Dissertation

zur Erlangung des Grades des
Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät der

Universität des Saarlandes

Optimal control for computing with near-term intermediate-scale quantum devices

Saarbrücken, 2024

Nicolas Wittler

UNIVERSITÄT
DES
SAARLANDES

Tag des Kolloquiums:	27.03.2025
Dekan:	Prof. Dr.-Ing. Dirk Bähre
Berichterstatter:	Prof. Dr. Frank K. Wilhelm-Mauch
	Prof. Dr. Michael Hartmann
Akademisches Mitglied:	Dr. (habil.) Philipp Hövel
Vorsitz:	Prof. Dr. Christoph Becher

Abstract

In the current era between early prototypes and large scale quantum computers, techniques to operate and mitigate potential errors during computation are still developing. In this thesis, we look at examples where an initial model description appears unable to explain experimentally observed behavior.

We see that two-level theory results cannot be applied to a real system where additional energy levels play a significant role. Consequently, optimizing controls for this oversimplified system will not work in experiment. Given a suitable model, complex controls can be derived and transferred to experiment, where they are further fine-tuned to account for small inaccuracies.

With $\mathbb{C}^3$ – Control, Calibration and Characterization – we add a third step: We use a record of measurement data to fine-tune the model to resemble the behavior of real experiments. This improved model serves as a basis for another loop of deriving controls which perform well on the experiment. Successively more detailed, matching models then enable the implementation of more complex control schemes.

Finally, we look to a theoretical study of possible models. Here, we tune the controls and model in tandem with the goal of achieving the best performance. Transferring sophisticated control schemes into practice can be hindered by exorbitant calibration requirements. We therefore limit our investigation to pulses with sparse descriptions and examine known design regimes.

Zusammenfassung

Während die Anwendung von Quantencomputern sich zwischen Grundlagenforschung und Prototypen befindet, ist eine der großen Herausforderungen ihre präzise Steuerung, für die eine gewissenhafte Charakterisierung notwendig ist. Wir befassen uns in dieser Arbeit mit Beispielen, bei denen durch die Vernachlässigung höher angeregter Zustände die experimentelle Beobachtung nicht quantitativ beschrieben und daher die Verwendung etablierter *optimal control*-Methoden verhindert wird.

Mit $\mathbb{C}^3$ – *Control, Calibration and Characterization* – präsentieren wir eine Methode, die systematisch Unterschiede zwischen Modellbeschreibung und Labordaten minimiert, und es ermöglicht, durch sukzessiv detailliertere Modellierung komplexere Steuerungsmethoden anzuwenden. Aufbauend auf bisherigen Methoden wie Ad-Hoc (*adaptive hybrid optimal control*), die bereits die numerisch optimierten Steuerpulse im Feedback mit dem Experiment anpasst, liefern diese Daten nun die Grundlage, das numerische Modell zu verfeinern und eine konsistente Charakterisierung vorzunehmen.

Anschließend stellen wir die Frage, welchen Vorteil es bringt Modellparameter gemeinsam mit den Parametern der Steuerpulse zu optimieren, um daraus Ableitungen für neue Entwürfe zu treffen. Als theoretische, numerische Studie betrachten wir ein wechselwirkendes Zweiqubitsystem in gängigen Parameterregimen, dispersiv bis stark gekoppelt, und optimieren Pulse mit durch wenige Parameter beschriebenen Ansätzen.

List of Publications

Characterizing two-photon transitions in a strongly driven Transmon
Nicolas Wittler, Shai Machnes and Frank K. Wilhelm
Update from Master's thesis, in preparation.
Appears here as Chapter 2.

Integrated Tool Set for Control, Calibration, and Characterization of Quantum Devices Applied to Superconducting Qubits
Nicolas Wittler, Federico Roy, Kevin Pack, Max Werninghaus, Anurag Saha Roy, Daniel J. Egger, Stefan Filipp, Frank K. Wilhelm, and Shai Machnes
Phys. Rev. Applied **15**, 034080 — Published 29 March 2021
Appears here as Chapter 3.

Co-designing Transmon devices for control with simple pulses
Nicolas Wittler, Shai Machnes and Frank K. Wilhelm
arXiv:2410.13619 – Submitted on October 17th, 2024
Appears here as Chapter 4.

Other works

An introduction into optimal control for quantum technologies
Frank K. Wilhelm, Susanna Kirchhoff, Shai Machnes, Nicolas Wittler, Dominique Sugny
Lecture notes for the 51st IFF Spring School
arXiv:2003.10132 – Submitted on March 23rd, 2020

Software tool-set for automated quantum system identification and device bring up
Anurag Saha Roy, Kevin Pack, Nicolas Wittler, Shai Machnes
Technical paper
arXiv:2205.04829 – Submitted on May 10th, 2022

Contents

List of Figures

List of Tables

Acknowledgments

Looking back at a little over five years, I'd like to thank the people that I worked with through these "interesting" times: a pandemic and the move from a rundown university building to a slightly less rundown but freshly painted building at one of the biggest research centers in Europe, following a stint in a metal box on the parking lot of the neighboring institute. During which the research group also expanded into a full institute; growing pains and all.

Starting with the Bachelor's thesis, I spent a good chunk of time in group of Frank Wilhelm-Mauch and colleagues. I think many of us agree, that what we value about his leadership style is the freedom he awards to every single one of us to pursue the topics that intrigue us the most. For a long time I worked with Shai Machnes as direct supervisor, who not once made me feel anything less than an equal collaborator. I enjoyed helping in teaching with Susanna, Kevin and Miralem. As people in the same boat – about to finish a Ph.D. – we shared a lot of the insights and frustrations in getting a thesis together with Gino, Francesco, Susanna, and Sven. Also thanks to David Bruschi, for many burgers, beer, and complaining about Deutsche Bahn.

Special thanks belong to Federico, whom I share the first author on our first paper with. How we managed to put all of that work together in a relatively short time still amazes me.

Finally, I'll thank the PGI-12 Skeleton Crew, Ashutosh, Paul and Gerhard, and the administrative assistants, Andrea, Luise, Michaela and Angelina, who keep the show on the road.

Hey man of science with your perfect rules of measure
Can you improve this place with the data that you gather?
– Brett Gurewitz

1 Introduction – Quantum devices as computers

The field of quantum computing is maturing. First fundamental discoveries have been made in laboratory devices over 25 years ago [1]. As part of the quantum flagship of the European Union [2] and German national programms, a lot of effort is made by research collaborations, such as OpenSuperQ [3] and GeQCoS [4], to produce practical demonstrators and by companies to present early products, that promise to exploit the laws of quantum mechanics, mainly the concepts of superposition and entanglement, to solve computational problems more effectively than a classical computer. These devices are called *Noisy Intermediate-Scale Quantum computers* (NISQ) [5] – noisy, because they are affected by interference from the environment, and intermediate-scale, because of the challenge in creating many quantum bits that still retain their quantum character.

In this thesis, we will look at advancing the operation and design of NISQ devices by applying the tools of optimal control. These present a clear formalism to investigate whether certain goals are realizable, how to mitigate limitations in current implementations, and thus bridge the gap between foundational experimental research and application driven device development.

1.1 Logical gates on quantum devices

In the context of quantum computing, it is a useful perspective to look at the basics of quantum physics as the *theory that explains the device operation.* As such, the register – a collection of qubits – of a quantum computer (or any state in general) is governed by the Schrödinger equation

$$i\hbar\partial_t |\Psi\rangle = H |\Psi\rangle \tag{1.1}$$

with a time-dependent Hamiltonian H. If the unitary operator U is the formal solution, i.e.,

$$U = \mathbb{T} \exp\left\{-i/\hbar \int dt H(t)\right\} \tag{1.2}$$

that maps an initial state $|\Psi(0)\rangle$ to a final state $|\Psi(T)\rangle = U|\Psi(0)\rangle$, then the computation task to solve a problem can be defined as:

For an input $|\Psi(0\rangle$, find $H(t)$ with $0 \leq t \leq T$ such that the corresponding state $|\Psi(T)\rangle$ at the final time T encodes the solution to the problem.

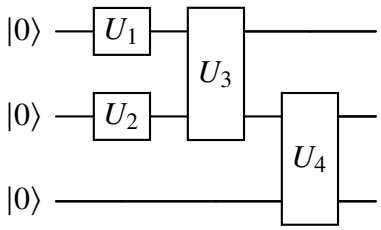

Figure 1.1: A representation of a quantum computation in the gate-based model as a circuit diagram.

Here, "encodes" means there is a way to extract the answer from the state, although that might not be trivial. This approach is more or less directly applied in *adiabatic quantum computing*[6]. Here, the problem is formulated by designing a "problem" Hamiltonian H_P where the ground state corresponds to the solution. Then, the Hamiltonian is tuned, starting from an initial Hamiltonian H_I with an easily preparable ground state to H_P. Following the adiabatic theorem, the system will end up in the ground state of H_P, as desired.

The *gate-based* model more closely resembles the classical notion of a computation. From a formal perspective, the time-evolution is decomposed into a sequence

$$U = U_N U_{N-1} \dots U_2 U_1 \tag{1.3}$$

of "elementary" gates $U_i \in \mathcal{G}$, selected from a gate-set $\mathcal{G}$ of pre-defined operations. Here, "elementary" means each U_i only acts on a subset of the full Hilbert space or that they span the complete set of operators but are themselves simple, as in Fig. 1.1.

If we consider the state of a single qubit as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, a two-level system with suggestively named basis states $|0\rangle$ and $|1\rangle$ to correspond to a classical bit, the state of a full register of bits is given by, e.g., $|\Psi\rangle = |10001110101\rangle$.

From this representation, we can appreciate the advantage of quantum computing: The register can be in a state like this that corresponds directly to a classical bit-string but also in a superposition, for example $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Such a state has no equivalent in classical computing. Many quantum algorithms begin with creating a register of every possible input string to then query a function only a single time, providing an exponential scaling with the number of bits compared to classical computing.

The gate-set to act on this register typically consists of single- or two-qubit operations, i.e., each U_i acts non-trivially on one or two subspaces, and as identity on the others. This is because they mostly rely on interaction between physical qubits, but also designs to create gates entangling more than two qubits are being pursued.

In the field, a number of specific two-qubit gates are named, e.g.

$$\begin{aligned}
\text{CNOT} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, & \text{iSWAP} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\
\sqrt{\text{iSWAP}} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, & \text{CZ} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.
\end{aligned} \tag{1.4}$$

1.2 Entanglement

What makes this kind of computing unique is the implication of the quantum behavior of the register on the complexity of a computation: Entanglement is one of the key ingredients for a fundamental advantage over classical computation. This section reviews some mathematical concepts to describe and categorize the entanglement of states and the operations that create it [7].

1.2.1 Locality in two-qubit gates

To introduce some concepts, we consider a system of two identical spins $H_Q = \frac{\omega}{2}\sigma_z^1 + \frac{\omega}{2}\sigma_z^2$ where we denote $\sigma_z^1 := \sigma_z \otimes \mathbb{1}$, etc. and the state of this system as a product state $|\Psi\rangle = |\psi_1\rangle^1 \otimes |\psi_2\rangle^2$.

Since both operators only act on one subspace, the Schrödinger equation for the two spins

$$i\partial_t\left(|\psi_1\rangle^1 \otimes |\psi_2\rangle^2\right) = \frac{\omega}{2}\sigma_z^1 |\psi_1\rangle^1 \otimes \frac{\omega}{2}\sigma_z^2 |\psi_2\rangle^2 \ . \tag{1.5}$$

can be solved completely independently.

When we add an interaction $H_I = \hbar g \sigma_x \otimes \sigma_x = \hbar g \sigma_x^1 \sigma_x^2$, we can create an entangled state, e.g., one of

$$\left\{\frac{|00\rangle + |11\rangle}{\sqrt{2}}, \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \frac{|01\rangle - |10\rangle}{\sqrt{2}}\right\} \tag{1.6}$$

where the state of one spin is completely determined by the measurement of other spin, and we can no longer express it as a product.

To generalize, if we look at the Hamiltonian H of a system and the time evolution U it generates, they are formally connected by the exponential map

$$\exp : A \mapsto e^{At} \tag{1.7}$$

with some real parameter t. The skew-hermitian operator A is an element of an *algebra* and the elements e^{At} form a *group*.

The mathematics that we use to describe quantum physics is Lie group theory. A central idea is the Lie bracket, a binary operation $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ on a vector space $\mathfrak{g}$ that

- is bilinear,
- satisfies that $[x, x] = 0, \forall x \in \mathfrak{g}$
- and the Jacobi identity $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \; \forall x, y, z \in \mathfrak{g}$.

A vector space $\mathfrak{g}$ over a field $\mathbb{F}$ together with a Lie bracket forms a *Lie algebra*.

If we consider a matrix U from the unitary group, named $\mathbf{U}(n)$, we know that $U^\dagger = U^{-1}$ or $e^{tX^\dagger} = e^{-tX}$, which means that the corresponding algebra $u(n)$ consists of skew-Hermitian matrices $X^\dagger = -X$. Since U is unitary, it has $|\det\{U\}| = \left|e^{i\varphi}\right| = 1$. The subgroup with $\det\{U\} = 1$, instead of a complex number with absolute value 1, is called the *special unitary group* $\mathbf{SU}(n)$ and its algebra is $su(n)$.

Limiting ourselves to the Hilbert space of a single two-state system $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$, with $\alpha, \beta \in \mathbb{C}$, we define the basis $|0\rangle, |1\rangle$ as the eigenstates of the Pauli matrix σ_z. Together, the Pauli matrices form a representation of the algebra $su(2) = \text{span}\frac{i}{2}\left\{\sigma_x, \sigma_y, \sigma_z\right\}$

As seen above, any single qubit unitary matrix is an element of $\mathbf{SU}(2)$ (or $\mathbf{U}(2)$ if we care about the global phase) and can thus be generated by the Lie algebra $su(2)$, which can be completely described by the (skewed) Pauli matrices $i\sigma_x$, $i\sigma_y$ and $i\sigma_z$.

Let us consider the Lie group of all two-qubit unitaries $\mathbf{SU}(4)$ and its subgroup $\mathbf{SU}(2) \otimes \mathbf{SU}(2)$, the Lie group of all simultaneous single qubit gates. To characterize the entanglement power of a given element U of $\mathbf{SU}(4)$, we can employ a Cartan decomposition [8] to find a suitable subalgebra of $su(4)$.

The algebra corresponding to $\mathbf{SU}(2) \otimes \mathbf{SU}(2)$ is

$$\mathfrak{k} = su(2) \otimes su(2) = \text{span}\frac{i}{2}\left\{\sigma_x^1, \sigma_y^1, \sigma_z^1, \sigma_x^2, \sigma_y^2, \sigma_z^2\right\} \tag{1.8}$$

We decompose an algebra $\mathfrak{g}$ into a subalgebra $\mathfrak{k}$ and its complement $\mathfrak{p} = \mathfrak{k}^\perp$ such that $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$. If $\mathfrak{p}$ and $\mathfrak{k}$ satisfy the commutation relations

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \quad [\mathfrak{p}, \mathfrak{k}] \subset \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k} \tag{1.9}$$

this is called a Cartan decomposition of $\mathfrak{g}$.

Let $\mathfrak{g} = su(4)$, then

$$\mathfrak{k} = \text{span}\frac{i}{2}\left\{\sigma_x^1, \sigma_y^1, \sigma_z^1, \sigma_x^2, \sigma_y^2, \sigma_z^2\right\} \tag{1.10}$$

$$\mathfrak{p} = \text{span}\frac{i}{2}\left\{\sigma_x^1\sigma_x^2, \sigma_x^1\sigma_y^2, \sigma_x^1\sigma_z^2, \sigma_y^1\sigma_x^2, \sigma_y^1\sigma_y^2, \sigma_y^1\sigma_z^2, \sigma_z^1\sigma_x^2, \sigma_z^1\sigma_y^2, \sigma_z^1\sigma_z^2\right\} \tag{1.11}$$

is a Cartan decomposition of $\mathfrak{g}$.

This allows us to make the following statement: Given K, with its algebra $\mathfrak{k}$, a compact subgroup of G and $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$ the Lie algebra of G. Let $\mathfrak{a} \subset \mathfrak{p}$ be a Cartan subalgebra of

the pair $(\mathfrak{g}, \mathfrak{k})$, which means that $\mathfrak{p} = \bigcup_{k \in \mathfrak{k}} [k, \mathfrak{a}]$. Then, we can obtain the full Lie group from $G = K \exp(\mathfrak{a}) K$.

Taking the diagonal elements of $\mathfrak{p}$, we can write the Abelian subalgebra

$$\mathfrak{a} = \operatorname{span} \frac{i}{2} \left\{ \sigma_x^1 \sigma_x^2, \sigma_y^1 \sigma_y^2, \sigma_z^1 \sigma_z^2 \right\} \tag{1.12}$$

which is a Cartan subalgebra of $\mathfrak{p}$, since $[\sigma_i^1, \sigma_j^1 \sigma_j^2] = 2i\epsilon_{ijk} \sigma_k^1 \sigma_j^2 \in \mathfrak{p}$.

Then, we can write any $U \in \mathbf{SU}(4)$ as

$$U = k_1 A k_2 = k_1 \exp\left\{ \frac{i}{2} \left[c_1 \sigma_x^1 \sigma_x^2 + c_2 \sigma_y^1 \sigma_y^2 + c_3 \sigma_z^1 \sigma_z^2 \right] \right\} k_2 \tag{1.13}$$

with $k_1, k_2 \in \mathbf{SU}(2) \otimes \mathbf{SU}(2)$ and the *Weyl coordinates* $c_1, c_2, c_3 \in \mathbb{R}$. Thus, we can specify an entangling gate with three real numbers, see also Table 1.1.

1.2.2 Entanglement of pure states

While the Weyl coordinates shown above define the entangling character of a gate, they are not easy to compute for a given operator U. We'll look at finding a different set of numbers, the Makhlin invariants [9], to contain the non-local properties.

Given a pure state $|\psi\rangle = \sum_{ab} c_{ab} |ab\rangle$, we define the entanglement $\operatorname{Ent} |\psi\rangle := c_{00} c_{11} - c_{01} c_{10}$ or $\operatorname{Ent} |\psi\rangle = \det \hat{\psi}$, where

$$\hat{\psi} = \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} \tag{1.14}$$

We quickly verify with $\operatorname{Ent}(|00\rangle + |10\rangle)/\sqrt{2} = (1 \cdot 0 - 0 \cdot 1)/2 = 0$ and $\operatorname{Ent}(|00\rangle + |11\rangle)/\sqrt{2} = (1 \cdot 1 - 0 \cdot 0)/2 = 1/2$ that CNOT creates entanglement.

Any state transformed by single qubit operations $k_1 \otimes k_2 \in \mathbf{SU}(2) \otimes \mathbf{SU}(2)$ does not change this measure, since $\det \hat{\psi}' = \det\left(k_1 \hat{\psi} k_2^T\right) = \det k_1 \det \hat{\psi} \det k_2^T = \det \hat{\psi}$.

We define the transformation of a matrix from the product basis into the Bell basis as $k \mapsto Q^\dagger k Q$ with

$$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & i \\ 0 & i & 1 & 0 \\ 0 & i & -1 & 0 \\ 1 & 0 & 0 & -i \end{pmatrix} \tag{1.15}$$

It can be shown [8] that the algebra of local gates $\mathfrak{k} = su(2) \otimes su(2)$ is isomorphic to the special orthogonal algebra $so(4)$, with the isomorphism $\phi : \mathfrak{k} \to so(4)$ given by $\phi(k) = Q^\dagger k Q$.

Gate	Makhlins	Weyl $/\pi$	Perfect entangler
local	$(1, 0, 3)$	$(0, 0, 0)$	x
CNOT	$(0, 0, 1)$	$(0.5, 0, 0)$	✓
CZ	$(0, 0, 1)$	$(0.5, 0, 0)$	✓
CR	$(0, 0, 1)$	$(0.5, 0, 0)$	✓
$\sqrt{\text{iSWAP}}$	$(0.25, 0, 1)$	$(0.25, 0.25, 0.25)$	✓
iSWAP	$(0, 0, 1)$	$(0.5, 0.5, 0.5)$	x

Table 1.1: Makhlin invariants and Weyl coordinates of common gates.

Transforming our subalgebra $\mathfrak{a}$ into the Bell basis

$$Q^\dagger \left\{ \sigma_x^1 \sigma_x^2, \sigma_y^1 \sigma_y^2, \sigma_z^1 \sigma_z^2 \right\} Q = \left\{ \sigma_z^1, -\sigma_z^2, \sigma_z^1 \sigma_z^2 \right\} \tag{1.16}$$

makes

$$F = Q^\dagger \exp\left\{ \frac{i}{2} \left[c_1 \sigma_x^1 \sigma_x^2 + c_2 \sigma_y^1 \sigma_y^2 + c_3 \sigma_z^1 \sigma_z^2 \right] \right\} Q \tag{1.17}$$

diagonal with eigenvalues

$$\left\{ e^{\frac{i}{2}(c_1 - c_2 + c_3)}, e^{\frac{i}{2}(c_1 + c_2 - c_3)}, e^{-\frac{i}{2}(c_1 + c_2 + c_3)}, e^{\frac{i}{2}(-c_1 + c_2 + c_3)} \right\}. \tag{1.18}$$

Now, we can transform the arbitrary unitary U to

$$U_B = Q^\dagger U Q = Q^\dagger k_1 A k_2 Q = Q^\dagger k_1 Q \; F \; Q^\dagger k_2 Q = O_1 F O_2 \tag{1.19}$$

and compute

$$m = U_B^T U_B = O_2^T \, F^2 \, O_2 \tag{1.20}$$

Since O_2 is orthogonal, the eigenvalues of F^2 are unchanged and completely determined by entangling operations.

The spectrum of $m = U_B^T U_B$ is given by its characteristic polynomial

$$|\lambda - m| = \lambda^4 + \operatorname{tr}(m)\lambda^3 + \frac{1}{2}\left[\operatorname{tr}^2(m) - \operatorname{tr}\left(m^2\right) \right] \lambda^2 - \operatorname{tr}(m)^* \lambda + 1 \tag{1.21}$$

which is determined by the quantities $\operatorname{tr}(m)$ and $\operatorname{tr}^2(m) - \operatorname{tr}\left(m^2\right)$.

From these, we define the *Makhlin invariants* of a unitary operation $U \in \mathbf{U}(4)$ with the Bell basis Q and $U_B = Q^\dagger U Q$ as

$$\begin{aligned} G_1 &= \operatorname{Re} \frac{\operatorname{tr}\left(U_B^T U_B\right)^2}{16 \det(U)} \quad , \quad G_2 = \operatorname{Im} \frac{\operatorname{tr}\left(U_B^T U_B\right)^2}{16 \det(U)} \\ G_3 &= \frac{\operatorname{tr}\left(U_B^T U_B\right)^2 - \operatorname{tr}\left((U_B^T U_B)^2\right)}{4 \det(U)} \end{aligned} \tag{1.22}$$

In contrast to the Weyl coordinates, these invariants are now easy to compute for a given unitary matrix. In summary, we now have the mathematical tool-set in hand to describe the entanglement characteristics of a quantum operation that we want to exploit for computation. Next, we turn to the physical systems that serve as the basis for a potential quantum computer.

1.3 Hardware

Several candidate platforms exist to implement a quantum computer. Any real-world device has multiple degrees of freedom and complicated dynamics, so the principal task is to prepare suitably isolated, well controlled subsystems to implement the qubits. The area with the longest tradition in applying quantum optimal control is nuclear magnetic resonance (NMR). Multiple practical applications exist, e.g., medical like magnetic resonance imaging (MRI) resulting in available mature technologies. The qubits themselves are already spin degrees of freedom, so ideal two-level systems. Additionally, time scales for the control electronics are favorable and allow implementing complex pulse shapes to steer the qubits. It makes sense then, that in this field also where complex optimal control techniques were also applied early on [10].

The same advantage of well characterized spectra is present in ion traps [11], where each qubit is guaranteed to be identical. Since they also share an atomic trap potential and associated degrees of motion, entangling multiple qubits with each other is also easily facilitated. However, the size of the traps also presents a natural limit to scaling up devices. When the limit is reached, other methods need to be applied to extend the computational space.

Recently, NV centers [12] have shown promising results by combing two time scales. Optical transitions and microwave activated transitions can be combined to create an architecture for quantum computing.

Many of the results driving big discoveries have been made with the use of superconducting devices [13, 14]. These have the double-edged distinction of made by humans. As such, these devices can be fabricated with specific goals in mind, with the price of having then to characterize them carefully. Tuning up these devices for computation is not as straight forward as e.g., using atomic transitions. The fabrication itself presents a number of challenges, the qubit properties depend on the process. Further, impurities can introduce unwanted, spurious resonances that couple to the computational degrees of freedom, even intermittently, and disturb device operation.

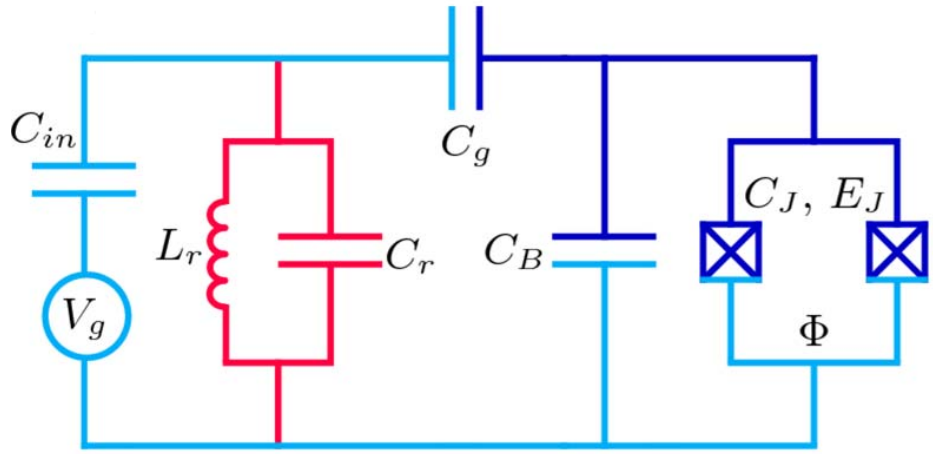

Figure 1.2: Reduced network for the Transmon device, coupled to a resonator. reproduced from [16].

1.3.1 Circuit quantization

To arrive at a quantum mechanical description of superconducting circuits, we need to perform a few steps known as circuit quantization [15]. Here, we'll show this procedure on the example to find the Hamiltonian of the Transmon, a specific superconducting device that has been used in many demonstrations of quantum utility.

In Fig. 1.2 we show the diagram of the circuit. Starting with the resonator, modeled as an electrical oscillator with capacity C_r and inductance L_r, we write Kirchhoff's laws for Voltage V and current I in the loop as

$$0 = I_L + I_C = -C_r\ddot{\phi} + \frac{1}{L_r}\phi \,. \tag{1.23}$$

We interpret this differential equation as an Euler-Lagrange equation

$$0 = -\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} + \frac{\partial \mathcal{L}}{\partial \phi} \tag{1.24}$$

and infer the Lagrangian

$$\mathcal{L}(\phi, \dot{\phi}) = -\frac{1}{2L_r}\phi^2 + \frac{C_r}{2}\dot{\phi}^2 \tag{1.25}$$

for the resonator, with the capacitance taking the role of a kinetic energy and the inductance energy the role of a potential.

An earlier version of this introduction to the Transmon already appeared in my master's thesis. It is reproduced and condensed here for completeness.

The canonically conjugate variable for the flux ϕ is the electric charge

$$Q_r = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = C_r \dot{\phi} \quad . \tag{1.26}$$

Performing a Legendre transformation to the conjugate charge and flux variables, we write

$$H_{\text{cavity}} = \dot{\phi}_r Q_r - \mathcal{L}(\phi_r, \dot{\phi}_r) = \frac{Q_r^2}{2C_r} + \frac{\phi_r^2}{2L_r}. \tag{1.27}$$

which we can interpret as the quantum mechanical Hamiltonian by considering the flux and charge as operators.

The variables Q_r and ϕ_r can be expressed using the bosonic ladder operators and the characteristic impedance $Z = \sqrt{L_r/C_r}$

$$\begin{aligned} a &= i\sqrt{\frac{1}{2Z}}\phi_r + \sqrt{\frac{Z}{2}}Q_r \\ a^\dagger &= -i\sqrt{\frac{1}{2Z}}\phi_r + \sqrt{\frac{Z}{2}}Q_r \end{aligned} \tag{1.28}$$

such that

$$a^\dagger a = \frac{1}{2Z}\phi_r^2 + \frac{Z}{2}Q_r^2 + i\frac{1}{2}[\phi_r, Q_r] \tag{1.29}$$

Flux and charge are conjugate variables, which means their commutator is i and

$$H_{\text{cavity}} = \frac{Q_r^2}{2C_r} + \frac{\phi_r^2}{2L_r} = \omega_r\left(a^\dagger a + \frac{1}{2}\right) \tag{1.30}$$

with the resonator frequency $\omega_r = \sqrt{\frac{1}{L_r C_r}}$ is the Hamiltonian that describes the cavity resonator as a harmonic oscillator with photon number $a^\dagger a$.

The Hamiltonian of the full circuit shown above is obtained by a similar computation [16], for now ignoring the cavity and just considering the flux ϕ_2 through the Josephson junction (dark blue in Fig. 1.2) and outside flux ϕ_1 (light blue) from the bias voltage V. In this loop the conservation of currents reads

$$\underbrace{(C_B + C_J)}_{C_1}\ddot{\phi}_1 + I_J \sin\phi_1 = \underbrace{\left(C_{\text{in}} + C_g\right)}_{C_2}\ddot{\phi}_2. \tag{1.31}$$

with the conservation of voltages the second flux $\dot{\phi}_2 = -V - \dot{\phi}_1$ can be eliminated, and the Lagrangian is

$$\mathcal{L} = \frac{C_1}{2}\dot{\phi}_1^2 + \frac{C_2}{2}\left(\dot{\phi}_1 + \dot{\phi}_V\right)^2 + I_J \cos\phi_1, \tag{1.32}$$

which determines the charge variables to be

$$\begin{aligned} Q_1 &= \frac{\partial \mathcal{L}}{\partial \dot{\phi}_1} = C_1\dot{\phi}_1 + C_2\left(\dot{\phi}_1 + \dot{\phi}_V\right) \\ Q_2 &= \frac{\partial \mathcal{L}}{\partial \dot{\phi}_V} = C_2\left(\dot{\phi}_1 + \dot{\phi}_V\right). \end{aligned} \tag{1.33}$$

Inverting this relation and performing the Legendre transformation yields the Hamiltonian

$$H = \frac{C_2(Q_1 - Q_2)^2}{2C_1C_2} + \frac{C_1Q_2^2}{2C_1C_2} - I_J\cos\phi_1 \tag{1.34}$$

With the charge and Josephson energies defined as

$$E_C = \frac{e^2}{2C_1} \quad \text{and} \quad E_J = \frac{I_J}{2e} \tag{1.35}$$

the Hamiltonian is written in an energy scale

$$H = E_C\left(\frac{Q_1^2}{e^2} - \frac{Q_2^2}{e^2}\right) + \frac{Q_2^2}{2C_2} - E_J\cos\phi_1\,. \tag{1.36}$$

Expressing the charges through number of cooper pairs $Q_1 = 2en$ and $Q_2 = 2en_g$

$$H_{\text{qubit}} = 4E_C(n - n_g)^2 - E_J\cos\phi \tag{1.37}$$

with the bias $n_g = \dfrac{C_gV_g}{2e}$, produces the same result as is obtained for a Cooper Pair Box [17].

The qubit is coupled to the resonator via the capacitor C_g, which means its current is added to equation Eq. (1.31) and the same steps are performed to calculate charges. This modifies

$$n_g = \frac{C_gV_g}{2e} + \frac{Q_r}{2e} \tag{1.38}$$

as the resonator introduces another offset. The capacitive coupling takes the form

$$H_{\text{coupling}} = \frac{\beta}{C_r}Q_1Q_r = \beta(2en)\sqrt{\frac{2\omega_r}{C_r}}\left(a^\dagger + a\right) \tag{1.39}$$

with $\beta = C_g/(C_1 + C_2)$.

1.3.2 The Transmon

In analogy to the harmonic oscillator the operators for phase ϕ and number of cooper pairs n can be written in terms of bosonic ladder operators

$$\phi = \frac{1}{\sqrt{2}}\left(\frac{8E_C}{E_J}\right)^{1/4}(b + b^\dagger) \quad \text{and} \quad n = -i\frac{1}{\sqrt{2}}\left(\frac{E_J}{8E_C}\right)^{1/4}(b - b^\dagger) \tag{1.40}$$

where the coefficients have to be chosen to match physical units.

The Transmon operates in a regime of small charge energy compared to Josephson energy in order to achieve low sensitivity to noise, while still retaining enough anharmonicity to allow selectively addressing transitions. This is made possible by the fact that charge noise is exponentially suppressed as E_J/E_C increases, while the anharmonicity scales with $\sqrt{E_J/E_C}$ [16].

This consideration allows for the cosine to be expanded to fourth order

$$\cos\phi = 1 - \frac{1}{2}\phi^2 + \frac{1}{24}\phi^4 + O(\phi^6) \tag{1.41}$$

and simplifying the Hamiltonian while eliminating the bias n_g with a transformation $U = \exp\{-in_g\phi\}$ to

$$\begin{aligned} H_{\text{qubit}} &= \sqrt{\frac{E_C E_J}{2}}\left[(b+b^\dagger)^2 - (b-b^\dagger)^2\right] - E_J - \frac{E_C}{12}(b+b^\dagger)^4 \\ &= \sqrt{8E_C E_J}\left(b^\dagger b + \frac{1}{2}\right) - E_J - \frac{E_C}{12}(b+b^\dagger)^4 \end{aligned} \tag{1.42}$$

The expression for the coupling of the qubit to the cavity

$$H_{\text{coupling}} = 2\beta e\sqrt{\frac{\omega}{2C_r}}n(a+a^\dagger) \tag{1.43}$$

can be simplified by noting that the operator n only acts on neighboring levels. Then the calculation of the matrix element

$$\left|\langle m+1|\, n\, |m\rangle_q\right| = \sqrt{\frac{m+1}{2}}\left(\frac{E_J}{8E_C}\right)^{1/4} \tag{1.44}$$

simplifies the interaction term to

$$\begin{aligned} H_{\text{coupling}} &= \sum_{m,m'} |m\rangle\langle m|_q\, 2\beta e\sqrt{\frac{\omega}{2C_c}}n\, |m'\rangle\langle m'|_q\, (a+a^\dagger) \\ &= \sum_m g_m\Big(\, |m+1\rangle\langle m|_q + |m\rangle\langle m+1|_q\,\Big)(a+a^\dagger) \end{aligned} \tag{1.45}$$

with the coupling constant

$$g_m = \beta e\sqrt{m}\sqrt{\frac{\omega}{2C_c}}\left(\frac{E_J}{8E_C}\right)^{1/4} = g\sqrt{m} \quad\Rightarrow\quad g_m^2 = mg \tag{1.46}$$

Scaling of the coupling with the square root of the photon number can be absorbed in the bosonic operator

$$b^\dagger = \sum_m \sqrt{m+1}\, |m+1\rangle\langle m|_q \tag{1.47}$$

and the coupling takes the familiar form

$$H_{\text{coupling}} = g(b+b^\dagger)(a+a^\dagger). \tag{1.48}$$

Controlling the Transmon system

The Transmon Hamiltonian can be more generally expressed in terms of a qubit frequency ω_q and the anharmonicity $\alpha_{m,m+1}$ between the m-th and the next level. These three parts (dropping the $\frac{1}{2}$ constant zero point energies)

$$\begin{aligned} H_{\text{cavity}} &= \omega_r a^\dagger a \\ H_{\text{qubit}} &= \omega_q b^\dagger b - \alpha_{12} \left|f\right\rangle\!\left\langle f\right| - \alpha_{23} \left|3_q\right\rangle\!\left\langle 3_q\right| - \alpha_{34} \left|4_q\right\rangle\!\left\langle 4_q\right| \ldots \\ H_{\text{coupling}} &= g(b + b^\dagger)(a + a^\dagger) \end{aligned} \tag{1.49}$$

add up to the full system Hamiltonian

$$H_0 = H_{\text{cavity}} + H_{\text{qubit}} + H_{\text{coupling}} \quad . \tag{1.50}$$

The cavity is driven classically by a control field $\Omega(t)$ so that the full system is

$$H = H_0 + \Omega(t) \left(a + a^\dagger\right). \tag{1.51}$$

The relatively small anharmonicity presents a challenge in the control of the Transmon. Using the appropriate control function $\Omega(t)$, we want to activate specific transitions to implement e.g., a quantum gate or population transfer. A single excitation from the ground $|g\rangle$ to the excited state $|e\rangle$ is driven by a resonant tone at the frequency ω_q. Since we act with a bosonic operator, we also need to consider the effect this drive has on the transition from $|e\rangle$ to $|f\rangle$ which is just one anharmonicity α_{ef} detuned.

1.3.3 DRAG – Derivative removal by adiabatic gate

We will one solution to the problem from the previous section along the procedure shown in [18]. Consider a 3-level-system that is controlled by a signal $u(t) = u_x(t)\cos(\omega_d t) + u_y(t)\sin(\omega_d t)$. The first two levels make up the computational subspace $|g\rangle, |e\rangle$ with transition frequency ω_q that we want to operate in and $|f\rangle$ accounts for the leakage. It is modeled by the Hamiltonian

$$H/\hbar = \omega_q \left|e\right\rangle\!\left\langle e\right| + (2\omega_q + \Delta) \left|f\right\rangle\!\left\langle f\right| + u(t)\hat{\sigma}^x_{g,e} + \lambda u(t)\hat{\sigma}^x_{e,f} \tag{1.52}$$

where the Pauli operators are $\hat{\sigma}^x_{j,k} = |j\rangle\langle k| + |k\rangle\langle j|$ and λ describes the coupling of the drive to the 1-2 transition. We expressed the frequency of the leakage state with the anharmonicity $\Delta = \omega_f - 2\omega_q$.

This introduction also appears as a section in "An introduction into optimal control for quantum technologies", Lecture notes for the 51st IFF Spring School. It was mainly written by N. Wittler.

Let's say we want to implement a simple Rabi pulse by choosing $u_x(t) = \Omega(t)$ and $u_y(t) = 0$. This gives rise to unwanted leakage out of the computational subspace with the term $\lambda\Omega(t)\hat{\sigma}^x_{e,f}$. The DRAG idea shows how we can counteract this leakage by choosing $u_y(t)$ appropriately.

We first express the Hamiltonian in the rotating frame with $R = \exp(i\omega_d |e\rangle\langle e| + 2i\omega_d |f\rangle\langle f|)$ following the rule $H^R = RHR^\dagger + i\hbar\dot{R}R^\dagger$ which gives

$$H^R/\hbar = \delta_1 |e\rangle\langle e| + \delta_2 |f\rangle\langle f| + \sum_{\alpha=x,y} \frac{u_\alpha}{2}(t)\hat{\sigma}^\alpha_{g,e} + \lambda\frac{u_\alpha}{2}(t)\hat{\sigma}^\alpha_{e,f} \,,$$

using the detunings $\delta_1 = \omega_q - \omega_d$ and $\delta_2 = \Delta + 2\delta_1$ between the drive and transition frequencies.

Applying an adiabatic transformation $V(t)$ by calculating $H^V = VHV^\dagger + i\hbar\dot{V}V^\dagger$ allows us to look at the system in a frame where the leakage and the y-component necessary to counteract it are visible. We take

$$V(t) = \exp\left[-i\frac{u_x(t)}{2\Delta}(\hat{\sigma}^y_{g,e} + \lambda\hat{\sigma}^y_{e,f})\right],$$

a transformation that depends on our intended signal u_x, and apply it to first order in u_x/Δ to find

$$\begin{aligned} H^V/\hbar = &\left(\delta_1 - \frac{(\lambda^2-4)u_x^2}{4\Delta}\right)|e\rangle\langle e| + \left(\delta_2 + \frac{(\lambda^2+2)u_x^2}{4\Delta}\right)|f\rangle\langle f| \\ &+ \frac{u_x}{2}\hat{\sigma}^x_{g,e} + \lambda\frac{u_x^2}{8\Delta}\hat{\sigma}^x_{g,f} + \left[\frac{u_y}{2} + \frac{\dot{u}_x}{2\Delta}\right](\hat{\sigma}^y_{g,e} + \lambda\hat{\sigma}^y_{e,f}) \end{aligned}$$

From this expression, we can see that our intended drive $u_x/2\hat{\sigma}^x_{g,e}$ is unchanged but if we also choose $u_y = -\dot{u}_x/\Delta$, we cancel the last term that is responsible for driving out of the computational subspace $\propto \lambda\hat{\sigma}^y_{e,f}$. The transformation also suggests detuning the drive by $\delta_1 = (\lambda^2 - 4)u_x^2/4\Delta$ to avoid stark shifting of the g-e transition.

1.4 Quantum optimal control

As a field, quantum optimal control is a subset that contains the more mathematical, theoretical apparatus of classical optimal control and approaches engineering with its most practical applications. As outlined before, this is where it neatly meets the demands of the NISQ era.

Formally, a (quantum) optimal control problem is the following:

- A dynamical system, i.e., a differential equation:

$$i\partial_t |\Psi\rangle = H |\Psi\rangle \tag{1.53}$$

- that is dependent on a control function $u(t)$, $H \equiv H(u(t), t)$
- and a cost functional $J \equiv J[|\Psi\rangle, u, t]$

The task is to find $u(t)$, such that J is maximal. Then, u is called the *optimal* control. In designing the functional J, we express the desired target.

1.4.1 Single spin example

A simple example to bring these concepts to life is the state transfer problem of a single spin. We write the Hamiltonian as

$$H = \frac{\omega}{2}\sigma_z + A\cos(\omega t)\sigma_x \tag{1.54}$$

where ω is the frequency of the qubit and A is the control amplitude of an AC drive on resonance with the qubit. We have taken the ansatz that A = const. and let the system start in the "0" eigenstate, $|\Psi(0)\rangle = |0\rangle$. To move the system to the excited state at the time T, we write the goal function as $J = |\langle 1|\Psi(T)\rangle|^2$.

Moving to the rotating frame with the qubit frequency, we can solve the simple Hamiltonian

$$H = \frac{A}{2}\sigma_x \tag{1.55}$$

in the rotating wave approximation. We can express the time-dependent state via a unitary transformation $|\Psi(t)\rangle = U(t)\,|\Psi(0)\rangle$, and obtain a differential equation for U

$$i\partial_t U(t) = HU(t)\,, \tag{1.56}$$

which has the formal solution

$$U(t) = \exp(-iHt) = \exp\left(-\frac{i}{2}A\sigma_x t\right) = \mathbb{1}\cos\left(\frac{A}{2}t\right) - i\sigma_x \sin\left(\frac{A}{2}t\right) \tag{1.57}$$

since H is not time-dependent. So, we see from

$$J = |\langle 1|\Psi(T)\rangle|^2 = |\langle 1|\,U(T)\,|0\rangle|^2 = \left|\sin\left(\frac{A}{2}T\right)\right|^2 \tag{1.58}$$

that for $AT = \pi$, J is at a maximum. In fact, J is maximal at $At = n\pi$ for $n \in \mathbb{N}$. This ambiguity can be resolved practically by tailoring the functional to prefer shorter times. In applied optimal control, particularly solutions found using numerics can often exhibit this behavior. To avoid this, care must be taken in analyzing the dynamics of the solution and adjusting the functional, if needed. It should also be pointed out, that by taking an ansatz of $u(t) = A\cos(\omega t)$, we have limited the possible solutions. In the next section, we'll look at a useful theorem that allows us to tackle more complicated problems, but also show that for the spin control problem, this particular choice is in fact optimal.

1.4.2 The Pontryagin Maximum Principle

Similar to canonical mechanics, where we define the action as a functional that produces physical trajectories, the Pontryagin maximum principle [19] can be used for statements about the optimality of controls.

From the formal definition in Eq. (1.53), different types of functional can be classified. In addition to the functional we used in the previous sections, a *final time cost* ϕ, we can define a running cost at some time t as $L(\Psi, u, t)$ so that a general functional has the form

$$J = \phi(\Psi(T), T) + \int_0^T L(\Psi, u, t)\mathrm{d}t \ . \tag{1.59}$$

We also define a *costate* $\langle\chi|$ that represents a backward propagated state with

$$\partial_t \langle\chi| = i \langle\chi| H \tag{1.60}$$

and with it, the so-called *control Hamiltonian* [1]

$$h(\langle\chi|, |\Psi\rangle, u) = \mathrm{Re}\ \langle\chi|H(u)|\Psi\rangle - L(\Psi, u) \ . \tag{1.61}$$

Now the statement of the theorem is that u is the optimal control, if it maximizes this function h for every $|\Psi\rangle$ and $\langle\chi|$.

By following this recipe, starting from a general Hamiltonian $H = \frac{\omega}{2}\sigma_z + u_x(t)\sigma_x + u_y(t)\sigma_y$ and enforcing limited power, i.e., $L = u_x^2(t) + u_y^2(t)$ it can be shown that the control fields are related by $\partial_t u_y(t) = k u_x(t)$ with some constant k. Thus, the optimal solutions are of the form $u_{x,y}(t) = A\cos(\omega t + \varphi)$ that we used as an ansatz in the simple example.

1.4.3 Gradient methods

When numerically optimizing functionals like Eq. (1.59), access to the gradient with respect to controls enables the use of powerful, efficient algorithms, e.g., L-BFGS [20]. There are several methods to obtain the gradient of a quantum time-evolution analytically, given some assumptions on the functional. A particularly useful concept here is the Shirmer derivative, also known by other names [21], that allows to compute the gradient of an exponential by a simple trick:

Given a matrix $A \equiv A(p, t)$ that is dependent on some parameter p and the time t, the gradient of its exponential $\partial_p e^A$ can be found alongside the exponential of A itself with

$$\exp\left[\begin{pmatrix} A & \partial_p A \\ 0 & A \end{pmatrix}\right] = \begin{pmatrix} e^A & \partial_p e^A \\ 0 & e^A \end{pmatrix} \tag{1.62}$$

[1] For the sake of readability, we use the $\mathbb{C}$-valued symbols of quantum mechanics. Some concepts shown here are however only defined for real-valued functions and quantities. The complex symbols can be mapped to a real field with extended dimension to more rigorously state these theorems. Details are e.g., in [19].

as opposed to approximating the gradient.

A method like GRAPE [10] exploits this relation in the setting of piece-wise constant controls, discretizing the time in slices $0 \leq t_j = j\Delta t \leq T$, such that the Hamiltonian is constant during one slice:

$$H(t_j) = H_0 + u_j H_c \tag{1.63}$$

Thus, we can expand the gradient of a state functional, writing $\partial_j := \partial_{u_j}$ for short:

$$\partial_j J = \partial_j [\, \langle\psi_T|U|\psi_0\rangle] = \partial_j \left[\langle\psi_T| \prod_{j=0}^{N} U_j |\psi_0\rangle \right] \tag{1.64}$$

Since only the j-th propagator depends on the control u_j, we collect $|\psi_k\rangle = \prod_{j=0}^{k} U_k$ and $\langle\chi_k|$ analogously, with $\langle\chi_N| = \langle\psi_T|$ and write

$$\partial_j J = \left\langle \chi_{j+1} \middle| \partial_j U_j \middle| \psi_{j-1} \right\rangle \ . \tag{1.65}$$

For each t_j, the gradient $\partial_j U_j$ can be found with Eq. (1.62):

$$\exp\left[-i \begin{pmatrix} H & H_c \\ 0 & H \end{pmatrix} \Delta t\right] = \begin{pmatrix} U & \partial_j U \\ 0 & U \end{pmatrix} \tag{1.66}$$

Note that the second row here is just the solution for the incremental propagator U at time slice j and the first can be obtained by taking the derivative of the Schrödinger equation with respect to u_j.

In GOAT [22] this approach is extended to analytic controls, i.e., controls that are not piece-wise constant. Since a single parameter changes the control field in more than a single time slice, Eq. (1.65) is not applicable. Instead, we can write a coupled system of equations for U and each derivative $\partial_p U$ and solve this system

$$i\partial_t \begin{pmatrix} \partial_p U \\ U \end{pmatrix} = \begin{pmatrix} H & \partial_p H \\ 0 & H \end{pmatrix} \begin{pmatrix} \partial_p U \\ U \end{pmatrix} \tag{1.67}$$

to obtain a time evolution of the gradients. Finally, we express the derivative of the functional to synthesize a gate V

$$J = \left| \frac{1}{\dim U} \operatorname{tr}\left(V^\dagger U\right) \right|^2 \tag{1.68}$$

in terms of $\partial_p U$ as

$$\partial_p J = -\operatorname{Re}\left(\frac{J^*}{|J|} \frac{1}{\dim U} \operatorname{tr}\left(V^\dagger \partial_p U\right) \right). \tag{1.69}$$

1.4.4 Automatic differentiation

A powerful concept borrowed from the machine learning world is the idea of computing derivatives numerically in a manner similar to how we first learn about the chain rule. As an

example, we borrow the introduction from [23] and look at the dependency of the resonance of a tunable Transmon with respect to the flux. The frequency is given by

$$\omega(\Phi) = \left(\omega^0 - \delta\right)\sqrt{\left|\cos\left(\frac{\Phi}{\Phi_0}\pi\right)\right|} + \delta \tag{1.70}$$

and we can compute the sensitivity to the flux, if we limit ourselves to the region where the cosine is positive with

$$\begin{aligned}\partial_\Phi\omega &= \left(\omega^0 - \delta\right)\partial_\Phi\sqrt{\cos\left(\frac{\Phi}{\Phi_0}\pi\right)} = \left(\omega^0 - \delta\right)\frac{\partial_\Phi\cos\left(\frac{\Phi}{\Phi_0}\pi\right)}{\sqrt{\cos\left(\frac{\Phi}{\Phi_0}\pi\right)}} \\ &= -\left(\omega^0 - \delta\right)\frac{\sin\left(\frac{\Phi}{\Phi_0}\pi\right)}{\sqrt{\cos\left(\frac{\Phi}{\Phi_0}\pi\right)}}\frac{\pi}{\Phi_0}\end{aligned} \tag{1.71}$$

To treat the same problem with automatic differentiation means introducing placeholder variables for each individual operation, so

$$\omega = x_1 = \left(\omega^0 - \delta\right)x_2 \quad , \quad x_2 = \sqrt{x_3} \quad , \quad x_3 = \cos\left(\frac{x_4}{\Phi_0}\pi\right) \quad \text{and} \quad x_4 = \Phi\,. \tag{1.72}$$

Each variable depends on new variables with increasing index until the final variable coincides with the desired derivative direction. To represent the example we carried out by hand, we can write generally

$$\frac{\partial\omega}{\partial\Phi} = \frac{\partial x_1}{\partial x_2}\,\frac{\partial x_2}{\partial x_3}\,\frac{\partial x_3}{\partial x_4} \tag{1.73}$$

By parsing and storing intermediate variables in this way, we can fully automatic compute the desired derivative, referring to a generic rule-set for elementary operations, such as multiplication, root-taking, etc. Since this is not a computer algebra system but sits in between numerics and CAS, this allows to take some derivatives otherwise not so easily expressed. In the example above, the automatic version works even when we keep the absolute. As the computation branches with the sign change in the cosine, so do the intermediate representations, resulting in the correct numerical derivative manual intervention.

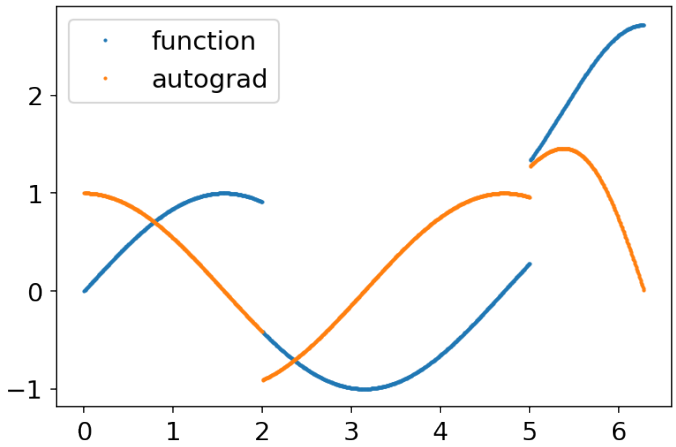

Figure 1.3: Demonstration of an automatic differentiation package. The function is piecewise defined, yet the derivative can be computed automatically.

The plot in Fig. 1.3 generated from the code

```
fig, ax = plt.subplots(1)
with tf.GradientTape() as t:
    t.watch(xs)
    ys = []
    for x in xs:
      if x < 2:
         ys.append(tf.sin(x))
      elif x > 5:
          ys.append(tf.exp(tf.cos(x)))
      else:
          ys.append(tf.cos(x))

ax.plot(xs, ys, label="function")
ax.plot(xs, t.gradient(ys, xs), label="autograd")
plt.legend()
```

to demonstrate the derivative of a mathematically problematic definition of a function.

1.5 Characterization and Benchmarking

Applying the concepts of optimal control to real-life experiments has two main problems: The dynamics will not be described by closed-system, coherent time-evolution U but the open quantum system analogue, a quantum channel. Second, every figure of merit has to be experimentally accessible. In principle, it is possible to reconstruct the full channel of a gate operation but at the price of an excessive amount of measurements.

To generalize the idea of a time-evolution, we extend the state description to the density

matrix $\rho = |\psi\rangle\langle\psi|$ and write the Liouville-von-Neumann equation

$$i\partial_t \rho = [H, \rho] \tag{1.74}$$

in place of the Schrödinger equation. In this description, we can separate $H = H_S + H_E + H_{SE}$ into a system Hamiltonian H_S (the dynamics we're interested in) and an environment H_E and their interaction H_{SE}. Then, applying some assumptions – markovianity and separating time scales – to the interaction and tracing out the degrees of freedom, we can capture the effects of the environment on the system in the Lindblad Master equation [24]

$$\partial_t \rho = -i[H, \rho] + \sum_j L_j \rho L_j^\dagger - \frac{1}{2}\left\{L_j L_j^\dagger, \rho\right\} = \mathcal{L}[\rho], \tag{1.75}$$

with the Lindbladian $\mathcal{L}$ and a set of *jump operators* L_j. The result is a quantum channel $\rho(t) = \Lambda_t[\rho(0)]$.

A more general way to express the effect of a quantum operation on the state of the system is with a linear map

$$\rho \to \frac{\Lambda(\rho)}{\mathrm{tr}\{\Lambda(\rho)\}} \tag{1.76}$$

where

$$\Lambda(\rho) = \sum_i A_i \rho A_i \tag{1.77}$$

with the Kraus operators A_i [25], that satisfy $\sum_i A_i^\dagger A_i = \mathbb{1}$. The choice of these A_i is not trivial. Practically, if a different, predetermined set of P_i is used, the channel can be written as

$$\Lambda(\rho) = \sum_{l,m} \chi_{l,m} P_l \rho P_m \tag{1.78}$$

with the *Chi matrix* $\chi_{l,m}$. Then, the entries of this matrix fully describe the quantum channel.

With this description, we can express the notion of a gate fidelity as used in the previous section. If U is some intended target gate operation, then the fidelity when comparing the actually implemented channel Λ is

$$F = \left(\mathrm{tr}\sqrt{\sqrt{\Lambda(\rho)}\mathcal{U}(\rho)\sqrt{\Lambda(\rho)}}\right)^2 \tag{1.79}$$

the typical overlap of two density matrices. We now want to characterize the gate by expressing the *average fidelity* [26]

$$\bar{F} = \int F(|\psi\rangle\langle\psi|)\mathrm{d}\psi \tag{1.80}$$

with can be compactly written with Eq. (1.78) as

$$\bar{F} = \frac{\chi_{0,0} d + 1}{d + 1}, \tag{1.81}$$

to express the fidelity of an operation on a $d \times d$ Hilbert space in general, instead of a lot of detail. In experimental application, this means significantly fewer measurements are needed to compute such a quantity.

1.5.1 Randomized Benchmarking

Often in application, the main goal is to improve gate operations without necessarily fully characterizing errors. The errors could be principally understood, but when parameters of a device drift over time, recalibration is still needed. If the procedure is also simple in operation, there is potential for automation.

Assume a single qubit of a quantum device has a set of gates $\mathcal{G}$ that we wish to characterize. We can use this gate-set to construct the so-called Clifford gates [27], a set of mainly 90 degree rotations on the Bloch sphere. Mathematically, the Clifford gates form a group which allows the following procedure:

Given a number n, randomly select n gates from the Clifford group $C_1, C_2, C_3, ..., C_n$ and assemble them to a sequence

$$C_1 C_2 C_3 \cdot .. \cdot C_n C_{n-1} \equiv \mathbb{1} \tag{1.82}$$

where C_{n+1} is chosen as the inverse of the previous sequence. By the group properties, this inverse always exists. If the gates that make up this sequence would be perfect, a measurement of

$$p_{00}(n) = |\langle 0|C_1 C_2 C_3 \cdot .. \cdot C_n C_{n+1}|0\rangle|^2 \tag{1.83}$$

would always return 1. Any imperfection in the qubit operation degrades this number increasingly with the selected sequence length n, as the errors are amplified. Additionally, this procedure has an effect called *twirling* [28] that maps any gate imperfections, including coherent errors like over- or under-rotation, to a depolarizing channel, which means the qubit approaches a fully mixed state with increasing sequence length.

To the resulting data, we can fit a curve of the form

$$p_{00}(n) = A(1-\epsilon)^n + B \tag{1.84}$$

where, for a qubit with two levels, A and B should be $1/2$ and ϵ is the *error per Clifford* [29]. The constants A and B represent respectively the state preparation error and finite temperature effects.

Even simpler and more elegantly, RB can be adapted to improve gate-set performance by picking a fixed length n and, in a feedback loop, tune-up gate parameters to maximize $p_{00}(n)$ in a procedure called ORBIT [30] (optimized randomized benchmarking for immediate tune-up).

1.5.2 Quantum Process tomography

When more detail about a quantum gate operation are of interest, a collection of tomography methods can be employed with increasing degree of effort in terms of number of measurements and processing [31]. It amounts to designing a scheme to measure every entry of the Chi matrix in Eq. (1.78). In case of a two qubit gate, we can choose the Pauli matrices as a

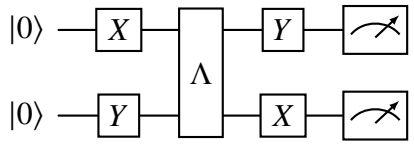

Figure 1.4: An example circuit to perform one measurement required for Quantum Process Tomography on a two qubit gate Λ.

basis, .i.e $P_l = \sigma_a \otimes \sigma_b$ and run a circuit as shown in Fig. 1.4. To reconstruct the channel, we thus need to prepare all permutations of I, X, Y and Z ($4^2 = 16$) for each qubit before and after applying the gate we want to characterize. Thus, we need to run $16^2 = 256$ circuits and we can visualize the resulting matrix as shown in Fig. 1.5.

1.5.3 Adaptation by Hybrid Optimal Control (Ad-HOC)

Transferring results from optimal control theory intro experimental practice is not an easy task. Within the optimal control community, sometimes slightly different vocabulary is used. Here, we'll take *open loop* control to mean techniques based on theoretical or numerical models, whereas *closed loop* involves communication with the experimental setup during optimization. A typical sequence consists of taking the textbook model description of an experimental system, plugging in the parameters gained from previous characterization and deriving optimal controls on this basis. This approach has been applied very successfully in NMR, e.g., [10] where models are well known and characterized to high precision. In other settings, such as superconducting qubits as complex solid state systems, the accuracy of model predictions is lower which makes the application of model bases control techniques difficult.

When performing optimal control in *closed loop*, the individual parameters of pulses are tuned mostly manually to maximize performance as indicated by some measurement result. The effort scales up the complexity of the control used, as there are more and more knobs to turn. In addition, if the model prediction is substantially different from the behavior of the real thing, the measurement might show very little sensitivity to changes in the controls.

The Ad-HOC procedure, see Fig. 1.6, connects the two steps. After a system is initially designed and fabricated, optimal control pulses are derived based on characterization measurements, such as finding qubit resonances, coupling strengths, etc. These pulses perform very well on the model, so even if there's a quantitative difference between the real experiment, the model-based solution has the right structure to work on this type of system. Then, fine-tuning the pulses with feedback from the experiment allows us to reach the best possible performance.

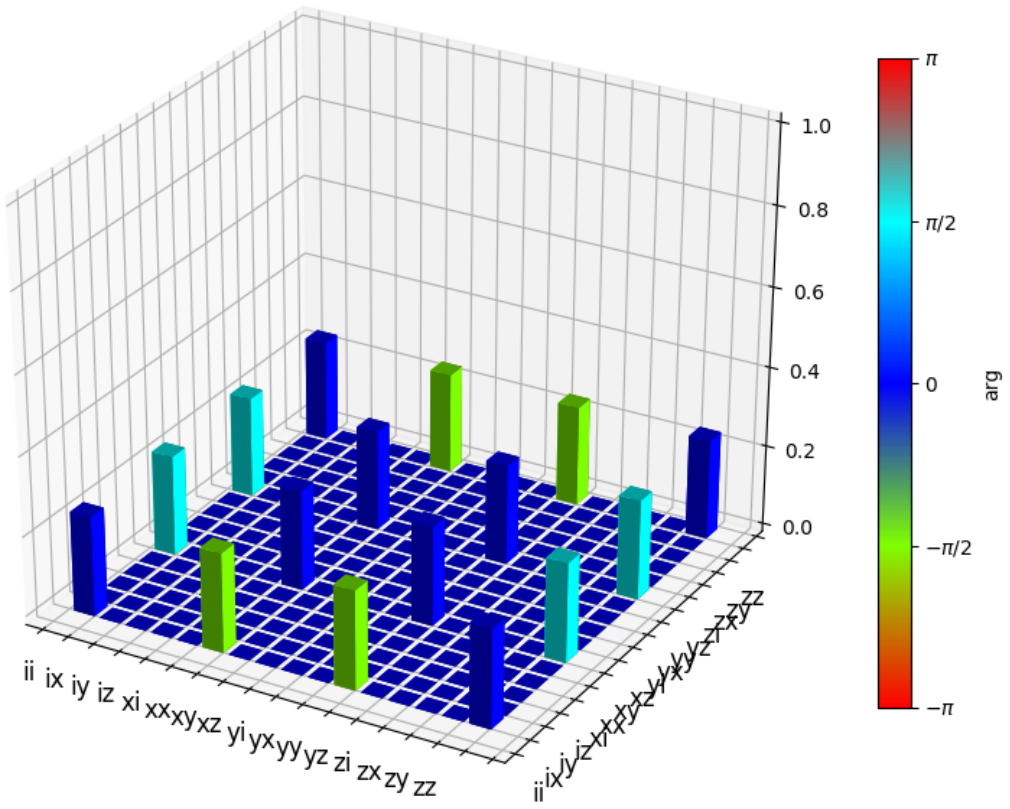

Figure 1.5: Visualization of the Chi matrix as a result of quantum process tomography experiments. Reproduced from [32].

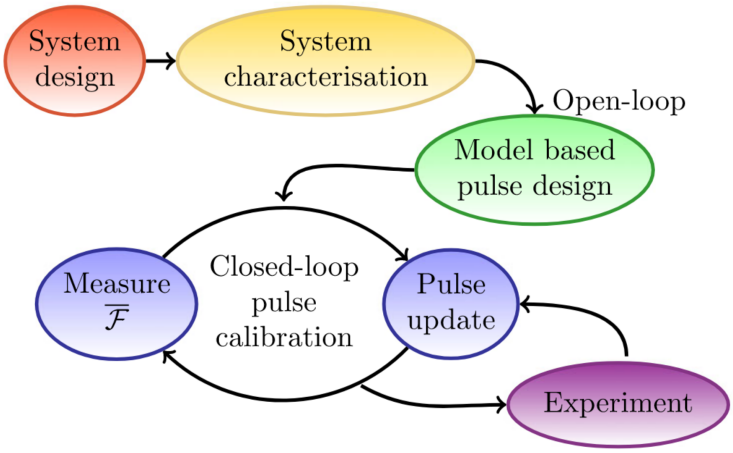

Figure 1.6: Ad-HOC protocol for finding optimal control pulses for quantum devices. After design and fabrication of the device, it is modeled based on measurable physical parameters. Based on this model, a control pulse is designed using analytic or numerical methods. This pulse is then tested on the physical system and adjusted according to the achieved fidelity. Reproduced from [33].

2 Characterizing two-photon transitions in a strongly driven Transmon

Any realistic application of quantum computing relies on having a qubit with a sufficiently long coherence time to perform a desired operation, from simple logic gates to complex quantum algorithms. Superconducting qubits [13] are usually embedded in or otherwise coupled to microwave cavities or waveguides in an architecture called Circuit Quantum Electrodynamics or cQED [34]. These resonators allow for the qubit to be controlled by external microwave fields for state generation, read-out, or coupling to other systems.

In three-dimensional circuit QED, microwave resonators are omnipresent – they are the way by which qubits are coupled to each other and to external controls. Cavity manufacturing has been technically well-developed, and it is possible to produce highly accurate 3D resonators with great coherence properties [35]. In such devices, a quantum state in the cavity can have significantly longer coherence times, than states in the qubit. It would therefore be convenient to use the cavity as a memory for the qubit: *write* and *read* operations swap the state of qubit and cavity, to store the state coherently for a comparatively long time.

Consider a computational subspace consisting of the ground and excited state of the qubit $|g\rangle, |e\rangle$ and two states of the cavity $|0\rangle, |1\rangle$. If we assume the cavity to be in the 0-state in the beginning of the process, the write operation on the combined system can be written as

$$\begin{aligned} U_{\text{write}} |e, 0\rangle &= |g, 1\rangle \\ U_{\text{write}} |g, 0\rangle &= |g, 0\rangle \end{aligned} \tag{2.1}$$

swapping the excitation between subsystems. The cavity state representing the logical 0 and 1 do not have to be the first two cavity Fock states. In fact, they don't even have to be Fock states. The only criterion is that the storage state of the cavity memory must be easily distinguished from the 'empty' state. However, in this work we will utilize the $|0\rangle$ and $|1\rangle$ Fock states, i.e., states of exactly 0 or 1 photon in the cavity.

This chapter represents an update on a section of my master's thesis. The analytic derivations have been redone so the result is more comparable to the literature. Numerical simulations have also been recoded and a comparison to analytics has been added. The introduction is reproduced here for completeness of the presentation and the text of the remaining sections rewritten. A publication is in preparation.

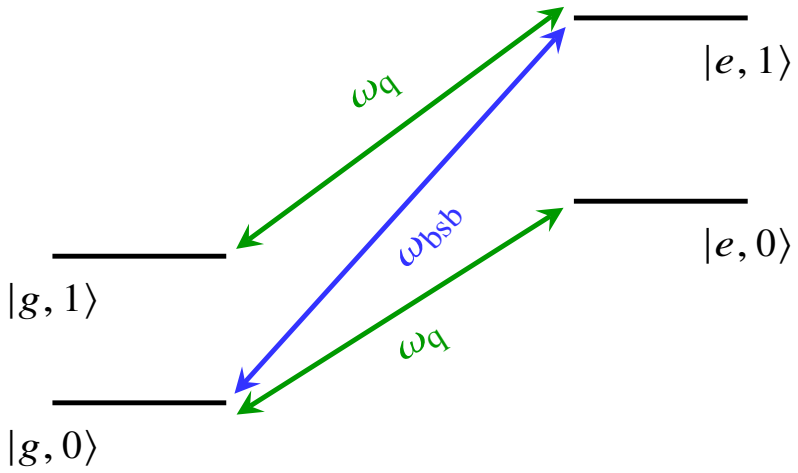

Figure 2.1: The write operation for a quantum memory, as suggested by [36], shown in an idealized computational subsystem of the participating states of the qubit $|g\rangle$ and $|e\rangle$ and the microwave cavity $|0\rangle$ and $|1\rangle$. The state of the qubit is transfered by a π-pulse on the blue sideband, followed by another π-pulse on the qubit transition.

The procedure to achieve this state transfer, suggested by [36], consists of driving a two-photon π-pulse at the so-called blue sideband transferring population from $|g,0\rangle$ to $|e,1\rangle$ and then another π-pulse on the qubit from $|e,1\rangle$ to $|g,1\rangle$ as shown in Figure Fig. 2.1.

Assuming these operations could be implemented ideally, the procedure acting on an arbitrary qubit state $\alpha|g\rangle + \beta|e\rangle$ starting from an empty cavity, results in

$$\begin{aligned} |\Psi_0\rangle &= (\alpha|g\rangle + \beta|e\rangle) \otimes |0\rangle \\ U_{\text{bsb}}|\Psi_0\rangle &= \alpha|e,1\rangle + \beta|e,0\rangle \\ U_{\text{q}}U_{\text{bsb}}|\Psi_0\rangle &= \alpha|g,1\rangle + \beta|g,0\rangle \\ &= |g\rangle \otimes (\alpha|1\rangle + \beta|0\rangle) \end{aligned} \tag{2.2}$$

where the state is indeed stored in the cavity as $\alpha|1\rangle + \beta|0\rangle$. Since the memory is not entangled with the qubit it can potentially be used for another operation, while the information of the saved qubit state rests in the cavity until it is required. Storage time is limited by the quality of the cavity, which can provide coherence times up to milliseconds [35].

2.1 Matching model with experiment

We take the parameters for the Transmon from [36] with the working transition frequency $\omega_q/2\pi$=6.234 GHz and anharmonicity $\delta/2\pi$=185 MHz that is coupled to the cavity mode

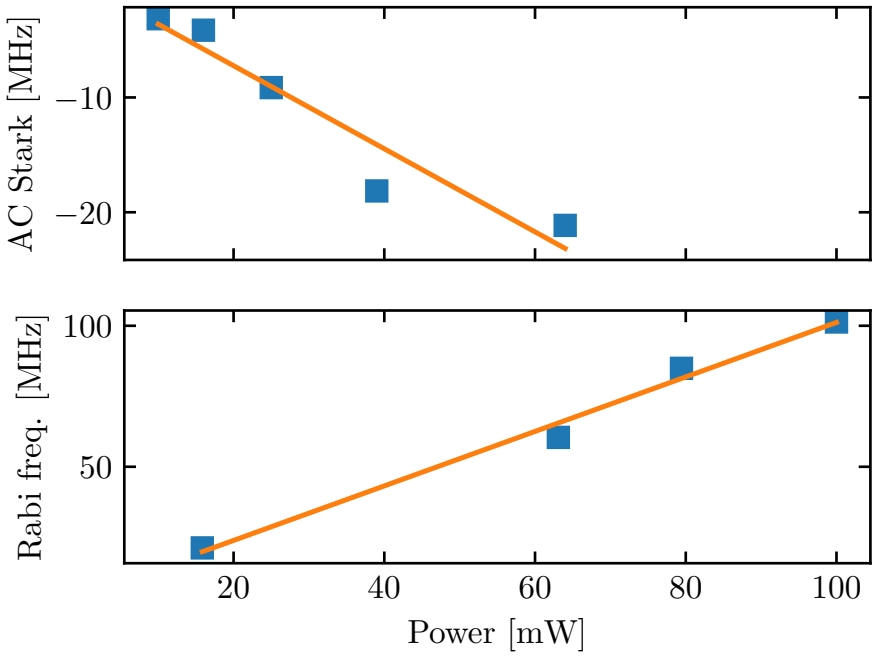

Figure 2.2: Rabi frequency and resonance for the blue sideband transition dependent on drive power. Linear fits, in agreement with theoretical predictions, are used to determine the dependency on drive power.

at $\omega_r = 8.708$ GHz with the coupling strength $g = 239$ MHz and write the Hamiltonian as

$$\begin{aligned} H^{(0)} =& \omega_r a^\dagger a + \omega_q b^\dagger b + \frac{\delta}{2}\Big(b^\dagger b - 1\Big)b^\dagger b \\ &+ g(a^\dagger + a)(b^\dagger + b) - \varepsilon(e^{-i\omega_d t}a^\dagger + e^{i\omega_d t}a)\,. \end{aligned} \tag{2.3}$$

The system is controlled by a drive on the cavity at ω_d with the amplitude ε. If we drive the cavity detuned from its resonance at ω_r, we create a coherent state. By applying a displacement transformation $D(\alpha) = \exp\{\alpha a^\dagger - \alpha^* a\}$ with the right choice of $\alpha(t)$, we can eliminate the drive on the cavity and find an effective expression for the drive on the Transmon [37]. In this description, the cavity is occupied by a mean photon number

$$n_{\text{offset}} = \frac{\varepsilon^2}{(\omega_d - \omega_r)^2} \tag{2.4}$$

and the Hamiltonian

$$\begin{aligned} H^{(1)} =& \omega_r a^\dagger a + \omega_q b^\dagger b + \frac{\delta}{2}\Big(b^\dagger b - 1\Big)b^\dagger b \\ &+ g(a^\dagger + a)(b^\dagger + b) - \Omega_R(e^{-i\omega_d t}b^\dagger + e^{i\omega_d t}b)\,. \end{aligned} \tag{2.5}$$

as a new drive term on the Transmon with Rabi frequency

$$\Omega_R = g\frac{\varepsilon}{\omega_r - \omega_d}\,. \tag{2.6}$$

For numerics, there's an advantage to simulate the system in this frame. If the cut-off is too low, the cavity in Eq. (2.3) will not behave harmonically but instead like a d-level system, where population can be reflected from the highest energy state. We expect to require a high drive amplitude to activate the two-photon transition.

The effect of drive power on the blue sideband transition is shown in Fig. 2.2. The predicted Rabi frequency for the sideband transition can be computed similarly, also from [37] as

$$\Omega_{\mathrm{BSB}} = g\frac{\Omega_R^2}{(\omega_q-\omega_d)^2} = g^3\frac{\varepsilon^2}{(\omega_r-\omega_d)^2(\omega_q-\omega_d)^2}\ . \tag{2.7}$$

It has a dependency on the square of the drive amplitude ε, which suggests a linear behavior with respect to power.

Estimating the actual drive amplitude that acts on the transition from the input power is not trivial. The total applied power of the signal that is sent to the cavity by the microwave generator is attenuated as well as shaped by different components of the experimental setup, and is then distributed across the different modes of the cavity. Instead, we can measure the Rabi frequency of a driven transition to gain insight into the effective power that arrives at the qubit.

When verifying the model, this is exactly the relation which is required to translate a theoretical drive amplitude into the power going into the experiment.

Also in Fig. 2.2, we see the AC Stark shift

$$\tilde{\omega}_q - \omega_q = \frac{1}{2}\frac{\Omega_R^2}{(\omega_q-\omega_d)} \tag{2.8}$$

of the transition frequency with increasing drive power. By combining the AC Stark shift and the Rabi frequency, we can obtain a characteristic curve for the system.

We eliminate drive power and relate the AC stark shift of the sideband transition to the Rabi frequency

$$\gamma_{\mathrm{Qubit}} := \frac{\Omega_{\mathrm{BSB}}}{(\tilde{\omega}_q-\omega_q)/2} = \frac{4g}{\omega_q-\omega_d}. \tag{2.9}$$

Two approximations are candidates for a discrepancy between model and experiment: The derivation of the sideband matrix element hinges on commutator properties of the Pauli operators describing the two-level of the qubit. Applying this result to a Transmon can require a correction. Driving a sideband requires selectivity of the qubit states at high drive power, such that the higher energy states of the Transmon will play a role, and treating the Transmon as a two-level system is not justified.

The second approximation is the model for the cavity. Describing the system in a Polaron frame obscures how strongly the cavity is actually driven.

2.2 Effective Hamiltonian for a d-level Transmon

Here, we extend the treatment detailed in [37] to include more than two levels of the Transmon, since the two-level approximation has not been sufficient to explain the behavior in seen in experiment. The anharmonicity of the Transmon does not isolate the qubit transition sufficiently from the rest of the spectrum, so leakage levels have to be considered in the derivation. We include a third energy state of the Transmon and find expressions for blue sideband Rabi frequency and Stark shift that show the effect of the presence of the third level.

Taking into account the typically low anharmonicity of Transmon circuits, strong driving of the system results in significant population of higher qubit levels. We thus extend the qubit displacement to a more general form

$$U = \exp\left\{\sum_m \beta_m^* \sigma_+^m - \beta_m \sigma_-^m\right\} = \exp\left\{\sum_m T_m\right\}. \tag{2.10}$$

where all neighboring Transmon levels are rotated by an angle β_m and we have shortened $\sigma_+^m := |m+1\rangle\langle m|$.

We apply this transformation, starting from Eq. (2.5)

$$H^{(2)} = U^\dagger H^{(1)} U - i U^\dagger \dot{U} \tag{2.11}$$

to compute a correction to the qubit subspace and provide more realistic estimates for AC-Stark shift and qubit Rabi frequency.

For the first part, we employ the Baker-Campell-Hausdorf formula

$$U^\dagger H^{(1)} U \simeq H^{(1)} + \sum_m \left[T_m, H^{(1)}\right] + \frac{1}{2}\sum_{m,n}\left[T_m, \left[T_n, H^{(1)}\right]\right] \tag{2.12}$$

where we keep terms up to $\propto \beta_m^2$.

The rotation β_n of the $\{|n+1\rangle, |n\rangle\}$ subspace is determined exactly as in [37]. The first order produces analogous expressions, this time with the energy of the corresponding Transmon transition

$$\left[\omega_q b^\dagger b + \frac{\delta}{2}\left(b^\dagger b - 1\right) b^\dagger b, \beta_n^* \sigma_+^n - \beta_n \sigma_-^n\right] = (\omega_q + n\delta)(\beta_n^* \sigma_+^n + \beta_n \sigma_-^n) \tag{2.13}$$

Again the commutator with the number expressions produce an additional coupling that will later cancel the direct drive. The drive commutator

$$\left[(b^\dagger + b), \beta_n^* \sigma_+^n - \beta_n \sigma_-^n\right] = \beta_n^*\left[b, \sigma_+^n\right] - \beta_n\left[b^\dagger, \sigma_-^n\right] = \sqrt{n}(\beta_n^* + \beta_n)\sigma_z^n \tag{2.14}$$

similarly produces $\sigma_z^n := |n\rangle\langle n| - |n+1\rangle\langle n+1|$ splitting between neighboring levels, induced by the drive.

The inertial term, we also expand to second order

$$-iU^\dagger \dot{U} \simeq \sum_m \left(\beta_m^* \dot{\beta}_m - \beta_m \dot{\beta}_m^*\right)\sigma_z^m - \dot{\beta}_m^* \sigma_+^m + \dot{\beta}_m \sigma_-^m \tag{2.15}$$

to obtain two contributions.

We compute the inertial term

$$U^\dagger \dot{U} \simeq \left(\beta_m \dot{\beta}_m^* - \beta_m^* \dot{\beta}_m\right)\sigma_z^m + \dot{T}_m \tag{2.16}$$

To determine β_m, we again collect the terms proportional to σ_+^m and find

$$(\omega_q + m\delta)\beta_m + \Omega_R \exp\{-i\omega_d t\} - i\dot{\beta}_m = 0 \tag{2.17}$$

which we can solve with

$$\beta_m = \frac{\Omega_R}{\omega_q + m\delta - \omega_d} \exp\{-i\omega_d t\} \tag{2.18}$$

So, the β_m are periodic functions with the drive frequency ω_d.

We can now identify resonances by inspecting the corresponding operators: In a rotating frame, each non-diagonal operator is dressed with a phase factor determined by the difference in eigenenergies E_m, e.g.,

$$\beta_m \sigma_+^m \to \beta_m \exp\{i(E_{m+1} - E_m)t\}\sigma_+^m \tag{2.19}$$

Hence, this transition is activated by the drive when $\omega_d = E_{m+1} - E_m$ is fulfilled.

We'll now consider the terms from the second order commutator. Applying the displacement to Eq. (2.13) gives a two-photon transition

$$\begin{aligned}\Big[(\omega_q + \delta)(\beta_n^* \sigma_+^n + \beta_n \sigma_-^n), \beta_m^* \sigma_+^m - \beta_m \sigma_-^m\Big] \\ = (\omega_q + n\delta)\Big(\beta_n^* \beta_m^* \left|n+2\right\rangle\!\left\langle n\right| - \beta_n \beta_m \left|n\right\rangle\!\left\langle n+2\right|\Big)\end{aligned} \tag{2.20}$$

resonant on a double excitation in the Transmon that will be neglected here.

The second order term that produces a sideband element is $\left[T_n, \left[T_m, b^\dagger + b\right]\right]$ from Eq. (2.14)

$$\left[(\beta_n^* + \beta_n)\sigma_z^n, \beta_m^* \sigma_+^m - \beta_m \sigma_-^m\right] = -\Big(\beta_m\left(\beta_{m+1} + \beta_{m+1}^* - \beta_m - \beta_m^*\right)\sigma_-^m + \text{h.c.}\Big) \tag{2.21}$$

since

$$\begin{aligned}&\sum_n \left(\beta_n^* + \beta_n\right)\beta_m^* \left[\sigma_z^n, \sigma_+^m\right] \\ &= \sum_n \left(\beta_n^* + \beta_n\right)\beta_m^* (\left|n\right\rangle\!\left\langle n\right| \left|m+1\right\rangle\!\left\langle m\right| - \left|m+1\right\rangle\!\left\langle m\right| \left|n\right\rangle\!\left\langle n\right|) \\ &\qquad - \left(\beta_n^* + \beta_n\right)\beta_m^* (\left|n+1\right\rangle\!\left\langle n+1\right| \left|m+1\right\rangle\!\left\langle m\right| + \left|m+1\right\rangle\!\left\langle m\right| \left|n+1\right\rangle\!\left\langle n+1\right|) \\ &= \left(\beta_{m+1}^* + \beta_{m+1}\right)\beta_m^* \left|m+1\right\rangle\!\left\langle m\right| - \left(\beta_m^* + \beta_m\right)\beta_m^* \left|m+1\right\rangle\!\left\langle m\right| \\ &\qquad - \left(\beta_m^* + \beta_m\right)\beta_m^* \left|m+1\right\rangle\!\left\langle m\right| + \left(\beta_{m-1}^* + \beta_{m-1}\right)\beta_m^* \left|m+1\right\rangle\!\left\langle m\right| \\ &= \beta_m^* (\beta_{m+1}^* + \beta_{m+1} - 2\left(\beta_m^* + \beta_m\right) + \beta_{m-1}^* + \beta_{m-1})\sigma_+^m\end{aligned} \tag{2.22}$$

and analogously for σ_-^m.

In summation, the second order Hamiltonian in the three level Transmon approximation is

$$H^{(2)} = H_0 + \sum_m \Omega_{\mathrm{BSB}}^m \left(a^\dagger + a\right)\left(\sigma_+^m \exp\{-2i\omega_d t\} + \sigma_-^m \exp\{2i\omega_d t\}\right) \tag{2.23}$$

with $H_0 = \omega_r a^\dagger a + \tilde{\omega}_q b^\dagger b + \frac{\delta}{2}(b^\dagger b - 1)b^\dagger b + \tilde{g}(a^\dagger + a)(b^\dagger + b)$.

We can compute the two-photon Rabi frequency from Eq. (2.21), collecting $\Delta_m := \omega_q + m\delta - \omega_d$

$$\Omega_{\mathrm{BSB}}^m = g\frac{\Omega_R^2}{2\Delta_m^2}\left(2 - \frac{1}{1+\delta/\Delta_m} - \frac{1}{1-\delta/\Delta_m}\right) \tag{2.24}$$

If the anharmonicity is small compared to the drive detuning, we can use the approximation on the right that is easy to compare to the qubit case in Eq. (2.8).

Here, we can appreciate the presence of higher levels in the Transmon Hamiltonian. The neighboring transitions differ by the anharmonicity δ which determines their addressability. It is also the factor that scales the Rabi frequency.

By plugging the solution for β_m into the first term in Eq. (2.15)

$$-i\left(\beta_m^*\dot{\beta}_m - \beta_m\dot{\beta}_m^*\right)\sigma_z^m = \omega_d\frac{\Omega_R^2}{\Delta_m^2}\sigma_z^m \tag{2.25}$$

and writing out the summation, we can find the Stark shifted Transmon frequencies

$$\tilde{\omega}_q^m = m\omega_q + \frac{\delta}{2}(m-1)m + \omega_d\frac{\Omega_R^2}{\Delta_m^2}\left(1 - \frac{\Delta_m^2}{(\Delta_m+\delta)^2}\right). \tag{2.26}$$

In summation, the second order Hamiltonian for two-photon transitions is $H^{(2)} = H_0 + H_d^{(2)}$

$$H_d^{(2)} = \sum_m \Omega_{\mathrm{BSB}}^m \left(a^\dagger + a\right)\left(\sigma_+^m \exp\{-2i\omega_d t\} + \sigma_-^m \exp\{2i\omega_d t\}\right) \tag{2.27}$$

with $H_0 = \omega_r a^\dagger a + \sum_m \tilde{\omega}_q^m |m\rangle\langle m| + \tilde{g}(a^\dagger + a)(b^\dagger + b)$.

Since the Transmon is weakly anharmonic, we can simplify these expressions by expanding them in δ/Δ_m. This is a commonly used correction term when applying two-level results to Transmon devices. We can also appreciate that when δ approaches the working transition frequencies, we recover the two-level expressions.

2.3 Floquet theory for the 3D-Transmon

We can try to support the perturbative results presented in the previous section with a numerical approach. In the control of quantum systems, usually a carrier signal with an

envelope that is shaped by an arbitrary waveform generator to achieve a desired pulse for a given operation. This envelope has a significantly lower bandwidth (around 500 MHz) than the local oscillator providing the drive tone (between 5 and 10 GHz). In this case, the envelope can be seen as quasi-constant during one oscillation of the carrier, and we can apply the Floquet method [38].

Applying the theory to the previous section the cavity drive takes the role of a perturbation

$$V^{(0)}(t) = \varepsilon \cos(\omega_d t)\left(a^\dagger + a\right) \tag{2.28}$$

that we can write as

$$V(t) = \Omega_R \cos(\omega_d t)\left(b^\dagger + b\right) \tag{2.29}$$

acting on the qubit. The benefit of this frame lies in expressing the indirect drive through the coupling with the cavity as an effective drive on the qubit. Also, as mentioned before, this provides a numerical advantage, since accurate simulations require a smaller Hilbert space.

The Floquet Hamiltonian

$$\mathcal{H}_F = \mathcal{H}^{(n-m)} + n\omega_d \delta_{n,m}. \tag{2.30}$$

consists of two types of blocks: For $n = 0$ the integral averages out the drive,

$$\begin{aligned} \mathcal{H}^{(0)} &= \frac{1}{T}\int_0^T \mathrm{d}t H(t) \\ &= H_0 + \varepsilon(a^\dagger + a)\frac{1}{T}\int_0^T \mathrm{d}t \cos(\omega_d t) = H_0 \end{aligned} \tag{2.31}$$

and the diagonal blocks of $\mathcal{H}_F = H_0 + n\omega_d$ are the unperturbed system Hamiltonian offset by the quasi-energy of the drive.

For $n - m = \pm 1$ only the drive will remain as a constant coupling

$$\begin{aligned} \mathcal{H}^{(\pm 1)} &= \frac{1}{T}\int_0^T \mathrm{d}t H(t) \\ &= \frac{1}{T}\Omega_R(b^\dagger + b)\int_0^T \mathrm{d}t \cos(\omega_d t)\exp\left(\pm \mathrm{i}\omega_d t\right) \\ &= \frac{\Omega_R}{2}(b^\dagger + b) =: \mathcal{V} \end{aligned} \tag{2.32}$$

as in a frame rotating with the drive frequency.

In total, the block structure of the Floquet Hamiltonian is

$$\mathcal{H}_{\mathcal{F}} = \begin{pmatrix} \ddots & \mathcal{V} & & & \\ \mathcal{V} & \boxed{H_0 - \omega_d} & \mathcal{V} & & \\ & \mathcal{V} & \boxed{H_0} & \mathcal{V} & \\ & & \mathcal{V} & \boxed{H_0 + \omega_d} & \mathcal{V} \\ & & & \mathcal{V} & \ddots \end{pmatrix}. \tag{2.33}$$

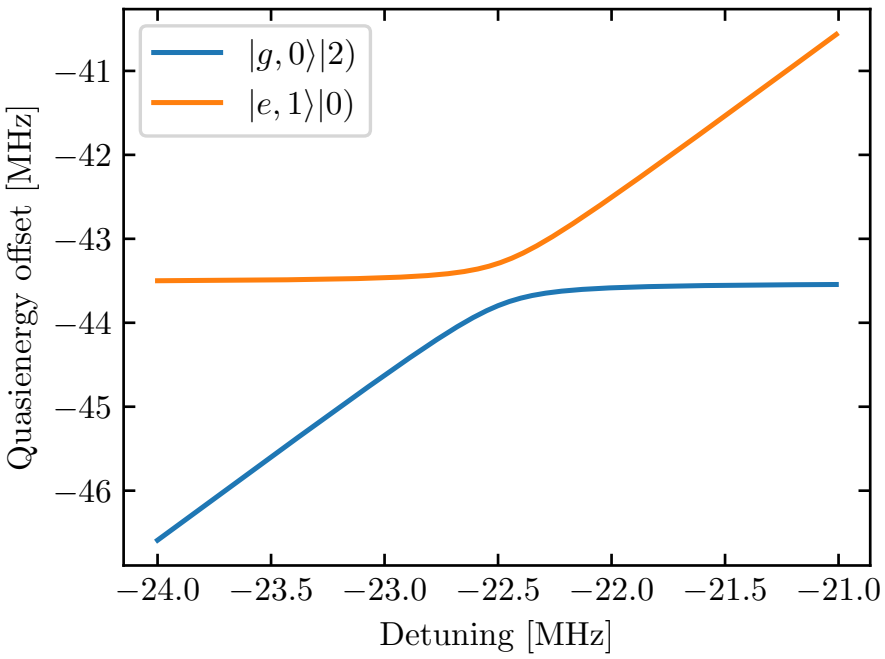

Figure 2.3: Eigenenergies calculated by the Floquet method for variable drive frequency with the detuning measured from half the bare blue sideband frequency. Labeling of the $|e1\rangle$ and $|g0\rangle$ states is chosen for left of the transition, where $|g0\rangle$ has a slope of $\sim 2\omega_d$.

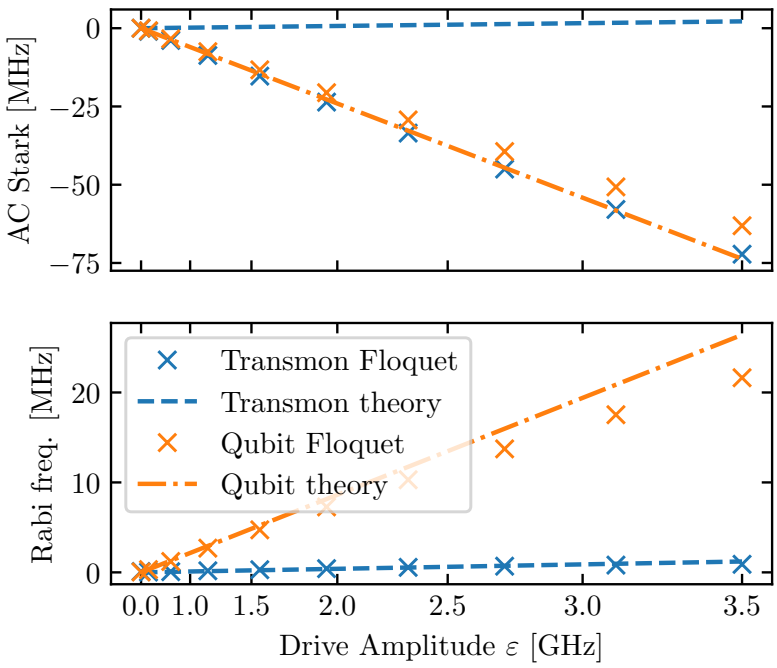

Figure 2.4: Comparison between the analytic results and Floquet simulation. With respect to the AC Stark shift, both systems behave the same, suggesting an effective separation of the targeted transitions. The extracted Rabi frequency however agrees with the theory prediction.

The two-photon sidebands, for example, connect blocks that differ by $2\omega_d$ in energy.

We simulate the Floquet Hamiltonian for a given drive strength, sweeping the drive frequency around the bare blue sideband transition at $\omega_{\mathrm{BSB}}/2 = (\omega_q + \omega_r)/2$ and compute the eigenvalues and eigenbasis. To characterize the transition, we identify those states with a significant overlap to those that interest us, namely $|e1\rangle$ and $|g0\rangle$. From the avoided crossing in Fig. 2.3, we can estimate both the resonance position and the effective Rabi frequency. Since the drive acts as a static coupling between the Floquet blocks, the strength of the interaction is given by the distance at the avoided crossing.

Now we compare the approximate sideband Rabi frequency from simulated avoided crossings like Fig. 2.3 with experimental data, to estimate of the relation between drive amplitude ε used in the model and power of the microwave source generating the carrier signal. In Fig. 2.4, we compare the two-level "Qubit" results with the correction "Transmon" from the previous section. Interestingly, the prediction for the Rabi frequency aligns well with the perturbative theory while the dynamic AC Stark shift behaves the same in both cases.

Using control theory, e.g. DRAG (see Section 1.3.3), it should be possible to drive the Transmon in such a way that it behaves like a two-level system, so either expression might be applicable.

When finally comparing to the experimental data in Fig. 2.5, we find a good agreement if

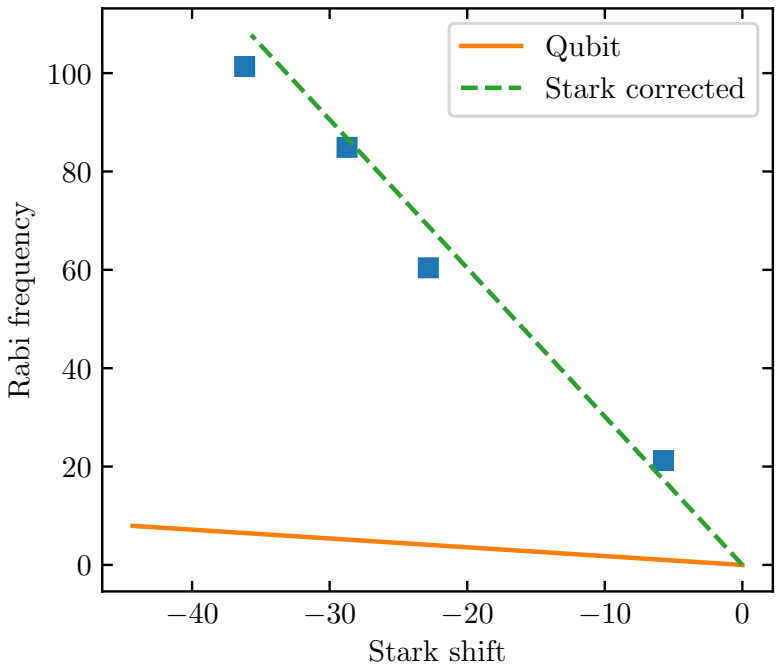

Figure 2.5: Comparing analytic expressions to the experimental data. Orange is the line for qubit expressions. If we add the Transmon correction δ/Δ_m to the AC Stark shift, it agrees with the experimental (green, dashed). However, this requires an unphysically high drive amplitude $\varepsilon \gg 1$ GHz to reach reported gate times under 100 ns. In this case, assumptions made during the derivation of Eq. (2.7) are not justified.

we apply the Transmon correction only to the Stark shift expression and use the Qubit Rabi frequency, contrary to the Floquet numerics.

2.4 Conclusion

In this chapter, we have highlighted a common problem in applied optimal control. Some systems, particularly human-made ones, are described with approximate models that serve well in qualitatively understanding their behavior.

Obtaining quantitatively matching expressions from theory for these kinds of systems has been a challenge for a while. Recently there are promising approaches [39] to deal with this kind of characterization problem.

Resolving the discrepancy between Floquet simulation, analytic results and experimental data is subject of future work. One clue is the estimated drive strength required in the simulation which is way into the strong driving regime. This is somewhat expected, looking at Eq. (2.7), both detunings are over 1 GHz and the coupling strength goes with the third power. So obtaining a Rabi frequency of around 100 ns results in a quite high amplitude ε. Since the analytic results are based on series expansions that require the ratio of ε and these detunings to be small , we don't expect them to be fully trustworthy, but only work as a guideline.

In order to apply sophisticated optimal control, the Ad-HOC (see Section 1.5.3) method requires somewhat comparable performance in model and experiment. We have also seen in this case, that the system characterization needs to be consistent. The different efforts using analytics and numerical Floquet analysis do not agree here. Consequently, a more systematic approach is desirable.

3 Integrated tool-set for control, calibration, and characterization of quantum devices applied to superconducting qubits

System characterization is often a manual, laborious task. Reliably teasing out the properties of a physical system is a fundamental task in understanding and describing its behavior. In application, this is less a matter of groundbreaking, novel technique but rather a means to an end: The construction of a working device. Efforts to scale-up quantum computation have reached a point where the principal limiting factor is not the number of qubits, but the entangling gate infidelity. However, the highly detailed system characterization required to understand the underlying error sources is an arduous process and impractical with increasing chip size. Open-loop optimal control techniques allow for the improvement of gates but are limited by the models they are based on. To rectify the situation, we provide an integrated open-source tool-set for Control, Calibration and Characterization ($\mathbb{C}^3$), capable of open-loop pulse optimization, model-free calibration, model fitting and refinement. We present a methodology to combine these tools to find a quantitatively accurate system model, high-fidelity gates and an approximate error budget, all based on a high-performance, feature-rich simulator. We illustrate our methods using simulated fixed-frequency superconducting qubits for which we learn model parameters with less than 1% error and derive a coherence limited cross-resonance (CR) gate that achieves 99.6% fidelity without need for calibration.

This chapter has been published as "Nicolas Wittler, Federico Roy, Kevin Pack, Max Werninghaus, Anurag Saha Roy, Daniel J. Egger, Stefan Filipp, Frank K. Wilhelm, and Shai Machnes, *Integrated Tool Set for Control, Calibration, and Characterization of Quantum Devices Applied to Superconducting Qubits* Phys. Rev. Applied **15**, 034080" by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. `https://doi.org/10.1103/PhysRevApplied.15.034080`
Unless stated otherwise, text and software were written in close collaboration by F. Roy and N. Wittler. Section 3.4.3 and Section 3.5 were authored by S. Machnes. The publication contains an appendix on gradient-free optimization methods that was written by K. Pack.

3.1 The problem

Scaling up quantum processing units (QPUs) is a monumental task, that requires the community to make progress on multiple fronts, most importantly improving gate fidelities and increasing the number of qubits. Over the past few years, significant emphasis has been placed on creating larger devices, yielding great success [14]. However, the number of qubits has outstripped the limits that fidelity places on their utility: In [40], a record quantum volume [41] of 64 was demonstrated, loosely translating to the device being able to perform $\log_2(64)^2 = 36$ entangling gates before fidelity drops below $2/3$, a relatively small number of gates for an array of six qubits; In [42] the circuit fidelity was 0.1% thus requiring 30 million repetitions to achieve the desired statistics. One could even argue that the two-qubit gate fidelities demonstrated in isolation in 2014 [43] (0.994) are comparable with those in 2019's [42] (0.9938), even though the latter are for simultaneous gates in a large 2D qubit array.

The relatively slow progress in improving gate fidelities can be traced back to an incomplete understanding of the sources of error. Indeed, characterization and calibration of QPUs to the desired accuracy is impractical and cumbersome, and operating on devices of increasing qubit number requires entangling gates to be fine-tuned for each individual pair to account for slightly varying properties. The resulting lack of detailed models makes it harder to identify where efforts must be focused to achieve higher fidelity gates [44, 45].

Given that "all models are wrong, but some are useful" [46], we describe a Good Model as follows:

> A *Good Model* is one that predicts the behavior of the system, for the operations we wish to perform, to accuracies we care about.

For a QPU, a Good Model has to have predictive power for the range of feasible gate-generating pulses and for long sequences of such gates, to a fidelity accuracy of the order of 10^{-5}. To the authors' knowledge, no such Good Model for a superconducting QPU has ever been published.

Since models serve as the basis to derive high-fidelity gates in open-loop optimal control [47–54], any inaccuracies of the model will inevitably degrade the experimental accuracy of the resulting gates. This problem is only partially ameliorated by the first-order insensitivity of optimized pulses to model inaccuracies [55, 56]. Methodologies such as the adaptive hybrid optimal control (Ad-HOC) protocol [57] – which combines a model-based open-loop optimization with a closed-loop experimental calibration [51, 58] – address this issue but leave one in an unsatisfactory position as the need for calibration proves the inadequacy of the model: the root causes of the remaining infidelities are unexplained.

Conversely, if a Good Model is known, gates generated by open-loop optimal control will, by definition, work on the experiment, not requiring further closed-loop calibration. This enables the use of complex pulses that would otherwise require time-consuming calibration.

Such a Good Model would also provide an error budget through a process of exploratory interrogation – evaluating the potential performance of the system where certain limitations have been removed, i.e. asking "what if ...?". Therefore, extracting a Good Model efficiently and in a highly automated manner is key to improving fidelities and a crucial step of QPU scale-up.

In this work we present $\mathbb{C}^3$, our proposed approach to control, calibrate and characterize QPUs. This chapter is structured as follows: We present the conceptual steps of $\mathbb{C}^3$ in Sec. 3.2 and illustrate the methodology by example in Sec. 3.3, showing how these steps are implemented. Sec. 3.4 includes a detailed description of the modeling, optimization procedures, the data comparison function and relevant prior work. We conclude in 3.7 with a discussion of the effort's current status and long-term directions.

3.2 $\mathbb{C}^3$ – Control, Calibration, and Characterization

Current methodology relies on tailored routines to extract individual parameters of the system's model (characterization) or fine-tune specific parameters of pulses used (calibration) [59, 60]. This approach becomes cumbersome and impractical as the number of model and pulse parameters increases. With $\mathbb{C}^3$ we propose a different paradigm: Optimizing a figure of merit that is sensitive to the set of parameters we care about. This eliminates the need to design per-parameter measurements, and thus provides a more general approach. $\mathbb{C}^3$ at its core is composed of three separate optimizations, respectively implementing the tasks of control, calibration, and characterization:

C$_1$ Given a model, find the pulse shapes that maximize fidelity with a target operation. Pulse shapes may be constrained by an ansatz or allow direct arbitrary waveform generator (AWG) parameterization.

C$_2$ Given pulse shapes, calibrate their parameters, if possible simultaneously, to maximize a figure of merit measured by the actual experiment, thus improving beyond the limits of a deficient model.

C$_3$ Given control pulses and their experimental measurement outcome, optimize model parameters to best reproduce the results. Enhance the model if needed.

The tasks of open-loop optimal control, C_1, and calibration, C_2, are fairly established in the community [47–51, 51–54, 57, 58]. To characterize the system and provide us with a Good Model, we introduce C_3, a tool to optimize model parameters by comparing model prediction to experimental data. We refer to this task as model learning. For this purpose, one requires an experimental data-set containing information about the implemented pulses and the corresponding measurement outcomes. To test the model accuracy, we reproduce the data-set, applying the same pulses to a simulation of the experiment, and compare the resulting outcomes: This provides a model match score to optimize. Initially, a candidate model is formulated based on previous information or intuition. If the model is suitable to explain the experiment, the optimization will converge to a near perfect match, thus

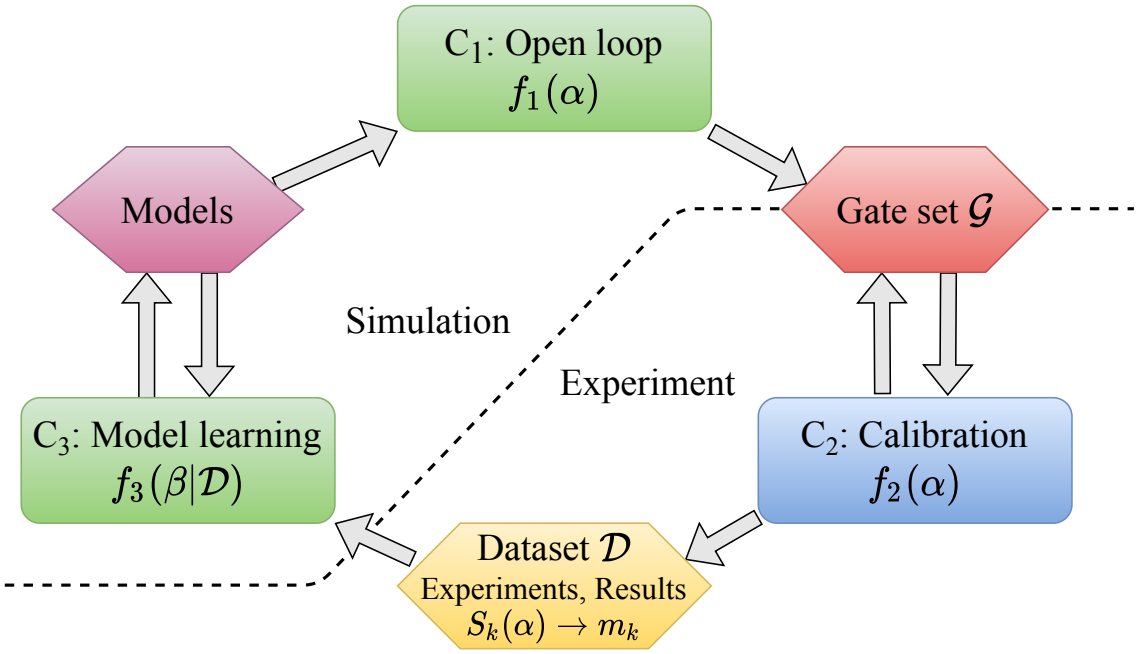

Figure 3.1: Diagram of the $\mathbb{C}^3$ tool-set in an integrated characterization loop. $\mathbf{C}_1$ is a tool for obtaining optimal pulses by finding the control parameters α that minimize a goal function $f_1(\alpha)$ in simulation. The **gate-set** $\mathcal{G}$ includes all the operation that one wishes to perform on the experiment, including the information of the ideal logical operations and the optimal pulse parameters α that implement them. $\mathbf{C}_2$ is a model-free experimental calibration procedure that optimizes pulse shapes with a gradient-free search to minimize an infidelity function $f_2(\alpha)$ by varying all parameters at once. A **data-set** is a collection of experiment/result pairs, including information about the pulses parameters used α, the sequences S_k performed and the final outcomes measured, m_k. $\mathbf{C}_3$ is a tool for model learning that determines the model parameters β that best explain the data-set. It minimizes a goal function $f_3(\beta|\mathcal{D})$ obtained by recreating experiments $S_k(\alpha)$ in simulation and comparing the results to the ones in the experiment. In $\mathbb{C}^3$ different parameterized models can be provided to represent various elements of the experiment to find the one that best describes it. After the learning, the resulting model can be the basis for another characterization loop, refining both model and controls.

providing numeric values for the model parameters. Instead, if the match is poor, the user supplies a new model, that is either an extension or modification of the previous candidate, and the optimization is repeated. Depending on user choice, learned values are carried over to the parameters of the new model or discarded.

As heterogeneous experimental data is the foundation for model learning, we suggest using the three tasks of $\mathbb{C}^3$ in sequence, as shown in Fig. 3.1. However, their application is by no means limited to this use case and one may choose to view them as stand-alone routines. The same tools used to realize C_1, C_2 and C_3 can also be used to further interrogate the system to obtain a sensitivity analysis of the optimized model in light of the experimental data and a breakdown of possible error sources.

We note that the intertwining of control and characterization has been raised in the more general context of control theory [61–64]. In quantum technology, there are some works which combine two of the three tasks: Ad-HOC [57] calls for optimal control followed by calibration; a combination of model-based gradient calculations and experimental calibrations is proposed by [65, 66], but the data gathered is not used to improve the system model; in [67] pulses are designed specifically for the purpose of reconstructing the noise spectrum.

3.3 Synthetic Application Example

The following synthetic example illustrates how $\mathbb{C}^3$ is used to obtain a Good Model in a realistic setting. We simulate a two-qubit QPU device using an underlying model, labeled the "real" model, which includes control discretization effects, electronics transfer functions, Markovian noise, and state preparation and measurement (SPAM) errors.

In this example, the simulated device is treated as a black-box, which we interrogate with $\mathbb{C}^3$. We derive (C_1) and calibrate (C_2) optimal control pulses and use the resulting data to extract a Good Model (C_3) by comparing the black-box to three candidates: We systematically enrich the model until it reproduces the behavior of the device observed in C_2. The recovered model is then used to design a two-qubit gate that performs well on the black-box without the need for tune-up.

3.3.1 The black-box device ("real" model)

The "real" model is composed of two coupled three-level Duffing oscillators, labeled by A and B, each directly driven by an external field $c_i(t)$. Initialization, dynamics and readout are performed in the dressed basis by solving the master equation in Lindblad form [68, 69].

$$\dot{\rho} = -i[H, \rho] + \sum_{\substack{i=A,B \\ j=\phi,\kappa}} L_{i,j}\rho L_{i,j}^{\dagger} - \frac{1}{2}\left\{L_{i,j}L_{i,j}^{\dagger}, \rho\right\} \tag{3.1}$$

with

$$H/\hbar = \sum_{i=A,B} \left[\omega_i b_i^\dagger b_i - \frac{\delta_i}{2}\left(b_i^\dagger b_i - 1\right) b_i^\dagger b_i \right] + \\ + g(b_A + b_A^\dagger)(b_B + b_B^\dagger) + \sum_{i=A,B} c_i(t)\left(b_i + b_i^\dagger\right), \tag{3.2}$$

where ω_i is the frequency of qubit i, δ_i is the anharmonicity, b_i ($b_i^\dagger$) is the lowering (raising) operator, and g is the coupling strength. Open-system effects are expressed by the dephasing and relaxation Lindblad operators $L_{i,\phi} = \sqrt{\frac{2}{T_i^{2*}}} b_i b_i^\dagger$ and $L_{i,\kappa} = \sqrt{\frac{1}{T_i^1}} b_i$ with decay rates $1/T_i^1$ and $1/T_i^{2*}$.
Given the input drive signals $\varepsilon_i(t)$, we calculate the effective control fields $c_i(t) = \varphi[\varepsilon_i(t)]$, where the transfer function φ [70] accounts for discretization introduced by the AWG, bandwidth limitations of hardware, and for a constant scaling φ_0, which translates input voltages to field amplitudes. We implement state preparation errors due to a non-zero initial temperature T by starting each experiment from the thermal state

$$\rho_{\text{init}} = \frac{1}{Z}\left[|0\rangle\langle 0| + \exp\left\{-\frac{E_1}{k_B T}\right\} |1\rangle\langle 1| \right. \\ \left. + \exp\left\{-\frac{E_2}{k_B T}\right\} |2\rangle\langle 2| \right] \tag{3.3}$$

where $Z = \sum_{k=0}^{2} \exp\{-E_k/k_B T\}$ is the partition function with energies $E_{0,1,2} = 0, \hbar\omega_q, \hbar(2\omega_q + \delta)$, and k_B is the Boltzmann constant. readout misclassification is included, measuring state $|n\rangle$ as state $|m\rangle$ with probability $p_{n\to m}$. For example the probability of measuring a state $\rho_\psi = |\psi\rangle\langle\psi|$ as $|0\rangle\langle 0|$ is

$$\Pi_0(\rho_\psi) = p_{0\to 0}\ \langle 0|\rho_\psi|0\rangle + \ p_{1\to 0}\ \langle 1|\rho_\psi|1\rangle \\ + \ p_{2\to 0}\ \langle 2|\rho_\psi|2\rangle\,. \tag{3.4}$$

Similarly to experiment, populations are estimated by averaging the results of multiple projective measurements, simulated as a multinomial draw from the distribution with probabilities $\{\Pi_n\}$, thus introducing noise stemming from a finite number of measurement repetitions (commonly known as 'shot noise'). The values of the "real" model parameters are summarized in Tab. 3.1.

3.3.2 Open-loop Optimal Control, $\mathbf{C}_1$

We assume that at the start of the $\mathbb{C}^3$ procedure the parameters of the system are only known to a rough precision, with its qubit frequencies and anharmonicities chosen to be within a few MHz of their "true" values. In the simple model, the qubits are *uncoupled* three-level Duffing oscillators, evolution follows *closed* systems dynamics, and state preparation and

Parameter	Real	Simple		Intermediate		Full (ORBIT)		ORBIT+QPT		Decoherence	
Learning	Model	Before	After	Before	After	Before	After	Before	After	Before	After
$\omega^{(A)}$ (MHz)	5000	−1.000	−0.886	−1.000	−0.230	−0.230	+0.004	+0.004	−0.016	"	"
$\delta^{(A)}$ (MHz)	210	+1.000	+0.702	+1.000	+0.281	+1.000	+0.400	+0.4008	+0.017	"	"
$\omega^{(B)}$ (MHz)	5600	+1.000	+0.592	+1.000	+0.013	+0.013	−0.003	−0.003	+0.006	"	"
$\delta^{(B)}$ (MHz)	240	+1.000	+0.981	+1.000	+4.32	+1.000	−0.016	−0.016	−0.005	"	"
φ_0 (MHz/V)	159.2	+1.592	−1.634	+1.592	−0.802	−0.802	+0.123	+0.123	+0.246	"	"
g (MHz)	20	–	–	+1.000	−0.665	−0.665	+0.046	+0.046	−0.119	"	"
T (mK)	50	–	–	–	–	+5.000	−3.172	−3.172	−0.216	"	"
$T_1^{(A)}$ (μs)	27	–	–	–	–	+4.000	+4.439	+4.439	+0.021	+0.021	+0.738
$T_{2*}^{(A)}$ (μs)	39	–	–	–	–	+2.000	+1.994	+1.994	−2.353	−2.353	−0.020
$T_1^{(B)}$ (μs)	23	–	–	–	–	+3.000	+4.543	+4.543	+5.704	+5.704	+0.666
$T_{2*}^{(B)}$ (μs)	31	–	–	–	–	+5.000	+6.183	+6.183	+4.716	+4.716	+0.897
$p_{0\to0}^{(A)}$ (%)	97	–	–	–	–	−2.00	−0.84	−0.84	−0.11	"	"
$p_{1\to1}^{(A)}$ (%)	96	–	–	–	–	+0.20	+0.38	+0.38	+0.02	"	"
$p_{0\to0}^{(B)}$ (%)	96	–	–	–	–	−2.00	−0.62	−0.62	−0.03	"	"
$p_{1\to1}^{(B)}$ (%)	95	–	–	–	–	+0.20	+0.08	+0.08	+0.01	"	"

Table 3.1: Overview of the parameters of the "real" model (reference values), and the candidate models, before and after the C_3 learning for different data-sets. Candidate model values are shown as difference from reference values. Dashes (–) indicate parameters not present in the model, quotation marks (") indicate parameters not being changed.

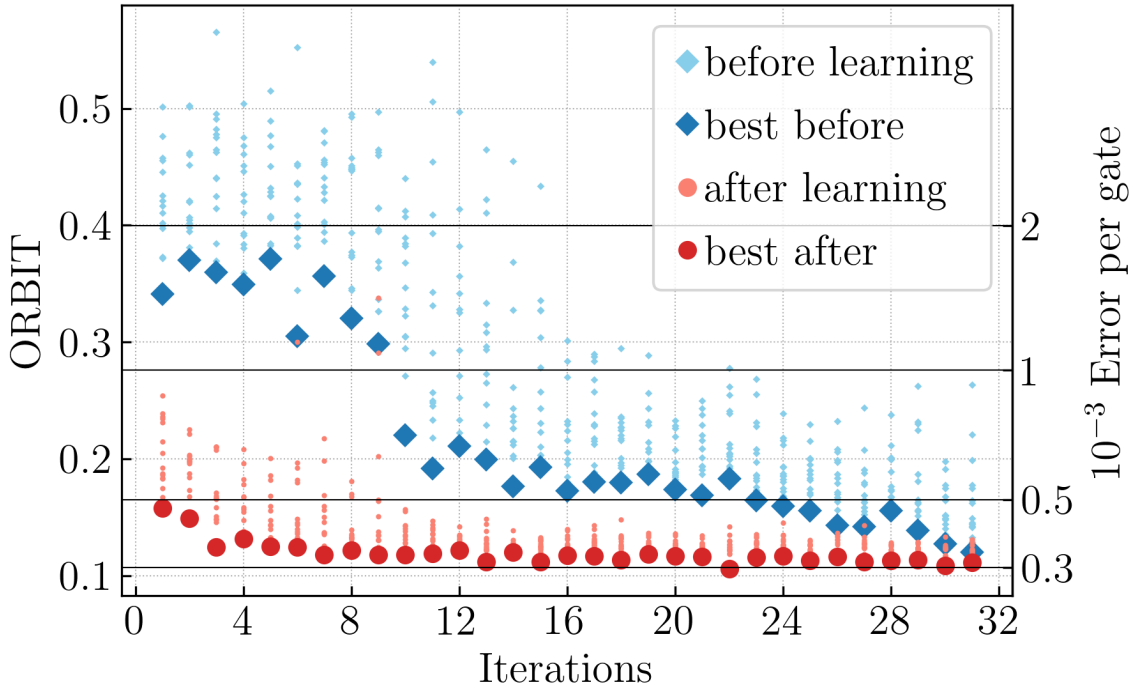

Figure 3.2: C_2 calibration on the device for single-qubit gates of qubit B. The initial point is suggested by C_1 before (after) learning of the model. The light blue diamonds (light red circles) represent the values of the ORBIT goal function, Eq. (3.9), for varying pulse parameters α as chosen by the search algorithm. The larger blue diamonds (larger red circles) highlight the best of 25 points generated and sampled at each iteration. In experimental practice, this batching helps reduce the overhead of loading pulses in AWG programming [71]. Both calibrations achieve the same final fidelity, however the optimal gates derived from the learned model provide a better initial guess. Assuming no SPAM errors the ORBIT value can be translated into an error per gate, indicated on the right axis. This is only meant to provide a rough estimate of the performance of the gate, noting that an ORBIT value of 0.5 represents maximum error per gate, i.e. completely depolarizing channels.

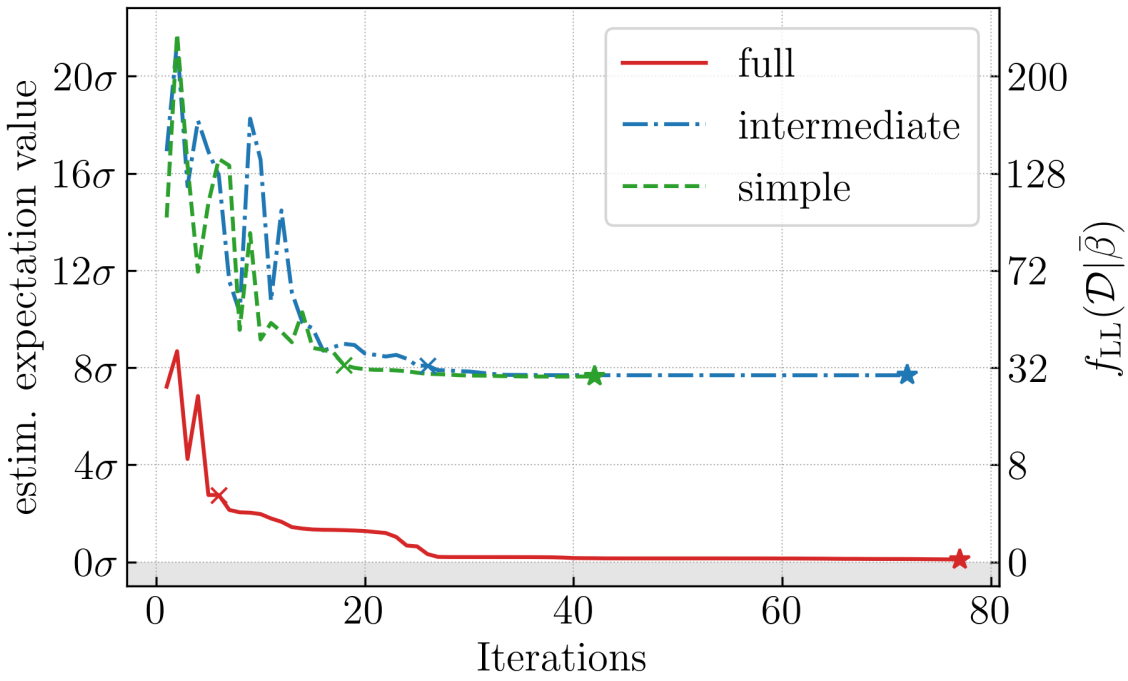

Figure 3.3: Progress of the C_3 optimization on a hierarchy of models: Simple model (green, dashed), intermediate model (blue, dot-dashed) and full model (red, solid), as described in the text. The Model match goal function $f_{\mathrm{LL}}(\beta)$ is defined in Eq. (3.11). The crosses show the switch-over from CMA-ES to L-BFGS. The CMA-ES algorithm evaluates a batch of points for each iteration (8, 9 and 12 for the simple, intermediate and full model respectively), only the best of each batch is shown. The L-BFGS algorithm takes on average approximately 1.2 evaluations per iteration, for all three models. The function f_{LL} is rescaled to express the match in terms of standard deviations of the binomial distribution that the experimental results are drawn from. The simple model is a close dispersive approximation of the intermediate model, demonstrated by their similar final match score. By including all relevant device properties the full model reaches an almost perfect match score.

measurement are assumed perfect. The Hamiltonian is

$$\begin{aligned} H/\hbar = &\sum_{i=A,B} \omega_i b_i^\dagger b_i - \frac{\delta_i}{2}\Big(b_i^\dagger b_i - 1\Big) b_i^\dagger b_i \\ &+ c_i(t)\Big(b_i + b_i^\dagger\Big). \end{aligned} \tag{3.5}$$

Assuming this model, we design pulses for single-qubit gates using C_1. To mitigate leakage, we choose a pulse ansatz with a Gaussian shape and a correction given by the derivative removal by adiabatic gate (DRAG) method [53],

$$\begin{aligned} \varepsilon(t) = &A\,\Omega_{\text{Gauss}}(t)\,\cos\big((\omega_d + \omega_{\text{off}})t + \phi_{xy}\big) \\ &- \frac{\eta}{\delta} A\,\dot{\Omega}_{\text{Gauss}}(t)\,\sin\big((\omega_d + \omega_{\text{off}})t + \phi_{xy}\big). \end{aligned} \tag{3.6}$$

Here, Ω_{Gauss} is a Gaussian envelope, $\dot{\Omega}_{\text{Gauss}}(t)$ is its time derivative, A is the amplitude of the drive, ω_{off} is a frequency offset and the DRAG parameter η can be adjusted to reduce leakage into the second excited state [53, 72]. The rotation axis can be freely chosen in the x-y plane by changing the phase of the drive signal $\omega_d t \rightarrow \omega_d t + \phi_{xy}$, implementing the unitary rotations $R(\phi_{xy}, \theta) = \exp\big\{-i(\cos\phi_{xy}\sigma_x + \sin\phi_{xy}\sigma_y)\theta\big\}$. By setting $\phi_{xy} = n\frac{\pi}{2}$ with $n = 0, 1, 2, 3$ and changing $\alpha = (A, \eta, \omega_{\text{off}})$ we aim to realize the single qubit gate-set

$$\mathcal{G} = \big\{X_{\pi/2}, Y_{\pi/2}, X_{-\pi/2}, Y_{-\pi/2}\big\}\,, \tag{3.7}$$

for each qubit separately, eight gates in total, where $X_{\pi/2} = \{R(0, \pi/2)\}$ and so on. With C_1 we use a gradient-descent method to find the parameters α that minimize the mean average gate infidelity

$$f_1(\alpha) = 1 - \frac{1}{|\mathcal{G}|}\sum_{U\in\mathcal{G}} f_{\text{av}}(U) = 1 - \frac{1}{|\mathcal{G}|}\sum_{U\in\mathcal{G}} \frac{\chi_{0,0}d + 1}{d + 1}, \tag{3.8}$$

where, $\chi_{0,0}$ is the $(0,0)$-th element of the Chi matrix representation of the gate error $U^\dagger \circ \tilde{U}(\alpha)$ between the ideal gate U and the implemented gate $\tilde{U}(\alpha)$ [73]. We optimize Gaussian pulses with a gate length of $t_g = 7$ ns, for both qubits, using the gradient-based L-BFGS algorithm [74]. The obtained optimal pulses yield a mean infidelity of $f_1(\alpha) = 6.6 \times 10^{-4}$ and $f_1(\alpha) = 4.9 \times 10^{-4}$ on the simple model for qubit A and qubit B respectively, realistic values for fast gates using this simple parametrization. Next, we compare the performance of these pulses on the black-box device, where the gates instead yield a mean infidelity of 2.4×10^{-3} for qubit A and 1.5×10^{-3} for qubit B. In fact, performing an experimentally realistic randomized benchmarking (RB) [75–78] measurement on the device yields an error per gate of 2.3×10^{-3} and 1.3×10^{-3} comparable with the theoretical average infidelity. The degradation of performance from optimal control simulation ($\approx 10^{-4}$) to experiment ($\approx 10^{-3}$) shows a clear mismatch between the device and the simple model.

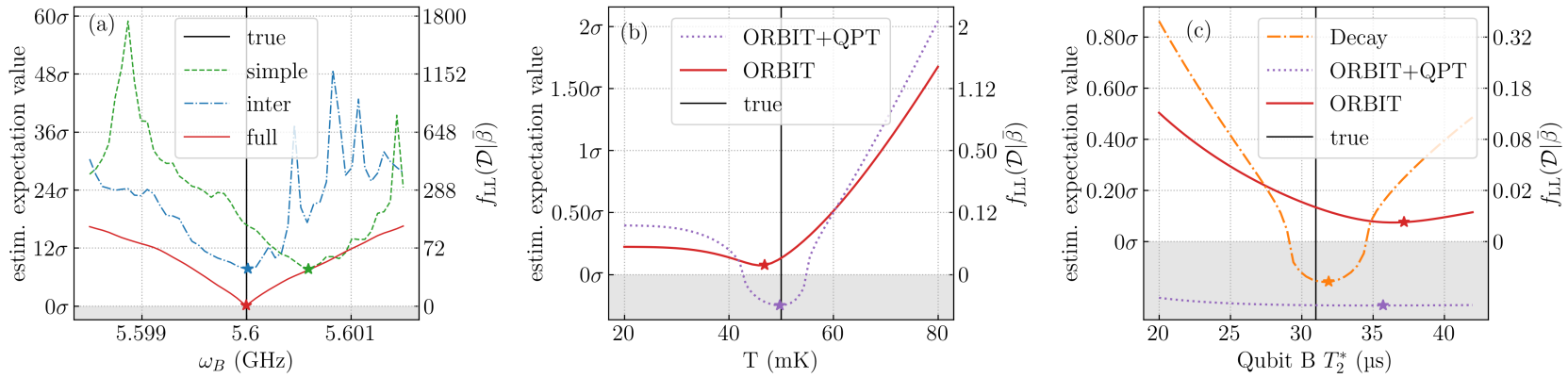

Figure 3.4: Sensitivity analysis of selected parameters for different models and data sets. In (a) we sweep the qubit frequency ω_B and evaluate the goal function f_{LL} on the learning data at each value. The star represents the optimal value returned by C_3. Intermediate (blue, dot-dashed) and Full (red, solid) models, show the same frequency value, while the value of the Simple model (green, dashed) is dispersively shifted, as expected. In (b) and (c) we perform the same sweep for the chip temperature and T_2^* of qubit B respectively, evaluated on the full model for different learning data-sets. The ORBIT data is the same used in the full model in (a). Introducing the Quantum Process Tomography (QPT) data (purple, dotted), allows a more precise definition of the temperature. To determine T_2^* we use relaxation and dephasing data (orange, dot-dashed). Match values below 0 can occur because of noisy data, finite sampling and deviation from the assumption of Gaussian distribution of the data.

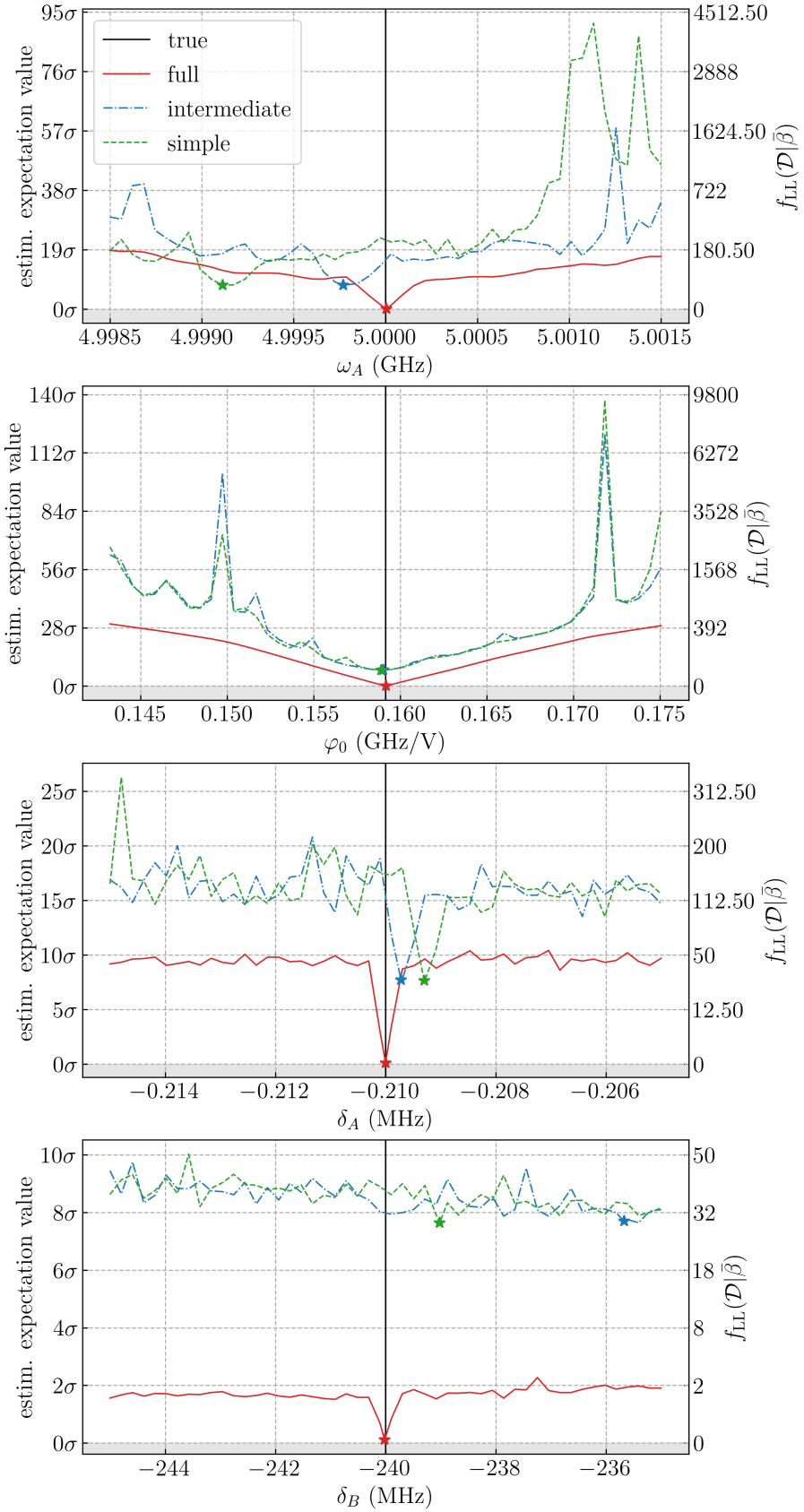

Figure 3.5: We compare the sensitivity of the hierarchy of models with respect to their Qubit frequency ω_A, field conversion factor φ_0 and the anharmonicities δ_A and δ_B.

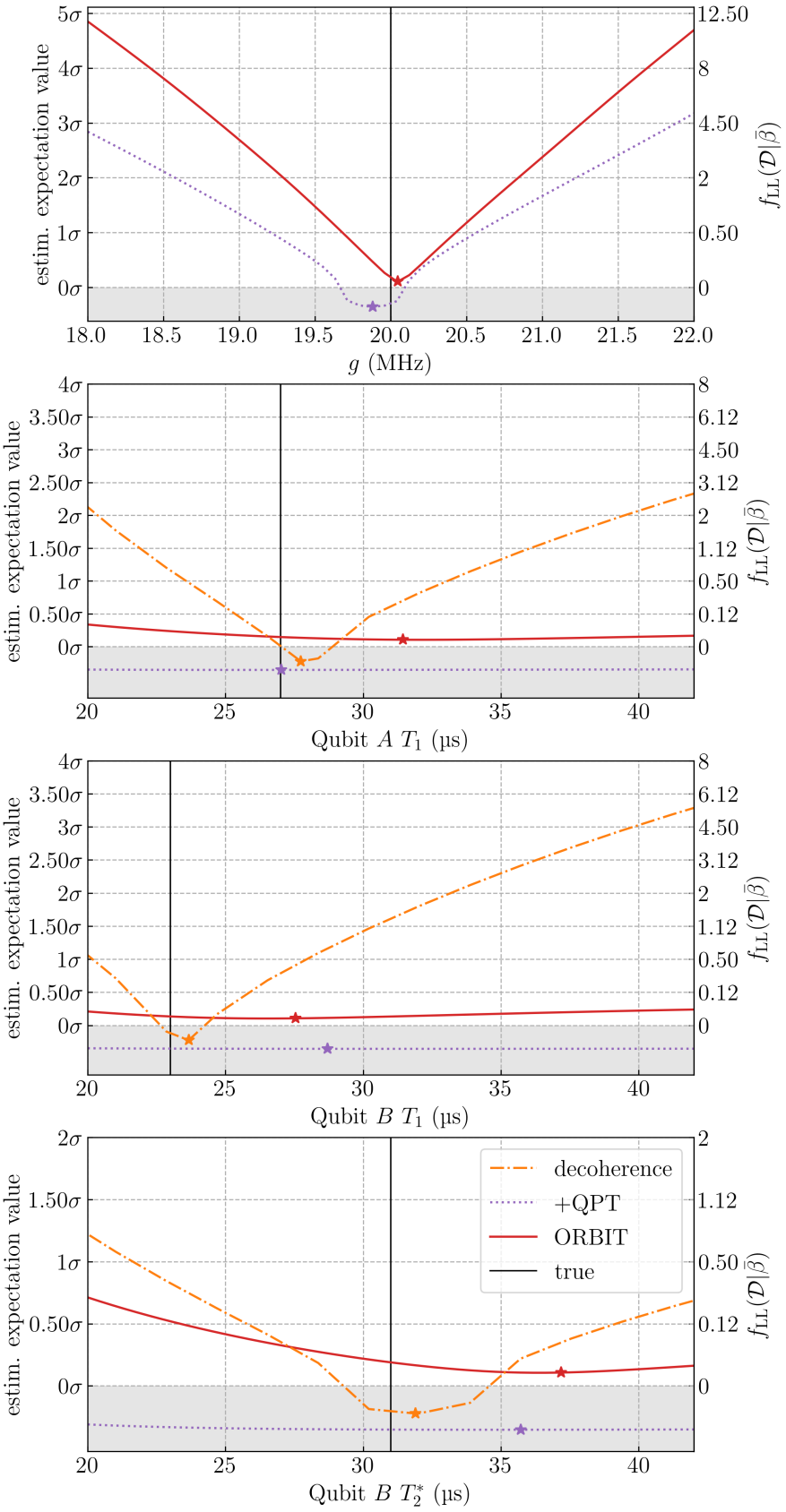

Figure 3.6: The sensitivity coupling strength g and relaxation and dephasing times T_1 and T_2^* based on different characterization methods.

3.3.3 Calibration, C_2

The next step is to calibrate the pulses derived by C_1 and improve their performance on the device. We use C_2 and employ a closed-loop, model-free, gradient-free optimization algorithm on an experimentally accessible figure of merit f_2. Since the goal is to evaluate a gate-set, we choose f_2 to be the ORBIT [58] (single-length RB) function

$$f_2(\alpha) = f_{\text{ORBIT}}(\alpha) = \frac{1}{N}\sum_{k=1}^{N}[1 - m_k(\alpha)], \tag{3.9}$$

averaging over N sequences. The survival probability, $m_k = \Pi_0\Big(S_k(\alpha)\rho_{\text{init}}S_k^\dagger(\alpha)\Big)$, is the probability to measure the state $|0\rangle$ (see Eq. 3.4) after applying random sequences

$$S_k(\alpha) := \left\{\prod_j^{L-1} C_{k,j}\right\} C_{\text{inv}} \tag{3.10}$$

composed of L Clifford gates [58], to the initial thermal state ρ_{init}. The $C_{k,j}$ are the random gates sampled from the Clifford group C (for a single qubit $|C| = 24$), and C_{inv} is chosen so that $S_k \equiv \mathcal{I}$ in the ideal case. We use the atomic operations $\mathcal{G}$ from Eq. (3.7) to construct the set of Clifford gates, e.g. $C_6 = X_{-\pi/2} \circ Y_{-\pi/2} \circ X_{\pi/2}$, and from them construct $N = 25$ RB sequences of length $L = 100$. The survival probabilities m_k are estimated by performing $s = 1000$ projective measurements and averaging.

To minimize f_2, we employ the CMA-ES [79] algorithm, a gradient-free search that samples the loss function in batches, and is fairly robust to local minima and noise [80]. See [71] for an experimental demonstration. The optimal pulse parameters from C_1 are used as the starting point of the optimization, and the parametrization is kept as in Eq. (3.6). We perform the calibration for each qubit independently, with similar results. See Fig. 3.2 for the ORBIT calibration data of qubit B. The initial point suggested by C_1 has an ORBIT infidelity of 0.50 and is improved by the optimization to 0.12. To account for SPAM errors, we perform a full RB measurement and estimate the infidelity of the gates before and after as 1.3×10^{-3} and 3.4×10^{-4} respectively. Qubit A shows a similar improvement of RB estimated error from 2.3×10^{-3} to 7.5×10^{-4}.

For the purpose of learning we define the data-set $\mathcal{D} := \big\{S_k(\alpha_j) \to m_{j,k}\big\}$, the collection of the experiments conducted during the C_2 calibration, consisting of pulse parameters α_j and gate sequences $S_k(\alpha_j)$, and the corresponding measurement outcomes $m_{j,k}$.

3.3.4 Characterization, C_3

In C_3, we use the data-set $\mathcal{D}$ obtained during ORBIT calibration to improve the model of the system. For each measurement result $m_{j,k}$ we compute the equivalent simulation result $\tilde{m}_{j,k}(\beta)$ by calculating the dynamics of the sequence $S_k(\alpha_j)$ given a set of model parameter values $\beta = (\omega_i, \delta_i, ...)$. Since simulating the whole data-set is computationally costly, for

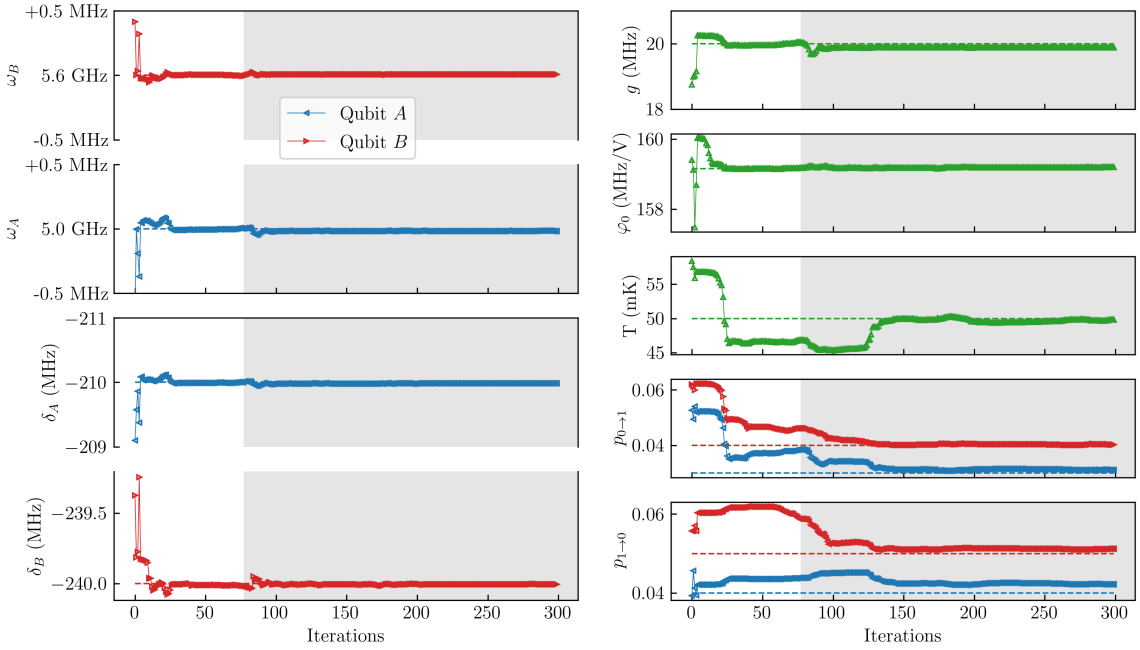

Figure 3.7: C_3 learning of the two-qubit model parameters. Blue, left (red, right) triangles indicate qubit A (B) parameters, respectively, while shared properties are shown with green upwards triangles. The true values of the "real" model are indicated as dashed lines. Learning begins using just ORBIT data (left white section) that fixes qubit frequencies, anharmonicities, coupling and line transfer function to their true values. Then tomography data from a two qubit experiment is added (right gray section), which allows better identification the chip temperature T and the misclassification constants $p_{0\to1}$, and $p_{1\to0}$.

the purpose of model learning we make a selection of eight pulse parameter sets j per qubit from the full data-set. Each parameter set includes $k = 1, .., 25$ sequences, meaning that we learn from a total of $N = 400$ measurement results, relabeled as m_n. We then construct a goal function

$$f_3(\beta) = f_{\mathrm{LL}}(\mathcal{D}|\beta) = \frac{1}{2N}\sum_{n=1}^{N}\left[\left(\frac{m_n - \tilde{m}_n}{\tilde{\sigma}_n}\right)^2 - 1\right] \tag{3.11}$$

that captures how well the model prediction $\tilde{m}_n$, with standard deviation $\tilde{\sigma}_n$, agrees with the recorded values m_n. Because of the finite number of measurements, the averaged m_n are noisy estimates of the population, with a mean μ_n and standard deviation σ_n. Thus, they cannot be matched perfectly even when all model parameters are exact. However, we can determine the expectation value of the goal function f_{LL} in the scenario where all $\tilde{m}_n$ are exactly a given number of standard deviations away from the underlying true value μ_n. A detailed mathematical discussion is presented in Sec. 3.4.3. To provide a more intuitive measure, we express the match f_{LL} in terms of the number of standard deviations that would result in the same score.

To minimize $f_3(\beta)$, we use a combination of two algorithms: Gradient-free (CMA-ES) to avoid local minima and gradient-based (L-BFGS) to converge quickly once the right minimum has been identified. Fig. 3.3 shows the convergence of the C_3 optimization for different models. The simple model is not able to reproduce the device's results, as the optimization ends at approximately 8 standard deviations away. This demonstrates that the experiment on the device includes behavior not captured by the simple model. Spectator effects might be significant even when performing only single qubit experiments, making the completely uncoupled model insufficient. Another source of this inconsistency might be SPAM errors not accounted for in the model, that might play a large role in actual measurement results. The parameter values resulting from this C_3 process and all following ones are shown in Table 3.1.

Going forward an informed decision has to be made about how to enhance the model. Since the true values of the parameters are not known in an experimental setting, we require a tool to determine the precision to which they are learned. We estimate the sensitivity to changes of model parameters around the optimal values β' by performing one-dimensional scans and observing the degradation in model match score, $f_{\mathrm{LL}}(\mathcal{D}|\beta' + \delta\beta)$. Fig. 3.4(a) shows that sweeping the value of frequency of qubit B produces a highly irregular landscape of the match score f_{LL}.

The simple model is then extended by adding the static coupling g of unknown exact value, resulting in the intermediate model. When repeating C_3, we initialize model parameters from the initial, rough values. We do not carry over the learned parameters from the simple model to the intermediate model because, by introducing a coupling, we expect slightly shifted frequencies compared to the bare frequencies of the uncoupled qubits. Nonetheless, convergence of the match score shows no improvement from the simple model, still only reaching within approximately 8 standard deviations from experiment results (Fig. 3.3) and resulting in a similar sensitivity landscape in Fig. 3.4(a). This suggests that the simple model is a close dispersive approximation of the intermediate model. Indeed, we observe a

dispersive shift [81] of 593 KHz, consistent with the expected $g^2/(\omega_B - \omega_A) \simeq 666$ KHz, given the coupling of $g \simeq 20$ MHz and the frequency difference $\omega_B - \omega_A \simeq 600$ MHz.

Finally, model complexity is increased by adding three relevant features: Markovian noise simulated by Lindblad master equation, initialization errors due to finite operating temperature and measurement errors in the form of misclassification. The system model is now of the same structure as the "real" model of the device. Starting from the best intermediate model parameters, the C_3 procedure converges satisfactorily, approaching the 0 standard deviations mark (Fig. 3.3)

In Fig. 3.7(a) we show the value of each parameter of the full model during optimization, as we introduce different learning data (in the next sections), and compare with their true value (dashed lines). By learning the model parameters with the ORBIT data (white left section of each plot) the model frequencies $\omega_{A/B}$, anharmonicities $\delta_{A/B}$, coupling g and line transfer function φ_0 converge to their true value. The temperature and misclassification parameters are not recovered, and we believe this is due to an extra degree of freedom that is not resolved by the experiments we have performed, as the effects of misclassification, Eq. (3.4), and initial thermal distribution, Eq. (3.3), are similar and can be partially exchanged. Dephasing and relaxation times (not shown) are also not recovered. Indeed, in Fig. 3.4(c) we show that the sensitivity of the data to dephasing time T_2^* of qubit B is minimal. RB sequences perform an effective random dynamical decoupling [82], providing a possible explanation to this result.

3.3.5 Validation of the learned model

After model matching on a subset of the data in the C_3 step, we now evaluate the predictive power of the learned models by computing the score on the rest of the data set (this is also known as a validation set in machine learning). This verifies that the selected subset captures all relevant behavior and alleviates the danger of overfitting.

Figures 3.8(a) to 3.8(c) depict the correlation between calibration data points $m_{j,k}$ and their model-based reconstructions $\tilde{m}_{j,k}$. We evaluate the goal function $f_{\mathrm{LL}}(\beta)$ over the validation set for the Simple, Intermediate and Full models and obtain values of 36.5 ($\approx 8.4\sigma$), 42.0 ($\approx 9.2\sigma$) and 0.028 ($\approx 0.2\sigma$) respectively. The conclusion is that, even though some parameters were not recovered by C_3, the learned full model is indeed a Good Model, as it reproduces the behavior of the system on *all previously recorded data points* to satisfying accuracy. This does not prevent additional measurement data to expose new behavior in the system: The notion of the Good Model is always tied to the underlying data-set.

Furthermore, we now repeat the C_1 procedure on the Good Model (yielding average gate infidelities of 6.3×10^{-4} and 1.1×10^{-3} for qubit A and B respectively) and show that the resulting pulses give a near optimal performance on the actual device and allow for faster C_2 convergence, as seen in Fig. 3.2. One would expect the pulses derived from the Good Model to be exactly optimal on the actual device. Even though it is not the case here, this is not because of an inaccurate model, but rather because of a disparity between the figures of

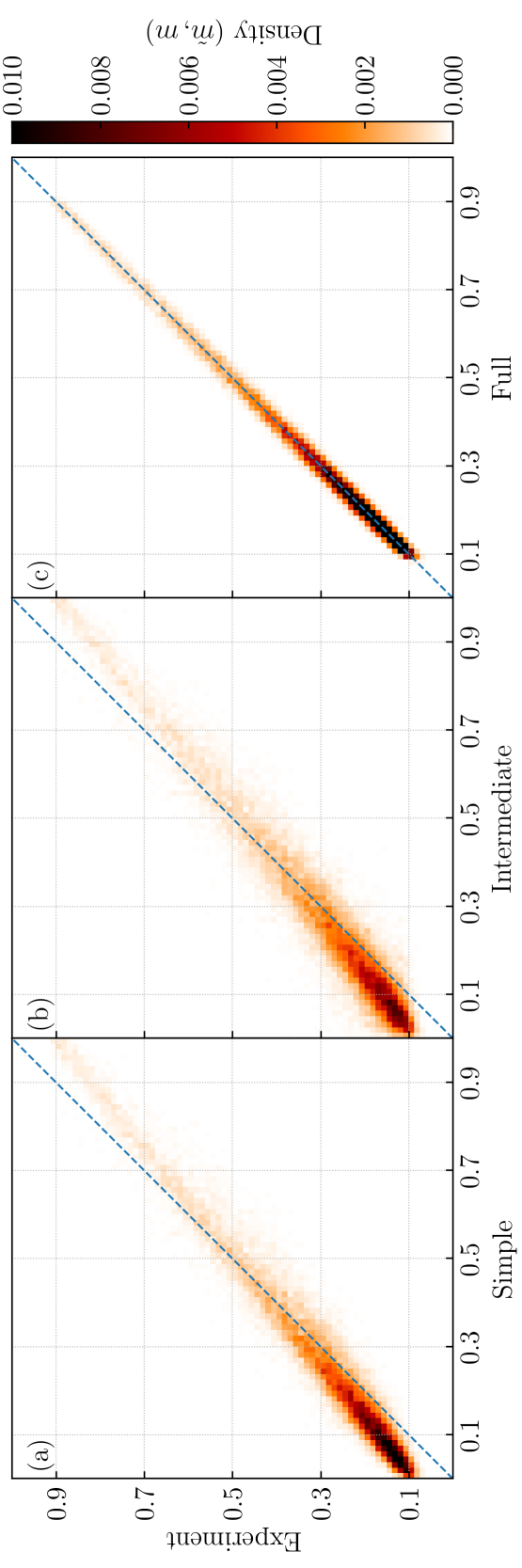

Figure 3.8: (a)-(c): Correlation between simulation and experiment expressed as a density of points $(\tilde{m}_k, m_k)$ for the Simple, Intermediate and Full model, respectively. The data-set is the so-called validation set: The data points k that were not used in optimization of the model parameters. The simple and intermediate models show poor correlation as the simulation predicts a wide distribution of measurement outcomes for each recorded value. They also exhibit a tilt that can be attributed to SPAM errors not considered in the models. Only the full model produces a consistently high density distribution centered around the diagonal, with minimal spread due to the noisy data.

merit used in C_1 (average infidelity) and C_2 (single-qubit ORBIT). Average fidelity captures effects of the whole system, including, in this case, an effective ZZ-coupling between the two qubits caused by a slight repulsion of the $|02\rangle$ and $|11\rangle$ states, that are 300 MHz apart. Minimizing a single-qubit ORBIT infidelity does not adjust for this effect, as we can verify by evaluating both RB (which captures only one qubit at a time) and average infidelity before and after calibration. Indeed, the average infidelity of qubit B is 1.2×10^{-3} (compatible with the performance of 1.1×10^{-3} on the Good Model) but the error per gate is estimated by RB as 4.1×10^{-4}. After the calibration the RB estimate is improved to 2.9×10^{-4} but the average infidelity is worsened to 1.9×10^{-3}. Performing simultaneous RB could resolve this issue.

3.3.6 Entangling gate

We further investigate the Good Model which was determined using only single-qubit calibration data by deriving a two-qubit cross-resonance (CR) gate [83, 84] with C_1. Both qubits are driven simultaneously at ω_B, the resonant frequency of qubit B, to accumulate a phase $\pm\pi/2$ conditioned on the state of qubit A [59]. Both drives are parameterized by flattop Gaussians. The resulting CR pulse has a gate infidelity of $f_{\text{av}} = 3.8 \times 10^{-3}$. When evaluated on the "real" model the gate has an infidelity of $f_{\text{av}} = 4.3 \times 10^{-3}$, again showing that the learned model predicts device behavior to high accuracy. Notably, the model learned using only single-qubit data was sufficient to accurately predict the performance of the two-qubit gate on the device. We suspect this to be caused by exchange interactions due to coupling and finite temperature: Even when performing only single-qubit gates, the finite temperature causes a partial excitation of higher states, which are then exchanged with the other qubit via the coupling and thus visible in the ORBIT data.

The performance of the gate on the device is verified with Quantum Process Tomography (QPT): We apply the CR gate preceded and followed by single-qubit gates to prepare and measure in the basis states, e.g. $S = \left(X_{\pi/2} \otimes Y_{\pi/2}\right) \circ \text{CR} \circ \left(X_{-\pi/2} \otimes Y_{\pi/2}\right)$ [85], and again collect these measurements into our data-set. We believe that the entangling gate lifts the degree of freedom between misclassification and initial thermal distribution discussed before, hence we now perform another C_3 optimization, using the QPT data (256 sequences) and one ORBIT parameter per qubit (2×25 sequences) as the learning data. Parameter convergence is shown in the gray areas of Fig. 3.7(a), where temperature and confusion matrix values are adjusted closer to the true values.

Fig. 3.4(b) substantiates the claim that the entangling gate data allows for a more precise learning of the chip temperature, exhibiting a narrower valley at the true value. However, we are still not able to learn the T_1 and T_2^* parameters, since the sequences in QPT are too short to be sensitive.

3.3.7 Relaxation and dephasing

To demonstrate how a specialized measurement is formulated within $\mathbb{C}^3$ we determine the values of T_1 and T_2^*, using simple established sequences that are known to be sensitive to these parameters. The decay lifetime T_1 is determined by preparing the excited state of the qubit, followed by increasing wait times and then measuring the ground state population. We write the sequence as

$$S_{T1}^{(n)} = X_{\pi/2} \circ X_{\pi/2} \circ \mathcal{I}^n \tag{3.12}$$

where $X_{\pi/2}$ is our previously optimized $\pi/2$ gate and $\mathcal{I}^n$ signifies n repetitions of the identity gate $\mathcal{I}$. Similarly

$$S_{T2*}^{(n)} = X_{\pi/2} \circ \mathcal{I}^{n/2} \circ X_{\pi/2} \circ X_{\pi/2} \circ \mathcal{I}^{n/2} \circ X_{-\pi/2} \tag{3.13}$$

defines a Ramsey echo sequence, used to measure the dephasing time T_2^*. We take 51 logarithmically spaced values of n between 100 and 10000 to capture the full decay curves.

Using this data-set we perform another C_3 optimization, freezing all model parameters learned until now and varying only the values of T_1 and T_2^*. By doing so we manage to determine their values to within 1μs of the true values (Fig. 3.9). This procedure is the $\mathbb{C}^3$ equivalent of a common exponential decay fit to the data. However, with $\mathbb{C}^3$ one does not require prior knowledge on the expected structure of the experimental results, i.e. an exponential decay. Hence, when matching the data $\mathbb{C}^3$ also accounts for SPAM errors without the need to adjust the fitting function.

Fig. 3.4(c) shows the sensitivity of f_{LL} to the value of T_2^* of qubit B. The new data shows a clear improvement in the accuracy of the value obtained and the minimum is better defined. For increased sensitivity one would require more decay data to learn from.

3.3.8 Sources of error

The Good Model allows us to break down which of the model properties are preventing higher gate fidelities. To this end, we investigate the Good Model for components limiting the performance of the CR gate by idealizing aspects of the model.

We investigate whether the Gaussian ansatz is limiting gate fidelities by further refining the optimal pulses using a piece-wise constant optimization with one pixel per AWG sample (as is done in [71]). Average infidelity improves only marginally from $f_{\mathrm{av}}^{\mathrm{DRAG}} = 3.8 \times 10^{-3}$ to $f_{\mathrm{av}}^{\mathrm{PWC}} = 3.6 \times 10^{-3}$, suggesting other factors are limiting fidelities.

To find out if performance is limited by decoherence effects, we re-optimize the CR gate while disabling Lindbladian dynamics. By open-loop optimization in this idealized coherent setting the error is decreased from 3.8×10^{-3} to 1.3×10^{-5}. Thus, the 100 ns CR gate considered here is coherence limited, as is the case in most experimental implementations [59, 86], making improvements in gate time essential [87].

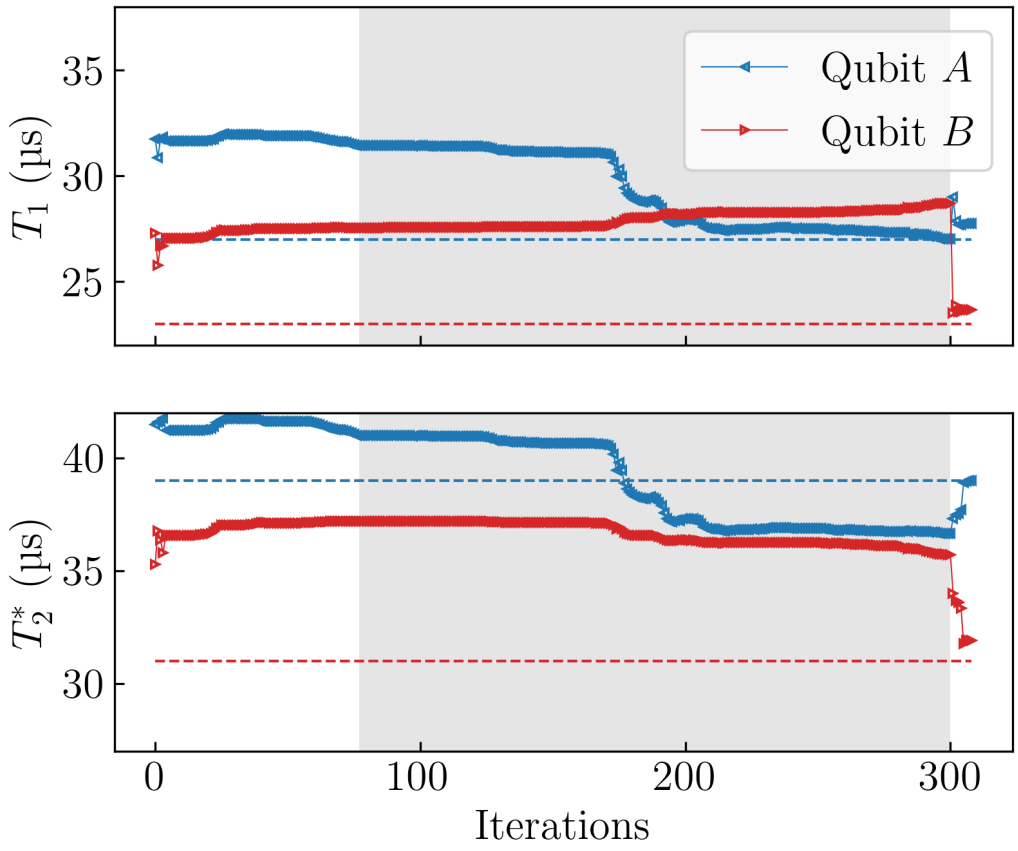

Figure 3.9: C_3 learning of the relaxation (T_1) and dephasing (T_2^*) parameters. Blue, left (red, right) triangles indicate qubit A (B) parameters, respectively. The true values of the "real" model are indicated as dashed lines. Background sections represent different learning datasets: just ORBIT data (left white section), a mix of ORBIT and QPT data (center gray section), decay and Ramsey data (right white section). The decay times are correctly identified only when specific data, sensitive to the decoherence effects, is used for learning, at which point they quickly converge to the real value.

3.4 $\mathbb{C}^3$ in-depth

Following is a detailed description of the $\mathbb{C}^3$ tool-set, its modeling capabilities and a general formulation of the optimization problems discussed in the previous section.

3.4.1 Experiment modeling

To combine control and characterization, $\mathbb{C}^3$ provides a detailed simulation that endeavors to encompass all relevant practical considerations of the experiment such as signal processing, SPAM errors, control transfer functions and Markovian noise. The simulator is used as the basis of the open-loop optimal control optimization (C_1) and the model parameter optimization (C_3). In both cases it is desirable to use gradient-based optimization algorithms [49, 88]. However, it is extremely cumbersome to manually derive the full analytical gradients of the quantum dynamics, especially when it includes the properties described above. Instead, $\mathbb{C}^3$ uses a numerics framework [89] which allows for automatic differentiation [90], making the tool-set more flexible and easily extendable. A similar approach is also used by [67, 91] for control and characterization.

Signal processing

The simulation allows for the specification of control signals $\varepsilon(t)$ as either analytical functions or as direct, piecewise constant AWG parameterization. Analytic parametrizations are sampled at the resolution of the waveform generator producing the envelope signal $\varepsilon_i = \varepsilon(t_i)$, representing voltages being applied to the control line, where the $\{t_i\}$ are the AWG sample times. The resulting signal will exhibit a *rise time* τ, due to the finite bandwidth of the control electronics. We model this by applying a convolution

$$\tilde{\varepsilon}(t) = \int_{t_0}^{t_f} \mathrm{interp}(\{\varepsilon_i\})(t) G\{t_f - t\}\, \mathrm{d}t \tag{3.14}$$

with

$$G(t) = \frac{1}{N} \exp\left\{-\frac{(t-\tau/2)^2}{8\tau^2}\right\}, \tag{3.15}$$

modeling a Gaussian filter, and

$$\mathrm{interp}(\{\varepsilon_i\})(t) = \{\varepsilon_i \,|\, t_i \leq t < t_{i+1}\} \tag{3.16}$$

interpolating the sampled signal to higher resolution for simulation. An I/Q-Mixer combines this envelope with a local oscillator signal of frequency ω_{lo} to

$$u(t) = I(t)\cos(\omega_{\mathrm{lo}} t) - Q(t)\sin(\omega_{\mathrm{lo}} t) \tag{3.17}$$

where the in-phase and quadrature components

$$\begin{aligned} I(t) &= \tilde{\varepsilon}(t)\cos\big(\phi_{xy} - \omega_{\text{off}}t\big) \\ Q(t) &= \tilde{\varepsilon}(t)\sin\big(\phi_{xy} - \omega_{\text{off}}t\big) \end{aligned} \tag{3.18}$$

are assigned by a control parameter ϕ_{xy}, and modulated to introduce a frequency offset ω_{off} on the drive. As noted in [92], in practice there will be additional errors during the mixing, which are not currently modeled. In transmitting this signal to the experiment, various distortions can occur, modeled by a response function φ, which also converts the field from line voltage to an amplitude $c(t) = \varphi[u(t)]$.

Time evolution

The system Hamiltonian is

$$H(t) = H_0 + \sum_k c_k(t)H_k \ , \tag{3.19}$$

with a drift H_0 and optional control Hamiltonians H_k. The dynamics of the system are described by the time-ordered propagator

$$U(t) = \mathcal{T}\exp\left\{-\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} H(t')\mathrm{d}t'\right\}, \tag{3.20}$$

given by solving the time-dependent Schrödinger equation, and approximated numerically by $U(t) \simeq \prod_{i=N}^{0} U_i$. Here, $U_i = \exp\{-\frac{\mathrm{i}}{\hbar}H(t_i)\Delta t\}$, and the total time is divided into N slices of length Δt that are fine enough so that the Hamiltonian can be considered constant in the interval.

In application, we will rarely perform a single gate or pulse in isolation. Experiments such as randomized benchmarking or the various flavors of tomography involve long pulse sequences, that are inefficient to simulate as a whole. Instead, the $\mathbb{C}^3$ simulator computes each propagator G of a defined gate-set $\mathcal{G}$ individually and compiles these matrix representations into sequences. This avoids the need to solve the equations of motions multiple times for the same exact pulses. As the propagators are calculated in the dressed laboratory frame (as opposed to the single-particle rotating frame), consecutive gates need to be adjusted to realign with the rotating frame of the drive signal, by applying a Z rotation with an angle of $(\omega_{\text{lo}} + \omega_{\text{off}})t_g$ [92].

To include open-system effects, we apply the equivalent procedure to obtain the process matrix

$$\mathcal{E}(t) = \mathcal{T}\exp\left\{\int_{t_0}^{t} \mathcal{L}(t')\mathrm{d}t'\right\} \tag{3.21}$$

by solving the master equation in Lindblad form,

$$\dot{\rho} = \mathcal{L}(\rho) = -i[H,\rho] + \sum_j L_j\rho L_j^\dagger - \frac{1}{2}\Big\{L_jL_j^\dagger, \rho\Big\}, \tag{3.22}$$

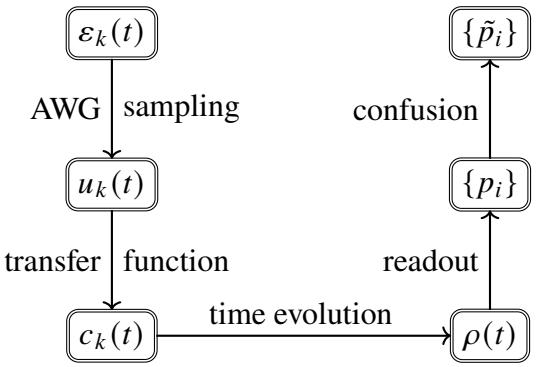

Figure 3.10: The process of simulating experimental procedure for signal processing and readout. The k-th control function is specified by some function $\varepsilon_k(t)$ and specifies the line voltage $u_k(t)$ by an arbitrary waveform generator (AWG) with limited bandwidth. Electrical properties of the setup, such as impedances, are expressed as a line transfer function φ, resulting in a control field $c_k(t) = \varphi[u_k(t)]$, as in Eq. (3.19). After solving the equation of motion for the system, readout and misclassification are modeled by applying rescaling and transformations to the simulated populations $p_i = |\rho_{ii}|^2$, according to Eq. (3.24).

where H is the Hamiltonian from Eq. (3.19), the L_js are Lindblad operators, and $\mathcal{L}$ is the generator in superoperator form [24]. The evolution of a state is obtained by applying the propagator as $\rho_f = U(t_g)\rho_i U^\dagger(t_g)$ for coherent evolution or $\rho_f = \mathcal{E}(t_g)[\rho_i]$ for incoherent evolution.

Initialization and readout

Given the temperature T of the device, the system is initialized in a mixed state

$$\rho_{\text{init}} = \sum_k \frac{1}{Z} |\phi_k\rangle\langle\phi_k| \exp\{-E_k/k_B T\} \tag{3.23}$$

where $\{|\phi_k\rangle\}$ is the eigenbasis of H_0 and the normalization is given by the canonical partition function $Z = \sum_k \exp\{-E_k/k_B T\}$.

We simulate readout by post-processing the final state ρ_f: From the density matrix, represented in the dressed basis, we obtain a vector of populations $\vec{p} = (p_k)$ by taking the absolute square of the diagonal. This is consistent with a slow (or weak) readout scheme in experiment. Measurement and classification errors are modeled with a misclassification (confusion) matrix $(p_{i\to j})_{ij}$ [93] such that the measured populations are

$$\tilde{p}_j = \sum_i p_{i\to j} p_i \ . \tag{3.24}$$

To simulate an experimental measurement with an average of l repetitions, we draw from a multinomial distribution of l trails and with probabilities $\tilde{p}_j$.

3.4.2 Optimizations

For open and closed-loop optimal control as well as model learning, performing optimization processes is required.

Open-loop Model-based Control: $\mathbf{C}_1$

In the typical setting of open-loop optimal control [47, 48], given a model of a system, we search for the optimal control pulses to drive the system to a desired state or generate a certain gate. Pulses are parameterized by an analytic ansatz (e.g. Gaussian pulse with DRAG correction [53] to remove Fourier components coupling to leakage levels), or by direct AWG samples. Constraints may be imposed to conform with experimental feasibility, such as power and bandwidth limitations. The goal function to be minimized is selected depending on the specified optimal control task, e.g. state infidelity for state transfer problems, or unitary trace infidelity for quantum gates [48, 73]. We suggest the use of average gate infidelity as the goal function, as it is experimentally accessible by performing RB or QPT, allowing comparison of performance in simulation and experiment.

Formally, the controls are parameterized as a vector of real numbers α. Given a goal function $f_1(\alpha)$, we search for $\min_\alpha f_1(\alpha)$. Optimal control methods such as GRAPE [50], Krotov [52, 94–96], and GOAT [49] have been devised to determine the gradient $\partial_\alpha f_1(\alpha)$ in order to facilitate convergence. These methods require a specific formulation of the problem and the analytical calculation of the gradient any additional elements in the model, whereas in $\mathbb{C}^3$, automatic differentiation allows to systematically account for any model feature, including, for example, line response functions or SPAM error.

The disadvantage of gradient-based algorithms is their propensity to get trapped in local minima. The severity of the problem is reduced by using a hierarchy of progressively more complex control ansätze. If this is insufficient, a short preliminary gradient-free search to find the convergence basin most often resolves the problem.

Closed-loop Model-free Calibration: $\mathbf{C}_2$

In calibration, a given pulse is optimized to improve a figure of merit $f_2(\alpha)$, computed from experimental measurement results. In addition to gradient-free optimization algorithms, there are methods to approximate the gradients (e.g. [97]), however, such approaches are generally less efficient than gradient-free algorithms [49, 98] as they require a high number of evaluations [99]. If the initial point of the optimization is given by C_1, this implements the already established Ad-HOC [57] method. During calibration, sets of control parameters α_j are sent to the experimental setup, alongside instructions of how to evaluate the current controls. For evaluating gate-sets, we suggest the ORBIT figure of merit, as it naturally performs a twirling of all sources of error, providing a single number to optimize. However, protocols tailored to specific needs can also be used, e.g. to obtain a desired conditional phase [44]. C_2 then optimizes the control parameters α_j to minimize a figure of merit.

While specialized measurements provide a straightforward way to fine-tune controls related to specific device properties, they do not generally account for interdependency. For more complex setups with many parameters, such calibrations cannot be done without extraordinary effort [100]. In contrast, $\mathbb{C}^3$ employs modern gradient-free optimization algorithms, such as CMA-ES, capable of optimizing dozens of parameters simultaneously, automating the task.

Model Learning: $\mathbf{C}_3$

Extracting the model from a data-set $\mathcal{D}$ can be thought of formally as analogous to the C_1 optimization task, where one varies the *model* parameters instead of the control parameters. For each measurement outcome m_k in the data-set,

$$\mathcal{D} = \{S_k \mapsto m_k\}_j \,, \tag{3.25}$$

the corresponding gate or pulse sequences $S_k(\alpha_j)$ with control parameters α_j are used to simulate the model's prediction $\tilde{m}_{j,k} = \tilde{m}\big(S_k(\alpha_j), \beta\big)$. The model learning goal function

$$f_3(\mathcal{D}|\beta) = f_3(\{\tilde{m}_k(\beta)\}, \{m_k\}) \tag{3.26}$$

quantifies the match between the data-set and the simulation of a system with parameters β. In this work, we use a rescaled log-likelihood

$$f_{\mathrm{LL}}(\mathcal{D}|\beta) = \frac{1}{2K} \sum_{k=1}^{K} \left(\left(\frac{m_k - \tilde{m}_k}{\tilde{\sigma}_k} \right)^2 - 1 \right) , \tag{3.27}$$

where the $\tilde{\sigma}_k$ is the standard deviation of a binomial distribution with mean $\tilde{m}_k$, resulting in a variation of the Mahalanobis distance [101]. This function is strictly correct under the Gaussian assumption and a two-level readout. See Sec. 3.4.3 for the extension for a multiple outcome readout. The measurement process on any physical device is noisy, i.e. each m_k is an estimate of a true underlying μ_k. Therefore, a realistic data-set $\mathcal{D}$ cannot be matched exactly by a deterministic simulation. The function f_{LL} is designed such that, for n data points, its expectation value is 0 when the model predicts the means μ_k correctly, and $\frac{1}{2}n^2$ if the distance is $\mu_k - \tilde{m}_k = n\sigma_k$ for all ks, according to Eq. (3.33). This provides a more intuitive measure of model match than the abstract value of f_{LL}, i.e. it allows us to make a statement like "the model differs from the experiment by approximately n standard deviations".

Due to the complexity of the physical systems, a potentially high number of interdependent parameters and complex features of the landscape, it is difficult for the optimization to converge to the global optimum. Therefore, we take the tried-and-tested experimental approach of starting with a simple model and iteratively refining it. We modify the model and repeat the C_3 fit, optionally retaining the optimized parameters which are shared by the previous and new model. Alternatively, we collect additional data and repeat the optimization on the same model. We emphasize that at each of these steps the physicists'

insights are required to evaluate the optimization's results, extend or discard models and decide whether collecting additional data is required. Furthermore, employing a gradient-based algorithm can, depending on the initial point, result in a local minimum. The optimizations presented here were successful when starting with a gradient-free CMA-ES search, known to be robust against local minima, switching over to the faster converging gradient-based L-BFGS method when a promising parameter region is identified. However, further research is required to find the best optimization strategy.

Outside the countless parameter-specific measurements, there are many approaches that aim to automate or generalize the task of characterization, such as: Bayesian learning (with Hamiltonian description [102–111], and more general-purpose [112–118]) , system-identification [119–124], compressed-sensing [109, 125, 126], neural network [127–129], and others [130].

In contrast, we note that in $\mathbb{C}^3$ a model takes explicit values for all its parameters, and is *not* represented as a high-dimensional distribution over model parameter space. This choice is driven by classical computation-load considerations: Because the $\mathbb{C}^3$ model is highly detailed, and, as consequence, associated simulations are non-trivial, we believe a full Bayesian approach to any of $\mathbb{C}^3$ optimizations is not computationally viable at this time.

3.4.3 C_3 Model Fitting Goal Function

When performing a series of experiments, $k \in [1, \ldots, K]$, on a quantum device, each experiment k is repeated a number of times and the normalized occurrences of the measurement outcomes are store in a result vector m_k. These are collected in $\{m_k\}_{k=1,\ldots,K}$, or $\{m_k\}$ for shorthand. Given a model and its parameters β, we aim to quantify how likely it is that it underlies the observed data with a function $f_3(\{m_k\}|\beta)$. Hence, we need to determine the distance between the experimental result, m_k, and the model prediction, $\tilde{m}_k(\beta)$. We define $p_k(m_k|\beta)$ be the model-predicted probability distribution function (PDF) for the result of experiment k. As the m_k are sampled from the readout distribution, we do not expect $m_k = \tilde{m}_k(\beta)$. Rather, we aim to define the function f such that its expectation value, $\mathrm{E}[f(\{m_k\}, \beta)]$, is zero if the underlying distributions from which the $\{m_k\}$ are drawn are the same as the model-predicted PDFs.

The Gaussian Assumption

To simplify calculation of $\mathrm{E}[f(\{m_k\}, \beta)]$, we can make some assumptions regarding the underlying distributions. The natural $p_k(m_k|\beta)$ PDF is multinomial, determined by the dimension of the qubit Hilbert space d_k (or binomial if dealing with a single qubit with no leakage levels). Under the assumption that for a large number of shots all possible readouts values are likely to appear, then by the central limit theorem (De Moivre–Laplace theorem), we can approximate p_k with a multivariate normal distribution. Although, the multinomial distribution has a non-diagonal covariance matrix, one can diagonalize the distribution and decompose it as a product of one-dimensional Gaussian distributions. Thus, we write the

PDF as a sum of $\sum_k (d_k - 1)$ such distributions, redefine K to equal the previous $\sum_k (d_k - 1)$ and the $\{m_k\}$ to be their means.

The model match distribution

We shall use $\{\tilde{\mu}_k(\beta)\}$ and $\{\tilde{\sigma}_k(\beta)\}$ to denote the mean and standard deviation of the model-predicted PDFs (after Gaussian assumption and multinomial diagonalization), and $\{\mu_k\}$ and $\{\sigma_k\}$ to denote the commensurate experimental values. We note that $\{\mu_k\}$ and $\{\sigma_k\}$ are unknown and unmeasured, and $\{m_k\}$ only provides an estimate of the mean. The simulation values $\{\tilde{m}_k\}$ on the other hand are deterministic and thus represent an exact estimate of the mean, hence $\{\tilde{m}_k \equiv \tilde{\mu}_k\}$.

The model-predicted PDF is given by the product of normalized Gaussian distributions, and gives the likelihood of the $\{m_k\}$ given the model parameters β as

$$p\left(\{m_k\}|\beta\right) = \prod_k p\left(m_k|\beta\right) \text{ , where} \tag{3.28}$$

$$p\left(m_k|\beta\right) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}_k} \exp\left(-\frac{1}{2}\left(\frac{m_k - \tilde{\mu}_k}{\tilde{\sigma}_k}\right)^2\right) \tag{3.29}$$

are the individual Gaussian distributions. We then construct the goal function as the average log-likelihood, rescaled to give the desired expectation value,

$$\begin{aligned} f_{\text{LL}}\left(\{m_k\}|\beta\right) &= -\log\left[\sqrt[K]{p\left(\{m_k\}|\beta\right)\prod_k \sqrt{2\pi}\tilde{\sigma}_k\sqrt{e}}\right] \\ &= \frac{1}{K}\sum_k \frac{1}{2}\left(\left(\frac{m_k - \tilde{\mu}_k}{\tilde{\sigma}_k}\right)^2 - 1\right) . \end{aligned} \tag{3.30}$$

Here the $\sqrt[K]{\cdot}$ gives the average of the log-likehoods, the $\sqrt{2\pi}\tilde{\sigma}_k$ removes the normalization of the Gaussians, such that they take value 1 when $m_k - \tilde{\mu}_k = 0$, and the log-likelihood is zero, and the $\sqrt{e}$ removes the residual part of the expectation caused by the noise in the $\{m_k\}$.

Then, in the general case, when the Gaussians determined by the model are not the same as the Gaussians in the experimental data:

$$\text{E}\left[f_{\text{LL}}\left(\{m_k\}|\beta\right)\right] = \frac{1}{2K}\sum_k \left(\left(\frac{\mu_k - \tilde{\mu}_k}{\tilde{\sigma}_k}\right)^2 + \left(\frac{\tilde{\sigma}_k}{\sigma_k}\right)^2 - 1\right) , \tag{3.31}$$

$$\text{Var}\left[f_{\text{LL}}\left(\{m_k\}|\beta\right)\right] = \frac{1}{K^2}\sum_k \left(\frac{\sigma_k}{\tilde{\sigma}_k}\right)^2 \left(\left(\frac{\mu_k - \tilde{\mu}_k}{\tilde{\sigma}_k}\right)^2 + \frac{1}{2}\left(\frac{\sigma_k}{\tilde{\sigma}_k}\right)^2\right) . \tag{3.32}$$

In the limit that both distributions have the same standard deviation $\sigma = \tilde{\sigma}$

$$\mathrm{E}\left[f_{\mathrm{LL}}^{\sigma\leftarrow\tilde{\sigma}}\left(\{m_k\}\,|\beta\right)\right] = \frac{1}{2K}\sum_k\left(\frac{\mu_k-\tilde{\mu}_k}{\tilde{\sigma}_k}\right)^2 \tag{3.33}$$

$$\mathrm{Var}\left[f_{\mathrm{LL}}^{\sigma\leftarrow\tilde{\sigma}}\left(\{m_k\}\,|\beta\right)\right] = \frac{1}{2K} + \frac{1}{K^2}\sum_k\left(\frac{\mu_k-\tilde{\mu}_k}{\tilde{\sigma}_k}\right)^2 . \tag{3.34}$$

Equation (3.33) then represents the square of the Mahalanobis distance (standardized Euclidean distance), giving an intuitive way to scale the f_{LL} function to understand the model match score. Indeed, when the model is exact and $\mu_k = \tilde{\mu}_k$ we get $\mathrm{E}\left[f_{\mathrm{LL}}^{\mathrm{exact}}\left(\{m_k\}\,|\beta\right)\right] = 0$. We note, however, that the function can take values below 0 as the variance for the exact case is $\mathrm{Var}\left[f_{\mathrm{LL}}^{\mathrm{exact}}\left(\{m_k\}\,|\beta\right)\right] = \frac{1}{2K}$. Such values indicate the standard deviation expected by the model, $\tilde{\sigma}_k$ is larger than the standard deviation observed experimentally, σ_k.

3.4.4 Model analysis

Both during and after the learning process, it is beneficial to interrogate the model to estimate its properties and their impact on the system behavior. As part of the $\mathbb{C}^3$ tool-set we perform sensitivity analysis for system parameters: Sweeping a single parameter, e.g. qubit frequency, across the range of interest, while keeping other parameters at their current best value, evaluating the model match score at each point, as seen in the example (Fig. 3.4(a)). The result is a 1-D cut through the optimization landscape that may exhibit a well-defined minimum, multiple extrema indicating a difficult optimization, or even appear flat in the case when a parameter does not affect the behavior of the current experiment. This landscape depends on both the selected model and data it is compared to. Depending on the ruggedness of the sensitivity, one might choose to utilize a gradient-based algorithm from the start or to first perform a gradient-free exploratory search to avoid local minima. In the case of a flat sensitivity, there are two courses of action: If the parameter is of little interest for successive experiments, it may be removed or set to a convenient value within the flat range; otherwise, one needs to design an experiment producing additional data that is sensitive to the parameter. The physicists' knowledge of common experimental practices (e.g. Rabi, Ramsey, Hahn echo sequences) and intuition guides the decision for the experiment design. When suspecting correlations between parameters, cuts in single dimensions are not enough and higher dimensional sweeps are necessary. After a successful learning process, the sensitivity analysis gives an estimate of the precision to which each parameter has been determined.

Furthermore, the simulation allows insight into the behavior of the system. Using well established methods such as time-resolved state and process tomography, it is possible to identify coherent errors, such as leakage out of the computational subspace, over-rotations, and the effects of noise. A Good Model also provides the basis for an error budget, as it contains the same limitations as the experiment it accurately predicts. The model can be used for extrapolation by idealizing certain aspects suspected as causes of infidelity (e.g. T_1

setting to infinity), and re-deriving control pulses using a C_1 optimization. The respective gain in fidelity gives an estimate of the error that this aspect is responsible for, suggesting areas of improvement for future devices.

3.5 Survey of parameter specific characterization

The task of characterization of quantum devices in general has received extensive attention. It would be presumptuous of us to even attempt a complete survey, therefore, we shall limit ourselves to a very limited look at a subset of model-specific methods we have subjectively found informative to our own work.

The standard approach at addressing the lack of a Good Model, as defined above, is to perform a long list of model-specific characterization experiments, each designed to measure a different parameter of the model: measure parameters of the readout resonator using frequency sweeps; qubit frequency measurements and relaxation time T_1 require Rabi experiments [131] (and with some extra effort the higher levels can be extracted); Ramsey [132] and Hahn echo measurements [133] provide dephasing data (under the Markovian assumption, which is known to be an over-simplification [134, 135]); measuring the control line response functions [136–139, 139, 140], the noise spectra [67, 135, 141, 142], continuous drifts in system parameters [143–146], and discontinuous jumps in parameters such as T_1 [147, 148]; state Preparation and Measurement Errors (SPAM) can be extracted from Randomized Benchmarking (RB; e.g. [75]) or dedicated procedures, such as [149]; qubit cross-talk can be measured by the method described in [150, 151] and many more. Model specific methods also exist for learning spin chain, lattice Hamiltonians and other multi-particle systems with a predefined network topology under limited access [152–156].

3.6 Open-source implementation

$\mathbb{C}^3$ is implemented as an open source project available at `https://q-optimize.org` under the Apache 2.0 license. The software is written in Python to interface conveniently with common experiment controllers, and has already been used in tandem with PycQED [157], Labber [158] and LabView [159]. The interface can occur at various levels of abstraction, from sharing control parameters to sampled waveform values. A modular design allows for Hamiltonian or Lindbladian descriptions of common physical systems (fixed and flux-tunable qubits, resonators, different types of coupling), specification of a list of devices to model the signal chain of the experiment (local oscillator, AWG, mixers, distortions and attenuations), different types of readout processing, and various fidelity functions. All components can be edited by the user or taken from reference libraries, accommodating to different needs. Configurations and data are stored as JSON files, and the full capabilities are accessible as command-line scripts, allowing for easy automation.

Numeric calculations are performed using TensorFlow [89]: The simulation of the dynamics and the pre and post processing are formulated as a network, with well-defined inputs (e.g.

control and model parameters) and outputs (goal function values), connected by many nodes, each performing a relatively simple operation (e.g. matrix exponentiation). TensorFlow enables the numerical computation of the Jacobian of a calculation – the gradient of each of the network outputs with respect to the network inputs (this capability is the evolution of what is known as back-propagation learning process in neural networks [160]). This process of automatic differentiation facilitates the modular structure, as any new component inherits this property, removing the need to analytically derive its gradient. Furthermore, the TensorFlow simulator is scalable, allowing deployment on a variety of high-performance computing hardware.

We note prior efforts simulating quantum circuits which allow for automatic differentiation, e.g. [91, 161], as well as large-scale simulations of quantum circuits, e.g. [162–164].

Each component of the control stack and model needs to conform to a general boilerplate that specifies what parameters it contains and how they are used. In this modular design, each class represents a component of the experiment that takes an input applies some parameter-dependent function to it and returns a result. For example, an envelope function for pulses would have this structure:

```
import tensorflow as tf
...
def my_envelope_fuction(t, parameters):
    amplitude = parameters["amp"]
    p2 = parameters["p2"]
...
return tf.some_math_function(amplitude, p2, t)
```

The only requirement to this code is that mathematical functions have to be taken from the TensorFlow package to allow for automatic differentiation. As an example of a control stack element, the finite rise time of an AWG is implemented with the following code:

```
class Response(Device):
def __init__(..., rise_time, ...):
    ...
    self.params['rise_time'] = rise_time

    def process(self, iq_signal):
            ...
            t = self.params['rise_time']
            sigma = t / 4
            ...
            # Convolution with a Gaussian
            ...
        return signal
```

A signal processing chain is represented by putting the output of one control stack element into the next. In calculating figures of merit, the user can choose from a library of functions or supply their own. For example, the infidelity of a state transfer process from $|\psi_0\rangle$ to $|\psi_{\text{ideal}}\rangle$, implemented by the simulated propagator U as follows:

```
def state_transfer_infid(U, psi_ideal, psi_0):
    psi_actual = tf.matmul(U, psi_0)
    overlap = tf_abs(
    tf.matmul(
        tf.linalg.adjoint(psi_ideal),
        psi_actual
        )
    )
    infid = 1 - overlap
return infid
```

At the time of writing, this release was the proof of concept for the software commercialization effort QRUISE [165], where is further developed in to a professional software package.

3.7 Discussion

In conclusion, we have described $\mathbb{C}^3$, an integrated methodology to improve quantum device performance that combines characterization, calibration and control. We have detailed its approach and implementation, demonstrating, on a synthetic QPU device, the individual methods and how they are synthesized into a more integrated concept. Analyzing single-qubit calibration data we successfully extracted an accurate model of the device, including realistic experimental considerations: line transfer functions, limitations of control electronics, readout error and finite operating temperature. From this model we were able to derive a working high-fidelity two-qubit gate, without requiring any further calibration.

This approach represents a holistic theoretical take on the experimental workflow of a complex quantum computing experiment, that takes into account interactions between different tasks of an experimental lab. $\mathbb{C}^3$ provides a path to achieve, starting from an incomplete understanding of the system, both high-fidelity pulses and an accurate model. It integrates the tasks of open-loop control (that would require an already accurate model) and of calibration (that would require an experiment-specific fine-tuning procedure). Most notably, it provides the tools to reflect on the experiment outcome and gate performance, improving the model description of the system and providing insight into its behavior. $\mathbb{C}^3$ is not a "black-box" experiment controller that replaces physicists or engineers – rather, it reduces tedious tasks allowing for interaction with the quantum device on a more structural level. Instead of simply producing high-fidelity operations, $\mathbb{C}^3$ provides meaningful output in the form of a Good Model of the system, and other insights such as an error budget and a sensitivity analysis. In this sense, $\mathbb{C}^3$ is not to be confused with any single optimal control or benchmarking technique, as it includes results from decades of research in these fields aimed at making controls that allow to actually reach high fidelities efficiently [47, 88], unifying them into one framework.

4 Co-designing Transmon devices for control with simple pulses

In the current NISQ era, there is demand for functional quantum devices to solve relevant computational problems, which motivates a utilitarian perspective on device design: The goal is to create a device that is able to run a given algorithm with state-of-the-art performance. In this work, we use optimal control tools to derive the gate set required by a toy algorithm and, in tandem, explore the model space of superconducting quantum computer design, from dispersively coupled to stronger interacting qubits, to maximize gate fidelity. We employ perfect entangler theory to provide flexibility in the search for a two-qubit gate on a given platform and to compare designs with different entangling mechanisms, e.g., CPHASE and $\sqrt{\texttt{iSWAP}}$. To ensure the applicability of our investigation, we limit ourselves to "simple" (i.e., sparse parametrization) pulses and quantify, where results differ from using the full complexity of piecewise constant controls.

4.1 Introduction

When designing control layouts for superconducting quantum computing devices, different approaches are being pursued. Some designs, e.g., from IBM, are trying to achieve as much as is possible with fixed-frequency qubits with microwave controls, avoiding the extra noise introduced by flux lines. This has the challenge of no control over operating sweet spots and couplings after fabrication. Adding flux lines incurs the cost of increased decoherence to be able to move the working points of qubits around at will. Following this design, adding additional tunable junctions to the chip to act as couplers between data qubit junctions increases the footprint of the chip, but promises to provide controlled interactions. In addition, this isolates the data qubits from flux noise [13].

Creating devices that are well suited for a given task, as opposed to more general, fundamental research, is desirable for the practical work on quantum algorithms. To this end, tools are created to explore and fine-tune quantum device designs in an integrated manner [166–169].

This chapter has been submitted as "Nicolas Wittler, Shai Machnes and Frank K. Wilhelm, *Co-designing Transmon devices for control with simple pulses*. The text was written entirely by N. Wittler.

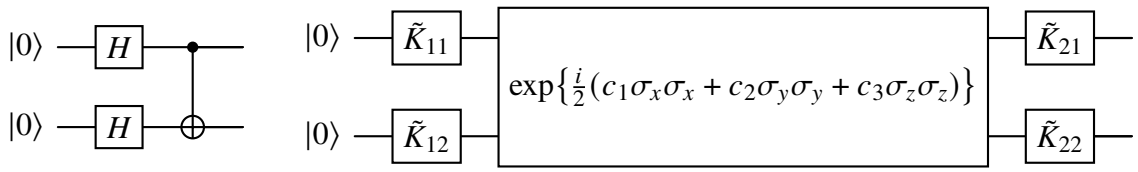

Figure 4.1: **Interpreting Lie group decomposition as circuit decomposition.** (a) Typical cell of a quantum computing circuit. (b) Decomposition with arbitrary entangling gate, given by the Weyl coordinates and modified local gates $\tilde{K}_{ij}$. Note that in a full algorithm circuit, layers of single qubit gates can be compressed.

It is our goal to systematically analyze designs for quantum computing devices that are fabricated for the purpose of executing quantum circuits or algorithms. We will present our procedure with the examples of two common design layouts: fixed qubits with fixed interaction (FQ), fixed qubits with a tunable coupler (TC). In this work, we'll consider the quantized Hamiltonian parameters – Transmon resonance frequencies, anharmonicities, coupling strengths – as the search space. Other efforts look to directly designing circuit quantities like charging and junction energies [168, 170] to produce devices with given properties, e.g. certain types of many qubit interactions terms.

Recent demonstrations have shown that even in restrictive control settings, almost all known gates are reachable [171]. This already presents a number of permutations of control schemes and platforms that need to be evaluated. Providing a level playing field is a non-trivial task, especially when trying to distill everything into a general procedure. Targeting a generalized perfect entangler can result in better outcomes than the "textbook" two-qubit gate [172–174]. When building a device for a certain application, there's an advantage in considering the requirements this poses to the gate-set [175, 176].

This is more relevant when we add the model parameters to the search. In [177], the authors coin the term "straddling" regime – the border between dispersive and strongly coupled – and show that this regime is optimal when employing full optimal control theory. We will employ these methods to investigate the advantage of a tunable coupler setup over fixed-frequency qubits, while limiting the control capabilities to simple pulses with few parameters. We'll see what the effect of limiting the pulse complexity has on the resulting model regime.

In Section 4.2, we reproduce some of the theory for creating entanglement. We present a general method to search for models that both facilitate entangling and local gates. As an initial application, we apply the method in Section 4.3 to a two-qubit chip with and without an extra qubit to act as a tunable coupler.

4.2 Method

For a functional quantum device, the fundamental tasks are: (1) Define and optimize local gates, (2) Define and optimize entangling gates, (3) Find optimal system parameters to

achieve both types of gates to high fidelity.

This presents a priori some conflicting interests: Entangling gates naturally profit from strong coupling, as opposed to local gates, which benefit from isolating subsystems. If the target is to run quantum circuits with high accuracy, we'll need to find a compromise. Here, we will present a procedure to achieve this, based on optimal control techniques.

The mechanism to generate entanglement for two-qubit gates generally depends on the platform, so choosing a perfect entangler as the optimization target is desirable to not limit possible solutions to a specific gate. For each setup, there exists a procedure to derive entangling gates; the cross-resonance (`CR`) gate for FQ and `CPHASE` or $\sqrt{\texttt{iSWAP}}$ for FQ and TQ, with AC or DC controls [178–180].

4.2.1 Perfect entangler theory

A known approach is to consider an operator $U \in SU(4)$ (the computational subspace, ignoring leakage levels for now) and characterize its non-local properties by computing the Makhlin invariants g_1, g_2 and g_3 [9].

We represent general gate U in the Bell basis as $U_B = Q^T U Q$, where

$$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & i \\ 0 & i & 1 & 0 \\ 0 & i & -1 & 0 \\ 1 & 0 & 0 & -i \end{pmatrix}. \tag{4.1}$$

In this basis, non-local gates are diagonal and local gates are orthogonal matrices.

The Makhlin invariants are then defined as

$$\begin{aligned} g_1 &= \mathrm{Re}\, \frac{\mathrm{Tr}\left(U_B^T U_B\right)^2}{16 \det(U)} \\ g_2 &= \mathrm{Im}\, \frac{\mathrm{Tr}\left(U_B^T U_B\right)^2}{16 \det(U)} \\ g_3 &= \frac{\mathrm{Tr}\left(U_B^T U_B\right)^2 - \mathrm{Tr}\left((U_B^T U_B)^2\right)}{4 \det(U)} \,. \end{aligned} \tag{4.2}$$

These invariants characterize equivalence classes, e.g., `CNOT`, `CPHASE` and `CR` share the same invariants, meaning they can be transformed into each other with local gates.

We can generalize further by combining the invariants $\vec{g} = (g_1, g_2, g_3)$ into a single functional,

$$d(\vec{g}) = g_3 \sqrt{g_1^2 + g_2^2} - g_1 \,, \tag{4.3}$$

which vanishes for the invariants on the surface of the volume of perfect entanglers. Inside the volume, the value has to be set to 0 manually [181].

This functional is straightforward to compute for a given unitary and represents a general measure for entangling performance. Some care must be taken tough, depending on which side of the volume the current $\bar{U}$ lies. We are interested in comparing the performance of single and two qubit operations during optimization and thus would prefer an equivalent fidelity of the form $\left|\mathrm{Tr}\left\{V^\dagger \bar{U}\right\}/\dim V\right|^2$, where V is the target gate and $\bar{U}$ the time evolution of the controlled system. One method is to use the perfect entangler functional for optimization and then evaluate the final performance based on fidelity. Since we aim to find suitable model parameters for both single qubit and entangling gates at the same time, we propose a direct numerical approach, the "tugboat" strategy.

4.2.2 Tugboat optimization for perfect entanglers

We write a general ideal target gate using the Cartan decomposition $V = K_1 A K_2$ with

$$A = \exp\left\{\frac{i}{2}(c_1\sigma_x\sigma_x + c_2\sigma_y\sigma_y + c_3\sigma_z\sigma_z)\right\} \tag{4.4}$$

where $K_{1,2} \in SU(2) \otimes SU(2)$. It can be shown, e.g., in [182] that the subalgebra $\{\sigma_x\sigma_x, \sigma_y\sigma_y, \sigma_z\sigma_z\}$ is sufficient to generate all of $SU(4)$ when combined with local rotations $K_{1,2}$. As an example, $\sigma_x\sigma_y$ can be created from $\sigma_x\sigma_x$ by a single qubit $\mathbb{1}\sigma_z$.

If we now confine the Weyl coordinates c_i to the volume of perfect entanglers (shaded blue in Fig. 4.2), we have a recipe to explicitly construct an arbitrary entangling gate. To this end, we propose the coordinate transformation

$$\begin{aligned} c_1 &= \frac{b_1 + b_2}{2} \\ c_2 &= \frac{b_1 - b_2}{2} \\ c_3 &= b_3(\pi/4 - |c_2 - \pi/4|) \end{aligned} \tag{4.5}$$

to conveniently express the boundary conditions for the perfect entangler volume [172] as

$$\begin{aligned} \pi/4 \le b_1 \le \pi/2 \\ 0 \le b_2 \le \pi/4 \\ 0 \le b_3 \le \pi. \end{aligned} \tag{4.6}$$

We will now consider a target gate, the "*tugboat*", $V(\vec{\gamma}, \vec{b})$, where $\vec{b} = (b_i)$ are the transformed Weyl coordinates, confined to the volume of perfect entanglers, and local rotations $\vec{\gamma} = (\gamma_i^{(j)})$ with $j = 1, 2$ enumerating the qubits and $i = x, y, z$ the rotation axis.

We assign an explicit fidelity error to a quantum gate $\bar{U}$ as

$$\epsilon_{\mathrm{II}}(\bar{U}) = \min_{\vec{b},\vec{\gamma}} \left\{1 - \left|\mathrm{Tr}\left\{V^\dagger(\vec{b}, \vec{\gamma})\bar{U}\right\}/\dim V\right|^2\right\} \tag{4.7}$$

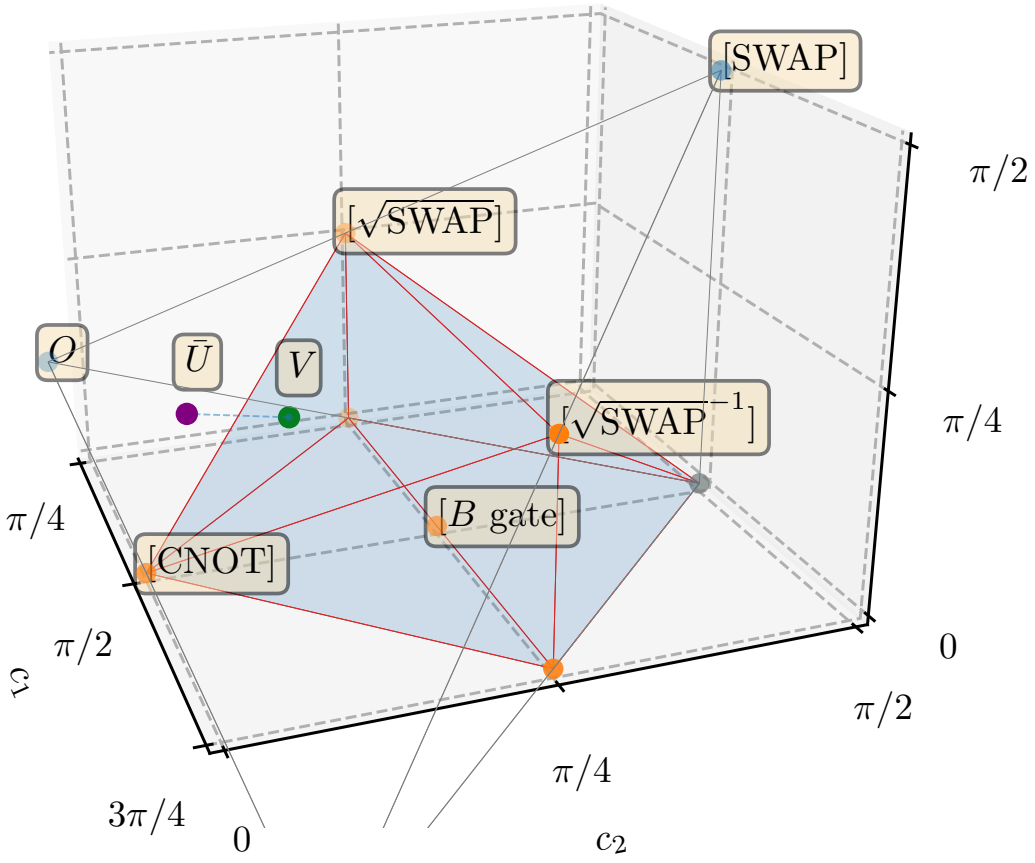

Figure 4.2: **Weyl chamber of named gates to illustrate the optimization procedure.** The labeled points indicate equivalence classes with respect to local rotations. The system dynamics $\bar{U}$ (projected into the computational space) are steered by the control and model parameters, while the target gate V is varied over the shaded volume, bounded by Eq. (4.6). This enables us to use $\left|\text{Tr}\left\{V^\dagger(\vec{b},\vec{\gamma})\bar{U}\right\}\right|$ in a regular distance measure Eq. (4.7), indicated by the dashed line.

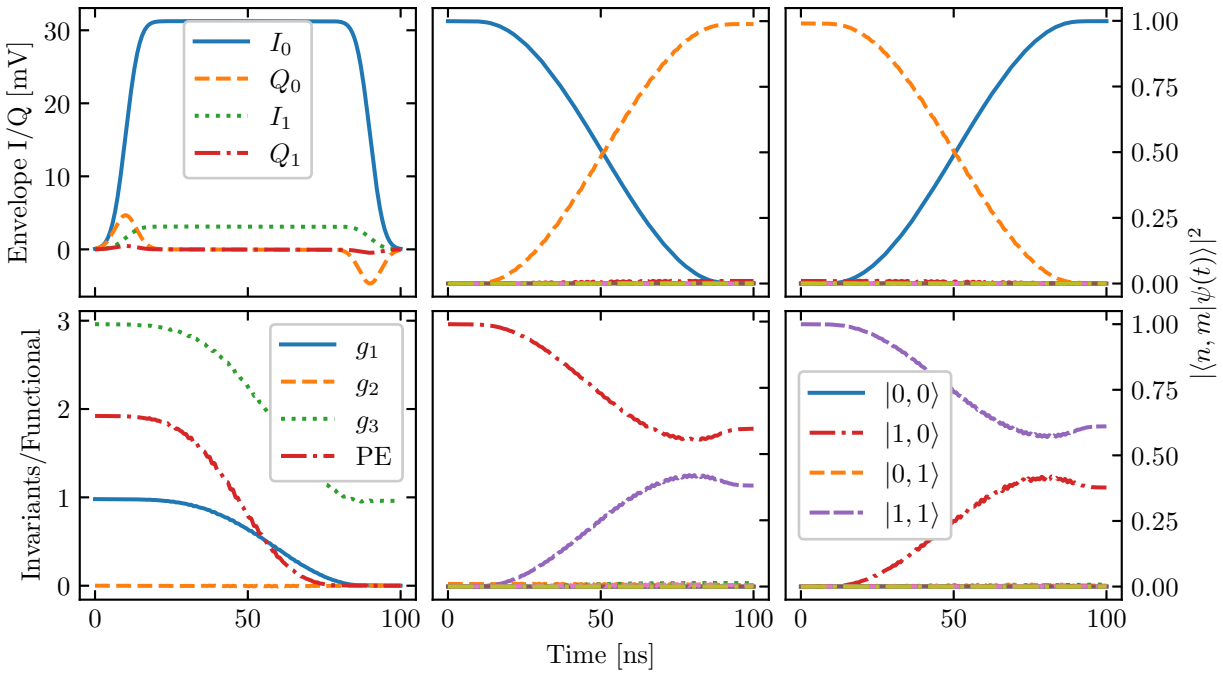

Figure 4.3: **Control fields and time-evolution of computational state populations of the two fixed Transmon device.** Both control fields are on resonance with the bare frequency of the second Transmon and the amplitude of the drive is increased by $(\omega_2 - \omega_1)/g$. Their I and Q envelopes are shown on the top left. The panels on the right show the time evolution of populations, initialized in the four computational basis states. We observe a Rabi oscillation for Transmon 2 (right index, the *target*), when Transmon 1 (left index, the *control*) is in the 0 state. When the control Transmon is in the 1 state, only a partial population transfer occurs at a higher general Rabi frequency, realizing a general cross-resonance effect. To verify the entangling property, we plot the Makhlin invariants and the value of the perfect entangler (PE) functional from Eq. (4.3).

where $\bar{U}$ is the time evolution of the system, projected into the computational subspace. It is not resource intensive to compute a unitary matrix V in this representation, compared to the system time evolution $\bar{U}$ which requires several orders more computational power. We can thus easily find the closest perfect entangler representation to the implemented gate $\bar{U}$ by varying $\vec{b}$ and $\vec{\gamma}$ at each optimization step. This way we assign a regular fidelity measure to $\bar{U}$ while retaining the generality of the perfect entangler goal. An added benefit of this procedure it that we also explicitly extract the Weyl coordinates and local rotations of the entangling gate that allow us to identify the equivalence class.

4.2.3 Optimal co-design

In this formalism, we can now also explore the model parameter space by including it in the optimization. Assume the model is described by a Hamiltonian $H(\vec{\alpha}, \vec{\beta})$, with $\vec{\beta}$ being the vector of model parameters and $\vec{\alpha}$ the controls. The system dynamics are then given by

$$U(\vec{\alpha}, \vec{\beta}) = \mathcal{T} \exp\left\{-\frac{i}{\hbar} \int dt H(\vec{\alpha}, \vec{\beta})\right\} \tag{4.8}$$

so formally the goal function depends on $\epsilon_{\mathrm{II}} \equiv \epsilon_{\mathrm{II}}(\vec{\alpha}, \vec{\beta}, \vec{b}, \vec{\gamma})$.

Textbook representations of two-qubit gates, e.g. $\texttt{CNOT} = |0\rangle\langle 0| \mathbb{1} + |1\rangle\langle 1| \sigma_x$, are implicitly written in the interaction picture according to the Hamiltonian in the measurement basis, denoted as $\tilde{H}$. We'll consider the readout basis to be the eigenbasis of the undriven system.

We obtain the simulation result $\bar{U} = \mathcal{T} \exp\left(-i \int H dt\right)$ in the product basis, since we will tune system parameters. Depending on the parameter regime, writing the transformation that diagonalizes H is not straightforward enough to facilitate computing the gradient of the goal function with respect to model properties.

To evaluate the optimization results, we can simulate $\tilde{H} = S^\dagger H S$, the diagonal (if applicable, dressed) Hamiltonian. We then write the goal unitary as $G = S\tilde{G}S^\dagger$ and, finally, the goal function as

$$\varepsilon = 1 - \left| \frac{1}{\dim G} \mathrm{Tr}\left\{G^\dagger \bar{U}\right\} \right|^2 \tag{4.9}$$

We can use as the transformation $S \equiv S(\beta)$ the closed, exact form in the computational subspace, depending on ω_1, ω_2 and g.

$$S = \begin{pmatrix} \cos(b) & 0 & 0 & \sin(b) \\ 0 & \cos(a) & \sin(a) & 0 \\ 0 & \sin(a) & \cos(a) & 0 \\ -\sin(b) & 0 & 0 & \cos(b) \end{pmatrix} \tag{4.10}$$

with $a = \arctan\left(\frac{g}{\omega_2 - \omega_1}\right)$ rotating the exchange interaction and $b = \arctan\left(\frac{g}{\omega_2 + \omega_1}\right)$ rotating the double excitations. Since we performed the RWA on the interaction terms, here $b = 0$.

As the fidelity of consecutive operations is given by their product, the error for single qubit gates (ϵ_{I}) followed by an entangling gate (with error ϵ_{II}) is

$$\bar{\epsilon} = 1 - (1 - \epsilon_{\mathrm{I}})(1 - \epsilon_{\mathrm{II}}) = \epsilon_{\mathrm{I}} + \epsilon_{\mathrm{II}} - \epsilon_{\mathrm{I}}\epsilon_{\mathrm{II}} \tag{4.11}$$

where we'll neglect the product $\epsilon_{\mathrm{I}}\epsilon_{\mathrm{II}}$ as its contribution vanishes for small errors below 1% and use $\epsilon = \epsilon_{\mathrm{I}} + \epsilon_{\mathrm{II}}$ as our goal function.

4.3 Application

We present the co-design procedure to explore designs for superconducting qubits. The model space consists of two Transmons with frequencies $\omega_{1,2}$, anharmonicities $\alpha_{1,2}$ and static interaction g_{12} collected in a five-dimensional vector $\vec{\beta} = (\omega_1, \omega_2, \alpha_1, \alpha_2, g)$. From the formal perspective, optimizing both model and controls at the same time constitutes a form of over-parametrization. For example, changes in drive tone and qubit frequency have the same effect on the goal function. Thus, we opt for simple controls, as opposed to a more general piecewise constant parametrization with hundreds of samples, so that the dimensions of $\vec{\alpha}$ and $\vec{\beta}$ are in the same order of magnitude.

We will first run through the process for the two-qubit chip with fixed interaction strength, and then add a third qubit to act as a tunable coupler. To characterize the model by a single number, we take the reduced coupling $\chi = g_{12}/(\omega_2 - \omega_1)$ and use the "straddling" regime $\chi = 0.1$ as a divider between the dispersive ($\chi < 0.1$) and strong coupling ($\chi > 0.1$). When there is a tunable coupler, we use the effective interaction strength, where the coupler is adiabatically eliminated to have a comparable quantity [183].

4.3.1 Local vs. entangling gates

For universal quantum computing, we need to derive a set of local, single-qubit gates and an entangling gate. As a first application, we consider a device made up of two fixed-frequency Transmons with a static coupling. With all driving and coupling resonators eliminated, the Hamiltonian is $H = \sum_{j=0,1,2,g} H_j$ with

$$\begin{aligned} H_0 &= \sum_i \omega_i a_i^\dagger a_i + \frac{\delta_i}{2}\Big(a_i^\dagger a_i - 1\Big)a_i^\dagger a_i \\ H_g &= g\Big(a_1^\dagger a_2 + a_1^\dagger a_2\Big) \\ H_i &= u_i(\vec{\alpha}, t)\Big(a_i^\dagger + a_i\Big)\,, \end{aligned} \tag{4.12}$$

where we are just considering the exchange interaction.

The control field on qubit i is $u(\vec{\alpha}, t) = A\left[I_i(t)\cos\big(\omega_i^d t + \phi_i\big) + Q_i(t)\sin\big(\omega_i^d t + \phi_i\big)\right]$ with a slow-varying, envelope I_i. As a basis function, we choose a flattop Gaussian shape

$$I_i(t) = \mathrm{erf}\left(-\frac{(t - t_{\mathrm{up}})}{\sigma}\right) - \mathrm{erf}\left(\frac{(t - t_{\mathrm{down}})}{\sigma}\right) \tag{4.13}$$

with amplitude A_i and ramps at t_{up} and t_{down} with a fixed width $\sigma = 5$ ns. We also add an out-of-phase DRAG [18] correction as $Q_i = \lambda_i \partial_t I_i$.

First, we optimize both controls $\vec{\alpha} = (A_i, \omega_i, \phi_i, t^i_{\mathrm{up}}, t^i_{\mathrm{down}}, \lambda_i)$ and model parameters $\vec{\beta} = (\omega_i, \delta_i, g)$ to find single qubit gates. We characterize the system by its reduced coupling strength $g/(\omega_2 - \omega_1)$. To find the entangling gate, we employ the procedure described in

Section 4.2.2. In Fig. 4.3 we show some example dynamics. We initialize the target gate V centered in the volume of perfect entanglers at $\vec{c}/\pi = (1/2, 1/4, 0.05)$ and the local rotations $\vec{\gamma}$ close to zero. After a set number of evaluations (here we found 4 to be a good value for convergence) in the gradient descent search, we update the target unitary coordinates $\vec{c}$ and $\vec{\gamma}$ and refresh the L-BFGS memory, if needed.

We want to include the effect of additional leakage in order to obtain realistic control solutions. For each subsystem, we simulate three Transmon levels labeled $|0\rangle$, $|1\rangle$ and $|2\rangle$. To ensure that system dynamics stay within the embedded computational subspace, we write a running cost

$$L = \sum_k \sum_{\lambda \in \Lambda} |\langle \lambda | \psi(t_k) \rangle|^2 \tag{4.14}$$

and add it to the goal function. The population in the leakage levels $\Lambda = \{|0,2\rangle, |1,2\rangle, |2,0\rangle, |2,1\rangle\}$ is evaluated at evenly spaced times t_i during propagation.

In Fig. 4.4, we show the infidelities ϵ during optimization. Starting in a dispersive setting where $g/(\omega_2 - \omega_1) < 0.1$, we observe that optimizing for single qubit gates trends towards isolated subsystems. When we target a perfect entangler, we observe convergence to a higher dispersivity, but still within the weak coupling regime. A combination of both goal functions, representing a single cell (Fig. 4.1) of some quantum computing algorithm, results in a compromise between the two edge cases, both in terms of dispersivity and reached fidelity. We also observe that the optimization trajectory in the "algorithm" target seems to be a superposition of the two layers. In this regime, the single qubit gates can be implemented to high fidelity, so the entangling and "algorithm" errors are close.

When initializing the system in the "straddling" regime, i.e., $g/(\omega_2 - \omega_1) = 0.1$ we find that we can recover the same convergence to single qubit gates. After a short exploration, the entangling gates return to the same effective coupling value and perform slightly better than in the dispersive initialization. The optimization for both however seems to get stuck realizing a compromise. This regime has been shown to be optimal in the presence of arbitrary control, but here, the chosen simple parametrization might be not expressive enough to realize this potential.

4.3.2 Flux-tunable coupler

As seen in the previous section, choosing a fixed frequency, fixed coupling design limits fidelities of algorithms, since the optimal parameter regimes for single and two qubit operations are different. The introduction of tunable elements presents a solution, by allowing different working points for both tasks. In principle, there are two designs: Adding a flux line to one of the qubits to tune its resonance frequency or add a tunable coupler, another non-data qubit with a tunable frequency. However, the added flux lines also provide a new avenue for noise to degrade coherence time of the qubits. To compare to the fixed frequency setup in the previous section, we choose to investigate a tunable coupler architecture and assume that the added noise channel on the coupler qubit is reasonably isolated from the "data" qubits.

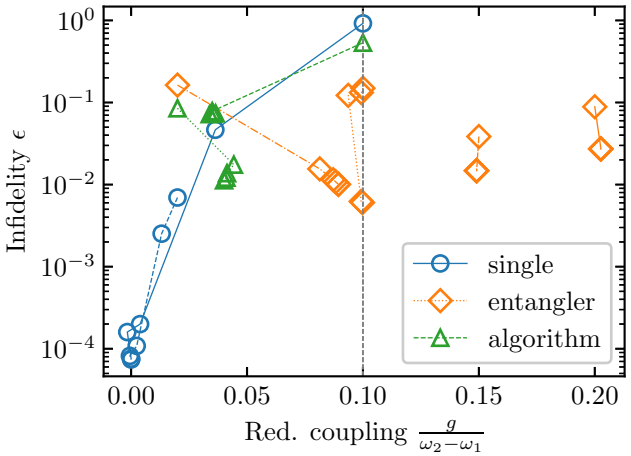

Figure 4.4: **Model space of a fixed frequency, two-qubit chip.** We optimize the parameters of two Transmons, their frequencies, anharmonicities and coupling, to implement local and non-local gates. Initializing the search around the straddling regime (dashed line), $\omega_2 - \omega_1 = 500$ MHz 2π and for different values of $g = 50$ MHz 2π, we observe that local gates trend towards the same convergence. Only trajectories that converge below an error of 0.1 are shown.

The flux-dependent frequency of a tunable element, is written as

$$\omega_i(\Phi) = \left(\omega_i^0 - \delta_i\right)\sqrt{\left|\cos\left(\frac{\Phi}{\Phi_0}\pi\right)\right|} + \delta_i \tag{4.15}$$

following the supplements of [184], where ω_i^0 is the sweet spot frequency and Φ the external flux. For simplicity, in this work we will directly consider the frequency offset $\omega_i^c(t)$ as the control parameter, such that $\omega_i = \omega_i^0 + \omega_i^c(t)$.

As a characteristic quantity, we look at the effective interaction strength

$$J = \frac{g_1 g_2}{2}\left(\frac{1}{\omega_1 - \omega_{\mathrm{TC}}^0} + \frac{1}{\omega_2 - \omega_{\mathrm{TC}}^0}\right) \tag{4.16}$$

induced between the two qubits [183]. Here, g_1 and g_2 being the coupling strength between qubit 1 and 2 and the coupler, respectively.

We put the idle point of the coupler ω_{TC}^0 above the two qubits in Eq. (4.16), around the same difference in frequency

$$\omega_{\mathrm{TC}}^0 - \omega_2 = \omega_2 - \omega_1 \ . \tag{4.17}$$

If, in addition, there's a static coupling g_{12} between the two qubits, the total effective interaction is

$$J' = g_{12} + \frac{g_1 g_2}{2}\left(\frac{1}{\omega_1 - \omega_{\mathrm{TC}}^0} + \frac{1}{\omega_2 - \omega_{\mathrm{TC}}^0}\right) \tag{4.18}$$

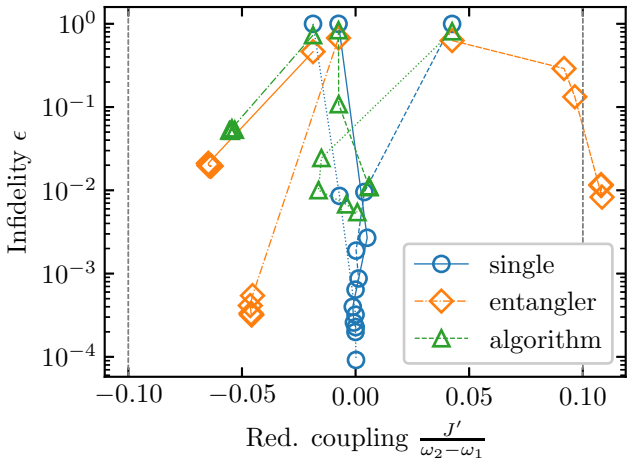

Figure 4.5: **Adding a tunable coupler to the model.** When we initialize the system in a sweet spot, where, according to Eq. Eq. (4.16) the interaction between the qubits vanishes, good single-qubit gates can be reached. Entangling gates are trending towards effective couplings similar to the straddling regime. Optimizations starting in stronger coupling regimes has either stalled, or recovered to regions of weak coupling. Only trajectories that converge below an error of 0.1 are shown.

For g_1 and g_2, we pick similar values to the coupling constant in the previous example, and we select the residual interaction g_{12} to initialize the optimization in different J' regimes. To entangle the qubits, we modulate ω_c with an AC tone close to their difference frequency [183], shaped by the same flattop envelopes as before. Again, we can implement single and entangling gates to good fidelities. Single qubit gates can reach good fidelities here, even exploring the region around the sweet spot and returning. Once the interaction strength is high enough, we can also entangle the qubits with a higher fidelity than the case without the tunable coupler. For the "algorithm" optimization, the same competition between local and non-local operations results in a stagnation at sub-par fidelity. The trajectory of the combined case is again alternating between the individual trajectories, see Fig. 4.5.

In both cases, we see that the approaching the strong coupling regime presents problems in realizing single-qubit gates, at least with the simple pulse shapes we explored here.

4.4 Summary and Outlook

We have shown a general procedure to investigate model design space for an application like superconducting qubits, where the properties of human-made devices can be targeted in fabrication. Instead of a full optimal control approach, we investigate if previous results about preferable regimes hold for simple pulse shapes. Compared to previous work, we choose a direct search in the Weyl space to find the closest perfect entangler to the gate

our test system can produce. This way, we avoid some mathematical difficulties, since the expression of the functional Eq. (4.3) is only defined outside the volume. Inside the volume, the value is 1 by definition. The direct use of Eq. (4.7) does not require these conditional expressions.

The application to the two-qubit model, optionally adding a tunable coupler, shows reliably that good single qubit gates can be reached when the effective interaction is turned almost off. This is as expected but also holds, when initializing a search some distance away from this regime. By the same reasoning, targeting entangling gates trends towards a similar point – the previously identified "straddling" regime from different initial values. There, we notice the necessity to restart optimizations, in the case of early termination. The use of simple pulses seems to limit the exploration of parameter space and prevents a good regime for the operation of both local and non-local quantum gates. To make a more systematic claim, a wider range of cases – control schemes and device designs – needs to be tested. For example, several schemes exist to mitigate the effect of spurious ZZ interactions to allow single qubit control on more strongly interacting systems [185, 186].

Even limited to the superconducting platform, several design paradigms to create quantum computing devices exist. The Transmon, with its specific E_C/E_J ratio, is only one possible design. It is the qubit design most demonstrations of applied quantum computing use. There are other promising paradigms, such as the Fluxonium, which combines insensitivity to external noise with a large anharmonicity [187]. So, a different search space can be explored with the method presented here by directly optimizing Josephson junctions, capacitances and inductances.

5 Conclusion

A lot of the tasks to set up a quantum computer are still based on manual procedures. Characterization and calibration are often thought of separately and performed in strict sequence. Advances in optimal control often are only adopted slowly in experimental, every day operation for similar reasons. In this thesis, we have looked at how to systematize some of these aspects from a theoretical perspective and produced software tools to implement general procedures in practice.

We proposed the $\mathbb{C}^3$ approach to bridge the gap between theory and practice by operating both in the same software environment. Seemingly simple issues, such as making sure that the same pulse shape definitions are used in both scenarios, are taken care of. Conversely, limitations of the real-world hardware are fed back into theoretical models to more accurately reproduce experimental data. Within this approach, we have seen that a systematically characterized model is can be sufficient to derive complex controls that perform well, even before calibration. In cases where it does not, that data is particularly useful to fill in missing gaps in the theoretical description.

Finally, we invert the problem: We ask, given a choice of physical system, what is the optimal basis to run a quantum algorithm? We presented some initial findings on a superconducting two-qubit device. Taking previous results, that show an optimal parameter regime for complex pulses, we investigate the same statement when simpler pulses with fewer parameters are used. Such pulses are already used in labs demonstrating quantum utility. We saw that this limitation can reduce the explorable model space and suggest a deeper investigation into the complexity of pulses vs. model parameter regimes.

In summary, we have shown the utility in applying advanced theoretical and software based methods to bridge the gap between demonstrators and early practical applications. The barrier of entangling gate fidelities needs to be crossed to advance quantum computing as a whole and enable error correction. This can be accomplished by the continued improvement of how we operate these devices and gain deeper understanding of their physics.

Bibliography

[1] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai, Nature **398**, 786 (1999).

[2] The future is Quantum, https://qt.eu/.

[3] A Quantum Computer for Europe | OpenSuperQPlus, https://opensuperqplus.eu/.

[4] German Quantum Computer based on Superconducting Qubits, https://geqcos.wmi.badw.de/home.

[5] G. Wendin, Reports on Progress in Physics **80**, 106001 (2017).

[6] T. Albash and D. A. Lidar, Reviews of Modern Physics **90**, 015002 (2018).

[7] B. C. Hall, *Lie Groups, Lie Algebras, and Representations*, Graduate Texts in Mathematics (Springer Cham, 2015).

[8] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Physical Review A **67**, 042313 (2003).

[9] Y. Makhlin, Quantum Information Processing **1**, 10.1023/A:1022144002391 (2002).

[10] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Journal of Magnetic Resonance **172**, 296 (2005).

[11] D. Kielpinski, C. Monroe, and D. J. Wineland, Nature **417**, 709 (2002).

[12] P. Rembold, N. Oshnik, M. M. Müller, S. Montangero, T. Calarco, and E. Neu, AVS Quantum Science **2**, 024701 (2020).

[13] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Applied Physics Reviews **6**, 021318 (2019), arXiv:1904.06560 [cond-mat, physics:physics, physics:quant-ph] .

[14] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley,

C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, Nature **574**, 505 (2019).

[15] M. H. Devoret *et al.*, Les Houches, Session LXIII **7** (1995).

[16] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Physical Review A **76**, 042319 (2007).

[17] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret, Science **296**, 886 (2002).

[18] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Physical Review Letters **103**, 10.1103/physrevlett.103.110501 (2009).

[19] D. D'Alessandro, *Introduction to Quantum Control and Dynamics* (Chapman and Hall/CRC, New York, 2008).

[20] R. H. Byrd, L. Peihuang, and J. Nocedal, A limited-memory algorithm for bound-constrained optimization, (1996).

[21] D. L. Goodwin and I. Kuprov, The Journal of Chemical Physics **143**, 084113 (2015).

[22] S. Machnes, E. Assémat, D. Tannor, and F. K. Wilhelm, Physical Review Letters **120**, 150401 (2018).

[23] M. H. Goerz, S. C. Carrasco, and V. S. Malinovsky, Quantum **6**, 871 (2022).

[24] D. A. Lidar, Lecture Notes on the Theory of Open Quantum Systems (2020), arXiv:1902.00967 [quant-ph] .

[25] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information: 10th Anniversary Edition, (2010).

[26] S. T. Flammia and Y.-K. Liu, Physical Review Letters **106**, 230501 (2011).

[27] D. Gottesman, Physical Review A **57**, 127 (1998).

[28] E. Magesan, J. M. Gambetta, and J. Emerson, Physical Review Letters **106**, 180504 (2011).

[29] M. Werninghaus, D. J. Egger, F. Roy, S. Machnes, F. K. Wilhelm, and S. Filipp, npj Quantum Information **7**, 14 (2021), arXiv:2003.05952 [quant-ph] .

[30] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, Physical Review Letters **112**, 240504 (2014).

[31] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Physical Review A **77**, 032322 (2008).

[32] J. R. Johansson, P. D. Nation, and F. Nori, Computer Physics Communications **184**, 1234 (2013), arXiv:1211.6518 [quant-ph] .

[33] D. J. Egger and F. K. Wilhelm, Physical Review Letters **112**, 10.1103/physrevlett.112.240503 (2014).

[34] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Physical Review A **69**, 10.1103/physreva.69.062320 (2004).

[35] M. Reagor, W. Pfaff, C. Axline, R. W. Heeres, N. Ofek, K. Sliwa, E. Holland, C. Wang, J. Blumoff, K. Chou, M. J. Hatridge, L. Frunzio, M. H. Devoret, L. Jiang, and R. J. Schoelkopf, Physical Review B **94**, 014506 (2016).

[36] E. Xie, F. Deppe, M. Renger, D. Repp, P. Eder, M. Fischer, J. Goetz, S. Pogorzalek, K. G. Fedorov, A. Marx, and R. Gross, Applied Physics Letters **112**, 202601 (2018).

[37] A. Blais, J. Gambetta, A. Wallraff, D. I. Schuster, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, Physical Review A **75**, (2007).

[38] S.-I. Chu and D. A. Telnov, Physics Reports **390**, 1 (2004).

[39] M. F. Dumas, B. Groleau-Paré, A. McDonald, M. H. Muñoz-Arias, C. Lledó, B. D'Anjou, and A. Blais, Physical Review X **14**, 041023 (2024).

[40] P. Jurcevic, A. Javadi-Abhari, L. S. Bishop, I. Lauer, D. F. Bogorin, M. Brink, L. Capelluto, O. Günlük, T. Itoko, N. Kanazawa, *et al.*, Quantum Science and Technology **6**, 025020 (2021).

[41] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta, Phys. Rev. A **100**, 032328 (2019).

[42] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, *et al.*, Nature **574**, 505 (2019).

[43] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and J. M. Martinis, Nature **508**, 500 (2014).

[44] M. C. Collodo, J. Herrmann, N. Lacroix, C. K. Andersen, A. Remm, S. Lazar, J.-C. Besse, T. Walter, A. Wallraff, and C. Eichler, Phys. Rev. Lett. **125**, 240502 (2020).

[45] M. Ganzhorn, G. Salis, D. J. Egger, A. Fuhrer, M. Mergenthaler, C. Müller, P. Müller, S. Paredes, M. Pechal, M. Werninghaus, and S. Filipp, Phys. Rev. Research **2**, 033447 (2020).

[46] G. E. P. Box, J. Am. Stat. Assoc. **71**, 791 (1976).

[47] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, D. Sugny, and F. K. Wilhelm, Eur. Phys. J. D **69**, 279 (2015).

[48] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Gruslys, S. Schirmer, and T. Schulte-Herbrüggen, Phys. Rev. A **84**, 022305 (2011).

[49] S. Machnes, E. Assémat, D. Tannor, and F. K. Wilhelm, Phys. Rev. Lett. **120**, 150401 (2018).

[50] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).

[51] P. Doria, T. Calarco, and S. Montangero, Phys. Rev. Lett. **106**, 190501 (2011).

[52] R. Kosloff, S. A. Rice, P. Gaspard, S. Tersigni, and D. Tannor, Chem. Phys. **139**, 201 (1989).

[53] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Phys. Rev. Lett. **103**, 110501 (2009).

[54] R. Schutjens, F. A. Dagga, D. J. Egger, and F. K. Wilhelm, Phys. Rev. A **88**, 052330 (2013).

[55] U. Dorner, T. Calarco, P. Zoller, A. Browaeys, and P. Grangier, J. Opt. B: Quantum Semiclassical Opt. **7**, S341 (2005).

[56] D. J. Egger and F. K. Wilhelm, Supercond. Sci. Technol. **27**, 014001 (2013).

[57] D. J. Egger and F. K. Wilhelm, Phys. Rev. Lett. **112**, 240503 (2014).

[58] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, Phys. Rev. Lett. **112**, 240504 (2014).

[59] S. Sheldon, E. Magesan, J. M. Chow, and J. M. Gambetta, Phys. Rev. A **93**, 060302(R) (2016).

[60] S. Sheldon, L. S. Bishop, E. Magesan, S. Filipp, J. M. Chow, and J. M. Gambetta, Phys. Rev. A **93**, 012301 (2016).

[61] M. Gevers, European Journal of Control **11**, 335 (2005).

[62] C. R. Rojas, J. S. Welsh, G. C. Goodwin, and A. Feuer, Automatica **43**, 993 (2007).

[63] L. Ljung, Annual Reviews in Control **34**, 1 (2010).

[64] X. Bombois, M. Gevers, R. Hildebrand, and G. Solari, Communications in Information and Systems **11**, 197 (2011).

[65] Q.-M. Chen, X. Yang, C. Arenz, R.-B. Wu, X. Peng, I. Pelczer, and H. Rabitz, Phys. Rev. A **101**, 032313 (2020).

[66] R.-B. Wu, B. Chu, D. H. Owens, and H. Rabitz, Phys. Rev. A **97**, 042122 (2018).

[67] H. Ball, M. J. Biercuk, A. Carvalho, J. Chen, M. Hush, L. A. D. Castro, L. Li, P. J. Liebermann, H. J. Slatyer, C. Edmunds, V. Frey, C. Hempel, and A. Milne, (2020), arXiv:2001.04060 [quant-ph] .

[68] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).

[69] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, J. Math. Phys. **17**, 821 (1976).

[70] B. Girod, R. Rabenstein, and A. Stenger, *Signals and systems* (Wiley, 2001).

[71] M. Werninghaus, D. J. Egger, F. Roy, S. Machnes, F. K. Wilhelm, and S. Filipp, npj Quantum Information **7**, 14 (2021).

[72] Z. Chen, J. Kelly, C. Quintana, R. Barends, B. Campbell, Y. 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, and J. M. Martinis, Phys. Rev. Lett. **116**, 020501 (2016).

[73] E. Magesan, R. Blume-Kohout, and J. Emerson, Phys. Rev. A **84**, 012309 (2011).

[74] J. Nocedal, Math. Comput. **35**, 773 (1980).

[75] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Phys. Rev. A **77**, 012307 (2008).

[76] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. Lett. **102**, 090502 (2009).

[77] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. Lett. **106**, 180504 (2011).

[78] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. A **85**, 042311 (2012).

[79] N. Hansen, S. D. Müller, and P. Koumoutsakos, Evol. Comput. **11**, 1 (2003).

[80] M. Sinanovic, Evaluation of calibration algorithms for quantum gates, B. Sc. Thesis (unpublished).

[81] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **69**, 062320 (2004).

[82] J. J. Wallman and S. T. Flammia, New J. Phys. **16**, 103032 (2014).

[83] C. Rigetti and M. Devoret, Phys. Rev. B **81**, 134507 (2010).

[84] J. M. Chow, A. D. Córcoles, J. M. Gambetta, C. Rigetti, B. R. Johnson, J. A. Smolin, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Phys. Rev. Lett. **107**, 080502 (2011).

[85] T. Alexander, N. Kanazawa, D. J. Egger, L. Capelluto, C. J. Wood, A. Javadi-Abhari, and D. C McKay, Quantum Science and Technology **5**, 044006 (2020).

[86] A. D. Córcoles, J. M. Gambetta, J. M. Chow, J. A. Smolin, M. Ware, J. Strand, B. L. T. Plourde, and M. Steffen, Phys. Rev. A **87**, 030301(R) (2013).

[87] S. Kirchhoff, T. Keßler, P. J. Liebermann, E. Assémat, S. Machnes, F. Motzoi, and F. K. Wilhelm, Phys. Rev. A **97**, 042348 (2018).

[88] F. K. Wilhelm, S. Kirchhoff, S. Machnes, N. Wittler, and D. Sugny, (2020), arXiv:2003.10132 .

[89] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, and M. Isard, (2015), software available from tensorflow.org.

[90] G. Kedem, ACM Transactions on Mathematical Software (TOMS) **6**, 150 (1980).

[91] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, Phys. Rev. A **95**, 042318 (2017).

[92] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Phys. Rev. A **96**, 022330 (2017).

[93] S. V. Stehman, Remote Sens. Environ. **62**, 77 (1997).

[94] A. P. Peirce, M. A. Dahleh, and H. Rabitz, Phys. Rev. A **37**, 4950 (1988).

[95] D. M. Reich, M. Ndong, and C. P. Koch, The Journal of Chemical Physics **136**, 104103 (2012).

[96] M. H. Goerz, D. Basilewitsch, F. Gago-Encinas, M. G. Krauss, K. P. Horn, D. M. Reich, and C. P. Koch, SciPost Phys. **7**, 80 (2019).

[97] Z. Leng, P. Mundada, S. Ghadimi, and A. Houck, (2019), arXiv:1910.03591 .

[98] *Numerical Optimization*, Springer Series in Operations Research and Financial Engineering (Springer New York, 2006).

[99] G. Feng, F. H. Cho, H. Katiyar, J. Li, D. Lu, J. Baugh, and R. Laflamme, Phys. Rev. A **98**, 052341 (2018).

[100] J. Kelly, P. O'Malley, M. Neeley, H. Neven, and J. M. Martinis, (2018), arXiv:1803.03226 .

[101] P. C. Mahalanobis (National Institute of Science of India, 1936) p. 49–55.

[102] S. G. Schirmer and D. K. L. Oi, Laser Phys. **20**, 1203 (2010).

[103] M. P. V. Stenberg, Y. R. Sanders, and F. K. Wilhelm, Phys. Rev. Lett. **113**, 210404 (2014).

[104] M. P. V. Stenberg, O. Köhn, and F. K. Wilhelm, Phys. Rev. A **93**, 012122 (2016).

[105] M. P. V. Stenberg and F. K. Wilhelm, Phys. Rev. A **94**, 052119 (2016).

[106] A. Sergeevich, A. Chandran, J. Combes, S. D. Bartlett, and H. M. Wiseman, Phys. Rev. A **84**, 052315 (2011).

[107] N. Wiebe, C. Granade, C. Ferrie, and D. G. Cory, Phys. Rev. Lett. **112**, 190501 (2014).

[108] N. Wiebe, C. Granade, C. Ferrie, and D. Cory, Phys. Rev. A **89**, 042314 (2014).

[109] N. Wiebe, C. Granade, and D. G. Cory, New J. Phys. **17**, 022005 (2015).

[110] J. Wang, S. Paesani, R. Santagati, S. Knauer, A. A. Gentile, N. Wiebe, M. Petruzzella, J. L. O'Brien, J. G. Rarity, A. Laing, *et al.*, Nat. Phys. **13**, 551 (2017).

[111] J. H. Cole, S. G. Schirmer, A. D. Greentree, C. J. Wellard, D. K. L. Oi, and L. C. L. Hollenberg, Phys. Rev. A **71**, 062312 (2005).

[112] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory, New J. Phys. **14**, 103013 (2012).

[113] M. Krenn, M. Malik, R. Fickler, R. Lapkiewicz, and A. Zeilinger, Phys. Rev. Lett. **116**, 090405 (2016).

[114] A. A. Melnikov, H. P. Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, Proceedings of the National Academy of Sciences **115**, 1221 (2018).

[115] M. Krenn, M. Erhard, and A. Zeilinger, Computer-inspired quantum experiments (2020), arXiv:2002.09970 [quant-ph] .

[116] A. A. Gentile, B. Flynn, S. Knauer, N. Wiebe, S. Paesani, C. E. Granade, J. G. Rarity, R. Santagati, and A. Laing, (2020), arXiv:2002.06169 .

[117] T. J. Evans, R. Harper, and S. T. Flammia, (2019), arXiv:1912.07636 .

[118] J. F. Ralph, S. Maskell, and K. Jacobs, Phys. Rev. A **96**, 052306 (2017).

[119] R. L. Kosut and H. Rabitz, IFAC Proceedings Volumes **35**, 397 (2002).

[120] S. G. Schirmer, A. Kolli, and D. K. L. Oi, Phys. Rev. A **69**, 050306(R) (2004).

[121] D. K. L. Oi and S. G. Schirmer, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **370**, 5386 (2012).

[122] D. Burgarth and K. Yuasa, Phys. Rev. Lett. **108**, 080502 (2012).

[123] J. Zhang and M. Sarovar, Phys. Rev. A **91**, 052121 (2015).

[124] S.-Y. Hou, H. Li, and G.-L. Long, Science Bulletin **62**, 863 (2017).

[125] A. Shabani, M. Mohseni, S. Lloyd, R. L. Kosut, and H. Rabitz, Phys. Rev. A **84**, 012107 (2011).

[126] K. Rudinger and R. Joynt, Phys. Rev. A **92**, 052322 (2015).

[127] B. Lusch, J. N. Kutz, and S. L. Brunton, Nat. Commun. **9**, 1 (2018).

[128] E. Flurin, L. S. Martin, S. Hacohen-Gourgy, and I. Siddiqi, Phys. Rev. X **10**, 011006 (2020).

[129] M. Dalgaard, F. Motzoi, J. J. Sørensen, and J. Sherson, npj Quantum Inf. **6**, (2020).

[130] S. Krastanov, S. Zhou, S. T. Flammia, and L. Jiang, Quantum Science and Technology **4**, 035003 (2019).

[131] I. I. Rabi, Phys. Rev. **51**, 652 (1937).

[132] N. F. Ramsey, Phys. Rev. **78**, 695 (1950).

[133] E. L. Hahn, Phys. Rev. **80**, 580 (1950).

[134] B. Pokharel, N. Anand, B. Fortman, and D. A. Lidar, Phys. Rev. Lett. **121**, 220502 (2018).

[135] C. Ferrie, C. Granade, G. Paz-Silva, and H. M. Wiseman, New J. Phys. **20**, 123005 (2018).

[136] M. A. Rol, L. Ciorciaro, F. K. Malinowski, B. M. Tarasinski, R. E. Sagastizabal, C. C. Bultink, Y. Salathe, N. Haandbaek, J. Sedivy, and L. DiCarlo, Appl. Phys. Lett. **116**, 054001 (2020).

[137] J. P. G. van Dijk, E. Kawakami, R. N. Schouten, M. Veldhorst, L. M. K. Vandersypen, M. Babaie, E. Charbon, and F. Sebastiano, Phys. Rev. Applied **12**, 044054 (2019).

[138] M. Jerger, A. Kulikov, Z. Vasselin, and A. Fedorov, Phys. Rev. Lett. **123**, 150501 (2019).

[139] I. N. Hincks, C. E. Granade, T. W. Borneman, and D. G. Cory, Phys. Rev. Applied **4**, 024012 (2015).

[140] S. Gustavsson, O. Zwier, J. Bylander, F. Yan, F. Yoshihara, Y. Nakamura, T. P. Orlando, and W. D. Oliver, Phys. Rev. Lett. **110**, 040502 (2013).

[141] R. S. Gupta, A. R. Milne, C. L. Edmunds, C. Hempel, and M. J. Biercuk, (2019), arXiv:1904.07225 .

[142] R. Harper, S. T. Flammia, and J. J. Wallman, (2019), arXiv:1907.13022 .

[143] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, *et al.*, Phys. Rev. A **94**, 032321 (2016).

[144] L. Casparis, T. W. Larsen, M. S. Olsen, F. Kuemmeth, P. Krogstrup, J. Nygård, K. D. Petersson, and C. M. Marcus, Phys. Rev. Lett. **116**, 150505 (2016).

[145] C. Miquel, J. P. Paz, and W. H. Zurek, Phys. Rev. Lett. **78**, 3971 (1997).

[146] R. C. Bialczak, R. McDermott, M. Ansmann, M. Hofheinz, N. Katz, E. Lucero, M. Neeley, A. D. O'Connell, H. Wang, A. N. Cleland, and J. M. Martinis, Phys. Rev. Lett. **99**, 187006 (2007).

[147] J. J. Burnett, A. Bengtsson, M. Scigliuzzo, D. Niepce, M. Kudra, P. Delsing, and J. Bylander, npj Quantum Inf. **5**, 1 (2019).

[148] M. Y. Niu, V. Smelyanskyi, P. Klimov, S. Boixo, R. Barends, J. Kelly, Y. Chen, K. Arya, B. Burkett, D. Bacon, Z. Chen, B. Chiaro, R. Collins, A. Dunsworth, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, T. Huang, E. Jeffrey, D. Landhuis, E. Lucero, A. Megrant, J. Mutus, X. Mi, O. Naaman, M. Neeley, C. Neill, C. Quintana, P. Roushan, J. M. Martinis, and H. Neven, (2019), arXiv:1912.04368 .

[149] A. Zhang, J. Xie, H. Xu, K. Zheng, H. Zhang, Y.-T. Poon, V. Vedral, and L. Zhang, Phys. Rev. Lett. **124**, 040402 (2020).

[150] P. Mundada, G. Zhang, T. Hazard, and A. Houck, Phys. Rev. Appl **12**, 054023 (2019).

[151] P. Murali, D. C. McKay, M. Martonosi, and A. Javadi-Abhari, (2020), arXiv:2001.02826 .

[152] D. Burgarth, K. Maruyama, and F. Nori, Phys. Rev. A **79**, 020305(R) (2009).

[153] C. Di Franco, M. Paternostro, and M. S. Kim, Phys. Rev. Lett. **102**, 187203 (2009).

[154] L. E. De Clercq, R. Oswald, C. Flühmann, B. Keitch, D. Kienzler, H.-Y. Lo, M. Marinelli, D. Nadlinger, V. Negnevitsky, and J. P. Home, Nat. Commun. **7**, 1 (2016).

[155] A. Sone and P. Cappellaro, Phys. Rev. A **95**, 022335 (2017).

[156] P. Yang, M. Yu, R. Betzholz, C. Arenz, and J. Cai, Phys. Rev. Lett. **124**, 010405 (2020).

[157] M. A. Rol, C. Dickel, S. Asaad, N. K. Langford, C. C. Bultink, R. Sagastizabal, N. K. Langford, G. de Lange, X. Fu, S. R. de Jong, F. Luthi, W. Vlothuizen, and M. S. Moreira, Pycqed_py3 (2019).

[158] Labber.org, Labber – software for instrument control and lab automation (2020).

[159] C. J. Kalkman, J. Clin. Monit. **11**, 51 (1995).

[160] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Nature **323**, 533 (1986).

[161] M. Abdelhafez, D. I. Schuster, and J. Koch, Phys. Rev. A **99**, 052327 (2019).

[162] D. Willsch, M. Nocon, F. Jin, H. De Raedt, and K. Michielsen, Phys. Rev. A **96**, 062302 (2017).

[163] D. Willsch, M. Willsch, F. Jin, H. De Raedt, and K. Michielsen, Phys. Rev. A **98**, 052348 (2018).

[164] D. Willsch, (2020), arXiv:2008.13490 .

[165] Qruise, https://qruise.com.

[166] J. Kunasaikaran, K. Mato, and R. Wille, A Framework for the Design and Realization of Alternative Superconducting Quantum Architectures (2023), arXiv:2305.07052 [quant-ph] .

[167] X. Ni, H.-H. Zhao, L. Wang, F. Wu, and J. Chen, npj Quantum Information **8**, 7 (2022).

[168] T. Rajabzadeh, A. Boulton-McKeehan, S. Bonkowsky, D. I. Schuster, and A. H. Safavi-Naeini, A General Framework for Gradient-Based Optimization of Superconducting Quantum Circuits using Qubit Discovery as a Case Study (2024), arXiv:2408.12704 [cond-mat, physics:physics, physics:quant-ph] .

[169] C. Guinn, S. Stein, E. Tureci, G. Avis, C. Liu, S. Krastanov, A. A. Houck, and A. Li, Co-Designed Superconducting Architecture for Lattice Surgery of Surface Codes with Quantum Interface Routing Card (2023), arXiv:2312.01246 [quant-ph] .

[170] T. Menke, F. Häse, S. Gustavsson, A. J. Kerman, W. D. Oliver, and A. Aspuru-Guzik, npj Quantum Information **7**, 8 (2021).

[171] K. X. Wei, I. Lauer, E. Pritchett, W. Shanks, D. C. McKay, and A. Javadi-Abhari, PRX Quantum **5**, 020338 (2024), arXiv:2310.12146 [quant-ph] .

[172] M. H. Goerz, G. Gualdi, D. M. Reich, C. P. Koch, F. Motzoi, K. B. Whaley, J. Vala, M. M. Müller, S. Montangero, and T. Calarco, Physical Review A **91**, (2015).

[173] S. F. Lin, S. Sussman, C. Duckering, P. S. Mundada, J. M. Baker, R. S. Kumar, A. A. Houck, and F. T. Chong, in *55th IEEE/ACM International Symposium on Microarchitecture, MICRO 2022, Chicago, IL, USA, October 1-5, 2022* (IEEE, 2022) pp. 1042–1058, arXiv:2208.13380 [quant-ph] .

[174] P. Kairys and T. S. Humble, in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, Broomfield, CO, USA, 2021) pp. 413–418.

[175] Google AI Quantum *et al.*, Physical Review Letters **125**, 120504 (2020).

[176] F. Setiawan, P. Groszkowski, and A. A. Clerk, Physical Review Applied **19**, 034071 (2023).

[177] M. H. Goerz, F. Motzoi, K. B. Whaley, and C. P. Koch, npj Quantum Information **3**, 1 (2017).

[178] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Nature **460**, 240 (2009).

[179] J. M. Chow, A. D. Córcoles, J. M. Gambetta, C. Rigetti, B. R. Johnson, J. A. Smolin, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Physical Review Letters **107**, 080502 (2011).

[180] J. Ghosh, A. Galiautdinov, Z. Zhou, A. N. Korotkov, J. M. Martinis, and M. R. Geller, Physical Review A **87**, 022309 (2013).

[181] P. Watts, J. Vala, M. M. Müller, T. Calarco, K. B. Whaley, D. M. Reich, M. H. Goerz, and C. P. Koch, Physical Review A **91**, (2015).

[182] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Physical Review A **67**, 042313 (2003).

[183] D. C. McKay, S. Filipp, A. Mezzacapo, E. Magesan, J. M. Chow, and J. M. Gambetta, Physical Review Applied **6**, 064007 (2016).

[184] M. A. Rol, F. Battistel, F. K. Malinowski, C. C. Bultink, B. M. Tarasinski, R. Vollmer, N. Haider, N. Muthusubramanian, A. Bruno, B. M. Terhal, and L. DiCarlo, Physical Review Letters **123**, 120502 (2019).

[185] P. Mundada, G. Zhang, T. Hazard, and A. Houck, Physical Review Applied **12**, 054023 (2019).

[186] A. Kandala, K. X. Wei, S. Srinivasan, E. Magesan, S. Carnevale, G. A. Keefe, D. Klaus, O. Dial, and D. C. McKay, Physical Review Letters **127**, 130501 (2021).

[187] F. Bao, H. Deng, D. Ding, R. Gao, X. Gao, C. Huang, X. Jiang, H.-S. Ku, Z. Li, X. Ma, X. Ni, J. Qin, Z. Song, H. Sun, C. Tang, T. Wang, F. Wu, T. Xia, W. Yu, F. Zhang, G. Zhang, X. Zhang, J. Zhou, X. Zhu, Y. Shi, J. Chen, H.-H. Zhao, and C. Deng, Physical Review Letters **129**, 010502 (2022).