlocal (quantum) crosstalk and capture its effects on elements of $\mathrm{SU}(2)^{\otimes n}$ on superconducting transmon qubits. Despite substantial local crosstalk, we show that error rates near the crosstalk-free limit are possible with modern control hardware. We further show how to tuneup simultaneous cross-resonance gates and, again, obtain dramatically lower error rates. Our results suggest that contrary to prevailing opinions [10, 30, 35, 29], crosstalk need not be a prohibitive limitation on noisy intermediate-scale quantum (NISQ) era devices [34]. Higher-quality quantum information processors may be made possible by using our techniques to better balance device fabrication and pulse design tradeoffs.

1.2 Measuring crosstalk

The second part of this thesis is dedicated to measuring crosstalk on experimental devices. Our approach involves analyzing the Hamiltonian for simultaneously driven uncoupled fixed-frequency transmons. Through careful analysis, we discover a new method called the double field method, which is a generalization of the Rabi experiment and is better suited for estimating crosstalk. With this new tool, we can determine the crosstalk parameters of a device with greater accuracy and fewer measurements than with standard Rabi experiments. We use the fit data from double field experiments to compare the theoretical predictions and experimental results, and we observe that they are in close agreement, validating our approach. Additionally, we determine the theoretical Hamiltonian corresponding to three driving fields to ensure the accuracy of our model.

Chapter 2

Background

This chapter briefly examines some of the background material necessary for describing our results.

2.1 Quantum theory

A system in classical mechanics is always well-defined. There exists a collection of variables that completely describes its current and future states. In a quantum mechanical setting, this is not the case. We must allow a superposition of possible states to describe a quantum system fully. Formally, a Hilbert space vector describes a quantum state. In quantum theory, we require that the Hilbert space is separable: the space has a countable orthonormal basis. The separability property ensures we can express any state as a linear combination of orthonormal basis elements. For example,

$$|\psi\rangle = \sum_k \alpha_k |\psi_k\rangle \ , \tag{2.1}$$

where α_k are complex numbers and $|\psi_k\rangle$ are basis elements. The coefficients α_k denote probability amplitudes.

The Schrödinger equation is a linear partial differential equation that governs the evolution of a quantum system. The appropriate form of the equation depends on the physical situation. In our case, it is natural to consider the most general time-dependent formulation. Concretely the equation reads

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle \ , \tag{2.2}$$

where H is the possibly time-dependent Hamiltonian of the system, and $\hbar$ is the reduced Planck constant. By working in natural units, we set the constant equal to one.

2.1.1 Unitary evolution

Given the system's initial state, we might want an operator that maps this state to the system's final state. This operator U is known as the time evolution operator,

$$|\psi\rangle = U\,|\psi_0\rangle\ . \tag{2.3}$$

It is unitary and, therefore, preserves the inner product between states. In other words, unitarity ensures the conservation of probability. For example,

$$\langle\psi|\psi\rangle = \langle\psi_0|\,U^\dagger U\,|\psi_0\rangle = 1\,. \tag{2.4}$$

2.1.2 Average Hamiltonian theory

A helpful technique for understanding Hamiltonians is the Magnus expansion. Applying this expansion yields effective Hamiltonians, which we can analyze to understand the behavior of an analytically intractable problem. Given a time-dependent Hamiltonian $H(t)$, we look for a time-independent Hamiltonian $\bar{H}$ such that $U(t) \approx \exp\bigl(-it\bar{H}\bigr)$. We expand the average Hamiltonian $\bar{H}$ as

$$\bar{H} = \bar{H}^{(1)} + \bar{H}^{(2)} + \bar{H}^{(3)} + \dots\,, \tag{2.5}$$

where the first two orders of the Magnus expansion are given by

$$\bar{H}^{(1)} = \frac{1}{t}\int_0^t dt_1 H(t_1)\,, \tag{2.6}$$

$$\bar{H}^{(2)} = \frac{1}{2it}\int_0^t dt_1 \int_0^{t_1} dt_2 [H(t_1), H(t_2)]\,. \tag{2.7}$$

2.2 Open quantum systems

It is impossible to completely isolate a quantum system that one might want to study from its environment. Therefore, developing a theory for open quantum systems for treating

these interactions is necessary to obtain an accurate description of a quantum system. We describe a few of these ideas that relate to the latter chapters.

A density operator ρ describes the state of a subsystem that interacts with its environment. The scalar product

$$\langle A \rangle = \sum_k a_k p_k = \mathrm{Tr}\left(\sum_k a_k \left| a_k \right\rangle \! \left\langle a_k \right| \rho \right) = \mathrm{Tr}(A\rho)\,, \tag{2.8}$$

yields the expectation value of a measurement, or an observable. It's impossible to determine if the entire system is pure based solely on the observables of the subsystem. This is especially true when the combined system has quantum entanglement, as the subsystem's state cannot be considered pure.

In order to investigate the behavior of open quantum systems over time, it is necessary to solve master equations that govern changes in the density matrix and associated observables. However, accurately modeling the system's environment as part of the larger system can be challenging due to its size and complexity. Energy and quantum coherence are two commonly studied observables, with particular attention paid to instances of quantum dissipation when energy is lost to the environment, and quantum decoherence when coherence is lost.

2.2.1 Decoherence & thermal relaxation

In this work we consider elementary, but quite reasonable models of decoherence and dissipation that affects quantum computers. These models are often referred to by their characteristic variables T_1 and T_2. They are critical concepts which refer to the duration it takes for a qubit to return its equilibrium state after being disturbed (or operated on). T_1, also known as the longitudinal relaxation time, is a measurement of the time it takes for a qubit to restore itself to its equilibrium state along the direction of the magnetic field. Meanwhile, T_2, or transverse relaxation time, refers to the length of time it takes for a qubit to lose coherence with other qubits due to random fluctuations in the magnetic field. We can express the effect on a qubit system in matrix form as

$$\rho(t) = \begin{bmatrix} 1 - \rho_{11} e^{-\Gamma_1 t} & \rho_{01} e^{-(\Gamma_1 + \Gamma_2)t/2} \\ \rho_{10} e^{-(\Gamma_1 + \Gamma_2)t/2} & \rho_{11} e^{-\Gamma_1 t} \end{bmatrix}, \tag{2.9}$$

where we defined $T_1 = 1/\Gamma_1$, and $T_2 = 2/(\Gamma_1 + \Gamma_2)$ so that

$$\frac{T_1}{T_2} \geq \frac{1}{2}\,. \tag{2.10}$$

In superconducting quantum computers one often has $T_2 < 2T_1$ which implies a combination of both effects.

2.3 Quantifying errors

A fundamental experimental and theoretical problem is to characterize the distance between an error-prone quantum process and an ideal target process. In this work, we use the average process fidelity to quantify errors and it is defined by

$$\Phi(\xi) = \int d\psi \, \langle\psi|\xi(\psi)|\psi\rangle \ , \tag{2.11}$$

where the integral is over the Haar measure $d\psi$ on state space, and it is normalized so $\int d\psi = 1$. ξ is a process that acts on some d-dimensional quantum system. The value of Φ quantifies how well ξ preserves quantum information. A value of one indicates perfect preservation, while a value of zero indicates poor preservation.

We can naturally extend Φ to measure how well a process ξ approximates an ideal process, U,

$$\Phi(\xi, U) = \int d\psi \, \langle\psi|U\xi(\psi)U^\dagger|\psi\rangle \ . \tag{2.12}$$

The process fidelity $\Phi(\xi, U) = 1$ if and only if ξ perfectly implements U. We can intuitively understand this statement by noting that $\Phi(\xi, U) = \Phi(U^\dagger \circ \xi)$.

The process fidelity has several desirable properties that make it useful in this work.

1.
$$\Phi(A_1 \otimes A_2, B_1 \otimes B_2) = \Phi(A_1, B_1)\Phi(A_2, B_2) \tag{2.13}$$

2. Φ is invariant under Pauli twirling.

Chapter 3

Theory of Crosstalk

This chapter focuses on material we published in [47]. Our objective is to understand the form of crosstalk affecting qubits during circuit execution. After defining its structure, we develop a scalable framework for accurately modeling idle and operation crosstalk on experimental devices. Our technique exploits the tensor product structure of local (classical) crosstalk to express its impact on gates efficiently. Through a perturbative expansion, we extend our ideas to nonlocal (quantum) crosstalk and capture its effects on elements of SU(2) on superconducting transmon qubits. Despite substantial local crosstalk, we show that modern control hardware makes error rates near the crosstalk-free limit possible. We further demonstrate how to tune up simultaneous cross-resonance gates and obtain dramatically lower error rates. Our results suggest that contrary to prevailing opinions, crosstalk need not be a prohibitive limitation on noisy intermediate-scale quantum NISQ era devices. Using our techniques to balance device fabrication and pulse design tradeoffs may make higher-quality quantum information processors possible.

3.1 What is crosstalk?

Crosstalk describes a broad range of effects that violate one of two assumptions: spatial locality and independence of operations [37, 38, 1]. Gates and other operations are supposed to act on disjoint subsets of qubits. However, unintended interactions can couple the qubits, producing nonlocal correlated noise. Even if an operation has a well-defined action on a particular subset of qubits, the effective noise might depend on its context – what operations affect other qubits.

Prior work has often approached the problem of implementing several operations on a collection of qubits by breaking it into a temporally disjoint sequence of gates. In contrast, Ref. [44] analyzed the problem of driving two spins with a homogenous field in the setting of NMR. However, it is unclear how to apply the method to multilevel systems such as transmons or trapped ions. Ref. [45] studied how to drive two transmons coupled to the same cavity suffering from spectral crowding with simultaneous X or Y gates with rotation angles π and $\pi/2$ (X, Y, Z denote Pauli matrices). In either case, these methods do not directly apply to many-qubit systems or handle nonlocal correlations. We aim to develop an efficient and systematic method for optimizing the implementation of nontrivial parallel operations under general crosstalk.

What crosstalk acts on physical qubits during idling or the implementation of gates (as opposed to preparation or measurement crosstalk), and how can we efficiently simulate and, consequently, try to mitigate it? It is natural to classify crosstalk as either local or nonlocal [38]. Local crosstalk can arise when a semiclassical drive field interacts with several qubits, causing unitary errors on supposedly idle qubits, but not entangling independent subsystems. Nonlocal crosstalk creates nonfactorizable correlations over system qubits and may originate from, for example, the residual static coupling between two qubits or miscalibration.

Quantifying and reducing crosstalk requires a figure of merit. Depending on the application, evaluating the average fidelity of one-qubit or two-qubit gates rather than the fidelity per clock cycle makes sense. Our ideas work in either case, but we focus on the former. Local error measures relate directly to fault-tolerance thresholds, are easier to estimate experimentally, and are more common in the literature. We show that the average local fidelity is especially simple to approximate.

3.1.1 Local crosstalk

Although local crosstalk (typically) produces correlated noise, it can be factorized and simulated efficiently on a digital computer. The induced correlations are classical and do not entangle the individual subsystems. We model local crosstalk via the Hamiltonian

$$H(t, \vec{x}) = \sum_k H_k(t, \vec{x})\,. \tag{3.1}$$

Each term H_k acts exclusively on subsystem k, and $\vec{x}$ denotes shared classical parameters that result in crosstalk. The vector $\vec{x}$ may, for example, contain the phases and amplitudes that specify drive fields. The average process fidelity Φ [31, 5] between a target

operation $U = U_1 \otimes \cdots \otimes U_n$ and the noisy implementation $\tilde{U} = \tilde{U}_1 \otimes \cdots \otimes \tilde{U}_n$, where $\tilde{U}_k = \mathcal{T} \exp\left[-i \int d\tau H_k(\tau, \vec{x})\right]$, can be expressed as

$$\Phi(U, \tilde{U}) = \prod_k \Phi(U_k, \tilde{U}_k) \,. \tag{3.2}$$

The equation holds more generally when $\{\tilde{U}_k\}$ are completely positive trace-preserving (CPTP) maps, for example, when a dissipative process also affects the system or the control parameters fluctuate over time.

3.1.2 Nonlocal crosstalk

Unlike local crosstalk, a digital computer cannot usually exactly simulate a large system affected by nonlocal crosstalk. Thus we develop a perturbative technique for simulating nonlocal crosstalk. Our approximation scheme characterizes a noise channel $\mathcal{E}$ by estimating some of the associated Pauli error rates $\{p_P\}$. The Pauli-twirled noise channel is

$$\mathcal{E}^{\mathcal{P}}(\rho) = \frac{1}{|\mathcal{P}^n|} \sum_{P \in \mathcal{P}^n} P^\dagger \mathcal{E}(P \rho P^\dagger) P \tag{3.3}$$

$$= \sum_{P \in \mathcal{P}^n} p_P P \rho P^\dagger \,, \tag{3.4}$$

where $\mathcal{P}^n$ is the Pauli group on n qubits. These error rates provide a partial description of the noise affecting a quantum system. On large experimental devices, we can scalably and estimate the parameters in a way that is robust to state preparation and measurement errors (SPAM) [12]. We might also combine the quantities to calculate holistic measures of device performance, such as the average two-qubit fidelity or global fidelity.

3.2 Graph theory model

It is helpful to sketch our approach using a graphical model of the noise (see, e.g., Ref. [11] for basic graph theory definitions). We construct a graph G where each node is a strongly interacting subsystem during an operation of interest, such as a qubit during a single-qubit gate or a two-qubit pair entangled by a cross-resonance interaction. The entire target operation is factorizable over the tensor product space partitioning defined by the

nodes. Edges denote nonlocal crosstalk that couples subsystems, and we only allow two-body coupling. We impose the constraint that the graph has limited connectivity (in a spatial spectral sense) since our approach relies on simulating subsystems. The constraint is satisfied in contemporary architectures where a majority of nodes have a degree of at most four.

3.2.1 Approximation theory

A pair of positive integers (d, o) specifies the expansion order of the noise approximation; d designates the 'environment' distance and o the maximum component order. We consider the set $\mathcal{G}_o$ of all components of all induced subgraphs of G such that the order of every component is less than or equal to o, and any component with an order less than o has the same edges as in G. I.e., we do not look at induced components with order less than o. The idea of the simulation scheme is to calculate the Pauli errors that occur on each component.

We approximate the behavior of a component $C \in \mathcal{G}_o$ by evolving it along with all vertices of distance at most d, generating a map $\mathcal{E}_{C,d}$. Next, we compute the diagonal $f_{C,d}$ of the Pauli-Liouville representation of the channel. A Walsh-Hadamard transformation W relates $f_{c,d}$ to the Pauli probability vector $\tilde{p}_{C,d}$, with $f_{C,d} = W\tilde{p}_{C,d}$ [12]. The vector $\tilde{p}_{C,d}$ is the error probability distribution for a Pauli twirled copy of $\mathcal{E}_{C,d}$. Marginalizing the error distribution over the environment produces an estimate of the local error distribution $\tilde{p}_C$ on the target component. After calculating the marginal distributions for all of the components in $\mathcal{G}_o$, we can use the theory of probabilistic graphical models [22] to construct an estimate of the entire Pauli error distribution up to some specified error weight. By truncating the distribution at some error weight, the size of the distribution scales polynomially in the number of qubits.

In practice, including the nearest environmental nodes is sufficient to compute the local error distribution with high relative precision.

3.3 Example perturbative approximations using our framework

This section details tangible examples employing our framework and presents a less abstract description of our perturbative technique for readers more interested in using our techniques. We describe examples that illustrate how our methods handle nonlocal crosstalk.

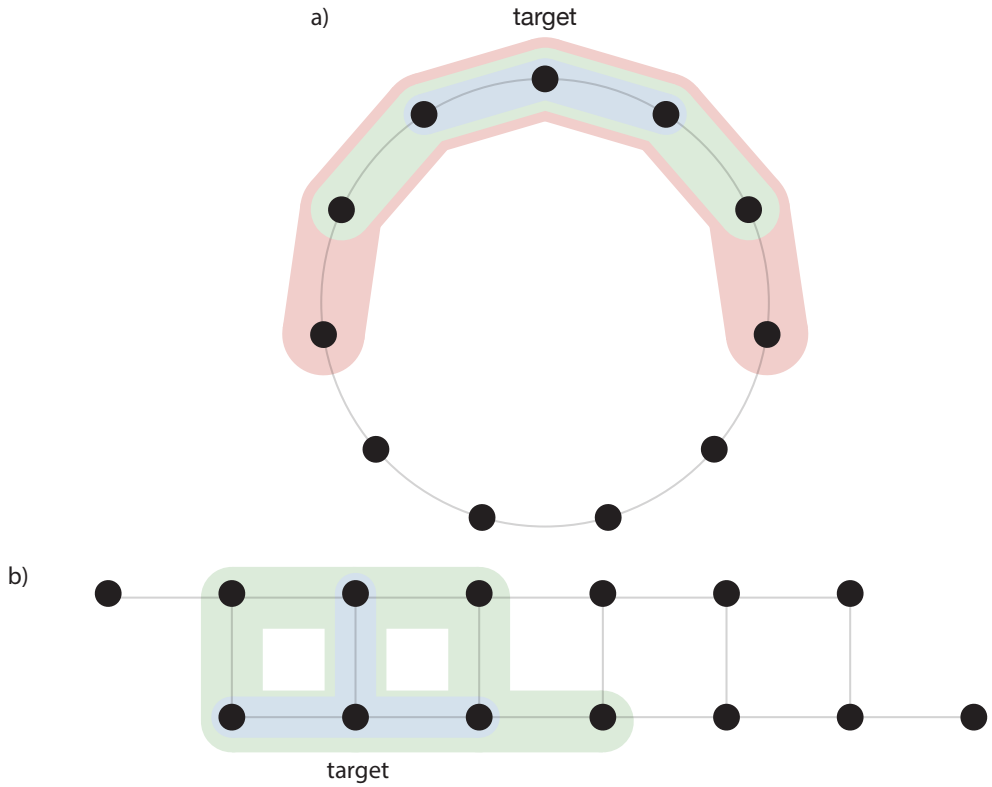

Figure 3.1: a) Connectivity graph of a ring of qubits, coupled through nearest-neighbor interactions (grey). We consider the task of simulating a target qubit. Blue, green, and red highlighting indicate the 'environment' qubits that we include in our simulation when $d = 1$, $d = 2$ and $d = 3$, respectively.
b) Connectivity graph of IBM Melbourne, with qubit-qubit couplings represented by grey lines between qubits. Note that we assume no other qubit-qubit coupling affects the device. We simulate a target qubit. Blue and green highlighting indicate the 'environment' qubits that we include in our simulation when $d = 1$, and $d = 2$, respectively.

We also bound the spatial computational complexity of each example. Local crosstalk factorizes over the tensor product of Hilbert spaces and is consequently compatible with these models, and we do not consider it below.

We introduce the symbol C to quantify the spatial complexity. We define it as the number of complex floating-point numbers in the matrix representation of the simulated unitary. Without an approximation scheme, the spatial complexity is

$$C \leq d_{\mathrm{s}}^{2N_{\mathrm{s}}}\,, \tag{3.5}$$

where d_{s} is the maximal subsystem dimension, and N_{s} is the number of subsystems. The coefficient 2 appears because we are interested in the size of a matrix rather than a vector. Alternatively, under our perturbative approximation, the spatial cost C_d when modeling neighbors of distance at most d is

$$C_d \leq N_{\mathrm{s}}(d_{\mathrm{s}})^{2[1+N_{\mathrm{n}}\sum_{i=0}^{d-1}(N_{\mathrm{n}}-1)^i]} \tag{3.6}$$

where N_{n} is the maximum degree of the graphical representation of the noise (the maximum number of neighbors of a subsystem). We also take o (the maximum component order) to be the largest subsystem's size. The target system has at most N_{n} neighbors, while each of its neighboring subsystems has at most $N_{\mathrm{n}}-1$ neighbors not already in the model. Hence the $N_{\mathrm{n}}-1$ in the equation.

For fixed $d_{\mathrm{s}}, N_{\mathrm{n}}$, and d, C_d scales linearly with the number of subsystems N_{s}, while C scales exponentially. For one or two qubits, the scaling differences are irrelevant. However, for a system with 10-20 qubits where none of the constituent qubits have a high degree (connected to many other qubits), our approach can be more than one million times more efficient in terms of spatial complexity. For 25+ qubits, we can efficiently simulate our perturbative approximation of a system's unitary on a standard laptop, while a naive simulation is typically beyond the largest supercomputers' reach.

Nevertheless, there are potential situations where our methods would not help. For example, if one qubit couples to many other (or most of the) qubits, then simulating it and its neighbors could be too difficult. A universal quantum information processor should probably not have this property, so it is not necessarily a realistic concern.

3.3.1 Single-qubit gates

We now look at concrete examples under our framework.

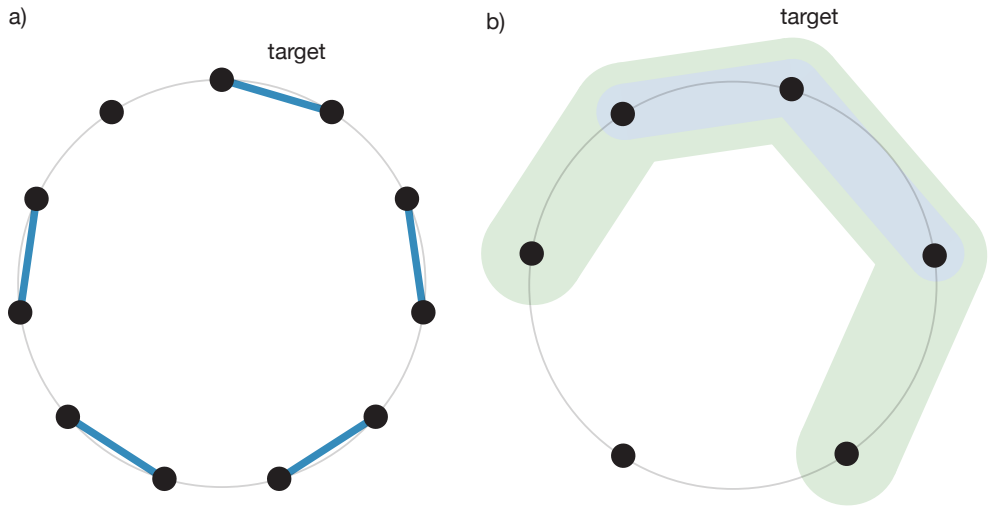

Figure 3.2: a) Connectivity graph of a ring of qubits, coupled through nearest-neighbor interactions (grey). We consider 5 simultaneous entangling gates. Blue lines specify the pairs of target qubits intentionally entangled by the gates.
b) Connectivity graph of the ring of qubits after abstracting away each pair of qubits targetted by the parallel entangling operations. We simulate a target pair of entangled qubits. Blue and green highlighting indicate the 'environment' subsystems that we include in our simulation when $d = 1$, and $d = 2$, respectively.

Ring of qubits

The first setup that we analyze is a ring of $N_s = 11$ qubits, where each qubit couples to its neighboring qubits. Fig. 3.1. a) schematically illustrates this setup. We consider the problem of implementing simultaneous single-qubit gates on every qubit and suppose that each qubit is a multilevel transmon and therefore incorporate a third energy level to model the effect of leakage ($d_s = 3$).

Without our approximation, the spatial complexity is $C = d_s^{2(11)} \approx 3.14 \times 10^{10}$. Performing a perturbative expansion with $d = 1$ and $d = 2$, we calculate improved spatial complexities, and the relative improvements are

$$C_1/C = N_s d_s^{2(1+2)}/C \approx 2.58 \times 10^{-7}\,, \tag{3.7}$$

$$C_2/C = N_s d_s^{2(1+2+2)}/C \approx 2.07 \times 10^{-5}\,. \tag{3.8}$$

IBM Melbourne

Next, we analyze the complexity of simulating single-qubit gates on IBM's Melbourne chip. We assume that IBM's public device information accurately reports all significant qubit-qubit coupling. Fig. 3.1. b) shows the device's connectivity. Taking $d_s = 3$ and $N_s = 14$ we compute a naive spatial complexity bound, $C = d_s^{2(14)} \approx 2.29 \times 10^{13}$. Performing a perturbative expansion with $d = 1$ and $d = 2$, we calculate improved spatial complexities, and the relative improvements are

$$C_1/C < N_s d_s^{2(1+3)}/C \approx 4.02 \times 10^{-9}\,, \tag{3.9}$$

$$C_2/C < N_s d_s^{2(1+3+4)}/C \approx 2.63 \times 10^{-5}\,. \tag{3.10}$$

3.3.2 Two-qubit gates

Ring of qubits

We analyze the same ring of 11 qubits described above, but now under simultaneous two-qubit gates. Fig 3.2. describes the setup and perturbative model. In this case, $N_s = 5$, $d_s = 9$. The bound on the complete simulation is the same with $C = 3.14 \times 10^{10}$. The relative improvements are

$$C_1/C < N_s (d_s)^{2(1+2)}/C \approx 1.02 \times 10^{-4} \tag{3.11}$$

$$C_2/C < N_s (d_s)^{2(1+2+2)}/C \approx 6.66 \times 10^{-1} \tag{3.12}$$

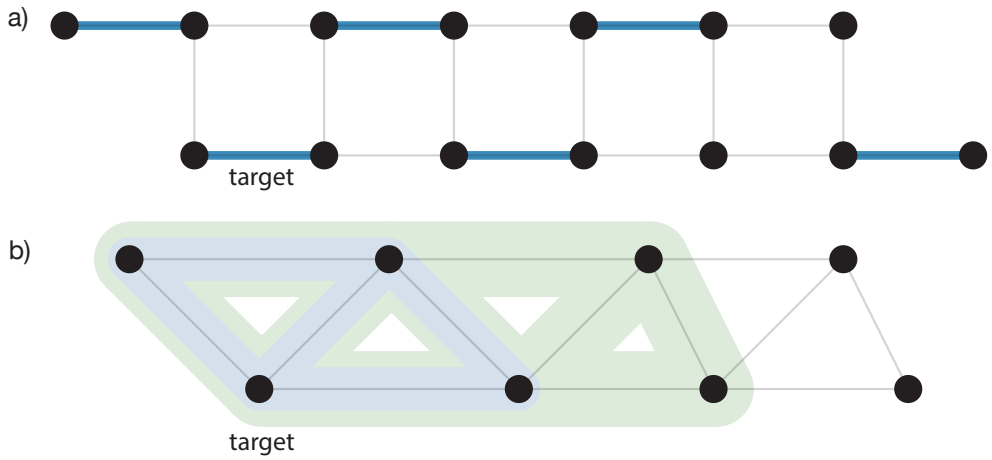

Figure 3.3: a) Connectivity graph of IBM Melbourne coupled through nearest-neighbor interactions (grey). Note that we assume no other qubit-qubit coupling affects the device. We consider 6 simultaneous entangling gates. Blue lines specify the pairs of target qubits intentionally entangled by the gates.
b) Connectivity graph of IBM Melbourne after abstracting away each pair of qubits targetted by the parallel entangling operations. We simulate a target pair of entangled qubits. Blue and green highlighting indicate the 'environment' subsystems that we include in our simulation when $d = 1$, and $d = 2$, respectively.

IBM Melbourne

Next, we model IBM's Melbourne chip under simultaneous two-qubit gates. Fig. 3.3 describes the setup and perturbative model. In this case, $N_\mathrm{s} = 8$, $d_\mathrm{s} = 9$. The bound on the complete simulation is the same with $C = 2.29 \times 10^{13}$.

$$C_1/C < N_\mathrm{s} d_\mathrm{s}^{2(1+3)}/C \approx 1.51 \times 10^{-5} \tag{3.13}$$
$$C_2/C = C/C = 1 \tag{3.14}$$

The first-order ($d = 1$) perturbative approximation simulates at most eight qubits and is, therefore, less costly to simulate than the entire device. Meanwhile, the second-order ($d = 2$) scheme necessarily simulates the entire device and affords no improvement.

3.4 Higher-order perturbative models

In the main text, we stated that simulating only directly neighboring subsystems usually yields a good approximation of a target system's behavior. Here we elaborate more on when this is true and focus on providing an intuitive explanation. It should also be possible to formally bound these effects with Lieb-Robinson bounds [24].

We can conceptually understand the limited-depth requirement from the observation that intermediate subsystems must mediate one subsystem's influence on a nonadjacent subsystem. In the high fidelity regime, which we are practically interested in, any significant such effect would typically couple the underlying subsystems and deviate significantly from the factorizable target operation.

For simplicity and clarity, we consider a time-independent Hamiltonian

$$H = H_\mathrm{s} \otimes I \otimes I + I \otimes H_\mathrm{n} \otimes I + I \otimes I \otimes H_\mathrm{nn} + \epsilon(H_\mathrm{s,n} \otimes I + I \otimes H_\mathrm{n,nn})\,, \tag{3.15}$$

where H_s, H_n, and H_nn are the internal Hamiltonians governing the target system, its nearest neighbors Hamiltonian, and the next-nearest neighbors, respectively. The interaction term $\epsilon H_\mathrm{s,n}$ couples the target with its neighbors and $\epsilon H_\mathrm{n,nn}$ couples the neighbors with the next-nearest neighbors. We define the unitaries

$$U_\mathrm{s,n}(t) = e^{-itH_\mathrm{s}} \otimes e^{-itH_\mathrm{n}} \otimes I\,, \tag{3.16}$$
$$U_\mathrm{n,nn}(t) = I \otimes e^{-itH_\mathrm{n}} \otimes e^{-itH_\mathrm{nn}}\,. \tag{3.17}$$

In the frame rotating according to the internal Hamiltonians of each of the three subsystems, the entire system evolves via the Hamiltonian

$$H_I(t) = \epsilon \left[U_{\mathrm{s,n}}^\dagger(t) H_{\mathrm{s,n}} U_{\mathrm{s,n}}(t) + U_{\mathrm{n,nn}}^\dagger(t) H_{\mathrm{n,nn}} U_{\mathrm{n,nn}}(t) \right] . \tag{3.18}$$

In the rotating frame, it is apparent that nearest-neighbor coupling is a first-order effect, while next-nearest neighbor coupling manifests as a second-order effect. Assuming that the magnitude of $H_{\mathrm{s,n}}$ and $H_{\mathrm{n,nn}}$ is less than or on the order of unity, t is small, and $\epsilon \ll 1$, we can neglect the effect of the next-nearest neighbors on the target system of interest. For example, if $\epsilon \lesssim 0.1$, then $\epsilon^2 \lesssim 0.01$. On real devices, where experiments can only correctly estimate the leading digit or two of a device parameter, the second-order contributions are nearly irrelevant.

There is a situation where nonadjacent subsystems may influence target dynamics significantly. Suppose a target qubit and nonadjacent qubit have similar frequencies. In that case, a far-detuned intermediate system could mediate strong dispersive coupling that entangles the two systems without itself appreciably coupling to either qubit. In all of our numerics, we observed no such effect, but it is a consideration that researchers using our techniques should keep in mind.

3.5 Single-qubit gate simulations

We review a typical implementation of single-qubit operations on transmons (see, e.g., [14, 23] for more details.) A local oscillator acts as a single tone microwave source outputting a constant signal $\cos(\omega' t)$ that is shaped by an arbitrary waveform generator via an in-phase and quadrature (IQ) mixer. A good description of a transmon qubit is an anharmonic oscillator driven by microwave pulses. In the lab frame, the relevant Hamiltonian is

$$H = \omega \hat{n} + \frac{\alpha}{2}(\hat{n} - 1)\hat{n} + \Omega(t)\cos(\omega' t - \gamma)(\hat{a} + \hat{a}^\dagger) , \tag{3.19}$$

where $\hat{a}$ is the annihilation operator of the oscillator, $\hat{n} = \hat{a}^\dagger \hat{a}$, α is the anharmonicity, γ is the drive phase, ω is the oscillator's resonant frequency, $\Omega(t)$ specifies the drive envelope, and we set $\hbar = 1$.

The lowest two energy levels form the qubit subspace. After making a rotating wave approximation (RWA) and moving into the rotating frame of the qubit, the Hamiltonian projected into the qubit subspace is [17]

$$H = \frac{1}{2}\Omega(t) e^{-i[\gamma + (\omega' - \omega)t]} \, |0\rangle\langle 1| + \text{h.c.} . \tag{3.20}$$

To see how the control induces single-qubit gates, consider a resonant pulse ($\omega = \omega'$). We also set the relative strength of the 0-1 transition to the 1-2 transition to 0), which corresponds to an ideal sufficiently long pulse. The control generates X and Y gates by modulating the coupling between the zero and one states, while the drive phase fixes the rotation axis in the XY-plane, and the pulse area sets the rotation angle. Rotations about the remaining Z-axis correspond to a change in the relative phase between the states. Rather than manipulating the transmon's state, it is equivalent to rotate the control with respect to the state, realizing a virtual-Z gate [18, 19, 27]. We accomplish this physically by adding a phase offset to all subsequent gates. A pulse with an area $\int dt\Omega(t) = \pi/2$ and a relative phase offset γ generates the unitary $V(\gamma) = Z_{-\gamma}X_{\pi/2}Z_{\gamma}$ with the notation $A_\theta = \exp(-i\theta A/2)$. Combining two of these phase-offset $\pi/2$ pulses and a final virtual-Z realizes any element of SU(2) [27].

Consider the problem of implementing an arbitrary element of $\mathrm{SU}(2)^{\otimes n}$ concurrently on an ensemble of qubits where their respective drive fields weakly interact with other qubits. The semiclassical Hamiltonian governing transmon k with local drive crosstalk is

$$\begin{aligned} H_k = \omega_k\hat{n} + \frac{\alpha_k}{2}(\hat{n}-1)\hat{n} \\ + \sum_j \beta_{jk}\Omega_j(t)\cos\big(\omega_j' t + \phi_j + \theta_{jk}\big)(\hat{a}+\hat{a}^\dagger)\,. \end{aligned} \tag{3.21}$$

The parameters β and θ characterize the crosstalk affecting the system. We focus on the case where each transmon has a local drive. The crosstalk parameters are $n \times n$ matrices, and we can set $\beta_{kk} = 1$ and $\theta_{kk} = 0$ without loss of generality by modifying Ω_k and ϕ_k. These constraints lead us to interpret β as the relative drive strength, and θ as the phase lag. Experimental data supports the model [48, 26, 49], and one can efficiently estimate the parameters with standard Rabi and Ramsey experiments.

For example, we simulate a system of $n = 100$ transmons that evolve under (3.21) and include the first three energy levels. The qubits are on a square grid with β_{jk} nonzero only for neighboring qubits. Qubits have a random frequency with $\omega_k/2\pi \sim \mathcal{N}(3\text{ GHz}, 500\text{ MHz})$, where $\mathcal{N}(\mu,\sigma)$ is a normal distribution centered at μ with a standard deviation of σ. All qubits have random anharmonicities: $\alpha_k/2\pi \sim \mathcal{N}(-330\text{ MHz}, 50\text{ MHz})$. In each iteration of the experiment, the target gate is chosen randomly from $\mathrm{SU}(2)^{\otimes n}$. The crosstalk phase lag parameter θ_{jk} are sampled randomly from the interval $[0, 2\pi)$, and we draw β_{jk} from a normal distribution centered at zero. There are two discrete periods of successive evolution, each taking time $t_{\pi/2}$. It is necessary to pick pulse shapes. On the one hand, we want pulses that yield error rates near the decoherence limit for short gate times. On the other hand, there are experimental realities, such as power-bandwidth constraints

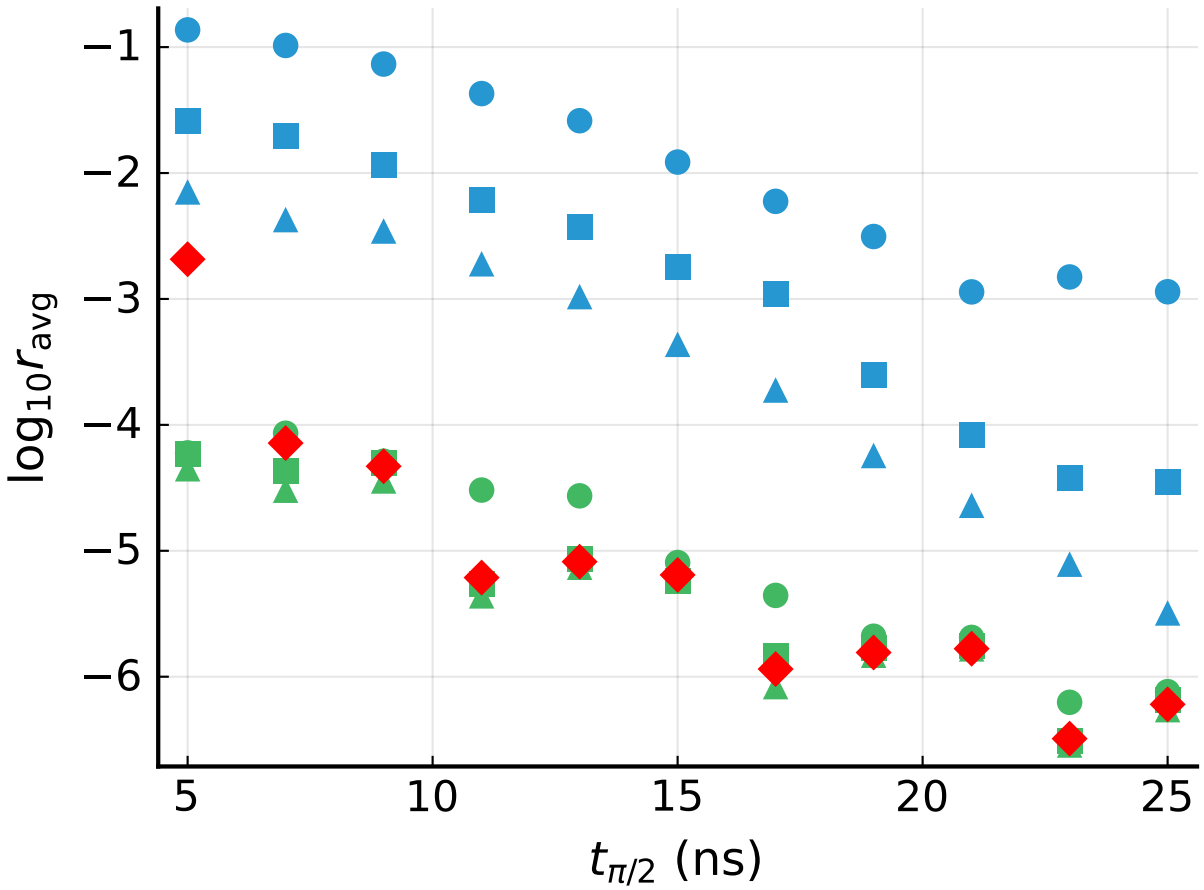

Figure 3.4: Plots illustrating a significant improvement in the average single-qubit process infidelity as a function of the time for a $\pi/2$ gate (the total simulation time is $2t_{\pi/2}$). There are 100 qubits in a square 2D array, and each qubit implements a random element of SU(2) via two $\pi/2$ pulses with intermediate phase offsets. The red diamonds denote the infidelity of the qubits with half-derivative DRAG corrections and no crosstalk. The blue points are infidelities obtained under the same control with crosstalk. Circles, squares, and triangles denote several relative crosstalk strengths β_{ij}, that are sampled from a normal distribution $\mathcal{N}(0, \sigma)$ with standard deviations $\sigma = 0.05$, $\sigma = 0.1$, and $\sigma = 0.25$ respectively. The green markers have identical crosstalk as their blue counterpart, but with optimized control parameters.

Crosstalk std β_{jk}	Original r_{avg}	Opt., $t_{\pi/2} = 2$ ns r_{avg}	Opt., $t_{\pi/2} = 5$ ns r_{avg}
0.05	6.02e-4	1.00e-4	1.86e-4
0.1	7.13e-4	1.03e-4	1.91e-4
0.25	2.13e-3	1.15e-4	1.84e-4
0.5	1.77e-2	1.07e-4	1.81e-4

Table 3.1: Data highlighting a dramatic reduction in the average single-qubit process infidelity for a simulation with realistic decoherence on a square array of 100 qubits and various levels of crosstalk (std β_{jk}). We model the same system considered in Fig. 3.4, but with $T_1 \sim \mathcal{N}(40\ \mu\text{s}, 5\ \mu\text{s})$ for each qubit, and $T_2 = 3T_1/2$. Naturally, there is an optimal gate time that minimizes the combined incoherent (increasing) and coherent (approximately decreasing) effects. We optimize the controls for $t_{\pi/2} = 1, 2, \ldots, 50$ ns. The 'Original' column corresponds to the optimal $t_{\pi/2}$ without control tuneup. For all values of std $\beta_{j,k}$, r_{avg} is minimized at $t_{\pi/2} = 2$ ns. On contemporary experimental devices, $t_{\pi/2} = 2$ ns may exceed accesible bandwidths, so we also report r_{avg} for $t_{\pi/2} = 5$ ns.

and the degree of calibration needed to implement complicated pulses accurately. Balancing these constraints, we pick Gaussian pulses with std $\Omega^{(x)} = t_{\pi/2}/4$, and half-derivative DRAG corrections $\Omega^{(y)} = -\dot{\Omega}^{(x)}/2\alpha$ [28, 13, 6].

Fig. 3.4 shows the average single-qubit process infidelity $r_{\text{avg}} = 1 - \langle \Phi_k \rangle$ as a function of $t_{\pi/2}$. Green diamonds denote the raw infidelity for a crosstalk-free system ($\beta_{jk} = \delta_{jk}$, where δ_{jk} is the Kronecker delta). The blue markers are infidelities obtained using the crosstalk-free control scheme but with various strengths of drive crosstalk. The red markers are infidelities obtained with optimized control and the same drive crosstalk as the blue markers. We optimize control pulses with the method of Ref. [25]. Applying the protocol requires the selection of appropriate optimization parameters. Sticking to our simple control ansatz, we tune the overall magnitude of the resonant $\Omega^{(x)}$ quadrature, off-resonant $\Omega^{(y)}$ quadrature and the carrier signal phase ϕ, for a total of $7n$ parameters. We observe approximately two orders of magnitude improvement in the infidelity with our crosstalk minimization technique.

In real experimental devices, decoherence significantly reduces the average error rates. Moreover, decoherence errors grow with time, whereas control errors typically decrease. These contrasting effects imply that there is an optimal gate time that minimizes their combined errors. We repeat the simulation implementing $\text{SU}(2)^{\otimes n}$ with decoherence added to the model. Table 3.1 presents data showing the potential benefit of our methods.

3.6 Two-qubit gate simulations

We continue our simulations using the ideal system of fixed-frequency transmons and the parameter values specified above. Our aim is to implement parallel CR gates [36, 8, 41, 26, 3], which are equivalent to CNOTs up to single-qubit operations. Constant capacitive coupling provides a mechanism for implementing entangling operations. Assuming equal coupling between all neighboring qubits in the system, the corresponding interaction Hamiltonian is

$$H_{\text{int}} = \sum_{\langle j,k\rangle=1} J_{jk}(a_j a_k^\dagger + a_j^\dagger a_k)\,, \tag{3.22}$$

where $\langle j,k\rangle = 1$ denotes a sum over all adjacent qubit pairs. The entire system evolves under $H_{\text{int}} + \sum_k H_k$.

The basic idea of the CR effect is that if we define the qubits in a dressed basis, local microwave drive fields drive both single and two-qubit gates. For two ideal coupled qubits, in the dressed basis, a drive applied to qubit 1 at the frequency of qubit 2 yields the effective Hamiltonian [8]

$$H_d = \Omega(t)\left(X_1 - \frac{J}{\Delta} Z_1 X_2\right)\,, \tag{3.23}$$

where $\Delta = \omega_1 - \omega_2$ is the difference of qubit frequencies and we made an RWA. The subscripts on the Pauli gates specify the affected qubit. Although we can decouple the direct qubit coupling, higher-levels of the transmon lead to additional terms in the effective Hamiltonian [26]. We can use the Z_1X_2 term to generate a maximally entangling gate.

Again, we simulate a system of $n = 100$ transmons on a grid and include the first three energy levels of each. We group adjacent qubits in pairs and try to implement 50 simultaneous maximally entangling gates using the CR effect. Our qubits have 8 distinct frequencies $3.0, 3.1, \ldots, 3.7$ GHz to ensure each CR pair is addressable. We set the frequencies so that no two neighbors of one qubit have the same frequency. Unlike in the first example, if certain qubit frequencies overlap the pair becomes extremely hard to manipulate independently. The target CNOT equivalent is determined using Cartan's KAK decomposition [46] and is invariant to local operations. The qubit coupling strength matrix J is symmetric and random with $J_{jk}/2\pi \sim \mathcal{N}(3.8\text{ MHz}, 1\text{ MHz})$. We realize qubit control with the same drives as above but with variable drive detuning and phase offset. We independently parameterize the resonant $\Omega^{(x)}$ and off-resonant $\Omega^{(y)}$ control envelopes with the first three Hanning window functions

$$\Omega_H(t) = \sum_{k=1}^{3} c_k \left[1 - \cos\left(\frac{2\pi k t}{t_{\text{CR}}}\right)\right]\,. \tag{3.24}$$

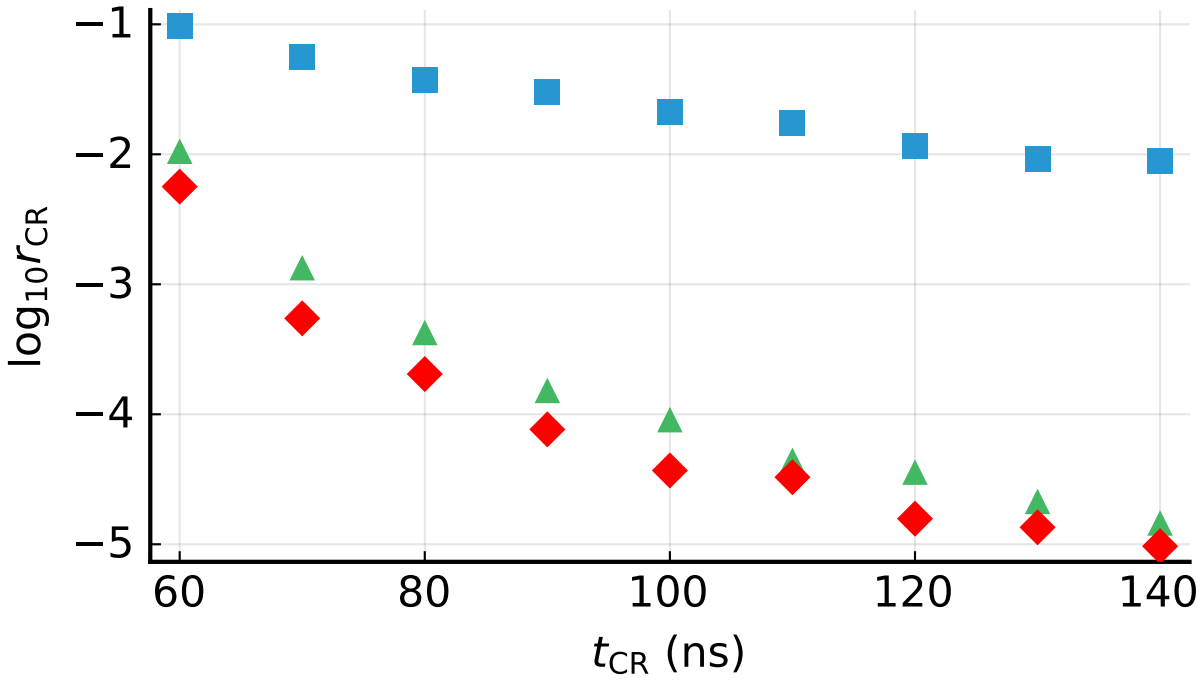

Figure 3.5: Plots showing a massive improvement in the average two-qubit process infidelity for a square array of transmons implementing 50 simultaneous maximally entangling gates via CR interactions. Each two-qubit pair approximates a CNOT-gate up to local operations. The red diamonds correspond to CR gate infidelities obtained without drive crosstalk or subsystem coupling. The blue squares are infidelities obtained with the same controls as the green diamonds but with drive crosstalk as in the single-qubit example ($\sigma = 0.1$) and constant nonlocal coupling between all adjacent qubits. The green triangles have the same crosstalk as the above model, but with optimized control parameters.

There are a total of $8n$ parameters that determine the n drive fields.

Fig. 3.5 shows the average two-qubit process infidelity of each entangling gate as a function of the gate duration t_{CR}. We compute all points with optimized pulse parameters [25] but under different system models. Red diamonds denote the infidelity obtained using a drive-crosstalk-free model and no undesirable J coupling. The blue squares are infidelities calculated using the crosstalk-free optimal control but with added drive crosstalk ($\sigma = 0.1$) and nonlocal coupling. The green triangles are infidelities obtained with controls tuned up under the crosstalk model. We approximate the nonlocal crosstalk effects with $d = 1$. The deviation caused by including additional neighbors is unresolvable on the plot.

Chapter 4

Crosstalk Experiments with Superconducting Qubits

This chapter focuses on the ongoing research that we are conducting with fixed-frequency superconducting transmon qubit-based quantum computers [21, 39]. We aim to measure crosstalk affecting experimental devices, with the ultimate objective of neutralizing it. In the first part, we developed a theory for modeling crosstalk and showed how optimal control could counteract it. We need to address experimental questions so that we can employ the theory.

In particular, we must consider the following questions:

1. How do we estimate the crosstalk parameters corresponding to a device?
2. How can we validate that the model accurately describes the device?

We approach these questions by first analyzing the Hamiltonian for simultaneously driven uncoupled fixed-frequency transmons. With careful analysis, we uncover a generalization of the Rabi experiment, the double field method, which is better suited for estimating crosstalk. With this new tool, we learn the crosstalk parameters corresponding to a device with greater precision and fewer measurements than could be achieved with standard Rabi experiments. We determine the theoretical Hamiltonian corresponding to three driving fields to validate the model. We compare experimental results and theoretical predictions using the fit data from double field experiments. We observe close agreement between theory and experiment that validates the correctness of our approach.

4.1 Hamiltonian analysis

In the following subchapter, we develop the basic theoretical tools needed to construct the double field method to estimate crosstalk parameters.

4.1.1 Deriving the rotating Hamiltonian

We begin by analyzing the Hamiltonian corresponding to a transmon subject to an arbitrary collection of semiclassical driving fields [14, 23]. The semiclassical Hamiltonian corresponding to transmon j with local drive crosstalk is

$$H^{(j)} = H_0^{(j)} + \sum_k H_1^{(jk)} \,, \tag{4.1}$$

where the transmon's Hamiltonian and drive Hamitonians are

$$H_0^{(j)} = \omega_j \hat{n} + \frac{\alpha_j}{2} \hat{n}(\hat{n} - 1) \,, \tag{4.2}$$

$$H_1^{(jk)} = \beta_{jk} \Omega_k (t - \tau_{jk}) \cos(\omega_k'(t - \tau_{jk}) - \theta_{jk} - \phi_k)(\hat{a} + \hat{a}^\dagger) \,. \tag{4.3}$$

The Hamiltonian $H_0^{(j)}$ describes transmon k (a Duffing oscillator), and the variables ω_j and α_j denote the transmon's frequency and anharmonicity, respectively. The driving Hamiltonian $H_1^{(jk)}$ is the effect induced by driving field k on qubit j. The variable β specifies the relative drive strength ($\beta_{jk} \geq 0$), θ is the relative drive phase, and ϕ_k is a phase shift set by software. By convention, when there is a one-to-one correspondence between drives and transmons, we should have $\beta_{jj} \approx 1$ and $\theta_{jj} = 0$. This imposes the notion that we intend drive j to control transmon j. We are free to fix $\theta_{jj} = 0$ since this specifies the measurement basis. However, since perfect calibration is impossible, β_{jj} is not necessarily exactly equal to 1. The function Ω_k is the pulse envelope, and ω_k' is the drive frequency. Since the time for a driving field to reach a qubit is not always constant, i.e., the signal path lengths differ, we introduce $\tau_{jk} = \tau_j - \tau_k$, which captures a relative time delay. The matrix τ is always skew-symmetric. Moreover, any column or row of τ completely specifies the rest of the matrix.

In the transmon's rotating frame, the drive Hamiltonians are

$$H_I^{(jk)} = U_0^{(j)\dagger} H_1^{(jk)} U_0^{(j)} = \sum_{m=0}^{\infty} a_{jk;m} \left|m\right\rangle\!\left\langle m+1\right| + a_{jk;m}^* \left|m+1\right\rangle\!\left\langle m\right| \,, \tag{4.4}$$

where

$$a_{jk;m} = \sqrt{m+1}\beta_{jk}\Omega_k(t_{jk})e^{-it\omega_j}e^{-it\alpha_j m}\cos(t_{jk}\omega'_k - \theta_{jk} - \phi_k)\,, \tag{4.5}$$

and we defined $t_{jk} = t - \tau_{jk}$. We assume $|\omega_j - \omega'_k| \ll \omega_j + \omega'_k$ and make a rotating wave approximation to get

$$a_{jk;m} = \frac{\sqrt{m+1}}{2}\beta_{jk}\Omega_k(t_{jk})\exp\Big[-i(t_{jk}\Delta\omega_{jk} + t\alpha_j m + \tilde{\theta}_{jk} + \phi_k)\Big]\,, \tag{4.6}$$

where we defined $\Delta\omega_{jk} = \omega_j - \omega'_k$ and the dressed phase shift $\tilde{\theta}_{jk} = \theta_{jk} + \omega_j\tau_{jk}$. We will see throughout our analysis that formulating the problem in terms of the dressed phase shift $\tilde{\theta}_{jk}$ is more natural. In any setting where θ_{jk} is relevant, it appears alongside $\omega_j\tau_{jk}$.

In our case, we only use the expressions for the two lowest transitions

$$a_{jk;0} = \frac{1}{2}\beta_{jk}\Omega_k(t_{jk})\exp\Big[-i(t_{jk}\Delta\omega_{jk} + \tilde{\theta}_{jk} + \phi_k)\Big]\,, \tag{4.7}$$

$$a_{jk;1} = \frac{1}{\sqrt{2}}\beta_{jk}\Omega_k(t_{jk})\exp\Big[-i(t_{jk}\Delta\omega_{jk} + \tilde{\theta}_{jk} + \phi_k + t\alpha_j)\Big]\,. \tag{4.8}$$

4.1.2 Cross-Rabi experiments

In a conventional Rabi experiment, we apply a single driving field to a qubit and measure. We repeat the procedure for different pulse areas by varying the duration or peak amplitude and fit a model to the resulting data. A cross-Rabi experiment is very similar. Instead of applying a driving field with which we intend to control a qubit a, we apply another field, b, resonantly. Such a setup allows us to measure the crosstalk strength and related parameters. Let us suppose that the pulse is sufficiently long and smooth so that we can neglect the effect of the third energy level. The Hamiltonian is then specified by

$$a_{ab;0} = \frac{1}{2}\beta_{ab}\Omega_k(t_{ab})\exp\Big(-i(\tilde{\theta}_{ab} + \phi_b)\Big)\,. \tag{4.9}$$

We can analytically solve this problem. The expectation values for the Pauli observables and initial state $|0\rangle$ are

$$\langle X\rangle = \sin(\beta_{ab}\langle\Omega_b\rangle t)\sin\Big(\tilde{\theta}_{ab} + \phi_b\Big)\,, \tag{4.10}$$

$$\langle Y\rangle = -\sin(\beta_{ab}\langle\Omega_b\rangle t)\cos\Big(\tilde{\theta}_{ab} + \phi_b\Big)\,, \tag{4.11}$$

$$\langle Z\rangle = \cos(\beta_{ab}\langle\Omega_b\rangle t)\,. \tag{4.12}$$

While the form of $\langle Z \rangle$ tells us that we can estimate β in a SPAM robust manner (provided it is not too small), measuring $\tilde{\theta}$ requires additional gates and is therefore not robust to SPAM. Moreover, this experiment does not depend on τ and will not yield any estimate.

4.1.3 Simultaneous Rabi experiments

A simultaneous Rabi experiment is a powerful generalization of the standard Rabi method. It makes it possible to measure very small β_{ab}, measure $\tilde{\theta}$ with only standard basis measurements and measure τ.

The most basic generalization is the double field Rabi experiment, where we apply two concurrent pulses. A *reference* Rabi pulse, which we generate by driving some qubit a of interest resonantly with drive b (similar results hold when the drive is not a), and a *perturbing* pulse we introduce by applying drive b. The perturbing pulse changes measurement outcomes in such a way that parameter confidence intervals are typically better. Moreover, variables that were previously unmeasurable matter. Because there are two concurrent pulses, we break the symmetry of the Rabi method.

In the following, we analyze the dynamics of a double field experiment. Because we are interested in estimating parameters and not maximizing gate fidelity, it is natural to consider smooth pulses and somewhat long durations. In this regime, the effect of the third energy level is small (provided we do not apply a pulse near the frequency of the 1-2 transition). The relevant Hamiltonian is entirely specified by

$$a_{aa;0} = \frac{1}{2}\beta_{aa}\Omega_j(t_{aa})\exp(-i\phi_a)\,, \tag{4.13}$$

$$a_{ab;0} = \frac{1}{2}\beta_{ab}\Omega_k(t_{ab})\exp\Big[-i(t_{ab}\Delta\omega_{ab} + \tilde{\theta}_{ab} + \phi_b)\Big]\,. \tag{4.14}$$

Resonant perturbing pulse

We first examine the simplest case when both drives are applied resonantly, $\Delta\omega_{ab} = 0$. It makes sense to split $\bar{H}^{(1)}$ into two parts: the zeroth-order term $\bar{H}^{(0)}$, which is the exact

solution when $a_{ab0} = 0$ and a first-order correction $\bar{H}^{(1)}$. It is straightforward to show that

$$\bar{H}^{(0)} = \frac{1}{2}\beta_{aa}\langle\Omega_a\rangle\cos\phi_a X + \frac{1}{2}\beta_{aa}\langle\Omega_a\rangle\sin\phi_a Y\,, \tag{4.15}$$

$$\bar{H}^{(1)} = \frac{1}{2}\beta_{ab}\langle\Omega_b\rangle\cos\Big(\tilde{\theta}_{ab}+\phi_b\Big)X + \frac{1}{2}\beta_{ab}\langle\Omega_b\rangle\sin\Big(\tilde{\theta}_{ab}+\phi_b\Big)Y\,, \tag{4.16}$$

$$\bar{H}^{(2)} = \frac{1}{4t}\beta_{aa}\beta_{ab}\sin\Big(\phi_a-\phi_b-\tilde{\theta}_{ab}\Big)\int_0^t dt_1\int_0^{t_1} dt_2\left[\Omega_a(t_1)\Omega_b(t_2-\tau_{ab})-\Omega_a(t_2)\Omega_b(t_1-\tau_{ab})\right]Z\,. \tag{4.17}$$

where X, Y, and Z are the standard Pauli operators.

We suppose that all driving envelopes are identical, $\Omega = \Omega_a = \Omega_b$. Such an assumption is reasonable and corresponds to an experiment where we apply, for example, Gaussian or flat-top sine envelopes with the same properties to each driving line. Furthermore, we consider an intial state $|0\rangle$.

$\tau_{ab} = 0$

It is clear that when $\tau_{ab} = 0$, $H^{(2)}$ and all higher order terms in the Magnus expansion vanish. Moreover, $H^{(1)}$ does not depend on τ_{ab}. Thus we can interpret $H^{(2)}$ as the leading order correction accounting for the effect of relative delays between pulses. If $\tau_{ab} = 0$, we can analytically solve the Schrödinger equation exactly and obtain

$$\langle X\rangle = \langle\Omega\rangle\left[\beta_{aa}\sin\phi_a + \beta_{ab}\sin\Big(\tilde{\theta}_{ab}+\phi_b\Big)\right]\frac{\sin\eta_{ab}t}{\eta_{ab}}\,, \tag{4.18}$$

$$\langle Y\rangle = \langle\Omega\rangle\left[\beta_{aa}\cos\phi_a + \beta_{ab}\cos\Big(\tilde{\theta}_{ab}+\phi_b\Big)\right]\frac{\sin\eta_{ab}t}{\eta_{ab}}\,, \tag{4.19}$$

$$\langle Z\rangle = \cos\eta_{ab}t\,, \tag{4.20}$$

where

$$\eta_{ab} = \langle\Omega\rangle\sqrt{\beta_{aa}^2+\beta_{ab}^2+2\beta_{aa}\beta_{ab}\cos\Big(\Delta\phi_{ab}-\tilde{\theta}_{ab}\Big)}\,, \tag{4.21}$$

and we defined $\Delta\phi_{ab} = \phi_a - \phi_b$. The quantity under the square root changes the period of a Rabi oscillation. If $\beta_{ab} \ll 1$, then

$$\eta_{ab} \approx \langle\Omega\rangle\left[\beta_{aa}+\beta_{ab}\cos\Big(\Delta\phi_{ab}-\tilde{\theta}_{ab}\Big)\right]\,. \tag{4.22}$$

When $\tau_{ab} = 0$, observe that varying the pulse area (total duration) and phase ϕ_a or ϕ_b and measuring Z is sufficient to fit β_{aa}, β_{ab}, and $\tilde{\theta}_{ab}$.

$\tau_{jk} \neq 0$

What should we do when τ_{ab} is nonzero? Let us consider the case where $\beta_{ab} \ll 1$. We can rewrite our Magnus expansion for the average Hamiltonian with terms up to second order in the form

$$H = \frac{\alpha}{2}\left(\cos\theta X + \sin\theta Y + \epsilon Z\right) , \tag{4.23}$$

where ϵ is close to zero. We can easily determine the expectations values for the Pauli observables under the evolution specified $U = \exp(-iH)$. Consider a Taylor series expansion of the solution about $\epsilon = 0$,

$$\langle X\rangle = \sin\alpha\sin\theta + (1-\cos\alpha)\cos\theta\epsilon + \frac{1}{2}(\alpha\cos\alpha - \sin\alpha)\sin\theta\epsilon^2 + \mathcal{O}\big(\epsilon^3\big) , \tag{4.24}$$

$$\langle Y\rangle = -\sin\alpha\cos\theta + (1-\cos\alpha)\sin\theta\epsilon - \frac{1}{2}(\alpha\cos\alpha - \sin\alpha)\cos\theta\epsilon^2 + \mathcal{O}\big(\epsilon^3\big) , \tag{4.25}$$

$$\langle Z\rangle = \cos\alpha + \left(1 - \cos\alpha - \frac{1}{2}\alpha\sin\alpha\right)\epsilon^2 + \mathcal{O}\big(\epsilon^4\big) . \tag{4.26}$$

Observe that there is no first-order term in the series for $\langle Z\rangle$. Thus measuring Z is not a practical experiment to learn τ_{ab} when β_{ab} is small. The optimal measurement that is maximally sensitive to ϵ and thus τ_{ab} lies in the XY plane. Suppose we set $\phi_a = 0$, then θ is close to 0 and $\langle X\rangle$ is a good measurement choice for learning τ_{ab}. Measuring Z is, of course, the easiest measurement to perform.

4.2 Double field experiments

In the previous subchapter, we developed the essential tools needed to construct a new approach to measuring crosstalk. Let us consider two qubits, Q_a and Q_b, and their separate driving fields, D_a and D_b. The crosstalk effect of D_b on Q_a is small when $\beta_{ab} \ll 1$. We assume Q_a starts in the ground state, apply drives D_a and D_b at frequency ω_a for some time t, and measure Q_a. Let us focus on the result of a $\langle Z\rangle$ measurement. If we neglect the third energy level of Q_a and truncate a Magnus expansion at second-order

$$\langle Z\rangle = \cos\langle\Omega\rangle\,\eta_{ab}t\,, \tag{4.27}$$

where

$$\eta_{ab} = \sqrt{\beta_{aa}^2 + \beta_{ab}^2 + 2\beta_{aa}\beta_{ab}\cos\Big(\Delta\phi_{ab} - \tilde{\theta}_{ab}\Big)} \tag{4.28}$$

$$\approx \beta_{aa} + \beta_{ab}\cos\Big(\Delta\phi_{ab} - \tilde{\theta}_{ab}\Big) . \tag{4.29}$$

Q_a	Q_b	β_{aa}	β_{ab}	θ_{ab}	χ^2/v
0	1	1.00 ± 3.63e-4	1.43e-1 ± 5.25e-4	6.22 ± 3.47e-3	1.32
0	2	1.00 ± 3.37e-4	9.17e-3 ± 4.66e-4	6.18 ± 5.30e-2	1.34
1	0	1.00 ± 3.80e-4	9.65e-2 ± 5.57e-4	5.05 ± 5.34e-3	0.72
1	2	1.00 ± 4.31e-4	3.54e-1 ± 5.97e-4	4.03 ± 1.65e-3	1.04
1	3	1.00 ± 4.49e-4	1.66e-1 ± 6.83e-4	3.63 ± 3.31e-3	1.43
2	0	1.00 ± 3.28e-4	3.03e-2 ± 4.72e-4	1.26 ± 1.51e-2	1.26
2	1	1.00 ± 3.33e-4	8.91e-2 ± 4.80e-4	1.29 ± 5.16e-3	0.82
2	3	1.00 ± 4.03e-4	2.55e-1 ± 6.19e-4	4.07 ± 2.00e-3	1.99
3	1	1.00 ± 3.56e-4	2.21e-2 ± 5.01e-4	2.57 ± 2.29e-2	1.08
3	2	1.00 ± 3.68e-4	9.25e-2 ± 5.19e-4	6.22 ± 5.61e-3	0.85

Table 4.1: Data obtained from double field experiments for various pairs of qubits on a chip. The values of χ^2/v are very close to one and indicate an excellent goodness of fit. The column Q_a denotes the measured qubit and Q_b denotes the perturbing drive. The columns β_{aa}, β_{ab}, and θ_{ab} are fit parameters along with error estimates obtained during least-squares fitting. The value of β_{aa} are all close to one and indicate that the drive field corresponding to a target qubit is correctly tuned. The crosstalk parameters β_{ab} and θ_{ab} span a wide range of values. We observe crosstalk strengths from 1% up to 35% on this device.

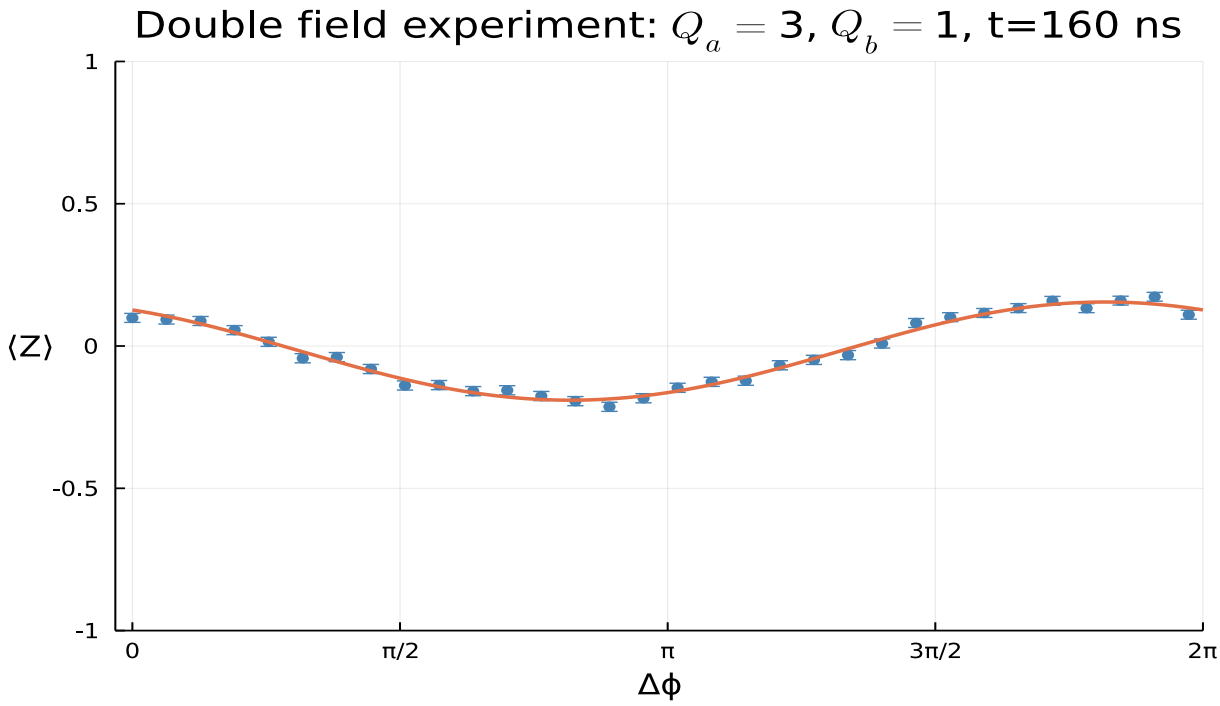

Figure 4.1: Plots showing the fit of our theoretical model and experimentally obtained data for a double field experiment. The blue error bars on experimental data designate the standard deviation of the error. The fit is close to theory across all values of $\Delta\phi$. Quantitatively, we have a reduced chi-square value of 1.08, meaning the fit is essentially perfect. We extract the values of crosstalk parameters from the fit: $\beta_{31} \approx$ 2.21e-2, $\theta_{31} \approx 2.57$, so the crosstalk is quite weak with a relative strength of 2%.

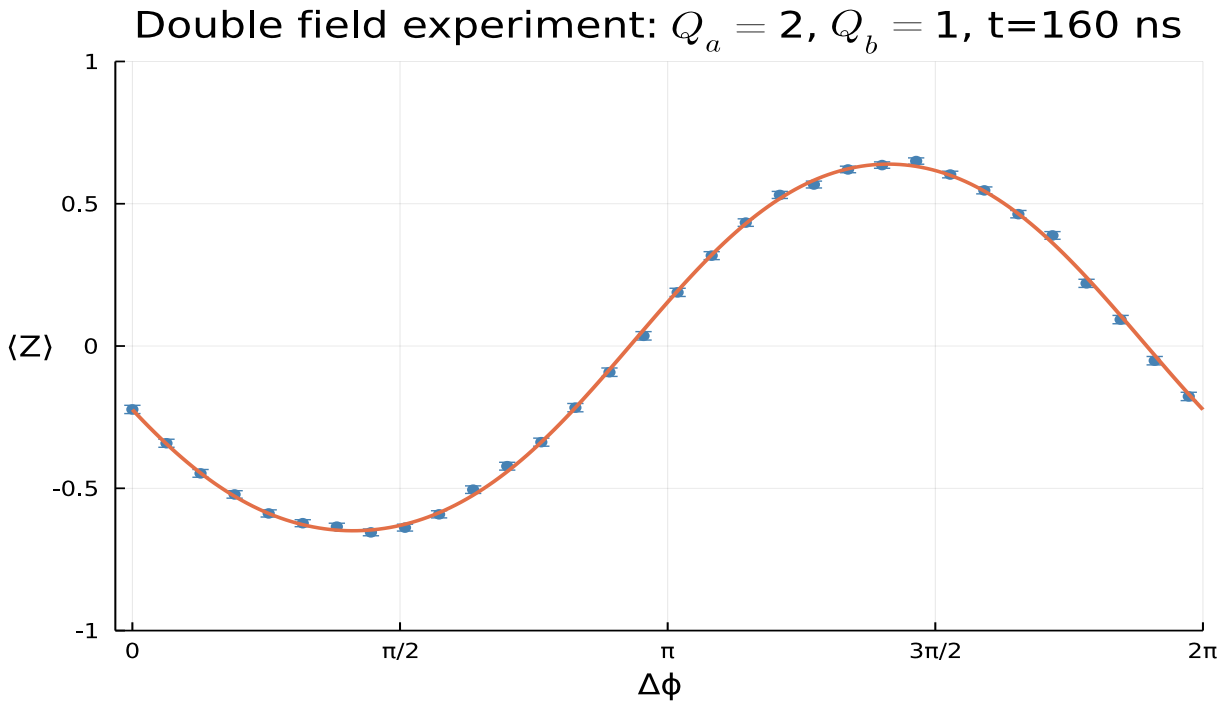

Figure 4.2: Plots showing the fit of our theoretical model and experimentally obtained data for a double field experiment. The blue error bars on experimental data designate the standard deviation of the error. The fit is close to theory across all values of $\Delta\phi$. Quantitatively, we have a reduced chi-square value of 0.82, meaning the fit is essentially perfect. We extract the values of crosstalk parameters from the fit: $\beta_{21} \approx$ 8.91e-2, $\theta_{21} \approx 1.29$, so the crosstalk is quite strong.

We aim to use this set of equations to devise an experimental procedure to estimate the crosstalk parameters β_{ab} and θ_{ab}. We want our procedure to be maximally sensitive to the perturbing driving field D_b. Therefore, we pick t and $\langle\Omega\rangle$ so that $\langle\Omega\rangle\,\beta_{aa}t = -\pi/2 \mod 2\pi$. Making this choice ensures that $\eta = 0$ if there is no crosstalk. We then obtain

$$\langle Z\rangle \approx \sin\Big(\beta_{ab}\,\langle\Omega\rangle\,t\cos\Big(\Delta\phi_{ab} - \tilde{\theta}_{ab}\Big)\Big)\,. \tag{4.30}$$

By examining the structure of this new equation, we are naturally led to an experimental procedure.

1. Set $\langle\Omega\rangle\,t = \pi/2 + n\pi$ and pick n so that it is as large as possible without decoherence having a substantial effect.
2. Select a range of $\Delta\phi_{ab}$ values, a natural choice is a uniform set of points on the interval $[0, 2\pi)$.
3. Perform the double field experiment with all pairs of qubits that we suspect have crosstalk between them.
4. Fit the theoretical model to experimental data using the method of least squares.
5. Compute reduced chi-squared values to assess the model's goodness of fit.

With the procedure established, we are now ready to conduct experiments. For these experiments, we consider four qubits from an eight-qubit chip. Three qubits, 0, 1, and 2, are adjacent and directly connected in a line by two resonators. A third qubit, 3, is farther away from 0, 1, and 2.

We performed the procedure between all possible pairs of these four qubits, and Table 4.1 contains the resulting data. Note that qubits 0 and 3 share a control line that is toggled between the two qubits with a switch. Consequently, we cannot perform the experiment with the pair (0, 3). However, performing it makes no sense since they cannot be simultaneously driven.

We observe excellent goodness of fit, which provides some validation that our model is reasonable. We chose $t = 160$ ns, but longer times are possible. By factoring in decoherence, it is likely possible to obtain accurate estimates of crosstalk on the order of 1e-5 based on T_1 and T_2 times. Nevertheless, crosstalk on our device is on the order of 1e-2, and such experiments provide little more value.

Fig. 4.1 and Fig. 4.2 plot data corresponding to two different double field experiments. We chose these figures because they showcase how our method looks in the case of weak and strong crosstalk, respectively. Note that the overall amplitude of the weak case could be made larger as it is proportional to the value of t.

4.3 Validating the model

The double field method shows promise for accurately learning the crosstalk parameters that characterize a superconducting chip. However, how can we ensure the model is correct for the entire chip with only pairwise experiments? We suppose crosstalk is linear, but verifying that the assumption is correct would be beneficial.

The most straightforward validation procedure is to measure what happens when we apply three concurrent driving fields. We can then compare the expected dynamics by combining the fit data for two double field experiments and the experimental data. Concretely, there is a natural generalization of (4.27) for three driving fields,

$$\langle Z \rangle = \cos(\langle \Omega \rangle \, \eta_{abc} t) \,, \tag{4.31}$$

where

$$\eta_{abc} = \sqrt{1 + \beta_{ab}^2 + \beta_{ac}^2 + 2\beta_{ab}\cos\Big(\Delta\phi_{ab} - \tilde{\theta}_{ab}\Big) + 2\beta_{ac}\cos\Big(\Delta\phi_{ac} - \tilde{\theta}_{ac}\Big) + 2\beta_{ab}\beta_{ac}\cos\Big(\Delta\phi_{bc} - \tilde{\theta}_{bc}\Big)} \,, \tag{4.32}$$

$$\tilde{\theta}_{bc} = \tilde{\theta}_{ab} - \tilde{\theta}_{ac} \,. \tag{4.33}$$

This equation gives us all that we need. Let us attenuate β_{ac} by a factor of $\sqrt{2}$. The attenuation is arbitrary but will provide more reassurance that linearity is correct. Fig. 4.3 and Fig. 4.4 plot data corresponding to two different validation experiments. We observe that our predictions and the experimental measurements are very close. Thus, our model appears to accurately describe crosstalk in a setting where no fitting is occurring.

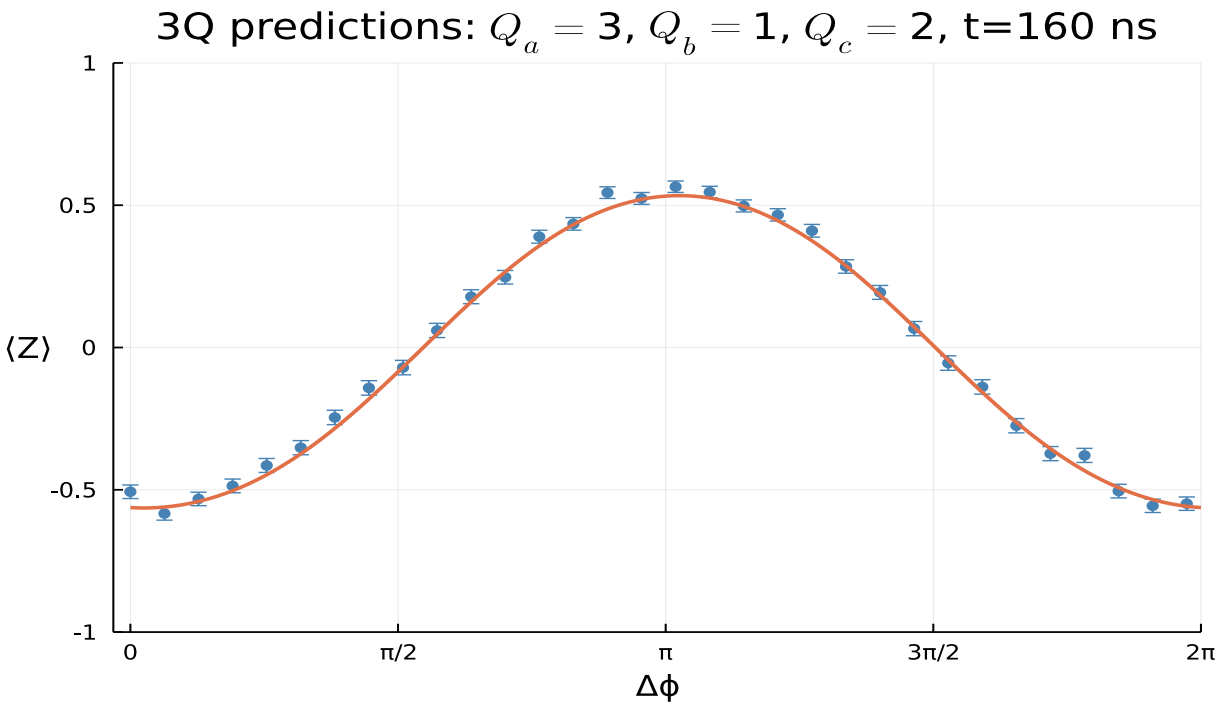

Figure 4.3: Plots comparing our theory for when three driving fields are simultaneously at the resonant frequency of $Q_a = 3$ versus experimentally obtained data. We attenuate the strength of the field labeled Q_c by a factor of $\sqrt{2}$. The blue error bars on experimental data designate the standard deviation of the error. We compute the value of the reduced chi-square as 1.15, which indicates that the theory and the observed data closely match. Note that the dynamics deviate significantly from a double field experiment with $Q_a = 3$, $Q_b = 1$ and an experiment with $Q_a = 3$, $Q_b = 2$. This observation provides strong evidence that our crosstalk model is appropriate and accurately describes the device.

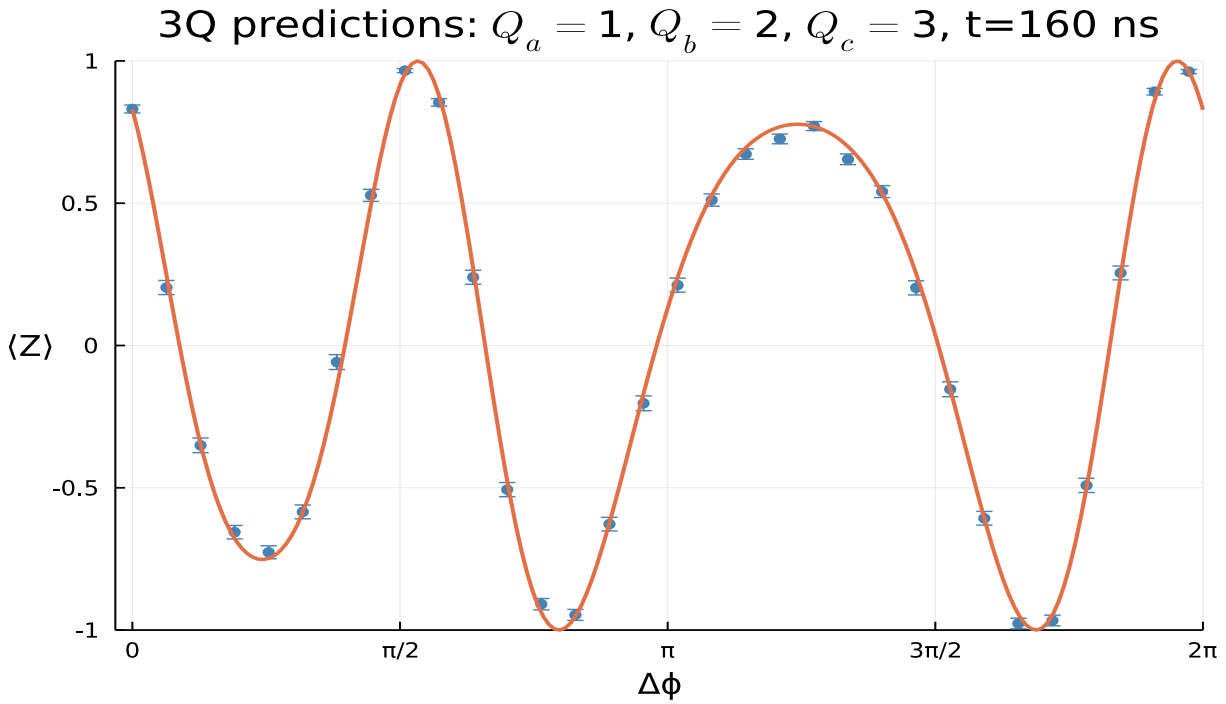

Figure 4.4: Plots comparing our theory for when three driving fields are simultaneously at the resonant frequency of $Q_a = 1$ versus experimentally obtained data. We attenuate the strength of the field labeled Q_c by a factor of $\sqrt{2}$. The blue error bars on experimental data designate the standard deviation of the error. We compute the value of the reduced chi-square as 1.07, which indicates that the theory and the observed data closely match. Note that the dynamics deviate significantly from a double field experiment with $Q_a = 1$, $Q_b = 2$ and an experiment with $Q_a = 1$, $Q_b = 3$. This observation provides strong evidence that our crosstalk model is appropriate and accurately describes the device.

Chapter 5

Conclusion

We have developed a framework for modeling crosstalk on experimental devices that have proven scalable and effective in capturing both local and nonlocal crosstalk. We achieved these results by taking advantage of the tensor product structure of local crosstalk and extending our approach to nonlocal crosstalk. Specifically, we observed the effects of crosstalk on superconducting transmon qubits. Despite significant local crosstalk, our experiments showed that error rates close to the crosstalk-free limit could be achieved using modern control hardware. Additionally, we demonstrated how tuning up simultaneous cross-resonance gates can yield significantly lower error rates.

Our findings have important implications for the development of quantum information processors. We have shown that crosstalk need not be a major limitation on noisy intermediate-scale quantum devices. Instead, our techniques can be used to balance device fabrication and pulse design tradeoffs better, potentially leading to higher-quality quantum information processors.

We have made considerable progress on the experimental side, but much more work is needed. Through our analysis of the Hamiltonian for simultaneously driven uncoupled fixed-frequency transmons, we have discovered a new and highly effective method for estimating crosstalk - the double field method. This innovative approach offers greater accuracy in determining crosstalk parameters with fewer measurements than standard Rabi experiments. Our use of fit data from double field experiments has validated our method, demonstrating close agreement between theoretical predictions and experimental results. Additionally, we have ensured the accuracy of our model by determining the theoretical Hamiltonian corresponding to three driving fields. These findings have significant implications for the development of more efficient and precise measurement techniques for

quantum devices.

There are still several outstanding problems that we hope to address:

1. Mitigating crosstalk - We have made substantial progress on the experimental side, but more technical work is needed to try and realize the theoretical improvements we showed.

2. Crosstalk on other devices - driving field crosstalk is much smaller on trapped ion quantum computers. Can our methods yield improvements with them? What other platforms have crosstalk, and can we develop a universal theoretical model that works on many different systems?

References

[1] Deanna M. Abrams, Nicolas Didier, Shane A. Caldwell, Blake R. Johnson, and Colm A. Ryan. Methods for measuring magnetic flux crosstalk between tunable transmons. *Physical Review Applied*, 12:064022, 12 2019.

[2] Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error rate. *SIAM Journal on Computing*, 38:1207–1282, 6 1999.

[3] Joseph L. Allen, Robert Kosut, and Eran Ginossar. Minimal time robust two qubit gates in circuit qed. 2019.

[4] Sergio Boixo, Sergei V. Isakov, Vadim N. Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J. Bremner, John M. Martinis, and Hartmut Neven. Characterizing quantum supremacy in near-term devices. *Nature Physics*, 14:595–600, 6 2018.

[5] Arnaud Carignan-Dugas, Joel J. Wallman, and Joseph Emerson. Bounding the average gate fidelity of composite channels using the unitarity. *New Journal of Physics*, 21:053016, 5 2019.

[6] Zijun Chen, Julian Kelly, Chris Quintana, R. Barends, B. Campbell, Yu Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Lucero, E. Jeffrey, A. Megrant, J. Mutus, M. Neeley, C. Neill, P. J.J. O'Malley, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Korotkov, John M. Martinis, P. J. J. O'Malley, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Korotkov, and John M. Martinis. Measuring and suppressing quantum state leakage in a superconducting qubit. *Physical Review Letters*, 116:020501, 1 2016.

[7] Jerry Chow, Oliver Dial, and Jay Gambetta. Ibm quantum breaks the 100-qubit processor barrier. *IBM Research Blog*, 2021.

[8] Jerry M. Chow, A. D. Córcoles, Jay M. Gambetta, Chad Rigetti, B. R. Johnson, John A. Smolin, J. R. Rozen, George A. Keefe, Mary B. Rothwell, Mark B. Ketchen, and M. Steffen. Simple all-microwave entangling gate for fixed-frequency superconducting qubits. *Physical Review Letters*, 107:080502, 8 2011.

[9] Isaac L. Chuang, Neil Gershenfeld, and Mark Kubinec. Experimental implementation of fast quantum searching. *Physical Review Letters*, 80:3408–3411, 4 1998.

[10] D. P. L. Aude Craik, N. M. Linke, M. A. Sepiol, T. P. Harty, J. F. Goodwin, C. J. Ballance, D. N. Stacey, A. M. Steane, D. M. Lucas, and D. T. C. Allcock. High-fidelity spatial and polarization addressing of ca+ 43 qubits using near-field microwave control. *Physical Review A*, 95:022337, 2 2017.

[11] Reinhard Diestel. Graph theory. *Graduate Texts in Mathematics*, 101, 2005.

[12] Steven T. Flammia and Joel J. Wallman. Efficient estimation of pauli channels. 2019.

[13] J. M. Gambetta, F. Motzoi, S. T. Merkel, and F. K. Wilhelm. Analytic control methods for high-fidelity unitary operations in a weakly nonlinear oscillator. *Physical Review A*, 83:012308, 1 2011.

[14] Jay M. Gambetta, Jerry M. Chow, and Matthias Steffen. Building logical qubits in a superconducting quantum computing system. *npj Quantum Information*, 3:2, 12 2017.

[15] Neil A Gershenfeld and Isaac L Chuang. Bulk spin-resonance quantum computation. *Science*, 275:350–356, 1 1997.

[16] Daniel Gottesman. Class of quantum error-correcting codes saturating the quantum hamming bound. *Physical Review A*, 54:1862–1868, 9 1996.

[17] Gilbert Grynberg, Alain Aspect, and Claude Fabre Fabre. *Introduction to quantum optics: from the semi-classical approach to quantized light.* Cambridge University Press, 2010.

[18] E. Knill, R. Laflamme, R. Martinez, and C.-H. Tseng. An algorithmic benchmark for quantum information processing. *Nature*, 404:368–370, 3 2000.

[19] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. *Physical Review A*, 77:012307, 1 2008.

[20] Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek. Resilient quantum computation. *Science*, 279:342–345, 1998.

[21] Jens Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Charge-insensitive qubit design derived from the cooper pair box. *Physical Review A*, 76:042319, 10 2007.

[22] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[23] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver. A quantum engineer's guide to superconducting qubits. *Applied Physics Reviews*, 6:021318, 6 2019.

[24] Elliott H. Lieb and Derek W. Robinson. The finite group velocity of quantum spin systems. *Communications in Mathematical Physics*, 28:251–257, 9 1972.

[25] Shai Machnes, Elie Assémat, David Tannor, and Frank K. Wilhelm. Tunable, flexible, and efficient optimization of control pulses for practical qubits. *Physical Review Letters*, 120:150401, 4 2018.

[26] Easwar Magesan and Jay M. Gambetta. Effective hamiltonian models of the cross-resonance gate. *arXiv preprint arXiv:1804.04073*, 4 2018.

[27] David C. McKay, Christopher J. Wood, Sarah Sheldon, Jerry M. Chow, and Jay M. Gambetta. Efficient z gates for quantum computing. *Physical Review A*, 96, 2017.

[28] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm. Simple pulses for elimination of leakage in weakly nonlinear qubits. *Physical Review Letters*, 103:110501, 9 2009.

[29] Prakash Murali, David C. McKay, Margaret Martonosi, and Ali Javadi-Abhari. Software mitigation of crosstalk on noisy intermediate-scale quantum computers. 1 2020.

[30] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, R. Barends, B. Burkett, Y. Chen, Z. Chen, A. Fowler, B. Foxen, M. Giustina, R. Graff, E. Jeffrey, T. Huang, J. Kelly, P. Klimov, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, H. Neven, and J. M. Martinis. A blueprint for demonstrating quantum supremacy with superconducting qubits. *Science*, 360:195–199, 4 2018.

[31] Michael A. Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, 303:249–252, 10 2002.

[32] Ch Piltz, B. Scharfenberger, A. Khromova, A. F. Varón, and Ch Wunderlich. Protecting conditional quantum gates by robust dynamical decoupling. *Physical Review Letters*, 110:200501, 5 2013.

[33] M. Pioro-Ladrière, T. Obata, Y. Tokura, Y.-S. S. Shin, T. Kubo, K. Yoshida, T. Taniyama, and S. Tarucha. Electrically driven single-electron spin resonance in a slanting zeeman field. *Nature Physics*, 4:776–779, 10 2008.

[34] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 8 2018.

[35] Matthew Reagor, Christopher B. Osborn, Nikolas Tezak, Alexa Staley, Guenevere Prawiroatmodjo, Michael Scheer, Nasser Alidoust, Eyob A. Sete, Nicolas Didier, Marcus P. da Silva, Ezer Acala, Joel Angeles, Andrew Bestwick, Maxwell Block, Benjamin Bloom, Adam Bradley, Catvu Bui, Shane Caldwell, Lauren Capelluto, Rick Chilcott, Jeff Cordova, Genya Crossman, Michael Curtis, Saniya Deshpande, Tristan El Bouayadi, Daniel Girshovich, Sabrina Hong, Alex Hudson, Peter Karalekas, Kat Kuang, Michael Lenihan, Riccardo Manenti, Thomas Manning, Jayss Marshall, Yuvraj Mohan, William O'Brien, Johannes Otterbach, Alexander Papageorge, Jean-Philip Paquette, Michael Pelstring, Anthony Polloreno, Vijay Rawat, Colm A. Ryan, Russ Renzas, Nick Rubin, Damon Russel, Michael Rust, Diego Scarabelli, Michael Selvanayagam, Rodney Sinclair, Robert Smith, Mark Suska, Ting-Wai To, Mehrnoosh Vahidpour, Nagesh Vodrahalli, Tyler Whyland, Kamal Yadav, William Zeng, and Chad T. Rigetti. Demonstration of universal parametric entangling gates on a multi-qubit lattice. *Science Advances*, 4:eaao3603, 2 2018.

[36] Chad Rigetti, Alexandre Blais, and Michel Devoret. Protocol for universal gates in optimally biased superconducting qubits. *Physical Review Letters*, 94:240502, 6 2005.

[37] Kenneth Rudinger, Timothy Proctor, Dylan Langharst, Mohan Sarovar, Kevin Young, and Robin Blume-Kohout. Probing context-dependent errors in quantum processors. *Physical Review X*, 9:021045, 6 2019.

[38] Mohan Sarovar, Timothy Proctor, Kenneth Rudinger, Kevin Young, Erik Nielsen, and Robin Blume-Kohout. Detecting crosstalk errors in quantum information processors. *Quantum*, 4:321, 9 2020.

[39] J. A. Schreier, A. A. Houck, Jens Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Suppressing charge noise decoherence in superconducting charge qubits. *Physical Review B*, 77:180502, 5 2008.

[40] R. Schutjens, F. Abu Dagga, D. J. Egger, and F. K. Wilhelm. Single-qubit gates in frequency-crowded transmon systems. *Physical Review A*, 88:052330, 11 2013.

[41] Sarah Sheldon, Easwar Magesan, Jerry M. Chow, and Jay M. Gambetta. Procedure for systematically tuning up cross-talk in the cross-resonance gate. *Physical Review A*, 93:060302, 6 2016.

[42] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52:R2493–R2496, 10 1995.

[43] Andrew M. Steane. Overhead and noise threshold of fault-tolerant quantum error correction. *Physical Review A*, 68:042322, 10 2003.

[44] Matthias Steffen, Lieven M.K. Vandersypen, and Isaac L. Chuang. Simultaneous soft pulses applied at nearby frequencies. *Journal of Magnetic Resonance*, 146:369–374, 10 2000.

[45] L. S. Theis, F. Motzoi, and F. K. Wilhelm. Simultaneous gates in frequency-crowded multilevel systems using fast, robust, analytic control shapes. *Physical Review A*, 93:012324, 1 2016.

[46] Robert R. Tucci. An introduction to cartan's kak decomposition for qc programmers. 7 2005.

[47] Adam Winick, Joel J. Wallman, and Joseph Emerson. Simulating and mitigating crosstalk. *Physical Review Letters*, 126:230502, 6 2021.

[48] T. Xia, M. Lichtman, K. Maller, A. W. Carr, M. J. Piotrowicz, L. Isenhower, and M. Saffman. Randomized benchmarking of single-qubit gates in a 2d array of neutral-atom qubits. *Physical Review Letters*, 114:100503, 3 2015.

[49] X. Xue, T. F. Watson, J. Helsen, D. R. Ward, D. E. Savage, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, S. Wehner, and Lieven M.K. Vandersypen. Benchmarking gate fidelities in a si / sige two-qubit device. *Physical Review X*, 9:21011, 2019.