Supercomputer simulations of transmon quantum computers

Dennis Willsch

JÜLICH
Forschungszentrum

Supercomputer simulations of transmon quantum computers

Dennis Willsch

Abstract

We develop a simulator for quantum computers composed of superconducting transmon qubits. The simulation model supports an arbitrary number of transmons and resonators. Quantum gates are implemented by time-dependent pulses. Nontrivial effects such as crosstalk, leakage to non-computational states, entanglement between transmons and resonators, and control errors due to the pulses are inherently included.

The time evolution of the quantum computer is obtained by solving the time-dependent Schrödinger equation. The simulation algorithm shows excellent scalability on high-performance supercomputers. We present results for the simulation of up to 16 transmons and resonators. Additionally, the model can be used to simulate environments, and we demonstrate the transition from an isolated system to an open quantum system governed by a Lindblad master equation. We also describe a procedure to extract model parameters from electromagnetic simulations or experiments.

We compare simulation results to experiments on several NISQ processors of the IBM Q Experience. We find nearly perfect agreement between simulation and experiment for quantum circuits designed to probe crosstalk in transmon systems. By studying common gate metrics such as the fidelity or the diamond distance, we find that they cannot reliably predict the performance of repeated gate applications or practical quantum algorithms. As an alternative, we find that the results from two-transmon gate set tomography have an exceptional predictive power. Finally, we test a protocol from the theory of quantum error correction and fault tolerance. We find that the protocol systematically improves the performance of transmon quantum computers in the presence of characteristic control and measurement errors.

Zusammenfassung

Wir entwickeln einen Simulator für Quantencomputer, die aus supraleitenden Transmon-Qubits bestehen. Das Simulationsmodell unterstützt eine beliebige Anzahl von Transmons und Resonatoren. Quantengatter werden durch zeitabhängige Pulse realisiert. Nichttriviale Effekte wie Crosstalk, Verlust in nicht rechnerische Zustände, Verschränkung zwischen Transmons und Resonatoren sowie Steuerungsfehler verursacht durch die Pulse sind automatisch miteinbezogen.

Die Zeitentwicklung des Quantencomputers wird durch Lösung der zeitabhängigen Schrödingergleichung bestimmt. Der Simulationsalgorithmus zeigt ausgezeichnete Skalierbarkeit auf Hochleistungs-Supercomputern. Wir präsentieren Ergebnisse für die Simulation von bis zu 16 Transmons und Resonatoren. Zusätzlich kann das Modell zur Simulation von Umgebungen verwendet werden. Wir demonstrieren den Übergang von einem isolierten System zu einem offenen Quantensystem, das von einer Lindblad-Mastergleichung bestimmt wird. Wir beschreiben außerdem ein Verfahren zur Extraktion von Modellparametern aus elektromagnetischen Simulationen oder Experimenten.

Wir vergleichen Simulationsergebnisse mit Experimenten auf mehreren NISQ-Prozessoren der IBM Q Experience. Wir finden eine nahezu perfekte Übereinstimmung zwischen Simulation und Experiment für Quantenschaltungen zur Untersuchung von Crosstalk in Transmon-Systemen. Durch Untersuchung gängiger Gatter-Metriken wie der Fidelity oder der Diamant-Distanz finden wir, dass sie die Leistung von wiederholten Gatteranwendungen oder praktischen Quantenalgorithmen nicht zuverlässig vorhersagen können. Als Alternative finden wir, dass die Ergebnisse einer Zwei-Transmon-Gattermengentomographie eine außergewöhnlich gute Vorhersagekraft aufweisen. Zum Schluss testen wir ein Protokoll aus der Theorie der Quantenfehlerkorrektur und Fehlertoleranz. Wir stellen fest, dass das Protokoll systematisch die Leistung von Transmon-Quantencomputern bei charakteristischen Steuerungs- und Messfehlern verbessert.

Contents

Chapter 1

Introduction

For over a century, humans have designed and built digital computing machines. The initial ideas can be traced back to the mid-1800s [Bab1837; Boo1847], but the actual construction started less than a century ago. In 1936, Zuse designed a floating point general-purpose computer [Zus1936] that led to the first programmable floating point machine in 1941, the Z3 [Cop2017]. Turing formalized the *universal computing machine* [Tur1937] that influenced the construction of the Colossus in 1943 [Ran1973], which was used to perform Boolean operations for cryptanalysis. Other computers of that time were the ABC [Ata1940] built in 1942 and the ENIAC built in 1945 [Ran1973].

Most of these early computers were based on vacuum tubes which made them large and unreliable. Universal digital computing only became *scalable* after the vacuum tubes were replaced by semiconductor devices such as transistors. Nowadays, computers are ubiquitous in everyday life; every mobile phone contains a general-purpose digital computer, and large-scale high-performance supercomputers are used routinely to solve some of the most difficult computational problems.

Similarly, for about a century, humans have developed and studied quantum theory. This physical theory has extraordinary descriptive power, also in numerous fields beyond physics [Khr2010]. The predictions of quantum theory are fundamentally stochastic, meaning that quantum theory can only predict probabilities for observable events [Bal1998]. In this sense, quantum theory is inherently linked to probability theory [Jay2003]. The mathematical framework of quantum theory is based on linear algebra and can be reduced to a few axioms [Neu1955; Bal1998; Nie2010].

Quantum computing

The essential idea of a *quantum computer* is to combine these two concepts, i.e., the universal computing machine and quantum theory. The goal is to build a computing machine that implements the equations of quantum theory for two-level systems. The elementary two-level systems of digital computers, the bits, are replaced by quantum bits, commonly known as *qubits*. As quantum theory only predicts probabilities, a program for a quantum computer basically determines a set of probabilities for the qubits.

An actual device, however, does not produce probabilities but individual bits. Therefore, a program is typically repeated multiple times to make the connection to the predicted probabilities. This is a common principle of all types of quantum computing: From

the user perspective, a quantum computer produces a large ensemble of individual results.

Two basic approaches to quantum computing are currently pursued by science and industry [Nat2019]. The first is called the *gate-based quantum computer*. Inspired by the gate model of digital computing, programs for a gate-based quantum computer are specified in terms of elementary *quantum gates* [Nie2010]. As every algorithm can be decomposed into a sequence of such elementary gates, the gate-based quantum computer is considered universal [Bar1995; Deu1995; DiV1995].

The other approach to quantum computing is *quantum annealing* [Fin1994; Kad1998; Far2000; Har2010; Joh2011]. Quantum annealers have turned out to be very useful for quickly producing an ensemble of close-to-optimal solutions to a given optimization or machine learning problem [Pud2012; Per2019; Orú2019; Wil2020b].

Over the past decades, research in gate-based quantum computing has evolved from an abstract, mathematical model of a computing machine [Ben1980; Deu1985] to a broad range of experimental devices. All of these pursue the idea of implementing the mathematical framework of quantum theory to gain a computational advantage over the mature technology of digital computers [Eke1996]. The strongest advantage is envisioned as an exponential speedup for a special set of mathematical problems, such as (1) the simulation of quantum mechanical systems which is believed to significantly aid in research and development [Fey1982; Bab2018; Küh2019], (2) the approximate solution of sparse linear systems in logarithmic time [Har2009] (albeit with a few caveats [Aar2015]), or (3) the polynomial-time factorization of integers [Sho1994; Sho1997] which might form a potential threat to the security of widely-used asymmetric cryptosystems such as RSA [Riv1978].

Currently, the two most advanced technologies for gate-based quantum computers use superconducting circuits [Rai2001; Vio2002; Bla2004; Wen2017] and trapped ions [Cir1995; Mon1995]. IBM and several other companies have made small superconducting quantum processors available to the community to explore the technology [IBM2016; Rig2017; Chi2018b; DWa2019; LaR2019]. Additionally, a superconducting quantum processor manufactured by Google has produced results for a well-defined class of problems that are beyond the reach of digital supercomputers [Aru2019], thereby achieving *quantum supremacy* [Pre2012; Boi2018].

Objectives

Despite the recent progress, it is still an open question if a universal, fully error-corrected quantum computer can be built. All current quantum processors belong to the class of *noisy intermediate-scale quantum* (NISQ) devices [Pre2018; Nat2019]. And although a fully error-corrected device is possible in theory [Sho1996; Aha1997; Ali2007], there is a priori no guarantee that it can also be built in practice. The essential questions are: How close do current NISQ devices come to the ideal, mathematical qubit model of a quantum computer? What are the main errors and limitations that would need to be overcome?

To address these questions, we carry out detailed supercomputer simulations of current NISQ devices. Additionally, we perform experiments on such devices to compare simulation results with experimental observations. We identify and analyze the main limitations, i.e., *leakage* and *crosstalk*, and study to what extent the induced errors can be corrected.

Supercomputer simulations

The aim of this thesis is to utilize the power of digital supercomputers to study the emerging technology of quantum computers. Supercomputer simulations are vital for the development and verification of quantum computers. For instance, massively parallel simulators such as the Jülich universal quantum computer simulator (JUQCS) [DeR2007; DeR2019a; Wil2020a] have been essential to verify Google's demonstration of quantum supremacy [Aru2019].

In this work, we develop a simulator of superconducting quantum processors. We focus on gate-based quantum computers with superconducting transmon qubits [Koc2007] because of the tremendous progress that has been reported recently for the transmon architecture [IBM2016; Rig2017; Chi2018b; Aru2019]. We devise a scalable method to simulate transmon systems with an arbitrary number of qubits and couplers, limited only by the available amount of physical memory and computing time. The simulation model includes the effects of higher transmon levels and generic time-dependent pulses. All model and pulse parameters are freely configurable.

Simulation method

The simulator solves the time-dependent Schrödinger equation (TDSE) with $\hbar = 1$,

$$i\frac{\partial}{\partial t}\left|\Psi(t)\right\rangle = H(t)\left|\Psi(t)\right\rangle, \tag{1.1}$$

where $H(t)$ is a generic, time-dependent model Hamiltonian representing the hardware of the quantum processor, including the transmon system and their electromagnetic environment.

From the solution $\left|\Psi(t)\right\rangle$, we can compute any physically relevant quantity of the system such as reduced density matrices with non-unitary dynamics describing the actual qubits. In this sense, the TDSE approach can be related to other common approaches based on master equations, perturbative studies, and completely positive trace-preserving maps. We study each of these relations in the course of this thesis.

The time dependence of $H(t)$ in Eq. (1.1) represents external microwave control pulses that are applied to the system. Each pulse is a time-dependent voltage signal designed to implement a certain quantum gate. We use an optimization procedure to find optimal pulse parameters for each simulated transmon system.

The simulator is based on the Suzuki-Trotter product-formula algorithm [DeR1987; DeR2000; DeR2006]. This allows the TDSE given by Eq. (1.1) to be solved on a sub-picosecond scale for time evolutions over several hundred microseconds, without making additional approximations. We formulate the TDSE in an appropriate basis that makes its solution amenable to large-scale supercomputer simulations. Most of the simulations presented in this thesis were performed on the supercomputers JURECA [Jül2018] and JUWELS [Jül2019].

Outline

In this thesis, we present results from the simulation of transmon systems with up to 16 transmons and couplers. As real NISQ devices of the same size and architecture are publicly accessible on the IBM Q Experience [IBM2016], we also perform some of the experiments on these quantum processors. This offers a great opportunity to relate the simulation results directly to experiments. We find that the main limitations revealed by the simulation, i.e., leakage and crosstalk due to additional transmon and resonator states, capture most of the errors observed in the corresponding NISQ devices.

This thesis is organized as follows. Chapter 2 reviews the mathematical model of a gate-based quantum computer, i.e., qubits, gates, circuits, quantum operations, and leakage. In Chapter 3, we define the supercomputer simulation method. After specifying the full model Hamiltonian, we describe in detail the numerical algorithm to solve the TDSE given by Eq. (1.1). We then define the primary model systems. Finally, we present a method to obtain model parameters for the simulation of electromagnetic environments.

Chapter 4 focuses on free, undriven time evolutions. This includes accuracy and performance benchmarks. Additionally, we relate our simulation approach to perturbative results and simulations of a Lindblad master equation.

In Chapter 5, we define the elementary single-qubit and two-qubit pulses used to implement quantum gates. We describe the optimization procedure used to find optimal pulse parameters and present optimization results for the larger transmon systems. Finally, we discuss the compilation process used to translate quantum circuits to pulse information for the simulator.

In Chapter 6, we characterize the optimized quantum gates in detail. After introducing the most prominent gate metrics, we study repeated gate applications on both the simulated systems and experimental devices. The chapter concludes with an application of gate set tomography (GST) and an assessment of its predictive power.

Chapter 7 combines the results from the previous chapters and applies them to a selected class of quantum circuits, executed using both the transmon simulator and experimental processors. We first design and implement a class of quantum circuits to study crosstalk in transmon systems. Secondly, we study quantum circuits designed to characterize the singlet state. Finally, we implement a full protocol from the theory of quantum fault tolerance to assess its potential to improve quantum computation in transmon architectures.

Chapter 8 contains our conclusions and an outlook on many interesting paths to continue the present work. Implementation details and separate mathematical proofs are given in Appendices A–J. Some results presented in this thesis have previously been published in [Wil2017; Wil2018b].

Chapter 2

Ideal gate-based quantum computing

In this chapter, we review the computational architecture of an ideal gate-based quantum computer and related concepts as formally defined in the literature [Yan2008; Nie2010; Wat2018]. We start by introducing the quantum bit as the fundamental unit of computation in Section 2.1. In Section 2.2, we define quantum gates as the basic operations that can be performed on a qubit. A combination of these operations is called a quantum circuit, which is introduced in Section 2.3. Qubits, quantum gates, and quantum circuits are the basic building blocks that are required to define algorithms for a gate-based quantum computer. Finally, in Section 2.4, we introduce the concept of quantum operations, which are used in a more general description of gate-based quantum computers in terms of mixed states.

Although it is still unclear when a large universal gate-based quantum computer can be built or if the envisioned exponential speedup can be delivered, the mathematical model of an ideal quantum computer is an interesting model to study. Advances in quantum algorithms with a theoretical speedup can inspire valuable discoveries in other areas. For instance, a quantum algorithm for recommendation systems (i.e., systems that are supposed to provide product suggestions to users based on past purchases) is known to provide an exponential speedup over previous classical algorithms [Ker2016]. Recently, this discovery has led to the development of novel, similarly efficient algorithms for digital computers [Tan2019]. Further examples of such *quantum-inspired* algorithms are given in [Tan2018; Gil2018; Chi2018a].

2.1 Quantum bits

A *quantum bit* or *qubit* is the generalization of a digital bit to the mathematical framework of quantum theory. The goal of this section is to understand the precise meaning of *generalization* in this context. This helps to understand why quantum computers have the theoretical potential to be more powerful than digital computers. We approach the concept of generalization by first defining the digital bit, and then highlighting the mathematical difference to the qubit.

2.1.1 Single qubits

A (digital) bit j is the fundamental unit of computation in every digital computer. The name "bit" stands for *binary digit*. It means that, at each point in a computation, j can only be either 0 or 1, i.e., $j \in \{0, 1\}$. If n bits are combined, the state of the computation is described by a bit string $J = j_0 \cdots j_{n-1}$. Again, each bit can only be either 0 or 1. The algebraic structure to describe all possible n-bit states J is given by the n-fold Cartesian product

$$J = j_0 \cdots j_{n-1} \in \{0, 1\}^n = \{0 \cdots 00, 0 \cdots 01, \ldots, 1 \cdots 11\}, \tag{2.1}$$

which has a finite number of elements $|\{0, 1\}^n| = 2^n$. A computation can be formally expressed as a function $f : \{0, 1\}^n \to \{0, 1\}^m$ for $m, n \in \mathbb{N}$. This formalism is known as Boolean algebra and was first introduced in [Boo1847].

A qubit $|\psi\rangle$, on the other hand, is defined as a two-level quantum system determined by two complex numbers $a_0, a_1 \in \mathbb{C}$ with $|a_0|^2 + |a_1|^2 = 1$. It is commonly written as

$$|\psi\rangle = a_0 |0\rangle + a_1 |1\rangle = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}, \tag{2.2}$$

where the *computational basis states* $|0\rangle$ and $|1\rangle$ represent the digital bit states 0 and 1, respectively. Thus, a qubit is not always either $|0\rangle$ or $|1\rangle$, but rather any complex linear combination of both. This concept is usually called *superposition* and is the first reason why a qubit can be considered a generalized bit.

If it were only this generalization, it would be easy to imagine why the computational model of quantum computing might have an advantage over digital computing. However, the drawback of this computational model is that the complex coefficients a_0 and a_1 defining the state given by Eq. (2.2) cannot be observed directly. Instead, as dictated by quantum theory, they only define probabilities $p_0 = |a_0|^2$ ($p_1 = |a_1|^2$) to observe the qubit in the binary state 0 (1). Such an observation is called *measurement*. The complex coefficients are correspondingly called *probability amplitudes*. The interpretation in the context of probability theory also requires that the amplitudes be normalized such that $|a_0|^2 + |a_1|^2 = 1$.

Thus, a constraint of the computational model of quantum computing is that, after a measurement, the state of the qubit is reset to the observed binary state. This mathematical constraint is the core of why algorithms for quantum computers have been notoriously hard to find, and that only a few key algorithms with a considerable theoretical speedup have been found [Aar2008].

2.1.2 Bloch sphere

The fact that the complex coefficients a_0 and a_1 in Eq. (2.2) are probability amplitudes can be used to find a convenient parametrization of a general single-qubit state: From $|a_0|^2 + |a_1|^2 = 1$, we know that there is an angle $\vartheta \in [0, \pi]$ such that $|a_0| = \cos(\vartheta/2)$ and $|a_1| = \sin(\vartheta/2)$. Furthermore, quantum theory states that the global phase of any state $|\psi\rangle$ is not observable, so we can set $a_0 = \cos(\vartheta/2) \geq 0$ and $a_1 = \exp(i\varphi) \sin(\vartheta/2)$, where
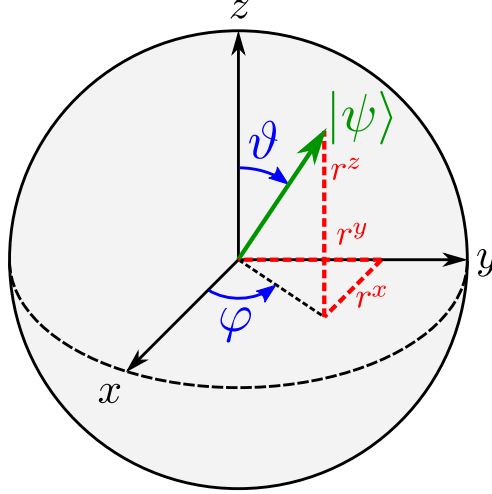
Figure 2.1: Bloch sphere representation of a pure single-qubit state $|\psi\rangle$. The azimuthal angle $\vartheta \in [0, \pi]$ and the polar angle $\varphi \in [0, 2\pi)$ are defined in Eq. (2.3), and the Cartesian coordinates r^x, r^y, and r^z are given by Eq. (2.5)

(a) state $|0\rangle$ (b) state $|1\rangle$ (c) state $|0\rangle + |1\rangle$ (d) state $|0\rangle + i|1\rangle$
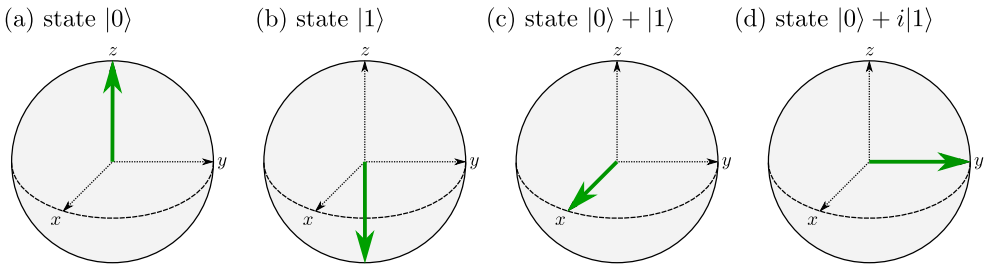
Figure 2.2: Bloch sphere representations of a few single-qubit states corresponding to eigenstates of the Pauli matrices defined in Eq. (2.4): (a) $+1$ eigenstate of σ^z, (b) -1 eigenstate of σ^z, (c) $+1$ eigenstate of σ^x, and (d) $+1$ eigenstate of σ^y. For simplicity, the normalization factor of $1/\sqrt{2}$ has been left out in the labels of (c) and (d).

$\varphi \in [0, 2\pi)$ encodes the relative phase between a_0 and a_1. The general single-qubit state $|\psi\rangle$ given by Eq. (2.2) thus becomes

$$|\psi\rangle = \cos\frac{\vartheta}{2}|0\rangle + e^{i\varphi}\sin\frac{\vartheta}{2}|1\rangle. \tag{2.3}$$

Due to the domain of the angles $\vartheta \in [0, \pi]$ and $\varphi \in [0, 2\pi)$, the single-qubit state $|\psi\rangle$ can be visualized on the surface of a three-dimensional sphere called *Bloch sphere* as shown in Fig. 2.1. Consequently, the representation defined by Eq. (2.3) is called the *Bloch sphere representation* and the three-dimensional vector $\vec{r}$ representing $|\psi\rangle$ on this sphere is called the *Bloch vector*. Its Cartesian coordinates r^x, r^y, and r^z can be calculated from the expectation values of $|\psi\rangle$ with respect to the Pauli matrices

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.4}$$

yielding

$$\vec{r} = \begin{pmatrix} r^x \\ r^y \\ r^z \end{pmatrix} = \begin{pmatrix} \langle\psi|\sigma^x|\psi\rangle \\ \langle\psi|\sigma^y|\psi\rangle \\ \langle\psi|\sigma^z|\psi\rangle \end{pmatrix} = \begin{pmatrix} 2\operatorname{Re} a_0^* a_1 \\ 2\operatorname{Im} a_0^* a_1 \\ |a_0|^2 - |a_1|^2 \end{pmatrix} = \begin{pmatrix} \sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta \end{pmatrix}. \tag{2.5}$$

The three Pauli matrices given in Eq. (2.4) are unitary, traceless, Hermitian matrices with eigenvalues ± 1. The corresponding eigenstates define the unit axes of the coordinate system of the Bloch sphere shown in Fig. 2.1. In particular, the eigenstates of σ^z are the computational basis states $|0\rangle$ and $|1\rangle$, whose Bloch vectors are given by the positive and the negative z axis, respectively (see Fig. 2.2(a) and (b)). The eigenstates of σ^x are denoted by $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ with Bloch vectors lying on the x axis, and the eigenstates of σ^y are given by $|\pm i\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$ with Bloch vectors lying on the y axis. Both σ^x and σ^y eigenstates corresponding to the eigenvalue $+1$ are visualized in Fig. 2.2(c) and (d), respectively.

2.1.3 Multiple qubits

The second reason why qubits can be seen as generalized bits becomes apparent when multiple qubits are combined. While an n-bit state in a digital computer is an element of the n-fold Cartesian product (see Eq. (2.1)), an n-qubit state in a quantum computer is an element of the n-fold *tensor product*. This means that an n-qubit state is a superposition of all 2^n bit states given in Eq. (2.1). We can therefore write an arbitrary multi-qubit state $|\psi\rangle$ as

$$|\psi\rangle = \sum_{j_0\cdots j_{n-1}\in\{0,1\}^n} a_{j_0\cdots j_{n-1}}|j_0\cdots j_{n-1}\rangle = \sum_{J=0}^{2^n-1} a_J|J\rangle = \begin{pmatrix} a_0 \\ \vdots \\ a_{2^n-1} \end{pmatrix}, \tag{2.6}$$

where the integer index J and its binary representation $j_0\cdots j_{n-1}$ can be used interchangeably.

Formally, a single-qubit state is an element of the two-dimensional complex Hilbert space $\mathcal{H}_2 = \mathrm{span}\{|0\rangle, |1\rangle\} \cong \mathbb{C}^2$. A multi-qubit state for n qubits is then an element of the 2^n-dimensional complex Hilbert space $\mathcal{H}_{2^n}$, given by the n-fold tensor product of $\mathcal{H}_2$,

$$\mathcal{H}_{2^n} = \mathcal{H}_2^{\otimes n} = \mathrm{span}\{|0\rangle, |1\rangle\}^{\otimes n} = \mathrm{span}\{ \bigcup_{\substack{j_0,\ldots,j_{n-1} \\ \in \{0,1\}}} |j_0 \cdots j_{n-1}\rangle\}$$

$$= \mathrm{span}\{ \underbrace{\overbrace{|0\cdots 0\rangle}^{n}, \ldots, \overbrace{|1\cdots 1\rangle}^{n}}_{2^n} \} \cong \mathbb{C}^{2^n}. \tag{2.7}$$

To keep the notation concise, we do not write the tensor products explicitly such that the states $|j_0 j_1 \cdots j_{n-1}\rangle$, $|j_0\rangle |j_1\rangle \cdots |j_{n-1}\rangle$, and $|j_0\rangle \otimes |j_1\rangle \otimes \cdots \otimes |j_{n-1}\rangle$ are understood to be the same. Note that the basis states $|j_0 \cdots j_{n-1}\rangle$ in the vector space $\mathcal{H}_{2^n}$ correspond to the same 2^n bit strings $j_0 \cdots j_{n-1}$ comprising the space for n-bit states given by Eq. (2.1). By analogy with single-qubit states, a multi-qubit state can be an arbitrary complex superposition of these basis states.

A property of Eq. (2.6) is that most of the states in the tensor-product space $\mathcal{H}_{2^n}$ cannot be written as tensor products themselves. Such states are called *entangled* states and the corresponding concept is called *entanglement*. Simple examples for entangled states in the two-qubit case $n = 2$ are the so-called Bell states

$$|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \tag{2.8a}$$

$$|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle), \tag{2.8b}$$

the last of which, $|\Psi^-\rangle$, is also known as the *singlet state*. The fact that these states are entangled can be proven by contradiction: Assuming that there exists a tensor-product state of the form $(a_0 |0\rangle + a_1 |1\rangle)(b_0 |0\rangle + b_1 |1\rangle)$ for any of the four Bell states leads to contradictory equations for $a_0, a_1, b_0, b_1 \in \mathbb{C}$.

The presence of entanglement is a consequence of the algebra (with complex tensor-product spaces) that we use in quantum theory to describe observations mathematically. It may seem peculiar in the sense that two separate qubits described by $(|00\rangle + |11\rangle)/\sqrt{2}$ have a strong correlation, i.e., when we measure one qubit we seem to know the state of the other qubit without measuring it. However, this sense of peculiarity is rather a consequence of the way we choose to describe the state.

For instance, suppose that we observe some process that can be described by a probability distribution $p(x, y)$. Then we empirically observe that one cannot express $p(x, y) = f(x)g(y)$ for any two functions f and g and therefore call the variables x and y *entangled*. The existence of such processes seems reasonable, but does not look very peculiar in this language. To make this example concrete in the present context, such a probability distribution for an entangled state $(|00\rangle + |11\rangle)/\sqrt{2}$ would be $p(x, y) = \delta_{xy}/2$ where $x, y \in \{0, 1\}$. It is obvious that δ_{xy} depends on both x and y and cannot be written as a product of some $f(x)$ and $g(y)$. For a non-entangled state such as $(|00\rangle + |10\rangle)/\sqrt{2}$, in contrast, the probability distribution would be $p(x, y) = \delta_{y0}/2$, and it can be easily expressed as a product $f(x)g(y)$ for $f(x) = 1$ and $g(y) = \delta_{y0}/2$.

So entanglement is a consequence of the fact that in quantum theory, we model physical states using complex tensor-product spaces (see Eq. (2.7)). This allows us to describe correlations between individual components of the full space, such as the individual qubits in Eqs. (2.8a) and (2.8b), which cannot be described by states in Cartesian-product spaces (see Eq. (2.1)). For this reason, there is no analogue of entanglement in the classical multi-bit states used to describe digital computers.

It is worth mentioning that there exist mathematical tests for entanglement, known as *Bell tests* in the context of *Bell inequalities* [Bel1964; Bel2004] and *separability criteria* [Hor1996; Ter2000]. If one wants to describe an experiment using a quantum theoretical model, such a Bell test yields an answer to the question if the observations have to be described in terms of an entangled state. Note, however, that such a test can never prove, in a mathematical sense, that the observations can *only* be described by a quantum theoretical model. An alternative "subquantum" model that can describe the individual events and not only their quantum theoretical statistics can be found in [DeR2005] (see [Wil2020f; DeR2020] for two particular applications of the model).

The constraint of normalizing the states is not included in the definition of the multi-qubit space given by Eq. (2.7). The purpose of this is to keep the linearity of the algebraic structures. Similarly, we have not made any efforts to eliminate the global phase of the complex coefficients of the general multi-qubit state given by Eq. (2.6). The concepts of normalization and global phase only play a role when the complex coefficients are to be interpreted as probability amplitudes, or when one wants to eliminate as many degrees of freedom as possible to find concise representations of quantum states. The latter was done to derive the Bloch sphere representation of a single-qubit state (see Eq. (2.3) and Fig. 2.1).

For general multi-qubit states $|\psi\rangle \in \mathcal{H}_{2^n}$, it is possible to visualize the individual qubits using one Bloch sphere per qubit. However, such a picture does not capture all the information required to describe the state $|\psi\rangle$, as it did for a single qubit. This can be understood from a simple counting argument: A single-qubit state $a_0 |0\rangle + a_1 |1\rangle$ is described by two complex numbers or, equivalently, four real numbers. Using the normalization constraint and eliminating the global phase, we were able to reduce these four real numbers to the two angles ϑ and φ, which can be visualized as a vector of length 1 on the Bloch sphere. For n qubits with one Bloch sphere per qubit, we would need $2n$ real numbers (or $3n$ if we do not fix the length of each Bloch vector to 1). However, a general n-qubit state such as the one given by Eq. (2.6) is described by 2^{n+1} real numbers. Even if we subtract two for the normalization constraint and the global phase, this is still much more than can be visualized using n spheres (see also [Ben2006]).

Nevertheless, it may still be helpful to visualize each single qubit on a separate Bloch sphere. To compute the Bloch vectors, we use the notation

$$\sigma_i^\alpha = I \otimes \cdots \otimes \sigma^\alpha \otimes \cdots \otimes I \tag{2.9}$$

for the corresponding Pauli matrices, where I denotes the identity matrix on the two-dimensional single-qubit space $\mathcal{H}_2$, $\alpha \in \{x, y, z\}$ labels the Pauli matrices given by Eq. (2.4), and $i \in \{0, \ldots, n-1\}$ denotes the position of σ^α in this tensor-product matrix using the same ordering as for the multi-qubit basis state $|j_0 \cdots j_{n-1}\rangle$. This means that

applying σ_i^α to this basis state only affects the qubit $|j_i\rangle$ such that

$$\sigma_i^\alpha |j_0 \cdots j_{n-1}\rangle = |j_0\rangle \cdots |j_{i-1}\rangle \left(\sigma^\alpha |j_i\rangle\right) |j_{i+1}\rangle \cdots |j_{n-1}\rangle. \tag{2.10}$$

Using σ_i^α, we can now compute Bloch vectors $\vec{r}_i \in \mathbb{R}^3$ for each qubit in the same manner as in Eq. (2.5) for the general multi-qubit state $|\psi\rangle \in \mathcal{H}_{2^n}$ given by Eq. (2.6). A short calculation yields

$$\vec{r}_i = \begin{pmatrix} r_i^x \\ r_i^y \\ r_i^z \end{pmatrix} = \begin{pmatrix} \langle\psi|\sigma_i^x|\psi\rangle \\ \langle\psi|\sigma_i^y|\psi\rangle \\ \langle\psi|\sigma_i^z|\psi\rangle \end{pmatrix} = \sum_{\substack{j_0\cdots j_{n-1} \\ \text{without } j_i}} \begin{pmatrix} 2\operatorname{Re}(a_{j_0\cdots 0\cdots j_{n-1}}^* a_{j_0\cdots 1\cdots j_{n-1}}) \\ 2\operatorname{Im}(a_{j_0\cdots 0\cdots j_{n-1}}^* a_{j_0\cdots 1\cdots j_{n-1}}) \\ |a_{j_0\cdots 0\cdots j_{n-1}}|^2 - |a_{j_0\cdots 1\cdots j_{n-1}}|^2 \end{pmatrix}, \tag{2.11}$$

exhibiting a similar structure as Eq. (2.5). This expression is used for the implementation of the visualizations discussed in Section 3.3.3 and Appendix A.

It is worth mentioning that for two-qubit states, some ideas have been proposed to visualize a general two-qubit state using three spheres. The requirement is that two of the three spheres shall correspond to the respective single-qubit Bloch vectors computed from Eq. (2.11). The remaining information about the amount of entanglement in the two-qubit state is then visualized on the third sphere. See [Wie2014] or [Rig2009] for more information.

2.1.4 Leakage

In practice, many physical realizations of qubits contain additional states beyond the computational basis states $|0\rangle$ and $|1\rangle$. For superconducting transmon qubits, for instance, this refers to higher excited states $|2\rangle, |3\rangle, \ldots$. Formally, these non-computational states are not described by the tensor-product structure of Eq. (2.7). Instead, they describe alternative states that belong to the individual qubits. This means that the single-qubit description from Eq. (2.2) needs to be extended by a direct sum structure

$$|\psi\rangle = a_0 |0\rangle + a_1 |1\rangle + a_2 |2\rangle + \cdots \in \mathcal{H}_2 \oplus \mathcal{H}_{\mathrm{L}}, \tag{2.12}$$

where $\mathcal{H}_{\mathrm{L}} = \operatorname{span}\{|2\rangle, |3\rangle, \ldots\}$. If the state $|\psi\rangle$ acquires a contribution of states from $\mathcal{H}_{\mathrm{L}}$, one generally speaks of *leakage*.

For the multi-qubit Hilbert space $\mathcal{H}_{2^n}$ defined in Eq. (2.7), the corresponding leakage space is constructed via

$$(\mathcal{H}_2 \oplus \mathcal{H}_{\mathrm{L}}) \otimes \cdots \otimes (\mathcal{H}_2 \oplus \mathcal{H}_{\mathrm{L}}) = \mathcal{H}_{2^n} \oplus \mathcal{H}_{\mathrm{Leak}}, \tag{2.13}$$

where $\mathcal{H}_{\mathrm{Leak}}$ is spanned by each state in which at least one part is an element of $\mathcal{H}_{\mathrm{L}}$.

The concept of leakage and its implications for evolutions of quantum systems are considered further in Section 2.4.2. Since leakage is a particularly important issue for superconducting transmon qubits, the concept will play an important role for the experiments studied throughout this thesis.

2.2 Quantum gates

Given a set of qubits as defined in the previous section, there are certain operations that can be performed on the qubits. These operations are called *quantum gates*. They are inspired by their analogue for digital computers, i.e., the digital logic gates.

A digital logic gate is an arbitrary function $f : \{0,1\}^n \to \{0,1\}^m$. It takes as input a bit string of length n such as the one defined in Eq. (2.1) and outputs another bit string of length m (not necessarily of the same length). There is, in principle, no further limitation for f.

Transferring this idea to quantum states, a quantum gate would need to be a function acting on the space $\mathcal{H}_{2^n}$ defined in Eq. (2.7), which is an immensely larger space. However, there are certain restrictions for quantum gates that impose some nontrivial limitations on these functions. In this section, we first discuss these limitations and then define the set of elementary quantum gates that are implemented in terms of pulses by the transmon simulator (see also Appendix B).

2.2.1 Unitary operators

Every quantum gate is a basic operation on multi-qubit states of the form given by Eq. (2.6). A quantum gate is mathematically defined as a linear map $U : \mathcal{H}_{2^n} \to \mathcal{H}_{2^n}$ with the important restriction that it has to be unitary (see Appendix C for a review of arguments why quantum theory requires unitary linear maps). Note that the linearity of the map U implies that expressions such as $U(|\psi\rangle + |\phi\rangle)$ can be evaluated as $U|\psi\rangle + U|\phi\rangle$ for any quantum gate U and any states $|\psi\rangle$ and $|\phi\rangle$.

As we only consider finite-dimensional Hilbert spaces, we identify unitary operators with their representation in terms of unitary matrices, i.e., complex invertible matrices $U \in \mathbb{C}^{2^n \times 2^n}$ that satisfy $U^{-1} = U^\dagger$, where $U^\dagger$ denotes the Hermitian conjugate (or conjugate transpose) of U. By definition, such matrices conserve the norm of quantum states, so the restriction to unitary matrices goes hand in hand with the interpretation of the complex coefficients in Eq. (2.6) as probability amplitudes.

From a computational perspective, the condition of unitary square matrices imposes severe limitations on the computational model of quantum computing. In particular, each quantum gate has to be invertible (i.e., *reversible*). This property is not fulfilled by many of the conventional digital logic gates. For instance, the logical AND gate takes two bits as input and produces only one bit as output (namely 1 if both of its inputs are 1, and 0 otherwise). Since it is not possible to deduce the two input bits given the output bit, the AND gate is not reversible. The same applies to the logical OR gate, the universal NAND gate, and many others. However, all these gates can, in principle, be made reversible by adding another output bit. This is the idea of classical, reversible computation, which bridges the gap to make the logical gates amenable to quantum computing (see [Fre1982; Ben1973] for more information).

2.2.2 Elementary quantum gates

An elementary single-qubit gate is given by a three-dimensional rotation on the Bloch sphere (see Fig. 2.1). General single-qubit gates are often represented as sequences of such single-qubit rotations. This is possible since for one qubit, the Bloch sphere yields a faithful representation of the state, in the sense that it captures the full information contained in the state. Furthermore, it has the advantage that single-qubit rotations directly relate to the pulses that are used on actual quantum processors (such as the IBM Q processors [IBM2016]) to implement the gates [Cro2017]. For this reason, we follow this convention and express all single-qubit gates as a sequence of rotations on the Bloch sphere.

A single-qubit rotation on the Bloch sphere is given by $R^\alpha(\vartheta) = \exp(-i\vartheta\sigma^\alpha/2)$, where α defines the axis of rotation, σ^α denotes the respective Pauli matrix given by Eq. (2.4), and ϑ is the angle of rotation. One can compute the matrix exponential to get the closed-form expressions

$$R^x(\vartheta) = e^{-i\vartheta\sigma^x/2} = \begin{pmatrix} \cos(\vartheta/2) & -i\sin(\vartheta/2) \\ -i\sin(\vartheta/2) & \cos(\vartheta/2) \end{pmatrix}, \tag{2.14a}$$

$$R^y(\vartheta) = e^{-i\vartheta\sigma^y/2} = \begin{pmatrix} \cos(\vartheta/2) & -\sin(\vartheta/2) \\ \sin(\vartheta/2) & \cos(\vartheta/2) \end{pmatrix}, \tag{2.14b}$$

$$R^z(\vartheta) = e^{-i\vartheta\sigma^z/2} = \begin{pmatrix} \exp(-i\vartheta/2) & 0 \\ 0 & \exp(i\vartheta/2) \end{pmatrix}. \tag{2.14c}$$

The operation $R^\alpha(\vartheta)\ket{\psi}$, where $\ket{\psi}$ is the single-qubit state shown in Fig. 2.1, corresponds to a rotation of the Bloch vector $\vec{r}$ (see Eq. (2.5)) around the axis α by an angle ϑ. The sense of rotation is given by the *right-hand rule*, set by the minus sign in the exponent of $R^\alpha(\vartheta)$.

We define three elementary combinations of $R^x(\pi/2)$ and $R^z(\vartheta)$ that are directly related to the hardware implementations of the IBM Q processors [Cro2017] (see also Section 5.1). They are given by

$$\mathrm{U1}(\lambda) = c_1 R^z(\lambda), \tag{2.15a}$$

$$\mathrm{U2}(\phi, \lambda) = c_2 R^z\left(\phi + \frac{\pi}{2}\right) R^x\left(\frac{\pi}{2}\right) R^z\left(\lambda - \frac{\pi}{2}\right), \tag{2.15b}$$

$$\mathrm{U3}(\theta, \phi, \lambda) = c_3 R^z(\phi + 3\pi) R^x\left(\frac{\pi}{2}\right) R^z(\theta + \pi) R^x\left(\frac{\pi}{2}\right) R^z(\lambda), \tag{2.15c}$$

where the complex phase factors $c_1 = \exp(i\lambda/2)$ and $c_2 = c_3 = \exp(i(\phi + \lambda)/2)$ are not essential for the operation of the gates and are given for reference only.

When a single-qubit gate is applied to one of the qubits of a multi-qubit state $\ket{\psi} \in \mathcal{H}_{2^n}$ (see Eq. (2.6)), the Pauli matrices σ^α in Eqs. (2.14a)–(2.14c) have to be replaced by σ^α_i defined in Eq. (2.9). Accordingly, we denote the corresponding rotations by

$$R^\alpha_i(\vartheta) = e^{-i\vartheta\sigma^\alpha_i/2} = I \otimes \cdots \otimes R^\alpha(\vartheta) \otimes \cdots \otimes I. \tag{2.16}$$

In the following chapters, we sometimes use the widespread alternative notations $X^\vartheta_i = R^x_i(\vartheta)$, $Y^\vartheta_i = R^y_i(\vartheta)$, $Z^\vartheta_i = R^z_i(\vartheta)$. Furthermore, we use the convention that $\vartheta = \pi$ if the "exponent" ϑ is not explicitly specified.

Any of the single-qubit gates used in this work can be expressed in terms of the single-qubit rotations defined in Eqs. (2.14a)–(2.14c) or the U gates defined in Eqs. (2.15a)–(2.15c) (see Tab. B.1 in Appendix B for a list of all gates, their matrix representations, and their relations to the elementary gates). Two particularly important single-qubit gates are the X gate (also known as bit flip or NOT gate) and the Hadamard gate. They are defined as

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad\qquad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \qquad (2.17)$$

As before, we extend these single-qubit gates to multi-qubit spaces using the notations X_i and H_i, as done in Eqs. (2.9) and (2.16). Note that the gate X_i corresponds to the single-qubit rotation X_i^ϑ for $\vartheta = \pi$ (up to a complex phase factor), which is in agreement with the above convention.

In quantum computing, there is a much larger variety of single-qubit gates than in digital computing (where the only nontrivial single-bit gate is the NOT gate). Nonetheless, it turns out that almost any two-qubit gate is sufficient to build a *universal gate set* from the single-qubit rotations [Deu1995; DiV1995] (see also [DiV2000]). In this context, universal means that any quantum gate on $\mathcal{H}_{2^n}$ can be represented as a finite sequence of gates from this set, using suitable angles for all single-qubit rotations.

One such two-qubit gate is the controlled NOT (CNOT) gate. It is defined as

$$\text{CNOT} = \begin{matrix} \begin{matrix} |00\rangle & |01\rangle & |10\rangle & |11\rangle \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}, \qquad (2.18)$$

such that $\text{CNOT} |j_0\rangle |j_1\rangle = |j_0\rangle |j_0 \oplus j_1\rangle$ where $\oplus$ denotes the XOR operation (or integer addition modulo 2). The effect of this gate is to flip the target qubit $|j_1\rangle$ if and only if the control qubit $|j_0\rangle$ is in state $|1\rangle$. In the multi-qubit space $\mathcal{H}_{2^n}$ (see Eq. (2.7)), we use the notation CNOT_{il} to denote a CNOT gate where qubit i is the control qubit and qubit l is the target qubit.

An important thing to realize is that the model of gate-based quantum computing with elementary gates given by the single-qubit rotations in Eqs. (2.14a)–(2.14c) is a model of analog computation. Given an angle $\vartheta \in [0, 2\pi)$, a high level of precision over the controlling pulse may be necessary to implement the rotation $R_i^\alpha(\vartheta)$ accurately. We study hardware implementations of quantum gates (and in particular the two-qubit CNOT gate) in more detail in the following chapters.

2.3 Quantum circuits

For a gate-based quantum computer, an algorithm is specified in terms of a *quantum circuit*. It consists of a sequence of quantum gates as defined in the previous section.
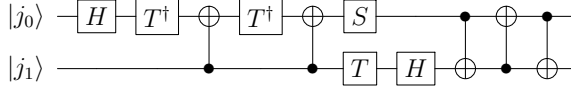
Figure 2.3: Circuit diagram for the two-qubit QFT. The boxes denote single-qubit gates (see Tab. B.1 in Appendix B), and vertical lines denote two-qubit CNOT gates as defined in Eq. (2.18). The solid circle in a CNOT gate denotes the control qubit, and the open circle denotes the target qubit. The matrix representation of this circuit is given by Eq. (2.19).

This model is inspired by the circuit model of digital computation, in which an algorithm for a digital computer can, in principle, be decomposed into a sequence of digital logic gates.

As each quantum gate is a unitary operator, a full quantum circuit is also a unitary operator (the set of unitary operators on a given Hilbert space forms a group). For an n-qubit system described by the Hilbert space $\mathcal{H}_{2^n}$ given by Eq. (2.7), a quantum circuit is thus a specification of a large but sparse unitary matrix $U \in \mathbb{C}^{2^n \times 2^n}$. If a quantum circuit is written as a sequence of gates from a certain gate set, this means that a (potentially large) unitary matrix is expressed as a product of (typically small) elementary unitary matrices. In this sense, a quantum circuit is just a decomposition of a large matrix into smaller matrices. A list of elementary gates and their matrix representations is given in Appendix B.

Quantum circuits are often expressed in a diagrammatic notation, where each qubit corresponds to a horizontal line. Gates on each qubit are specified using corresponding circuit symbols. An example for a simple circuit diagram representing a two-qubit quantum Fourier transform (QFT) is shown in Fig. 2.3. It is taken from [Nie2010] after rewriting some of the gates in terms of the elementary gate set used in this thesis.

Circuit diagrams are defined such that time increases from left to right. Therefore, when converting between circuit diagrams and their matrix representations, the order of the matrices has to be reversed. For instance, the unitary matrix corresponding to the circuit shown in Fig. 2.3 is computed as

$$U^{(\text{QFT2})} = \text{CNOT}_{01}\, \text{CNOT}_{10}\, \text{CNOT}_{01}\, H_1\, T_1\, S_0\, \text{CNOT}_{10}\, T_0^\dagger\, \text{CNOT}_{10}\, T_0^\dagger\, H_0$$

$$= \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}. \tag{2.19}$$

The QFT can be generalized to an arbitrary number of qubits. Its generic property is that it maps a computational basis state to a uniform superposition over all states with relative phases dependent on the original state. This can be seen from its definition for a

basis state $|J\rangle \in \mathcal{H}_{2^n}$ [Nie2010],

$$U^{(\text{QFT}n)} |J\rangle = \frac{1}{2^{n/2}} \sum_{K=0}^{2^n-1} e^{2\pi i JK/2^n} |K\rangle, \tag{2.20}$$

where we identified J and its binary representation $j_0 \cdots j_{n-1}$ as done in Eq. (2.6).

Mathematically, the QFT corresponds to a discrete Fourier transform over the additive group of integers modulo 2^n. Interestingly, a universal quantum computer using the gate set defined in Appendix B can do a QFT in a number of steps *polynomial* in n, as opposed to the *Fast Fourier Transform* that requires $\mathcal{O}(n2^n)$ steps [Coo1965]. This discovery forms the basis for most of the quantum algorithms for which an exponential speedup is known.

One such algorithm is Shor's factorization algorithm [Sho1994] in which the QFT is basically used to find the period of a suitable function (note that finding periods is a generic feature of any Fourier transform). The common algebraic problem solved by most algorithms with an exponential speedup is the *hidden subgroup problem* [Joz2001].

The difficulty in finding these algorithms, however, is that even for an ideal gate-based quantum computer, the complex phases in Eq. (2.20) are not directly accessible (see also the discussion at the end of Section 2.1.1). Furthermore, actual implementations of the QFT require an extremely precise control over the phases of the pulses that are used to implement the gates. These phases are related to the relative phase factors present in Eq. (2.20). We study the problem of controlling the phases and implementing gates through pulses in more detail in the following chapters.

2.4 Quantum operations

In some situations, it is expedient to consider not a single, pure quantum state $|\psi\rangle \in \mathcal{H}_{2^n}$ representing the n qubits, but a probability distribution $\{p_i\}$ over several quantum states $|\psi_i\rangle \in \mathcal{H}_{2^n}$. Such states are called *mixed states*. They can be represented by a density matrix

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \tag{2.21}$$

This representation is called *ensemble representation*. The fact that $\{p_i\}$ represents a probability distribution means that each $p_i \geq 0$, and $\text{Tr}\,\rho = \sum_i p_i = 1$. Further properties of ρ are $\rho = \rho^\dagger$ and that ρ is a positive (semidefinite) matrix, i.e., $\langle\phi|\rho|\phi\rangle \geq 0$ for all $|\phi\rangle$. Note that this characterization of density matrices is both necessary and sufficient, since each positive matrix ρ with $\text{Tr}\,\rho = 1$ can be written in the form of Eq. (2.21) by taking its spectral decomposition such that $p_i \geq 0$ are the eigenvalues and $|\psi_i\rangle$ are the eigenvectors of ρ.

The model of a gate-based quantum computer can be extended to a description in terms of mixed states. The result of applying a quantum gate U, which maps a pure quantum state $|\psi\rangle$ to $U |\psi\rangle$, is then described by

$$\rho \mapsto U\rho U^\dagger = \sum_i p_i U |\psi_i\rangle\langle\psi_i| U^\dagger. \tag{2.22}$$

For mixed states, however, one often considers more general transformations called *quantum operations* [Kra1971]. We denote such an operation by a map $\mathcal{E}$ that transforms ρ according to

$$\rho \mapsto \mathcal{E}(\rho). \tag{2.23}$$

Typical constraints on $\mathcal{E}$ are that it be linear (cf. Appendix C), Hermiticity-preserving $(\mathcal{E}(\rho)^\dagger = \mathcal{E}(\rho^\dagger))$, and completely positive. The latter means formally that if A is a positive matrix on an extended Hilbert space of arbitrary dimensionality, then also the extended map $\mathcal{E} \otimes \mathbb{1}$ preserves the positivity of A [Sti1955]. This property ensures that a density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ with non-negative probabilities p_i is mapped to another density matrix $\mathcal{E}(\rho)$ that also represents non-negative probabilities. Furthermore, the "completely" in completely positive ensures that this preservation of positivity also applies if the description is extended to another system (such as extending the description of a single qubit to two qubits, or one qubit and an environment). If $\mathcal{E}$ is also trace-preserving (i.e., $\mathrm{Tr}\,\mathcal{E}(\rho) = \mathrm{Tr}\,\rho$), we call the completely positive trace-preserving (CPTP) map $\mathcal{E}$ *quantum channel* or *error channel* (cf. Section 7.2.2).

2.4.1 Representations of quantum operations

There are several ways of representing a quantum operation $\mathcal{E}$. Since it is a linear map, one obvious way is to write it as a matrix with elements $\mathcal{E}_{(ij),(kl)} = \mathrm{Tr}\,|i\rangle\langle j|^\dagger\,\mathcal{E}(|k\rangle\langle l|) = \langle i|\mathcal{E}(|k\rangle\langle l|)|j\rangle$. For an n-qubit system of dimension $N = 2^n$, this matrix has $N^2 \times N^2$ complex elements. It is referred to as the matrix representation of $\mathcal{E}$ in the standard basis. However, this representation is not tuned to the particular properties of $\mathcal{E}$. For instance, the property of complete positivity is not easily expressed by this matrix.

A more useful representation is the so-called *Kraus representation*, $\mathcal{E}(\rho) = \sum_\alpha A_\alpha \rho B_\alpha^\dagger$, where $A_\alpha, B_\alpha \in \mathbb{C}^{N \times N}$. One can show that $\mathcal{E}$ is completely positive if and only if $B_\alpha = A_\alpha$ [Kra1971; Cho1975]. This means that for completely positive maps $\mathcal{E}$, a Kraus representation is given by

$$\mathcal{E}(\rho) = \sum_{\alpha=1}^R E_\alpha \rho E_\alpha^\dagger, \tag{2.24}$$

where $E_\alpha \in \mathbb{C}^{N \times N}$. The smallest number R of terms in the Kraus representation is called the *Kraus rank* of $\mathcal{E}$. Using the Kraus representation, one has that $\mathcal{E}$ is trace-preserving if and only if $\sum_\alpha E_\alpha^\dagger E_\alpha = \mathbb{1}$.

Another commonly used matrix representation of $\mathcal{E}$ is the so-called *Choi matrix* $J(\mathcal{E})$. It is defined as

$$J(\mathcal{E}) = \frac{1}{N}\sum_{i,j=0}^{N-1} |i\rangle\langle j| \otimes \mathcal{E}(|i\rangle\langle j|) = \frac{1}{N}\begin{pmatrix} \mathcal{E}(|0\rangle\langle 0|) & \mathcal{E}(|0\rangle\langle 1|) & \cdots \\ \mathcal{E}(|1\rangle\langle 0|) & \mathcal{E}(|1\rangle\langle 1|) & \\ \vdots & & \ddots \end{pmatrix}, \tag{2.25}$$

and it represents a matrix with $N^2 \times N^2$ complex coefficients. A useful property of the Choi matrix is that $\mathcal{E}$ is completely positive if and only if $J(\mathcal{E})$ is positive semidefinite [Cho1975].

Furthermore, the rank of the Choi matrix (i.e., the number of non-zero singular values) yields the minimum number R of terms in the Kraus representation given by Eq. (2.24). For this reason, the Kraus rank R is also called the *Choi rank* of $\mathcal{E}$.

The Choi matrix is often written in compact form as $J(\mathcal{E}) = (\mathbb{1} \otimes \mathcal{E})(\Phi)$, where $\Phi = |\Phi\rangle\langle\Phi|$ and $|\Phi\rangle = \sum_j |jj\rangle / \sqrt{N}$ is the maximally-entangled state on an extended Hilbert space. Note also that the order of the tensor-product factors is sometimes reversed such that $J(\mathcal{E}) = (\mathcal{E} \otimes \mathbb{1})(\Phi)$ is used instead. Both definitions are equivalent, but the matrix representation of the one given in Eq. (2.25) appears more canonical.

Finally, there is a matrix representation that is particularly convenient for numerical work called the *Pauli transfer matrix G*. It is the matrix representation of the linear map $\mathcal{E}$ with respect to the Pauli basis

$$\mathcal{P} = \{I, \sigma^x, \sigma^y, \sigma^z\}^{\otimes n}. \tag{2.26}$$

We denote the elements of $\mathcal{P}$ by P_i for $i = 0, \ldots, N^2 - 1$, where $P_0 = I \otimes \cdots \otimes I$, $P_1 = I \otimes \cdots \otimes I \otimes \sigma^x$, $P_2 = I \otimes \cdots \otimes I \otimes \sigma^y$, and so on. The basis $\mathcal{P}$ is orthogonal with respect to the Hilbert-Schmidt inner product, $\operatorname{Tr} P_i^\dagger P_j = N\delta_{ij}$. The corresponding normalized basis elements are given by $\widehat{P}_i = P_i/\sqrt{N}$. In this basis, we obtain the Pauli transfer matrix of $\mathcal{E}$ as

$$G_{ij} = \operatorname{Tr} \widehat{P}_i \mathcal{E}(\widehat{P}_j) = \frac{1}{N}\operatorname{Tr} P_i \mathcal{E}(P_j). \tag{2.27}$$

The matrix G obeys the typical properties expected from matrix representations of linear maps, namely that the columns contain the images of the basis elements, and the composition of two maps $\mathcal{E}_1 \circ \mathcal{E}_2$ corresponds to the matrix product $G_1 G_2$ of their Pauli transfer matrices. Furthermore, a quantum operation $\mathcal{E}$ preserves Hermiticity, so the matrix G consists only of real numbers, which makes it convenient for numerical work. For trace-preserving and trace-decreasing quantum operations, the matrix elements G_{ij} are in the range $[-1, 1]$. We also have $G_{0j} = \operatorname{Tr} \mathcal{E}(P_j)/N$, so the first row of G is given by $(1, 0, \ldots, 0)$ if and only if $\mathcal{E}$ is trace-preserving. Another property of the Pauli transfer matrix is that if the first column of G is given by $(1, 0, \ldots, 0)^T$, the map $\mathcal{E}$ is unital (meaning that $\mathcal{E}(\mathbb{1}) = \mathbb{1}$). Finally, if all rows and all columns of G contain exactly one non-zero entry of magnitude 1, the map $\mathcal{E}$ is known as a *Clifford gate*, which means that it maps Pauli operators to Pauli operators. Clifford gates are useful as it can be shown that simulating quantum circuits containing only Clifford gates is much less complex than simulating universal quantum circuits [Got1998a].

A modern perspective including a graphical calculus for the common ways of representing quantum operations is given in [Woo2015]. For the present work, the (generalized) Kraus representation and the Pauli transfer matrix are the most useful representations. The former is used in various places in Chapter 6 and Chapter 7. The latter, because of the convenient properties discussed above, will prove particularly useful for the tomography experiments studied in Section 6.3.

2.4.2 Transformations of subsystems and leakage

In Section 2.2, we said that quantum theory typically considers unitary transformations (see Appendix C for the reasons behind this). This notion is also contained in the formalism of quantum operations, although it is not immediately apparent from the completely positive map $\mathcal{E}$ given in Eq. (2.24). To illustrate this connection, we extend the computational Hilbert space $\mathcal{H}_{2^n}$ given by Eq. (2.7) with another Hilbert space $\mathcal{H}_{\mathrm{Env}}$,

$$\mathcal{H} = \mathcal{H}_{2^n} \otimes \mathcal{H}_{\mathrm{Env}}, \tag{2.28}$$

which can be interpreted as a simple system-environment model. The idea of this model is that the system and the environment, which are initially described by a product state $\rho \otimes \rho_{\mathrm{Env}}$, undergo a joint unitary transformation U, i.e., $\rho \otimes \rho_{\mathrm{Env}} \mapsto U(\rho \otimes \rho_{\mathrm{Env}})U^\dagger$. The final state of the system is then fully described by

$$\mathcal{E}(\rho) = \mathrm{Tr}_{\mathrm{Env}}\big(U(\rho \otimes \rho_{\mathrm{Env}})U^\dagger\big), \tag{2.29}$$

where $\mathrm{Tr}_{\mathrm{Env}}$ denotes the partial trace over the environment's degrees of freedom. In this context, "fully described" means that the expectation value of each observable A on $\mathcal{H}_{2^n}$ is given by $\mathrm{Tr}A\mathcal{E}(\rho)$.

To relate Eq. (2.29) to the Kraus representation given by Eq. (2.24), we write the initial state of the environment as a pure state $\rho_{\mathrm{Env}} = |e_0\rangle\langle e_0|$ (note that the space H_{Env} can always be chosen large enough so that ρ_{Env} can be expressed as a pure state [Nie2010]). Choosing $E_\alpha = (\mathbb{1}\otimes\langle e_\alpha|)U(\mathbb{1}\otimes|e_\alpha\rangle)$, where $\{|e_\alpha\rangle\}$ is a basis of $\mathcal{H}_{\mathrm{Env}}$ completing $|e_0\rangle$, yields the Kraus representation of $\mathcal{E}$ given by Eq. (2.24). The fact that the Kraus operators are given by E_α and $E_\alpha^\dagger$, respectively, shows that each model of the form of Eq. (2.29) is automatically completely positive (see [Bre2007] for a more comprehensive discussion).

The system-environment model can also be extended to a description of leakage (cf. Section 2.1.4) by supplementing the Hilbert space given by Eq. (2.28) with a direct sum for higher, non-computational states,

$$\mathcal{H} = (\mathcal{H}_{2^n} \oplus \mathcal{H}_{\mathrm{Leak}}) \otimes \mathcal{H}_{\mathrm{Env}}, \tag{2.30}$$

where $\mathcal{H}_{\mathrm{Leak}}$ is defined in Eq. (2.13). Since $\mathcal{H}_{2^n}$ and $\mathcal{H}_{\mathrm{Leak}}$ form a direct sum and not a direct product, one cannot trace out $\mathcal{H}_{\mathrm{Leak}}$ to obtain a description of the computational subspace. Instead, one can project the result of Eq. (2.29) onto H_{2^n}. The corresponding quantum operation for a density matrix ρ on $\mathcal{H}_{2^n}$ is then given by

$$\overline{\mathcal{E}}(\rho) = P\,\mathrm{Tr}_{\mathrm{Env}}\big(U(\rho \otimes \rho_{\mathrm{Env}})U^\dagger\big)P = \sum_\alpha \overline{E}_\alpha \rho \overline{E}_\alpha^\dagger, \tag{2.31}$$

where $\overline{E}_\alpha = P(\mathbb{1}\otimes\langle e_\alpha|)U(\mathbb{1}\otimes|e_\alpha\rangle)P$, and P denotes the projection onto the computational subspace H_{2^n}. As before, the state $\overline{\mathcal{E}}(\rho)$ fully describes the final state, in the sense that expectation values for observables on $\mathcal{H}_{2^n}$ can be evaluated with $\overline{\mathcal{E}}(\rho)$. Furthermore, since $\overline{\mathcal{E}}(\rho)$ can be written in Kraus form (see Eq. (2.24)), it is automatically completely positive. The only difference is that it may no longer be trace-preserving due to the projection P.

Note that, in a description of an experiment, it may be that leakage can still be described by trace-preserving quantum operations. This depends on how higher, non-computational states show up in the measurement. For instance, if a measurement reports each non-computational state as $|1\rangle$, the quantum operation would need to map each non-computational state to $|1\rangle$ (which is a non-unitary operation), but the corresponding quantum operation would be trace-preserving. If, however, a non-computational state shows up randomly as 0 or 1, or if the measured event is classified as "wrong" and discarded, the procedure would be described in terms of a trace-decreasing quantum operation. In this case, the resulting probability distribution may need to be renormalized, which is sometimes also be modeled by a nonlinear, trace-preserving quantum operation of the form $\rho \mapsto \mathcal{E}(\rho)/\mathrm{Tr}\,\mathcal{E}(\rho)$.

Finally, we remark that there are also simple quantum systems whose evolution cannot be described by quantum operations of the form of Eq. (2.24), i.e., systems that cannot be described by completely positive maps. In the context of Eq. (2.29), this may be the case if the system and the environment do not start in a separable state such as $\rho \otimes \rho_{\mathrm{Env}}$. See [Nie2010] for a simple example of such a system. Further characterizations of quantum systems beyond completely positive maps are given in [Car2008; Dom2016].

We utilize the formalism of quantum operations on mixed states in Chapter 6 to introduce error metrics on quantum gates and the procedure of gate set tomography. Furthermore, practical applications of quantum channels are given in Chapter 7 for simple, effective error models and example circuits from the theory of quantum fault tolerance.

Chapter 3

Simulating superconducting transmon qubits

Over the past decades, superconducting circuits have emerged as a convenient platform to engineer quantum mechanical systems with very few degrees of freedom. This is remarkable in the sense that such quantum mechanical systems are usually given by atoms or single electrons, i.e., *microscopic* objects that cannot be easily perceived. Superconducting circuits, however, are visible to the naked eye. And although these macroscopic electrical systems are composed of a huge number of atoms, they exhibit collective phenomena that can be accurately described with Hamiltonians that have only a small number of charge and phase variables. In this context, the field is commonly called *macroscopic quantum mechanics* and the systems are often referred to as *artificial atoms*. It is this property that makes superconducting circuits ideal candidates to engineer quantum mechanical two-level systems that serve as qubits for quantum information processors, and a huge variety of different candidates has been studied in the literature. A review of prominent superconducting architectures for quantum information processors and their theoretical modeling can be found in [Wen2017].

In this chapter, we introduce and describe the transmon simulator that is used for most of the simulations presented in this work. Section 3.1 reviews the quantum mechanical modeling of superconducting circuits with an emphasis on the architecture used for transmon qubits. In Section 3.2, we define the generic model Hamiltonian used to describe the dynamics of a transmon quantum computer. Section 3.3 describes the simulation packages developed for this work, including the implementations of the numerical algorithms. The most important transmon model systems used for the results presented in the remainder of this thesis are defined in Section 3.4. Finally, in Section 3.5, we describe a procedure to extract suitable model parameters from experiments or electromagnetic simulations of the experimental devices.

3.1 Superconducting circuits

A superconducting circuit is an electronic circuit in which the circuit elements are superconducting. This means that they conduct electricity with practically zero resistance. Just like conventional electronic circuits, a superconducting circuit includes basic circuit

elements such as capacitors or inductors. A circuit element of particular importance for quantum information processors is the Josephson junction. It consists of two superconducting metals with an insulating barrier in between. A key observation was that such a system can be described by quantum mechanical tunneling processes of the superconducting charge carriers, the Cooper pairs [Jos1962].

3.1.1 Quantum and classical descriptions

The quantum mechanical description of superconducting circuits emerged from a course given by Devoret at the Les Houches School of Physics [Dev1997] and has recently been reviewed and updated in [Voo2017]. The first step is to obtain a classical Hamiltonian (or a Lagrangian) describing the dynamics of the superconducting circuit. For basic circuit elements such as capacitors or inductors, the equations of motion determined by the Hamiltonian are the corresponding differential equations of classical electrodynamics (see [Jac1999]). Typically, the resulting Hamiltonians are a set of harmonic oscillators with potential anharmonicities from the Josephson junctions.

The quantum description in terms of a quantum Hamiltonian is obtained by quantizing the harmonic oscillators. The quantized Hamiltonian is the key element to model dissipationless superconducting circuits for quantum information processors. It can directly be used to obtain the dynamics of the system as described by the TDSE given by Eq. (1.1).

For the quantum mechanical description of superconducting circuits to be appropriate, the following physical conditions are considered necessary [Dev1997; Voo2017]:

(a) The characteristic wave lengths corresponding to the oscillation frequencies need to be larger than the dimensions of the chip. In this case, the circuit is in the *lumped element limit* and can be described by only a few collective degrees of freedom such as the charge or the flux (which could still follow the classical equations of motion).

(b) The temperature surrounding the system needs to be sufficiently low such that thermal fluctuations are much smaller than the spacing of the energy levels (although a high temperature does not by itself invalidate the quantum description).

(c) The widths of the energy levels (caused by dissipative elements such as resistors that induce damping and reduce quality factors) need to be much smaller than the spacing of the energy levels.

For quantum harmonic oscillators, a well-known property is that the expectation values still follow the classical equations of motion. Therefore, quantum phenomena would only be observable in second-order expectation values (such as the variance as a function of temperature). The presence of at least one nonlinear circuit element (such as a Josephson junction) can make quantum effects more directly observable [Dev1985; Mar1987; Cla1988].

Essentially, however, the superconducting systems under investigation are macroscopic objects in the sense that they contain a large number of constituents (such as aluminum atoms). Ultimately, one cannot prove that a quantum description for a certain experiment is necessary. The reason is that conceptually, one can never formally prove that a
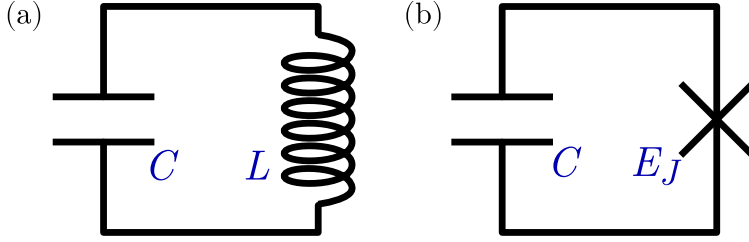
Figure 3.1: Lumped element circuit diagrams for (a) an LC resonator with capacitance C and inductance L and (b) a Josephson junction with capacitance C and Josephson energy E_J.

(potentially unknown) model will not be able to describe the observations properly (see also [Jay1996]).

In the present case, descriptions based solely on classical electrodynamics are occasionally explored. A nice overview concerning Josephson junctions is given in [Bla2016]. Further ventures in this direction can be found in [Mic2005; Mar2007; Grø2010; Kad2016; Bla2017; Iva2018], the last of which is particularly constructive. It would be interesting to investigate the descriptive potential of these classical models for the quantum computing systems under investigation.

This project, however, is concerned with the quantum description. We show that the discrete energy-level structure of the quantum mechanical description makes the problem amenable to large-scale simulations.

To introduce the quantum descriptions, we consider two elementary circuit components of a quantum information processor with transmon qubits. The first is a simple LC resonator and the second is a Josephson junction. The lumped element circuits for isolated, ideal versions of both are shown in Fig. 3.1. We summarize the main relations; a more detailed review can be found in [Voo2017].

3.1.2 LC resonator

The LC resonator shown in Fig. 3.1(a) is a typical example for an harmonic oscillator in classical electrodynamics. Its characteristic frequency is $\omega_{\mathrm{LC}} = 1/\sqrt{LC}$, where C is the capacitance and L is the inductance. The dynamical variables are the charge Q of the capacitor and the magnetic flux Φ of the inductor. They obey the typical differential equations for an harmonic oscillator, i.e., $\ddot{Q} + \omega_{\mathrm{LC}}^2 Q = 0$ and $\ddot{\Phi} + \omega_{\mathrm{LC}}^2 \Phi = 0$. The equations are Hamilton's equations of motion for the Hamiltonian

$$H_{\mathrm{LC}} = \frac{Q^2}{2C} + \frac{\Phi^2}{2L}. \tag{3.1}$$

A quantum mechanical description of the system can be obtained by replacing the variables Q and Φ with operators on the Hilbert space $L^2(\mathbb{R})$ of square-integrable functions over $\mathbb{R}$. The operators have the spectrum $\mathbb{R}$ and they obey a commutation relation $[Q, \Phi] \propto i$. They have the same algebraic properties as the position and momentum

operators used in the treatment of the quantum harmonic oscillator [Bal1998]. Thus the system can be diagonalized algebraically by introducing raising and lowering operators $\hat{a}^\dagger$ and $\hat{a}$, respectively, such that (up to a constant)

$$H_{\text{LC}} = \omega_{\text{LC}} \hat{a}^\dagger \hat{a}. \tag{3.2}$$

In quantum optics, the operator $\hat{a}^\dagger \hat{a}$ is called the photon number operator and its eigenvalues $k \in \mathbb{N}_0$ represent the number of photons in the corresponding monochromatic electric field of frequency ω_{LC} [Fox2006].

3.1.3 Josephson junction

The circuit shown in Fig. 3.1(b) represents the capacitance C and the Josephson energy E_J of a Josephson junction. The capacitance originates directly from the superconductor-insulator-superconductor geometry. It determines one of two characteristic energy scales of a Josephson junction, namely the charging energy

$$E_C = \frac{e^2}{2C}, \tag{3.3}$$

where e is the electron charge. E_C can be interpreted as the electrostatic energy required to charge the capacitor with an additional electron (although, to be precise, the superconducting charge carriers are Cooper pairs with a charge of $2e$, so the electrostatic energy changes in multiples of $4E_C$). The other characteristic energy scale is associated with the tunneling current through the junction, namely the Josephson energy

$$E_J = \frac{I_c}{2e}, \tag{3.4}$$

where I_c is the critical current representing the maximum tunneling current. E_J can be interpreted as the energy required for tunneling processes through the insulating barrier of the Josephson junction.

An important Hamiltonian for the circuit shown in Fig. 3.1(b), analogous to H_{LC} given in Eq. (3.1), reads

$$H_{\text{JJ}} = \frac{Q^2}{2C} - E_J \cos(\Phi/\phi_0), \tag{3.5}$$

where Q denotes the charge on the capacitor C, Φ stands for the magnetic flux, and $\phi_0 = \hbar/2e$ is the reduced flux quantum. The common characterization of a Josephson junction as a weakly anharmonic oscillator can be seen from Eq. (3.5) by expanding the cosine in powers of Φ (see also Section 3.5). The leading terms are then the same as for the harmonic oscillator represented by H_{LC} in Eq. (3.1).

An important difference, however, is that the flux Φ in H_{LC} ranges from $-\infty$ to ∞. In H_{JJ}, the dependence on Φ is periodic such that restricting $\Phi/\phi_0 \in [0, 2\pi)$ is sufficient to describe the system. A deeper understanding for the periodicity of Φ can be gained by describing the superconductors on both sides of the Josephson junction in the BCS theory of superconductivity [Bar1957]. This approach links the magnetic flux Φ directly to the

phase difference of the collective Cooper pair ground states (see [Wil2016; Voo2017] for more information).

In the quantum mechanical version of H_{JJ}, the charge Q and flux Φ are again represented by Hermitian operators. They are often made dimensionless such that $\hat{n} = Q/2e$ denotes the charge operator and $\hat{\varphi} = \Phi/\phi_0$ denotes the phase operator. The Hamiltonian then becomes

$$H_{\text{JJ}} = 4E_C\hat{n}^2 - E_J\cos\hat{\varphi}. \tag{3.6}$$

Note that the spectra of the operators $\hat{n}$ and $\hat{\varphi}$ are not the same as for the harmonic oscillator described in the previous section. In fact, the spectrum of $\hat{n}$ is $\mathbb{Z}$, where both positive and negative integers are physically meaningful since $\hat{n}$ represents the difference in the number of net charges on both sides of the capacitor. Consequently, the conjugate operator $\hat{\varphi}$ has a bounded spectrum given by $[0, 2\pi)$, representing the periodic dependence on Φ in Eq. (3.5).

The different spectra can lead to apparent mathematical paradoxes or contradictions such as $1 = 0$ (cf. [Dir1927; Car1968; Los1992; Bar2007]). The contradictions stem from applying the operators to states that lie outside their domain. This problem does not occur if the operator domains are evaluated and adhered to properly (see [Wil2016]).

3.1.4 Cooper pair box

A system known as the Cooper pair box (CPB) [Bou1998] can be obtained by applying an external voltage bias $V_g(t)$ to the Josephson junction shown in Fig. 3.1(b). The voltage bias can be used to control the number of charges (i.e., Cooper pairs) stored on the capacitor of the Josephson junction. The external voltage is modeled by a time-dependent offset to the number of charges given by $n_g(t) = C_gV_g(t)/2e$, where C_g is the capacitance of the gate through which the voltage is applied. The Hamiltonian H_{JJ} given in Eq. (3.6) then needs to be replaced by the time-dependent Hamiltonian

$$H_{\text{CPB}} = 4E_C(\hat{n} - n_g(t))^2 - E_J\cos\hat{\varphi}. \tag{3.7}$$

An important property of the CPB is that the dynamics of the system can be controlled externally through an electromagnetic pulse described by $n_g(t)$. In quantum computing systems, $n_g(t)$ represents the pulses that are used to implement quantum gates (see Chapter 5).

CPBs have been used to engineer qubits for quantum information processors for almost 20 years [Vio2002; Bla2004], where two low-energy states of the multi-level system are used to define the qubits. CPB qubits can be classified into two groups by means of the energy scales E_J and E_C, namely the charge qubit where $E_C \gg E_J$ (such that the charging energy dominates), and the transmon qubit where $E_C \ll E_J$ [Koc2007] (the name *transmon* is derived from *transmission-line shunted plasma oscillation* and refers to a large shunt capacitance that reduces E_C, see Eq. (3.3)). The transmon has turned out to be much more coherent [Pai2011; Rig2012] and easier to control [Cho2012]. It is now employed in many experimental quantum computing platforms [Nat2019] by companies such as IBM [IBM2016], Google [Aru2019], and Rigetti Computing [Rig2017].

3.2 Transmon quantum computer model

The topic of this thesis is the simulation of a system of transmons and resonators by means of solving the TDSE,

$$i\frac{\partial}{\partial t}\left|\Psi(t)\right\rangle = H\left|\Psi(t)\right\rangle, \tag{3.8}$$

for a generic model Hamiltonian H. The solution to Eq. (3.8), namely the state $\left|\Psi(t)\right\rangle$, can be used to compute any physically relevant quantity. The model Hamiltonian H includes N_{Tr} transmons described as CPBs (see Eq. (3.7)) to represent the qubits, N_{Res} transmission-line resonators described as LC oscillators (see Eq. (3.2)), and various ways of coupling transmons and resonators.

Throughout this work, we use units with $\hbar = 1$ unless otherwise stated. Often, time and energy are the only necessary physical quantities that appear in this thesis. Typically, time is given in nanoseconds, and energies and frequencies are given in gigahertz.

3.2.1 Hamiltonian

The full model Hamiltonian reads

$$H = H_{\text{Tr}} + H_{\text{Res}} + H_{\text{Int}}, \tag{3.9a}$$

where

$$H_{\text{Tr}} = \sum_{i=0}^{N_{\text{Tr}}-1} \left[4E_{Ci}(\hat{n}_i - n_{gi}(t))^2 - E_{Ji}\cos\hat{\varphi}_i \right], \tag{3.9b}$$

$$H_{\text{Res}} = \sum_{r=0}^{N_{\text{Res}}-1} \Omega_r \hat{a}_r^\dagger \hat{a}_r + \sum_{r=0}^{N_{\text{Res}}-1} \Omega_r \epsilon_r(t)(\hat{a}_r + \hat{a}_r^\dagger), \tag{3.9c}$$

$$H_{\text{Int}} = \sum_{r=0}^{N_{\text{Res}}-1} \sum_{i=0}^{N_{\text{Tr}}-1} G_{ri}\hat{n}_i(\hat{a}_r + \hat{a}_r^\dagger) \tag{3.9d}$$

$$+ \sum_{0 \le r < l < N_{\text{Res}}} \lambda_{rl}(\hat{a}_r + \hat{a}_r^\dagger)(\hat{a}_l + \hat{a}_l^\dagger) \tag{3.9e}$$

$$+ \sum_{0 \le i < j < N_{\text{Tr}}} E_{Ci,Cj}\hat{n}_i\hat{n}_j. \tag{3.9f}$$

The transmon Hamiltonian H_{Tr} given by Eq. (3.9b) is a sum of N_{Tr} CPB Hamiltonians (see Eq. (3.7)). Each transmon $i = 0, \ldots, N_{\text{Tr}} - 1$ is defined by its charging energy E_{Ci} and Josephson energy E_{Ji} (see Eqs. (3.3) and (3.4)). $\hat{n}_i$ is the number operator and $\hat{\varphi}_i$ is the phase operator of transmon i. $n_{gi}(t)$ represents a time-dependent external pulse applied to the transmon.

The resonator Hamiltonian H_{Res} given by Eq. (3.9c) is a sum of N_{Res} LC resonators (see Eq. (3.2)). Each resonator $r = 0, \ldots, N_{\text{Res}} - 1$ is defined by its frequency Ω_r. The raising and lowering operators of resonator r are given by $\hat{a}_r^\dagger$ and $\hat{a}_r$, respectively. $\epsilon_r(t)$ denotes a time-dependent pulse applied to resonator r.

Table 3.1: Characteristic values for the parameters of the model Hamiltonian defined in Eqs. (3.9a)–(3.9f). The numbers are usually determined by the corresponding experimental setups. The complexity of the simulation grows with N_{Tr} and N_{Res}, so these quantities are limited to keep the computational cost reasonable (see below). The exact numerical values of the energies, however, are irrelevant for the computational cost of the simulation.

N_{Tr}	N_{Res}	$E_{Ci}/2\pi$	$E_{Ji}/2\pi$	$\Omega_r/2\pi$	$G_{ri}/2\pi$	$\lambda_{rl}/2\pi$	$E_{Ci,Cj}/2\pi$
1–20	1–20	0.1–1 GHz	10–15 GHz	4–7 GHz	10–100 MHz	1–20 MHz	10–100 MHz

The interaction between transmons and resonators is modeled by H_{Int} given by Eqs. (3.9d)–(3.9f). The first term given in Eq. (3.9d) represents a coupling between each resonator r and each transmon i with coupling strength G_{ri}. This coupling originates from the capacitive interaction between the transmons and resonators (see [Koc2007]) since the number operator $\hat{n}_i$ describes the amount of charges stored in the CPB's capacitor and $\hat{a}_r + \hat{a}_r^\dagger$ represents the electric field of the resonator. In the simulation model, in principle, each resonator r can be coupled to each transmon i such that the matrix G_{ri} is dense. However, this is hard to realize in experiments for a system of more than a few transmons. The architecture based on coupling different transmons via resonators is characteristic of the processors available on the IBM Q Experience [IBM2016].

The second interaction term given by Eq. (3.9e) models an electric dipole interaction between the resonators. It is typically used to describe photonic interactions when a bath of resonators is used as a model for an environment [Koc2007] (see Section 3.5).

Finally, the term given by Eq. (3.9f) describes a capacitive coupling between the transmons, where $E_{Ci,Cj}$ is the capacitive coupling energy between transmon i and j. A coupling mechanism of this form is used by Rigetti Computing [Did2018] (see also [Rea2018; Cal2018]) and Google [Bar2013; Nei2018; Aru2019].

The simulation model defined by Eqs. (3.9a)–(3.9f) contains a large set of parameters and time-dependent functions which can be set to arbitrary values. However, note that in the practically relevant scenarios studied in this thesis, only a small subset of all parameters is non-zero. Typically, the non-zero parameters are set to values which have been measured in corresponding experimental setups. All energies are usually specified in gigahertz and in units of 2π (using $\hbar = 1$). Consequently, the characteristic time scale for these systems is nanoseconds. See Tab. 3.1 for characteristic values of the model parameters.

3.2.2 Choice of the basis

To solve the TDSE given by Eq. (3.8) on a digital computer, it is necessary to choose a basis for the state $|\Psi(t)\rangle$ such that the solution of the TDSE corresponds to a set of complex coefficients to be determined numerically.

For the resonators described by H_{Res} (see Eq. (3.9c)), an obvious choice for the basis vectors are the eigenstates of the photon number operator $\hat{a}_r^\dagger \hat{a}_r$ for resonator r. These are

the photon number states (or Fock states), denoted by $|k_r\rangle$ for $k_r \in \mathbb{N}_0$ such that

$$\hat{a}_r^\dagger \hat{a}_r = \sum_{k_r} k_r |k_r\rangle\langle k_r| . \tag{3.10}$$

All other terms in H_{Res} are proportional to the operator $\hat{a}_r + \hat{a}_r^\dagger$, which corresponds to the electric field in resonator r. Its representation with respect to the photon number states is

$$\hat{a}_r + \hat{a}_r^\dagger = \sum_{k_r} \sqrt{k_r + 1}(|k_r\rangle\langle k_r + 1| + |k_r + 1\rangle\langle k_r|) = \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & \sqrt{2} & & \\ & \sqrt{2} & 0 & \sqrt{3} & \\ & & \sqrt{3} & 0 & \ddots \\ & & & \ddots & \ddots \end{pmatrix}, \tag{3.11}$$

i.e., a tridiagonal symmetric matrix with zeros on the diagonal.

For the transmons described by H_{Tr} (see Eq. (3.9b)), there are two possible choices of basis states that come to mind. One option is given by the eigenstates of the charge number operators $\hat{n}_i$. They are called *charge states* and are denoted by $|n_i\rangle$ for $n_i \in \mathbb{Z}$ such that

$$\hat{n}_i = \sum_{n_i} n_i |n_i\rangle\langle n_i| . \tag{3.12}$$

The interpretation of the state $|n_i\rangle$ is that the capacitor of the Josephson junction is charged with a net charge of n_i Cooper pairs (note that the net charge can be both positive and negative such that $n_i \in \mathbb{Z}$). An advantage of this basis is that each operator in Eqs. (3.9a)–(3.9f) associated with transmon i has a straightforward representation with respect to this basis. Specifically, the operator $\cos \hat{\varphi}_i$ is given by

$$\cos \hat{\varphi}_i = \sum_{n_i} \frac{1}{2}(|n_i\rangle\langle n_i + 1| + |n_i + 1\rangle\langle n_i|). \tag{3.13}$$

It couples the charge states $|n_i\rangle$ and $|n_i \pm 1\rangle$. In other words, it describes the tunneling processes of Cooper pairs from one side of the Josephson junction to the other, thereby changing the net charge on the capacitor by one. A CPB qubit simulator solving the TDSE in this basis was studied in [Wil2016].

However, for transmon simulations, a much more appropriate basis is given by the transmon eigenstates. They are denoted by $|m_i\rangle$ for $m_i \in \mathbb{N}_0$ and correspond to the eigenstates of the full Josephson junction Hamiltonian H_{JJ} given in Eq. (3.6). This means that for each transmon i, we have

$$4E_{Ci}\hat{n}_i^2 - E_{Ji} \cos \hat{\varphi}_i = \sum_{m_i} E_{i,m_i}^{\text{Tr}} |m_i\rangle\langle m_i| , \tag{3.14}$$

where E_{i,m_i}^{Tr} denotes the corresponding eigenvalues (we typically shift the eigenvalues E_{i,m_i}^{Tr} by the respective ground-state energy $E_{i,0}^{\text{Tr}}$ such that $E_{i,0}^{\text{Tr}} = 0$ for all i). The ground state

$|m_i = 0\rangle$ and the first excited state $|m_i = 1\rangle$ are the so-called qubit states or computational states of each transmon. Correspondingly, the energy difference $\tilde{\omega}_i = E_{i,1}^{\mathrm{Tr}} - E_{i,0}^{\mathrm{Tr}}$ between the two lowest states is called the qubit frequency. Note that in practice, it may be beneficial to use a slightly shifted frequency ω_i for pulse control because of the presence of other components such as resonators (see the discussion around Eq. (3.39) or Section 5.3.2). More insight into the spectrum can be gained by observing that the $\hat{\varphi}_i$ dependence of $\cos \hat{\varphi}_i$ in leading order is $\hat{\varphi}_i^2$ (up to a constant). This means that Eq. (3.14) can be seen as an harmonic oscillator with anharmonic higher-order corrections. It can be shown that the spectrum $\{E_{i,m_i}^{\mathrm{Tr}}\}$ slightly deviates from an equidistant spectrum by an anharmonicity $\alpha_i = E_{i,2}^{\mathrm{Tr}} - E_{i,1}^{\mathrm{Tr}} - \tilde{\omega}_i \approx -E_{Ci}$ [Koc2007]. The energy difference between higher levels $m_i' + 1$ and m_i' is approximately reduced by $m_i'|\alpha_i|$ such that $E_{i,m_i'+1}^{\mathrm{Tr}} - E_{i,m_i'}^{\mathrm{Tr}} \approx \tilde{\omega}_i + \alpha m_i'$ [Gam2013].

To set up a simulation for Eqs. (3.9a)–(3.9f) in the transmon basis $|m_i\rangle$, we need to find a representation for the charge operator $\hat{n}_i$ given in Eq. (3.12) with respect to this basis. One approach is given in [Did2018], where the authors perform a systematic perturbation theory up to 25th order in the parameter $\overline{\xi}_i = \sqrt{2E_{Ci}/E_{Ji}}$. For typical device parameters (see Tab. 3.1), $\overline{\xi}_i$ takes values between 0.1 and 0.5.

However, since this work is based on a computer simulation where we have access to all numerical values of the parameters, we take a different approach to obtain the representation of $\hat{n}_i$ in the transmon basis: We construct the tridiagonal matrix for the transmon Hamiltonian given by Eq. (3.14) in terms of the charge states $|n_i\rangle$ and use numerical diagonalization to obtain the coefficients $\langle n_i|m_i\rangle$, i.e., the representation of the charge states in the transmon basis. The global phase of the transmon states is chosen such that the coefficients $\langle n_i|m_i\rangle \in \mathbb{R}$. The characteristic distribution of these coefficients is shown in Fig. 3.2. Typically, a truncation of the tridiagonal matrix given by Eq. (3.14) to 50 charge states below and above $n_i = 0$ suffices to obtain the coefficients $\langle n_i|m_i\rangle$ for the lowest states $m_i = 0, 1, 2, \ldots$ to machine precision.

Given the coefficients $\langle n_i|m_i\rangle$, we can obtain the matrix representation of the charge operator $\hat{n}_i$ in the transmon basis. The matrix is symmetric since $\langle n_i|m_i\rangle \in \mathbb{R}$, and its characteristic form is

$$\hat{n}_i = \sum_{m_i m_i'} n_i^{(m_i,m_i')} |m_i\rangle\langle m_i'| = \begin{pmatrix} & n_i^{(0,1)} & & n_i^{(0,3)} & \\ n_i^{(0,1)} & & n_i^{(1,2)} & & \cdots \\ & n_i^{(1,2)} & & n_i^{(2,3)} & \\ n_i^{(0,3)} & & n_i^{(2,3)} & & \ddots \\ & \vdots & & \ddots & \end{pmatrix}, \qquad (3.15)$$

where the coefficients $n_i^{(m_i,m_i')} = \sum_{n_i} n_i \langle m_i|n_i\rangle \langle n_i|m_i'\rangle$ are non-zero only if m_i and m_i' differ by an odd number. Moreover, if this odd number is large (i.e., the coefficient $n_i^{(m_i,m_i')}$ is far away from the diagonal), the matrix element $n_i^{(m_i,m_i')}$ tends to zero. The coefficients $n_i^{(m_i,m_i\pm1)}$ on the subdiagonal are approximately equal to those of the tridiagonal matrix $\hat{a}_r + \hat{a}_r^\dagger$ shown in Eq. (3.11). In fact, for the approximation of the transmon as an anharmonic oscillator often found in the literature, $\hat{n}_i$ is effectively replaced by the operator $-(E_{Ji}/32E_{Ci})^{1/4}(\hat{b}_i + \hat{b}_i^\dagger)$, where $\hat{b}_i + \hat{b}_i^\dagger$ is represented by the same tridiagonal matrix (see
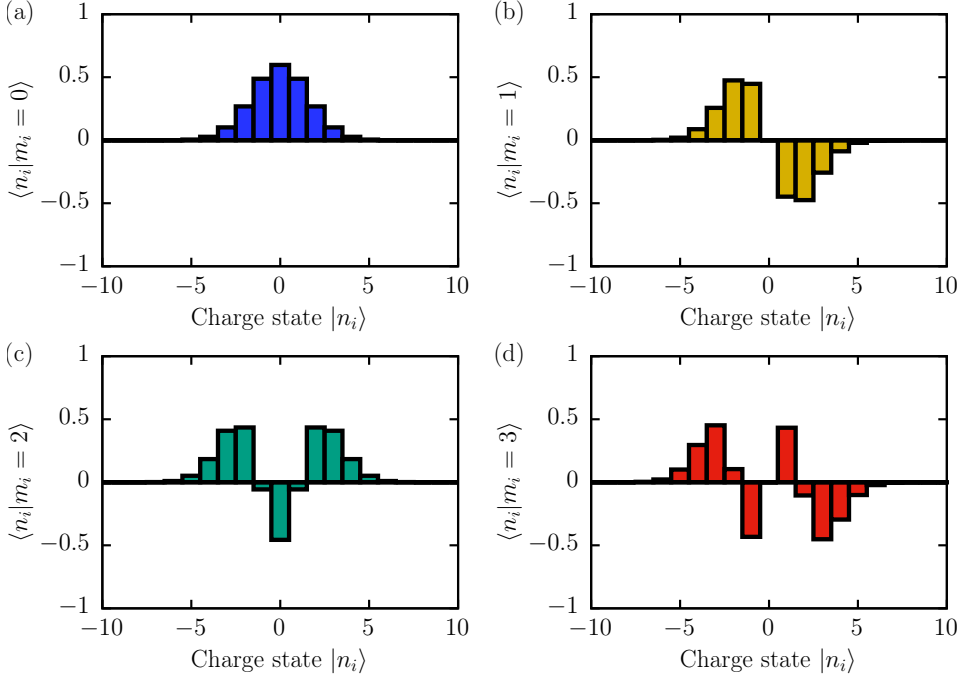
Figure 3.2: Characteristic distribution of the four lowest transmon eigenstates (a) $|m_i = 0\rangle$, (b) $|m_i = 1\rangle$, (c) $|m_i = 2\rangle$, (d) $|m_i = 3\rangle$, in terms of the charge basis $|n_i\rangle$. Shown are the coefficients $\langle n_i | m_i \rangle$ obtained from a numerical diagonalization of the tridiagonal matrix given by the left-hand side of Eq. (3.14) for $E_{Ci} = 2\pi \times 0.222\,\text{GHz}$ and $E_{Ji} = 2\pi \times 12.61\,\text{GHz}$ (the KIT system, cf. Tab. 3.2). 101 charge states $n_i = -50, \ldots, 50$ have been used for the diagonalization. Note that the sign (i.e., the global phase) of $\langle n_i | m_i \rangle$ is irrelevant, but the relative signs of the bars with even and odd parity around $n_i = 0$ is characteristic.

Section 4.2.1). The coefficient $n_i^{(0,3)}$, which is dropped in this approximation, is typically smaller than the other matrix elements by a factor of 10–50. Nevertheless, we observed that the coefficient is still significant for an accurate simulation of the time evolution. See Section 4.2 and, in particular, Fig. 4.7 for an empirical investigation of approximations of this type.

The total Hilbert space of N_{Tr} transmons and N_{Res} resonators is given by

$$\mathcal{H}_{\text{total}} = \left(\bigotimes_{r=0}^{N_{\text{Res}}-1} \bigoplus_{k_r \in \mathbb{N}_0} \text{span}\{|k_r\rangle\} \right) \otimes \left(\bigotimes_{i=0}^{N_{\text{Tr}}-1} \bigoplus_{m_i \in \mathbb{N}_0} \text{span}\{|m_i\rangle\} \right), \qquad (3.16)$$

which is an instance of the system-environment model with leakage defined in Eq. (2.30). Although the dimension of $\mathcal{H}_{\text{total}}$ is infinite, the advantage of the chosen basis is that only a small number of basis states needs to be implemented to describe the dynamics of the

system. To be precise, a transmon-resonator system initialized in the ground state usually requires four states from the transmon basis for each transmon (see Section 4.1.1 below), whereas in the charge basis $\{|n_i\rangle\}$, at least 17 states $n_i = -8, \ldots, 8$ need to be taken into account to obtain the dynamics to sufficient precision [Wil2016]. A similar argument holds for the Fock states $|k_r\rangle$ used for each of the resonators.

Therefore, for the majority of this work, we restrict the basis to four states for each of the $N_{\mathrm{Tr}} + N_{\mathrm{Res}}$ subsystems. We found that this basis is sufficient to describe the dynamics for almost all applications. Furthermore, it enables a large-scale high-performance implementation of the simulation (see Section 3.3.1). The basis states are given by the four lowest energy eigenstates $|m_i\rangle$ of each transmon and four Fock states $|k_r\rangle$ starting at some $k_r^{\mathrm{offset}} \in \mathbb{N}_0$ for each of the resonators. The offset will be set to $k_r^{\mathrm{offset}} = 0$ for most of the simulations such that the resonator parts of the basis consist of the four lowest energy Fock states for each resonator. Thus, the effective Hilbert space is given by

$$\mathcal{H} = \mathrm{span}\{|k_0\rangle|k_1\rangle\cdots|k_{N_{\mathrm{Res}}-1}\rangle|m_0\rangle|m_1\rangle\cdots|m_{N_{\mathrm{Tr}}-1}\rangle\}, \tag{3.17}$$

where each $k_r \in \{k_r^{\mathrm{offset}}, k_r^{\mathrm{offset}}+1, k_r^{\mathrm{offset}}+2, k_r^{\mathrm{offset}}+3\}$ and each $m_i \in \{0, 1, 2, 3\}$. The total number of states (i.e., the dimension of the Hilbert space) is thus $\dim(\mathcal{H}) = 4^{N_{\mathrm{Res}}+N_{\mathrm{Tr}}}$.

3.3 Simulation toolkit

The software package developed for this thesis consists of several programs for the simulation of transmon quantum computers. The central tool, `solver`, computes the solution $|\Psi(t)\rangle$ of the TDSE given in Eq. (3.8), by which any physically relevant quantity of the system can be evaluated. Additionally, the toolkit contains several convenience programs to either prepare a run for `solver` or to evaluate its results. The individual software components are:

- `solver`: Compute the numerical solution $|\Psi(t)\rangle$ of the TDSE given by Eq. (3.8) for the full model Hamiltonian H given by Eqs. (3.9a)–(3.9f). H is characterized by a set of numerical values for the parameters and pulse shapes for the time-dependent functions. The result $|\Psi(t)\rangle$ is the time evolution of a given initial state $|\Psi(0)\rangle$.

- `evaluator`: Given the solution $|\Psi(t)\rangle$ produced by `solver`, compute expectation values such as the Bloch vectors given in Eq. (2.11) or perform basis transformations to rotating frames. The program can also be used to evaluate the accuracy and the overlap between different solutions.

- `visualizer`: Generate 3D visualizations of the time evolution of the Bloch vectors computed from the solution $|\Psi(t)\rangle$ using real-time rendering.

- `optimizer`: Optimize quantum gates by finding suitable pulses (see Chapter 5). This program invokes `solver` repeatedly for varying pulse parameters and evaluates the success of implementing a certain gate by studying the time evolution produced by `solver`.

- `compiler`: Given a quantum circuit specified as a set of quantum gates in a certain file format such as OpenQASM [Cro2017] or the JUQCS instruction set defined in [DeR2019a], generate the pulse shape information required for `solver`. The program can use the results produced by `optimizer` and can also produce appropriate configuration files to set up the environment for `solver`.

In this section, we give a detailed description of the algorithms underlying `solver`, `evaluator`, and `visualizer`. Advanced functionalities implemented by `optimizer` and `compiler` are discussed in the Sections 5.3 and 5.4, respectively.

3.3.1 Numerical algorithm: `solver`

The task of `solver` is to simulate the time evolution $|\Psi(t)\rangle$ of a given initial state $|\Psi(0)\rangle$. Formally, the time evolution can be expressed as

$$|\Psi(t)\rangle = \mathcal{U}(t,0) |\Psi(0)\rangle, \tag{3.18}$$

where $\mathcal{U}$ denotes the unitary time-evolution operator of the system given by

$$\mathcal{U}(t_1,t_0) = \mathcal{T} \exp\left(-i \int_{t_0}^{t_1} H(\tilde{t}) \, \mathrm{d}\tilde{t}\right). \tag{3.19}$$

In this expression, $H(\tilde{t})$ is the time-dependent model Hamiltonian defined in Eqs. (3.9a)–(3.9f) and $\mathcal{T}$ is the time-ordering symbol (see [Bru2004]). The time-evolution operator satisfies the group property $\mathcal{U}(t_1,t_0) = \mathcal{U}(t_1,t') \mathcal{U}(t',t_0)$ such that the time evolution from 0 to t can be written as a product of successive time evolutions with a time step τ, i.e.,

$$|\Psi(t)\rangle = \mathcal{U}(t,t-\tau) \cdots \mathcal{U}(2\tau,\tau) \mathcal{U}(\tau,0) |\Psi(0)\rangle. \tag{3.20}$$

In each step, the current state vector $|\Psi(t_0)\rangle$ is updated by computing

$$|\Psi(t_0+\tau)\rangle = \mathcal{U}(t_0+\tau,t_0) |\Psi(t_0)\rangle, \tag{3.21}$$

starting from $t_0 = 0$. Here, we have chosen the time step τ to be constant, but in principle it may be updated dynamically for each time step to speed up the simulation, if the problem allows it. In general, the time step τ needs to be sufficiently small such that $H(t_0)$ and $H(t_0+\tau)$ are well approximated by $H(t_0+\tau/2)$ (this condition is affected by the time dependence of the pulses $n_{gi}(t)$ and $\epsilon_r(t)$ in Eqs. (3.9b) and (3.9c), respectively). In this case, the time-evolution operator $\mathcal{U}(t_0+\tau,t_0)$ (which is equal to Eq. (3.19) for $t_1 = t_0+\tau$) simplifies to $\mathcal{U}(t_0+\tau,t_0) = e^{-i\tau H(t_0+\tau/2)}$. The update rule for the time evolution given by Eq. (3.21) then becomes

$$|\Psi(t_0+\tau)\rangle = e^{-i\tau H(t_0+\tau/2)} |\Psi(t_0)\rangle. \tag{3.22}$$

The core of the transmon simulator implements this operation to propagate the state vector according to Eq. (3.20). Obviously, there are many different ways to implement Eq. (3.22), but whether a simulation for large systems is feasible or not heavily depends

on the choice of basis and the particular algorithm. A suitable choice for the basis is the product basis of resonator eigenstates $|k_r\rangle$ and transmon eigenstates $|m_i\rangle$ discussed in Section 3.2.2.

In this basis, the solution $|\Psi(t)\rangle$ is determined by its complex expansion coefficients $\psi_{k_0 k_1 \cdots m_0 m_1 \cdots}(t)$ defined by

$$|\Psi(t)\rangle = \sum_{\substack{k_0 k_1 \cdots \\ m_0 m_1 \cdots}} \psi_{k_0 k_1 \cdots m_0 m_1 \cdots}(t) \, |k_0 k_1 \cdots m_0 m_1 \cdots\rangle. \tag{3.23}$$

As each k_r and m_i in the truncated Hilbert space $\mathcal{H}$ (see Eq. (3.17)) can take one of four different values, $|\Psi(t)\rangle$ is described by an array of $4^{N_{\mathrm{Res}}+N_{\mathrm{Tr}}}$ complex coefficients. The index $k_0 k_1 \cdots m_0 m_1 \cdots$ can thus be efficiently encoded in an integer of $2(N_{\mathrm{Res}} + N_{\mathrm{Tr}})$ bits. We introduce the notation

$$\mathtt{KM} = \underbrace{\overset{\overset{\textstyle k_0\ k_1}{\textstyle \uparrow\ \uparrow}}{00\,00} \cdots 00}_{\mathtt{K}}\, \underbrace{\overset{\overset{\textstyle m_0 m_1}{\textstyle \uparrow\ \uparrow}}{00\,00} \cdots 00}_{\mathtt{M}},\ \ldots,\ \underbrace{11 \cdots 11}_{\mathtt{K}}\, \underbrace{11 \cdots 11}_{\mathtt{M}} \tag{3.24}$$

for this integer, where each group of two bits encodes the respective value of $k_r \in \{k_r^{\mathrm{offset}}, k_r^{\mathrm{offset}} + 1, k_r^{\mathrm{offset}} + 2, k_r^{\mathrm{offset}} + 3\}$ and $m_i \in \{0, 1, 2, 3\}$. Using this notation, the expansion in Eq. (3.23) reads

$$|\Psi(t)\rangle = \sum_{\mathtt{KM}} \psi_{\mathtt{KM}}(t) \, |\mathtt{KM}\rangle. \tag{3.25}$$

We choose unsigned 64-bit integers to represent the index $\mathtt{KM}$ in the implementations (see Appendix D). This choice is reasonable since the number of bits required for N_{Tr} transmons and N_{Res} resonators is $2(N_{\mathrm{Tr}} + N_{\mathrm{Res}})$. This means that 32-bit integers would impose the unnecessary constraint $N_{\mathrm{Tr}} + N_{\mathrm{Res}} \leq 16$ even though larger simulations are technically feasible. Integers larger than 64 bits are not required since there exists no system today that can store more than 2^{64} complex numbers.

We have already studied the matrix representation with respect to $|\mathtt{KM}\rangle$ for parts of the Hamiltonian in Section 3.2.2. This knowledge can be used to derive an algorithm to implement the operation given by Eq. (3.22).

Suzuki-Trotter product-formula algorithm

The algorithm that we use to implement Eq. (3.22) is the second-order Suzuki-Trotter product-formula algorithm [DeR1987]. It is based on generalizations of the Lie-Trotter formula [Lie1888; Tro1959; Suz1976; Suz1985]. The algorithm belongs to a family of explicit and unconditionally stable algorithms for linear parabolic difference equations and has been applied numerous times to solve the TDSE, especially also in the context of quantum computing [DeR2000; DeR2002; DeR2006; Wil2016; Wil2017; Wil2018b; Wil2020d; Wil2020e]. Moreover, rigorous error bounds have been proven to assess the accuracy of the algorithm [DeR1987; Huy1990]. Recently, the bounds have been extended to tight error bounds for observables [Wil2020d]. We test these bounds in Section 4.1.

The first step in the derivation of the algorithm is to split the Hamiltonian H given by Eqs. (3.9a)–(3.9f) at time $\tilde{t} = t_0 + \tau/2$ into a part H_0 that is diagonal with respect to $|\mathtt{KM}\rangle$ and the remaining part W. We obtain up to an irrelevant constant (which would only lead to a global phase in the solution),

$$H = H_0 + W, \tag{3.26a}$$

$$H_0 = \sum_{i=0}^{N_{\mathrm{Tr}}-1} (4E_{Ci}\hat{n}_i^2 - E_{Ji}\cos\hat{\varphi}_i) + \sum_{r=0}^{N_{\mathrm{Res}}-1} \Omega_r \hat{a}_r^\dagger \hat{a}_r, \tag{3.26b}$$

$$W = \sum_{i=0}^{N_{\mathrm{Tr}}-1} -8E_{Ci}n_{gi}(\tilde{t})\hat{n}_i \tag{3.26c}$$

$$+ \sum_{r=0}^{N_{\mathrm{Res}}-1} \Omega_r \varepsilon_r(\tilde{t})(\hat{a}_r + \hat{a}_r^\dagger) \tag{3.26d}$$

$$+ \sum_{r=0}^{N_{\mathrm{Res}}-1} \sum_{i=0}^{N_{\mathrm{Tr}}-1} G_{ri}\hat{n}_i(\hat{a}_r + \hat{a}_r^\dagger) \tag{3.26e}$$

$$+ \sum_{0 \le r < l < N_{\mathrm{Res}}} \lambda_{rl}(\hat{a}_r + \hat{a}_r^\dagger)(\hat{a}_l + \hat{a}_l^\dagger) \tag{3.26f}$$

$$+ \sum_{0 \le i < j < N_{\mathrm{Tr}}} E_{Ci,Cj}\hat{n}_i\hat{n}_j. \tag{3.26g}$$

The second-order product-formula decomposition for $\mathcal{U}(t_0 + \tau, t_0) = e^{-i\tau(H_0+W)}$ is given by

$$\widetilde{\mathcal{U}} = e^{-i\tau H_0/2} \, e^{-i\tau W} \, e^{-i\tau H_0/2}. \tag{3.27}$$

This decomposition is equal to $\mathcal{U}(t_0 + \tau, t_0)$ up to second order in τ. It is the only approximation apart from the discretization in time, and the error can be well controlled by the time step τ using the rigorous bounds given in [DeR1987; Wil2020d]. To apply the decomposition to the state vector given in Eq. (3.25), we need to derive its action on the basis state $|\mathtt{KM}\rangle$.

For the diagonal part $e^{-i\tau H_0/2}$, we make use of the spectral representations given in Eqs. (3.10) and (3.14), yielding

$$H_0 = \sum_{\mathtt{KM}} \left(k_0\Omega_0 + k_1\Omega_1 + \cdots + E_{0,m_0}^{\mathrm{Tr}} + E_{1,m_1}^{\mathrm{Tr}} + \cdots\right) |\mathtt{KM}\rangle\langle\mathtt{KM}|. \tag{3.28}$$

Thus we obtain the explicit expression

$$e^{-i\tau H_0/2}|\mathtt{KM}\rangle = \exp(-i\tau(k_0\Omega_0 + k_1\Omega_1 + \cdots + E_{0,m_0}^{\mathrm{Tr}} + E_{1,m_1}^{\mathrm{Tr}} + \cdots)/2)|\mathtt{KM}\rangle, \tag{3.29}$$

which can be implemented on a computer by multiplying each coefficient $\psi_{\mathtt{KM}}$ with the corresponding phase factor. Since all operations are independent, this operation can be easily parallelized.

The action of the non-diagonal part $e^{-i\tau W}$ on $|\text{KM}\rangle$ is more complicated. One option would be to apply the Suzuki-Trotter product-formula decomposition again to each of the terms contained in W (see Eqs. (3.26c)–(3.26g)) and evaluate the smaller matrix exponentials explicitly. Such an approach was used in [Wil2016]. However, in the present case, a non-diagonal term like $\hat{n}_i \hat{n}_j$ in Eq. (3.26g) would result in 16-component updates (cf. Eq. (3.15)) of the state vector given by Eq. (3.25). Furthermore, additional second-order decompositions of the matrix exponential $e^{-i\tau W}$ would introduce additional errors of order τ^3 such that the time step τ would need to be reduced.

Therefore, we follow a different route by constructing a transformation V to change to the eigenbasis of W such that $\Lambda = V^\dagger W V$ is diagonal. This means that the operator $e^{-i\tau W}$ in Eq. (3.27) can be implemented as $V\, e^{-i\tau\Lambda}\, V^\dagger$. Using this, the decomposition used to implement the time step Eq. (3.21) reads

$$\widetilde{\mathcal{U}} = e^{-i\tau H_0/2}\, V\, e^{-i\tau\Lambda}\, V^\dagger\, e^{-i\tau H_0/2}. \tag{3.30}$$

Since the Hilbert space given by Eq. (3.17) is a product of multiple spaces with only four dimensions each, V can be written as a tensor product of complex 4×4 matrices,

$$V = \bigotimes_{r=0}^{N_{\text{Res}}-1} V_r^{(a)} \bigotimes_{i=0}^{N_{\text{Tr}}-1} V_i^{(n)}, \tag{3.31}$$

where $V_r^{(a)}$ diagonalizes the matrix representation of $\hat{a}_r + \hat{a}_r^\dagger$ in the Fock basis (see Eq. (3.11) and $V_i^{(n)}$ diagonalizes the matrix representation of $\hat{n}_i$ in the transmon basis (see Eq. (3.15)). We obtain these 4×4 matrices numerically by diagonalizing the corresponding 4×4 matrix representations given in Eq. (3.11) and Eq. (3.15), respectively. The matrices $V_r^{(a)}$ and $V_i^{(n)}$ only need to be computed once, i.e., before the actual time evolution starts. Furthermore, the matrices $V_r^{(a)}$ are all equal for different r, so only one of them needs to be stored in memory. In summary, we have

$$V_r^{(a)}\, \Lambda_r^{(a)}\, V_r^{(a)\dagger} = \begin{pmatrix} & \sqrt{1+k_r^{\text{offset}}} & & \\ \sqrt{1+k_r^{\text{offset}}} & & \sqrt{2+k_r^{\text{offset}}} & \\ & \sqrt{2+k_r^{\text{offset}}} & & \sqrt{3+k_r^{\text{offset}}} \\ & & \sqrt{3+k_r^{\text{offset}}} & \end{pmatrix}, \tag{3.32a}$$

$$V_i^{(n)}\, \Lambda_i^{(n)}\, V_i^{(n)\dagger} = \begin{pmatrix} & n_i^{(0,1)} & & n_i^{(0,3)} \\ n_i^{(0,1)} & & n_i^{(1,2)} & \\ & n_i^{(1,2)} & & n_i^{(2,3)} \\ n_i^{(0,3)} & & n_i^{(2,3)} & \end{pmatrix}, \tag{3.32b}$$

where $n_i^{(m_i, m_i')} = \sum_{n_i} n_i \langle m_i | n_i \rangle \langle n_i | m_i' \rangle$ (see Eq. (3.15)), and the corresponding eigenvalues are contained in the diagonal matrices $\Lambda_r^{(a)}$ and $\Lambda_i^{(n)}$, respectively.

A proof that V defined in Eq. (3.31) diagonalizes W given by Eqs. (3.26c)–(3.26g) can

be obtained by explicitly computing $\Lambda = V^\dagger W V$:

$$
\Lambda = \sum_{i=0}^{N_{\mathrm{Tr}}-1} -8E_{Ci} n_{gi}(\tilde{t}) \Lambda_i^{(n)} + \sum_{r=0}^{N_{\mathrm{Res}}-1} \Omega_r \varepsilon_r(\tilde{t}) \Lambda_r^{(a)} + \sum_{r=0}^{N_{\mathrm{Res}}-1} \sum_{i=0}^{N_{\mathrm{Tr}}-1} G_{ri} \Lambda_r^{(a)} \otimes \Lambda_i^{(n)}
$$
$$
+ \sum_{0 \le r < l < N_{\mathrm{Res}}} \lambda_{rl} \Lambda_r^{(a)} \otimes \Lambda_l^{(a)} + \sum_{0 \le i < j < N_{\mathrm{Tr}}} E_{Ci,Cj} \Lambda_i^{(n)} \otimes \Lambda_j^{(n)}, \tag{3.33}
$$

which is a direct expression for the eigenvalues of W. It can be used to implement the operation $e^{-i\tau\Lambda}$ in Eq. (3.30) in the same way as $e^{-i\tau H_0/2}$ in Eq. (3.29).

The only thing left for the implementation of $\widetilde{\mathcal{U}}$ given in Eq. (3.30) is the implementation of the basis transformation V. Each 4×4 component $V_{r/i}^{(a/n)}$ of V in Eq. (3.31) results in four-component updates of the coefficients $\psi_{\mathtt{KM}}$ of the form

$$
\begin{pmatrix} \psi_{*\cdots*00*\cdots*} \\ \psi_{*\cdots*01*\cdots*} \\ \psi_{*\cdots*10*\cdots*} \\ \psi_{*\cdots*11*\cdots*} \end{pmatrix} \leftarrow V_{r/i}^{(a/n)} \begin{pmatrix} \psi_{*\cdots*00*\cdots*} \\ \psi_{*\cdots*01*\cdots*} \\ \psi_{*\cdots*10*\cdots*} \\ \psi_{*\cdots*11*\cdots*} \end{pmatrix}, \tag{3.34}
$$

where the notation $*\cdots*$ indicates that the 4×4 transformation $V_r^{(a)}$ $(V_i^{(n)})$ needs to be done in a loop over $\mathtt{KM} = 0, \ldots, 4^{N_{\mathrm{Tr}}+N_{\mathrm{Res}}} - 1$ where the two bits corresponding to k_r (m_i) are fixed (cf. Eq. (3.24)).

We study three alternatives to implement this loop over $\mathtt{KM}$ on a supercomputer. There is a priori no guarantee which of the implementations performs best on which processor. It is reasonable to focus mainly on optimizing this part since it makes the largest contribution to the run time of the algorithm (see Fig. 4.5 below). For this reason, we compare the alternative implementations empirically in Section 4.1.2. A `C++` sample implementation for each is given in Listings D.1–D.3 in Appendix D.

Implementation 0: Complete single loop with branches

The simplest approach consists of a complete loop over all $\mathtt{KM}$ from 0 to $\dim(\mathcal{H}) - 1 = 4^{N_{\mathrm{Tr}}+N_{\mathrm{Res}}} - 1$. In each iteration, we test if the two bits corresponding to the current transformation are 0 (i.e., if $\mathtt{KM} = *\cdots*00*\cdots*$). If they are, we perform the 4×4 update of the coefficients $\psi_{\mathtt{KM}}$. In other words, we iterate over all $\mathtt{KM}$, but only do something every fourth iteration.

This implementation might seem naive since the inner loop has four times as many iterations as it needs. Moreover, the test $\mathtt{KM} = *\cdots*00*\cdots*$ introduces branches in the code which may interrupt the sequential flow in the instruction pipeline of the processor.

However, modern processors use branch predictors to detect patterns in the evaluation of conditional structures [Smi1981; Mit2019]. This means that instructions which are likely to follow the branch are already loaded into the pipeline before knowing if they are really going to be executed. Since the branch under consideration (see lines 6 and 13 of Listing D.1) has an easy pattern (it evaluates to true every four iterations), it could be that branch prediction effectively removes the overhead. Therefore, the implementation is worth studying in more detail, also to assess the impact that optimizations on this level can have with modern compilers and processors.

Implementation 1: Reduced single loop with bitwise operations

The next implementation explicitly reduces the number of iterations in the inner loop by a factor of four. The price to be paid is that the actual index KM needs to be computed from the reduced iteration count. This is done by means of additional bitwise operations (see lines 8 and 18 of Listing D.2).

   This implementation has the smallest amount of branches, at the cost of additional computation required for the bitwise operations to obtain the index KM. Furthermore, it requires the largest amount of code and is less readable than the other implementations. This approach might have been the first choice for older architectures where branches in performance-critical code directly result in an increased run time. It is interesting to see if this intuition also holds for modern processors.

Implementation 2: Reduced nested loops

The last implementation divides the loop over KM into two separate, nested loops over the higher part K and the lower part M of the index KM (see Eq. (3.24)). A potential problem of this implementation is that the loops themselves also introduce branches, and the evaluation of the tests may not be as predictable as the branches in implementation 0.

   However, modern processors are well tuned to the execution of loops with simple conditions and increments, and separate loops over K and M are easy to parse and parallelize. This implementation conveys the programmer's intent more clearly and does not contain as many explicit bitwise operations as implementation 1 or additional iterations as implementation 0.

Storage of the results and the computational subspace

The results produced by `solver` are the complex coefficients $\psi_{\text{KM}}(t)$ of the state vector $|\Psi(t)\rangle$ given in Eq. (3.25) at certain times $t \in \mathbb{T}$ during the time evolution. The set $\mathbb{T}$ usually consists of the times after each pulse implementing a certain quantum gate. The maximum number of times in $\mathbb{T}$ is determined by the total duration of the time evolution divided by the time step τ. The particular set of times $\mathbb{T}$ at which the coefficients are saved is often much smaller than the total number of time steps.

   The coefficients $\psi_{\text{KM}}(t)$ are typically stored in separate text files containing the modulus abs($\psi_{\text{KM}}(t)$) and the argument arg($\psi_{\text{KM}}(t)$) of the complex numbers. Additionally, the current state vector can be saved in binary format. This is useful when the simulation is interrupted and needs to be continued at a later point in time; for instance, if the simulation takes longer than the maximum time that a job can allocate on a supercomputer.

   Sometimes, it may not be feasible or necessary to store the coefficients $\psi_{\text{KM}}(t)$ for all $4^{N_{\text{Res}}+N_{\text{Tr}}}$ values of the index KM (see Eq. (3.24)). It is often sufficient to consider only the projection of $|\Psi(t)\rangle$ on the so-called *computational subspace* of the Hilbert space $\mathcal{H}$ given by Eq. (3.17). The projection can be formally written in terms a projection operator

defined by

$$P_{\mathcal{H}_{2^n}} |\mathtt{KM}\rangle = \begin{cases} |\mathtt{KM}\rangle & \text{if } k_0, k_1, \ldots = 0 \text{ and } m_0, m_1, \ldots \in \{0, 1\} \\ 0 & \text{otherwise} \end{cases}, \qquad (3.35)$$

where $n = N_{\mathrm{Tr}}$. Hence $P_{\mathcal{H}_{2^n}}$ keeps only those states $|\mathtt{KM}\rangle$ for which the resonator part $\mathtt{K} = 0$ and each m_i in the transmon part $\mathtt{M}$ is either 0 or 1.

Since the range of the operator $P_{\mathcal{H}_{2^n}}$ is a subspace of dimension 2^n, it can be identified with the multi-qubit space $\mathcal{H}_{2^n}$ defined in Eq. (2.7). This is the computational subspace of $\mathcal{H}$. Consequently, we call all other states $|\mathtt{KM}\rangle$ for which $P_{\mathcal{H}_{2^n}} |\mathtt{KM}\rangle = 0$ *higher levels* or *non-computational states* of $\mathcal{H}$ (see also Section 2.4.2). This concept will play an important role in Chapter 5, when we study and optimize pulses $n_{gi}(t)$ to implement a set of quantum gates on the computational subspace.

Note that a projection on the computational subspace may not be sufficient to compute any arbitrary observable for the system, since transmons may suffer from leakage or become entangled with resonators. In the latter case, a partial trace over the resonators' degrees of freedom would be more appropriate (see e.g. Eq. (3.38) below). However, for the optimization of quantum gates, where a pulse is explicitly optimized to render the transmon within the computational subspace, the projection defined by Eq. (3.35) is adequate.

3.3.2 Evaluation of the results: `evaluator`

The first step after setting up a new simulation (i.e., after specifying the device parameters for the Hamiltonian including potential time-dependent pulses) is to configure the simulation parameters to ensure that the results are accurate up to a certain numerical precision. Afterwards, qubit-specific properties such as the qubit frequency or the Bloch vectors can be calculated from the results. These tasks are provided by `evaluator`. They typically need to be done before pulses are optimized to implement the quantum gates.

Adjusting the time step by monitoring overlap and error

A crucial parameter of the simulation is the time step τ used to solve the TDSE (see the discussion below Eq. (3.20)). Despite the existence of rigorous error bounds (see [DeR1987; Wil2020d]), it is often crucial to tweak τ such that the simulation produces results equal to the mathematical solution of the TDSE (up to some desired precision), but still runs in reasonable time on a (super)computer. In practice, one usually starts with a small time step and gradually increases τ as long as the resulting state vectors $|\Psi(t)\rangle$ (or certain desired expectation values) effectively stay the same. For most of the simulations, we use time steps $\tau \in \{10^{-3}\,\mathrm{ns}, 10^{-4}\,\mathrm{ns}\}$.

To check whether two state vectors $|\Psi^{\tau_1}(t)\rangle$ and $|\Psi^{\tau_2}(t)\rangle$ resulting from simulations with different time steps $\tau_1 < \tau_2$ are effectively the same, `evaluator` provides an option to compute the respective overlap given by

$$\mathrm{overlap}(t) = \frac{|\langle \Psi^{\tau_1}(t)|\Psi^{\tau_2}(t)\rangle|^2}{\langle \Psi^{\tau_1}(t)|\Psi^{\tau_1}(t)\rangle \langle \Psi^{\tau_2}(t)|\Psi^{\tau_2}(t)\rangle} \in [0, 1]. \qquad (3.36)$$

This quantity has the advantage of being independent of a global phase difference between each of the state vectors. A difference in the global phase is typically the first observable numerical error caused by an increased time step, but it is irrelevant for computing physically meaningful quantities. Although the simulation result $|\Psi(t)\rangle$ is always normalized since the Suzuki-Trotter product-formula algorithm is unitary by definition [DeR1987], Eq. (3.36) includes an explicit normalization of both state vectors in the denominator. The reason for this is that `evaluator` can then also be applied to a projection of the state vectors on the computational subspace (see Eq. (3.35)).

A larger time step $\tau_2 > \tau_1$ is sufficient for the simulation if overlap(t) is 1 for all times $t \in \mathbb{T}$. For convenience, `evaluator` also computes the average error between $|\Psi^{\tau_1}(t)\rangle$ and $|\Psi^{\tau_2}(t)\rangle$ given by

$$1 - \frac{1}{|\mathbb{T}|} \sum_{t \in \mathbb{T}} \text{overlap}(t), \qquad (3.37)$$

such that only a single number needs to be monitored when configuring the time step.

We present results from applying this procedure in practice at the end of Section 4.1.1. Additionally, we study the behavior of local and global errors with respect to rigorous error bounds (see Figs. 4.1 and 4.2). However, note that a study of the error bounds requires small, undriven systems, whereas the practical procedure outlined in this section also works for larger systems with time-dependent Hamiltonians.

Computing Bloch vectors

Given the coefficients $\psi_{\texttt{KM}}(t)$ of the resulting state vector $|\Psi(t)\rangle$, one could in principle compute its projection on the computational subspace (see Eq. (3.35)) and evaluate the Bloch vectors $\vec{r}_i(t)$ for each qubit according to Eq. (2.11).

However, a characteristic problem of transmon qubits is that a significant part of the state may lie outside the computational subspace during the time evolution. This problem is known as leakage [Che2016; Woo2018; Wil2017] (see also Section 2.1.4). In this case, a simple projection as defined by Eq. (3.35) may not be sufficient to compute the observables $\vec{r}_i(t) = \langle \Psi(t)|\vec{\sigma}_i|\Psi(t)\rangle$. Instead, we need to trace over all other degrees of freedom of the Hilbert space. Hence, Eq. (2.11) becomes

$$\vec{r}_i(t) = \sum_{k_0 k_1 \cdots} \sum_{\substack{m_0 m_1 \cdots \\ \text{without } m_i}} \begin{pmatrix} 2\,\text{Re}(\psi^*_{k_0 k_1 \cdots m_0 m_1 \cdots (m_i=0) \cdots}(t)\psi_{k_0 k_1 \cdots m_0 m_1 \cdots (m_i=1) \cdots}(t)) \\ 2\,\text{Im}(\psi^*_{k_0 k_1 \cdots m_0 m_1 \cdots (m_i=0) \cdots}(t)\psi_{k_0 k_1 \cdots m_0 m_1 \cdots (m_i=1) \cdots}(t)) \\ |\psi_{k_0 k_1 \cdots m_0 m_1 \cdots (m_i=0) \cdots}(t)|^2 - |\psi_{k_0 k_1 \cdots m_0 m_1 \cdots (m_i=1) \cdots}(t)|^2 \end{pmatrix}, \quad (3.38)$$

where all indices $\texttt{KM} = k_0 k_1 \cdots m_0 m_1 \cdots$ without m_i enumerate all other basis states included in the simulation (cf. Eq. (3.17)).

Determining qubit frequencies

The computational basis states for each qubit i are given by the respective lowest-energy transmon eigenstates $|m_i = 0\rangle$ and $|m_i = 1\rangle$. The energy difference between these states is called the qubit transition frequency ω_i, because it corresponds to the frequency that

an externally applied pulse needs to have to drive transitions between the states. It may not be equal to the qubit frequency $\tilde{\omega}_i$ obtained from diagonalization (see the discussion below Eq. (3.14)) due to the presence of other transmons and resonators in the system.

One option to measure this frequency makes use of the fact that the time evolution of the state $|1\rangle$ results in a relative phase factor $e^{-i\omega_i t}$ between the states $|0\rangle$ and $|1\rangle$. If the qubit is prepared in the uniform superposition $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, its time evolution yields a Bloch vector that rotates around the z axis with a frequency of ω_i (see Eq. (2.5)), i.e.

$$\vec{r}_i^{\,(\text{theory})}(t) = \begin{pmatrix} \cos(\omega_i t) \\ -\sin(\omega_i t) \\ 0 \end{pmatrix}. \tag{3.39}$$

We can thus infer a good candidate for the effective frequency of the qubit by fitting ω_i in this equation to the time evolution of the Bloch vector of qubit i computed using Eq. (3.38). Specifically, this means that we first prepare the qubit in the state $(|m_i = 0\rangle + |m_i = 1\rangle)/\sqrt{2}$, and then have it evolve freely for a certain time T. The other qubits are all prepared in the state $|0\rangle$ (this protocol is in agreement with the experimental procedure; see, for instance, the red squares in Fig. 7.3(b) obtained from experiments on the `ibmqx4` processor [IBM2018b]). At each time t, we then compute the respective qubit's Bloch vector according to Eq. (3.38), and finally fit Eq. (3.39) to the data. The squared error for this fit is given by

$$\chi^2(\omega_i) = \sum_{n=1}^{N_{\text{data}}} \left[(r_i^x(t_n) - \cos(\omega_i t_n))^2 + (r_i^y(t_n) + \sin(\omega_i t_n))^2 \right], \tag{3.40}$$

where N_{data} is the number of points included in the fit (usually much smaller than the total number of time steps required for the simulation), t_n is the nth point in time (comprising the set $\mathbb{T}$ above), and $r_i^x(t_n)$ ($r_i^y(t_n)$) is the x (y) component of $\vec{r}_i(t)$ given by Eq. (3.38) at time t_n. Note that it is advantageous to include the data for both "quadratures" $r_i^x(t_n)$ and $r_i^y(t_n)$ in the error function, instead of only fitting a cosine function to $r_i^x(t_n)$. Otherwise the function can have additional extrema and be harder to minimize properly.

We apply this procedure to determine frequencies since it emulates a typical experimental procedure to infer qubit frequencies. An alternative exact diagonalization of the full system to determine its eigenenergies is usually not feasible for larger systems. The time T is typically chosen on the order of 1000 ns and the number of data points included in the fit is $N_{\text{data}} = 10000$. The inferred qubit frequency ω_i will later serve as an initial value for the drive frequency of the pulse optimizations to implement quantum gates.

A technical difficulty of the procedure is that the error function $\chi^2(\omega_i)$ given by Eq. (3.40) is a strongly oscillating function of ω_i with many local minima and only one sharply peaked global minimum (see for instance Fig. 3.6 below). Therefore, standard fitting routines might have problems in locating the right minimum.

We apply a method called *Golden Section Search* which is designed to handle the worst possible case of one-dimensional function minimization [Pre2007]. The method brackets the minimum by maintaining a triplet of points $\omega_a < \omega_b < \omega_c$, and chooses the next point

to be the golden mean point (closer to ω_b) within the larger segment of $\omega_b - \omega_a$ and $\omega_c - \omega_b$. In each step, the bracketing interval $\omega_c - \omega_a$ will be a factor of $(\sqrt{5} - 1)/2 \approx 0.61803$ (the inverse golden section) smaller than the preceding interval. This particular ratio stems from an optimality condition for function minimization similar to the bisection method for finding zeros (see [Pre2007] for more information).

Transformation to the rotating frame

During a free time evolution, the Bloch vector $\vec{r}_i(t)$ of qubit i describes rotations around the z axis (see Eq. (3.39)) at the frequency ω_i of the qubit. For transmon qubits, for which ω_i can be around $2\pi \times 5\,\mathrm{GHz}$, typical quantum gate implementations (such as rotations of $\vec{r}_i(t)$ around the x or z axis) may take approximately $80\,\mathrm{ns}$ (see Eq. (5.9) in Section 5.1). Thus, $\vec{r}_i(t)$ performs a large number of rotations during the time needed to apply one gate.

For this reason, it is convenient to describe the qubit in a basis rotating at the qubit's frequency, both for the description of the pulses to implement quantum gates and also for the purpose of visualization. This basis is commonly called the *rotating frame* (the other basis is often referred to as the *lab frame* in this context). The rotating frame is defined as a change to a time-dependent basis according to $(|0\rangle, |1\rangle) \mapsto (|0\rangle, \exp(-i\omega_i t)|1\rangle)$, effectively removing the relative phase between the computational basis states mentioned above.

For the coefficients $\psi_{k_0 k_1 \cdots m_0 m_1 \cdots}(t)$ of the solution $|\Psi(t)\rangle$ of the transmon simulation, this change of basis amounts to replacing

$$\psi_{k_0 k_1 \cdots m_0 m_1 \cdots}(t) \mapsto e^{it \sum_i \omega_i m_i} \psi_{k_0 k_1 \cdots m_0 m_1 \cdots}(t). \tag{3.41}$$

Note that this transformation only removes the relative phase factors between states corresponding to $m_i = 0$ and $m_i = 1$. It does not completely remove relative phases of higher non-computational states $m_i > 1$ since transmon eigenenergies are not exactly equidistant (see e.g. [Koc2007]). However, this is also not required as we are only interested in the computational states when describing transmon qubits in a rotating frame. Furthermore, the transformation Eq. (3.41) does not affect the probability $|\psi_{k_0 k_1 \cdots m_0 m_1 \cdots}(t)|^2$.

The geometrical effect of the rotating frame is that the Bloch vector $\vec{r}_i(t)$ given in Eq. (3.38), after replacing the coefficients according to Eq. (3.41), effectively stands still. As shown below, this only holds on average since the influence of higher levels and crosstalk will typically make the Bloch vectors shrink, wiggle, or continue to rotate slowly in time (see, for instance, Fig. 7.2(b)). The goal of the pulse optimizations is then to tune the pulse parameters to capture this effect such that the result of applying a quantum gate also displaces the qubit's Bloch vector into the desired position.

3.3.3 Visualization of the results: `visualizer`

Quantum gate implementations for transmon qubits typically consist of time-dependent pulses whose purpose is to rotate the qubits' Bloch vectors in their respective Bloch spheres (see Fig. 2.1). This is a continuous, analog operation which can be easily visualized, especially for single-qubit gates. To engineer and assess the effect of different pulses, it is instructive to study the time evolution of each $\vec{r}_i(t)$ computed from Eq. (3.38) (potentially in a rotating frame, see Eq. (3.41)).

For this purpose, `visualizer` processes the coordinates of each $\vec{r}_i(t)$ and renders a three-dimensional scene that can be navigated in time and space using mouse and keyboard. The program makes use of the high-performance cross-platform open-source engine Irrlicht [Geb2005]. It is written in `C++` and builds on the real-time renderer that has been developed for the work presented in [Wil2016]. See Appendix A for some example renderings.

3.4 Definition of the model systems

In this section, we give a definition of the most important transmon systems used in the following chapters. Each system is characterized by its parameters for the full model Hamiltonian given by Eqs. (3.9a)–(3.9f). The values of the device parameters and the topology of the systems are inspired by various transmon systems used for quantum processors in experiments.

3.4.1 Single transmon-resonator system

The simplest system studied in this work consists of a single transmon coupled to a single resonator. It models a system used in an experiment that has been conducted at the Karlsruhe Institute of Technology (KIT). The parameters are given in Tab. 3.2. This system is used for two purposes in Section 4.2 of the following chapter, namely to study the validity of perturbative approaches and to assess the influence of higher photon numbers in the resonator. It is also part of the environment model defined in the next section.

3.4.2 Transmon-resonator system coupled to a bath

We consider the single transmon-resonator system from Section 3.4.1, coupled to a separate bath of harmonic oscillators. The form of the model is sketched in Fig. 3.3. Although only the resonator is explicitly connected to the bath, the model is completely general: as we show in Section 3.5, it is uniquely related to the Foster representation of a superconducting environment [Nig2012]. This model is used extensively in Section 4.3, where we use the bath approach to study the transition from an isolated system to an open quantum system described by a quantum master equation. The particular model parameters chosen for this study are given in Tab. 3.3.

Table 3.2: Model parameters for a device operated at KIT with a single transmon coupled to a readout resonator [Rie2019], simulated by solving the TDSE for the model Hamiltonian given by Eqs. (3.9a)–(3.9f). The indices $i = 0$ and $r = 0$ have been dropped for simplicity. All energies are expressed in GHz ($\hbar = 1$). Unspecified parameters are set to zero. An estimate for the qubit frequency ω and the anharmonicity α are obtained by diagonalizing the transmon Hamiltonian given in Eq. (3.14). They are given for reference only.

N_{Tr}	N_{Res}	$E_C/2\pi$	$E_J/2\pi$	$\Omega/2\pi$	$G/2\pi$	$\tilde{\omega}/2\pi$	$\alpha/2\pi$
1	1	$0.222\,\text{GHz}$	$12.61\,\text{GHz}$	$5.821\,\text{GHz}$	$0.0349\,\text{GHz}$	$4.498\,\text{GHz}$	$-0.252\,\text{GHz}$
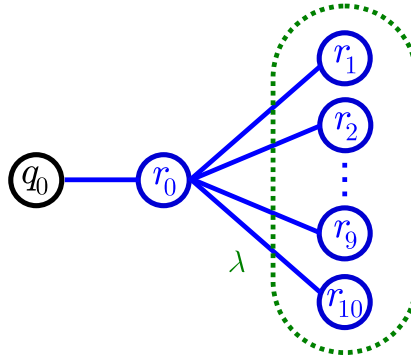
Figure 3.3: Setup of a system with one transmon and one resonator, coupled to a bath of 10 additional resonators. Such a setup is a generic model for a general linear superconducting environment (see Section 3.5). The system is simulated by solving the TDSE for the model Hamiltonian given in Eqs. (3.9a)–(3.9f) (see Tab. 3.3 for the model parameters).

Table 3.3: Model parameters for the KIT device specified in Tab. 3.2 coupled to a bath of 10 resonators (see Fig. 3.3). The relation between the specified resonator parameters and the parameters of the model Hamiltonian given in Eqs. (3.9a)–(3.9f) is $(\Omega_{r=0}, \Omega_{r=1}, \dots, \Omega_{r=10}) \leftrightarrow (\Omega, W_{l=1}, \dots, W_{l=10})$ and $\lambda_{rl} \leftrightarrow \lambda_l$ (for simplicity, we do not write the index $r = 0$ of the central resonator). The notation Gaussian(μ,σ) means that the frequencies are drawn from a Gaussian distribution with mean $\Omega = 2\pi \times 5.821\,\text{GHz}$ and standard deviation $\sigma = 2\pi \times 1\,\text{GHz}$. The notation Uniform(0,λ) means that the coupling strengths are drawn from a uniform distribution between 0 and $\lambda \in 2\pi \times \{5\,\text{MHz}, 10\,\text{MHz}, 20\,\text{MHz}\}$ (see Section 4.3).

N_{Tr}	N_{Res}	E_C, E_J, Ω, G	$W_l/2\pi$	$\lambda_l/2\pi$
1	11	(see Tab. 3.2)	Gaussian(Ω,σ)	Uniform(0,λ)

3.4.3 Two-transmon system

The smallest nontrivial system that allows for a simulation of a quantum computer needs at least two qubits [Gar2004] such that the study of two-qubit gates is possible. The parameters of such a system with two transmons coupled by one resonator are defined in Tab. 3.4. We use this system frequently throughout the following chapters to analyze two-qubit gates and small quantum algorithms. It has also been used for the results published in [Wil2017].

Table 3.4: Model parameters for a system with $N_{\mathrm{Tr}} = 2$ transmons and $N_{\mathrm{Res}} = 1$ resonator, simulated by solving the TDSE for the model Hamiltonian given by Eqs. (3.9a)–(3.9f). The parameters are inspired by device parameters of the quantum processors available on the IBM Q Experience between December 2016 and September 2017 [IBM2016] and have been obtained using an electromagnetic HFSS simulation [Sol2016] (see Section 3.5 for more information on this procedure). All energies are expressed in GHz ($\hbar = 1$). Unspecified parameters are set to zero. Estimates for the qubit frequencies $\tilde{\omega}_i$ and the anharmonicities α_i are obtained by diagonalizing the transmon Hamiltonian given by Eq. (3.14) (see also the actual frequencies given in Eqs. (4.31a) and (4.31b)). The resonator operates at frequency $\Omega_{r=0}/2\pi = 7\,\mathrm{GHz}$.

Transmon i	$E_{Ci}/2\pi$	$E_{Ji}/2\pi$	$G_i/2\pi$	$\tilde{\omega}_i/2\pi$	$\alpha_i/2\pi$
0	0.301	13.349	0.07	5.350	−0.350
1	0.301	12.292	0.07	5.120	−0.353

3.4.4 Small five-transmon system

The fourth system considered in this work is a setup inspired by the five-transmon device that was available on the IBM Q Experience in 2016 [IBM2016] and was further characterized in [Tak2017]. It consists of five transmons coupled by two resonators, each of which connects to three of the five transmons. The system is sketched in Fig. 3.4 and its model parameters are defined in Tab. 3.5 and Tab. 3.6. We use this system primarily for simulations of the quantum circuit experiments studied in Chapter 7. In particular, we study a circuit designed to reveal frequency shifts due to crosstalk for which we also perform the corresponding experiment on a quantum processor (see Section 7.1).

3.4.5 Large five-transmon system

The largest system used for simulations of actual quantum algorithms consists of five transmons and six resonators (larger systems are only considered for free evolutions and benchmarks). It is sketched in Fig. 3.5 as a subset of IBM's 16-qubit device, extended with an additional resonator. The model parameters of the system are given in Tab. 3.7 and Tab. 3.8. The qubit frequencies ω in Tab. 3.7 have been obtained by the procedure described above (see Eq. (3.40)); the corresponding squared errors defined for each qubit are shown in Fig. 3.6.
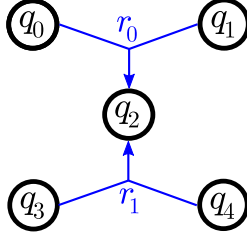
Figure 3.4: Setup of a system with five transmon qubits and two resonators inspired by the five-transmon device that was available on the IBM Q Experience in 2016 [IBM2016]. The system is described by the Hamiltonian given by Eqs. (3.9a)–(3.9f) (see Tab. 3.5 and Tab. 3.6 for the model parameters).

Table 3.5: Model parameters for a system of $N_{\mathrm{Tr}} = 5$ transmons and $N_{\mathrm{Res}} = 2$ resonators, simulated by solving the TDSE for the model Hamiltonian given by Eqs. (3.9a)–(3.9f). The system is sketched in Fig. 3.4. All values are given in GHz ($\hbar = 1$). The qubit frequencies ω have been obtained by minimizing the function given in Eq. (3.40) (see also Fig. 3.6). The drive frequencies f have been obtained by the single-qubit pulse-optimization procedure (see Sections 5.1 and 5.3). The parameters of the resonators are given separately in Tab. 3.6.

	q_0	q_1	q_2	q_3	q_4
$E_C/2\pi$	0.301	0.301	0.301	0.301	0.301
$E_J/2\pi$	13.3511	13.1446	12.2942	12.7882	12.0903
$\omega/2\pi$	5.34732	5.30259	5.11382	5.22509	5.07094
f	5.34697	5.30232	5.11345	5.22506	5.07065

Table 3.6: Model parameters of the resonators coupling the transmon qubits specified in Tab. 3.5. The parameters determine the resonator Hamiltonian defined in Eq. (3.9c). All values are given in GHz ($\hbar = 1$).

	r_0	r_1
$\Omega/2\pi$	7.01	6.63
$G/2\pi$	0.07	0.07
Coupled to	q_0, q_1, q_2	q_2, q_3, q_4
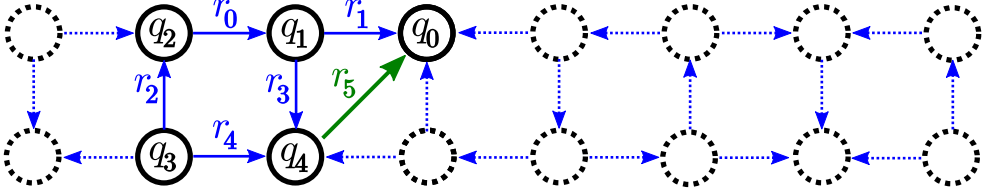
Figure 3.5: Setup of a system with five transmons and six resonators. The system is described by the Hamiltonian given by Eqs. (3.9a)–(3.9f) (see Tab. 3.7 and Tab. 3.8 for the model parameters). It represents a subset of the 16-qubit device `ibmqx5` [IBM2016] indicated by the dashed lines. An additional resonator r_5 has been added to the simulation model to enable the implementation of all circuits required for the fault-tolerance experiment considered in Section 7.3.

Table 3.7: Model parameters for a system of $N_{\text{Tr}} = 5$ transmons and $N_{\text{Res}} = 6$ resonators, simulated by solving the TDSE for the model Hamiltonian given by Eqs. (3.9a)–(3.9f). The system is sketched in Fig. 3.5. All values are given in GHz ($\hbar = 1$). The qubit frequencies ω have been obtained by minimizing the function given in Eq. (3.40) (see the plot in Fig. 3.6). The drive frequencies f have been obtained by the single-qubit pulse-optimization procedure (see Sections 5.1 and 5.3). The parameters of the resonators are given separately in Tab. 3.8.

	q_0	q_1	q_2	q_3	q_4
$E_C/2\pi$	0.301	0.301	0.301	0.301	0.301
$E_J/2\pi$	11.6671	12.1273	13.003	12.2456	11.1943
$\omega/2\pi$	4.97154	5.07063	5.26657	5.10145	4.86036
f	4.97164	5.07043	5.26634	5.10147	4.86055

Table 3.8: Model parameters of the resonators coupling the transmon qubits specified in Tab. 3.7. The parameters determine the resonator Hamiltonian defined in Eq. (3.9c). All values are given in GHz ($\hbar = 1$).

	r_0	r_1	r_2	r_3	r_4	r_5
$\Omega/2\pi$	6.45	6.25	6.65	6.65	6.45	6.85
$G/2\pi$	0.07	0.07	0.07	0.07	0.07	0.07
Coupled to	q_1, q_2	q_0, q_1	q_2, q_3	q_1, q_4	q_3, q_4	q_0, q_4
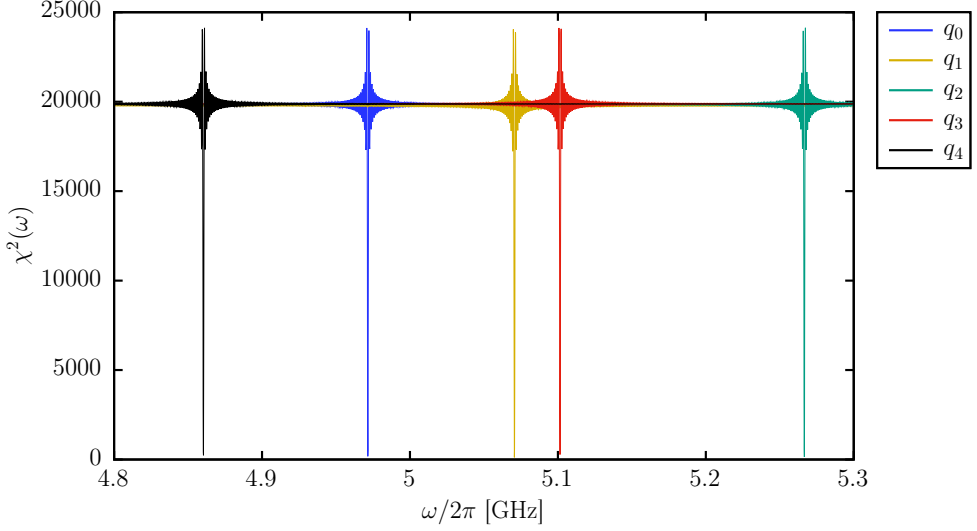
Figure 3.6: Squared error $\chi^2(\omega)$ defined in Eq. (3.40) for $N_{\text{data}} = 10000$ points from the free time evolution of the system sketched in Fig. 3.5 up to $T = 1000$ ns. Each line corresponds to a separate simulation in which the transmon q_i (whose frequency is to be determined) is initialized in the state $|+\rangle$ and the other transmons are initialized in the state $|0\rangle$. The time step used for the simulations is $\tau = 10^{-3}$ ns. The minima of the sharply peaked functions correspond to the qubit frequencies listed in Tab. 3.7. They are determined using the Golden Section Search method described above with $\omega_a/2\pi = 4.8$ GHz and $\omega_c/2\pi = 5.3$ GHz. All simulations were performed on the supercomputer JURECA [Jül2018].

We mainly use the large five-transmon system to test a fault-tolerant protocol from the field of quantum error correction (see Section 7.3). The same model has also been used for the results published in [Wil2018b].

3.5 Modeling electromagnetic environments

In this section, we describe a general method to find suitable model parameters for simulating electromagnetic environments with the simulation method defined in Section 3.2. The parameters can be obtained either by directly probing the experimental system or by performing an electromagnetic simulation of the device.

The following construction is inspired by the black box quantization method [Nig2012; Ans2019] and makes use of Foster's theorem [Fos1924]. It is also related to the method for extracting circuit Hamiltonians described in [Bou2012]. Furthermore, it can be extended to lossy electromagnetic environments to capture dissipative dynamics and predict relaxation rates (see [Sol2014; Sol2015]). Following a similar approach, the authors in [Sol2019] demonstrate how to extract the device parameters, including transmon-transmon and transmon-resonator couplings, for the 16-qubit device `ibmqx5` [IBM2018a] sketched in Fig. 3.5.

To make the construction concrete, we consider in detail the transmon-resonator-bath system defined in Section 3.4.2 for a general number L of bath resonators. In particular, we show that the system is a sufficiently general model for a Josephson junction coupled to a linear but otherwise arbitrary electromagnetic environment. This implies that there is no loss of generality by modeling a bath of uncoupled harmonic oscillators that only interact with a central resonator which is in turn coupled to a transmon. In principle, the method can be extended to multiple transmons by following the corresponding generalization in [Nig2012; Ans2019].

The transmon-resonator-bath Hamiltonian reads (cf. Eqs. (3.9a)–(3.9f))

$$H = H_{\mathrm{Tr}} + H_{\mathrm{Res}} + H_{\mathrm{Bath}}, \tag{3.42a}$$

$$H_{\mathrm{Tr}} = 4E_C\hat{n}^2 - E_J\cos\hat{\varphi}, \tag{3.42b}$$

$$H_{\mathrm{Res}} = \Omega\hat{a}^\dagger\hat{a} + G\hat{n}(\hat{a} + \hat{a}^\dagger), \tag{3.42c}$$

$$H_{\mathrm{Bath}} = \sum_{l=1}^{L} W_l\hat{b}_l^\dagger\hat{b}_l + \sum_{l=1}^{L} \lambda_l(\hat{a} + \hat{a}^\dagger)(\hat{b}_l + \hat{b}_l^\dagger). \tag{3.42d}$$

The idea is to treat all linear contributions canonically, and later work out the connection to the nonlinear parts (note that *(non)linear* in this context means *(an)harmonic*). The only nonlinear electromagnetic contribution to H comes from the Josephson junction described by the transmon Hamiltonian H_{Tr}. We split the Hamiltonian H into all linear and purely nonlinear parts by expanding the cosine in H_{Tr},

$$H_{\mathrm{Tr}} = 4E_C\hat{n}^2 + \frac{E_J}{2}\hat{\varphi}^2 + H_{\mathrm{Nonlin}}, \tag{3.43}$$

where the nonlinear part reads (up to a constant),

$$H_{\mathrm{Nonlin}} = E_J(1 - \cos\hat{\varphi}) - \frac{E_J}{2}\hat{\varphi}^2 = E_J\sum_{n=2}^{\infty}\frac{(-1)^{n+1}}{(2n)!}\hat{\varphi}^{2n} = -\frac{E_J}{24}\hat{\varphi}^4 + \cdots. \tag{3.44}$$
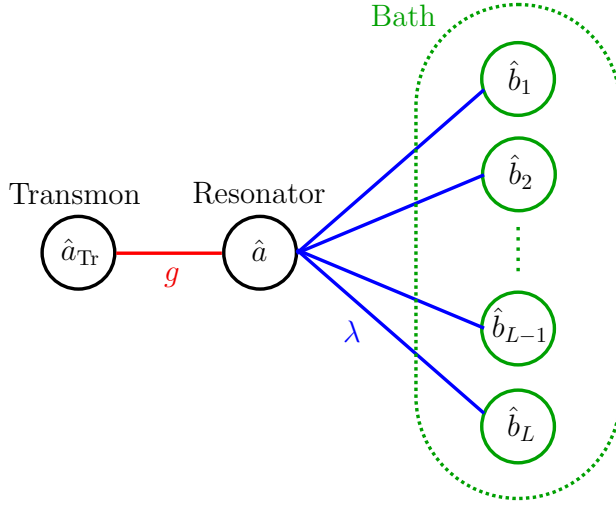
Figure 3.7: Setup of a system with one transmon and one central resonator, coupled to a bath of L noninteracting resonators. The coupling strength between the central resonator and the transmon (the bath) is g (λ_l for $l = 1, \ldots, L$). Although only the central resonator is connected to each other component, we show that such a setup is sufficiently general to model a generic case.

To diagonalize the linear part of H_{Tr}, we introduce operators $\hat{a}_{\text{Tr}}$ and $\hat{a}_{\text{Tr}}^\dagger$ such that

$$\hat{n} = -\frac{1}{\sqrt{2}} \left(\frac{E_J}{8E_C} \right)^{1/4} (\hat{a}_{\text{Tr}} + \hat{a}_{\text{Tr}}^\dagger), \tag{3.45a}$$

$$\hat{\varphi} = \frac{i}{\sqrt{2}} \left(\frac{8E_C}{E_J} \right)^{1/4} (\hat{a}_{\text{Tr}} - \hat{a}_{\text{Tr}}^\dagger). \tag{3.45b}$$

Note that no approximation is required for this step (see the discussion in Section 4.2.1). Using Eqs. (3.45a) and (3.45b), the linear part of the full Hamiltonian $H = H_{\text{Lin}} + H_{\text{Nonlin}}$ becomes

$$H_{\text{Lin}} = \Omega_{\text{Tr}} \hat{a}_{\text{Tr}}^\dagger \hat{a}_{\text{Tr}} + \Omega \hat{a}^\dagger \hat{a} + \sum_{l=1}^{L} W_l \hat{b}_l^\dagger \hat{b}_l$$

$$+ g(\hat{a}_{\text{Tr}} + \hat{a}_{\text{Tr}}^\dagger)(\hat{a} + \hat{a}^\dagger)$$

$$+ \sum_{l=1}^{L} \lambda_l (\hat{a} + \hat{a}^\dagger)(\hat{b}_l + \hat{b}_l^\dagger), \tag{3.46}$$

where $\Omega_{\text{Tr}} = \sqrt{8E_C E_J}$ denotes the frequency of the linear part of the transmon (which is different from the actual qubit frequency; see the discussion below Eq. (3.14)), and $g = -(E_J/32E_C)^{1/4}G$. The model is sketched in Fig. 3.7.

(a)

(b)
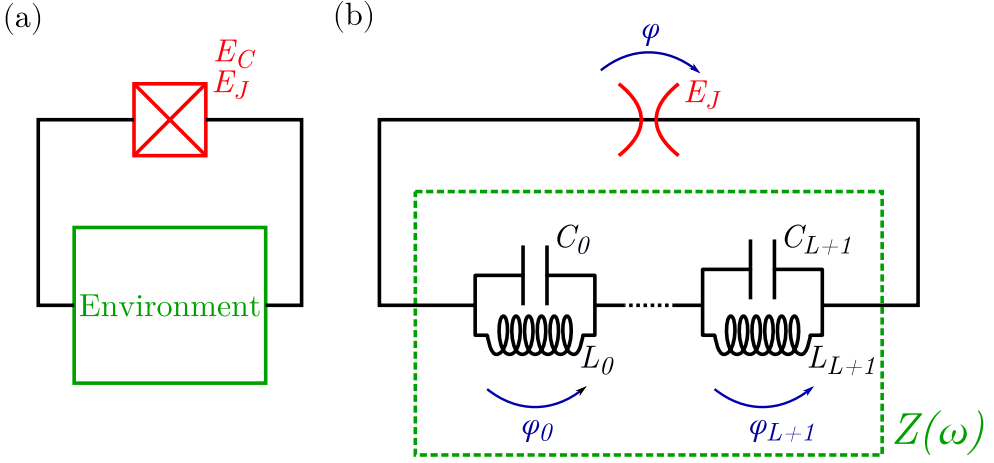
Figure 3.8: (a) Schematics of a transmon qubit (represented by a Josephson junction with capacitive energy E_C and inductive energy E_J) coupled to an electromagnetic environment. (b) The environment is characterized by its complex impedance $Z(\omega)$ and assumed to be linear, passive, and lossless. It is expressed in terms of capacitances C_j and inductances L_j using the Foster representation of the first kind (see Eq. (3.48)). Note that all linear parts of the Josephson junction are combined with the rest of the environment and lumped into $Z(\omega)$. The only nonlinear part, represented by the red spider symbol, corresponds to the Hamiltonian H_{Nonlin} given in Eq. (3.44).

In what follows, we first review the Foster representation of an electromagnetic environment, which gives rise to a procedure for extracting suitable model parameters. We then construct the mapping to the parameters of the Hamiltonian H_{Lin} in the rotating wave approximation (RWA). Finally, we give the full symplectic transformation to relate the model parameters to the Hamiltonian without the RWA.

3.5.1 The Foster representation of an electromagnetic environment

We consider a Josephson junction that is connected to a general, linear electromagnetic environment (see Fig. 3.8(a)). Such an environment is characterized by its electrical impedance $Z(\omega)$ or, equivalently, by its admittance $Y(\omega) = 1/Z(\omega)$. Note that for the treatment of several Josephson junctions, the impedance $Z(\omega)$ is replaced by an impedance matrix [Nig2012].

As we consider superconducting systems, we assume in the following analysis that the environment is lossless, i.e., no dissipative elements such as resistors are present. This means that the impedance $Z(\omega)$ is purely imaginary. Note, however, that the formalism can be extended to dissipative elements (see [Sol2014; Sol2015]).

Foster's theorem [Fos1924] states that for a lossless, passive network, $Z(\omega)$ is a complex

meromorphic function of the form

$$Z(\omega) = -iz_0 \frac{(\overline{\omega}_0^2 - \omega^2) \cdots (\overline{\omega}_L^2 - \omega^2)}{(\omega_0^2 - \omega^2) \cdots (\omega_{L+1}^2 - \omega^2)}, \tag{3.47}$$

where $z_0 \geq 0$ is a constant and $\overline{\omega}_j$ (ω_j) are the resonant (antiresonant) frequencies, which obey $0 \leq \omega_0 \leq \overline{\omega}_0 \leq \cdots \leq \overline{\omega}_L \leq \omega_{L+1}$. In particular, the function $Z(\omega)$ has poles at each ω_j (for simplicity, we do not consider poles at zero or infinity). By writing $\omega_j = 1/\sqrt{L_j C_j}$ for suitable inductances L_j and capacitances C_j and using a pole decomposition of $Z(\omega)$, Foster showed that $Z(\omega)$ can be synthesized by the series of LC-oscillators shown in Fig. 3.8(b). This representation is sometimes called the *Foster representation of the first kind* [Fel2009]. Explicitly, we obtain

$$Z(\omega) = \sum_{j=0}^{L+1} Z_j = \sum_{j=0}^{L+1} \frac{i\omega/C_j}{\omega_j^2 - \omega^2}, \tag{3.48}$$

where we used the relation $Z_j = 1/(1/Z_{L_j} + 1/Z_{C_j})$ for the impedance of a parallel combination of an inductance $Z_{L_j} = i\omega L_j$ and a capacitance $Z_{C_j} = 1/i\omega C_j$.

Relation to experiments

The impedance of the environment $Z(\omega)$ can, in principle, be extracted from current-voltage measurements. In practice, however, it is much more convenient to obtain $Z(\omega)$ from a finite-element high-frequency structure simulator (HFSS) to solve Maxwell's equations for a specification of the device's geometry [Nig2012].

After finding $Z(\omega)$, the next step is to determine the frequencies ω_j corresponding to the poles of the impedance (see Eq. (3.48)) or, equivalently, the zeros of the admittance $Y(\omega) = 1/Z(\omega)$. This can be done numerically. Additionally, we obtain the capacitances C_j and inductances L_j of the Foster representation shown in Fig. 3.8(b) using the formulas

$$C_j = \frac{|\operatorname{Im} Y'(\omega_j)|}{2}, \tag{3.49a}$$

$$L_j = \frac{2}{\omega_j^2 |\operatorname{Im} Y'(\omega_j)|}. \tag{3.49b}$$

These expressions can be proven by observing that the residue of $Z(\omega)$ at $\omega = \omega_j$ is $\operatorname{Res}(Z, \omega_j) = \lim_{\omega \to \omega_j}(\omega - \omega_j)Z(\omega) = i/2C_j$. Furthermore, since $Z(\omega) = f(\omega)/g(\omega)$ with f and g holomorphic and $g(\omega_j) = 0$ and $g'(\omega_j) \neq 0$ (cf. Eq. (3.47)), we have $\operatorname{Res}(Z, \omega_j) = \lim_{\omega \to \omega_j}(\omega - \omega_j)f(\omega)/g(\omega) = f(\omega_j)/g'(\omega_j)$. Also, we have $Y'(\omega_j) = g'(\omega_j)/f(\omega_j)$ and thus $Y'(\omega_j) = 1/\operatorname{Res}(Z, \omega_j) = -2iC_j$, which yields Eq. (3.49a). Equation (3.49b) then follows from $\omega_j^2 = 1/L_j C_j$.

Using Eqs. (3.49a) and (3.49b), we also find an expression for the *characteristic impedance of a lossless resonator*,

$$Z_j^{\text{eff}} = \sqrt{\frac{L_j}{C_j}} = \frac{2}{\omega_j |\operatorname{Im} Y'(\omega_j)|}. \tag{3.50}$$

Quantization

After extracting all fundamental modes ω_j of the environment depicted in Fig. 3.8(b), we can directly write down a Hamiltonian of $L + 2$ quantum harmonic oscillators for the electromagnetic environment of the system,

$$H_{\text{Env}} = \sum_{j=0}^{L+1} \omega_j \hat{c}_j^\dagger \hat{c}_j, \tag{3.51}$$

where $\hat{c}_l$ and $\hat{c}_l^\dagger$ are the corresponding ladder operators (see Section 3.1.1 for some general remarks on quantum and classical descriptions).

Note that the Hamiltonian in Eq. (3.51) describes the same system as the model Hamiltonian H_{Lin} given by Eq. (3.46). In essence, H_{Env} is the diagonal version of H_{Lin}. However, it is not obvious that the model parameters in H_{Lin} (i.e., the set of frequencies and coupling coefficients) are sufficiently generic to capture all instances of H_{Env}. Constructing the mapping to the parameters of H_{Lin} is the purpose of this section.

Up to this point, we have a Hamiltonian H_{Env} to describe the linear environment (dashed green box in Fig. 3.8(b)), and another Hamiltonian H_{Nonlin} (see Eq. (3.44)) for the nonlinear part represented by the spider symbol in Fig. 3.8(b). To make the connection between both, we need to relate the operator $\hat{\varphi} \propto (\hat{a}_{\text{Tr}} - \hat{a}_{\text{Tr}}^\dagger)$ given in Eq. (3.45b) to the ladder operators $\hat{c}_j$ and $\hat{c}_j^\dagger$ in Eq. (3.51). As indicated in Fig. 3.8(b), this relation can be made through the phase $\hat{\varphi}$ using the conservation of the total magnetic flux (cf. [Nig2012]):

$$\hat{\varphi} = \sum_j \hat{\varphi}_j = \sum_j \frac{i}{\sqrt{2}} \xi_j (\hat{c}_j - \hat{c}_j^\dagger), \tag{3.52}$$

where $\xi_j = \sqrt{Z_j^{\text{eff}}}/\phi_0$ is a dimensionless coefficient, Z_j^{eff} is given by Eq. (3.50), and ϕ_0 is the reduced flux quantum (note that in SI units, $\xi_j = \sqrt{\hbar Z_j^{\text{eff}}}/\phi_0$ and $\phi_0 = \hbar/2e$). The convention $\hat{\varphi}_j \propto i(\hat{c}_j - \hat{c}_j^\dagger)$ instead of $\hat{c}_j + \hat{c}_j^\dagger$ (as used in [Nig2012]) is only for compatibility with Eqs. (3.45a) and (3.45b) and does not affect the resulting dynamics.

3.5.2 Mapping to the model Hamiltonian

In this section, we construct the mapping from H_{Env} given by Eq. (3.51) to the model Hamiltonian H_{Lin} given by Eq. (3.46). For simplicity, we first consider the RWA version of the model Hamiltonian,

$$H_{\text{Lin}}^{\text{RWA}} = \Omega_{\text{Tr}} \hat{a}_{\text{Tr}}^\dagger \hat{a}_{\text{Tr}} + \Omega \hat{a}^\dagger \hat{a} + g(\hat{a}_{\text{Tr}}^\dagger \hat{a} + \hat{a}_{\text{Tr}} \hat{a}^\dagger) + \sum_{l=1}^{L} W_l \hat{b}_l^\dagger \hat{b}_l + \sum_{l=1}^{L} \lambda_l (\hat{a}^\dagger \hat{b}_l + \hat{a} \hat{b}_l^\dagger), \tag{3.53}$$

and generalize the mapping to the full model Hamiltonian H_{Lin} afterwards.

Derivation within the RWA

First, we write both $H_{\text{Lin}}^{\text{RWA}}$ and H_{Env} in matrix form,

$$H_{\text{Lin}}^{\text{RWA}} = \hat{\boldsymbol{a}}^{\dagger}\boldsymbol{\Omega}\hat{\boldsymbol{a}}, \qquad \boldsymbol{\Omega} = \begin{pmatrix} \Omega_{\text{Tr}} & g & & & \\ g & \Omega & \lambda_1 & \cdots & \lambda_L \\ & \lambda_1 & W_1 & & \\ & \vdots & & \ddots & \\ & \lambda_L & & & W_L \end{pmatrix}, \qquad \hat{\boldsymbol{a}} = \begin{pmatrix} \hat{a}_{\text{Tr}} \\ \hat{a} \\ \hat{b}_1 \\ \vdots \\ \hat{b}_L \end{pmatrix}, \qquad (3.54\text{a})$$

$$H_{\text{Env}} = \hat{\boldsymbol{c}}^{\dagger}\boldsymbol{\omega}\hat{\boldsymbol{c}}, \qquad \boldsymbol{\omega} = \begin{pmatrix} \omega_0 & & \\ & \ddots & \\ & & \omega_{L+1} \end{pmatrix}, \qquad \hat{\boldsymbol{c}} = \begin{pmatrix} \hat{c}_0 \\ \vdots \\ \hat{c}_{L+1} \end{pmatrix}. \qquad (3.54\text{b})$$

The goal is to construct a linear transformation $\hat{\boldsymbol{c}} = U\hat{\boldsymbol{a}}$, or in components,

$$\hat{\boldsymbol{c}}_j = \sum_{j'} U_{jj'}\hat{\boldsymbol{a}}_{j'}, \qquad (3.55)$$

such that $H_{\text{Lin}}^{\text{RWA}} = H_{\text{Env}}$. Here, U is a matrix of dimension $(L+2) \times (L+2)$. There are three conditions that U needs to meet:

$$\text{(I)} \qquad\qquad \boldsymbol{\Omega} = U^{\dagger}\boldsymbol{\omega}U, \qquad\qquad (3.56\text{a})$$

$$\text{(II)} \qquad\qquad \mathbb{1} = U^{\dagger}U, \qquad\qquad (3.56\text{b})$$

$$\text{(III)} \qquad\qquad \hat{a}_{\text{Tr}} \propto \sum_j \xi_j \hat{c}_j. \qquad\qquad (3.56\text{c})$$

Condition (I) comes from the requirement $H_{\text{Lin}}^{\text{RWA}} = H_{\text{Env}}$ and yields direct expressions for the model parameters $\boldsymbol{\Omega}$ in terms of the results from Section 3.5.1. Condition (II) follows from the fact that both operators $\hat{\boldsymbol{a}}_j$ and $\hat{\boldsymbol{c}}_j$ for $j = 0, \ldots, L+1$ need to satisfy the bosonic commutation relations,

$$\delta_{ij} = [\hat{\boldsymbol{c}}_i, \hat{\boldsymbol{c}}_j^{\dagger}] = \sum_{i'j'} U_{ii'}U_{jj'}^*[\hat{\boldsymbol{a}}_{i'}, \hat{\boldsymbol{a}}_{j'}^{\dagger}] = (UU^{\dagger})_{ij}, \qquad (3.57)$$

which means that U is unitary. Note that this is the point where the RWA simplifies the construction; in general, a transformation between bosonic operators is not necessarily unitary (see below). Finally, by condition (III), we can make sure that the transformation is compatible with the already existing relation between $\hat{a}_{\text{Tr}}$ and $\hat{c}_j$ following from Eqs. (3.45b) and (3.52), namely that $(\hat{a}_{\text{Tr}} - \hat{a}_{\text{Tr}}^{\dagger}) \propto \hat{\varphi} \propto \sum_j \xi_j(\hat{c}_j - \hat{c}_j^{\dagger})$.

In what follows, we construct the elements of U. Let $\boldsymbol{u}_j$ for $j = 0, \ldots, L+1$ denote the columns of U, i.e., $U = (\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{L+1})$. Although Eq. (3.56a) states that U diagonalizes $\boldsymbol{\Omega}$, the construction is not straightforward, because we do not know the elements of $\boldsymbol{\Omega}$ yet. Instead, we need to start from the eigenvalues ω_j and obtain the elements of $\boldsymbol{\Omega}$ from Eq. (3.56a),

$$\boldsymbol{\Omega}_{ij} = \boldsymbol{u}_i^{\dagger}\boldsymbol{\omega}\boldsymbol{u}_j. \qquad (3.58)$$

Furthermore, Eq. (3.56b) requires that the vectors $(\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{L+1})$ form an orthonormal basis, and the fact that $\boldsymbol{\Omega}$ in Eq. (3.54a) is symmetric allows us to choose $\boldsymbol{u}_j \in \mathbb{R}^{L+2}$ (which makes U an orthogonal matrix).

For the first vector, $\boldsymbol{u}_0$, we obtain using Eqs. (3.55) and (3.57), $\boldsymbol{u}_0^\dagger \hat{\boldsymbol{c}} = \hat{\boldsymbol{a}}_0 = \hat{a}_{\mathrm{Tr}}$. To satisfy condition (III) in Eq. (3.56c), we set

$$\boldsymbol{u}_0 = \frac{1}{\sqrt{\sum_j \xi_j^2}} \begin{pmatrix} \xi_0 \\ \vdots \\ \xi_{L+1} \end{pmatrix}, \tag{3.59}$$

where $\xi_j = \sqrt{Z_j^{\mathrm{eff}}}/\phi_0 \in \mathbb{R}$ is given below Eq. (3.52). It is not uncommon to find $\xi_0 \gg \xi_j$ for $j \geq 1$ since the slope of the admittance $|\mathrm{Im}\, Y'(\omega_0)|$ in the denominator of Eq. (3.50) is typically small for the lowest frequency ω_0 (see also Fig. 2 in the supplementary material of [Nig2012]). Hence, the first component of $\boldsymbol{u}_0$ is often the largest. Equation (3.58) then yields the first model parameter,

$$\Omega_{\mathrm{Tr}} = \boldsymbol{u}_0^\dagger \boldsymbol{\omega} \boldsymbol{u}_0 = \frac{\sum_j \xi_j^2 \omega_j}{\sum_j \xi_j^2}, \tag{3.60}$$

i.e., a weighted average of ω_j, where the weights are given by the effective characteristic impedances $\xi_j^2 \propto Z_j^{\mathrm{eff}}$. We can also determine E_C and E_J at this point: By inserting $\hat{a}_{\mathrm{Tr}} = \boldsymbol{u}_0^\dagger \hat{\boldsymbol{c}}$ and $\hat{a}_{\mathrm{Tr}}^\dagger = \hat{\boldsymbol{c}}^\dagger \boldsymbol{u}_0$ into Eq. (3.45b) and comparing with Eq. (3.52), we obtain $\sqrt{8 E_C/E_J} = \sum_j \xi_j^2$. The second equation for E_C and E_J comes from Eq. (3.60) by using $\Omega_{\mathrm{Tr}} = \sqrt{8 E_C E_J}$ (see Eq. (3.46)). Combining both, we find

$$E_C = \frac{1}{8} \sum_j \xi_j^2 \omega_j, \tag{3.61a}$$

$$E_J = \frac{\sum_j \xi_j^2 \omega_j}{(\sum_j \xi_j^2)^2}. \tag{3.61b}$$

The choice of the next vector, $\boldsymbol{u}_1$, requires a little trick: We want all elements $\boldsymbol{\Omega}_{j0} = 0$ for $j \geq 2$ (the first column of $\boldsymbol{\Omega}$ in Eq. (3.54a)). These elements are given by $\boldsymbol{\Omega}_{j0} = \boldsymbol{u}_j^\dagger \boldsymbol{\omega} \boldsymbol{u}_0$ (see Eq. (3.58)). If we choose $\boldsymbol{u}_1$ such that

$$\mathrm{span}\{\boldsymbol{u}_0, \boldsymbol{u}_1\} = \mathrm{span}\{\boldsymbol{u}_0, \boldsymbol{\omega} \boldsymbol{u}_0\}, \tag{3.62}$$

we know that all $\boldsymbol{u}_j$ for $j \geq 2$ will be orthogonal to $\boldsymbol{\omega} \boldsymbol{u}_0$. Therefore, $\boldsymbol{\Omega}_{j0} = \boldsymbol{u}_j^\dagger \boldsymbol{\omega} \boldsymbol{u}_0 = 0$. As $\boldsymbol{\Omega}_{0j} = \boldsymbol{\Omega}_{j0}$ (see Eq. (3.58)), this choice also produces the zeros in the first row of $\boldsymbol{\Omega}$ (see Eq. (3.54a)). A real unit vector $\boldsymbol{u}_1$ that accomplishes Eq. (3.62) and is orthogonal to $\boldsymbol{u}_0$ is given by

$$\boldsymbol{u}_1 = \frac{\boldsymbol{\omega} \boldsymbol{u}_0 - \Omega_{\mathrm{Tr}} \boldsymbol{u}_0}{|\boldsymbol{\omega} \boldsymbol{u}_0 - \Omega_{\mathrm{Tr}} \boldsymbol{u}_0|}, \tag{3.63}$$

where we used that the frequencies ω_j in $\boldsymbol{\omega}$ are, in general, different from Ω_{Tr}. Equation (3.58) then yields the next set of model parameters,

$$\Omega = \boldsymbol{u}_1^\dagger \boldsymbol{\omega} \boldsymbol{u}_1, \tag{3.64a}$$

$$g = \boldsymbol{u}_0^\dagger \boldsymbol{\omega} \boldsymbol{u}_1. \tag{3.64b}$$

The remaining columns $(\boldsymbol{u}_2, \ldots, \boldsymbol{u}_{L+1})$ need to be orthogonal to $\mathcal{S} = \text{span}\{\boldsymbol{u}_0, \boldsymbol{u}_1\}$. Furthermore, they need to satisfy $\boldsymbol{\Omega}_{ij} = \boldsymbol{u}_i^\dagger \boldsymbol{\omega} \boldsymbol{u}_j = 0$ for $i, j \geq 2$ with $i \neq j$ because the bottom-right $L \times L$ block of $\boldsymbol{\Omega}$ in Eq. (3.54a) is diagonal. One way to achieve this is to complete a set of linearly independent vectors to an orthonormal basis $(\boldsymbol{u}_0, \boldsymbol{u}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{L+1})$ of $\mathbb{R}^{L+2}$ using the Gram-Schmidt procedure, followed by a rediagonalization of $\boldsymbol{\omega}$ in the space spanned by the vectors $\boldsymbol{v}_j$.

Another way to achieve this is to project $\boldsymbol{\omega}$ directly onto the orthogonal complement $\mathcal{S}^\perp$ of $\mathcal{S}$ using the projector

$$P = \mathbb{1} - \boldsymbol{u}_0 \boldsymbol{u}_0^\dagger - \boldsymbol{u}_1 \boldsymbol{u}_1^\dagger, \tag{3.65}$$

and to diagonalize the projected matrix $P \boldsymbol{\omega} P$. Since P reduces the rank of $\boldsymbol{\omega}$ from $L+2$ to L, the matrix $P \boldsymbol{\omega} P$ has only L non-zero eigenvalues. These are exactly the frequencies W_l for $l = 1, \ldots, L$ of the L noninteracting oscillators of the bath (cf. Fig. 3.7). The corresponding eigenvectors make up the remaining columns $\boldsymbol{u}_{l+1}$ of U,

$$(P \boldsymbol{\omega} P) \boldsymbol{u}_{l+1} = W_l \boldsymbol{u}_{l+1}. \tag{3.66}$$

Thus we find for the remaining model parameters W_l and λ_l of $H_{\text{Lin}}^{\text{RWA}}$ for $l = 1, \ldots, L$,

$$W_l = \boldsymbol{u}_{l+1}^\dagger \boldsymbol{\omega} \boldsymbol{u}_{l+1}, \tag{3.67a}$$

$$\lambda_l = \boldsymbol{u}_{l+1}^\dagger \boldsymbol{\omega} \boldsymbol{u}_1. \tag{3.67b}$$

Derivation without the RWA

Without the RWA, the model Hamiltonian H_{Lin} given in Eq. (3.46) also contains quadratic terms of the form $\hat{a}_j^2$ and $(\hat{a}_j^\dagger)^2$ (the operators $\hat{a}$ are defined in Eq. (3.54a)). In this more general case, the Hamiltonians read

$$H_{\text{Lin}} = \begin{pmatrix} \hat{\boldsymbol{a}}^\dagger & \hat{\boldsymbol{a}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Omega}/2 & \boldsymbol{\Omega}/2 - \boldsymbol{D} \\ \boldsymbol{\Omega}/2 - \boldsymbol{D} & \boldsymbol{\Omega}/2 \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{a}} \\ \hat{\boldsymbol{a}}^\dagger \end{pmatrix}, \tag{3.68a}$$

$$H_{\text{Env}} = \begin{pmatrix} \hat{\boldsymbol{c}}^\dagger & \hat{\boldsymbol{c}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\omega}/2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\omega}/2 \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{c}} \\ \hat{\boldsymbol{c}}^\dagger \end{pmatrix}, \tag{3.68b}$$

where $\boldsymbol{D} = \text{diag}(\boldsymbol{\Omega}/2)$, and $\hat{\boldsymbol{c}}$ and $\boldsymbol{\omega}$ are defined in Eq. (3.54b). This case would require a more general transformation,

$$\begin{pmatrix} \hat{\boldsymbol{c}} \\ \hat{\boldsymbol{c}}^\dagger \end{pmatrix} = T \begin{pmatrix} \hat{\boldsymbol{a}} \\ \hat{\boldsymbol{a}}^\dagger \end{pmatrix}, \tag{3.69}$$

that also mixes the operators in $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{a}}^\dagger$. The most general transformation T that preserves the bosonic commutation relations is a Bogoliubov transformation, which is also known as a Bogoliubov-Valatin transformation as it was studied independently by Bogoliubov and Valatin to find solutions of the BCS theory of superconductivity [Bog1958; Val1958]. Bogoliubov transformations have to be *para-unitary*, i.e.,

$$T^\dagger \begin{pmatrix} \mathbb{1} & \mathbf{0} \\ \mathbf{0} & -\mathbb{1} \end{pmatrix} T = \begin{pmatrix} \mathbb{1} & \mathbf{0} \\ \mathbf{0} & -\mathbb{1} \end{pmatrix}. \tag{3.70}$$

Bogoliubov transformations have been studied in great detail by Colpa [Col1978]. It is not directly obvious that a suitable para-unitary transformation producing $H_{\mathrm{Lin}} = H_{\mathrm{Env}}$ exists, as the matrices on the diagonal and off-diagonal blocks in Eq. (3.68a) need to be equal except for the diagonal $\boldsymbol{D}$. Nonetheless, it is possible to construct a para-unitary transformation for the present problem by following Section 5 of [Col1978].

A simpler mapping that yields $H_{\mathrm{Lin}} = H_{\mathrm{Env}}$, however, can be obtained by working in an appropriate position-momentum representation, i.e., by writing the bosonic operators $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{c}}$ as linear combinations of Hermitian operators. Specifically, we define Hermitian operators $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{q}}$, and $\hat{\boldsymbol{p}}$ such that for $j = 0, \ldots, L+1$,

$$\hat{\boldsymbol{a}}_j = \sqrt{\alpha_j}\, \hat{\boldsymbol{x}}_j + \frac{i}{\sqrt{\alpha_j}}\, \hat{\boldsymbol{y}}_j, \tag{3.71a}$$

$$\hat{\boldsymbol{c}}_j = \sqrt{\omega_j}\, \hat{\boldsymbol{q}}_j + \frac{i}{\sqrt{\omega_j}}\, \hat{\boldsymbol{p}}_j, \tag{3.71b}$$

where $\alpha_0 = \Omega_{\mathrm{Tr}}$, $\alpha_1 = \Omega$, $\alpha_{l+1} = W_l$ for $l = 1, \ldots, L$, and ω_j are the frequencies of the linear environment (see Eq. (3.51)). The inverse relations are given by

$$\hat{\boldsymbol{x}}_j = \frac{\hat{\boldsymbol{a}}_j + \hat{\boldsymbol{a}}_j^\dagger}{2\sqrt{\alpha_j}}, \qquad\qquad \hat{\boldsymbol{y}}_j = \frac{\hat{\boldsymbol{a}}_j - \hat{\boldsymbol{a}}_j^\dagger}{2i/\sqrt{\alpha_j}}, \tag{3.72a}$$

$$\hat{\boldsymbol{q}}_j = \frac{\hat{\boldsymbol{c}}_j + \hat{\boldsymbol{c}}_j^\dagger}{2\sqrt{\omega_j}}, \qquad\qquad \hat{\boldsymbol{p}}_j = \frac{\hat{\boldsymbol{c}}_j - \hat{\boldsymbol{c}}_j^\dagger}{2i/\sqrt{\omega_j}}. \tag{3.72b}$$

In this representation, H_{Lin} in Eq. (3.46) and H_{Env} in Eq. (3.54b) amount to

$$H_{\mathrm{Lin}} = \begin{pmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} \end{pmatrix} \begin{pmatrix} \boldsymbol{X} & \mathbf{0} \\ \mathbf{0} & \mathbb{1} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{x}} \\ \hat{\boldsymbol{y}} \end{pmatrix}, \tag{3.73a}$$

$$H_{\mathrm{Env}} = \begin{pmatrix} \hat{\boldsymbol{q}} & \hat{\boldsymbol{p}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\omega}^2 & \mathbf{0} \\ \mathbf{0} & \mathbb{1} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{q}} \\ \hat{\boldsymbol{p}} \end{pmatrix}, \tag{3.73b}$$

where

$$\boldsymbol{X} = \begin{pmatrix} \Omega_{\mathrm{Tr}}^2 & 2g\sqrt{\Omega_{\mathrm{Tr}}\Omega} & & & \\ 2g\sqrt{\Omega_{\mathrm{Tr}}\Omega} & \Omega^2 & 2\lambda_1\sqrt{\Omega W_1} & \cdots & 2\lambda_L\sqrt{\Omega W_L} \\ & 2\lambda_1\sqrt{\Omega W_1} & W_1^2 & & \\ & \vdots & & \ddots & \\ & 2\lambda_L\sqrt{\Omega W_L} & & & W_L^2 \end{pmatrix}, \tag{3.74}$$

and $\boldsymbol{\omega}^2 = \mathrm{diag}(\omega_0^2, \ldots, \omega_{L+1}^2)$. Note that the coefficients in Eqs. (3.71a) and (3.71b) were chosen such that the bottom-right blocks in Eqs. (3.73a) and (3.73b) are identity matrices, and the commutation relations,

$$[\hat{\boldsymbol{x}}_j, \hat{\boldsymbol{y}}_{j'}] = [\hat{\boldsymbol{q}}_j, \hat{\boldsymbol{p}}_{j'}] = \frac{i}{2}\delta_{jj'}, \tag{3.75}$$

take the same form for each j and j'.

To find a mapping that yields $H_{\mathrm{Lin}} = H_{\mathrm{Env}}$, we now search for a transformation

$$\begin{pmatrix} \hat{\boldsymbol{q}} \\ \hat{\boldsymbol{p}} \end{pmatrix} = S \begin{pmatrix} \hat{\boldsymbol{x}} \\ \hat{\boldsymbol{y}} \end{pmatrix}. \tag{3.76}$$

The transformation S needs to be real (to preserve Hermiticity) and *symplectic*, i.e.

$$S^T \begin{pmatrix} \boldsymbol{0} & \mathbb{1} \\ -\mathbb{1} & \boldsymbol{0} \end{pmatrix} S = \begin{pmatrix} \boldsymbol{0} & \mathbb{1} \\ -\mathbb{1} & \boldsymbol{0} \end{pmatrix}, \tag{3.77}$$

to preserve the commutation relations in Eq. (3.75). The symplectic condition is analogous to the para-unitary condition defined in Eq. (3.70) for Bogoliubov transformations in the ladder-operator representation (see also [Mey2009a; Dop2009]). As an ansatz, we try

$$S = \begin{pmatrix} O & \boldsymbol{0} \\ \boldsymbol{0} & O \end{pmatrix} \tag{3.78}$$

with an orthogonal matrix O, i.e., $O^T O = \mathbb{1}$. By testing Eq. (3.77), S is easily seen to be symplectic. Inserting Eqs. (3.76) and (3.78) into Eq. (3.73b), we see that for $H_{\mathrm{Lin}} = H_{\mathrm{Env}}$, O has to satisfy

$$\boldsymbol{X} = O^T \boldsymbol{\omega}^2 O, \tag{3.79}$$

which is the analogous condition to Eq. (3.56a). Note that the ansatz Eq. (3.78) only works because the coefficients in Eqs. (3.71a) and (3.71b) are such that the bottom-right blocks in Eqs. (3.73a) and (3.73b) are identity matrices.

We now construct the elements of the orthogonal matrix O. Let $\boldsymbol{o}_j$ for $j = 0, \ldots, L+1$ denote the columns of O, i.e., $O = (\boldsymbol{o}_0, \ldots, \boldsymbol{o}_{L+1})$. Since the matrices $\boldsymbol{X}$ and $\boldsymbol{\omega}^2$ in Eq. (3.79) have the same structure as $\boldsymbol{\Omega}$ and $\boldsymbol{\omega}$ in Eqs. (3.54a) and (3.54b), we can use almost the same construction for O as we did for U (note that the resulting elements of U were real, so U was already orthogonal). The major difference lies in the first column $\boldsymbol{o}_0$ defined by $\hat{\boldsymbol{y}}_0 = \boldsymbol{o}_0^T \hat{\boldsymbol{p}}$ since it has to be compatible with the flux conservation Eq. (3.52). This equation reads in terms of the new coordinates

$$\hat{\varphi} = -\sum_j \sqrt{\frac{2}{\omega_j}} \xi_j \hat{\boldsymbol{p}}_j. \tag{3.80}$$

On the other hand, we have from Eq. (3.45b),

$$\hat{\varphi} = -\sqrt{\frac{2}{\Omega_{\mathrm{Tr}}}} \left(\frac{8E_C}{E_J}\right)^{1/4} \hat{\boldsymbol{y}}_0. \tag{3.81}$$

Therefore, we choose the first normalized column of O as (cf. Eq. (3.59))

$$\boldsymbol{o}_0 = \frac{1}{\sqrt{\sum_j \xi_j^2/\omega_j}} \begin{pmatrix} \xi_0/\sqrt{\omega_0} \\ \vdots \\ \xi_{L+1}/\sqrt{\omega_{L+1}} \end{pmatrix}. \tag{3.82}$$

The remaining construction of $O = (\boldsymbol{o}_0, \boldsymbol{o}_1, \ldots, \boldsymbol{o}_{L+1})$ is completely analogous to the construction of $U = (\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{L+1})$ above: We set (cf. Eq. (3.63))

$$\boldsymbol{o}_1 = \frac{\boldsymbol{\omega}^2 \boldsymbol{o}_0 - \Omega_{\mathrm{Tr}}^2 \boldsymbol{o}_0}{|\boldsymbol{\omega}^2 \boldsymbol{o}_0 - \Omega_{\mathrm{Tr}}^2 \boldsymbol{o}_0|}, \tag{3.83}$$

where $\Omega_{\mathrm{Tr}}^2 = \boldsymbol{o}_0^T \boldsymbol{\omega}^2 \boldsymbol{o}_0$. Then we define the projector $P = \mathbb{1} - \boldsymbol{o}_0 \boldsymbol{o}_0^T - \boldsymbol{o}_1 \boldsymbol{o}_1^T$, diagonalize $P \boldsymbol{\omega}^2 P$, and find all non-zero eigenvalues W_l^2 for $l = 1, \ldots, L$. The corresponding eigenvectors $\boldsymbol{o}_{l+1}$ constitute the remaining columns of O.

Given $O = (\boldsymbol{o}_0, \ldots, \boldsymbol{o}_{L+1})$, Eqs. (3.74) and (3.79) then yield for the model parameters of H_{Lin} in Eq. (3.46):

$$\Omega_{\mathrm{Tr}} = \sqrt{\boldsymbol{o}_0^T \boldsymbol{\omega}^2 \boldsymbol{o}_0} = \left(\frac{\sum_j \frac{\xi_j^2}{\omega_j} \omega_j^2}{\sum_j \frac{\xi_j^2}{\omega_j}} \right)^{1/2}, \tag{3.84a}$$

$$\Omega = \sqrt{\boldsymbol{o}_1^T \boldsymbol{\omega}^2 \boldsymbol{o}_1}, \tag{3.84b}$$

$$W_l = \sqrt{\boldsymbol{o}_{l+1}^T \boldsymbol{\omega}^2 \boldsymbol{o}_{l+1}}, \tag{3.84c}$$

$$g = \frac{\boldsymbol{o}_1^T \boldsymbol{\omega}^2 \boldsymbol{o}_0}{2\sqrt{\Omega_{\mathrm{Tr}} \Omega}}, \tag{3.84d}$$

$$\lambda_l = \frac{\boldsymbol{o}_{l+1}^T \boldsymbol{\omega}^2 \boldsymbol{o}_1}{2\sqrt{\Omega W_l}}. \tag{3.84e}$$

Finally, from Eqs. (3.80)–(3.82), $\hat{\boldsymbol{y}}_0 = \boldsymbol{o}_0^T \hat{\boldsymbol{p}}$, and $\Omega_{\mathrm{Tr}} = \sqrt{8E_C E_J}$, we obtain expressions for the remaining model parameters of the full Hamiltonian H in Eq. (3.42b),

$$E_C = \frac{1}{8} \sum_j \xi_j^2 \omega_j, \tag{3.85}$$

$$E_J = \left(\sum_j \frac{\xi_j^2}{\omega_j} \right)^{-1}, \tag{3.86}$$

$$G = -\left(\frac{32 E_C}{E_J} \right)^{1/4} g. \tag{3.87}$$

Note that the procedure yields all model parameters of H, including the transmon energies E_C and E_J, because all linear components have been included in the electromagnetic environment shown in Fig. 3.8.

Chapter 4

Free time evolution

A straightforward application of the simulation framework presented in the previous chapter is the free (undriven) evolution of various transmon-resonator systems. This is interesting for three reasons: First, it does not require the specification of pulses such that the systems can be easily scaled up and benchmarked with respect to accuracy and performance. Second, small free systems are simple enough such that one can compare to perturbative results often used in analytical works. And third, an undriven evolution may be easier to study in laboratory experiments.

A free time evolution mathematically means that in the model Hamiltonian given by Eqs. (3.9a)–(3.9f), the time-dependent fields $n_{gi}(t)$ and $\epsilon_r(t)$ are zero. The model Hamiltonian considered in this chapter can therefore be reduced to

$$H^{\text{free}} = H^{\text{free}}_{\text{Tr}} + H^{\text{free}}_{\text{Res}}, \tag{4.1a}$$

$$H^{\text{free}}_{\text{Tr}} = \sum_{i=0}^{N_{\text{Tr}}-1} \left[4E_{Ci}\hat{n}_i^2 - E_{Ji}\cos\hat{\varphi}_i \right], \tag{4.1b}$$

$$H^{\text{free}}_{\text{Res}} = \sum_{r=0}^{N_{\text{Res}}-1} \Omega_r \hat{a}_r^\dagger \hat{a}_r + \sum_{r=0}^{N_{\text{Res}}-1} \sum_{i=0}^{N_{\text{Tr}}-1} G_{ri}\hat{n}_i(\hat{a}_r + \hat{a}_r^\dagger)$$
$$+ \sum_{0 \le r < l < N_{\text{Res}}} \lambda_{rl}(\hat{a}_r + \hat{a}_r^\dagger)(\hat{a}_l + \hat{a}_l^\dagger). \tag{4.1c}$$

We use this Hamiltonian to test the implementation of the main part of the simulation algorithm (see Section 3.3). This includes accuracy and performance benchmarks discussed in Section 4.1, where we use small systems to compare the simulation results to exact diagonalization, and larger systems to assess weak and strong scaling performance on parallel supercomputers. In Section 4.2, we simulate a single transmon-resonator system to analyze the accuracy of known analytical results based on perturbation theory. Section 4.3 is centered around a larger system of resonators that play the role of a bath as a model for an open quantum system. The results are compared to those obtained from a Lindblad master equation. In particular, we consider the time evolution of a transmon as a function of the number of photons in its readout resonator. A corresponding experiment has been conducted at the Karlsruhe Institute of Technology (KIT), and we use the same device parameters to define the particular instance of the model Hamiltonian H^{free} (see

Tab. 3.2 and Tab. 3.3). Finally, in Section 4.4, we study the free evolution of two coupled transmon qubits for different initial states, by which we characterize state-dependent frequency shifts caused by the resonator-mediated exchange interaction.

4.1 Accuracy and performance benchmarks

In this section, we study the accuracy and the performance of the simulation algorithm described in Section 3.3.1. For the former, the number of transmons N_{Tr} and resonators N_{Res} in the system is kept small to compare the results to exact diagonalization, which is infeasible if the size of the Hilbert space is too large. For the latter, we study increasing system sizes with a focus on the scaling of the algorithm. All simulations were performed on the supercomputer JURECA [Jül2018].

4.1.1 Accuracy

A mandatory step for any implementation of a numerical algorithm is to verify its correctness. In this context, correctness concerns two points: First, the algorithm should be stable with respect to different initial conditions (i.e., different initial states for the TDSE in Eq. (3.8)). This is guaranteed as the Suzuki-Trotter product-formula algorithm is unconditionally stable, since all updates of the state $|\Psi(t)\rangle$ in the decomposition are unitary by construction (see Eqs. (3.19) and (3.30)). Second, the result of the algorithm has to agree, up to a certain controllable precision, with the exact solution of the mathematical problem. Specifically, this means that the resulting coefficients of the solution $|\Psi(t)\rangle$ have to agree with the mathematical solution of the TDSE in Eq. (3.8). Verifying this is the purpose of the present section.

The systems used to study the accuracy of the simulation are small such that they can still be diagonalized exactly: a single transmon-resonator device produced at KIT (see Section 3.4.1), and a device with two transmons coupled by a single resonator inspired by one of the early processors available on the IBM Q Experience (see Section 3.4.3).

For the Suzuki-Trotter product-formula algorithm, the numerical error can be controlled by the time step τ used to integrate the TDSE. Specifically, we test the following local and global error bounds for the second-order product-formula algorithm:

$$\|\Psi - \overline{\Psi}\| \leq \text{const} \cdot N_\tau \cdot \tau^3, \tag{4.2a}$$

$$1 - |\langle\Psi|\overline{\Psi}\rangle|^2 \leq \text{const} \cdot N_\tau^2 \cdot \tau^6, \tag{4.2b}$$

where Ψ represents the coefficients of the state vector after N_τ time steps of size τ, and $\overline{\Psi}$ represents the exact solution at time $N_\tau\tau$. Equation (4.2a) was proven in [DeR1987; Huy1990], along with an expression for the constant in terms of commutators of H_0 and W given in Eqs. (3.26a)–(3.26g). To prove Eq. (4.2b), note that $1 - |\langle\Psi|\overline{\Psi}\rangle|^2 \leq 1 - (\text{Re}\,\langle\Psi|\overline{\Psi}\rangle)^2 \leq 2(1 - \text{Re}\,\langle\Psi|\overline{\Psi}\rangle) = \|\Psi - \overline{\Psi}\|^2$.

The local error is obtained by performing a single time step ($N_\tau = 1$) for different values of τ. The global error, on the other hand, is obtained by keeping τ fixed and performing $N_\tau\tau$ as a function of N_τ. Note that a nice property of the algorithm is that the bound on the global error in Eq. (4.2a) only grows linearly with the number of time steps N_τ.

Figure 4.1: Local error between the result of the simulation $|\Psi\rangle$ and the exact result $|\overline{\Psi}\rangle$ after a single time step τ. The purpose of this local error plot is to verify the scaling laws given in Eqs. (4.2a) and (4.2b) as a function of the time step τ for (a) the KIT system defined in Section 3.4.1, and (b) the IBM system defined in Section 3.4.3. Additionally, the plots show errors for the observable X given by Eq. (4.4) and the respective bounds given by Eqs. (4.3a) and (4.3b).

Both local and global errors are shown in Figs. 4.1 and 4.2, respectively. The initial state of the simulations is set to $|+\rangle$ ($|++\rangle$) for the KIT (IBM) system. This means that we study a worst-case scenario because of the fast, observable rotations performed by the Bloch vectors (see Eq. (3.39)). We obtain the mathematical solution $\overline{\Psi}$ by exact diagonalization of the full system, which can be done up to machine precision [Dem2008].

For the local error shown in Fig. 4.1, we used four states for each transmon and each resonator to obtain $\overline{\Psi}$ in order to verify the power-law scaling in Eqs. (4.2a) and (4.2b) as a function of τ. Figure 4.1 shows that the scaling laws are satisfied, as soon as the errors leave the range of machine precision around 10^{-15} where the numbers are practically zero.

For the global error shown in Fig. 4.2, we used 10 states to obtain $\overline{\Psi}$ such that $\overline{\Psi}$ is equal to the mathematical result (up to machine precision). For this reason, Fig. 4.2 also includes the error made by truncating the Hilbert space given in Eq. (3.17). We see that even for the largest time step used in this work ($\tau = 10^{-3}$ ns), the errors are sufficiently small.

Accuracy of observables

Often, one is not interested in the full state vector itself (especially in cases where the system is so large that it is not practical to store the coefficients on disk anymore).

(a) KIT

(b) IBM

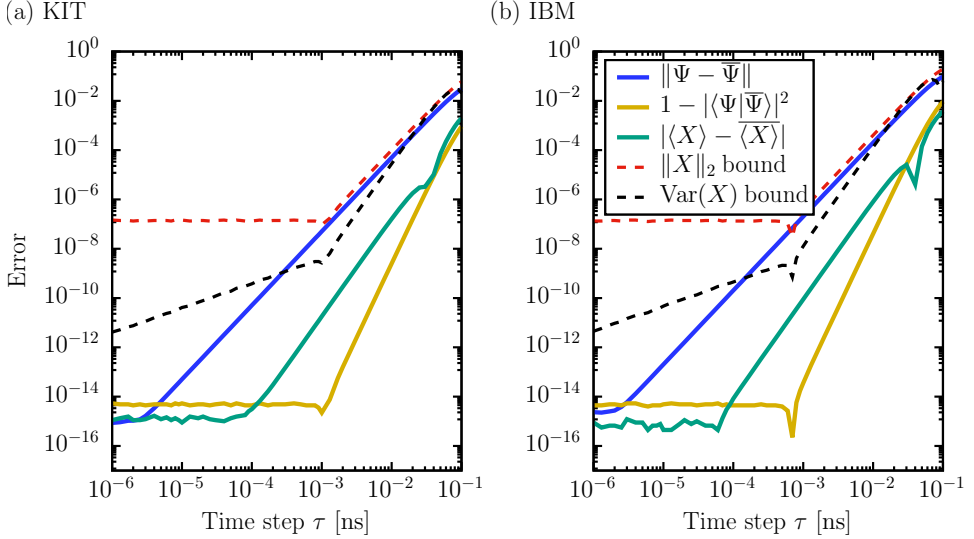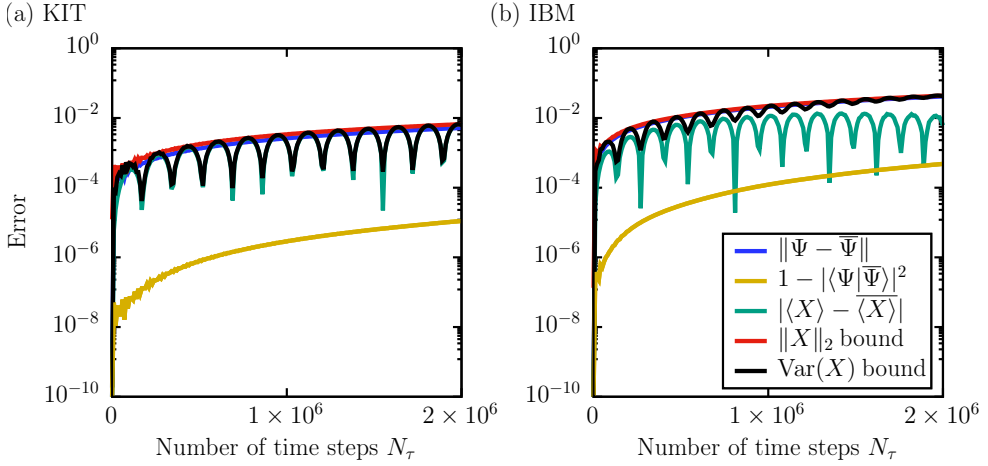Figure 4.2: Global error between the result of the simulation $|\Psi\rangle$ and the exact result $|\overline{\Psi}\rangle$. (a) the KIT system defined in Section 3.4.1, (b) the IBM system defined in Section 3.4.3. Additionally, the plots show errors for the observable X given in Eq. (4.4) and the respective bounds given in Eqs. (4.3a) and (4.3b). The total time evolution corresponds to $2\,\mu s$ using a time step of $\tau = 10^{-3}\,ns$.

Rather, the goal of the simulation is to produce the expectation value $\langle X\rangle = \langle\Psi|X|\Psi\rangle$ of an observable X. In practice, an expectation value may often be more accurate than suggested by the error bounds given in Eqs. (4.2a) and (4.2b). A reason for this is that the error $\|\Psi - \overline{\Psi}\|$ is often dominated by a difference in the global phase, which does not lead to errors in expectation values.

Therefore, we additionally consider the actual error on the observable $|\langle X\rangle - \overline{\langle X\rangle}|$, where $\overline{\langle X\rangle} = \langle\overline{\Psi}|X|\overline{\Psi}\rangle$. This quantity gives a measure of the discriminability of different states. For any Hermitian operator X, one can prove the following bounds (see Appendix E and [Wil2020d] for more information):

$$|\langle X\rangle - \overline{\langle X\rangle}| \leq 2\sqrt{\Delta}\,\|X\|_2, \tag{4.3a}$$

$$|\langle X\rangle - \overline{\langle X\rangle}| \leq 2\sqrt{\Delta}\,\sqrt{\mathrm{Var}_\Psi(X)}\,|\langle\Psi|\overline{\Psi}\rangle| + 2\Delta\,\|X\|_2, \tag{4.3b}$$

where $\Delta = 1 - |\langle\Psi|\overline{\Psi}\rangle|^2$ is the *distinguishability* between $|\Psi\rangle$ and $|\overline{\Psi}\rangle$, $\|X\|_2$ denotes the spectral norm (largest singular value) of X, and $\mathrm{Var}_\Psi(X) = \langle X^2\rangle - \langle X\rangle^2$ is the variance of X with respect to the state $|\Psi\rangle$. Note that both bounds are general bounds for how well different quantum states can be distinguished. However, the second bound may be preferable because it depends, to leading order in $\sqrt{\Delta}$, only on the variance of the observable X in the respective state, instead of the state-independent operator norm $\|X\|_2$. Both bounds can be related to the time step τ by using $\Delta \propto \tau^6$ (see Eq. (4.2b)), without knowledge of the exact result $|\overline{\Psi}\rangle$. Furthermore, they can be estimated by evaluating the expressions in [DeR1987; Huy1990] for the constant prefactor in Eq. (4.2b).

As an application, we study the observables

$$X = \begin{cases} \sigma_0^x & \text{(KIT system)} \\ \sigma_0^x \sigma_1^x & \text{(IBM system)} \end{cases}.$$ (4.4)

In both cases, we have $\|X\|_2 = 1$ and $\text{Var}_\Psi(X) = 1 - \langle X \rangle^2$. Since the initial state of the simulation is set to $|+\rangle$ ($|++\rangle$) for the KIT (IBM) system, the observables X in Eq. (4.4) measure the fast rotations performed by the Bloch vectors (see Eq. (3.39)). Therefore, we expect errors to surface quickly.

Indeed, Fig. 4.1 shows that the local error of $|\langle X \rangle - \langle \overline{X} \rangle|$ increases earlier than $1 - |\langle \Psi | \overline{\Psi} \rangle|^2$, but is still a factor of 10–100 smaller than the bounds given in Eqs. (4.3a) and (4.3b). Interestingly, though, the $\text{Var}_\Psi(X)$ bound is tight for the global error shown in Fig. 4.2(a), even after two million time steps of size $\tau = 10^{-3}$ ns (corresponding to 2 µs). For the slightly larger systems studied in Fig. 4.2(b), we see that the global error of $|\langle X \rangle - \langle \overline{X} \rangle|$ saturates after 1 µs while the respective bounds increase.

Finally, we evaluate the error defined by Eq. (3.37) between results obtained with two different time steps $\tau_1 = 10^{-4}$ ns and $\tau_2 = 10^{-3}$ ns. We consider the overlap in Eq. (3.36) between $|\Psi^{\tau_1}(t)\rangle$ and $|\Psi^{\tau_2}(t)\rangle$ after every 10 ns for a total time of 2 µs. This means that the set $\mathbb{T}$ in Eq. (3.37) contains 200 items. For the KIT system, we obtain an error of 3.34×10^{-8}. The IBM system yields an error of 9.50×10^{-7}. Hence, we conclude that a time step of 10^{-3} ns is sufficient for both systems. In contrast to the rigorous error bounds studied above, this procedure to determine the time step does not require knowledge of the exact result $|\overline{\Psi}\rangle$, so it can also be used for much larger systems with time-dependent Hamiltonians.

4.1.2 Performance

Having verified the accuracy of the simulation method, the next step is to assess the performance of the transmon simulator. This requires a simulation of larger systems such that most of the time spent goes into updating the state vector $|\Psi\rangle$ for each time step of size τ. As discussed in Section 3.3 (see Eqs. (3.22) and (3.30)), this update is given by

$$|\Psi\rangle \leftarrow \underbrace{e^{-i\tau H_0/2} V e^{-i\tau \Lambda} V^\dagger e^{-i\tau H_0/2}}_{\tilde{\mathcal{U}}} |\Psi\rangle .$$ (4.5)

The transformation associated with the operator V is the most complicated part (see Eq. (3.31); the other operators are diagonal). Therefore, we expect the bottleneck to be in this transformation. We compare three different implementations of this transformation (see Appendix D).

Performance benchmarks are usually obtained by measuring the run time as a function of the system size $N_{\text{Tr}} + N_{\text{Res}}$ and the number of parallel threads N_{Threads}. We consider two different system configurations. For configuration (1), we vary N_{Tr} while keeping $N_{\text{Res}} = 1$ fixed, and for configuration (2), we vary both $N_{\text{Tr}} = N_{\text{Res}}$. The model parameters are given in Tab. 4.1. For each of the following benchmarks, we perform 1000 time steps of size $\tau = 10^{-3}$ ns. All simulations were performed on the JURECA cluster [Jül2018] at the Jülich Supercomputing Centre.

Table 4.1: Model parameters for performance benchmarks for various system sizes $N_{\mathrm{Tr}} + N_{\mathrm{Res}}$ between 1 and 16 (cf. Eqs. (4.1a)–(4.1c)). For configuration (1), N_{Tr} is variable and $N_{\mathrm{Res}} = 1$. For configuration (2), $N_{\mathrm{Tr}} = N_{\mathrm{Res}}$ are both variable. The symbols u_i^I, u_r^I, and u_{ri}^I denote uniform random numbers drawn from the interval I, where $i = 0, \ldots, N_{\mathrm{Tr}} - 1$ and $r = 0, \ldots, N_{\mathrm{Res}} - 1$. The parameters are inspired by the IBM device specified in Tab. 3.4. All energies are expressed in GHz ($\hbar = 1$). Unspecified parameters are set to zero.

Configuration	$E_{Ci}/2\pi$	$E_{Ji}/2\pi$	$\Omega_r/2\pi$	$G_{ri}/2\pi$
(1) N_{Tr} variable, $N_{\mathrm{Res}} = 1$	0.301	$10 + u_i^{[0,5)}$	$5.5 + u_r^{[0,3)}$	$0.055 + u_{ri}^{[0,0.03)}$
(2) $N_{\mathrm{Tr}} = N_{\mathrm{Res}}$ variable	0.301	$10 + u_i^{[0,5)}$	$5.5 + u_r^{[0,2)}$	$0.055 + u_{ri}^{[0,0.02)}$

For a given system size $N_{\mathrm{Tr}} + N_{\mathrm{Res}}$, the memory needed to store the state vector $|\Psi\rangle$ is

$$\dim(\mathcal{H}) \times 16 \, \mathrm{bytes} = 4^{N_{\mathrm{Tr}} + N_{\mathrm{Res}} + 2} \, \mathrm{bytes}, \tag{4.6}$$

where the size of the Hilbert space is $\dim(\mathcal{H}) = 4^{N_{\mathrm{Tr}} + N_{\mathrm{Res}}}$ (see Eq. (3.17)) and the factor of 16 is due to the use of complex double-precision floating point numbers. Therefore, we expect the computational work to grow exponentially with the system size $N_{\mathrm{Tr}} + N_{\mathrm{Res}}$, and the goal is to work against this by increasing the number of threads.

As a simple check, we measure the run time T as a function of the system size, for each of the three implementations (see Appendix D) in both single-threaded and multi-threaded runs. The result for configuration (1) is shown in Fig. 4.3. The run times for configuration (2) are almost the same (data not shown). We see that in the single-threaded case, the run time T grows as expected already for very small systems. In the multi-threaded case, the scaling only starts at system sizes around 7. This effect is due to the overhead of managing 48 threads, compared to the relatively small computational work required for $N_{\mathrm{Tr}} + N_{\mathrm{Res}} < 7$. At this stage, one cannot see an essential difference between the three implementations yet.

Weak and strong scaling

The performance of a parallel implementation is usually evaluated using two scaling metrics, namely the *weak scaling* performance and the *strong scaling* performance. The weak scaling performance is obtained by simultaneously increasing the system size $N_{\mathrm{Tr}} + N_{\mathrm{Res}}$ and the number of threads N_{Threads}, while keeping the computational work per thread constant. For the present algorithm, increasing the system size $N_{\mathrm{Tr}} + N_{\mathrm{Res}}$ by 1 increases the size of the state vector $|\Psi\rangle$ by a factor of 4 (see Eq. (4.6)). However, the total computational work is increased by *more than* a factor of 4. This effect is easily seen in Listings D.1–D.3 in Appendix D: the number of operations in the inner loop grows by a factor of 4 and, additionally, the number of iterations in one of the outer loops grows by 1 (the reason is that increasing $N_{\mathrm{Tr}} + N_{\mathrm{Res}}$ adds additional four-component updates to the transformation V given in Eq. (3.31)). Therefore, the total computational work is proportional to $(N_{\mathrm{Tr}} + N_{\mathrm{Res}}) 4^{N_{\mathrm{Tr}} + N_{\mathrm{Res}}}$. Since the number of threads N_{Threads} can only be increased
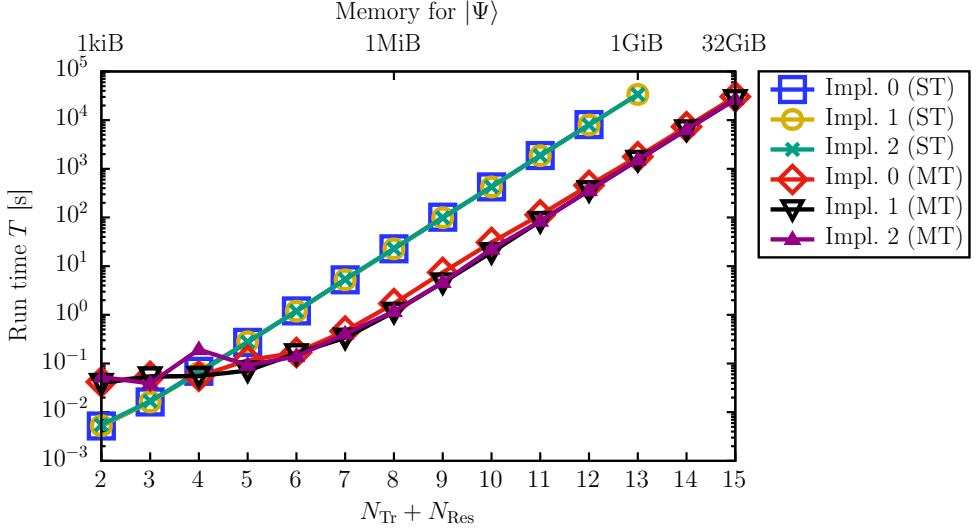
Figure 4.3: Run time T as a function of the system size $N_{\mathrm{Tr}} + N_{\mathrm{Res}}$, for all three implementations (see Appendix D) in the single-threaded case (ST) using one core and the multi-threaded case (MT) using 48 threads on 24 cores. The memory required to store the state vector $|\Psi\rangle$ is linked to the system size via Eq. (4.6). Lines are drawn to guide the eye.
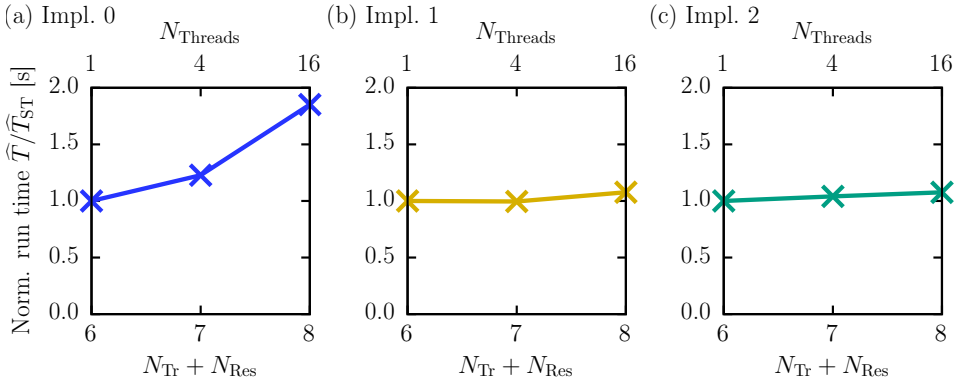
Figure 4.4: Weak scaling performance of the normalized run time $\widehat{T}/\widehat{T}_{\mathrm{ST}}$ using (a) implementation 0 where $\widehat{T}_{\mathrm{ST}} = 1.202\,\mathrm{s}/6$, (b) implementation 1 where $\widehat{T}_{\mathrm{ST}} = 1.188\,\mathrm{s}/6$, (c) implementation 2 where $\widehat{T}_{\mathrm{ST}} = 1.188\,\mathrm{s}/6$. Lines are drawn to guide the eye.

by a factor of 4, we consider the rescaled run time $\widehat{T} = T/(N_{\mathrm{Tr}} + N_{\mathrm{Res}})$. We normalize this quantity by the time $\widehat{T}_{\mathrm{ST}}$ obtained for the single-threaded case at $N_{\mathrm{Tr}} + N_{\mathrm{Res}} = 6$.

The weak scaling performance is shown in Fig. 4.4(a)–(c) for each of the three implementations, using configuration (1) since the system size in configuration (2) can only be increased in steps of 2 (cf. Tab. 4.1). We find that implementations 1 and 2 show almost ideal scaling behavior while implementation 0 does not scale as favorably. This means that relying on branch predictors, as done for implementation 0, can hamper parallel scalability. In particular, it is the additional non-loop branches present in implementation 0 (cf. Listing D.1) that destroy the scalability. In Fig. 4.4, the difference between the three implementations is much more evident than in Fig. 4.3, where the fact that implementation 0 is slower can only be seen in the slightly higher red diamonds and the missing blue square at $N_{\mathrm{Tr}} + N_{\mathrm{Res}} = 13$ (in this case, the computation did not finish within the allotted time of 24 hours).

The strong scaling performance is obtained by increasing the number of threads N_{Threads} while keeping the computational work (i.e., the system size) fixed. We choose a moderate system of size 10 from configuration (2) (see Tab. 4.1) with $N_{\mathrm{Tr}} = 5$ transmons and $N_{\mathrm{Res}} = 5$ resonators. The number of threads N_{Threads} is varied from 1 to 48 in steps of 1.

The strong scaling performance is shown in Fig. 4.5 for implementation 2 (the results for implementation 1 are almost identical). A node on JURECA has 24 physical cores which can each process two threads using hyper-threading (also called simultaneous multithreading) [Jül2018]. Therefore, we find two separate domains, namely the domain $1 \leq N_{\mathrm{Threads}} \leq 24$ where hyper-threading is off and the domain $25 \leq N_{\mathrm{Threads}} \leq 48$ where hyper-threading is on. In each domain, a fit of the function $f(x) = ax^b$ to the observed run times yields an almost ideal strong scaling exponent of $b \approx -1$. Although the run time decreases at the hyper-threading threshold from $N_{\mathrm{Threads}} = 24$ to $N_{\mathrm{Threads}} = 25$, we find that the simulation performs best when the full capacity of the node with $N_{\mathrm{Threads}} = 48$ is used.

To analyze which part of the update rule given in Eq. (4.5) takes the longest fraction of the run time, we perform a breakdown of the computational work. For this purpose, we measure the run time for each of the transformations included in $\widehat{\mathcal{U}}$ as a function of the system size $N_{\mathrm{Tr}} + N_{\mathrm{Res}}$. For each transformation, we compute the median of 10 run times, measured every 100 time steps. The result is shown in Fig. 4.6 for both the single-threaded and the multi-threaded case, using configuration (1) and implementation 2. As expected, the bottleneck for larger systems is the transformation V. The irregularities for smaller system sizes $N_{\mathrm{Tr}} + N_{\mathrm{Res}} < 7$ in the multi-threaded case shown in Fig. 4.6(b) again reflect the above-mentioned overhead required for managing too many threads for smaller systems.

We conclude that, among the three implementations for V (cf. Appendix D), implementations 1 and 2 are preferable since they are faster and show almost ideal weak and strong scaling performance. We use implementation 2 for the simulation work presented in the remainder of this thesis (see also the discussion at the end of Section 3.3.1).
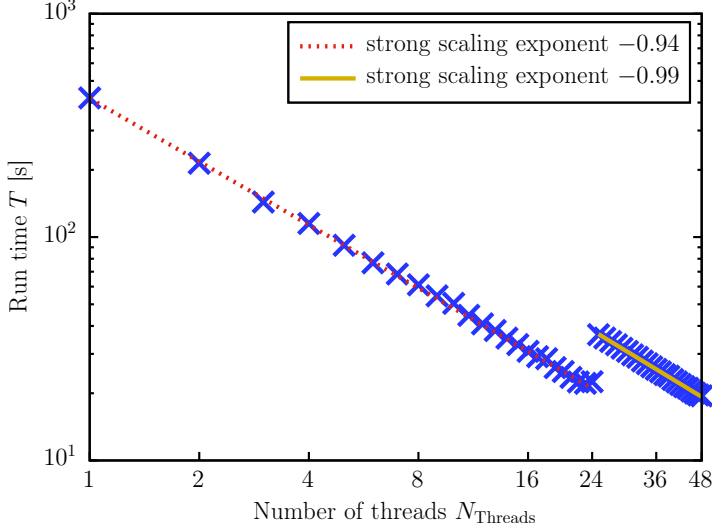
Figure 4.5: Strong scaling performance for a system of $N_{\text{Tr}} = 5$ transmons and $N_{\text{Res}} = 5$ resonators. Blue crosses show the run time T as a function of the number of threads N_{Threads}. Lines indicate a fit of the function $f(x) = ax^b$ to the data, where b is the strong scaling exponent (ideally -1) and a is a constant. The line is dotted in the domain $1 \leq N_{\text{Threads}} \leq 24$ without hyper-threading, and solid in the domain $25 \leq N_{\text{Threads}} \leq 48$ with hyper-threading.
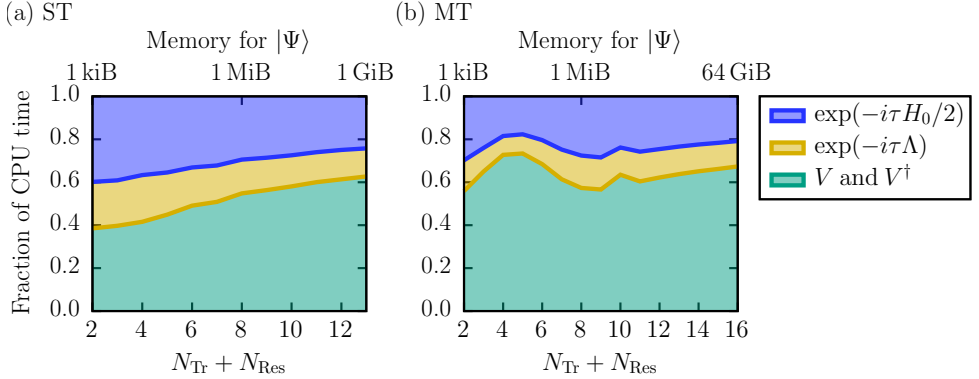
Figure 4.6: Breakdown of the time required for the different transformations for the update rule given in Eq. (4.5); (a) single-threaded case using one core, (b) multi-threaded case using 48 threads on 24 cores. The run-time fractions are shown for both $\exp(-i\tau H_0/2)$ transformations (blue), for the $\exp(-i\tau\Lambda)$ transformation (yellow), and for both V and $V^\dagger$ (green). The memory required to store the state vector $|\Psi\rangle$ (top axis) is linked to the system size (bottom axis) via Eq. (4.6).

4.2 Single transmon-resonator system

In this section, we study the elementary case of a single transmon coupled to a single resonator. Such a system is small enough to make detailed analytical investigations possible. It has often been considered in the literature (e.g., in the initial proposals of the circuit QED architecture [Bla2004] and the transmon [Koc2007]). Also, the system can be related to the famous Jaynes-Cummings model [Jay1963]. Therefore, this simple setup is an ideal candidate to relate results from the transmon simulator to previous analytical work.

The Hamiltonian of the single transmon-resonator system is given by

$$H_{\text{Single}} = 4E_C \hat{n}^2 - E_J \cos \hat{\varphi} + \Omega \hat{a}^\dagger \hat{a} + G\hat{n}(\hat{a} + \hat{a}^\dagger), \tag{4.7}$$

and corresponds to H^{free} given in Eqs. (4.1a)–(4.1c) for $N_{\text{Tr}} = N_{\text{Res}} = 1$. We use the model parameters of a corresponding system manufactured at KIT (see Tab. 3.2). We first give a brief summary of some analytical results for this model and then check their range of validity with the simulation.

4.2.1 Overview of known perturbative results

The Hamiltonian H_{Single} given in Eq. (3.16) expressed in the transmon basis $\{|m\rangle\}$ reads

$$H_{\text{Single}} = \sum_m E_m^{\text{Tr}} |m\rangle\langle m| + \Omega \hat{a}^\dagger \hat{a} + \sum_{mm'} G n^{(m,m')} |m\rangle\langle m'| (\hat{a} + \hat{a}^\dagger), \tag{4.8}$$

where E_m^{Tr} denotes the transmon eigenenergies (see Eq. (3.14)), and $n^{(m,m')}$ denotes the matrix elements of the charge operator $\hat{n}$ in the transmon basis (see Eq. (3.15)). In this representation, H_{Single} resembles the Jaynes-Cummings model [Jay1963], which is why it is often called *generalized Jaynes-Cummings Hamiltonian* [Koc2007].

Several approximations to this Hamiltonian are frequently found in the literature. The first step in most of them is an approximation of the transmon as an anharmonic oscillator (AO), based on the observation that the energy levels E_m^{Tr} are almost equidistant. To apply this approximation, we introduce operators $\hat{a}_{\text{Tr}}$ and $\hat{a}_{\text{Tr}}^\dagger$ (cf. Eqs. (3.45a) and (3.45b)) such that

$$\hat{n} = -\frac{1}{\sqrt{2}} \left(\frac{E_J}{8E_C} \right)^{1/4} (\hat{a}_{\text{Tr}} + \hat{a}_{\text{Tr}}^\dagger), \tag{4.9a}$$

$$\hat{\varphi} = \frac{i}{\sqrt{2}} \left(\frac{8E_C}{E_J} \right)^{1/4} (\hat{a}_{\text{Tr}} - \hat{a}_{\text{Tr}}^\dagger), \tag{4.9b}$$

with $[\hat{a}_{\text{Tr}}, \hat{a}_{\text{Tr}}^\dagger] = 1$. The operators $\hat{n}$ and $\hat{\varphi}$ can always be written in this way because they are Hermitian and obey $[\hat{n}, \hat{\varphi}] = i$ (note that, strictly speaking, the commutator $[\hat{n}, \hat{\varphi}] = i$ is only well defined on the domain of periodic functions; see Appendix B in [Wil2016]). In the literature, the phase factors -1 and i are sometimes put in different places (cf. [Dev1997; Koc2007; Gam2013]), but this does not change the resulting time evolutions.

For the AO approximation, the operators $\hat{a}_{\text{Tr}}$ and $\hat{a}_{\text{Tr}}^{\dagger}$ are effectively replaced by *ladder operators* $\hat{b} = \sum_m \sqrt{m+1}\,|m\rangle\langle m+1|$ and $\hat{b}^{\dagger} = \sum_m \sqrt{m+1}\,|m+1\rangle\langle m|$, respectively (see [Koc2007]). This means that the matrix representation of $\hat{n}$ in the transmon basis (see Eq. (3.15)) is approximated by a tridiagonal matrix.

We first study the accuracy for the case in which the AO approximation is only done for the interaction term $G\hat{n}(\hat{a}+\hat{a}^{\dagger})$, but in combination with the rotating wave approximation (RWA), $(\hat{b}+\hat{b}^{\dagger})(\hat{a}+\hat{a}^{\dagger}) \approx \hat{b}\hat{a}^{\dagger} + \hat{b}^{\dagger}\hat{a}$. We define (cf. Eq. (4.8))

$$H_{\text{Single}}^{\text{AOIntRWA}} = \sum_m E_m^{\text{Tr}}\,|m\rangle\langle m| + \Omega\hat{a}^{\dagger}\hat{a} + g(\hat{b}\hat{a}^{\dagger} + \hat{b}^{\dagger}\hat{a}), \tag{4.10}$$

where

$$g = -\left(\frac{E_J}{32E_C}\right)^{1/4} G \tag{4.11}$$

is the rescaled transmon-resonator coupling.

Since g is small compared to the other energy scales (cf. Tab. 3.2), the Hamiltonian in Eq. (4.10) is often diagonalized in first-order perturbation theory (PT) for the eigenstates $|km\rangle$ of the diagonal part (note that $\hat{a}^{\dagger}\hat{a} = \sum_k k\,|k\rangle\langle k|$ is also diagonal in this basis). One obtains [Koc2007]

$$\overline{|km\rangle} = \frac{1}{\mathcal{N}_{km}}\left(|km\rangle + \frac{g}{\Delta_m}\hat{a}\hat{b}^{\dagger}\,|km\rangle - \frac{g}{\Delta_{m-1}}\hat{a}^{\dagger}\hat{b}\,|km\rangle\right), \tag{4.12}$$

where $\Delta_m = E_{m+1}^{\text{Tr}} - E_m^{\text{Tr}} - \Omega$, and the factor $\mathcal{N}_{km}$ is chosen such that $\overline{|km\rangle}$ is normalized. The states $\overline{|km\rangle}$ are often called *dressed* states since the transmon states are dressed by a small photonic component [Fox2006] (see also [Pom2020]). Using Eq. (4.12), we define the second approximation under investigation by

$$H_{\text{Single}}^{\text{AOIntPT}} = \sum_{km}(E_m^{\text{Tr}} + k\Omega)\overline{|km\rangle\langle km|}, \tag{4.13}$$

where the eigenvalues $E_m^{\text{Tr}} + k\Omega$ are the same as for the original Hamiltonian given in Eq. (4.8), because they are not affected by the off-diagonal interaction operator in first-order PT. Note that the same result can be obtained using the Schrieffer-Wolff transformation [Sch1966; Bra2011] (see [Ric2013] for the calculation).

A third alternative is to extend the AO approximation (which was only done for the interaction before) to the full transmon (cf. [Gam2013]),

$$H_{\text{Single}}^{\text{AOFull}} = \bar{\omega}\hat{b}^{\dagger}\hat{b} + \frac{\bar{\alpha}}{2}\hat{b}^{\dagger}\hat{b}(\hat{b}^{\dagger}\hat{b}-1) + \Omega\hat{a}^{\dagger}\hat{a} + g(\hat{b}+\hat{b}^{\dagger})(\hat{a}+\hat{a}^{\dagger}), \tag{4.14}$$

where $\bar{\omega} = \sqrt{8E_C E_J} - E_C$ and $\bar{\alpha} = -E_C$ are the resulting approximations to the qubit frequency and the anharmonicity, respectively (see also the discussion below Eq. (3.14)). In this form, the transmon's character as an AO is evident: Using $\hat{b}^{\dagger}\hat{b} = \sum_m m\,|m\rangle\langle m|$, Eq. (4.14) states that the eigenvalues E_m^{Tr} are approximated by $\bar{\omega}m + \bar{\alpha}m(m-1)/2$ such

that successive eigenvalues differ by $E_{m+1}^{\text{Tr}} - E_m^{\text{Tr}} \approx \bar{\omega} + \bar{\alpha}m$. Note that unlike Eq. (4.10), the RWA has not been used in Eq. (4.14).

Finally, we consider the two-level approximation (TLA) of the transmon. This is the crudest approximation since it retains only the two lowest-energy eigenstates $|m = 0\rangle$ and $|m = 1\rangle$ of Eq. (4.14). Expressing the ladder operators in terms of Pauli matrices $\sigma^z = -|0\rangle\langle 0| + |1\rangle\langle 1|$ and $\sigma^x = |0\rangle\langle 1| + |1\rangle\langle 0|$, we obtain (up to a constant)

$$H_{\text{Single}}^{\text{TLA}} = -\frac{\bar{\omega}}{2}\sigma^z + \Omega\hat{a}^\dagger\hat{a} + g\sigma^x(\hat{a} + \hat{a}^\dagger), \tag{4.15}$$

which resembles the Jaynes-Cummings model of an atom coupled to an electric field (see [Jay1963; Ger2005]), except that the RWA has not been used. The reason for not using the RWA is that all deviations from the exact time evolution can then be attributed to the TLA.

In summary, we list all of these approximations in terms of the transmon states $\{|m\rangle\}$ or the dressed transmon-resonator states $\{\overline{|km\rangle}\}$, respectively:

$$H_{\text{Single}}^{\text{AOIntRWA}} = \sum_m E_m^{\text{Tr}}\,|m\rangle\langle m| + \Omega\hat{a}^\dagger\hat{a}$$
$$+ \sum_m g\sqrt{m+1}(|m\rangle\langle m+1|\,\hat{a}^\dagger + |m+1\rangle\langle m|\,\hat{a}), \tag{4.16a}$$

$$H_{\text{Single}}^{\text{AOIntPT}} = \sum_{km}(E_m^{\text{Tr}} + k\Omega)\overline{|km\rangle\langle km|}, \tag{4.16b}$$

$$H_{\text{Single}}^{\text{AOFull}} = \sum_m(\bar{\omega}m + \frac{\bar{\alpha}}{2}m(m-1))\,|m\rangle\langle m| + \Omega\hat{a}^\dagger\hat{a}$$
$$+ \sum_m g\sqrt{m+1}(|m\rangle\langle m+1| + |m+1\rangle\langle m|)(\hat{a} + \hat{a}^\dagger), \tag{4.16c}$$

$$H_{\text{Single}}^{\text{TLA}} = \bar{\omega}\,|m = 1\rangle\langle m = 1| + \Omega\hat{a}^\dagger\hat{a}$$
$$+ g(|m = 0\rangle\langle m = 1| + |m = 1\rangle\langle m = 0|)(\hat{a} + \hat{a}^\dagger). \tag{4.16d}$$

4.2.2 Comparison to simulation results

We measure the success of the different approximations given in Eqs. (4.16a)–(4.16d) by how well they predict the time evolution. For every Hamiltonian $\widetilde{H} \in \{H_{\text{Single}}^{\text{AOIntRWA}}, H_{\text{Single}}^{\text{AOIntPT}}, H_{\text{Single}}^{\text{AOFull}}, H_{\text{Single}}^{\text{TLA}}\}$ in Eqs. (4.16a)–(4.16d), we obtain the time evolution by numerical diagonalization. This is done by diagonalizing the matrix representation of $\widetilde{H}$ with respect to the joint transmon-resonator basis $\{|km\rangle\}$ for $k, m \in \{0, \ldots, 9\}$. We then use the resulting eigenvalues and eigenvectors of $\widetilde{H}$ to compute the time-evolution operator $\widetilde{U}(t) = \exp(-i\widetilde{H}t)$. Note that this procedure is only possible since the matrices are small enough and the Hamiltonians do not contain time-dependent terms.

We compute the probability $p_{km}(t)$ for $0 \le t \le 100\,\text{ns}$ to measure the system in the state $|km\rangle$ if it has been initialized in $|km\rangle$ at $t = 0$. Using the time-evolution operator $\widetilde{U}(t)$, this probability is given by

$$p_{km}(t) = |\langle km|\widetilde{U}(t)|km\rangle|^2. \tag{4.17}$$
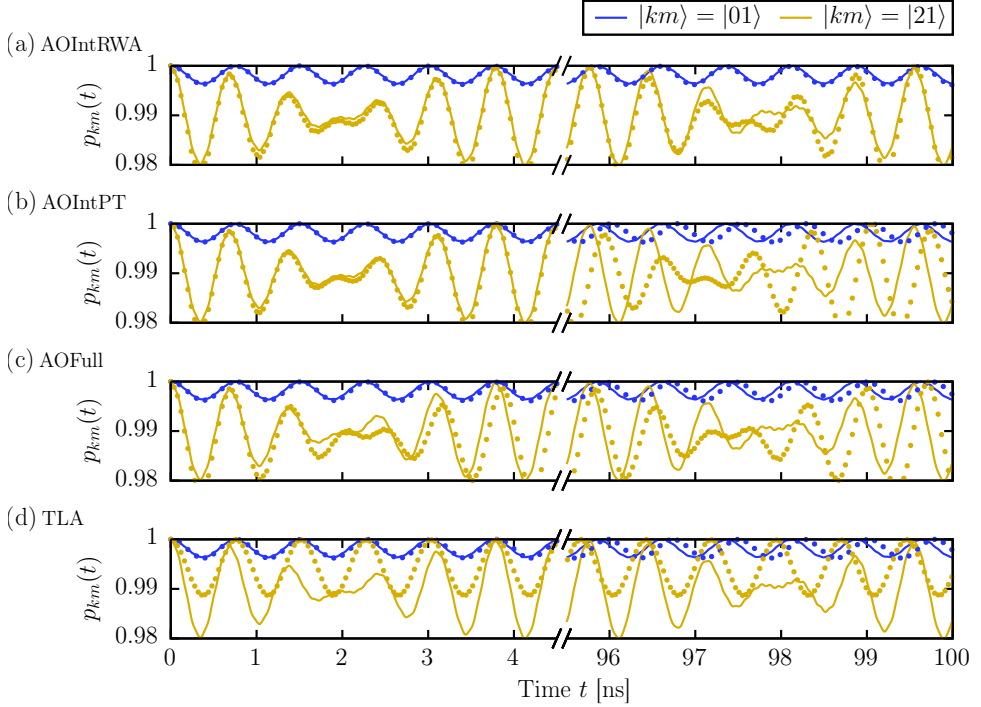
Figure 4.7: Time evolution of two different initial states $|km\rangle = |01\rangle, |21\rangle$ to compare the exact solution (solid lines) and various common approximations (dots): (a) using an anharmonic oscillator for the interaction term combined with the RWA (see Eq. (4.10)), (b) the same but in first-order PT (see Eq. (4.13)), (c) using an anharmonic oscillator for the whole transmon (see Eq. (4.14)), (d) using a two-level system to describe the transmon (see Eq. (4.15)). Shown is the probability $p_{km}(t)$ to measure the system in the state $|km\rangle$ (see Eq. (4.17)). The exact time evolution corresponds to the Hamiltonian H_{Single} given in Eq. (4.7), and the approximations (a)–(d) correspond to the Hamiltonians given in Eqs. (4.16a)–(4.16d)

The reference time evolution is the exact result governed by the Hamiltonian given in Eq. (4.7). It is obtained using the transmon simulator defined in Section 3.3 (note that the error for this system can be made arbitrarily small by scaling the time step τ; see Fig. 4.1(a)). Explicitly, using the notation from Eq. (3.23), we set $\psi_{k_0 m_0}(0) = \delta_{k_0 k} \delta_{m_0 m}$, perform the time evolution, and compute $p_{km}(t) = |\psi_{km}(t)|^2$.

The result is shown in Fig. 4.7 for the initial states $|km\rangle = |01\rangle, |21\rangle$, i.e., the transmon is always initialized in the excited qubit state $|m = 1\rangle$ and the resonator is populated with either $k = 0$ or $k = 2$ photons. We see that if the AO approximation is only used for the interaction, both the short-term and the long-term evolution are described reasonably well (see Fig. 4.7(a)). We checked that in this case, the result is the same whether the RWA is used for this term or not. The perturbative result shown in Fig. 4.7(b) can describe the short-term evolution equally well, but acquires a shift for longer time evolutions. The same applies if the whole transmon is approximated as an anharmonic oscillator (see Fig. 4.7(c)). However, in this case a slight drift for the $k = 2$ case is also already observable at $t = 4$ ns. Still, in all approximations for which the transmon is described by more than two levels, the average amplitude of the oscillations is correct. As shown in Fig. 4.7(d), this is not true anymore for the two-level approximation. It suffers both from a drift and also from a reduced amplitude in the $k = 2$ case. The strength of the interaction for $k \neq 0$ is thus underestimated.

We can conclude from this that higher transmon states play an important role in mediating the interaction between the transmon and the resonator. We find that none of the approximations are suitable for the optimization of quantum gate pulses which take $\mathcal{O}(100\,\text{ns})$. The reason is that all approximations develop a drift at this time scale, and drifts are related to inaccurate relative phases which are required to be very precise in order to implement e.g. the CNOT gate. As shown in Section 4.4, however, there still exist nontrivial cases in which even a two-level approximation describes the time evolution rather well.

4.3 Transmon-resonator system coupled to a bath

In practice, it is almost impossible to keep readout resonators completely void of photons (see, for instance, [Sur2015; Bul2016]). This means, in the context of this work, that it is difficult to prepare a resonator exactly in the state $|k = 0\rangle$ (cf. Eq. (3.23)). An interesting question is how much a transmon qubit is affected by the presence of photons in the resonator.

We study this question using three complimentary approaches. First, we consider an isolated transmon-resonator system as defined in Section 3.4.1. Next, we study its coupling to a heat bath using the model defined in Section 3.4.2 (this model is the same that can be used to describe electromagnetic environments, see Section 3.5). These first two approaches are based on solving the TDSE given in Eq. (3.8) using the transmon simulator described in Section 3.3. For the third approach, we use a quantum master equation approach. This allows us to address the transition from a closed system over the system-bath model to the effective evolution described by a quantum master equation.

4.3.1 Simulation models

The central system considered in this section is the single transmon-resonator system defined in Section 3.4.1. We supplement the system with a bath of $L = 10$ resonators to model an open quantum system. The bath Hamiltonian essentially represents a collection of harmonic oscillators, i.e., a boson bath. Thus, the model Hamiltonian for the TDSE is given by $H = H_{\text{Single}} + H_{\text{Bath}}$, where

$$H_{\text{Single}} = 4E_C\hat{n}^2 - E_J\cos\hat{\varphi} + \Omega\hat{a}^\dagger\hat{a} + G\hat{n}(\hat{a} + \hat{a}^\dagger), \tag{4.18a}$$

$$H_{\text{Bath}} = \sum_{l=1}^{L} W_l\hat{b}_l^\dagger\hat{b}_l + \sum_{l=1}^{L} \lambda_l(\hat{a} + \hat{a}^\dagger)(\hat{b}_l + \hat{b}_l^\dagger). \tag{4.18b}$$

For clarity, we use the symbols $\hat{b}_l$ to distinguish the bath resonators from the central resonator. Thus, the mapping to the model Hamiltonian $H_{\text{Res}}^{\text{free}}$ given in Eq. (4.1c) is $\hat{a}_{r=0} \leftrightarrow \hat{a}$ and $\hat{a}_l \leftrightarrow \hat{b}_l$ for $l = 1, \ldots, L$. The specification of the bath parameters and their relation to the parameters of the full model Hamiltonian used in the simulation framework (see Eqs. (3.9a)–(3.9f)) is given in Tab. 3.3. The topology of the system is sketched in Fig. 3.3.

Isolated system

The first approach focuses on the isolated transmon-resonator system. As in Section 4.2, we solve the TDSE

$$i\frac{\partial}{\partial t}\ket{\Psi(t)} = H_{\text{Single}}\ket{\Psi(t)}, \tag{4.19}$$

where H_{Single} is given in Eq. (4.18a), corresponding to H^{free} given in Eqs. (4.1a)–(4.1c) for $N_{\text{Tr}} = N_{\text{Res}} = 1$. We use the model parameters of a transmon system manufactured at KIT (see Tab. 3.2).

The system's initial state is given by $\ket{\Psi(0)} = \ket{k, m = 0}$ such that the transmon is initialized in its ground state and the resonator is populated with k photons. We intentionally do not initialize the system in an eigenstate of the full Hamiltonian (such as the dressed states given in Eq. (4.12)). The reason for this is that in an eigenstate, the time evolution would be trivial for each k. Instead, the goal is to assess the impact of k additional photons on the transmon-resonator interaction and its consequences for the time evolution of the transmon system.

We analyze the free time evolution for various k between 0 and 180. At each time t, we evaluate the probability $p_{m\neq0}(t)$ to find the transmon in a higher excited state $\ket{m}$ with $m \neq 0$. This probability is given by

$$p_{m\neq0}(t) = 1 - \bra{\Psi(t)}(\ket{m=0}\bra{m=0})\ket{\Psi(t)}, \tag{4.20}$$

where $\ket{\Psi(t)}$ is given in Eq. (3.23). Since the Hilbert space given in Eq. (3.17) includes only four states for the resonator, we make use of the parameter k^{offset} to study larger values of k. Specifically, we set $k^{\text{offset}} = k - 2$ (for $k \geq 2$) such that two Fock states

below $|k\rangle$ and one Fock state above $|k\rangle$ are taken into account. We have verified by exact diagonalization that with this choice, the quantities studied for this system are accurate up to three significant digits.

System with bath

For the second approach, we use the transmon simulator to solve the TDSE

$$i\frac{\partial}{\partial t}|\Psi(t)\rangle = (H_{\text{Single}} + H_{\text{Bath}})|\Psi(t)\rangle, \tag{4.21}$$

where H_{Single} and H_{Bath} are given in Eqs. (4.18a) and (4.18b), respectively. We consider various coupling strengths λ_l that are chosen uniformly from $[0, \lambda]$, where $\lambda \in 2\pi \times \{5\,\text{MHz}, 10\,\text{MHz}, 20\,\text{MHz}\}$. Furthermore, the frequencies W_l of the bath modes are chosen randomly from a Gaussian distribution centered around Ω (see Tab. 3.3). The choice of random bath parameters is motivated by the observation that for such large, generic models, randomness in the bath parameters is required to model generic effects [Jin2013]. Note that in principle, this particular system-bath setup is a sufficiently general model for a superconducting environment (see Section 3.5, where we also describe a procedure to extract the bath parameters W_l and λ_l from electromagnetic HFSS simulations).

We solve Eq. (4.21) for an initial state given by a product of the state $|k, m = 0\rangle$ (as before) and the zero temperature ground state of the bath. Note that the state vectors $|\Psi(t)\rangle$ in Eqs. (4.19) and (4.21) are defined on different Hilbert spaces. Specifically, the Hilbert space for Eq. (4.19) includes one transmon and one resonator, whereas the one for Eq. (4.21) contains 11 resonators and is thus much larger (cf. Eq. (3.23)). In the case $\lambda = 0$, however, the solution of Eq. (4.21) coincides with the solution of Eq. (4.19) after projection onto the smaller Hilbert space. All simulations for the approach given by Eq. (4.21) were performed on the supercomputers JURECA [Jül2018] and JUWELS [Jül2019].

As before, we study the effect of a larger number of photons k in the system's resonator by evaluating Eq. (4.20). As opposed to the isolated case, the photons now directly couple to the environment. This means that for a larger number of photons k, energy can leak out of the resonator into the environment. Furthermore, energy exchange between the transmon and the environment can be virtually mediated by the resonator. Therefore, a reasonable expectation would be that the average transmon excitation $\langle p_{m\neq0}(t)\rangle$ is reduced.

We remark that simulating Eq. (4.21) also allows for a study of the theory used to describe the Purcell effect for transmons, which starts from a bath model similar to Eq. (4.18b) to derive the transmon relaxation rate (see [Koc2007]).

Master equation

As a third approach, we consider a Lindblad master equation (also known as GKLS equation) [Gor1976; Lin1976] for the system's reduced density matrix $\rho(t)$,

$$\frac{\partial}{\partial t}\rho(t) = -i[H_{\text{Single}}, \rho(t)] + \sum_{\gamma} \kappa_{\gamma} \mathcal{D}[A_{\gamma}](\rho(t)), \tag{4.22}$$

where the dissipator $\mathcal{D}[A_\gamma]$ is defined by $\mathcal{D}[A_\gamma](B) = A_\gamma B A_\gamma^\dagger - (A_\gamma^\dagger A_\gamma B + B A_\gamma^\dagger A_\gamma)/2$ for some operator B. This type of master equation is the most general form for the generator of the quantum dynamical semigroup [Bre2007]. The associated time evolution is automatically CPTP (cf. Section 2.4). In this context, the reference to the algebraic structure of a semigroup is due to the fact that an inverse time evolution is not required [Kos1972] (unitary time evolutions, on the other hand, form a group because they are reversible). A comprehensive historical account of the events that led to Eq. (4.22), including a survey of the results, is given in [Chr2017].

For a single dissipator $\mathcal{D}[\hat{a}]$, the master equation in Eq. (4.22) reads

$$\frac{\partial}{\partial t}\rho(t) = -i[H_{\text{Single}}, \rho(t)] + \frac{\kappa}{2}\left(2\hat{a}\rho(t)\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\rho(t) - \rho(t)\hat{a}^\dagger\hat{a}\right), \tag{4.23}$$

where $\kappa = 2\pi \times 2.7\,\text{MHz}$ denotes the photon loss rate obtained from a corresponding experiment at KIT [Rie2019]. In addition to the single dissipator $\mathcal{D}[\hat{a}]$, one frequently considers effects of the form $\mathcal{D}[\hat{a}^\dagger]$ for environment-induced photon excitation, or corresponding versions for qubit relaxation, excitation, and dephasing (see e.g. [Sur2015]). The motivation to consider $\mathcal{D}[\hat{a}]$ the most dominant dissipator is that the readout resonator is directly coupled to the "outside world" (i.e, the transmission line, whose temperature may be much higher than the effective temperature of the transmons). Therefore, photons in the resonator can easily leak out and take energy away from the system. This is especially true if the number of photons k in the resonator is large.

In the case of the master equation, the solution is given by the system's density matrix $\rho(t)$. Therefore, the probability from Eq. (4.20) to find the transmon in a higher excited state $|m\rangle$ with $m \neq 0$ is given by

$$p_{m\neq 0}(t) = 1 - \langle m = 0|\rho(t)|m = 0\rangle. \tag{4.24}$$

We emphasize that the bath approach in Eq. (4.21) is more general than the other two: The solution reduces to the one of Eq. (4.19) if the bath coupling strength $\lambda = 0$, and a Markovian master equation of the form of Eq. (4.23) for the reduced density matrix $\rho(t) = \text{Tr}_{\text{Bath}}|\Psi(t)\rangle\langle\Psi(t)|$ can describe the more complicated dynamics generated by a system-bath model only under certain conditions [Bre2007] (see [Zha2016; DeR2017] for detailed investigations of this point).

4.3.2 Results

The time evolution of $p_{m\neq 0}(t)$ for $k = 0, 20, \ldots, 180$ photons is shown in Fig. 4.8. For the isolated system, Fig. 4.8(a) shows oscillations of the transmon between $|m = 0\rangle$ and higher excited states. For $k = 180$ (red curve), this oscillation is quite strong such that after $0.33\,\text{ns}$, $p_{m\neq 0}(t)$ already reaches 46%. A closer inspection (data not shown) yields that at this point in time, the probability to find $|m = 1\rangle$ ($|m = 2\rangle$) is 37% (9%). Hence, an increased number of photons can also excite the qubit to higher, non-computational states. Figure 4.8 also shows that with increasing k, the frequency of the oscillation between $|m = 0\rangle$ and higher states grows.
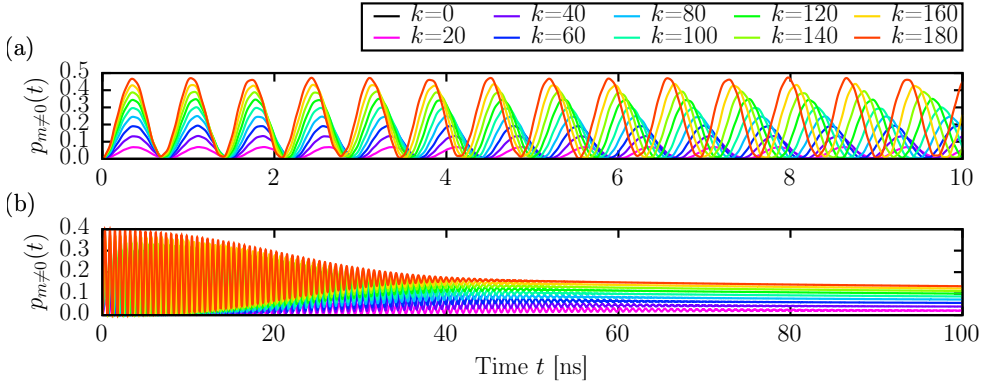
Figure 4.8: Time evolution of the probability $p_{m \neq 0}(t)$ to find the transmon in any excited state $|m\rangle$ for $m \neq 0$. $p_{m \neq 0}(t)$ is obtained by solving (a) the TDSE given in Eq. (4.19) (see Eq. (4.20)), (b) the master equation given in Eq. (4.23) (see Eq. (4.24)). Different colors correspond to a different number of photons k in the resonator.
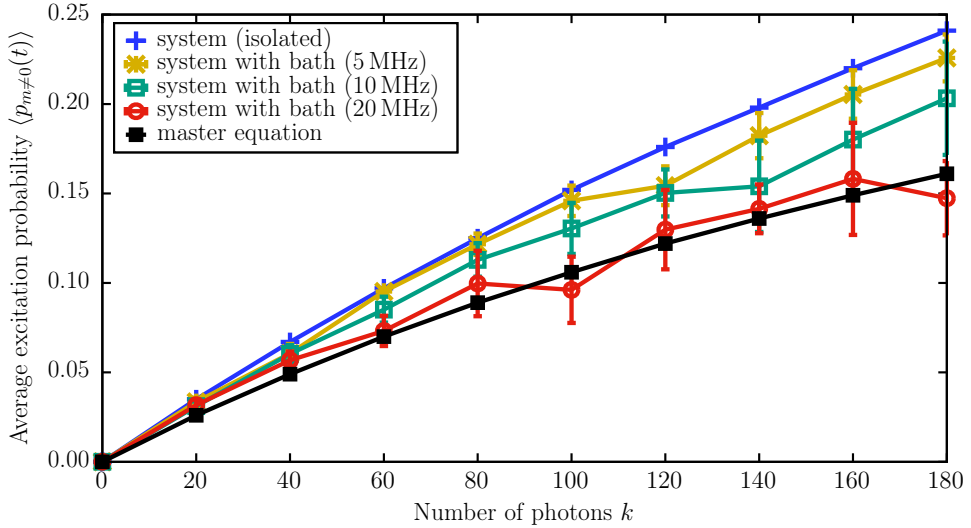
Figure 4.9: Average probability to find the transmon in any excited state $|m\rangle$ for $m \neq 0$ as a function of the number of photons k in the resonator, using the three approaches described in Section 4.3.1. Each point corresponds to the time average $\langle p_{m \neq 0}(t) \rangle$ given in Eq. (4.25). For the isolated system and the master equation, the corresponding probabilities $p_{m \neq 0}(t)$ are shown in Fig. 4.8. For the bath simulations, each point (error bar) is the mean (standard deviation) of 10 independent results for $\langle p_{m \neq 0}(t) \rangle$, each of which uses different, random bath parameters that are distributed as specified in Tab. 3.3.

In Fig. 4.8(b), we show the corresponding probability $p_{m \neq 0}(t)$ as described by the Lindblad master equation given in Eq. (4.23). For the first 10 ns, the time evolution is almost equal to the isolated transmon-resonator case shown in Fig. 4.8(a). After that, the photon loss modeled by the dissipator $\mathcal{D}[\hat{a}]$ becomes observable, and the oscillations of $p_{m \neq 0}(t)$ decay. Note that the state at the end of the depicted time evolution is not the steady state, since the photon-loss mechanism in this model would continue to take energy from the system until the resonator is completely depleted of photons.

We average the probability $p_{m \neq 0}(t)$ shown in Fig. 4.8 over a period $0 \leq t \leq T$ to obtain the average excitation probability

$$\langle p_{m \neq 0}(t) \rangle = \frac{1}{T} \int\limits_0^T p_{m \neq 0}(t) \, \mathrm{d}t, \tag{4.25}$$

for each number of photons $k = 0, 20, \ldots, 180$ and each of the three approaches introduced above. For the TDSE-based approaches given by Eqs. (4.19) and (4.21), we take $T = 20$ ns such that enough oscillations contribute to the average. For the master-equation approach given by Eq. (4.23), we take $T = 100$ ns, because this is roughly the time scale of the measurement process in the corresponding experiment [Rie2019]. The result is shown in Fig. 4.9.

For the bath simulations (yellow, green, and red curves in Fig. 4.9), we additionally average $\langle p_{m \neq 0}(t) \rangle$ over ten independent runs using different bath parameters. Since for each k, we also have random parameters, the averages give a clear indication of the generic trend. However, we also see fluctuations (represented by error bars) for different baths. See below for an analysis of this point.

Generically, we see that the average excitation probability increases with the number of photons k. The overall effect is most pronounced for the isolated system (blue line in Fig. 4.9). This makes sense since without an environment, energy can only be exchanged between the transmon and the resonator (see Eq. (4.18a)). Qualitatively, this interaction is proportional to $\hat{a} + \hat{a}^\dagger \sim \sqrt{k+1}$ (cf. Eq. (3.11)). Therefore, a larger number of photons k can lead to a higher excitation of the transmon.

In the system-bath models, however, energy from the resonator and (in second order) from the transmon can dissipate into the environment. Therefore, we see in Fig. 4.9 that the transmon excitation decreases. In particular, for increasing system-bath couplings λ, the results approach the purely dissipative situation modeled by the Lindblad master equation (black line). Note that in general, TDSE dynamics of a system coupled to a bath can exhibit much more complicated, non-Markovian behavior that is incompatible with a Lindblad master equation [DeR2017].

Bath fluctuations

The remarkably nice transition from an isolated system (TDSE) to a dissipative system (master equation) by gradually increasing the system-bath coupling λ cannot be observed in every simulation. This is the reason for the fluctuations represented by the error bars in Fig. 4.9, especially for stronger system-bath couplings (see the red line corresponding to $\lambda = 2\pi \times 20$ MHz).
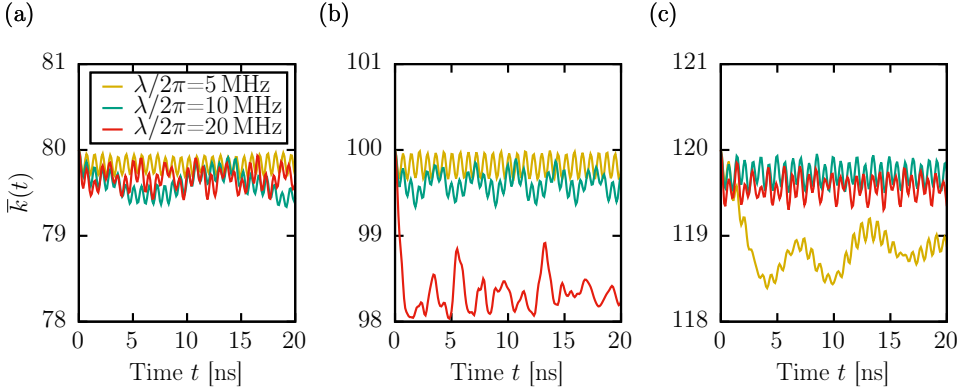
Figure 4.10: Time evolution of the average number of photons $\overline{k}(t)$ (see Eq. (4.26)) in the resonator in the presence of the bath described by Eq. (4.18b). The model is schematically shown in Fig. 3.3. For each initial number of photons $\overline{k}(0) = k$ with (a) $k = 80$, (b) $k = 100$, and (c) $k = 120$, we show one time evolution from the 10 independent runs that yield a point on the yellow, green, and red lines in Fig. 4.9. The same colors yellow, green, and red are used to represent different coupling strengths λ.

A closer look at the data shows that most bath configurations follow the average, but a few particular configurations cause a dip in the excitation probability. This dip is observable in the red line at $k = 100$ and $k = 180$ in Fig. 4.9. To understand the reason for this, we evaluate the time evolution of the average photon number in the resonator,

$$\overline{k}(t) = \langle \Psi(t)| \, \hat{a}^\dagger \hat{a} \, |\Psi(t)\rangle \,, \tag{4.26}$$

where $|\Psi(t)\rangle$ is the solution of the TDSE given by Eq. (4.21).

Figure 4.10(a) shows a representative result for $\overline{k}(t)$ for each λ, taken from one of the ten runs for each λ corresponding to $k = 80$ in Fig. 4.9. We see that over the course of the time evolution, $\overline{k}(t)$ stays nicely within the four simulated Fock states (cf. Eq. (3.17)).

However, the red line in Fig. 4.10(b) shows an extreme case for $\lambda = 2\pi \times 20\,\mathrm{MHz}$, in which the average photon number $\overline{k}(t)$ immediately drops and hits the computational boundary at $k = 98$. This case corresponds to a particular configuration of the bath that causes a dip in the corresponding red line in Fig. 4.9 at $k = 100$. A similar situation is depicted in Fig. 4.10(c) for $\lambda = 2\pi \times 5\,\mathrm{MHz}$ (yellow line). This is the cause of the slightly less pronounced dip in the corresponding yellow line at $k = 120$ in Fig. 4.9.

To understand this immediate drop in the photon number, we investigated the corresponding spectral properties of the bath. In many cases, there is a certain bath mode with a frequency W_l close to the resonator frequency Ω and a particularly strong coupling λ_l (see Eq. (4.18b)). It seems reasonable that such a resonant condition leads to a special situation. However, this explanation does not hold for each instance that exhibits this behavior, and it is complicated to recognize the resonant pathway in all cases. This points out an opportunity to improve the model and suggests an interesting venue for further research.

4.3.3 Additional ways to improve the models

For the results presented above, we initialized the resonator in a Fock state $|k\rangle$ with k photons. A more adequate model of the experimental situation would be to consider a coherent state [Gla1963; Fox2006], given by

$$|\alpha\rangle = \sum_{k'} e^{-|\alpha|^2/2} \frac{\alpha^{k'}}{k'!} |k'\rangle \,. \tag{4.27}$$

For such a state, $\alpha = \sqrt{k}$ would represent the electromagnetic field in the resonator with an average photon number k. We ran simulations including up to 300 Fock states (data not shown) and found that the average excitation probability $\langle p_{m\neq0}(t)\rangle$ shown in Fig. 4.9 is only marginally reduced (from 24% to 22% at the maximum for $k = 180$). For the master equation, the effect is even weaker, with a decrease by less than 0.1%. The time evolution $p_{m\neq0}(t)$, however, is more interesting, with sharply peaked oscillations at a 5–10 times higher frequency on top of the curves shown in Fig. 4.8.

Future work will go into extending the model such that larger baths with additional Fock states can be simulated. This would allow a precise classification of the bath configurations that lead to the immediate drop in the photon number illustrated in Fig. 4.10(b) for $\lambda = 2\pi \times 20\,\text{MHz}$ (red line). Furthermore, it enables an initialization of the bath in a thermal state at a certain temperature $T \neq 0$ such that finite temperature effects can be studied.

From a statistical physics point of view, the transition from the isolated system over the system-bath model to the master equation shown in Fig. 4.9 could be scrutinized. In this respect, it would be interesting to study a more general type of master equation including photon excitations $\mathcal{D}[\hat{a}^\dagger]$ and additional dissipators for the transmon itself, which are often used to describe experimental observations (see e.g. [Sur2015]).

In the context of modeling experiments, it would be an exciting idea to use, instead of random bath parameters, the frequencies W_l and couplings λ_l representing the superconducting environment of the particular device. A procedure to extract these parameters from experiments or electromagnetic HFSS simulations of the device is described in Section 3.5. We plan to continue in this direction for a new sample manufactured at KIT [Rie2019].

Finally, an interesting conceptual question is to what extent the generic features are specific to the bosonic bath considered in this section. For instance, an alternative model for an environment would be a spin bath to model a system of two-level defects in materials [Mül2009; Mül2019; Wil2020d; Wil2020e]. Similarly, one could replace the bosonic bath with a fermionic bath of superconducting (Bogoliubov) quasiparticles [Kiv1990], which are also considered to be sources of relaxation, decoherence, and electromagnetic dissipation [Cat2011; Cat2012; Pop2014].

4.4 Effective ZZ interaction for coupled transmons

We consider a pair of transmon qubits coupled by a resonator. This way of coupling transmon qubits is the primary architecture studied in this work. The effective transmon coupling mediated by a resonator has been frequently studied in the literature (see for instance [Bla2004; Maj2007; Li2008; Gam2013; Ric2013; Bil2015; Wil2017; Ku2020]). Analytical calculations often use perturbation theory to obtain dominant effective couplings such as $\sigma_0^x \sigma_1^x$ or $(\sigma_0^x \sigma_1^x + \sigma_0^y \sigma_0^y)/2$. In this section, we study a much weaker coupling of the type $\sigma_0^z \sigma_1^z$. Albeit very small, this coupling is still relevant for experiments; we will later construct a quantum circuit by which its effects can be directly observed in the IBM Q processors (see Section 7.1). Note that it is also possible to obtain this coupling from a complete microwave description of the system [Sol2019].

A coupling of the type $\sigma_0^z \sigma_1^z$, also known as longitudinal coupling, makes the frequency of one qubit depend on the state of the other qubit. To see this, consider a Hamiltonian describing two qubits (Q0, Q1) of the form

$$H_{ZZ} = -\frac{\omega_0'}{2}\sigma_0^z - \frac{\omega_1'}{2}\sigma_1^z + J\sigma_0^z \sigma_1^z. \tag{4.28}$$

When Q1 is in state $|0\rangle$, the frequency of Q0 is $\omega_0' - 2J$ (given by the difference between the eigenvalues of $|10\rangle$ and $|00\rangle$). However, when Q1 is in state $|1\rangle$, the frequency of Q0 is $\omega_0' + 2J$. This means that the frequency of Q0 depends on the state of Q1. In the same manner, the frequency of Q1 depends on the state of Q0.

To study this effect for a pair of transmon qubits, we simulate a system of $N_{\mathrm{Tr}} = 2$ transmons coupled by $N_{\mathrm{Res}} = 1$ resonator (see Tab. 3.4 for the model parameters). We determine the frequency of Q0 for three different initial states of Q1, namely $|0\rangle$, $|+\rangle$, and $|1\rangle$. To determine the frequency, we make use of the procedure described in Section 3.3.2. Thus, we initialize Q0 in the state $|+\rangle$ and obtain the frequency from the time evolution of its Bloch vector $\vec{r}_0(t)$ (see Eqs. (3.39) and (3.40)). The resonator is always initialized in its ground state $|k = 0\rangle$.

Figure 4.11 shows a plot of the time evolution of the Bloch vector's x component $r_0^x(t)$ for the three different initial states of Q1, along with the frequencies determined by the procedure described in Section 3.3.2. As expected, the frequency of Q0 depends on the state of Q1. Since the difference in frequency is on the order of $\Delta f \approx 0.0001\,\mathrm{GHz}$, we need to simulate the time evolution up to $1\,\mu s$ (corresponding to $1\,\mu s/\tau = 10^6$ time steps) to observe the difference in f.

As can be seen, the frequencies used for the cosine functions (lines) describe the simulation results (points) very accurately. However, on closer inspection, we see that the amplitude of the oscillation corresponding to the case where Q1 is in state $|+\rangle$ (yellow circles) decreases over time, which is not described accurately by $\cos(2\pi ft)$. The reason for this is that the magnitude of the Bloch vector $\vec{r}_0(t)$ becomes smaller which corresponds to entanglement building up in the state.

Interestingly, the effect is described correctly by the effective Hamiltonian H_{ZZ} given
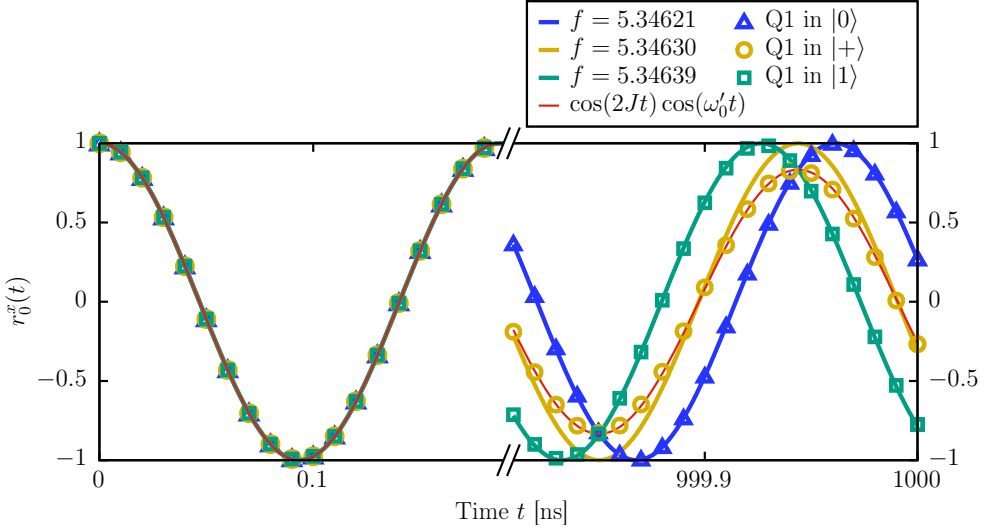
Figure 4.11: Time evolution of the x component of the Bloch vector $\vec{r}_0(t)$ of Q0 (cf. Eq. (3.38)), for different initial states of Q1 (indicated by blue, yellow, and green colors). Points represent simulation results. Blue, yellow, and green lines correspond to the functions $\cos(2\pi f t)$, where the frequency f is given in GHz in the legend. The red line corresponds to the function $\cos(2Jt)\cos(\omega_0' t)$ (see Eq. (4.30)), where J and ω_0' are given in Eqs. (4.31a) and (4.31c), respectively. The difference in the frequencies is only observable after a long time evolution (right panel of the plot).

by Eq. (4.28). To see this, we compute the corresponding time evolution under H_{ZZ},

$$e^{-it H_{ZZ}} |{++}\rangle = \frac{1}{2}\Big(e^{it(\omega_0' + \omega_1' - 2J)/2} |00\rangle + e^{it(\omega_0' - \omega_1' + 2J)/2} |01\rangle$$
$$+ e^{it(-\omega_0' + \omega_1' + 2J)/2} |10\rangle + e^{it(-\omega_0' - \omega_1' - 2J)/2} |11\rangle\Big). \tag{4.29}$$

Evaluating the expectation value $r_0^x(t) = \langle \sigma_0^x \rangle$ for this state yields

$$r_0^x(t) = \cos(2Jt)\cos(\omega_0' t). \tag{4.30}$$

The parameters ω_0', ω_1', and J can be obtained from the time evolution of the four states ($|{+}0\rangle, |{+}1\rangle, |0{+}\rangle, |1{+}\rangle$). Using the same procedure as before, we extract the corresponding four frequencies $f_i^{\pm}$ from the data for $\vec{r}_i(t)$ (see Eqs. (3.39) and (3.40)). For instance, the frequency f_0^- (f_0^+) obtained from the time evolution of $|{+}0\rangle$ ($|{+}1\rangle$) corresponds to the blue (green) line in Fig. 4.11. The frequencies $f_i^{\pm}$ are related to the parameters ω_0', ω_1', and J via $2\pi f_i^{\pm} = \omega_i' \pm 2J$ (see Eq. (4.28)). Therefore, we have $\omega_i' = 2\pi(f_i^+ + f_i^-)/2$ and $J = 2\pi(f_i^+ - f_i^-)/4$ (which is the same for both $i = 0$ and $i = 1$ up to 10^{-10} GHz),

which evaluates to

$$\omega_0' = 2\pi \times 5.346300 \, \text{GHz}, \tag{4.31a}$$

$$\omega_1' = 2\pi \times 5.116707 \, \text{GHz}, \tag{4.31b}$$

$$J = 2\pi \times 46.6 \, \text{kHz}. \tag{4.31c}$$

The frequencies ω_i' are shifted with respect to the individual qubit frequencies $\tilde{\omega}_i$ given in Tab. 3.4 due to the presence of the resonator. Note the large number of significant digits that is required to resolve the value of J.

The function $\cos(2Jt)\cos(\omega_0't)$ given in Eq. (4.30) is shown as a red line in Fig. 4.11. We see that it describes the decrease in amplitude very accurately, despite the extremely small value of J. In Section 6.3, we will reproduce the same value for J (see Eq. (6.32)) using a much more sophisticated procedure called *gate set tomography* (GST).

It is worth mentioning that the effective Hamiltonian H_{ZZ} given in Eq. (4.28) can also be derived from the original transmon-resonator Hamiltonian by doing a perturbative diagonalization. Such a calculation is given in [Gam2013] and, using a more recent technique, in [Mag2020]. The starting point of the calculation is a two-transmon version of the anharmonic oscillator Hamiltonian $H_{\text{Single}}^{\text{AOIntRWA}}$ given in Eq. (4.10). Furthermore, in [Bil2015], a derivation starting from a two-qubit version of $H_{\text{Single}}^{\text{TLA}}$ given in Eq. (4.15) is presented, without resorting to the RWA. Both calculations yield the correct type of longitudinal ZZ interaction. Furthermore, the order of magnitude of the frequency corrections is right: The difference between ω_i' given in Eqs. (4.31a) and (4.31b) and the original frequencies $\tilde{\omega}_i$ given in Tab. 3.4 is approximately equal to the Lamb shift $-g_i^2/(\Omega - \tilde{\omega}_i) \in \{2\pi \times -0.0035 \, \text{GHz}, 2\pi \times -0.0029 \, \text{GHz}\}$, where $g_i = -(E_{Ji}/32E_{Ci})^{1/4}G_i$. However, the respective values for the longitudinal coupling strength J given in Eq. (4.31c) are different:

$$J^{\text{AOIntRWA}} \approx 2\pi \times 240 \, \text{kHz}, \tag{4.32a}$$

$$J^{\text{TLA}} \approx 2\pi \times 5 \, \text{kHz}. \tag{4.32b}$$

The interaction strength J^{AOIntRWA} is stronger because it refers to the eigenbasis obtained after the perturbative diagonalization. The reason for J^{TLA} being too small is that the TLA discards contributions from higher levels. Thus, higher transmon states play an important role in mediating the exchange interaction between transmon qubits. Independent of these deviations, however, all perturbative calculations yield the important scaling law $J \propto G^4$, showing that the resonator-mediated exchange interaction depends sensitively on the respective transmon-resonator couplings.

Although the magnitude of the coupling is extremely small (see Eq. (4.31c)), one can find circuits by which its consequences, namely the state-dependent frequencies, can be directly observed in the IBM Q processors (see Section 7.1). Furthermore, the effective evolution described by H_{ZZ} given in Eq. (4.28), including the same value for J given in Eq. (4.31c), is also found, without using prior information, by the black box model of GST (see Eq. (6.32) in Section 6.3).

4.5 Conclusions

In this chapter, we studied undriven time evolutions. After verifying known error bounds for the transmon simulator (see Eqs. (4.2a) and (4.2b)), we found that new error bounds for observables [Wil2020d] (see Eqs. (4.3a) and (4.3b)) are indeed tight. In detailed benchmarks, we observed that the simulation algorithm exhibits nearly ideal weak and strong scaling behavior (see Figs. 4.3–4.6) on the supercomputer JURECA [Jül2018] for implementations 1 and 2 (cf. Appendix D). We compared the time evolution produced by the transmon simulator to known perturbative results, and found that the perturbative results qualitatively predict the time evolution properly, but all develop a drift after a short time (see Fig. 4.7), which makes them unsuitable for accurate pulse optimization. By coupling an isolated system to a bosonic heat bath, we observed that system-environment models based on the solution of the TDSE can, under certain conditions, be effectively described by the CPTP dynamics generated by a Lindblad master equation (see Fig. 4.9). Finally, we characterized the resonator-mediated exchange interaction between coupled transmons. Although the magnitude of this interaction is extremely small (see Eq. (4.31c)), we note that it is accurately reproducible using a black box model (see Section 6.3.2). Furthermore, its consequences such as the state-dependent frequency shifts of neighboring qubits shown in Fig. 4.11, can be directly observed in the IBM Q processors (see Section 7.1).

Chapter 5

Optimizing pulses for quantum gates

For gate-based quantum computers, a quantum gate is implemented by a certain external action on the system. A natural way of interacting with a superconducting system is to apply an electromagnetic pulse. In the transmon systems considered in this thesis, such a pulse is a microwave voltage pulse, applied to each qubit through its respective transmission line (see Section 3.1.4). A voltage pulse is modeled by the external time-dependent functions $n_{gi}(t)$ in the Hamiltonian given by Eqs. (3.9a)–(3.9f). We consider a generic sum of microwave voltage pulses

$$n_{gi}(t) = \sum_j \Omega_{ij}(t) \cos(2\pi f_{ij} t - \gamma_{ij}), \qquad (5.1)$$

where $\Omega_{ij}(t)$ is the envelope of pulse j on qubit i, f_{ij} is the corresponding drive frequency, and γ_{ij} is an offset phase. The generic expression for microwave pulses given by Eq. (5.1) is motivated by the form of the microwave signals used in typical experiments to implement quantum gates (see [McK2017] for more information).

To implement a particular quantum gate, the pulses $n_{gi}(t)$ given by Eq. (5.1) must be chosen such that the time evolution of the full system corresponds to the unitary transformation representing the desired quantum gate. In other words, if $U : \mathcal{H}_{2^n} \to \mathcal{H}_{2^n}$ is the unitary operator corresponding to the desired quantum gate (see Section 2.2), the functions $n_{gi}(t)$ for $0 \le t < T$ need to be chosen such that the time-evolution operator of the full system $\mathcal{U}(T, 0)$ (see Eq. (3.19)) implements U (potentially in a certain rotating frame, meaning that the columns of $\mathcal{U}(T, 0)$ are transformed according to Eq. (3.41)).

The challenge, however, is that the time-evolution operator $\mathcal{U}(T, 0)$ acts on the much larger Hilbert space $\mathcal{H}$ of all transmons and resonators given by Eq. (3.17). Therefore, the operators U and $\mathcal{U}(T, 0)$ can only be equal after projecting $\mathcal{U}(T, 0)$ on the smaller computational subspace $\mathcal{H}_{2^n}$, yielding

$$M = P_{\mathcal{H}_{2^n}} \mathcal{U}(T, 0) P_{\mathcal{H}_{2^n}}, \qquad (5.2)$$

where $P_{\mathcal{H}_{2^n}}$ denotes the projection operator defined in Eq. (3.35). In almost all practical cases, the projected matrix M is not unitary anymore, so it is, strictly speaking, impossible to make M equal to the desired quantum gate U. Intuitively, this means that non-computational states affect the time evolution of the total system, a particular problem for transmon qubits known as *leakage* [Che2016; Wil2017; Woo2018] (see also

Section 2.1.4). The best one can hope for is to find a pulse resulting in a transformation M that approximates the desired quantum gate U as closely as possible. The important question is whether such a fundamentally imperfect implementation is sufficient in practice.

Therefore, the aim of this chapter is to develop an optimization scheme for a set of pulse parameters for Eq. (5.1) to implement the closest approximation to U. The particular set of parameters depends on the kind of quantum gate to be optimized. In Sections 5.1 and 5.2, we specify this set of parameters for single- and two-qubit gates, respectively. In Section 5.3, we describe the optimization procedure and present results for some of the model systems used for this work. Section 5.4 gives an example of the compilation process to translate a quantum circuit into a sequence of pulses. Finally, we give a brief overview of alternative optimization techniques in Section 5.5. Applying the optimized pulses to actual quantum circuits and comparing their performance to experimental implementations is the topic of the following chapters.

5.1 Single-qubit pulses

Applying a pulse of the form of Eq. (5.1) has the effect that the qubit represented by transmon i is rotated around its Bloch sphere. Specifically, within the RWA, one can show [Gam2013] that a pulse of the form

$$\Omega(t)\cos(2\pi ft - \gamma), \tag{5.3}$$

on transmon i, where $f = \omega_i/2\pi$ is given by the qubit frequency, corresponds to a rotation by an angle

$$\vartheta = b_i \int_0^T \Omega(t)\, \mathrm{d}t, \tag{5.4}$$

where T is the duration of the pulse, and

$$b_i = 8E_{Ci}\left(\frac{E_{Ji}}{32E_{Ci}}\right)^{1/4} \tag{5.5}$$

is the energy scale of the dimensionless amplitudes. In other words, the area under the envelope $\Omega(t)$ determines the angle of rotation. The axis of rotation is defined by the phase γ in Eq. (5.3). In particular, $\gamma = 0$ ($\gamma = \pi/2$) corresponds to the x (y) axis. Furthermore, one can show that choosing $f \neq \omega_i/2\pi$ results in additional rotations around the z axis. See [Gam2013; Wil2016] for more information on these properties.

For the above relations, the coefficients of the state vector have to be expressed in the rotating frame (see Eq. (3.41)). Note that, although the choice of frame does not affect the result of the final measurement, it matters when we want to interpret or visualize the coefficients of the intermediate state vector $|\Psi(t)\rangle$.

As discussed in Section 2.2.2, every single-qubit gate can be expressed in terms of rotations on the Bloch sphere. We implement the particular set of single-qubit rotations

$X^{\pi/2} = R^x(\pi/2)$ (a $\pi/2$ rotation of the qubit around the x axis) and $Z^\vartheta = R^z(\vartheta)$ (an arbitrary rotation by an angle ϑ around the z axis). These gates represent the elementary building blocks of the U1, U2, and U3 gates given by Eqs. (2.15a)–(2.15c), which in turn can be used to express all standard single-qubit gates (see Tab. B.1 in Appendix B). This choice of elementary single-qubit gates is the same that was made for the IBM Q processors [Cro2017].

5.1.1 The VZ gate

In principle, the information given above is sufficient to find candidates for pulses to implement both $X^{\pi/2}$ and Z^ϑ. However, one can simplify the hardware implementation further by using the concept of a virtual Z gate (VZ gate). This concept is common practice in the transmon architecture under investigation [McK2017]. Therefore, we also implement this concept in the transmon simulator.

The VZ gate is based on the fact that the phase γ in Eq. (5.3) defines the axis of rotation in the xy plane. In other words, γ determines the frame of reference in which the qubit is defined. Therefore, whenever the next gate in a sequence of gates is Z^ϑ or U1(λ), instead of applying a pulse, we rotate our personal frame of reference. This affects the phases $\{\gamma\}$ of *all the following pulses* according to a given rule.

For single-qubit pulses, this rule corresponds to an exchange of operations according to the scheme

$$\text{pulse}^{(m)}(\gamma) \cdots \text{pulse}^{(1)}(\gamma)\, Z^\vartheta \,|\Psi\rangle = Z^\vartheta \,\text{pulse}^{(m)}(\gamma - \vartheta) \cdots \text{pulse}^{(1)}(\gamma - \vartheta)\, |\Psi\rangle. \tag{5.6}$$

The advantage of this scheme is that no time on the hardware is required to implement the family of rotations Z^ϑ. This is the reason why the VZ gate is called a *virtual* gate.

However, the downside is that during the time evolution, we have to keep track of the VZ phases γ_i for each transmon i. Furthermore, for every elementary quantum gate, we need to define how it commutes with Z^ϑ and how ϑ affects the phases of the underlying pulses. For the single-qubit pulse to implement $X^{\pi/2}$ and the more complicated two-qubit pulses, this rule is given in the following sections (see Eq. (5.10) and Fig. 5.3, respectively).

5.1.2 The GD pulse

To implement $X^{\pi/2}$ with a VZ phase parameter γ, we use a microwave pulse of the form

$$\Omega_G(t) \cos(2\pi f t - \gamma), \tag{5.7}$$

where the envelope $\Omega_G(t)$ is a Gaussian defined as

$$\Omega_{\mathrm{G}}(t) = \Omega_X \frac{\exp\left(-\frac{(t - T_X/2)^2}{2\sigma^2}\right) - \exp\left(-\frac{T_X^2}{8\sigma^2}\right)}{1 - \exp\left(-\frac{T_X^2}{8\sigma^2}\right)}. \tag{5.8}$$

Here, Ω_X is the amplitude that needs to be chosen such that the angle ϑ given by Eq. (5.4) equals $\pi/2$, T_X is the duration of the pulse (typically around 80 ns), and $\sigma = T_X/4$

characterizes the width of the Gaussian. Note that the Gaussian is shifted vertically such that $\Omega_G(0) = \Omega_G(T) = 0$.

For transmon qubits, simple Gaussian pulses like the one given in Eq. (5.7) can drive the qubit out of the computational subspace such that higher levels $|m\rangle$ with $m \geq 2$ are excited. To mitigate this effect, a technique known as DRAG has become standard [Mot2009; Cho2010; Gam2011; The2018]. There are several alternatives of implementing DRAG (see [Gam2013; Wil2016] for a comparison), but the common concept is that a shifted microwave pulse with an amplitude given by the derivative $\dot{\Omega}_G(t)$ of the Gaussian in Eq. (5.8) is added to Eq. (5.7). The prefactor of this term is the so-called DRAG coefficient β_X. We also implement this concept for the transmon simulator.

The single-qubit pulse to implement $X^{\pi/2}$ with DRAG correction is given by

$$\text{GD}^{\pi/2}(\gamma) : \quad \Omega_G(t)\cos(2\pi f t - \gamma) + \beta_X \dot{\Omega}_G(t)\cos(2\pi f t - (\gamma + \pi/2)), \qquad (5.9)$$

and is a characterized by four pulse parameters $(f, T_X, \Omega_X, \beta_X)$. These parameters are tuned in the pulse optimization procedure discussed in Section 5.3. Initial values for the optimization are either given by theory or taken from experiments. Specifically, the drive frequency f is initialized to the qubit frequency $f_i = \omega_i/2\pi$ (determined using the procedure described in Section 3.3.2); the time T_X is typically kept fixed at around 80 ns (given by the corresponding processor, see e.g. [IBM2018a]); the drive strength Ω_X is determined from Eq. (5.4); and the DRAG coefficient β_X is set to $-1/2\alpha$, where α is the anharmonicity (see the text below Eq. (3.14)). The phase γ in Eq. (5.9) is used to implement VZ gates according to the rule

$$\text{GD}^{\pi/2}(\gamma)\, Z^\vartheta\, |\Psi\rangle = Z^\vartheta\, \text{GD}^{\pi/2}(\gamma - \vartheta)\, |\Psi\rangle . \qquad (5.10)$$

By analogy with the notation used for single-qubit gates in multi-qubit systems (see Eq. (2.16)), we denote a GD pulse on qubit i by $\text{GD}_i^{\pi/2}(\gamma)$. In addition to $\text{GD}_i^{\pi/2}(\gamma)$, we also define a pulse $\text{GD}_i^{\pi}(\gamma)$ that is supposed to implement a full X_i^π rotation (i.e., a bit flip). This pulse is used mainly as a building block for the two-qubit pulses defined in the next section; a single-qubit X_i^π gate is typically implemented as $\text{U3}_i(\pi, 0, \pi)$ (see also Eq. (2.15c)), i.e., in terms of two $\text{GD}_i^{\pi/2}(\gamma)$ pulses as done for the IBM Q processors [Cro2017].

Technically, a pulse for the X^π rotation differs from the $X^{\pi/2}$ rotation in that the drive amplitude Ω_X is twice as large, since the angle of rotation given by Eq. (5.4) is directly proportional to Ω_X. However, the other parameters may also come out differently in the parameter-optimization process. In general, it is nontrivial to predict the best set of parameters $(f, T_X, \Omega_X, \beta_X)$ for a full transmon-resonator system.

5.1.3 The zero pulse

We define a particular pulse representing an undriven time evolution for a time T, denoted by zero(T). It is used as an explicit expression for an identity gate. Typically, $T = T_X$ is the time used for single-qubit gates (see above).

In the transmon computer model defined in Section 3.2, a zero pulse on qubit i corresponds to setting $n_{gi}(t) = 0$. Usually, it is not necessary to specify the zero pulse explicitly,

since $n_{gi}(t) = 0$ whenever no pulse is specified. However, it can be used to configure free time evolutions as studied in Chapter 4, or to have the simulation run freely for some time after the last actual pulse has been applied.

Note that for some quantum computing systems, an identity gate is implemented by explicit pulses with zero net effect. For instance, a second-order dynamical decoupling sequence of the form $X^\pi Y^\pi X^\pi Y^\pi$ [Kho2009] was found to improve the performance of identity gates in a recent trapped-ion qubit system [Blu2017]. However, we choose to implement the identity gate in terms of an undriven time evolution, which is also done for the IBM Q processors [IBM2016; Cro2017]. Additionally, this implementation allow us to study and understand the emerging effects in coupled transmon systems (see Sections 4.4 and 7.1, and also the gate set tomography experiments discussed in Section 6.3.2).

5.2 Two-qubit pulses

The universal two-qubit gate used for the quantum computer simulations presented in this thesis is the CNOT gate defined by Eq. (2.18). A prominent pulse to implement the CNOT gate is the cross-resonance (CR) pulse [Rig2010; Cho2011; Gro2012]. Its idea was first proposed in [Par2006], and the knowledge and methods about how to use it for transmon quantum computers have continuously improved over the past years [Cór2013; She2016b; Tak2017; Tri2019; Mag2020; Mal2020].

5.2.1 CNOT gates based on the CR effect

We consider a CNOT gate between a control qubit i_C and a target qubit i_T. The basic CR pulse is defined as a microwave pulse applied to the control qubit at the resonance frequency $f_{i_T} = \omega_{i_T}/2\pi$ of the target qubit. Since the frequencies of adjacent transmon qubits typically differ by $100 - 300\,\mathrm{MHz}$, the CR pulse is a slightly off-resonant pulse on the control qubit. For the time-dependent pulses $n_{gi}(t)$ given by Eq. (5.1), we define the basic CR pulse (denoted by CR0) as a flat-topped Gaussian microwave pulse

$$\mathrm{CR0}(\gamma): \quad \Omega_{GF}(t)\cos(2\pi f_{i_T} t - \gamma), \tag{5.11}$$

where γ is the VZ phase (cf. Section 5.1.1), and the envelope $\Omega_{GF}(t)$ of the CR pulse is a flat-topped Gaussian. The latter is formally defined as

$$\Omega_{GF}(t) = \begin{cases} \Omega_G(t) \left| \begin{pmatrix} \Omega_X \\ T_X \\ \sigma \end{pmatrix} \mapsto \begin{pmatrix} \Omega_{\mathrm{CR}} \\ 2T_{\mathrm{rise}} \\ T_{\mathrm{rise}}/3 \end{pmatrix} \right. & (0 \leq t \leq T_{\mathrm{rise}}) \\ \Omega_{\mathrm{CR}} & (T_{\mathrm{rise}} \leq t \leq T_{\mathrm{rise}} + T_{\mathrm{CR}}) \\ \Omega_G(t - T_{\mathrm{CR}}) \left| \begin{pmatrix} \Omega_X \\ T_X \\ \sigma \end{pmatrix} \mapsto \begin{pmatrix} \Omega_{\mathrm{CR}} \\ 2T_{\mathrm{rise}} \\ T_{\mathrm{rise}}/3 \end{pmatrix} \right. & (T_{\mathrm{rise}} + T_{\mathrm{CR}} \leq t \leq 2T_{\mathrm{rise}} + T_{\mathrm{CR}}) \end{cases}, \tag{5.12}$$

where Ω_{CR} is the amplitude of the CR pulse, T_{CR} is the time of the flat part in the middle of the pulse, and $T_{\mathrm{rise}} = 15\,\mathrm{ns}$ is the time of the Gaussian rise at the beginning and the Gaussian fall at the end of the pulse. The total duration of the CR pulse is thus $T_{\mathrm{tot}} = T_{\mathrm{CR}} + 30\,\mathrm{ns}$. The parameters for the Gaussian $\Omega_G(t)$ given by Eq. (5.8) have to be replaced as indicated in Eq. (5.12). The basic CR0 pulse given by Eq. (5.11) is

schematically shown in Fig. 5.1(a). It works as a building block for the pulse sequences defined below to implement the CNOT gate.

The effect of the basic CR0 pulse can be illustrated in terms of the effective two-qubit Hamiltonian

$$H_{\text{eff}} = \frac{J_{IX}}{2}\sigma_{i_T}^x + \frac{J_{ZX}}{2}\sigma_{i_C}^z\sigma_{i_T}^x + \frac{J_{ZI}}{2}\sigma_{i_C}^z, \tag{5.13}$$

where i_C is the index of the control qubit (to which the pulse is applied), i_T is the index of the target qubit (which determines the frequency of the pulse), and the coefficients J_{IX}, J_{ZX}, and J_{ZI} represent the strength of the two-qubit terms. Both J_{IX} and J_{ZX} are approximately proportional to Ω_{CR} (see below). After an application of the CR0 pulse, the implemented transformation is approximately given by

$$\exp(-iH_{\text{eff}}T_{\text{tot}}) \propto \begin{pmatrix} \overset{\text{Control in }|0\rangle}{R^x((J_{IX}+J_{ZX})T_{\text{tot}})} & 0 \\ 0 & \overset{\text{Control in }|1\rangle}{R^x((J_{IX}-J_{ZX})T_{\text{tot}})e^{i\eta}} \end{pmatrix}, \tag{5.14}$$

where $R^x(\vartheta)$ is the matrix of the single-qubit x rotation defined in Eq. (2.14a) and $e^{i\eta}$ is a phase factor. The operation expressed by Eq. (5.14) is an x rotation of the target qubit by an angle $\vartheta = (J_{IX}+J_{ZX})T_{\text{tot}}$ ($\vartheta = (J_{IX}-J_{ZX})T_{\text{tot}}$) if the control qubit is in state $|0\rangle$ ($|1\rangle$). See Fig. 5.2(a) and (b) for a Bloch-sphere visualization of the time evolution, computed from the simulation of the two-transmon system defined in Section 3.4.3 under the application of a CR0 pulse. In each plot, the number of arrows per time plotted is constant, so larger spacings between successive arrows represent a faster relative time evolution. Furthermore, the vectors are not renormalized, so the fact that the magnitude of all vectors is almost one means that the states evolve as almost unentangled states if the initial state is a computational basis state.

Note that the effective Hamiltonian given by Eq. (5.13) does not accurately model the intermediate time evolution during the application of the pulse including the finite rise and fall time. For instance, the control qubit does not stand still during the evolution, as indicated by the red arrows in Fig. 5.2. However, the result of the pulse application is described well enough by Eq. (5.14): an x rotation of the target qubit, for which the angle depends on the state of the control qubit. Since this operation maps the target qubit to different states depending on whether the control qubit is in state $|0\rangle$ or $|1\rangle$, it is an entangling operation and can be used to assemble a CNOT gate.

We consider three candidates, denoted by CR1, CR2, and CR4, to implement a CNOT gate based on the elementary CR0 pulse. They are schematically plotted in Fig. 5.1(b)–(d). For each pulse, the time evolution of a two-transmon system is shown in Fig. 5.2(c)–(h). While the target qubit (blue) can be easily seen to end up in the proper state, the time evolution of the control qubit (red) can be much more complicated.

The CR1 pulse

As a first candidate to implement the CNOT gate, we consider the CR1 pulse shown in Fig. 5.1(b). This pulse sequence was introduced in [Wil2017] and a demonstration of how to implement it on an IBM Q processor by means of *Qiskit Pulse* is given in [Ale2020].
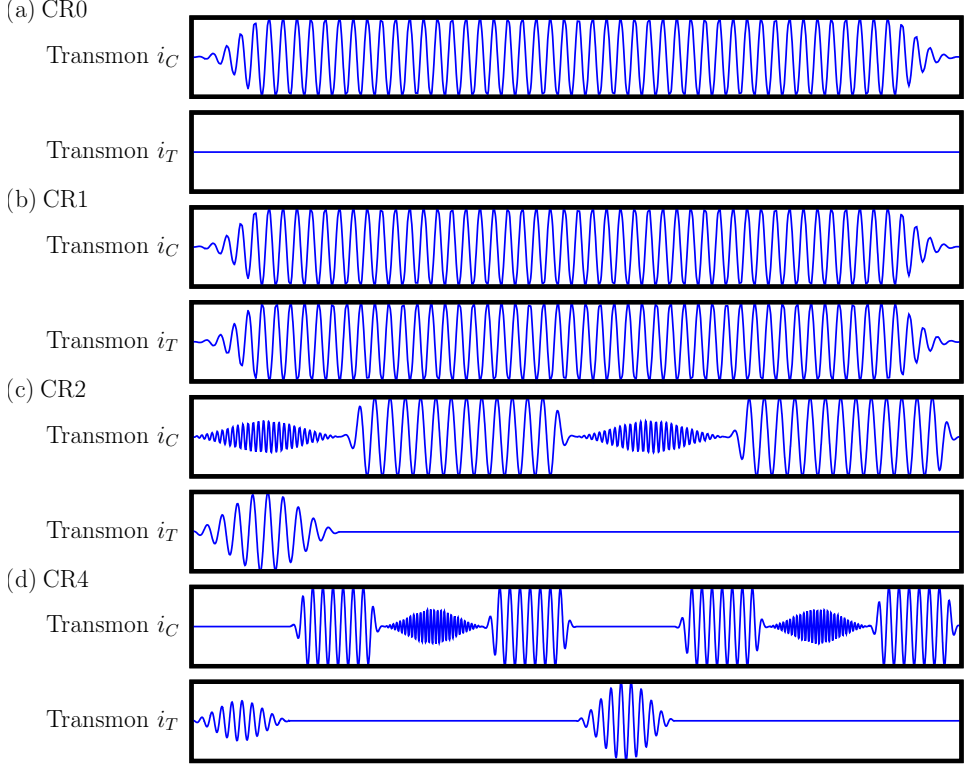
Figure 5.1: CR pulse sequences as a function of time, corresponding to $n_{gi}(t)$ given by Eq. (5.1) where $i = i_C$ ($i = i_T$) denotes the control (target) qubit, (a) basic CR0 pulse on the control qubit at the frequency of the target qubit (see Eq. (5.11)); (b)–(d) three different realizations CR1, CR2, and CR4 of a CNOT gate using combinations of CR pulses (see Eq. (5.11)) and single-qubit GD pulses (see Eq. (5.9)). Gaussians represent the GD pulses and implement $X^{\pi/2}$ and X^π rotations. Flat-topped Gaussians represent the CR pulses. The CR1 gate consists only of flat-topped Gaussian pulses at the target frequency. The CR2 gate is an echoed CR gate containing two additional X^π pulses on the control qubit and one $X^{\pi/2}$ pulse on the target qubit. The CR4 gate is a four-pulse echoed CR gate that contains an additional X^π pulse on the target qubit. See Fig. 5.3 for the full pulse specifications.

(a) CR0 $|00\rangle$

(b) CR0 $|10\rangle$

(c) CR1 $|00\rangle$

(d) CR1 $|10\rangle$

(e) CR2 $|00\rangle$

(f) CR2 $|10\rangle$

(g) CR4 $|00\rangle$

(h) CR4 $|10\rangle$
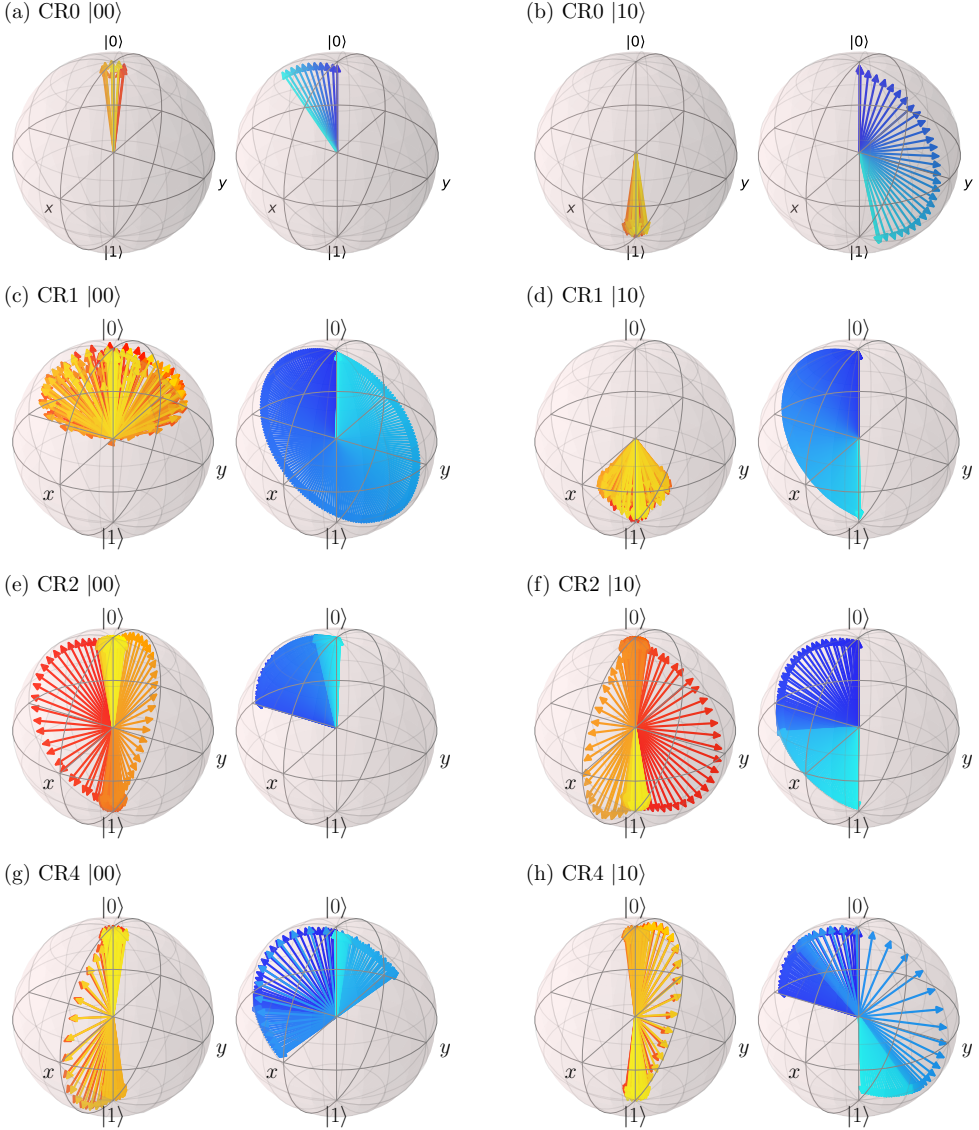
Figure 5.2: Bloch-sphere representation of the time evolutions of two transmons under the application of the CR pulses shown in Fig. 5.1; (a) and (b) basic CR0 pulse with $\Omega_{\mathrm{CR}} = 0.01$ and $T_{\mathrm{CR}} = 270\,\mathrm{ns}$; (c)–(h) CR1, CR2, and CR4 pulse implementing a CNOT gate. The time t is encoded in the color of the arrows, i.e., from red to yellow (blue to cyan) for the control (target) qubit. The pulse parameters for (c)–(h) result from the optimization procedure described in Section 5.3 and are listed in Appendix F. The model parameters of the simulated transmon system are given in Tab. 3.4. The Bloch vectors are computed by Eq. (3.38) (not renormalized) in a frame rotating at the frequencies given in Tab. F.1. The data has been visualized with `QuTiP` [Joh2012; Joh2013].
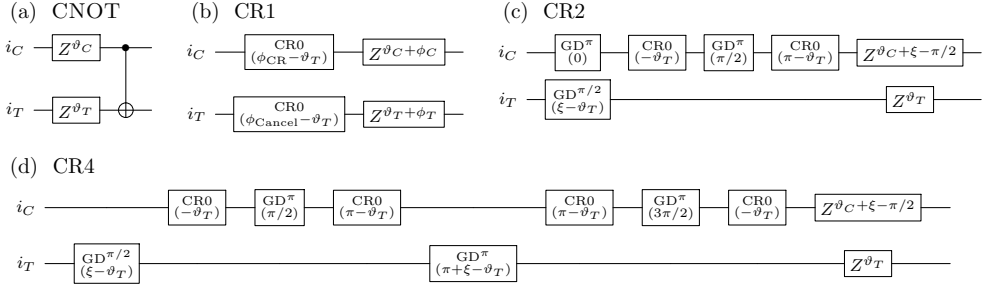
Figure 5.3: Specifications of the pulse sequences shown in Fig. 5.1(b)–(d) to implement the CNOT gate; (a) elementary CNOT gate with preceding local z rotations; (b)–(d) CR1, CR2, and CR4 schemes to implement the CNOT gate with local z rotations, where the z gates have been propagated to the end of the sequence to make the pulse scheme compatible with the VZ gate. Note that the VZ phases of the control qubit are unaffected by ϑ_C because of Eq. (5.21). The single-qubit GD pulse is defined in Eq. (5.9) and the basic CR0 pulse is defined in Eq. (5.11).

The CR1 pulse consists of two simultaneous elementary CR0 pulses: one on the control qubit and one on the target qubit. Both pulses have the frequency f_{i_T} of the target qubit. The amplitude of the second pulse is denoted by Ω_{Cancel} and is typically smaller than the control pulse amplitude Ω_{CR}. Since the second pulse on the target qubit is a resonant driving, its effect is a simple x rotation whose angle is determined by Eq. (5.4). Therefore, Ω_{Cancel} directly changes the strength J_{IX} of the unconditional x rotation given in Eq. (5.13). This behavior is also confirmed by the simulation results discussed in the next section (see Fig. 5.4(c) and (d)).

We can use this fact in the following way: The elementary CR0 pulse with amplitude Ω_{CR} on the control qubit sets the magnitude of J_{IX} and J_{ZX}. By changing Ω_{Cancel} on the target qubit, we can adjust J_{IX} independently such that

$$(J_{IX} + J_{ZX})T_{\mathrm{tot}} \bmod 2\pi = 0, \tag{5.15a}$$

$$(J_{IX} - J_{ZX})T_{\mathrm{tot}} \bmod 2\pi = \pi, \tag{5.15b}$$

and the implemented transformation in Eq. (5.14) becomes proportional to a CNOT gate (up to local z rotations). The effect of Eq. (5.15a) can be seen in Fig. 5.2(c), where the target qubit undergoes a full 2π rotation and remains effectively unchanged. Similarly, the effect of Eq. (5.15b) is shown in Fig. 5.2(d), where the target qubit undergoes a π rotation such that it ends up in the $|1\rangle$ state.

There may still be spurious local z rotations. One is represented by the phase factor $e^{i\eta}$ in Eq. (5.14), which stems from the coefficient J_{ZI}. Because of such phase errors, the

93

operation on the computational subspace actually amounts to

$$\begin{pmatrix} e^{i\chi_1} & & & \\ & e^{i\chi_2} & & \\ & & e^{i\chi_3} & \\ & & & e^{i\chi_4} \end{pmatrix}. \tag{5.16}$$

However, the phase errors χ_i can be corrected with phase shifts ϕ_{CR} and ϕ_{Cancel} in the basic CR0 pulses defined in Eq. (5.11), followed by local z rotations $Z_{i_C}^{\phi_C} \otimes Z_{i_T}^{\phi_T}$.

In summary, the $\mathrm{CNOT}_{i_C i_T}$ implementation using the CR1 pulse is

$$\mathrm{CNOT}_{i_C i_T} = (Z_{i_C}^{\phi_C} \otimes Z_{i_T}^{\phi_T})(\mathrm{CR0}_{i_C}(\phi_{\mathrm{CR}}) \otimes \mathrm{CR0}_{i_T}(\phi_{\mathrm{Cancel}})). \tag{5.17}$$

The CR1 gate depends on 7 parameters $(f_{i_T}, T_{\mathrm{CR}}, \Omega_{\mathrm{CR}}, \Omega_{\mathrm{Cancel}}, \phi_{\mathrm{CR}}, \phi_{\mathrm{Cancel}}, \phi_C, \phi_T)$ to be optimized in the optimization procedure. In principle, one phase parameter could be eliminated; however, we found that keeping an additional phase parameter helps in mitigating phase errors caused by other components of the system. The full pulse sequence including VZ phases is specified in Fig. 5.3(b).

The CR2 pulse

The CR2 pulse implements the CNOT gate using a two-pulse echo scheme. The idea has been analyzed in [Cór2013] and further specified in the supplementary material of [Tak2017]. It is also currently used for the processors on the IBM Q Experience [IBM2016] (see the IBM Q backend specifications). The sequence of pulses is schematically shown in Fig. 5.1(c).

The idea of the echo scheme is that the CR0 pulse is split into two parts with opposite amplitudes Ω_{CR}. Both parts are defined to have the duration $T_{\mathrm{CR}} + 30\,\mathrm{ns}$. Between these two parts, the control qubit is inverted using a GD^π pulse. In this way, the J_{IX} component in Eq. (5.13) is canceled, whereas the desired J_{ZX} component is doubled. Besides canceling the J_{IX} component, this scheme also addresses the J_{ZI} component in Eq. (5.13) and the residual longitudinal interaction of the form given in Eq. (4.28).

The amplitude Ω_{CR} and the time T_{CR} of each of the two CR0 parts are chosen such that

$$J_{ZX} T_{\mathrm{tot}} = \frac{\pi}{4}. \tag{5.18}$$

This means that the combined effect of the two CR0 pulses and the intermediate GD^π pulse is a $\pi/2$ rotation of the target qubit in one direction if the control qubit is in state $|0\rangle$, and in the other direction if the control qubit is in state $|1\rangle$. The CNOT gate is completed with an additional GD^π pulse on the control qubit (whose VZ phase can be used to take care of J_{ZI}) and a $\mathrm{GD}^{\pi/2}$ pulse on the target qubit. These pulses have been moved to the beginning of the pulse sequence shown in Fig. 5.1(c).

The time evolution generated by the CR2 pulse sequence is shown in Fig. 5.2(e) and (f): First, the target qubit is rotated by $\pi/2$ to the negative y axis. Then the first CR0 pulse rotates it either back towards $|0\rangle$ (Fig. 5.2(e)) or further on towards $|1\rangle$ (Fig. 5.2(f)).

The second CR0 pulse finishes this rotation. Note also that the angular velocity of the target qubit is different for both CR0 parts (indicated by the abrupt transition from blue to cyan). This is due to the different magnitudes of $|J_{IX} \pm J_{ZX}|$.

In practice, we use a slightly more sophisticated relation to obtain good initial values from Eq. (5.18) for the time and the amplitude of the CR pulses. The idea is to account for the finite rise and fall of $\Omega_{GF}(t)$ given by Eq. (5.12) by integrating over time. The area under the envelope then yields the angle of rotation, analogous to the single-qubit result in Eq. (5.4). The relation reads

$$\int\limits_0^{T_{\text{CR}}+30\,\text{ns}} J_{ZX}^{\text{theory}}|_{\Omega_{\text{CR}}\mapsto\Omega_{GF}(t)} \, \mathrm{d}t = \text{sign}(J_{ZX}^{\text{theory}})\frac{\pi}{4}, \tag{5.19}$$

where J_{ZX}^{theory} is given by the first term of the perturbative result given in Eq. (5.22b) below. Typically, we set $\Omega_{\text{CR}} = 0.01$ and solve Eq. (5.19) for T_{CR} to get initial values for the optimization.

The set of parameters specifying the CR2 pulse is $(f_{i_C}, f_{i_T}, T_{\text{CR}}, \Omega_{\text{CR}})$. Implicitly, the CR2 pulse also depends on the parameters $(T_X^\pi, \Omega_X^\pi, \beta_X^\pi)_C$ of the GD$^\pi$ pulses on the control qubit, as well as the parameters $(T_X^{\pi/2}, \Omega_X^{\pi/2}, \beta_X^{\pi/2})_T$ of the GD$^{\pi/2}$ pulse on the target qubit. Additionally, a phase shift $\xi = \pi$ is required if $J_{ZX} < 0$ for $\Omega_{\text{CR}} > 0$. This can happen when $\omega_{i_C} < \omega_{i_T}$ or $\omega_{i_C} > \omega_{i_T} + |\alpha_{i_C}|$ (see Eq. (5.22b) or Fig. 5.12 in [Wil2016] for more information). The full pulse sequence is specified in Fig. 5.3(c) and has a total duration of $2(T_{\text{CR}} + 30\,\text{ns}) + (T_X^\pi)_C + \max\{(T_X^\pi)_C, (T_X^{\pi/2})_T\}$.

The CR4 pulse

The CR4 pulse splits the basic CR0 pulse into four parts and contains another set of GD pulses to echo out additional phase errors. It has been used in [Tak2016] and further refined in [Tak2017] (see the corresponding supplementary material). The full pulse is sketched in Fig. 5.1(d).

The time evolution of a two-transmon system under the application of a CR4 pulse is shown in Fig. 5.2(g) and (h). We see that, as in CR2, the target qubit undergoes an initial $\pi/2$ rotation followed by two CR0 pulses. Then, a π pulse moves it to the opposite side of the Bloch sphere. Another set of two CR0 pulses finally rotates it back to $|0\rangle$ if the control qubit is in state $|0\rangle$ (see Fig. 5.2(g)), or to $|1\rangle$ if the control qubit is in state $|1\rangle$ (see Fig. 5.2(h)).

Similar to the CR2 pulse, a CR4 pulse is determined by the amplitude Ω_{CR} and the time T_{CR} for each of the four CR0 parts. Implicitly, the CR4 pulse also depends on the parameters of all intermediate GD pulses. The full pulse sequence is specified in Fig. 5.3(d).

Implementation of the VZ gate in CR pulses

All CNOT pulses need to be compatible with the VZ gate introduced in Section 5.1.1. This means that we have to define how the pulses commute with Z gates in the spirit of

Eq. (5.6). Since the CRn pulses are two-qubit pulses, the scheme has to take into account the two phases ϑ_C and ϑ_T of the control and the target qubit, respectively:

$$\text{CNOT}_{i_C i_T} \left(Z_{i_C}^{\vartheta_C} \otimes Z_{i_T}^{\vartheta_T} \right) |\Psi\rangle = \left(Z_{i_C}^{\vartheta_C'} \otimes Z_{i_T}^{\vartheta_T'} \right) \text{CR}n(\{\gamma'\}) |\Psi\rangle. \tag{5.20}$$

The VZ phases $\{\gamma'\}$ of the intermediate pulses and the resulting phases ϑ_C' and ϑ_T' are summarized in Fig. 5.3(b)–(d).

Note that only the target phase ϑ_T has an influence on the VZ phases of the control qubit. The reason for this is that

$$\text{CNOT}_{i_C i_T} Z_{i_C}^{\vartheta_C} = Z_{i_C}^{\vartheta_C} \text{CNOT}_{i_C i_T}, \tag{5.21}$$

meaning that the CNOT gate commutes with z rotations on the control qubit.

5.2.2 Analysis of IX and ZX interactions

The effective interaction strengths J_{IX} and J_{ZX} can be both estimated analytically or extracted from simulations (such as the one shown in Fig. 5.2(a) and (b)). The procedure for the simulation emulates the procedure used in experiments [She2016b]. In this section, we explore both routes and systematically compare the results.

Analytic expressions for J_{IX} and J_{ZX} can be derived perturbatively. See [Mag2020] for an extensive perturbative calculation, yielding analytic expressions up to third order in the drive strength Ω_{CR} for the coefficients J_{IX} and J_{ZX}. The results are

$$J_{IX}^{\text{theory}} = -\frac{J_{\text{xch}}}{\alpha_{i_C} + \Delta} b_{i_C} \Omega_{\text{CR}} + \frac{J_{\text{xch}} \alpha_{i_C} \Delta}{(\alpha_{i_C} + \Delta)^3 (\alpha_{i_C} + 2\Delta)(3\alpha_{i_C} + 2\Delta)} (b_{i_C} \Omega_{\text{CR}})^3, \tag{5.22a}$$

$$J_{ZX}^{\text{theory}} = -\frac{J_{\text{xch}} \alpha_{i_C}}{\Delta(\alpha_{i_C} + \Delta)} b_{i_C} \Omega_{\text{CR}} + \frac{J_{\text{xch}} \alpha_{i_C}^2 (3\alpha_{i_C}^3 + 11\alpha_{i_C}^2 \Delta + 15\alpha_{i_C}\Delta^2 + 9\Delta^3)}{2\Delta^3 (\alpha_{i_C} + \Delta)^3 (\alpha_{i_C} + 2\Delta)(3\alpha_{i_C} + 2\Delta)} (b_{i_C} \Omega_{\text{CR}})^3, \tag{5.22b}$$

where $\Delta = \omega_{i_C} - \omega_{i_T}$ is the difference between the qubit frequencies, α_{i_C} is the anharmonicity of the control qubit, b_{i_C} is the conversion factor between energies and dimensionless amplitudes (see Eq. (5.5)), and J_{xch} is the effective transmon-exchange coupling. The latter can be approximated as

$$J_{\text{xch}} = \frac{g_{i_C} g_{i_T} (\omega_{i_C} + \omega_{i_T} - 2\Omega)}{2(\omega_{i_C} - \Omega)(\omega_{i_T} - \Omega)}, \tag{5.23}$$

where g_{i_C} and g_{i_T} are the rescaled transmon-resonator couplings given by Eq. (4.11), and Ω is the frequency of the resonator. Both the linear terms and the cubic corrections in Eqs. (5.22a) and (5.22b) are shown as dotted and dashed lines, respectively, in Fig. 5.4(a) and (b). We remark that Eqs. (5.22a) and (5.22b) can also be obtained as a special case in an extensive theoretical study using the energy-basis representation of a transmon [Mal2020].

From the simulations, we extract J_{IX} and J_{ZX} by measuring the oscillations of the target qubit conditional on the control qubit being in state $|0\rangle$ and state $|1\rangle$. This means that for
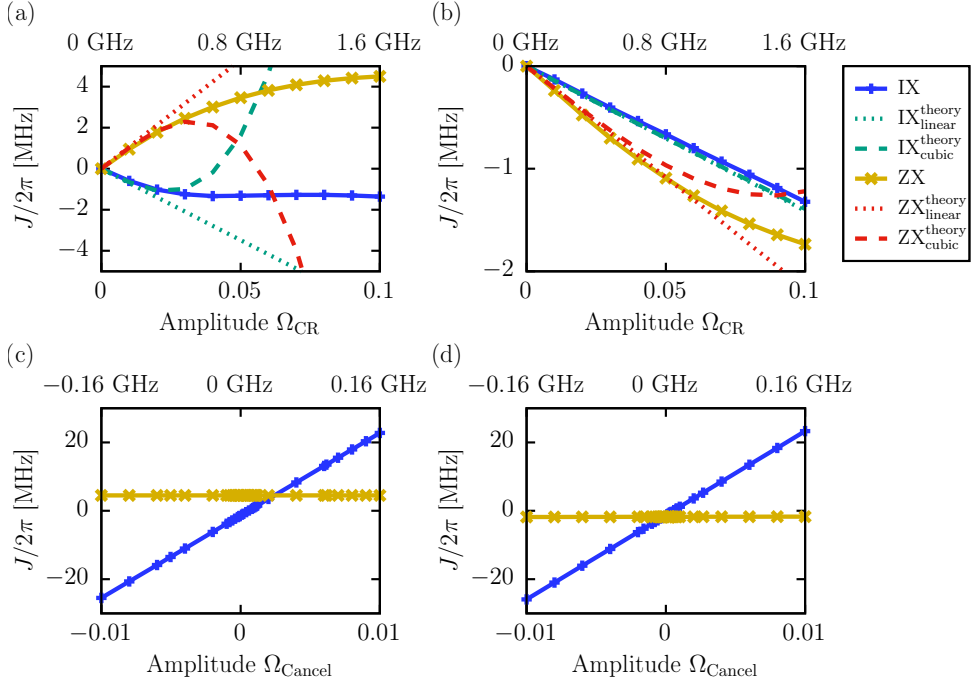
Figure 5.4: Interaction strengths J_{IX} and J_{ZX} (see Eq. (5.13)) as a function of the CR drive amplitudes Ω_{CR} and Ω_{Cancel}. (a) and (b) represent applications of a CR0 pulse (see Fig. 5.1(a)) with amplitude Ω_{CR} on the control qubit; (c) and (d) represent applications of a CR1 pulse (see Fig. 5.1(b)) with $\Omega_{CR} = 0.1$ for the control qubit fixed and Ω_{Cancel} for the target qubit variable. Each CR pulse has a total duration of $T_{CR} + 30\,\text{ns} = 500\,\text{ns}$. For (a) and (c), the control qubit is $i_C = 0$; for (b) and (d), the control qubit is $i_C = 1$. The dimensionless amplitudes can be converted to the strength of the drive by multiplying them with the conversion factor given in Eq. (5.5) (shown on top of the figures). Each point in the figures results from two simulations of the system defined in Section 3.4.3: one where the control qubit is in state $|0\rangle$ and one where the control qubit is in state $|1\rangle$. The theory predictions given by Eqs. (5.22a) and (5.22b) are shown as dotted (dashed) lines for the linear (cubic) approximations.

each amplitude Ω_{CR}, we simulate two time evolutions for both states of the control qubit. The resulting evolutions of the target qubit are described by Eq. (5.14) and visualized in Fig. 5.2(a) and (b). The procedure to obtain J_{IX} and J_{ZX} from the data is described in [She2016b] (see also Fig. 5.12 in [Wil2016]).

As shown in Fig. 5.4(a) and (b), the linear terms in Eqs. (5.22a) and (5.22b) correctly describe the regime of weak driving $\Omega_{\mathrm{CR}} \lesssim 0.3$. Although the cubic correction to the theoretical results properly captures the sign of the curvature, it diverges quickly from the numerical results. The fact that sometimes the numerical result is not exactly between the linear and cubic theory predictions for $0 \leq \Omega_{\mathrm{CR}} \leq 0.05$ may be due to the approximation made for the exchange coupling in Eq. (5.23).

For fixed $\Omega_{\mathrm{CR}} = 0.01$, we additionally apply a CR0 pulse at the target frequency on the target qubit. Technically, this combination of pulses corresponds to the CR1 pulse scheme sketched in Fig. 5.1(b). The amplitude Ω_{Cancel} of the second pulse is varied between -0.01 and 0.01. The same analysis as before is used to obtain the coefficients J_{IX} and J_{ZX}. The result is plotted in Fig. 5.4(c) and (d). As can be seen, the additional pulse on the target qubit linearly displaces J_{IX} and does not affect J_{ZX}. This property has been used to satisfy the CR1 pulse conditions given by Eqs. (5.15a) and (5.15b).

5.3 Optimization of pulse parameters

The functionality discussed in this section is part of the `optimizer` module of the software toolkit developed for this thesis (cf. Section 3.3). Its task is to optimize a set of pulse parameters $\vec{x}$ to implement a quantum gate U. This is done by first generating the appropriate pulse information for the time-dependent functions $n_{gi}(t)$ in Eqs. (3.9a)–(3.9f), and then invoking `solver` with this pulse information for different initial states from the computational basis. From the results, `optimizer` infers new pulse parameters $\vec{x}'$ and invokes `solver` again. Eventually, the procedure converges to a set of pulse parameters that can then be used by `compiler` to translate arbitrary quantum circuits into pulse information. In what follows, we outline the procedure to determine the pulse parameters for a particular quantum gate U.

Given a certain voltage pulse $n_{gi}(t)$ of the form of Eq. (5.1), the goal is to implement a quantum gate U on the computational subspace. We denote the actual transformation of the computational subspace after the application of the pulse by the matrix M defined in Eq. (5.2).

The matrix M depends on the set of parameters $\vec{x}$ defining the particular pulse. For the single- and two-qubit pulses of interest, these parameters have been specified in the previous sections. They typically consist of times, amplitudes, and phases. The goal is to optimize the parameters $\vec{x}$, starting from some initial values suggested by theory, with the objective to make M as close as possible to U.

Quantitatively, we measure closeness between M and U in terms of the matrix distance

$$\Delta(M, U) = \|M - zU\|_F^2, \tag{5.24}$$

which is induced by the Frobenius norm $\|A\|_F^2 = \sum_{ij}|A_{ij}|^2$. The phase factor $z = \pm\sqrt{\mathrm{Tr}(MU^\dagger)/\mathrm{Tr}(MU^\dagger)^*}$ is chosen such that the difference due to the global phases of

M and U is minimal, since quantum gates are considered equivalent if they only differ by a global phase (the two candidates for z can be derived by minimizing Eq. (5.24) w.r.t. $z = e^{i\zeta}$ for $\zeta \in \mathbb{R}$). In principle, one could use other, more sophisticated quantities to measure closeness between quantum gates. Obvious examples include the common gate metrics studied in Section 6.1, e.g., the fidelity or the diamond distance. However, we found that for practical purposes, the choice of the distance function does not significantly change the quality of the resulting quantum gates. Furthermore, Eq. (5.24) is numerically well suited for optimization and yields a reasonably fast convergence.

We construct M by initializing the system in each of the computational basis states at $t = 0$ and simulating its time evolution $|\Psi(t)\rangle$ under the application of the pulse for $0 \leq t \leq T$. Each final state vector $|\Psi(T)\rangle$ is transformed to the rotating frame (see Eq. (3.41)) and projected onto the computational subspace to obtain the columns of M.

The size of the matrix M is, in principle, determined by the computational subspace on which the quantum gate U shall be implemented. Specifically, this means that M and U are 2×2 complex matrices for single-qubit gates, and 4×4 complex matrices for two-qubit gates. However, if numerically feasible, we sometimes optimize M on the whole computational subspace, as done for the transmon-resonator system studied in [Wil2017].

For a set of pulse parameters $\vec{x}$, the evaluation of the objective function $\Delta(M, U)$ given by Eq. (5.24) is a complicated procedure that involves several simulations of the time evolutions of a joint transmon-resonator system. It is therefore nontrivial to find suitable gradients of $\Delta(M, U)$ with respect to $\vec{x}$. Fortunately, there is a multidimensional, gradient-free algorithm well suited for the optimization of a few parameters in the case where the most complicated step is the evaluation of the objective function: the Nelder–Mead algorithm [Nel1965; Pre2007].

5.3.1 The Nelder–Mead algorithm

Many minimization algorithms in multiple dimensions are based on the evaluation or estimation of gradients of the objective function. This either requires analytic expressions for the function's reaction to changes in the parameters, or repeated function evaluations to trace changes in the function values back to changes in the parameters. Furthermore, typical multidimensional minimization algorithms base their computational strategy on well-known minimization algorithms in a single dimension. Popular examples of these are quasi-Newton methods such as BFGS [Pre2007] or L-BFGS-B [Zhu1997; Mor2011].

The Nelder–Mead algorithm (also known as the downhill simplex method) is based on a completely different approach that does not require the evaluation of gradients. It is an entirely self-contained, direct search method based on geometrical heuristics. In principle, it can also be applied to constrained or discrete optimization problems [Lue2004; Aud2018]. Due to its conciseness, it is often used as a first step to produce reasonable results and may become the method of choice if only a few parameters need to be optimized and the evaluation of the objective function is rather complicated. This is the case for the present work, where less than 10 pulse parameters need to be optimized but the evaluation of the objective function may take several minutes on a supercomputer (e.g. for systems with more than 10 transmons or resonators, cf. Fig. 4.3).
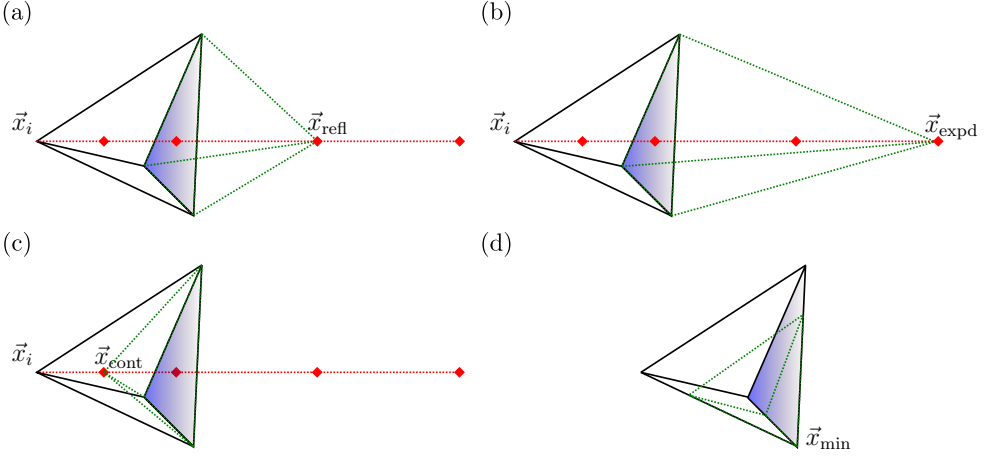
Figure 5.5: Illustration of the four central operations of the Nelder–Mead algorithm in $N = 3$ dimensions. The simplex is defined by $N + 1 = 4$ points. Shown are the basic Nelder–Mead operations which (a) reflect, (b) expand, (c) contract, or (d) shrink the simplex. The solid, black simplex represents the initial simplex. The dashed, green simplex represents the simplex after the corresponding operation.

To find a minimum of a function $F(\vec{x})$ in N real dimensions (i.e., $\vec{x} \in \mathbb{R}^N$), the Nelder–Mead method maintains a set of $N + 1$ points $S = (\vec{x}_0, \ldots, \vec{x}_N)$. Geometrically, S defines a simplex in N dimensions. At each step in the minimization, the "highest" point $\vec{x}_i$, for which $F_i = F(\vec{x}_i) = \max_j F(\vec{x}_j)$, is reflected along a line through the opposite face of the simplex (i.e., through the centroid of the face spanned by all the other points $\vec{x}_j$ for $j \neq i$, see Fig. 5.5(a)).

After this reflection, four things can happen:

(a) If the new point $\vec{x}_{\mathrm{refl}}$ is better than at least one $\vec{x}_j$ for $j \neq i$, but not better than the current minimum, it is taken as the new $\vec{x}_i$.

(b) If $\vec{x}_{\mathrm{refl}}$ is even better than the current minimum, it is expanded further along the line to $\vec{x}_{\mathrm{expd}}$ (see Fig. 5.5(b)). The better point of $\vec{x}_{\mathrm{refl}}$ and $\vec{x}_{\mathrm{expd}}$ is taken as the new $\vec{x}_i$.

(c) Otherwise, if $\vec{x}_{\mathrm{refl}}$ is worse than all $\vec{x}_j$ for $j \neq i$, it is contracted back along the line to $\vec{x}_{\mathrm{cont}}$ (see Fig. 5.5(c)). If this point is better than the previous maximum $\vec{x}_i$, it is taken as the new $\vec{x}_i$.

(d) Only if $\vec{x}_{\mathrm{cont}}$ is still not better than the previous maximum $\vec{x}_i$, the simplex is shrunk towards the current minimum $\vec{x}_{\mathrm{min}}$ as shown in Fig. 5.5(d) (this will be the exception).

Most of the time, the algorithm will either take the reflected point $\vec{x}_{\mathrm{refl}}$ or the contracted point $\vec{x}_{\mathrm{cont}}$ (cf. Figs. 5.6(b) and 5.7(b) below). See [Pre2007] for an example implementation or [Wil2018a] for a modular implementation of the algorithm.

As with many optimization algorithms, the choice of the initial parameters is crucial. Therefore, it is a good idea to use initial values from theory (such as those outlined in Sections 5.1 and 5.2) or from an initial scan of parameters. Each point of the simplex is then initialized to

$$\vec{x}_0 = \vec{x}_{\text{init}}, \tag{5.25a}$$

$$\vec{x}_1 = \vec{x}_{\text{init}} + \delta\vec{x} \cdot (1\,0\,\cdots\,0\,0)^T, \tag{5.25b}$$

$$\vdots \tag{5.25c}$$

$$\vec{x}_N = \vec{x}_{\text{init}} + \delta\vec{x} \cdot (0\,0\,\cdots\,0\,1)^T, \tag{5.25d}$$

where $\vec{x}_{\text{init}}$ is the initial set of parameters, and $\delta\vec{x}$ is a set of characteristic scales for each parameter.

We terminate the optimization when the fractional range of all function values $\vec{F} = (F(\vec{x}_0), \ldots, F(\vec{x}_N))$ of the simplex becomes smaller than a certain tolerance ε [Pre2007]. Quantitatively, this means that

$$\delta = 2\frac{|\max \vec{F} - \min \vec{F}|}{|\max \vec{F}| + |\min \vec{F}| + \texttt{tiny}} < \varepsilon, \tag{5.26}$$

where $\texttt{tiny} = 10^{-10}$ is used to prevent division by zero. Typically, we choose $\varepsilon = 10^{-4}$ to make sure that the optimization does not run indefinitely because of limited resolution in $\Delta(M, U)$ due to the finite time step τ.

In general, no multidimensional minimization algorithm can guarantee to find the global optimum; but a local minimum may still produce close-to-optimal results. For the Nelder–Mead algorithm, the stability of the solutions can be tested by restarting the algorithm from the previous minimum with a new set of characteristic scales $\delta\vec{x}$ that may be chosen a factor of 10 smaller than the previous characteristic scales. This step may also help to escape from local minima that are not yet close to optimal.

5.3.2 Optimization results

We use the `optimizer` module (cf. Section 3.3) to optimize quantum gate pulses for the two-transmon system defined in Section 3.4.3, the small five-transmon system defined in Section 3.4.4, and the large five-transmon system defined in Section 3.4.5. The pulses are used for the quantum circuit simulations described in the following chapters. All resulting pulse parameters are summarized in Appendix F.

As an example, we present results for the optimization process of gate pulses for the large five-transmon system (see Section 3.4.5) used for the fault-tolerance experiments discussed in Section 7.3. We consider two sets of gate pulses that are optimized separately. For the first set, labeled *with frequency tuning* (or `withf` in the tables given in Appendix F), the drive frequency f of the single-qubit pulses defined by Eq. (5.9) is included in the set of parameters to optimize. The idea is that slightly off-resonant driving (i.e. $f \neq f_i$) can mitigate phase errors since it induces rotations about the z axis [Gam2013]. The second gate set, labeled *without frequency tuning*, keeps the frequencies f fixed at the
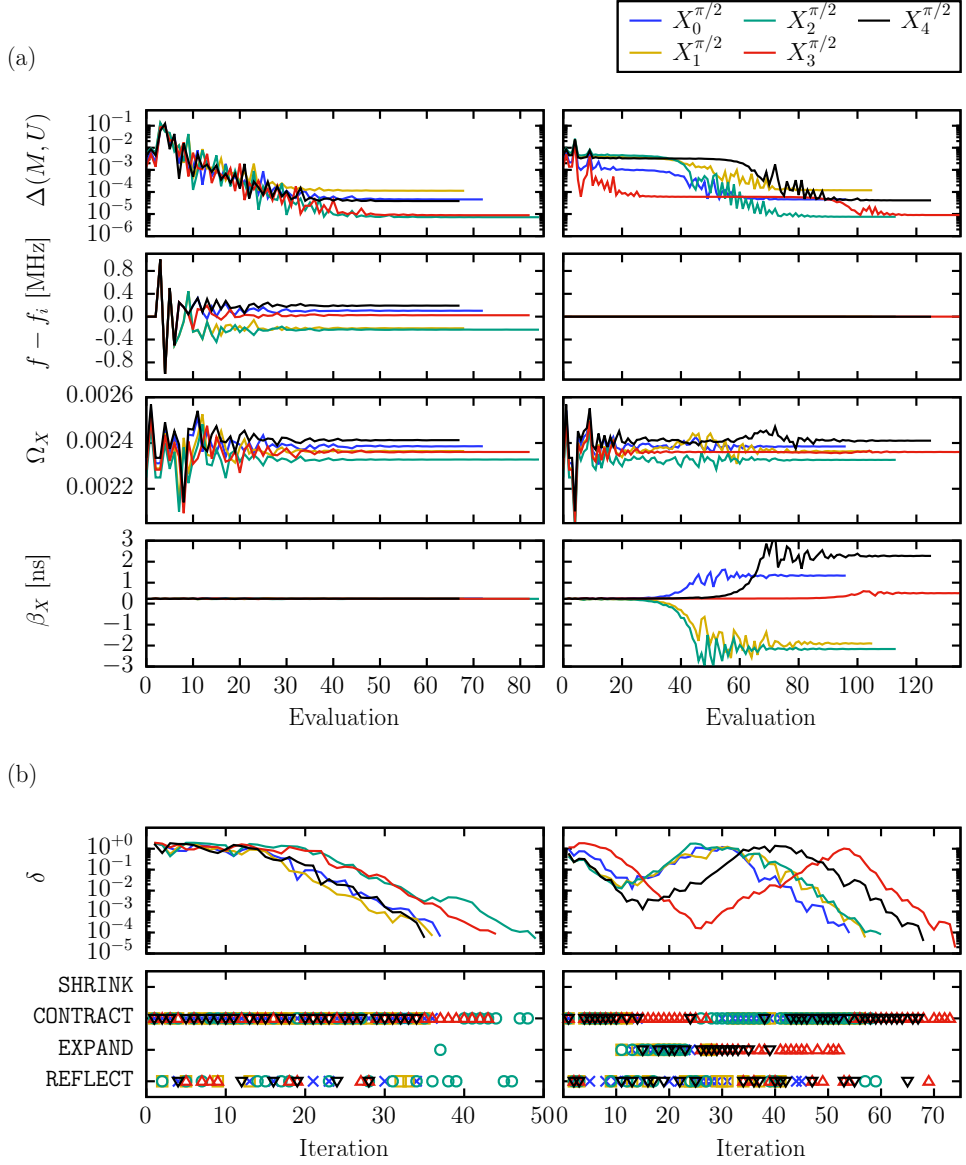
Figure 5.6: Optimization of single-qubit GD pulses for a system with five transmons and six resonators with frequency tuning (left panels) and without (right panels); (a) values of the pulse parameters specified in Eq. (5.9) at each evaluation of $\Delta(M, U)$ given by Eq. (5.24); (b) convergence criterion δ given by Eq. (5.26) and Nelder–Mead operation (see Fig. 5.5) at each iteration of the optimization. The system is sketched in Fig. 3.5 (see also Tab. 3.7 and Tab. 3.8). The resulting parameters are given in Tab. F.5. All optimizations were performed on the supercomputer JURECA [Jül2018].
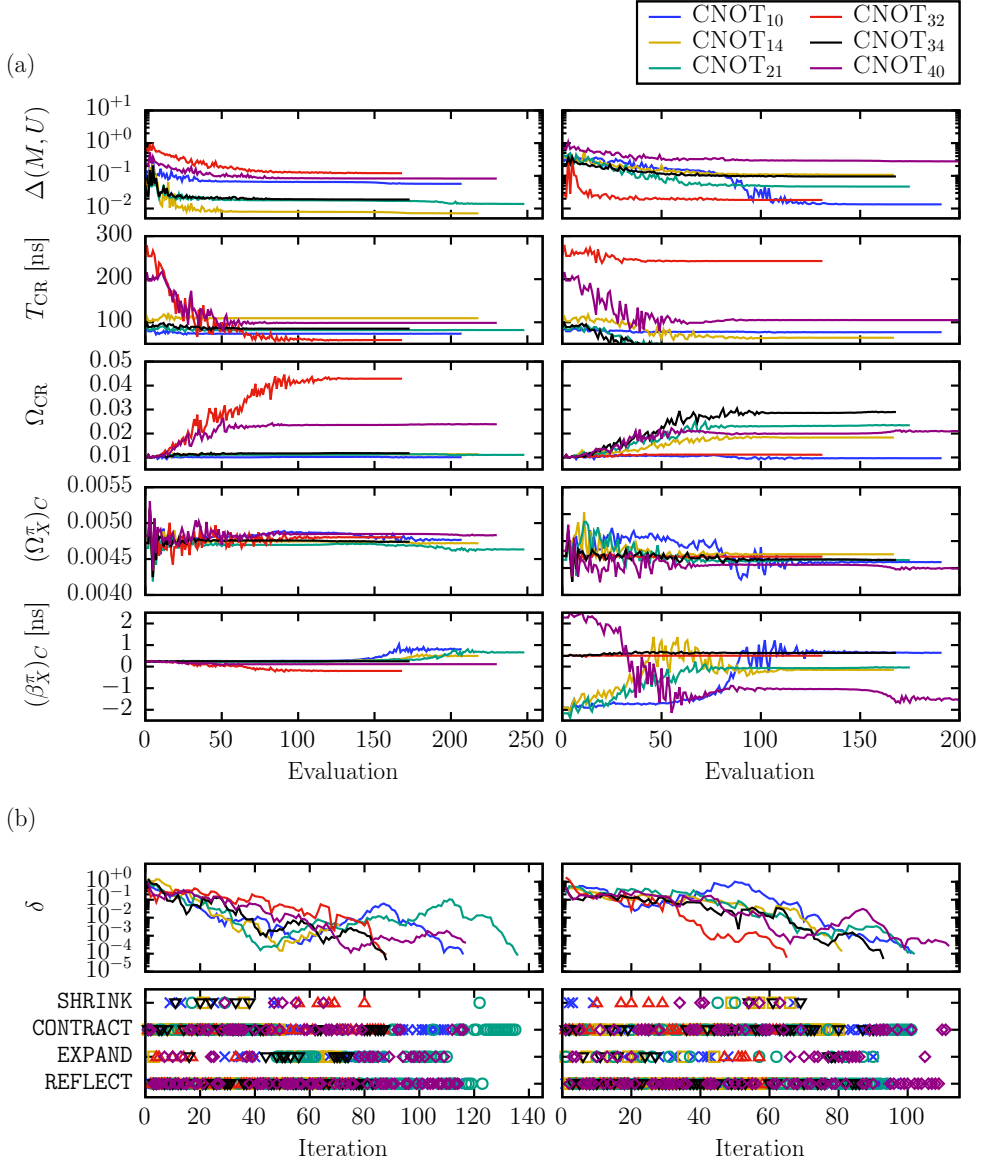
Figure 5.7: The same as in Fig. 5.6 but for the optimization of the two-qubit CR2 pulses specified in Section 5.2.1, using the GD pulses with frequency tuning (left panels) and without (right panels). The resulting parameters are given in Tab. F.6 in Appendix F. All optimizations were performed on the supercomputer JURECA [Jül2018].

qubit frequencies f_i. The optimization process of the first (second) set is shown on the left (right) panels in Figs. 5.6 and 5.7. Note that the relation between the two different scales on the x axis, i.e., the number of objective function evaluations and the number of Nelder–Mead iterations, is not one-to-one since some Nelder–Mead operations require several function evaluations.

Figure 5.6 illustrates the optimization process of the single-qubit GD pulses. As discussed in Section 5.1, the GD pulses are used to implement $X_i^{\pi/2} = R_i^x(\pi/2)$ (the target matrix is defined in Eq. (2.16)). At the first evaluation, Fig. 5.6(a) shows that the initial values mentioned below Eq. (5.9) already yield reasonable candidates, for which the matrix distance $\Delta(M, R^x(\pi/2))$ defined by Eq. (5.24) is between 10^{-2} and 10^{-3}. The following three evaluations correspond to the initialization of the simplex according to Eqs. (5.25a)–(5.25d). After that, the optimization begins its search through the parameter space. As shown in Fig. 5.6(b), the minimization with frequency tuning (left panels) converges after less than 50 iterations for each transmon i. We see that in most iterations, the simplex is contracted or reflected (cf. Fig. 5.5). Note that the DRAG coefficients β_X do not change significantly for the gate set with frequency tuning.

For the gate set without frequency tuning, i.e., when the drive frequency $f = f_i$ is fixed, this behavior is completely different (see the right panels of Fig. 5.6): When the convergence criterion δ reaches a local minimum after approximately 20–30 iterations, the DRAG coefficients β_X start to diverge significantly from their initial values. This means that the optimization without frequency tuning tries to compensate for the inability to change f by drastically changing β_X. From Eq. (5.9), we see that β_X controls a pulse whose phase is shifted by $\pi/2$. Thus it can be exploited to trigger effective rotations about the z axis. In this way, changing β_X can have the same effect as the slightly off-resonant driving included in the optimizations with frequency tuning. Note also that the simplex is frequently expanded during the drastic increase of $|\beta_X|$ (cf. Fig. 5.6(b)).

Eventually, the optimizations with and without frequency tuning converge to pulse parameters that implement the $X^{\pi/2}$ gate with a similar matrix distance. To make further statements about the quality of the different single-qubit gate sets, we study the corresponding gate metrics in Section 6.1.4.

Figure 5.7 shows the optimization of two-qubit CR2 pulses to implement CNOT gates between each pair of qubits with a corresponding resonator (cf. Fig. 3.5). The optimization process looks similar to the single-qubit case shown in Fig. 5.6 except that more parameters need to be optimized. Furthermore, the time evolution due to the CR2 pulse is more complicated (cf. Fig. 5.2(e) and (f)). As shown in Fig. 5.7(b), this is reflected by the fact that the convergence criterion δ goes a more convoluted path, on which the Nelder–Mead optimization occasionally needs to do the SHRINK operation illustrated in Fig. 5.5(d).

For some gates, such as CNOT_{34} with frequency tuning, the distance $\Delta(M, \mathrm{CNOT})$ shown in Fig. 5.7(a) (black line on the left panels) does not turn out much better than the initial point. This suggests that in this case, the theory outlined in Section 5.2 already provides good pulse parameters. However, for CNOT_{32} and CNOT_{40} with frequency tuning (red and purple lines on the left panels), the optimization finds that making the pulses much shorter (by reducing T_{CR}) and stronger (by increasing Ω_{CR}) yields better

gates. For the gates without frequency tuning shown on the right panels of Fig. 5.7(a), we again see the effect observed for the single-qubit pulses, namely that the DRAG coefficients $(\beta_X^\pi)_C$ are modified drastically by the optimization process to compensate for the fixed drive frequency.

As with the single-qubit gates, we study the quality of the resulting two-qubit gates in more detail by analyzing various gate error rates in Section 6.1.4. The metrics of the particular two-qubit gates discussed here are presented in Tab. 6.3.

The result of the `optimizer` module is a set of parameters defining each elementary quantum gate pulse for a system. This serves as input for the `compiler` module discussed in the following section.

5.4 Compiling quantum circuits

The `compiler` module is the last component of the simulation toolkit introduced in Section 3.3. It takes as input a specification of the elementary pulses (as produced by `optimizer`) and a quantum circuit to compile. The circuit can be in one of several file formats such as OpenQASM [Cro2017] or the JUQCS instruction set [DeR2019a]. The output is a sequence of pulse information describing the time-dependent functions in the model Hamiltonian given by Eqs. (3.9a)–(3.9f). This is used by the transmon simulator to perform the simulation.

In this section, we discuss an example of the compilation process. The model system is the two-transmon system defined in Section 3.4.3. We consider the compilation of the simple circuit shown in Fig. 5.8. It contains one single-qubit gate and one two-qubit gate and creates the maximally entangled state $(|00\rangle + |11\rangle)/\sqrt{2}$.

To compile the circuit into a sequence of pulses, `compiler` needs a specification of the elementary pulse parameters produced by `optimizer`. For the system of interest, the input file is given in Listing 5.1. It represents a subset of the full set of pulse parameters given in Tab. F.1 and Tab. F.2 in Appendix F. The parameters of the single-qubit GD pulses are defined by Eq. (5.9). The parameters of the two-qubit CR2 pulse are defined in Section 5.1.2 (see the text below Eq. (5.19)).

The first step to translate a circuit into a sequence of pulses is to express all single-qubit gates in terms of the elementary single-qubit U gates defined in Eqs. (2.15a)–(2.15c). In the present case, this means that we need to write the gate `-Y` as

$$\boxed{\text{-Y}} = \text{U2}(0,0) = Z^{\pi/2}\, X^{\pi/2}\, Z^{-\pi/2}. \tag{5.27}$$

The first gate acting on the qubits is $Z^{-\pi/2}$. Since Z^ϑ gates are implemented by means of the VZ gate defined in Section 5.1.1, it only affects the VZ phase of the following pulses. The next gate is $X^{\pi/2}$, which is implemented by $\text{GD}^{\pi/2}(0)$. According to Eq. (5.10), we have

$$Z^{\pi/2}\, \text{GD}^{\pi/2}(0)\, Z^{-\pi/2}\, |\Psi\rangle = Z^{\pi/2}\, Z^{-\pi/2}\, \text{GD}^{\pi/2}(\pi/2)\, |\Psi\rangle = \text{GD}^{\pi/2}(\pi/2)\, |\Psi\rangle, \tag{5.28}$$

so the first pulse resulting from the compilation is $\text{GD}^{\pi/2}(\pi/2)$. It is specified in line 2 of Listing 5.2: a pulse for the first 83 ns with phase $\pi/2$ and Gaussian envelope $\Omega_G(t)$ (see Eq. (5.8)) defined by its duration `T`, amplitude `A`, and width `sigma`.
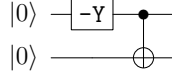
Figure 5.8: Circuit diagram for an example of the process performed by the `compiler` module. See Tab. B.1 in Appendix B for a definition of the circuit elements.

Listing 5.1 Optimized pulse parameters for the system defined in Section 3.4.3

```
1  #name     type f       TX OmegaX   betaX
2  xpih-0    GD   5.3463 83 0.002221 0.2309
3  xpih-1    GD   5.1167 83 0.002269 0.2891
4
5  #name     type fiC    fiT    TCR       OmegaCR TX OmegaXpiC betaXpiC OmegaXpihT betaXpihT
6  cnot-0-1  CR2  5.3463 5.1167 102.9746 0.01111 83 0.004444  0.2193   0.002269   0.2891
```

From the second term in the definition of the GD pulse in Eq. (5.9), we obtain line 3 of Listing 5.2: a DRAG pulse at the same time as the first pulse, characterized by a phase of π, the identifier `gaussdot`, and a much smaller amplitude `A` than the main Gaussian pulse. Since the Z gates in Eq. (5.28) cancel, no VZ phase is carried over to the following pulses.

The next gate in Fig. 5.8 is a two-qubit CNOT gate which is implemented in terms of a CR2 pulse shown in Fig. 5.1(c). As specified in Fig. 5.3(c), this requires a $\text{GD}^\pi(0)$ pulse on qubit 0 (defined in lines 4 and 5) and a $\text{GD}^{\pi/2}(0)$ pulse on qubit 1 (defined in lines 6 and 7). Since no VZ phase was carried over from the previous pulses, we have $\vartheta_C = \vartheta_T = 0$. In line 8, we find the first flat-topped Gaussian with phase 0 and a duration of $T_{\text{CR}} + 2T_{\text{rise}}$ with $T_{\text{rise}} = 15\,\text{ns}$. After another $\text{GD}^\pi(0)$ pulse on qubit 0, the CR2 pulse is finalized by the second flat-topped Gaussian defined in line 11. The echo scheme implemented by the CR2 pulse is reflected by the phase difference of π between line 8 and line 11. This is the full information needed to describe the CR2 pulse visualized in Fig. 5.2(e) and (f).

As the last line of Listing 5.2 shows, the full pulse sequence takes 514.950 ns. This is the time for which `solver` needs to compute the time evolution to simulate the circuit shown in Fig. 5.8.

Listing 5.2 Compiled pulse information for the circuit in Fig. 5.8

```
1   #i Tstart   Tend     f       phase    pulse     T       A         sigma (Trise)
2   0  0        83       5.3463 1.57080  gauss     83      0.0022212 20.75
3   0  0        83       5.3463 3.14159  gaussdot  83      0.0005129 20.75
4   0  83       166      5.3463 0        gauss     83      0.0044440 20.75
5   0  83       166      5.3463 1.57080  gaussdot  83      0.0009744 20.75
6   1  83       166      5.1167 0        gauss     83      0.0022686 20.75
7   1  83       166      5.1167 1.57080  gaussdot  83      0.0006558 20.75
8   0  166      298.975  5.1167 0        gaussflat 132.975 0.0111083 5     15
9   0  298.975  381.975  5.3463 1.57080  gauss     83      0.0044440 20.75
10  0  298.975  381.975  5.3463 3.14159  gaussdot  83      0.0009744 20.75
11  0  381.975  514.950  5.1167 3.14159  gaussflat 132.975 0.0111083 5     15
```

5.5 Alternative gate optimization techniques

The goal of obtaining the best pulse parameters to implement a desired quantum gate can be addressed by many different optimization techniques. In Section 5.3.1, we have discussed why a Nelder–Mead optimization is a reasonable approach for the present case, in which the evaluation of the objective function by simulating the time evolution is the most expensive step.

However, given the vast number of optimization and machine learning techniques, one may pursue other approaches in the hope of finding better pulse parameters or completely new pulse shapes. Such efforts belong to the field of quantum control and are often addressed by applying the techniques directly to the experiment. Especially the popularity of deep reinforcement learning (DeepRL) [Goo2016], which has stemmed from the recent successes in playing Atari games [Mni2015] or board games such as Go [Sil2016], may inspire efforts to apply the framework to quantum control. For this reason, a few research groups have recently implemented strategies to apply DeepRL to control optimization [Niu2019; An2019] (see also [Pal2017]). It would be interesting to study such methods to obtain completely new pulses, and to test them using both the transmon simulator and the real processor by means of the recently released OpenPulse interface [McK2018].

For the purpose of the present work, however, we choose not to optimize the pulses further. One reason is that the pulse parameters found by the Nelder–Mead method are in the same range that is used in experiments, so the resulting pulses resemble the experimental approach. In fact, as we shall see in the following chapter, the pulses often exhibit a much better performance than their experimental equivalents. Hence we leave the endeavor of studying more optimization methods for pulses in simulated transmon quantum computers for future work.

5.6 Conclusions

In this chapter, we studied how time-dependent pulses can be used and optimized to implement quantum gates. Two results obtained in this analysis are noteworthy: First, the third-order perturbative results for the CR interaction strengths given by Eqs. (5.22a) and (5.22b) are useful for obtaining initial pulse parameters, but the predictions quickly diverge from the actual interaction strengths for larger CR drive amplitudes (see Fig. 5.4). Secondly, when the drive frequencies are not tuned during the pulse optimization, the DRAG coefficients β_X can be (mis)used to compensate for phase errors such that equally good quantum gates are still possible (cf. the fourth row in Fig. 5.6(a)).

However, it is important to realize that, because of the presence of higher transmon states in the time evolution of the full system, the projection of the full time-evolution operator on the computational subspace (see Eq. (5.2)) is inherently non-unitary. Therefore, even in theory, it is impossible to realize perfect quantum gates on the computational subspace (unless certain resonance conditions are met exactly, which is practically impossible). This implies that bare quantum gate implementations are always faulty. The question is whether the associated error rates can still be made small enough to allow for a reasonable operation of the devices. This question is studied in the following chapters.

Chapter 6

Errors in quantum gates

The goal of this chapter is to characterize the optimized quantum gate pulses and compare their performance to experiments. In general, neither the pulses used in experiments nor the optimized pulses used for the transmon simulator can implement perfect quantum gates (see Section 5.6). The advantage of the simulator, though, is that we have full access to the complex coefficients of the quantum state. This means that a much deeper analysis of the intrinsic errors can be performed.

Note that all errors observed in the simulation are inherently part of the unitary time evolution of the full transmon-resonator system. However, this does not mean that they have to be unitary maps on the computational subspace (such as systematic over- or underrotations). Often, unitary maps on a large system are not unitary on a smaller part of the system (see also Section 2.4).

We study quantum gate errors in four complementary ways:

(a) by evaluating common gate metrics such as the fidelity [Nie2002], the diamond distance [Kit1997], and the unitarity [Wal2015b];

(b) by analyzing repeated gate applications;

(c) by performing gate set tomography using the black box model of a quantum computer;

(d) by assessing the performance of actual quantum algorithms.

Results for (a) and (b) are presented in Sections 6.1 and 6.2, respectively. These findings can be directly compared to experiments on the IBM Q Experience [IBM2016]. Section 6.3 contains results for (c), i.e., an extensive gate set tomography of the simulated two-transmon device. Real quantum algorithms (d) are partially tested in this chapter (such as the quantum Fourier transform in Section 6.2 or the circuits for gate set tomography in Section 6.3), but this is mainly the topic of the following chapter. Some of the results presented in Sections 6.1 and 6.2 of this chapter have been published in [Wil2017].

6.1 Evaluation of gate metrics

The performance of an implementation for a certain quantum gate is often measured in terms of various gate metrics. Such metrics are meant to be single numbers representing the degree of success or failure to which the quantum gate has been implemented. Some of these, such as the diamond distance, are sufficient in the mathematical sense, meaning that a value of 0 implies a perfect implementation of the gate. Of course, in an experiment, it is not possible to unambiguously prove a perfect implementation, simply because a particular quantity cannot be estimated with zero error in practice. Nevertheless, a good value may still inspire confidence in the underlying implementation.

Gate metrics are defined in terms of quantum operations, i.e., completely positive linear maps on the space of density matrices on the computational subspace $\mathcal{H}_{2^n}$ (see Section 2.4). We define two particular quantum operations,

$$\mathcal{G}_{id}(|\psi\rangle\langle\psi|) = U\,|\psi\rangle\langle\psi|\,U^\dagger, \tag{6.1a}$$

$$\mathcal{G}_{ac}(|\psi\rangle\langle\psi|) = M\,|\psi\rangle\langle\psi|\,M^\dagger. \tag{6.1b}$$

Here, $\mathcal{G}_{id}$ denotes the *ideal* quantum operation, where U is the intended unitary quantum gate (cf. Section 2.2), and $\mathcal{G}_{ac}$ denotes the *actual* operation performed on the computational subspace. The matrix M in Eq. (6.1b) is the transformation of the computational subspace obtained after applying the quantum gate pulse to the full system. Formally, M is given by the projection in Eq. (5.2), i.e., $M = P_{\mathcal{H}_{2^n}}\mathcal{U}(T,0)P_{\mathcal{H}_{2^n}}$, where $\mathcal{U}(T,0)$ is the time evolution operator resulting from the simulation (potentially expressed in a rotating frame according to Eq. (3.41)), and $P_{\mathcal{H}_{2^n}}$ is the projector onto the computational subspace. Note that n denotes the number of qubits involved in the quantum gate U, which may be smaller than the total number of qubits.

We construct the matrix M in Eq. (6.1b) using the same procedure that was used for the optimization of the pulses (see Section 5.3). Specifically, for single-qubit gates, we have $M \in \mathbb{C}^{2\times2}$, so constructing M requires two simulations of the time evolutions under the particular pulse (one for the initial state $|0\rangle$ and one for the initial state $|1\rangle$). Similarly, constructing M for a two-qubit gate requires four simulations.

Because of the projection in Eq. (5.2), M is generally not a unitary matrix. Nevertheless, the quantum operations defined in Eqs. (6.1a) and (6.1b) are completely positive maps. This can be seen since $\langle\phi|(\mathcal{G}_{ac} \otimes \mathbb{1})(A)|\phi\rangle = \langle\phi'|A|\phi'\rangle \geq 0$ for all positive operators A (using $|\phi'\rangle = (M\otimes\mathbb{1})^\dagger\,|\phi\rangle$). We remark that an alternative definition using M^{-1} instead of $M^\dagger$ in Eq. (6.1b) would not preserve Hermiticity, so the computed gate metrics would not be real numbers.

6.1.1 Average gate fidelity

The average gate fidelity between the quantum operations defined in Eqs. (6.1a) and (6.1b) is given by [Hor1999; Nie2002]

$$F_{\text{avg}} = \int \mathrm{d}|\psi\rangle \, \langle\psi|\,\mathcal{G}_{ac}(\mathcal{G}_{id}^{-1}(|\psi\rangle\langle\psi|))\,|\psi\rangle\,, \tag{6.2}$$

where the integral is taken over normalized pure states. In practice, Eq. (6.2) means that the overlap of a state with itself after applying $\mathcal{G}_{ac}$ and $\mathcal{G}_{id}^{-1}$ is averaged over random pure states from the unit sphere [Ben2006], yielding an average measure of agreement between the ideal and the actual gate operation.

Note that some authors [Bon2010; Sek2014] use a different formula for the fidelity in the case of trace-decreasing quantum operations. These formulas are effectively based on trace-preserving extensions of the quantum operation. However, we decide to use Eq. (6.2) also for trace-decreasing quantum operations. The reason is that, if a quantum operation does not preserve the trace, it suffers from leakage, which should be reflected accordingly by a reduced fidelity. Otherwise, we could have quantum operations with leakage out of the computational subspace that still attain fidelities unreasonably close to unity.

In practice, one could evaluate the average gate fidelity in Eq. (6.2) by sampling the integrand $\langle\psi|\,\mathcal{G}_{ac}(\mathcal{G}_{id}^{-1}(|\psi\rangle\langle\psi|))\,|\psi\rangle = |\langle\psi|\,MU^{\dagger}\,|\psi\rangle|^2$ for a sufficiently large number of random states $|\psi\rangle$ [Ben2006]. This procedure is used for the unitarity defined below.

For the fidelity, however, we can use an alternative closed-form expression proven in Appendix G,

$$F_{\text{avg}} = \frac{|\text{Tr}(MU^{\dagger})|^2 + \text{Tr}(M^{\dagger}M)}{N(N+1)}, \tag{6.3}$$

where $N = 2^n$ denotes the dimension of the computational subspace. It can either be derived by generalizing a well-known relation between the average gate fidelity and the entanglement fidelity [Hor1999; Nie2002] to the case of non-trace-preserving quantum operations (see Section G.2), or by evaluating the integral in Eq. (6.2) directly (see Section G.3). Equation (6.3) can also be found in [Ped2007].

In experiments, the average gate fidelity is typically estimated using randomized benchmarking (RB) [Eme2005; Kni2008; Mag2012]. However, such results must be treated with caution as it has been shown that RB cannot measure F_{avg} [Qi2019], but typically produces numbers that overestimate the performance of the gates [Pro2017; Lin2019].

6.1.2 Diamond distance

The diamond distance is a mathematical construct that measures the difference between quantum operations. It was introduced in [Kit1997; Aha2008] and is the relevant quantity for many threshold theorems in the theory of fault-tolerant quantum computation [Ter2005; Ali2006; Ali2007; Aha2008; Ng2009; San2016]. Unfortunately, evaluating the diamond distance in practice is nontrivial [Hen2010; Wat2018]. Furthermore, the relation to more accessible quantities such as the average gate fidelity given by Eq. (6.2) is not straightforward [Wal2014; San2016]. Therefore, we look at these issues in more detail.

Definition of the diamond distance

The diamond distance between the two quantum operations $\mathcal{G}_{id}$ and $\mathcal{G}_{ac}$ is defined as

$$\eta_{\diamond} = \frac{1}{2} \left\| \mathcal{G}_{ac} \circ \mathcal{G}_{id}^{-1} - \mathbb{1} \right\|_{\diamond}, \tag{6.4}$$

where $\|\cdot\|_\Diamond$ denotes the diamond norm. For a general superoperator $\mathcal{T}$, the diamond norm is defined as [Kit1997; Aha1998]

$$\|\mathcal{T}\|_\Diamond = \sup_{X \neq 0} \frac{\|(\mathcal{T} \otimes \mathbb{1})(X)\|_{\mathrm{Tr}}}{\|X\|_{\mathrm{Tr}}}, \tag{6.5}$$

where $\|\cdot\|_{\mathrm{Tr}}$ denotes the trace norm defined as $\|X\|_{\mathrm{Tr}} = \mathrm{Tr}\sqrt{X^\dagger X}$, i.e., the sum of the singular values of X. The identity operator $\mathbb{1}$ in Eq. (6.5) acts on a space that is at least as large as the space of matrices that $\mathcal{T}$ acts on. Interestingly, one can show that it does not need to be larger [Gil2005; Joh2009]. This means that the supremum can be computed by extending the Hilbert space with another Hilbert space of the same dimensionality.

The fact that the diamond norm is the same even if the Hilbert space is extended by a much larger space is a stability property that is sometimes expressed in an alternative definition of the diamond norm (see e.g. [San2016]),

$$\|\mathcal{T}\|_\Diamond = \sup_{\mathcal{H}'} \sup_{\rho \in \mathrm{dens}(\mathcal{H} \otimes \mathcal{H}')} \|(\mathcal{T} \otimes \mathbb{1})(\rho)\|_{\mathrm{Tr}}, \tag{6.6}$$

where $\mathcal{H}'$ is an arbitrary ancillary Hilbert space and $\mathrm{dens}(\mathcal{H} \otimes \mathcal{H}')$ denotes the set of all density matrices on the joint Hilbert space $\mathcal{H} \otimes \mathcal{H}'$. Equation (6.6) means that the diamond distance $\eta_\Diamond$ corresponds to the worst-case error, since the trace norm is maximized over all ancillary Hilbert spaces $\mathcal{H}'$ added to $\mathcal{H}$ and any input density matrix ρ on this joint space. The diamond distance is thus a very strong measure, in the sense that a small value for this quantity is not easy to achieve for implementations of quantum gates. Note that the diamond distance can be interpreted as the distance between ideal and actual probability distributions, since the trace norm is directly related to the total variation distance [Gil2005; San2016].

We note that in the mathematical literature, the diamond norm is also known as the completely bounded trace norm [Pau2003]. Many properties of this norm in relation to its use in quantum information can be found in [Wat2018].

Computation of the diamond distance

We consider two ways to compute the diamond distance given by Eq. (6.4). The first method is based on a direct evaluation of the definition given in Eq. (6.5), so it includes a maximization problem. The second method makes use of a minimization algorithm presented in [Joh2009]. If both methods (i.e., the minimization and the maximization) yield the same numeric quantity up to a reasonable number of digits, we have estimated the true diamond distance with sufficient accuracy.

Historically, the first deep connection between the evaluation of the diamond distance between two quantum channels and a convex optimization problem was made in [Gil2005]. Apart from the methods we pursue in this work, similar algorithms have been mentioned in [Zar2006; Ben2010], and a class of algorithms based on semidefinite programming has been given in [Wat2009; Wat2013; Wat2018], including proofs of their efficiency.

Maximization algorithm

The definition of the diamond norm given by Eq. (6.5) is valid for a general linear map $\mathcal{T}$. However, in almost all practical situations, $\mathcal{T}$ is at least Hermiticity-preserving, i.e., $\mathcal{T}(\rho)^\dagger = \mathcal{T}(\rho^\dagger)$. This is also true in the present case, in which

$$\mathcal{T}(\rho) = \mathcal{G}_{ac}(\mathcal{G}_{id}(\rho)) - \rho = W\rho W^\dagger - \rho, \tag{6.7}$$

where $W = MU^\dagger$ (cf. Eqs. (6.1a) and (6.1b)). In this case, the supremum in Eq. (6.5) is attained by a pure state $X = |x\rangle\langle x|$ with $|x\rangle \in \mathcal{H}_{2^n} \otimes \mathcal{H}_{2^n}$ [Wat2018]. Since $\|X\|_{\mathrm{Tr}} = \langle x|x\rangle = 1$, Eq. (6.4) becomes

$$\eta_\Diamond = \frac{1}{2} \max_{|x\rangle} \left\| (W \otimes \mathbb{1}) |x\rangle\langle x| (W^\dagger \otimes \mathbb{1}) - |x\rangle\langle x| \right\|_{\mathrm{Tr}}. \tag{6.8}$$

This expression requires the evaluation of the trace norm of a rank-2 matrix of the form $R = \alpha |v\rangle\langle v| - \beta |x\rangle\langle x|$, where $\alpha, \beta \geq 0$ and $|v\rangle$ is a normalized pure state. For the trace norm $\|R\|_{\mathrm{Tr}}$ of such a matrix, one can derive a closed-form expression:

$$\| \alpha |v\rangle\langle v| - \beta |x\rangle\langle x| \|_{\mathrm{Tr}} = \sqrt{(\alpha + \beta)^2 - 4\alpha\beta|\langle v|x\rangle|^2}. \tag{6.9}$$

This can be shown by noting that a rank-2 matrix has at most two non-zero singular values. Since the rank-2 matrix R is Hermitian (and thus normal), its singular values are the absolute values of its eigenvalues $\mu_\pm$, for which a short calculation yields

$$\mu_\pm = \frac{\alpha - \beta}{2} \pm \frac{1}{2}\sqrt{(\alpha + \beta)^2 - 4\alpha\beta|\langle v|x\rangle|^2}. \tag{6.10}$$

As both eigenvalues $\mu_\pm$ have opposite signs (meaning $|\mu_\pm| = \pm\mu_\pm$ since $\alpha, \beta > 0$ and $0 \leq |\langle v|x\rangle| \leq 1$), we obtain the trace norm $\|R\|_{\mathrm{Tr}} = |\mu_+| + |\mu_-| = \mu_+ - \mu_-$.

To apply this result to Eq. (6.8), we set $\alpha = \langle x|W^\dagger W \otimes \mathbb{1}|x\rangle \in [0,1]$, $\beta = 1$, $|v\rangle = (W \otimes \mathbb{1}) |x\rangle / \sqrt{\alpha}$ and $|\langle v|x\rangle|^2 = |\langle x|W \otimes \mathbb{1}|x\rangle|^2/\alpha$. Thus we obtain

$$\eta_\Diamond = \frac{1}{2} \max_{|x\rangle} \sqrt{(\langle x|W^\dagger W \otimes \mathbb{1}|x\rangle + 1)^2 - 4|\langle x|W \otimes \mathbb{1}|x\rangle|^2}. \tag{6.11}$$

For the one- and two-qubit quantum operations of interest, this result describes a quadratic optimization problem that can directly be solved on a computer [Boy2004].

Furthermore, we show in Appendix H that if both quantum operations $\mathcal{G}_{ac} = M \cdot M^\dagger$ and $\mathcal{G}_{id} = U \cdot U^\dagger$ are unitary, we can obtain an explicit result from Eq. (6.11). This yields an elementary proof of the statements given in [Aha1998] and [Joh2009]. In most situations considered in the present work, however, M is not exactly unitary due to leakage out of the computational subspace (see below).

Minimization algorithm

For the second algorithm to compute the diamond norm $\|\mathcal{T}\|_\diamond$, we use a slightly modified version of the minimization algorithm presented in [Joh2009]. The algorithm is based on minimizing over all generalized Kraus representations

$$\mathcal{T}(\rho) = \sum_l A_l \rho B_l, \tag{6.12}$$

where A_l and B_l are generalized Kraus operators of $\mathcal{T}$ (cf. Eq. (2.24) in Section 2.4). We have

$$\|\mathcal{T}\|_\diamond = \inf_{A_l, B_l} \left\{ \left\| \sum_l A_l^\dagger A_l \right\|_2^{1/2} \left\| \sum_l B_l B_l^\dagger \right\|_2^{1/2} \right\}, \tag{6.13}$$

where $\|\cdot\|_2$ denotes the spectral norm (largest singular value). Note that this expression differs from the one given in [Joh2009] by the position of the Hermitian conjugate. The reason for this is that the completely bounded (spectral) norm computed in [Joh2009] and the diamond norm computed here are related to each other by replacing $A_l \mapsto A_l^\dagger$ and $B_l \mapsto B_l^\dagger$.

To compute $\eta_\diamond$ given by Eq. (6.4), we set $\mathcal{T} = \mathcal{G}_{ac} \circ \mathcal{G}_{id}^{-1} - \mathbb{1}$. We obtain a generalized Kraus representation of $\mathcal{T}$ by inserting Eqs. (6.1a) and (6.1b):

$$\mathcal{T}(\rho) = \mathcal{G}_{ac}(\mathcal{G}_{id}^{-1}(\rho)) - \rho = MU^\dagger \rho U M^\dagger - \rho, \tag{6.14}$$

where we can identify $A_1 = MU^\dagger, A_2 = \mathbb{1}, B_1 = UM^\dagger$, and $B_2 = -\mathbb{1}$. We assume that both (A_1, A_2) and (B_1, B_2) are linearly independent, since otherwise we would have $MU^\dagger = \gamma\mathbb{1}$ for some γ such that we could directly obtain $\eta_\diamond = |1 - |\gamma|^2|/2$ (which is consistent with Eq. (6.11) for $W = \gamma\mathbb{1}$).

All other generalized Kraus representations of $\mathcal{T}$ are related to (A_1, A_2) and (B_1, B_2) by an invertible complex 2×2 matrix S. Thus, the infimum over A_l and B_l in Eq. (6.13) can be written as an infimum over all invertible $S \in \mathbb{C}^{2\times 2}$ after replacing $(A_1, A_2)^T \mapsto S^{-1}(A_1, A_2)^T$ and $(B_1, B_2) \mapsto (B_1, B_2)S$. Combining this with Eq. (6.13), we obtain

$$\eta_\diamond = \frac{1}{2} \inf_S \left\{ \left\| (UM^\dagger, \quad \mathbb{1}) \, S^{-\dagger} S^{-1} \begin{pmatrix} MU^\dagger \\ \mathbb{1} \end{pmatrix} \right\|_2^{1/2} \left\| (UM^\dagger, \quad -\mathbb{1}) \, SS^\dagger \begin{pmatrix} MU^\dagger \\ -\mathbb{1} \end{pmatrix} \right\|_2^{1/2} \right\}, \tag{6.15}$$

where the infimum is taken over all invertible $S \in \mathbb{C}^{2\times 2}$.

In practice, we evaluate Eq. (6.15) by sampling over 10^4 random S and then applying the minimization method described in Section 5.3.1 to the eight real coefficients of the best S sampled. We verified that this procedure, together with the maximization method discussed above, produces reliable values for $\eta_\diamond$ in all relevant cases.

Relation to the fidelity

The relation between the average gate fidelity given by Eq. (6.2) and the diamond distance given by Eq. (6.4) is not obvious. As discussed in [San2016], this has led some groups to make partly unjustified claims about reaching fault-tolerance thresholds [Cho2012; Bar2014]. As a consequence, the connection between fidelity and diamond distance has been further studied in the literature [Wal2014; Wal2015a; Kue2016]. In these references, two bounds relating both quantities have been proven. We have

$$\eta_\Diamond^{\text{Pauli}} \leq \eta_\Diamond \leq \eta_\Diamond^{\text{ub}}, \tag{6.16}$$

where, for trace-preserving quantum operations,

$$\eta_\Diamond^{\text{Pauli}} = \frac{N+1}{N}(1 - F_{\text{avg}}), \tag{6.17a}$$

$$\eta_\Diamond^{\text{ub}} = \sqrt{N(N+1)(1 - F_{\text{avg}})}. \tag{6.17b}$$

For trace-decreasing quantum operations, however, $\eta_\Diamond$ can actually be lower than $\eta_\Diamond^{\text{Pauli}}$. Therefore, we prove a new lower bound, which also holds for trace-decreasing quantum operations, in Appendix I. Applying the result in Eq. (I.8) to $\mathcal{E} = MU^\dagger \cdot UM^\dagger$ yields

$$\eta_\Diamond^{\text{lb}} = \eta_\Diamond^{\text{Pauli}} - \frac{N+2}{2N}\left(1 - \frac{\text{Tr}\, M^\dagger M}{N}\right), \tag{6.18}$$

which is a proper lower bound on $\eta_\Diamond$. Note that this bound coincides with Eq. (6.17a) in the trace-preserving case $\text{Tr}\, M^\dagger M = N$.

The lower bound $\eta_\Diamond^{\text{Pauli}}$ given by Eq. (6.17a) is saturated if the actual operation $\mathcal{G}_{ac}$ can be represented as a Pauli channel [San2016]. This means that $\mathcal{G}_{ac}$ has a Kraus representation in which each Kraus operator is proportional to a tensor product of Pauli matrices. Such quantum operations are typically considered to represent errors that are easy to correct. We shall see that for the simulations considered in this work, we often find $\eta_\Diamond \gg \eta_\Diamond^{\text{Pauli}}$ (cf. Tab. 6.1).

The upper bound $\eta_\Diamond^{\text{ub}}$ given by Eq. (6.17b) grows exponentially with the number of qubits, which is the reason for the fact that impressively high fidelities do not generally imply a sufficiently small diamond distance [San2016]. Furthermore, the authors in [San2016] have shown that the upper bound is asymptotically tight, which means there cannot be any bound on $\eta_\Diamond$ in terms of N and F_{avg} that scales better than exponentially in the number of qubits.

6.1.3 Unitarity

Because of the projection required to obtain the transformation M of the computational subspace under a certain pulse, the actual operation $\mathcal{G}_{ac}$ given by Eq. (6.1b) is often not unitary. This means that $M^\dagger \neq M^{-1}$, implying that the quantum operation is not trace-preserving such that we typically have $\text{Tr}(\mathcal{G}_{ac}(\rho)) < \text{Tr}(\rho)$. This is the mathematical manifestation of the fact that the systems under investigation suffer from *leakage* into non-computational states (see also the discussion below Eq. (5.2)).

To quantify the effects of leakage for a particular quantum gate pulse, a quantity called *unitarity* has been introduced [Wal2015b; Wal2016]. It is defined as

$$u = \frac{N}{N-1} \int \mathrm{d}|\psi\rangle \, \mathrm{Tr}\big[\mathcal{G}'_{ac}(|\psi\rangle\langle\psi|)^\dagger \mathcal{G}'_{ac}(|\psi\rangle\langle\psi|)\big], \tag{6.19}$$

where

$$\mathcal{G}'_{ac}(|\psi\rangle\langle\psi|) = \mathcal{G}_{ac}(|\psi\rangle\langle\psi| - \mathbb{1}/N) - \mathrm{Tr}\left[\frac{\mathcal{G}_{ac}(|\psi\rangle\langle\psi| - \mathbb{1}/N)}{\sqrt{N}}\right] \mathbb{1}. \tag{6.20}$$

The rationale behind this definition is that Eq. (6.19) corresponds to the average purity $\mathrm{Tr}(\rho^\dagger \rho)$ where ρ is the output of the quantum operation acting on a pure state. The reason that the quantum operation is $\mathcal{G}'_{ac}$ instead of $\mathcal{G}_{ac}$ is that otherwise, one could define explicitly trace-decreasing or non-unital quantum operations with $u = 1$ (see [Wal2015b]). As for the average gate fidelity given by Eq. (6.2), the integral is taken over random states from the unit sphere. Since we have no closed-form expression in this case, we evaluate u by sampling over 10^5 random states. A typical procedure to generate random states is given by setting $|\psi\rangle = \sum_j (a_j + ib_j)|j\rangle$, where the coefficients a_j and b_j are first drawn from a normal distribution and subsequently normalized [Ben2006].

As the definition of the unitarity u given by Eq. (6.19) involves only the actual quantum operation $\mathcal{G}_{ac}$ and not the ideal gate $\mathcal{G}_{id}$, the value $u = 1$ does not imply equivalence between the two operations. In that sense, it differs from other metrics such as the diamond distance introduced in Section 6.1.2. Nevertheless, the unitarity is a useful quantity to measure how incoherent the errors appear on the computational subspace. Therefore, it is an interesting metric to study in the present case, in which all errors are actually systematic and coherent on the total transmon-resonator Hilbert space.

6.1.4 Results

We evaluate the average gate fidelity F_{avg} given by Eq. (6.2), the diamond distance $\eta_\Diamond$ given by Eq. (6.4), and the unitarity u given by Eq. (6.19) for all three transmon systems for which we have optimized quantum gate pulses (cf. Section 5.3.2). For the two-transmon system, we additionally evaluate the bounds given in Eqs. (6.17a)–(6.18).

Two-transmon system

The gate metrics for the two-transmon system defined in Section 3.4.3 are given in Tab. 6.1. We see that the overall performance of the gates is reasonably good, in that the fidelity F_{avg} and the unitarity u are close to one, and the diamond distance $\eta_\Diamond$ is close to zero. The fidelities are in the same range as those reported in experiments for similar pulse schemes [Kel2014; Bar2014; She2016a; She2016b; Gam2017]. In fact, top fidelities reported in experiments are sometimes better; this is an artifact of the RB procedure used to measure the fidelities: As shown in [Pro2017], the number reported by RB often overestimates the performance of the gate. Indeed, we shall see below (cf., for instance, Fig. 6.2(a)) that in practice, the performance of the pulses optimized for the transmon simulator is often much better.

Table 6.1: Gate metrics for the two-transmon system defined in Section 3.4.3. The pulse types are defined in Section 5.1 and Section 5.2 (see Appendix F for the optimized pulse parameters). The label in the third column refers to the internal name used for the compiler (see Listing 5.1 in Section 5.4). Δ is the distance objective given by Eq. (5.24). The reported gate metrics are the average gate fidelity F_{avg} (see Eq. (6.2)), the diamond distance $\eta_\Diamond$ (see Eq. (6.4)), and the unitarity u (see Eq. (6.19)). The bounds $\eta_\Diamond^{\mathrm{lb}}$, $\eta_\Diamond^{\mathrm{Pauli}}$, and $\eta_\Diamond^{\mathrm{ub}}$ are given by Eqs. (6.17a)–(6.18), respectively.

Gate	Pulse	Label	Δ	F_{avg}	$\eta_\Diamond$	$\eta_\Diamond^{\mathrm{lb}}$	$\eta_\Diamond^{\mathrm{Pauli}}$	$\eta_\Diamond^{\mathrm{ub}}$	u
$X_0^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-0	2.2×10^{-3}	0.9946	0.027	0.003	0.007	0.33	0.990
$X_1^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-1	2.3×10^{-3}	0.9942	0.028	0.003	0.007	0.34	0.989
X_0^{π}	GD^{π}	xpi-0	1.3×10^{-3}	0.9949	0.020	0.003	0.006	0.32	0.990
X_1^{π}	GD^{π}	xpi-1	1.5×10^{-3}	0.9943	0.023	0.003	0.007	0.34	0.989
CNOT_{01}	CR1	cnot-0-1	1.3×10^{-3}	0.9842	0.029	0.008	0.020	0.56	0.969
CNOT_{10}	CR1	cnot-1-0	2.3×10^{-3}	0.9951	0.033	0.003	0.006	0.31	0.991
CNOT_{01}	CR2	cnot-0-1	6.1×10^{-3}	0.9943	0.048	0.004	0.007	0.34	0.991
CNOT_{10}	CR2	cnot-1-0	5.6×10^{-3}	0.9947	0.048	0.003	0.007	0.32	0.992
CNOT_{01}	CR4	cnot-0-1	5.4×10^{-3}	0.9934	0.049	0.004	0.008	0.36	0.989
CNOT_{10}	CR4	cnot-1-0	4.5×10^{-3}	0.9946	0.044	0.003	0.007	0.33	0.991

For all gates (including the single-qubit gates), we have optimized the transformation M of the entire, four-dimensional computational subspace (see Eq. (5.2)). In this respect, the two-transmon case differs from the optimizations for the five-transmon systems discussed below, for which only the relevant part of the computational subspace is considered.

Note that a large fidelity does not always correspond to a small diamond distance. Especially for the CNOT gates, we observe similar fidelities as for the single-qubit gates but an almost twice as large diamond distance (see also [San2016]). An extreme case is the CR1-type CNOT_{01}: The fidelity $F_{\mathrm{avg}} = 0.9842$ is much worse than for the other gates, whereas the diamond distance $\eta_\Diamond = 0.029$ has the best value found for all two-qubit gates.

A decrease in fidelity often corresponds to a decrease in unitarity. From this, we conclude that leakage due to non-computational states in the transmons and the resonator is the dominant source of error for the pulses optimized for this system, even though the pulse-shaping techniques DRAG [Mot2009] and VZ corrections [McK2017] have been used.

The diamond distances given in Tab. 6.1 are always within the bounds given in Eq. (6.16). In most cases, we also find $\eta_\Diamond \gg \eta_\Diamond^{\mathrm{Pauli}}$, suggesting that the systematic errors included in the pulses are inherently different from the simple Pauli-type errors (see [San2016]).

Small five-transmon system

The gate metrics of the elementary pulses for the small five-transmon system defined in Section 3.4.4 are given in Tab. 6.2. For the pulse optimizations, the matrices M

Table 6.2: Gate metrics for the small five-transmon system defined in Section 3.4.4. The pulse types are defined in Section 5.2 (see Appendix F for the optimized pulse parameters). Δ is the distance objective given by Eq. (5.24), F_{avg} is the average gate fidelity given by Eq. (6.2), $\eta_\diamond$ is the diamond distance given by Eq. (6.4), and u is the unitarity given by Eq. (6.19).

Gate	Pulse	Label	Δ	F_{avg}	$\eta_\diamond$	u
$X_0^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	`xpih-0-withf`	1.30×10^{-4}	0.9879	0.011	0.9759
$X_1^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	`xpih-1-withf`	1.28×10^{-4}	0.9879	0.011	0.9758
$X_2^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	`xpih-2-withf`	2.20×10^{-4}	0.9837	0.015	0.9675
$X_3^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	`xpih-3-withf`	7.76×10^{-6}	0.9965	0.003	0.9930
$X_4^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	`xpih-4-withf`	2.71×10^{-4}	0.9828	0.016	0.9658
CNOT_{02}	CR2	`cnot-0-2-withf`	5.04×10^{-2}	0.9691	0.136	0.9584
CNOT_{12}	CR2	`cnot-1-2-withf`	2.92×10^{-2}	0.9708	0.107	0.9534
CNOT_{32}	CR2	`cnot-3-2-withf`	1.46×10^{-2}	0.9848	0.082	0.9754
CNOT_{42}	CR2	`cnot-4-2-withf`	6.63×10^{-2}	0.9774	0.177	0.9812

and U used in Eq. (5.24) are 2×2 matrices for all single-qubit gates $X_i^{\pi/2}$, and 4×4 matrices for all two-qubit gates CNOT_{ij}. This is different from the two-transmon results given in Tab. 6.1, for which always the full four-dimensional computational subspace was considered. Although this does not really affect the quality of the optimized pulses, it can be seen in the objective function Δ used for the optimization (see Tab. 6.2), because it typically differs by about two orders of magnitude between single-qubit and two-qubit pulses.

Another difference to the two-transmon system is that also the drive frequency f has been optimized (cf. Section 5.3.2). This is indicated by the label `withf` in Tab. 6.2. The reason is that we found that when keeping the drive frequency f fixed, especially the two-qubit gates yield diamond distances approximately twice as large and fidelities as low as 0.94 (data not shown).

The resulting gate metrics in Tab. 6.2 are slightly worse than those obtained for the two-transmon system (see Tab. 6.1). This is reasonable since every transmon is coupled to at least one additional transmon through a resonator (see the topology of the system in Fig. 3.4). This means that a much larger number of states are present in the joint time evolution of the system and affect the gate operation. In particular, the CNOT gates, for which the target qubit is coupled to three additional qubits (typically called "spectator" qubits [Tak2017]), suffer from this aspect in that they yield diamond distances larger by a factor of 2–4 compared to the two-transmon system.

A particularly noteworthy result is the single-qubit gate $X_3^{\pi/2}$, which attains by far the best values for all gate metrics, also in comparison to the two-transmon results shown in Tab. 6.1. However, when looking at an actual application (see the fourth Bloch vector visualized in Fig. A.1 in Appendix A), the performance of $X_1^{\pi/2}$ is better despite much worse gate metrics.

Table 6.3: The same as in Tab. 6.2 for the large five-transmon system defined in Section 3.4.5. As indicated by the label in the third column, all gate pulses have been optimized both with and without frequency tuning (cf. Section 5.3.2).

Gate	Pulse	Label	Δ	F_{avg}	$\eta_\diamond$	u
$X_0^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-0	4.60×10^{-5}	0.9930	0.007	0.9860
$X_1^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-1	1.19×10^{-4}	0.9884	0.011	0.9770
$X_2^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-2	7.52×10^{-6}	0.9962	0.002	0.9925
$X_3^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-3	8.99×10^{-6}	0.9965	0.003	0.9930
$X_4^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-4	4.17×10^{-5}	0.9934	0.006	0.9868
$X_0^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-0-withf	4.59×10^{-5}	0.9930	0.007	0.9860
$X_1^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-1-withf	1.14×10^{-4}	0.9887	0.011	0.9774
$X_2^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-2-withf	7.20×10^{-6}	0.9963	0.002	0.9927
$X_3^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-3-withf	8.85×10^{-6}	0.9965	0.003	0.9930
$X_4^{\pi/2}$	$\mathrm{GD}^{\pi/2}$	xpih-4-withf	3.87×10^{-5}	0.9936	0.006	0.9873
CNOT_{10}	CR2	cnot-1-0	1.34×10^{-2}	0.9852	0.071	0.9758
CNOT_{14}	CR2	cnot-1-4	1.08×10^{-1}	0.9621	0.177	0.9668
CNOT_{21}	CR2	cnot-2-1	4.68×10^{-2}	0.9714	0.119	0.9615
CNOT_{32}	CR2	cnot-3-2	1.83×10^{-2}	0.9852	0.088	0.9777
CNOT_{34}	CR2	cnot-3-4	9.54×10^{-2}	0.9671	0.179	0.9720
CNOT_{40}	CR2	cnot-4-0	2.78×10^{-1}	0.9347	0.284	0.9783
CNOT_{10}	CR2	cnot-1-0-withf	5.70×10^{-2}	0.9751	0.149	0.9728
CNOT_{14}	CR2	cnot-1-4-withf	7.13×10^{-3}	0.9841	0.056	0.9712
CNOT_{21}	CR2	cnot-2-1-withf	1.38×10^{-2}	0.9806	0.081	0.9668
CNOT_{32}	CR2	cnot-3-2-withf	1.21×10^{-1}	0.9644	0.207	0.9764
CNOT_{34}	CR2	cnot-3-4-withf	1.88×10^{-2}	0.9832	0.090	0.9740
CNOT_{40}	CR2	cnot-4-0-withf	8.27×10^{-2}	0.9739	0.168	0.9806

Large five-transmon system

The gate metrics of the elementary pulses for the large five-transmon system defined in Section 3.4.5 are given in Tab. 6.3. As for the small five-qubit system, the optimized matrix M is a 2×2 matrix for all single-qubit gates $X_i^{\pi/2}$, and a 4×4 matrix for all two-qubit gates CNOT_{ij}.

However, in this case, we consider both sets of gate pulses, i.e., with and without frequency tuning. The reason for this is that the difference in the gate metrics of the CNOT gates with and without frequency tuning is not as pronounced as for the small five-qubit system. Furthermore, the single-qubit gates show almost no difference, so it is reasonable to compare their performance in actual quantum circuits instead.

Compared to the results for the small five-qubit system presented in Tab. 6.2, the single-qubit gates emerge considerably better from the optimization. Only the $X_1^{\pi/2}$ gate appears less promising. Therefore, although the parameter space is quite large, the gate metrics alone seem to suggest that the topology of the larger system shown in Fig. 3.5 (where only two transmons are connected by a resonator) is better suited for single-qubit

gates than the topology of the smaller system depicted in Fig. 3.4.

The two-qubit gate metrics reported in Tab. 6.2 show much larger fluctuations. Based solely on these metrics, there is no clear winner between the gates with and without frequency tuning. This makes sense regarding the much more complicated implementation of the two-qubit CR2 gates (see Fig. 5.1 in Section 5.2).

The pulse parameters of the gates with and without frequency tuning come out very differently: As can be seen in Tab. F.5 in Appendix F, for the gates without frequency tuning, the optimization often finds exotic values for the DRAG coefficients β_X. As the resulting gate metrics given in Tab. 6.3 are almost the same, though, it seems that the DRAG coefficients are effectively used to compensate for incorrect drive frequencies. This observation agrees with the conclusions from the previous chapter (see Section 5.6).

We remark that in practical applications, it turns out that the gate set with frequency tuning comes closer to experimental results and also shows better performance on average (see [Wil2018b]).

6.2 Repeated gate applications

The simplest practical test of the performance of quantum gates is to apply them repeatedly. Such an experiment yields a first quantitative assessment about how the accumulation of errors materializes in actual quantum algorithms. We first consider the evolution of the diamond distance under repeated pulse applications, and then analyze the evolution of observable error rates in both simulations and experiments.

6.2.1 Evolution of the diamond distance

For the first experiment on repeated gate applications, we use the two-transmon system because many errors are most clearly understood when reduced to the smallest reproducible case. Specifically, we test $r = 1, \ldots, 20$ repetitions of gate pulses from the elementary gate set optimized for the two-transmon system. This involves the single-qubit $X^{\pi/2}$ and X^{π} rotations corresponding to the elementary GD pulses (see Section 5.1), and the two-qubit CNOT gates corresponding to the CR1, CR2, and CR4 pulses (Section 5.2).

For each gate pulse, we run the two-transmon simulation with the corresponding pulses applied 20 times. Given the duration T of a certain pulse (cf. Tab. F.1 and Tab. F.2 in Appendix F), this means that the time evolution is simulated for $0 \leq t \leq 20T$ for each initial state from the computational basis ($|00\rangle, |01\rangle, |10\rangle, |11\rangle$). At each time $t = rT$ after r pulse applications, we extract the transformation matrix $M(rT) = P_{\mathcal{H}_{2^n}} \mathcal{U}(rT, 0) P_{\mathcal{H}_{2^n}}$ of the computational subspace according to Eq. (5.2) (see also Section 5.3).

If the implementation of a particular quantum gate U was perfect, the transformation $M(rT)$ would be equivalent to U^r. In this section, we measure the error rate between $M(rT)$ and U^r in terms of the diamond distance $\eta_\Diamond$ defined in Eq. (6.4), where the ideal operation and the actual operation are given by

$$\mathcal{G}_{id}(\rho) = U^r \rho (U^\dagger)^r, \tag{6.21a}$$

$$\mathcal{G}_{ac}(\rho) = M(rT) \rho M(rT)^\dagger. \tag{6.21b}$$
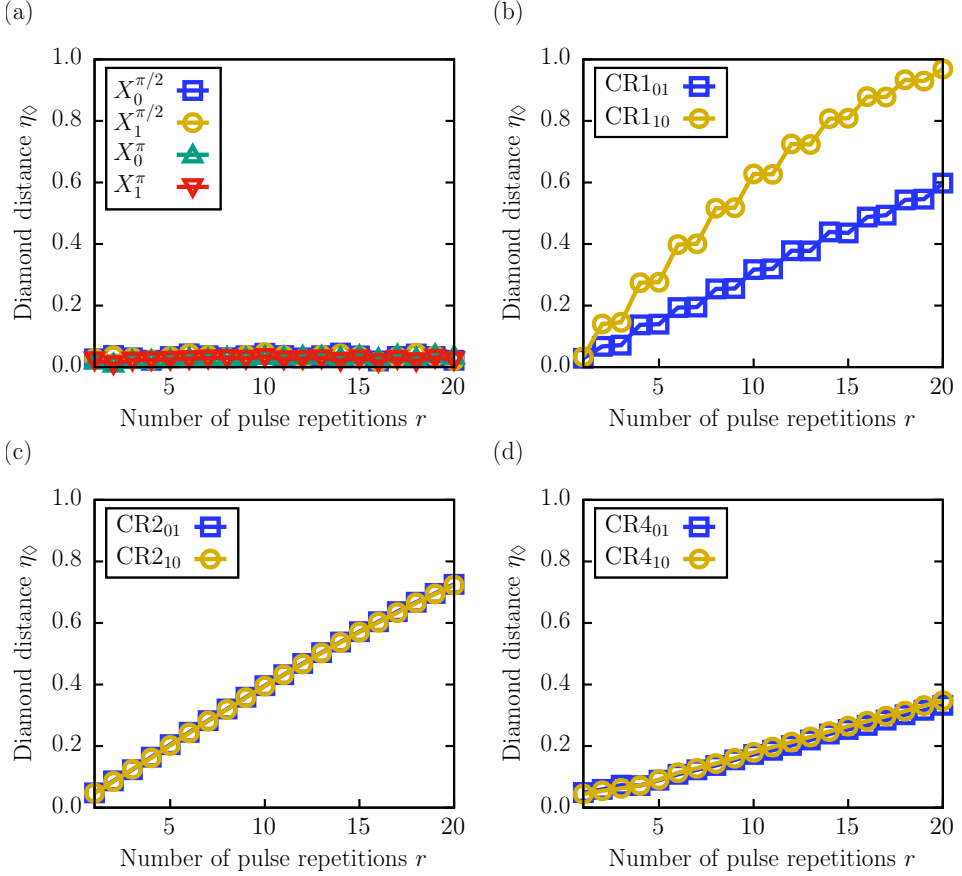
Figure 6.1: Error rate after repeated application of (a) single-qubit X gates and (b)–(d) two-qubit CNOT gates (the indicated CR pulse sequences are given in Fig. 5.1). Shown is the evolution of the diamond distance $\eta_\diamond$ (see Section 6.1.2) between the ideal operation $\mathcal{G}_{id}$ given by Eq. (6.21a), corresponding to r applications of the respective quantum gate U, and the actual operation $\mathcal{G}_{ac}$ given by Eq. (6.21b). The simulated system is the two-transmon system defined in Section 3.4.3. The corresponding gate metrics are given in Tab. 6.1.

Table 6.4: Comparison of the CR2 and CR4 pulses (cf. Fig. 5.1) for a single $CNOT_{01}$ gate, 20 successive $CNOT_{01}$ gates, and four successive QFT applications (one QFT contains five CNOT pulses and two GD pulses; see Fig. 2.3 for the circuit and Section 5.4 for how gates are mapped to pulses). The numbers reported are the diamond distances defined in Section 6.1.2, but the same qualitative trends are true for the other gate metrics (data not shown). The results for $CNOT^1$ are taken from Tab. 6.1. The results for $CNOT^1$ and $CNOT^{20}$ can also be seen as blue squares in Fig. 6.1(c) and (d).

Pulse	$CNOT^1$	$CNOT^{20}$	QFT^4
CR2	0.048	0.73	0.27
CR4	0.049	0.33	0.32

The results are presented in Fig. 6.1. We note that the same qualitative results for each of the curves can be seen for the other gate metrics studied in Section 6.1 (data not shown).

The performance of the single-qubit gates shown in Fig. 6.1(a) is reasonably good. The error rate of around 2% for $r = 1$ (which corresponds to the results for $\eta_\diamond$ given in Tab. 6.1) does not grow after repeated gate applications.

This is different for the two-qubit gates shown in Fig. 6.1(b)–(d), for which the error rate after repeated applications grows approximately linearly with r for all pulse schemes. For the CR1-type CNOT, we also observe a different performance if control and target qubit are interchanged. This error seems to have been canceled by the echo schemes used for the CR2 and CR4 pulses.

Comparing Fig. 6.1(b) and (c), we see that the CR1-type gate $CNOT_{01}$ (blue squares in Fig. 6.1(b)) achieves the smallest error rate after repeated applications. Interestingly, this gate pulse has the worst results for the average gate fidelity and the unitarity (see Tab. 6.1). This means that the gate metrics do not reflect the actual performance of the pulses in repeated applications.

The fact that we find the best performance for the CR4 pulse scheme (see Fig. 6.1(d)) agrees with experimental observations [Tak2017]. Note in particular that this is in contrast to the results suggested by the gate metrics in Tab. 6.1. For instance, the CR4-type gate $CNOT_{01}$ has the worst diamond distance $\eta_\diamond = 0.049$ after a single pulse application, but performs best of all in repeated applications (blue squares in Fig. 6.1(d)). Still, this observation can change again when looking at the performance of the gates in a practical quantum algorithm: In Tab. 6.4, we see that the same CR4 pulse performs worse than CR2 when used 20 times in a quantum Fourier transform (QFT). Therefore, the gate metrics can also not reliably predict the performance in practical applications.

Finally, it is worth mentioning that the transformation $M(rT)$ after $r > 1$ pulse applications is always closer to the ideal result U^r than the product $M(T)^r$ (data not shown). This means that the pulses have been tuned to the full time evolution of the transmon-resonator system including non-computational states; in a sense, the pulses are capable of using the more complicated dynamics on the larger space for improved results on the computational subspace. It also means that the effect of higher levels on the time evolution cannot be neglected. This is in agreement with the findings reported in Fig. 4.7(d),

which show that the two-level approximation cannot sufficiently describe the dynamics of full transmon systems. However, as we show in Section 6.3, there exist maps on the two-qubit reduced density matrix that can reliably describe the dynamics under repeated pulse applications.

6.2.2 Relation to experiments

The goal of this section is to relate the gate metrics obtained in Section 6.1.4 and their evolution studied in Section 6.2.1 to the performance of a real device. When executing a quantum circuit on a real device with two qubits, the result is a two-bit string $j_0 j_1$. To relate the results to the ideal, theoretical probabilities $p_{j_0 j_1}^{(\mathrm{id})}$, one has to execute the same circuit repeatedly and estimate the experimental relative frequencies $p_{j_0 j_1}^{(\mathrm{exp})}$ of each bit string $j_0 j_1$ by sampling. The number of samples is called *shots*. For the simple experiments studied in this section and also the more complicated circuits in Chapter 7, we always consider the maximum number of 8192 shots.

To compare two probability distributions $p_{j_0 j_1}$ and $\tilde{p}_{j_0 j_1}$, we use the statistical distance

$$D = \frac{1}{2} \sum_J |p_J - \tilde{p}_J|, \tag{6.22}$$

where the sum is over all two-bit strings $J = j_0 j_1$. In what follows, the distribution p_J is always the result $p_{j_0 j_1}^{(\mathrm{id})}$ expected from an ideal, gate-based quantum computer as introduced in Chapter 2. The other distribution, $\tilde{p}_J$, is either given by relative frequencies $p_{j_0 j_1}^{(\mathrm{exp})}$ measured in an experiment, or probabilities $p_{j_0 j_1}^{(\mathrm{sim})} = |\langle m_0 = j_0, m_1 = j_1 | \Psi \rangle|^2$ obtained from the state vector $|\Psi\rangle$ (see Eq. (3.23)) produced by the transmon simulator.

This distance measure D given by Eq. (6.22) is also known as the *total variation distance*. It can be interpreted as the minimum fraction of samples that must be altered to achieve $\tilde{p}_J \to p_J$. Furthermore, the total variation distance induces the diamond distance $\eta_\diamond$ defined in Section 6.1.2 on the space of quantum channels. See [San2016] for more information.

We perform 20 repetitions of the CNOT gate on the input states $|00\rangle$ and $|10\rangle$ using both a real device and also different pulse schemes on the simulated systems. For the experiment on the real processor, we used qubits Q3 and Q4 of the five-qubit processor that was available on the IBM Q Experience [IBM2016] on August 17, 2017. The CNOT gate was implemented by means of the CR2 pulse defined in Fig. 5.1(c) (see also Fig. 5.3(c)). The RB error rates reported by the processor at the time of execution were 0.0376 for the CNOT gate, 0.0031 and 0.0016 for the single-qubit gates, and 0.033 and 0.06 for the readout errors.

The results of these conceptually simple experiments are shown in Fig. 6.2(a) and (b). We see that on the real processor, the CNOT gate on the state $|00\rangle$ yields a reasonable performance if applied twice, but degrades rapidly for more repetitions. For the initial state $|10\rangle$ shown in Fig. 6.2(b), we see a larger offset of about 0.2 if applied twice; however, the actual error increases only very slightly over repeated applications of the gate.

For the simulations, we first use the small two-transmon system defined in Section 3.4.3 to compare the different pulse implementations CR1, CR2, and CR4 of the CNOT gate
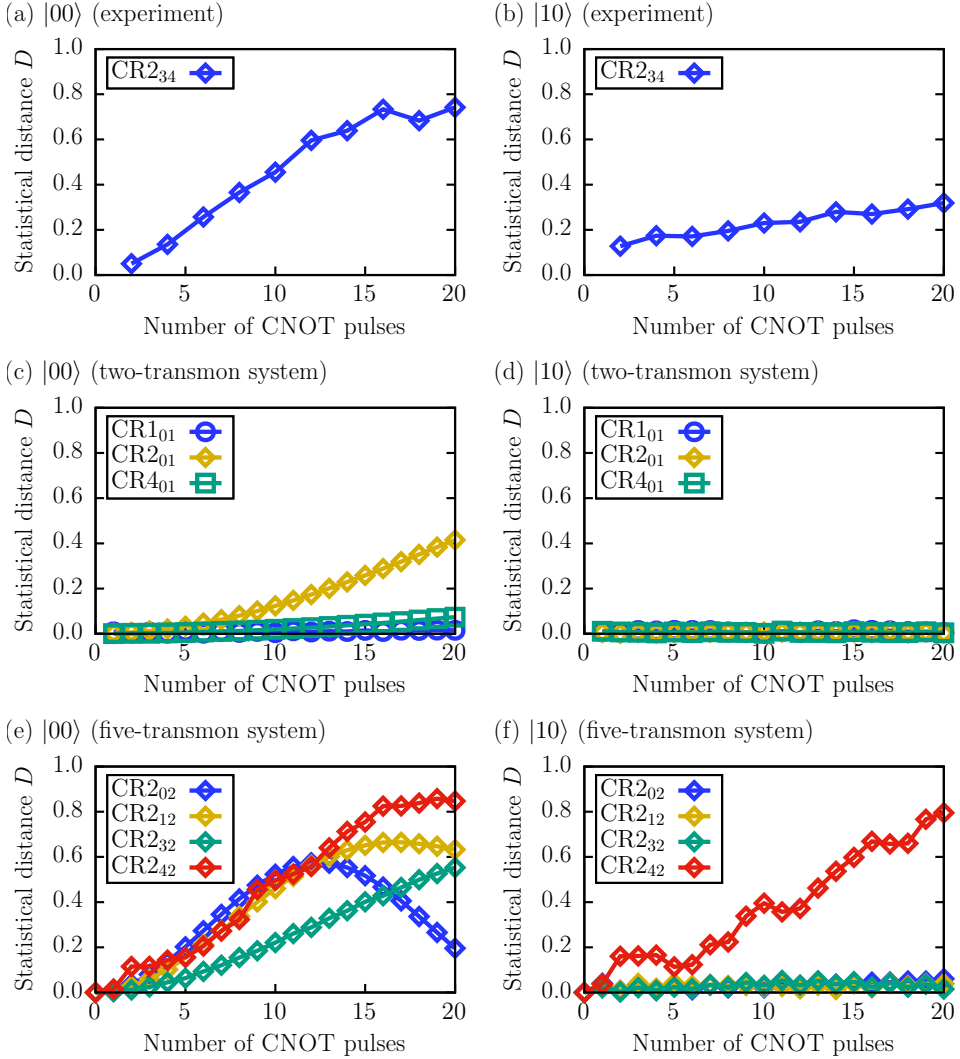
Figure 6.2: Evolution of the statistical distance D given by Eq. (6.22), measured after repeated applications of the two-qubit CNOT gate. (a) and (b) show results from an IBM Q processor (see text), (c) and (d) have been obtained from the two-transmon simulation defined in Section 3.4.3, and (e) and (f) have been obtained from the five-transmon simulation defined in Section 3.4.4. The gate metrics for the simulated systems are given in Tab. 6.1 and Tab. 6.2, respectively. The left (right) panels correspond to the initial state $|00\rangle$ ($|10\rangle$). Different markers correspond to different cross-resonance pulses (see Fig. 5.1). Different colors correspond to pulses between different qubits. All simulations of the five-transmon system were performed on the supercomputer JURECA [Jül2018].

(see Fig. 5.1). The results shown in Fig. 6.2(c) and (d) demonstrate that only the two-pulse echoed CNOT gate, CR2, shows a performance similar to the experiment. This is reasonable as this is the same pulse scheme that is used for the experiment. However, the offset of about 0.2 for the initial state $|10\rangle$ (see Fig. 6.2(b)) cannot be observed in the simulation. Although it can be modeled by including a simple measurement error of 0.2 (in the spirit of Eq. (7.19) in Section 7.3), it is more likely due to environmental effects that are not included in the two-transmon simulation.

Note especially that the other pulse schemes, CR1 and CR4, perform much better than CR2 (see Fig. 6.2(c)). A similar trend has been observed in experiments with CR1 [Ale2020] and CR4 [Tak2017]. This is not at all reflected by the gate metrics shown in Tab. 6.1. In particular, the fidelity F_{avg} is nearly the same for all pulses, despite the strong difference in the actual gate performance. In fact, the only gate with a comparably bad fidelity, $CR1_{01}$, performs unexpectedly well (see Fig. 6.2(c) and (d)). Also the diamond distances shown in Fig. 6.1 do not suggest that. We note that, since the diamond distance is related to the worst-case statistical distance [San2016], it is possible that for another input state, the statistical distances would show a comparable increase for repeated pulses.

In Fig. 6.2(e) and (f), we present simulation results for all CNOT gates on the small five-qubit system defined in Section 3.4.4. Again, the errors are not properly captured by the corresponding gate metrics given in Tab. 6.2. Especially the performance of the pulse $CR2_{42}$ applied to the input state $|10\rangle$ (the red line in Fig. 6.2(f)) degrades quickly, although this pulse scored the best unitarity and the second-best fidelity of all two-qubit gates (see Tab. 6.2). In this case, only the diamond distance suggests the bad performance.

Apart from this, we observe that the performance of the simulated transmon system presented in Fig. 6.2(e) comes much closer to the result of the experiment shown in Fig. 6.2(a). Note that the agreement can only be qualitative because the simulated five-transmon system corresponds to a different device that does not have a CNOT gate between transmons 3 and 4 (cf. Section 3.4.4 and Tab. F.4). Still, the observation that the five-transmon simulation comes much closer to the experiment than the two-transmon simulation gives positive evidence for the hypothesis that the additional features included in the five-transmon simulation, i.e., leakage and crosstalk due to additional transmons and resonators, can capture most of the errors observed in the experiment.

6.3 Gate set tomography

The results presented in the previous sections demonstrate that common gate metrics such as the average fidelity or the diamond norm cannot reliably predict the performance of a sequence of quantum gates (cf. [Iye2018]). It has also been recognized in the literature that the RB-number, i.e., the quantity produced by randomized benchmarking [Pro2017], suffers from the same problems [Blu2017]. The obvious conclusion is that a quantum gate is too complicated to be characterized by a single number.

Therefore, in this section, we explore a much more sophisticated approach to characterize implementations of quantum gates called *gate set tomography* (GST) [Blu2013; Gre2015]. In short, the idea of GST is to treat a quantum device as a black box that takes only digital input and produces only digital output. One then fits the best quan-
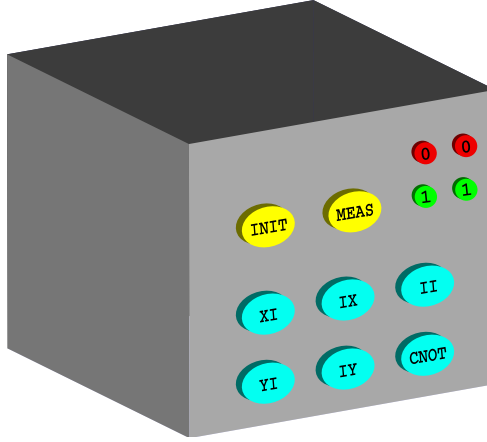
Figure 6.3: Black box model of a quantum computer. For GST, a quantum computer is considered as a black box with only digital input and output, agnostic to the physical details of the implementation. The black box provides buttons for initialization and measurement (yellow), and several gate operations (cyan; see Tab. 6.5 for the meaning of the labels). On each measurement, the black box produces binary output in the form of light signals indicating a bit string. In GST, the best quantum theoretical two-level description is fitted to the frequency of bit strings produced by the black box.

tum theoretical description to experiments performed with the black box. An example black box model for a two-qubit device is shown in Fig. 6.3. A recent application to a quantum processor based on ion traps is presented in [Blu2017], for which an open-source implementation of GST called `pyGSTi` [Nie2018] has been developed. This package has also been used for the GST experiments presented in this thesis.

Our goal is to test the GST procedure for output generated by the transmon simulation developed for this work (see Section 3.3). In particular, the objective is to study whether the resulting quantum theoretical description has a better predictive power than the common gate metrics. As shown below, the answer to this question is affirmative: The two-level description produced by GST can reliably predict the performance of gates when used repeatedly in quantum algorithms.

6.3.1 The idea of GST

In quantum theory, we describe the general state of a system in terms of a density matrix ρ. The evolution of the state is represented by a linear map $\rho \mapsto \mathcal{G}(\rho)$, where $\mathcal{G}$ is typically a quantum operation as introduced in Section 2.4. After the evolution, the system is measured and the probability to obtain outcome J is given by

$$p_J = \text{Tr}\, E_J \mathcal{G}(\rho). \tag{6.23}$$

Here, the set of operators $\{E_J\}$ represents the measurement. The minimum requirements on $\{E_J\}$ are that the E_J are positive semidefinite and $\sum_J E_J = \mathbb{1}$ such that the p_J are

valid probabilities. In the literature, a set $\{E_J\}$ with these properties is called POVM [Nie2010] and Eq. (6.23) is a generalization of the Born rule (see also [Fuc2001; Fuc2002]).

The problem of tomography

The central idea of quantum tomography is to construct, from experimentally observed relative frequencies p_J, an element of the quantum theoretical description given by Eq. (6.23). If the element to reconstruct is the density matrix ρ, one speaks of *quantum state tomography* [Smi1993; Lei1996]. If the objective is to find a representation of the measurement $\{E_J\}$, the procedure is known as *quantum measurement tomography* [Lui1999]. And lastly, characterizing the map $\mathcal{G}$ is called *quantum process tomography* [Chu1997; Poy1997].

All these kinds of tomography have a fundamental flaw, namely that they require the other elements of Eq. (6.23) to be postulated. For instance, quantum process tomography requires the system to be prepared in various known states ρ and measured with various known operators E_J, which are typically implemented using the very same gates $\mathcal{G}$ that process tomography tries to characterize [Blu2013]. In that sense, all three kinds of tomography are self-referential and circular. Note that this problem also affects more recent tomography proposals [Xin2017; Hel2019].

The solution provided by GST

The idea of GST is to solve this problem by self-consistently estimating the state ρ, the process $\mathcal{G}$, and the measurement $\{E_J\}$ of Eq. (6.23) at the same time, using only experimentally observed frequencies p_J. The philosophy is that all elements of the description in Eq. (6.23) should be accessible from the observed data. If the system is prepared in a certain state ρ, the data should reveal it; similarly, if the measurement is in a different basis, the data should suggest it; etc. The central viewpoint is that obtaining enough data by playing around with the black box should be sufficient to predict its future behavior [Blu2013].

To frame this idea, the system is treated as a black box with very limited control. Such a black box for a two-qubit quantum computer is shown in Fig. 6.3. It has several buttons for interaction and two sets of lights signaling the output, but the actual physical implementation is irrelevant. The mapping of the buttons to the mathematical description in Eq. (6.23) is as follows. The yellow `INIT` button is to be described by the mathematical object ρ. Each cyan button is to be described by a map in the set $\{\mathcal{G}_{\texttt{XI}}, \mathcal{G}_{\texttt{YI}}, \mathcal{G}_{\texttt{IX}}, \mathcal{G}_{\texttt{IY}}, \mathcal{G}_{\texttt{II}}, \mathcal{G}_{\texttt{CNOT}}\}$ (the target operation indicated by the indices is given in Tab. 6.5 below). The process $\mathcal{G}$ is then an arbitrary sequence of elements in this set, corresponding to the sequence of buttons pressed. Finally, the yellow `MEAS` button makes the system produce an event $J \in \{00, 01, 10, 11\}$, which is to be described by $\{E_J\}$. An experiment in this model takes a simple form:

1. Press the `INIT` button to initialize the system.

2. Press a sequence of gate buttons from the set $\{\texttt{XI}, \texttt{YI}, \texttt{IX}, \texttt{IY}, \texttt{II}, \texttt{CNOT}\}$.

3. Press the `MEAS` button and record the outcome.

For each experiment corresponding to the same sequence of buttons pressed, the relative frequencies p_J are computed from the outcomes. The best description in terms of ρ, $\mathcal{G}$, and E_J is then obtained from the data.

It is important to recognize that only the frequencies p_J of each event J from the data are used. All other information about the individual events need to be discarded. This approach is typical of all quantum theoretical models, in the sense that quantum theory cannot model the individual events [DeR2019b] or long-time correlations between individual events [Wil2020c].

Mathematical framework

To fit a mathematical description in terms of E_J, $\mathcal{G}$, and ρ to the measured relative frequencies p_J, it is useful to introduce a vector representation of Eq. (6.23). We obtain such a representation by expanding the $N \times N$ matrices E_J and ρ in the normalized Pauli basis $\mathcal{P}$ defined in Eq. (2.26),

$$\rho = \sum_i \rho_i \widehat{P}_i, \tag{6.24a}$$

$$E_J = \sum_i e_{Ji} \widehat{P}_i, \tag{6.24b}$$

where $\rho_i = \text{Tr}\,\widehat{P}_i \rho$ and $e_{Ji} = \text{Tr}\,\widehat{P}_i E_J$. We arrange the coefficients $\{\rho_i\}$ and $\{e_{Ji}\}$ in N^2-dimensional vectors denoted by $|\rho\rangle\!\rangle$ and $|E_J\rangle\!\rangle$, respectively.

In the literature, the vector space corresponding to this vector representation of matrices is sometimes called Hilbert-Schmidt space. The inner product on this space is given by the Hilbert-Schmidt inner product $\langle\!\langle X|Y\rangle\!\rangle = \text{Tr}\,X^\dagger Y$ for complex matrices X, Y.

Inserting Eqs. (6.24a) and (6.24b) into Eq. (6.23) yields

$$p_J = \sum_{ij} e_{Ji} \left(\text{Tr}\,\widehat{P}_i \mathcal{G}(\widehat{P}_j) \right) \rho_j = \langle\!\langle E_J|G|\rho\rangle\!\rangle, \tag{6.25}$$

where the matrix G representing the map $\mathcal{G}$ is the Pauli transfer matrix defined in Eq. (2.27). We use a similar notation to distinguish between the maps $\{\mathcal{G}_{\text{XI}}, \ldots\}$ and their matrix representations $\{G_{\text{XI}}, \ldots\}$.

Due to the use of the Pauli basis, all coefficients in the vectors $|\rho\rangle\!\rangle$ and $|E_J\rangle\!\rangle$ and the matrices G are real. Still, the total number of coefficients for two qubits ($N = 4$) is $16 + 4 \times 16 + 6 \times 16^2 = 1616$ (16 for ρ, 16 for each E_J, and 16^2 for each of the gate matrices $G_{\text{XI}}, G_{\text{YI}}, G_{\text{IX}}, G_{\text{IY}}, G_{\text{II}}$, and G_{CNOT}). This means that a considerable number of experiments needs to be run to obtain enough relative frequencies p_J to determine the coefficients accurately.

These experiments are not chosen arbitrarily. Rather, they are constructed systematically to yield a sufficient amount of information about the coefficients, while still targeting operations such as single-qubit rotations that are commonly used in current quantum information processors [Blu2017]. This construction works as follows: The description of data from an experiment, which corresponds to a sequence $(s_1, \ldots, s_L)$ of pressed cyan

buttons in Fig. 6.3, has the structure

$$p_J^{(s_1,\ldots,s_L)} = \langle\!\langle E_J | \underbrace{G_{s_L}\cdots G_{s_*}}_{F^{(M)}} \underbrace{G_{s_*}\cdots G_{s_*}}_{g^l} \underbrace{G_{s_*}\cdots G_{s_1}}_{F^{(P)}} |\rho\rangle\!\rangle, \tag{6.26}$$

where $F^{(M)}$ is called a measurement fiducial, g^l is called a germ power, and $F^{(P)}$ is called a preparation fiducial. The fiducials are chosen to prepare an *informationally complete* set of states (for instance, the eigenstates of the Pauli matrices, four of which are shown in Fig. 2.2(a)–(d)). The germs g, raised to logarithmically spaced integer powers such as $l \in \{1, 2, 4, 8, 16, 32\}$, are chosen to amplify certain systematic errors expected from the gates. All germs and fiducials are sequences of the six gate matrices $G_{\mathtt{XI}}, G_{\mathtt{YI}}, G_{\mathtt{IX}}, G_{\mathtt{IY}}, G_{\mathtt{II}}$, and $G_{\mathtt{CNOT}}$. They explicitly include empty sequences such that also "bare" experiments corresponding to $\langle\!\langle E_J | F_i^{(M)} F_j^{(P)} |\rho\rangle\!\rangle$ and, in particular, $\langle\!\langle E_J |\rho\rangle\!\rangle$ are represented by the data. See [Blu2013; Blu2017] for more information on these design choices.

When describing experimentally obtained data $\{p_J^{(s_1,\ldots,s_L)}\}$, the mathematical objects contained in Eq. (6.26) can only be determined up to an invertible matrix M. The reason for this is that the $p_J^{(s_1,\ldots,s_L)}$ stay the same if all objects in Eq. (6.26) are replaced according to $\langle\!\langle E_J | \mapsto \langle\!\langle E_J | M^{-1}$, $G \mapsto MGM^{-1}$, and $|\rho\rangle\!\rangle \mapsto M|\rho\rangle\!\rangle$. This fundamental freedom in GST is called the *GST gauge*.

There are several ways to fix a certain gauge transformation M. For instance, if one requires the description of the state ρ to be independent of the operations that are performed on it (as typically required by consistency in quantum theoretical descriptions, see [DeR2019b]), one could choose a gauge M that always maps the fitted ρ to some fixed representation. In general, it can be problematic to fix a certain gauge and interpret the resulting, gauge dependent error rates (see also [Lin2019]). For the following analysis, however, the goal is to predict the $p_J^{(s_1,\ldots,s_L)}$, so the choice of the gauge is irrelevant. Therefore, we use the default gauge optimization performed by `pyGSTi`.

The actual self-consistent fitting procedure involves several steps. The first step is to construct a matrix `AB`, where $\mathtt{AB}_{J+Ni,j}$ contains the relative frequency corresponding to $\langle\!\langle E_J | F_i^{(M)} F_j^{(P)} |\rho\rangle\!\rangle$, i.e., obtained from the "bare" experiments. The (pseudo-)inverse of this matrix yields initial estimates for ρ, E_J, and G (this step requires the fiducials to be be informationally complete). This part of the procedure is called *linear inversion GST* and is described in detail in [Blu2013; Gre2015]. A difference between the single-qubit GST described in these references and the two-qubit GST studied here is that all data for $J \in \{00, 01, 10, 11\}$ is included in `AB`. The full iterative procedure to refine the initial estimates self-consistently is explained in [Blu2017] (see Fig. 1 of this work).

We remark that, because of the vast number of parameters to be estimated in GST, it would be interesting to explore techniques from deep learning to improve the fitting procedure. These techniques have specifically been invented to fit statistical models with millions of parameters to observed data [Goo2016].

6.3.2 Running GST

We apply GST to the two-transmon simulation model defined in Section 3.4.3. This requires 58990 quantum circuits (see below) to be compiled into pulses (cf. Section 5.4)

Table 6.5: Summary of the pulses used to implement the gates corresponding to the buttons of the black box shown in Fig. 6.3 by means of the two-transmon simulation model defined in Section 3.4.3. The general single-qubit GD pulses are defined in Section 5.1.2. The implementation of $Y^{\pi/2} = $ `-Y` in particular is derived in Eqs. (5.27) and (5.28). The identity gate is implemented as an undriven time evolution for $T_X = 83\,\text{ns}$, i.e., the duration of the single-qubit pulses (see Section 5.1.3). The two-qubit CNOT gate is implemented using the CR2 pulse (see Section 5.2). All pulse parameters are listed in Appendix F, and the target gates are defined in Tab. B.1. Gate metrics for the elementary pulses are given in Tab. 6.1.

Button	Pulse	Target gate	Description
XI	$\text{GD}_0^{\pi/2}(0)$	$X_0^{\pi/2}$	$\pi/2$ rotation of qubit 0 about the x axis
YI	$\text{GD}_0^{\pi/2}(\pi/2)$	$Y_0^{\pi/2}$	$\pi/2$ rotation of qubit 0 about the y axis
IX	$\text{GD}_1^{\pi/2}(0)$	$X_1^{\pi/2}$	$\pi/2$ rotation of qubit 1 about the x axis
IY	$\text{GD}_1^{\pi/2}(\pi/2)$	$Y_1^{\pi/2}$	$\pi/2$ rotation of qubit 1 about the y axis
II	$\text{zero}(T_X)$	I	identity gate
CNOT	CR2	CNOT_{01}	CNOT between qubit 0 (control) and 1 (target)

and simulated by solving the TDSE. From the final state vector $|\Psi\rangle$, only the relative frequencies for the two-qubit states $|00\rangle, |01\rangle, |10\rangle$, and $|11\rangle$ are extracted. In this way, the simulation represents the implementation that is hidden inside the black box shown in Fig. 6.3. The particular pulses used to implement the cyan buttons are summarized in Tab. 6.5.

We use the software package `pyGSTi` [Nie2018] to produce the input (the quantum circuits) to the black box and to analyze the output (the relative frequencies p_J). The list of simulated quantum circuits comprises 58990 gate sequences with a maximum of 38 gates per sequence. `pyGSTi` respects the structure of Eq. (6.26), using 16 preparation fiducials $F^{(P)}$, 11 measurement fiducials $F^{(M)}$, and 89 germs g raised to the power $l \in \{1, 2, 4, 8, 16, 32\}$. Since the output p_J is passed to `pyGSTi` in terms of actual integer counts for each event $J \in \{00, 01, 10, 11\}$, we compute counts for 1000 samples. This number reflects the expected accuracy of the product-formula algorithm for $\tau = 10^{-3}\,\text{ns}$ (cf. Fig. 4.2(b)). Furthermore, `pyGSTi` offers several modes for estimating the objects ρ, $\mathcal{G}$, and E_J in Eq. (6.23):

Full: Fully unconstrained maps,

TP: Constrain the maps to be trace-preserving quantum operations (cf. Section 2.4),

CPTP: Constrain the maps to be CPTP quantum channels (cf. Section 2.4).

In this section, we report the CPTP estimates, because they proved to be the most reliable estimates when trying to predict the performance of quantum gate circuits (see below). The estimated initial density matrix reads

$$\rho = |00\rangle\langle 00| + .009(|00\rangle\langle 01| + |01\rangle\langle 00|) + \mathcal{O}(10^{-3}), \tag{6.27}$$

and the estimated operators describing the measurement are

$$E_{00} = |00\rangle\langle00| + .009(|00\rangle\langle01| + |01\rangle\langle00|) + \mathcal{O}(10^{-3}), \tag{6.28a}$$

$$E_{01} = |01\rangle\langle01| - .009(|00\rangle\langle01| + |01\rangle\langle00|) + \mathcal{O}(10^{-3}), \tag{6.28b}$$

$$E_{10} = |10\rangle\langle10| + \mathcal{O}(10^{-3}), \tag{6.28c}$$

$$E_{11} = |11\rangle\langle11| + \mathcal{O}(10^{-3}). \tag{6.28d}$$

Up to the small off-diagonal component in ρ, E_{00}, and E_{01}, these estimates match the expected result given by the first term. The small deviation may be a consequence of the gauge transformation found by `pyGSTi`, since we still have $\mathrm{Tr}\, E_{00}\rho = 1 + \mathcal{O}(10^{-4})$ and $\mathrm{Tr}\, E_{01}\rho = 0 + \mathcal{O}(10^{-4})$.

Visualization of the Pauli transfer matrices

In Fig. 6.4, we show the resulting estimates for the gate processes $\mathcal{G}$ in terms of their Pauli transfer matrices G (see Eq. (2.27)), which consist of 16×16 real numbers in the range $[-1, 1]$. One thing to note is that for all matrices G, the first row $i = 0$ and the first column $j = 0$ (both corresponding to the axis label `II`) contain only one non-zero entry at $G_{00} = 1$. This reflects the fact that the estimated maps are both trace-preserving and unital.

For each gate, the expected target gate would be represented by the same image with all small blue and yellow bars replaced by gray areas such that only one large bar occurs in each row and each column in Fig. 6.4. This reflects the fact that the target gates are Clifford gates. The result shows that the operations performed by the pulses are best described by non-Clifford gates.

The largest deviations (i.e., the largest bars that should have been gray areas) can be observed for the identity gate G_{II} shown in Fig. 6.4(e) and the CNOT gate shown in Fig. 6.4(f). For the identity gate, the deviations occur systematically on the antidiagonal. This effect is analyzed in the following section. It corresponds to an intrinsic error of the form $\sigma_0^z \otimes \sigma_1^z$. A related behavior can also be observed in experiments on the IBM Q processors (see Section 7.1).

Axis-angle decompositions of the estimated gates

To gain further insight into the GST results, we apply a decomposition algorithm to express the gate maps $\mathcal{G}$ in terms of a Hamiltonian generator,

$$\mathcal{G}(\rho) \approx e^{-iH} \rho\, e^{iH}. \tag{6.29}$$

Details on the decomposition algorithm are explained in Appendix J. A similar algorithm is implemented by the `pyGSTi` package [Nie2018]. Alternative approaches to find effective Hamiltonians have been studied in [Ric2013; Wil2016; Wil2020d].

We consider a Hamiltonian H expressed in the Pauli basis (see Eq. (2.26)),

$$H = \sum_{k=0}^{d-1} h_k P_k/2, \tag{6.30}$$

(a) $G_{\mathtt{XI}}$

(b) $G_{\mathtt{YI}}$

(c) $G_{\mathtt{IX}}$

(d) $G_{\mathtt{IY}}$

(e) $G_{\mathtt{II}}$

(f) $G_{\mathtt{CNOT}}$
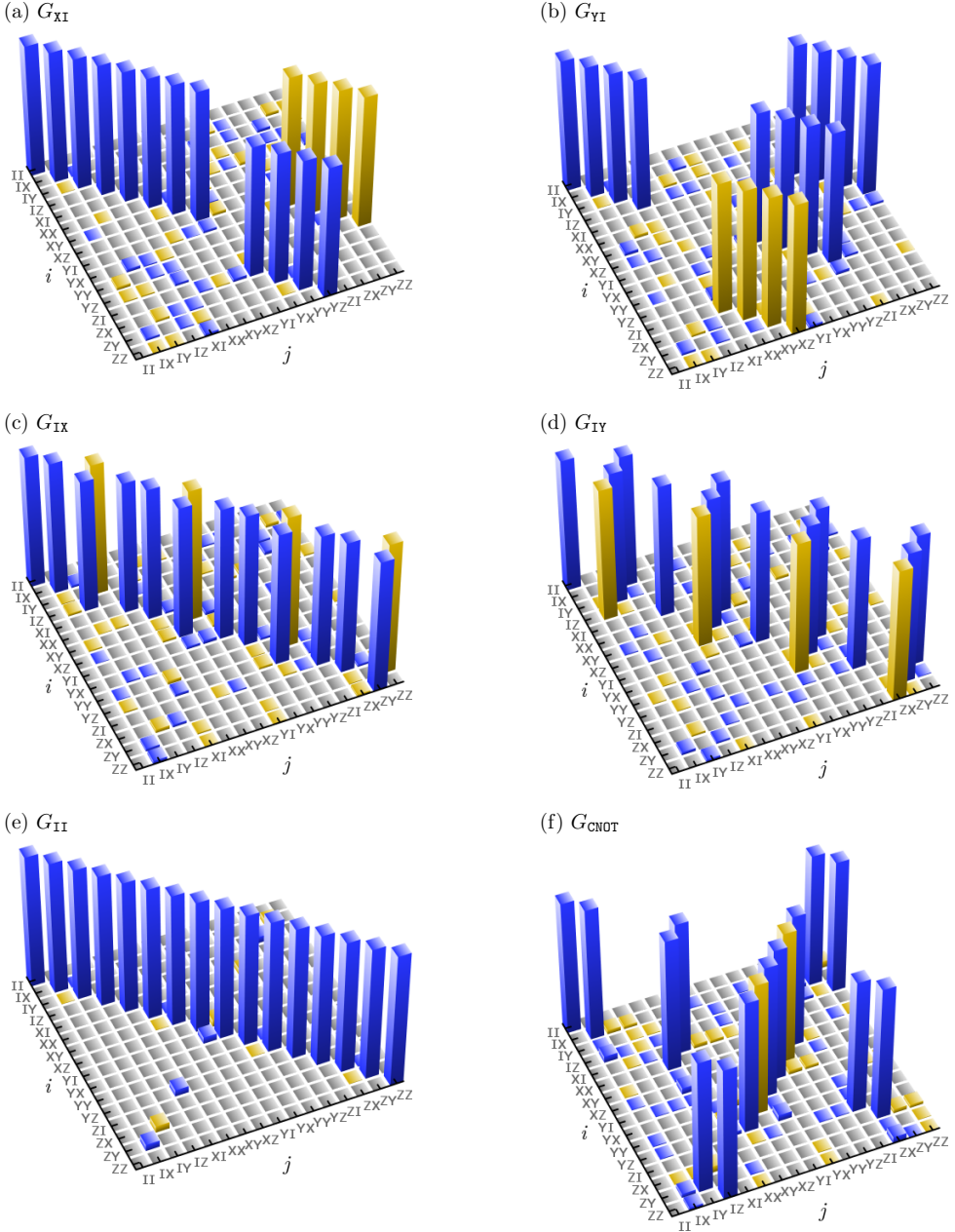
Figure 6.4: Estimated gates resulting from GST experiments on a two-transmon simulation of 58990 circuits implemented by pulses. Shown are the Pauli transfer matrices G_s (see Eq. (2.27)), where the subscript s indicates the button of the black box model shown in Fig. 6.3. Blue (yellow) bars indicate positive (negative) values. Gray areas indicate absolute values below 10^{-3}. All large bars represent values close to ± 1.

132

where $d = N^2 = 4^n$ and $h_k \in \mathbb{R}$. The reason for this parametrization is that given h_k, the action of e^{-iH} can be interpreted as a rotation characterized by an angle φ and an axis $\widehat{h}$ according to

$$\varphi = \sqrt{\sum_k h_k^2}, \tag{6.31a}$$

$$\widehat{h}_k = \frac{h_k}{\sqrt{\sum_k h_k^2}}. \tag{6.31b}$$

For instance, an ideal $\pi/2$ rotation about the x axis for qubit 1 ($X_1^{\pi/2}$) corresponds to $\varphi = \pi/2$ and $\widehat{h}_k = \delta_{k1}$ since $P_1 = I \otimes \sigma^x$ (cf. Eq. (2.26)).

The results of the axis-angle decomposition are given in Tab. 6.6. For the single-qubit gates $\mathcal{G}_{\mathtt{XI}}$, $\mathcal{G}_{\mathtt{YI}}$, $\mathcal{G}_{\mathtt{IX}}$, and $\mathcal{G}_{\mathtt{IY}}$, GST almost perfectly reproduces the $\pi/2$ rotations that the GD pulses have been designed to implement. Furthermore, the decomposition errors γ are reasonable small, and the computed diamond distances $\eta_\lozenge$ agree with the gate metrics reported in Tab. 6.1.

The most interesting results are given by the decompositions of the identity gate $\mathcal{G}_{\mathtt{II}}$ and the CNOT gate. For the identity gate, GST extracted, without prior knowledge, the exact same type of $\sigma_0^z \otimes \sigma_1^z$ interaction studied in Fig. 4.11. From the parameters given in Tab. 6.6, we can also compute the interaction strength of the effective Hamiltonian H_{ZZ} introduced in Eq. (4.28),

$$J^{\mathrm{GST}} = \frac{\varphi \widehat{h}_{15}}{2T_X} = 2\pi \times 46.6\,\mathrm{kHz}, \tag{6.32}$$

where $T_X = 83\,\mathrm{ns}$ is the duration of the identity gate, and $\widehat{h}_{15} = 0.9973$ is the ZZ entry of the axis $\widehat{h}$. This result perfectly matches the result given by Eq. (4.31c). Note that this ZZ effect is still qualitatively compatible with $\mathcal{G}_{\mathtt{II}}$ implementing an identity gate since the angle $\varphi = 0.0155\pi$ is close to 0. In Section 7.1, we construct a circuit to observe related crosstalk effects in an IBM Q processor.

For the estimated CNOT gate $\mathcal{G}_{\mathtt{CNOT}}$, the axis-angle decomposition reported in Tab. 6.6 yields an effective Hamiltonian of the form

$$\frac{1}{2}0.865\pi(.58[\sigma_0^z + \sigma_1^x - \sigma_0^z\sigma_1^x] - .01(\sigma_1^y - \sigma_1^z + \sigma_0^z\sigma_1^y) + \mathcal{O}(10^{-3})). \tag{6.33}$$

This Hamiltonian agrees very well with the effective Hamiltonian expected from the CR pulse (see Eq. (5.13)). Furthermore, the next-order terms of the form IY, IZ, and ZY resemble the contributions observed experimentally in [She2016b] (see also Fig. 5.12 in [Wil2016]). It is remarkable that GST reliably resolves all these effects given only the measured relative frequencies p_J, without any prior knowledge about the transmon dynamics such as the CR pulse used internally to implement the gate.

However, one should also be careful not to put too much trust in the decomposition of $\mathcal{G}_{\mathtt{CNOT}}$, since in this case, the decomposition error γ in Tab. 6.6 is larger than for the other gates. One reason for this could be that a minimal Kraus representation of $\mathcal{G}_{\mathtt{CNOT}}(\rho)$

Table 6.6: Axis-angle decompositions and errors of the estimated gates $\mathcal{G}$. The decompositions are obtained from the Hamiltonian given by Eq. (6.30), from which the angle φ and the axis $\widehat{h}$ are extracted via Eqs. (6.31a) and (6.31b). The notation "$x\mathtt{m}y$" means $x \times 10^{-y}$. Blue (yellow) colors highlight significant positive (negative) coefficients. The decomposition error γ measures the error of the approximation given by Eq. (6.29) (see Eq. (J.11) for the precise definition). The target gates are defined in Tab. B.1. The target error $\eta_\diamond$ is the diamond distance between the full estimated map $\mathcal{G}$ (not only its axis-angle decomposition) and the target gate (cf. Section 6.1.2).

Gate	Angle φ	Axis $\widehat{h}$				Error γ	Target	Target error $\eta_\diamond$

$\mathcal{G}_{\mathtt{XI}}$, Angle $\varphi = 0.5001\pi$, Error $\gamma = 2.8 \times 10^{-5}$, Target $X_0^{\pi/2}$, Target error $\eta_\diamond = 0.023$

Axis $\widehat{h}$	I	X	Y	Z
I	0	7m5	2m4	−2m3
X	**1.0**	5m5	−3m5	6m4
Y	2m4	−.01	.01	−3m3
Z	4m4	1m4	−3m3	.02

$\mathcal{G}_{\mathtt{YI}}$, Angle $\varphi = 0.5002\pi$, Error $\gamma = 2.0 \times 10^{-5}$, Target $Y_0^{\pi/2}$, Target error $\eta_\diamond = 0.023$

Axis $\widehat{h}$	I	X	Y	Z
I	0	8m5	3m4	−2m3
X	2m4	.01	−.01	3m3
Y	**1.0**	1m4	8m5	6m4
Z	−2m4	1m4	3m3	.02

$\mathcal{G}_{\mathtt{IX}}$, Angle $\varphi = 0.5001\pi$, Error $\gamma = 2.1 \times 10^{-4}$, Target $X_1^{\pi/2}$, Target error $\eta_\diamond = 0.029$

Axis $\widehat{h}$	I	X	Y	Z
I	0	**1.0**	.01	−.01
X	2m4	−6m5	−.01	−3m3
Y	2m4	−1m4	.01	−3m3
Z	8m5	5m4	−3m3	.02

$\mathcal{G}_{\mathtt{IY}}$, Angle $\varphi = 0.5001\pi$, Error $\gamma = 1.2 \times 10^{-5}$, Target $Y_1^{\pi/2}$, Target error $\eta_\diamond = 0.022$

Axis $\widehat{h}$	I	X	Y	Z
I	0	−.01	**1.0**	3m7
X	2m4	.01	9m5	9m5
Y	2m4	−.01	3m5	5m4
Z	6m5	3m3	5m5	.02

$\mathcal{G}_{\mathtt{II}}$, Angle $\varphi = 0.0155\pi$, Error $\gamma = 2.4 \times 10^{-5}$, Target I, Target error $\eta_\diamond = 0.026$

Axis $\widehat{h}$	I	X	Y	Z
I	0	3m3	.01	−.07
X	.01	5m4	−.01	1m3
Y	.01	5m3	1m3	5m5
Z	1m3	.01	−3m3	**1.0**

$\mathcal{G}_{\mathtt{CNOT}}$, Angle $\varphi = 0.8655\pi$, Error $\gamma = 4.8 \times 10^{-2}$, Target CNOT_{01}, Target error $\eta_\diamond = 0.041$

Axis $\widehat{h}$	I	X	Y	Z
I	0	**.58**	−.01	.01
X	8m5	−5m4	−3m3	3m3
Y	−4m6	2m4	3m3	4m3
Z	**.58**	**−.58**	−.01	2m3

(cf. Eq. (2.24)) needs more than one term such that Eq. (6.29) is not easily achievable. However, we examined the corresponding Kraus rank by studying the singular values of the Choi matrix $J(\mathcal{G}_{\texttt{CNOT}})$ (see Eq. (2.25)) and found that all but one singular value are smaller than 3×10^{-4}. Instead, what happens in this case is that the corresponding Lindblad operator $\mathcal{L} = \log \mathcal{G}_{\texttt{CNOT}}$ does not exactly have the form required to write it in terms of a Hamiltonian (see Eq. (J.9) in Appendix J for more information).

Intuitively, this decomposition error expresses the lingering effect of non-computational levels on the result after the pulse application. This additional information, which is properly captured by the GST result visualized in Fig. 6.4(f), is essential to reliably predict the effect of many repeated pulse applications. This is demonstrated in the following section.

6.3.3 Predicting repeated pulse applications

To test the predictive power of GST, as opposed to that of the gate metrics studied in Section 6.2, we simulate $r = 1, \ldots, 1000$ repeated gate pulses. We compute the statistical distance D given by Eq. (6.22), where p_J is the distribution for an ideal, gate-based quantum computer $p_J^{(\text{id})}$, and $\tilde{p}_J$ is either given by $p_{j_0 j_1}^{(\text{sim})} = |\langle m_0 = j_0, m_1 = j_1|\Psi\rangle|^2$, obtained from the transmon simulation model defined in Section 3.2, or by a GST prediction $p_J^{(\text{GST})}$.

The distribution predicted by GST is computed through Eq. (6.23). In vector form, this means that

$$p_J^{(\text{GST})} = \langle\!\langle E_J|(G_s)^r G_{(\text{prep})}|\rho\rangle\!\rangle, \tag{6.34}$$

where $|\rho\rangle\!\rangle$ is the vectorized initial density matrix (see Eq. (6.24a)), $G_{(\text{prep})}$ denotes the gates used to prepare a certain initial state, $(G_s)^r$ represents r repetitions of a certain gate $s \in \{\texttt{XI}, \texttt{YI}, \texttt{IX}, \texttt{IY}, \texttt{II}, \texttt{CNOT}\}$, and $\langle\!\langle E_J|$ is a vectorized measurement operator (see Eq. (6.24b)).

Note that also the initial state preparation $G_{(\text{prep})}$ explicitly uses the GST estimated gates G_s, because the purpose of GST is to self-consistently describe all preparation, gate, and measurement processes. We test four particular initial states corresponding to the following gate sequences:

$$|00\rangle : \qquad G_{(\text{prep})}|\rho\rangle\!\rangle = |\rho\rangle\!\rangle, \tag{6.35a}$$

$$|10\rangle : \qquad G_{(\text{prep})}|\rho\rangle\!\rangle = G_{\texttt{XI}}G_{\texttt{XI}}|\rho\rangle\!\rangle, \tag{6.35b}$$

$$|{+}{+}\rangle : \qquad G_{(\text{prep})}|\rho\rangle\!\rangle = G_{\texttt{YI}}G_{\texttt{IY}}|\rho\rangle\!\rangle, \tag{6.35c}$$

$$|\Psi^-\rangle : \qquad G_{(\text{prep})}|\rho\rangle\!\rangle = G_{\texttt{CNOT}}G_{\texttt{YI}}G_{\texttt{XI}}G_{\texttt{XI}}G_{\texttt{IX}}G_{\texttt{IX}}|\rho\rangle\!\rangle, \tag{6.35d}$$

where $|\Psi^-\rangle \propto |01\rangle - |10\rangle$ is the singlet state, and G_s denotes the Pauli transfer matrix of the gate map $\mathcal{G}_s$ (see Eq. (2.27)). For the transmon simulation, the relation between gates and the pulses used to implement them is given in Tab. 6.5.

We test two particular GST predictions. The first prediction is based on the CPTP estimate, for which the preparation ρ and the measurements E_J are given by Eqs. (6.27)–(6.28d), and the gates G_s correspond to the Pauli transfer matrices shown in Fig. 6.4. The second prediction uses, for each gate G_s, its corresponding decomposition in terms

of an effective Hamiltonian given in Tab. 6.6 (see also Appendix J). Besides the CPTP estimate, we also analyzed the "Full" estimate and the TP estimate introduced above. However, the CPTP estimates proved to be the most reliable (data not shown).

The results of the simulation, the GST prediction, and the decomposed GST prediction are shown in Fig. 6.5. The overall agreement between the simulation and the GST prediction up to $r = 500$ repetitions is remarkable, especially since only sequences of up to 38 gates of the form of Eq. (6.26) have been used to run GST. Evidently, the predictive power slightly decreases for $r = 1000$ pulse applications, but it remains sufficient. In general, it is far superior to the predictive power of the gate metrics, which in some cases even fail to predict the performance for two successive pulse applications (cf. Fig. 6.1).

There are tiny oscillations in the simulation results that are most apparent on the red lines in Fig. 6.5, corresponding to results for the initial state $|\Psi^-\rangle$. These oscillations are smoothed out in the corresponding GST predictions. Furthermore, small deviations in amplitude between simulation and GST prediction can be seen for the identity gate and the CNOT gate (last two rows in Fig. 6.5). These effects stem from entanglement with non-computational states during the time evolution which are not completely captured by the GST model. They are most pronounced for the state $|\Psi^-\rangle$ because of the CNOT used in the state preparation (cf. Eq. (6.35d)). Still, the overall performance is sufficiently well described by the GST results.

The right column in Fig. 6.5 might suggest that a similar conclusion is appropriate for the predictive power of the effective Hamiltonians given in Tab. 6.6, obtained from the axis-angle decompositions. Indeed, for all but the CNOT gate, the corresponding decomposition error γ is quite low. However, for the CNOT gate, we see that the periods of the blue, green, and red lines in the bottom-right panel of Fig. 6.5 are wrong. This means that already for $r = 20$ repetitions, the prediction can be wrong. Specifically, for the initial state $|00\rangle$ (blue line), we have $D_{\texttt{CNOT}} \approx 0.41$ for the simulation and the GST prediction, while $D_{\texttt{CNOT}} \approx 0.28$ for the decomposed GST prediction. Hence the decomposition given by Eq. (6.29) is not suitable to obtain reliable predictions for the two-qubit gate.

From the GST results, another interesting conclusion about the accuracy of the simulation algorithm can be drawn. For $D_{\texttt{CNOT}}$ (the last row in Fig. 6.5), $r = 1000$ repetitions correspond to a simulation of the time evolution over 433 µs. Using a time step of $\tau = 10^{-3}$ ns for the product-formula algorithm (cf. Eq. (3.30)), this corresponds to more than 4×10^8 time steps. GST, in contrast, uses only data from at most 38 gate pulses, corresponding to a time evolution for at most 14 µs. The agreement between simulation results and the GST prediction also after such a long time evolution gives confidence in the accuracy of the product-formula algorithm in actual applications. This was not obvious from the error analysis of the simulated system (see Fig. 4.2(b)).

In summary, GST and the underlying black box philosophy provide a reliable model to effectively describe and predict quantum gate applications in complicated two-transmon systems. It solves the problem of circular reasoning present in many alternative proposals of tomography by self-consistently fitting descriptions of the preparation, the gates, and the measurement to the data. Of course, one may argue where the state preparation procedure actually starts (see [Bal1998]), and to what extent repeated pulses really generate
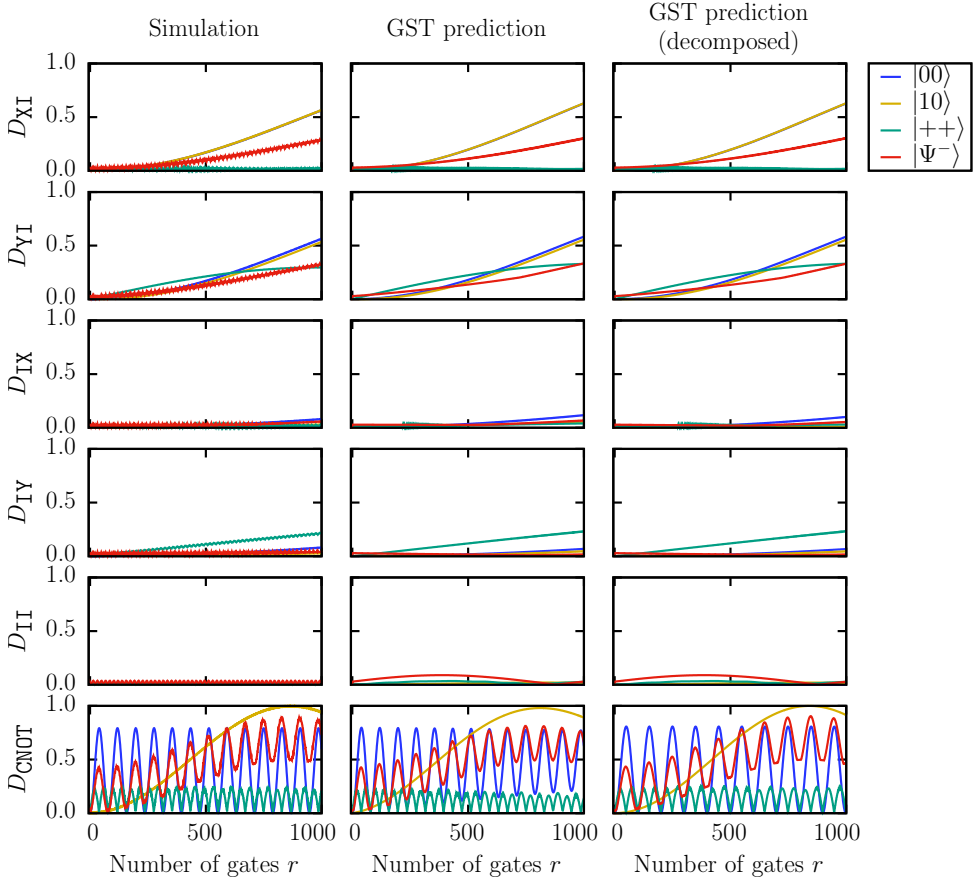
Figure 6.5: Test of the predictive power of GST for up to 1000 gates implemented through pulses. Shown is the statistical distance D_s to the ideal result (cf. Eq. (6.22)), where the subscript s indicates the gate (see Tab. 6.5), corresponding to a certain button of the black box model shown in Fig. 6.3. The left column shows simulation results for the transmon model described in Section 3.4.3. The middle column shows the GST prediction computed through Eq. (6.34). The right column shows the same except that not the full estimates $\mathcal{G}_s$, but their Hamiltonian axis-angle decompositions are used (see Tab. 6.6). Different colors indicate different initial states that are also prepared using the estimated gates, as stated in Eqs. (6.35a)–(6.35d). For all but $D_{\texttt{CNOT}}$, only every second data point is plotted (otherwise, there would be small oscillations on some of the lines that would render them too thick to be distinguishable). Note that the first 20 points on the blue and yellow lines in the bottom-left panel correspond to the yellow lines in Fig. 6.2(c) and (d).

the same time evolution or if non-Markovian effects dominate the system. Furthermore, GST is not scalable to arbitrarily sized qubit systems due to the exponential growth of the mathematical objects required in the description. However, efforts at applying the idea of GST to systems with more than two qubits are being explored [Nie2017; Son2019; Gov2020]. And most importantly, GST shows that it is possible to obtain reliable descriptions for gate-based quantum computers by only pushing well-chosen sequences of buttons on a black box and analyzing the digital output.

6.4 Conclusions

The purpose of this chapter was to characterize the performance of quantum gate pulses in transmon systems. We first studied common gate metrics such as the average gate fidelity given by Eq. (6.2), the diamond distance given by Eq. (6.4), and the unitarity given by Eq. (6.19). Two byproducts of the analysis were (1) an explicit expression for the fidelity proven in Appendix G and (2) a new lower bound on the diamond distance proven in Appendix I; both of which apply, unlike previous results, also to trace-decreasing quantum operations.

From the evaluated gate metrics, we concluded that the errors are systematic and inherently different from simple Pauli-type errors (compare $\eta_\diamond$ and $\eta_\diamond^{\mathrm{Pauli}}$ in Tab. 6.1). Furthermore, we found that a large part of the reduction in fidelity can be attributed to non-unitary evolutions of (and in particular, leakage out of) the computational subspace (compare F_{avg} and u in Tab. 6.2 and Tab. 6.3). Regarding the non-unitarity, repeated applications showed that the optimized gates have been tuned to the participation of higher transmon states in the time evolution (see the final remark in Section 6.2.1).

Although the gate metrics help in characterizing the errors, we found that none of the metrics are suitable for predicting the performance in repeated pulse applications. For instance, the pulse CR4$_{01}$ has the worst diamond distance (see Tab. 6.1), but performs best in repeated applications (see Fig. 6.1(d)). Also, the gate $X_3^{\pi/2}$ has by far the best gate metrics of all optimized transmon pulses (see Tab. 6.2), but the resulting Bloch vector is still more tilted than for $X_1^{\pi/2}$ after repeated applications (see Fig. A.1 in Appendix A). The conclusion is that the gate metrics cannot reliably predict the performance in practical applications, and that quantum gate pulses are generally much too complicated to be characterized by a single number.

By evaluating error rates for experiments on an IBM Q processor, we found that the CNOT gate performs much worse when the control qubit is in state $|0\rangle$. In the simulation, this systematic error could also be observed for the same CR2 pulse that was used in the experiment (see Fig. 6.2(c)). The optimized CR1 and CR4 pulses shown in Fig. 5.1, however, did not suffer from the same systematic error. Furthermore, we found that only by simulating a larger system with five transmons and two resonators, we could observe similar deviations from the ideal result in both simulation and experiment (see Fig. 6.2(e)). This implies that the crosstalk between transmons, which is inherently part of the full transmon system's time evolution, is responsible for most of the errors (see also the explicit crosstalk experiment in Section 7.1).

Finally, by performing an extensive GST, we found that the resulting CPTP estimates provide a much more accurate, effective discrete description of the evolution generated by the optimized gate pulses. The estimates were obtained only from the relative frequencies observed in a black box model of the system. They showed that the implemented quantum gates, unlike the intended target gates, are best described by non-Clifford operations (see Fig. 6.4). We found that GST was capable of reproducing the exact same ZZ interaction during the identity gate (see Eq. (6.32)) that was also found in Section 4.4. Secondly, it accurately reproduced the effective evolution expected from the CR pulses without prior knowledge of the theory (see Eq. (6.33)). And lastly, the GST results exhibited an exceptional predictive power for up to 1000 pulse applications (see Fig. 6.5) that is far superior to the conventional gate metrics.

Chapter 7

Selected quantum circuit experiments

In this chapter, we combine the results from the previous chapters, i.e., the transmon simulator developed in Chapter 3 and benchmarked in Chapter 4, the optimized gate pulses described in Chapter 5, and the actual quantum gates characterized in Chapter 6, and apply them to several selected classes of quantum circuits.

In principle, some more or less complicated circuits have already been simulated for the results of the previous chapters (e.g., the gate set tomography results in Section 6.3 required 58990 quantum circuits with time evolutions up to 433 µs).

Unlike the previous chapters, however, the present chapter focuses entirely on relating the simulation results to experiments on quantum processors such as the `ibmqx4` [IBM2018b] or the `ibmqx5` [IBM2018a] which are available on the IBM Q Experience [IBM2016].

In particular, we consider:

(a) an observation of crosstalk and the induced systematic errors;

(b) a characterization of the singlet state $(|01\rangle - |10\rangle)/\sqrt{2}$;

(c) a test of a protocol from the theory of quantum fault tolerance.

For each of these experiments, we run the corresponding quantum circuits on a quantum processor and compare the results to data produced by the transmon simulator. The experiment corresponding to (a) consists of a particular family of circuits, inspired by the systematic effects observed in the simulation (see Section 4.4). For this experiment, we obtain almost perfect agreement between the simulator and an IBM Q processor (see Section 7.1). The results for (b) extend previous work published in [Mic2017; Wil2017], where the agreement between experiment and simulation was only qualitative. We consider several alternatives such as modified pulse parameters or effective error models for the environment (see Section 7.2). Finally, in Section 7.3, we implement a protocol from the theory of quantum error correction and fault tolerance. We find that the protocol provides a systematic way to improve the results in both the simulation and the real processor, suggesting that the dominant errors are of the same nature. Some of these results have previously been published in [Wil2018b].

7.1 Crosstalk experiments

In typical quantum computer experiments, individual qubits tend to interact even if no gate is performed between them. This kind of *crosstalk* is an always-on coupling that is inherently part of the free evolution of the quantum system. In this section, we examine such crosstalk using both the five-transmon model defined in Section 3.4.4 and the five-qubit processor `ibmqx4` of the IBM Q Experience [IBM2018b]. We study a particular circuit designed to amplify the interaction and compare results from the simulation with data obtained from the experiment.

Crosstalk effects during the free evolution of a two-transmon system have been studied in Section 4.4. In this simple case, the interaction resulted in state-dependent frequency shifts (see Fig. 4.11). They could be described in terms of an effective ZZ interaction of the form of Eq. (4.28). In Section 6.3, we obtained the same effective two-level description from the black box model of GST (see Eq. (6.32)).

However, the five-transmon system defined in Section 3.4.4 has many additional states in the transmons and the resonators that take part in the full time evolution. Furthermore, the interaction topology between the qubits is significantly more complicated (see Fig. 3.4). For this reason, it may be hard to derive a similarly simple, effective model analytically, and GST is not easily doable anymore. Therefore, it seems a good opportunity to compare the simulation results to an experiment based on the same architecture. Obviously, this requires a certain family of quantum circuits to characterize the effect.

7.1.1 Circuit and simulation results

In the free evolution studied in Section 4.4, the resonator-mediated interaction caused the uniform superposition $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ of a qubit to evolve differently depending on the state of its neighboring qubit. This is also the reason that the interaction could be quantified accurately in the black box model of GST (see Section 6.3). Therefore, we study a family of quantum circuits that prepare one qubit in the state $|+\rangle$ and characterize its free evolution as a function of the state of the other qubits. The general form of the circuits is given in Fig. 7.1.

Conceptually, the circuit is very simple. The first part of the circuit prepares the system in a product state, where the state of qubit 0 is given by $|+\rangle$. In the simulation, this is implemented using the gate $H = \mathrm{U2}(0, \pi) = Z^{\pi/2} X^{\pi/2} Z^{\pi/2}$, which is compiled into a single-qubit GD pulse according to Eqs. (5.9) and (5.10). The other qubits are prepared in one of the states $\{|0\rangle, |+\rangle, |1\rangle\}$, using one of the gates $P_i \in \{I, H, X\}$. In the simulation, these gates are implemented using the pulse zero(T_X) (see Section 5.1.3), one GD pulse, or two GD pulses, respectively (cf. Tab. B.1 and Section 5.1.2). The idle gate is implemented by a free time evolution for a duration T_{idle}, corresponding to the pulse zero(T_{idle}) (see Section 5.1.3). Finally, the gate $Y^{\pi/2}$ on qubit 0 is implemented as one GD pulse as described explicitly in Section 5.4.

Ideally, the idle gate would not affect any of the qubits. Therefore, the ideal circuit
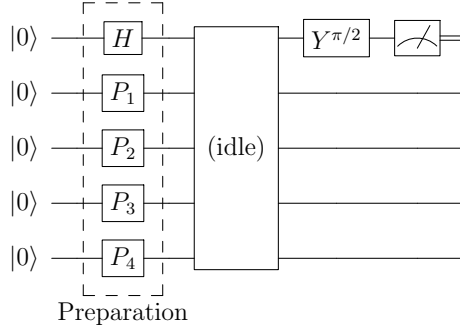
Figure 7.1: Circuit diagram for crosstalk experiments on a five-transmon system. Qubit 0 is prepared in the state $|+\rangle$ using the H gate. Qubits 1 to 4 are prepared in one of the states $\{|0\rangle, |+\rangle, |1\rangle\}$, corresponding to the gates $P_i \in \{I, H, X\}$. The "idle" gate denotes a free time evolution for a time T_{idle}, without the application of any pulses or any externally generated interaction between the qubits. The $Y^{\pi/2}$ gate rotates the state $|+\rangle$ to the state $|1\rangle$. Therefore, ideally, the measurement of qubit 0 at the end should always return 1, corresponding to a Bloch vector $\vec{r}_0 = (0, 0, -1)^T$ (see Eq. (7.1)). The standard gates used in this circuit are defined in Tab. B.1 in Appendix B.

outcome is given by

$$|00000\rangle \overset{\text{Prep.}}{\mapsto} |+q_1 q_2 q_3 q_4\rangle \overset{\text{Idle}}{\mapsto} |+q_1 q_2 q_3 q_4\rangle \overset{Y_0^{\pi/2}}{\mapsto} |1 q_1 q_2 q_3 q_4\rangle \overset{\text{Meas.}}{\mapsto} 1, \tag{7.1}$$

where $|q_i\rangle = P_i |0\rangle \in \{|0\rangle, |+\rangle, |1\rangle\}$. This corresponds to a Bloch vector $\vec{r}_0$ that is aligned with the negative z axis, i.e., $r_0^z = -1$.

However, crosstalk between the transmons during the idle gate is expected to change the state of qubit 0 in a way that depends on the state of the other qubits. We simulate the five-transmon system by solving the TDSE given by Eq. (3.8) with the time step $\tau = 10^{-3}$ ns as described in Section 3.3. A few representative examples of the evolution of the qubits' Bloch vectors are shown in Fig. 7.2.

In Fig. 7.2(a), all qubits are rotated to the positive x axis using an H gate. The idle gate is absent such that the final $Y^{\pi/2}$ gate rotates qubit 0 to the negative z axis as expected. In this case, crosstalk between the transmons only makes the Bloch vectors wiggle slightly, which can be seen in the not perfectly straight motion of the arrows.

When the duration of the idle gate is set to $T_{\text{idle}} = 1600$ ns, however, each Bloch vector keeps on rotating clockwise about the z axis (see Fig. 7.2(b)). Furthermore, the magnitude of the Bloch vectors shrinks, which represents entanglement forming between the qubits. Since the Bloch vector of qubit 0 ends up near the negative y axis, the final $Y^{\pi/2}$ gate does not rotate the qubit, resulting in a 50% error of the final measurement.

In Fig. 7.2(c) and (d), we see that the crosstalk effect on qubit 0 is even stronger if qubits 1 to 4 are prepared in the state $|1\rangle$. The same idle time $T_{\text{idle}} = 1600$ ns renders the Bloch vector of qubit 0 on the negative x axis. The final $Y^{\pi/2}$ gate then rotates the

(a) $Y_0^{\pi/2} H_0 H_1 H_2 H_3 H_4 \, |00000\rangle$

(b) $Y_0^{\pi/2}$ (idle) $H_0 H_1 H_2 H_3 H_4 \, |00000\rangle$

(c) $Y_0^{\pi/2} H_0 X_1 X_2 X_3 X_4 \, |00000\rangle$

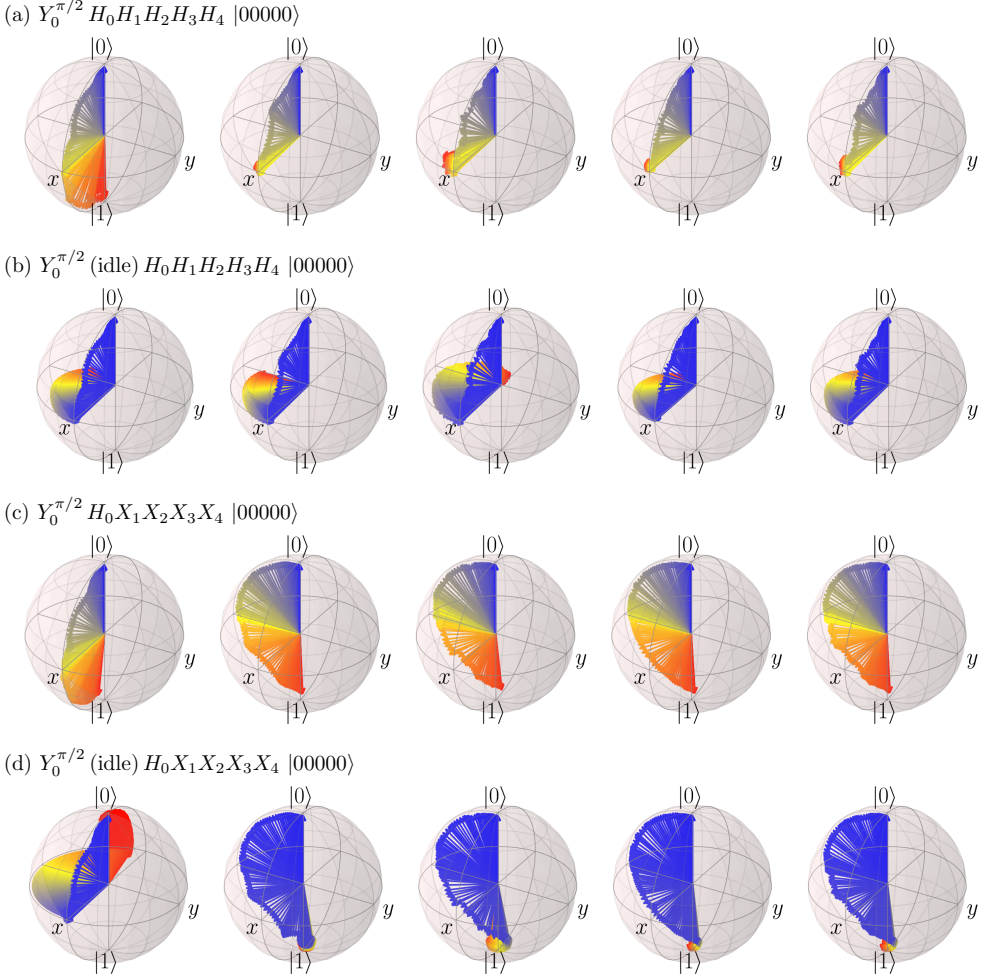(d) $Y_0^{\pi/2}$ (idle) $H_0 X_1 X_2 X_3 X_4 \, |00000\rangle$

Figure 7.2: Bloch-sphere representation of the time evolution of five transmon qubits during the application of the pulses corresponding to the gate circuit given in Fig. 7.1 with (a) $P_i = H$, $T_{\text{idle}} = 0\,\text{ns}$, (b) $P_i = H$, $T_{\text{idle}} = 1600\,\text{ns}$, (c) $P_i = X$, $T_{\text{idle}} = 0\,\text{ns}$, (d) $P_i = X$, $T_{\text{idle}} = 1600\,\text{ns}$. The time t is encoded in the color of the arrows (from blue to red). The model parameters of the simulated transmon system are given in Tab. 3.5 and Tab. 3.6. The Bloch vectors $\vec{r}_i(t)$ at time t are computed according to Eq. (3.38) in a frame rotating at the frequencies f_i given in Tab. 3.5. The simulation results have been visualized with `QuTiP` [Joh2012; Joh2013].

qubit back to $|0\rangle$ (see Fig. 7.2(d)), while it should have ended up in state $|1\rangle$. Using this effect, one could, in principle, engineer a circuit with 100% error rate. Note that this effect qualitatively agrees with the increased frequency observed in Fig. 4.11, since if the qubit's frequency is larger than that of the rotating frame, the Bloch vector performs a clockwise rotation about the z axis. In the next section, we study to what extent the same effect can be observed in a real five-transmon processor.

7.1.2 Comparison with experiments on the IBM Q Experience

When executing the circuit in Fig. 7.1 on a real five-qubit processor, the result is given by the relative frequencies $p_{0/1}^{(\text{exp})}$, corresponding to the relative number of events that the measurement of qubit 0 returned 0 or 1. From the frequencies, we compute the z component of the qubit's Bloch vector according to Eq. (2.5) such that $r_0^z = p_0^{(\text{exp})} - p_1^{(\text{exp})} = 2p_0^{(\text{exp})} - 1$. As the ideal result for each instance of the circuit in Fig. 7.1 is $|1\rangle$ (i.e., $r_0^z = -1$), the error rate in terms of the statistical distance D (see Eq. (6.22)) is given by $p_0^{(\text{exp})}$. In terms of the coordinates of the Bloch vector $\vec{r}_0$, we have

$$D = \frac{1 + r_0^z}{2}. \tag{7.2}$$

In the simulation, we compute the Bloch vector $\vec{r}_0$ through Eq. (3.38). Since this formula includes a projection onto the computational subspace, using Eq. (7.2) for the error effectively yields $D = p_0^{(\text{sim})} + p_{\geq 2}^{(\text{sim})}/2$. Conceptually, this means that a measurement of the transmon in a higher, non-computational state is interpreted as a 50% chance to find the qubit in state $|0\rangle$ or $|1\rangle$. Another reasonable alternative would be to always count a higher state as $|1\rangle$. In practice, however, this conceptual choice only has a negligible effect on the result, because $p_{\geq 2}^{(\text{sim})} < 0.002$ in all cases under investigation.

We use the five-qubit processor `ibmqx4` on the IBM Q Experience [IBM2018b] to execute the circuit given in Fig. 7.1. In the experiment, the idle gate is implemented by $n_{\text{idle}} = 0, \ldots, 120$ identity gates. All experiments were run between February 21, 2018 and March 19, 2018 with 8192 shots.

In Fig. 7.3(a), we show results from both simulation and experiment for the circuit $Y_0^{\pi/2}(\text{idle}) |{+}{+}{+}{+}{+}\rangle$, corresponding to the Bloch-vector evolutions in Fig. 7.2(a) and (b). The agreement between simulation and experiment is remarkable. Both are equally far away from the ideal result $r_0^z = -1$. The only observable difference between simulation and experimental results for r_0^z is the slightly reduced amplitude and the tiny oscillations between successive circuit simulations. These are apparently not resolved by the experiment and appear to be smoothed out. This points out a limitation in the degree to which the device could be engineered to implement the laws of quantum theory on a macroscopic scale. The most obvious explanation would be that the deviation is caused by influences from the environment that are not included in the simulation model (cf. also Fig. 7.6, where a similarly reduced amplitude can be described in terms of environmental effects). Further investigation by including the qubits' readout resonators or an effective environment in the simulation (in the spirit of Section 3.5) could shed light on this aspect. Nevertheless, the otherwise excellent agreement seems to suggest that for this
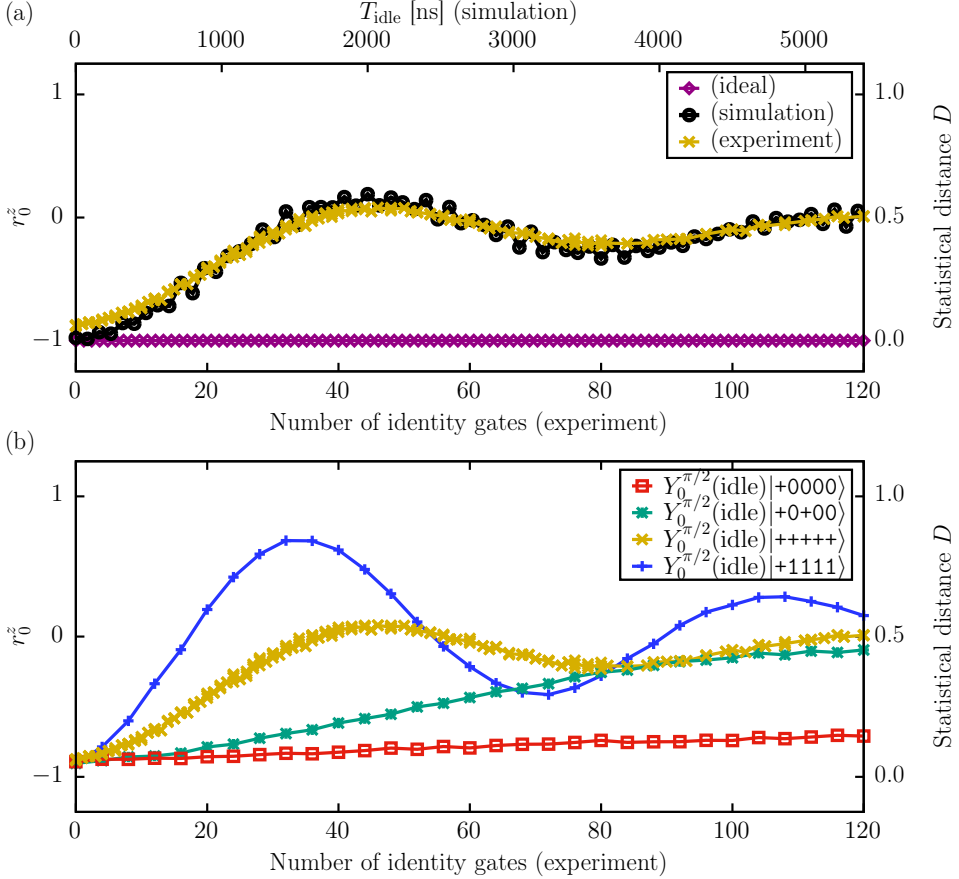
Figure 7.3: Results of the crosstalk experiments based on the circuit given in Fig. 7.1; (a) ideal, simulation, and experimental results for the circuit $Y_0^{\pi/2}(\text{idle}) \, |{+}{+}{+}{+}{+}\rangle$; (b) experimental results for the same type of circuit with different state preparations P_i as indicated by the legend. Shown is the z component of the Bloch vector $\vec{r}_0 = \langle \vec{\sigma}_0 \rangle$ of qubit 0. The right axis shows the error rate in terms of the statistical distance D defined in Eq. (6.22), which is linked to r_0^z via Eq. (7.2). In the experiment, the "idle" part is implemented by a certain number of identity gates given on the bottom axis. In the simulation, the "idle" part is implemented by the pulse zero(T_{idle}) (see Section 5.1.3), where T_{idle} is shown on the top axis. The simulation results in (a) at $T_{\text{idle}} = 0$ ns ($T_{\text{idle}} = 1600$ ns) correspond to the time evolutions visualized in Fig. 7.2(a) (Fig. 7.2(b)).

particular experiment, the most dominant sources of error are the crosstalk effects that are inherently included in the unitary evolution of the simulated transmon model defined in Section 3.2.

Comparing the time scale used in the simulation (top axis in Fig. 7.3(a)) with the number of identity gates used in the experiment (bottom axis in Fig. 7.3(a)), we find that a single identity gate corresponds to $T_{\text{idle}} = 45$ ns. According to the calibration data obtained from the processor, the single-qubit gate duration at the time of the experiment was 83.3 ns with a buffer of 6.7 ns between gates. A difference in the exact time scales is to be expected because the device parameters of the simulated transmon system (see Fig. 3.4 and Tab. 3.5) do not exactly match those of `ibmqx4` [IBM2018b]. Furthermore, the time scale depends sensitively on the transmon-resonator couplings G given in Tab. 3.6, in the sense that the effective longitudinal interaction J given by Eq. (4.28) is proportional to G^4 [Bil2015]. In the experiment, G is not measured but estimated from simulations [Sol2016]. The fact that the transmon simulator and the IBM Q processors still agree so well suggests that this effect is independent of the exact values of the parameters of the full Hamiltonian given by Eqs. (3.9a)–(3.9f).

The agreement between simulation and experiment shown in Fig. 7.3(a) suggests that also other trends seen in the simulation are observable in the experiment. For instance, in Fig. 7.2(d), we found that the crosstalk-induced rotation during the idle gate grows in speed when all other transmons are prepared in state $|1\rangle$. To test this hypothesis, we execute the circuit given in Fig. 7.1 for three additional state preparations on the processor `ibmqx4` [IBM2018b]. The results are shown in Fig. 7.3(b).

Indeed, we see that with all other qubits in state $|1\rangle$, the deviation from the ideal result, indicated by the deviation of the blue plusses from $r_0^z = -1$, becomes even stronger. Furthermore, if only the central qubit is prepared in the state $|+\rangle$ (green crosses), the impact is not as pronounced. This also demonstrates that crosstalk effects go beyond the simple two-qubit picture studied in Section 4.4.

Finally, with all other qubits in state $|0\rangle$, the Bloch vector of qubit 0 effectively stands still during the idle gate (red squares in Fig. 7.3(b)). As this is the same situation that has been chosen to determine qubit frequencies in Section 3.3.2 (see the text below Eq. (3.39)), it shows that our method is in agreement with the experiment.

7.2 Characterization of the singlet state

In this section, we consider a conceptually simple experiment that is commonly used to test whether a system exhibits quantum behavior, in the sense that it produces correlations described by an entangled state. Specifically, we consider the singlet state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \qquad (7.3)$$

which is one of the maximally entangled Bell states defined in Eqs. (2.8a) and (2.8b). A property of this particular Bell state is that independent measurements of both qubits always yield opposite results. This property is called *perfectly anti-correlated*. From Eq. (7.3), we see that this holds when measuring in the computational basis, i.e., when
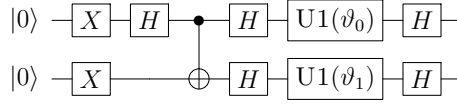
Figure 7.4: Circuit diagram for experiments on the singlet state $|\Psi^-\rangle$ given by Eq. (7.3). The first four gates prepare the state. The last six gates are used to implement a measurement of the state in the eigenbasis of $(\vec{\sigma}_0 \cdot \vec{a})(\vec{\sigma}_1 \cdot \vec{b})$ as a function of ϑ_0 and ϑ_1 (cf. Eq. (7.6)). All circuit elements are defined in Tab. B.1 in Appendix B.

measuring the observables σ_0^z and σ_1^z: whenever one qubit is found in state $|0\rangle$ (with eigenvalue $+1$ since $\sigma^z |0\rangle = +|0\rangle$), the other qubit is found in state $|1\rangle$ (with eigenvalue -1 since $\sigma^z |1\rangle = -|1\rangle$). One can show [Nie2010] that this property holds for any other pair of observables $\vec{\sigma}_0 \cdot \vec{v}$ and $\vec{\sigma}_1 \cdot \vec{v}$, where $\vec{v} \in \mathbb{R}^3$ is the measurement direction with $\|\vec{v}\| = 1$, and $\vec{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ is the vector of Pauli matrices on qubit i. Moreover, if $\vec{v}$ is replaced by $\vec{a}$ for qubit 0 and $\vec{b}$ for qubit 1, a short calculation yields

$$\langle \Psi^- | (\vec{\sigma}_0 \cdot \vec{a})(\vec{\sigma}_1 \cdot \vec{b}) | \Psi^- \rangle = -\vec{a} \cdot \vec{b} = -\cos \vartheta, \tag{7.4}$$

where $\vartheta = \sphericalangle(\vec{a}, \vec{b})$ is the angle between the measurement directions $\vec{a}$ and $\vec{b}$. Thus, ϑ determines the degree of correlation expected in the measurements of both qubits. Furthermore, we have

$$\langle \Psi^- | \vec{\sigma}_0 \cdot \vec{a} | \Psi^- \rangle = 0, \tag{7.5a}$$

$$\langle \Psi^- | \vec{\sigma}_1 \cdot \vec{b} | \Psi^- \rangle = 0. \tag{7.5b}$$

This means that, when measuring only one of the two observables, a measurement is expected to produce an equal number of $+1$'s and -1's.

7.2.1 Experiment

We perform an in-depth characterization of the singlet state by studying the expectation values given in Eqs. (7.4)–(7.5b) for various measurement directions $\vec{a}$ and $\vec{b}$. Specifically, we choose $\vec{a} = (0, \sin \vartheta_0, \cos \vartheta_0)^T$ and $\vec{b} = (0, \sin \vartheta_1, \cos \vartheta_1)^T$ such that $-\vec{a} \cdot \vec{b} = -\cos(\vartheta_1 - \vartheta_0)$. The quantum gate circuit to implement this experiment is given in Fig. 7.4. It consists of a few gates used to prepare the singlet state (using the fact that $|\Psi^-\rangle = \text{CNOT}_{01} H_0 X_0 X_1 |00\rangle$), and a set of gates to implement the measurement. For the latter, note that the gate sequence $H U1(\vartheta_0) H$ can be used to transform between the computational basis (i.e. the eigenbasis of σ^z) and the eigenbasis of $\vec{\sigma} \cdot \vec{a}$ since

$$(H U1(\vartheta_0) H)^\dagger \sigma^z (H U1(\vartheta_0) H) = \sin \vartheta_0 \, \sigma^y + \cos \vartheta_0 \, \sigma^z = \vec{\sigma} \cdot \vec{a}. \tag{7.6}$$

A similar relation holds for ϑ_1 and $\vec{b}$. One can compute the ideal probability distribution of the final state of the circuit shown in Fig. 7.4 as

$$p(j_0 j_1 | \vartheta_0 \vartheta_1) = |\langle j_0 j_1 | (H\mathrm{U1}(\vartheta_0)H) \otimes (H\mathrm{U1}(\vartheta_1)H)|\Psi^-\rangle|^2$$
$$= \frac{1 - (-1)^{j_0+j_1}\cos(\vartheta_1 - \vartheta_0)}{4}. \tag{7.7}$$

To verify that this distribution complies with Eqs. (7.4)–(7.5b), one can compute the expectation values as

$$E_{01}(\vartheta_0, \vartheta_1) = \sum_{j_0 j_1}(-1)^{j_0}(-1)^{j_1}p(j_0 j_1|\vartheta_0 \vartheta_1) = -\cos(\vartheta_1 - \vartheta_0), \tag{7.8a}$$

$$E_0(\vartheta_0, \vartheta_1) = \sum_{j_0 j_1}(-1)^{j_0}p(j_0 j_1|\vartheta_0 \vartheta_1) = 0, \tag{7.8b}$$

$$E_1(\vartheta_0, \vartheta_1) = \sum_{j_0 j_1}(-1)^{j_1}p(j_0 j_1|\vartheta_0 \vartheta_1) = 0, \tag{7.8c}$$

where we used the fact that j_0 (j_1) labels the eigenvalue $(-1)^{j_0}$ ($(-1)^{j_1}$) of $\vec{\sigma}_0 \cdot \vec{a}$ ($\vec{\sigma}_1 \cdot \vec{b}$).

When using devices such as an IBM Q processor for this experiment, one does not obtain a probability distribution $p(j_0 j_1|\vartheta_0 \vartheta_1)$ directly. Instead, the device produces a number of samples (usually 8192) for each setting of ϑ_0 and ϑ_1. The resulting frequencies for each bit string,

$$f(j_0 j_1) = \frac{\# \text{ of samples } j_0 j_1}{8192}, \tag{7.9}$$

are then used as estimators for the probabilities $p(j_0 j_1|\vartheta_0 \vartheta_1)$. We denote the corresponding estimates for the expectation values given in Eqs. (7.8a)–(7.8c) by $F_{01}(\vartheta_0, \vartheta_1)$, $F_0(\vartheta_0, \vartheta_1)$, and $F_1(\vartheta_0, \vartheta_1)$, respectively. The goal of the experiment is to see how well F_{01}, F_0, and F_1 agree with the theoretical result E_{01}, E_0, and E_1.

Some aspects of the experiment have already been studied. In [Mic2017], we performed the experiment on an IBM Q processor as part of a collection of benchmarking circuits (the results are plotted as hollow circles in Figs. 7.5 and 7.7 below). In [Wil2017], we simulated the experiment using the optimized pulses discussed in Chapter 5 (the corresponding results are shown in Fig. 7.5(a) and (b)). The particular pulse parameters are given in Tab. F.1 and Tab. F.2 in Appendix F.

As can be seen in Fig. 7.5(a) and (b), there is quite some difference in the results, in the sense that the IBM Q processor produces rather large, systematic deviations while the simulation performs reasonably well in comparison with the ideal result (dashed black lines). Note that the gate metrics given in Tab. 6.1 are almost equal to the error rates of the processor, so the deviations between simulation and experiment are not captured by the gate metrics (cf. also the conclusions in Section 6.4).

Adjusting the pulse parameters

Since the gates implemented by the optimized pulses for the two-transmon simulation model perform much better, an interesting question is whether the errors can be reproduced solely from the unitary control errors caused by the time-dependent pulses. To
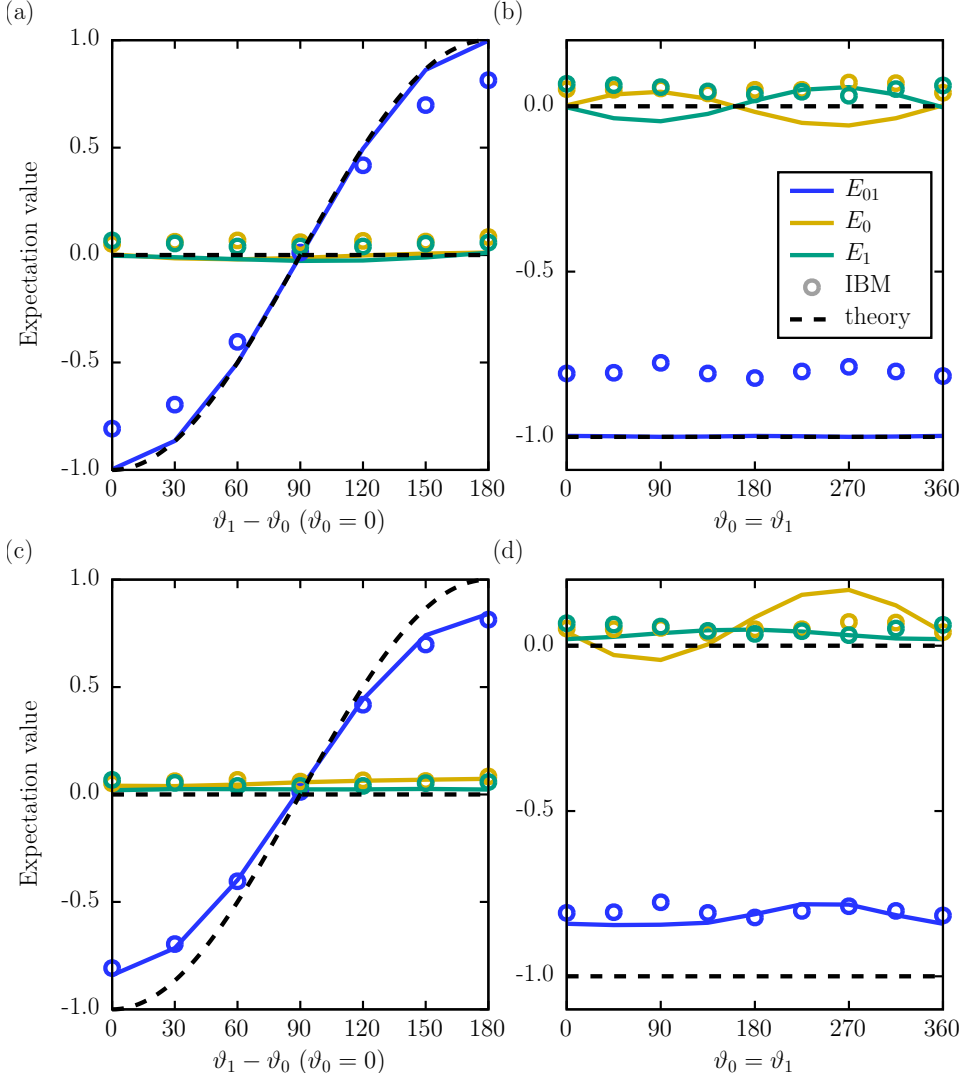
Figure 7.5: Results for the singlet-state characterization using (a),(b) the original optimized pulse parameters given in Appendix F and (c),(d) the modified pulse parameters that have been fitted to the experimental results (see Tab. 7.1). Lines show the expectation values E_{01}, E_0, and E_1 defined in Eqs. (7.8a)–(7.8c), computed using the probabilities $p(j_0 j_1 | \vartheta_0 \vartheta_1)$ from the two-transmon simulation (cf. Section 3.4.3). In all figures, hollow circles represent the expectation values F_{01}, F_0, and F_1 obtained from experiments on the IBM Q processor that was available on February 16, 2017 (see [Mic2017]), computed using the relative frequencies $f(j_0 j_1 | \vartheta_0 \vartheta_1)$ given by Eq. (7.9). Dashed lines represent the theoretical result, computed from the ideal probability distribution given by Eq. (7.7).

Table 7.1: Pulse parameters used for the simulation results shown in Fig. 7.5. The original values represent the optimized pulses discussed in Chapter 5 (the corresponding results are shown in Fig. 7.5(a) and (b)). They are taken from the pulse parameters given in Tab. F.1 and Tab. F.2 in Appendix F. The fitted values result from minimizing the objective function given by Eq. (7.10) (the corresponding results are shown in Fig. 7.5(c) and (d)). The duration T_X of the single-qubit GD pulses is the same for all gates. The two-qubit CNOT pulse is based on the CR2 scheme (see Fig. 5.1(c)).

Pulse name	Parameter	Original value	Fitted value
`xpih-0`	f_0 [GHz]	5.3463	5.34647
	$(\Omega_X^{\pi/2})_0$	0.002221	0.001980
	$(\beta_X^{\pi/2})_0$ [ns]	0.2309	0.2584
`xpih-1`	f_1 [GHz]	5.1167	5.11691
	$(\Omega_X^{\pi/2})_1$	0.002269	0.002296
	$(\beta_X^{\pi/2})_0$ [ns]	0.2891	0.4698
`cnot-0-1`	T_{CR} [ns]	102.9746	102.9720
	Ω_{CR}	0.01111	0.00979
	$(\Omega_X^{\pi})_C$	0.004444	0.003415
	$(\beta_X^{\pi})_C$ [ns]	0.2193	0.2239
	$(\Omega_X^{\pi/2})_T$	0.002269	0.002032
	$(\beta_X^{\pi/2})_T$ [ns]	0.2891	0.2742
(for all pulses)	T_X [ns]	83	82.893

address this question, we try to fit the pulse parameters such that simulation results and experimental results match. If this is possible, it would suggest that the errors for this experiment are largely due to the pulses chosen to implement the gates. Otherwise, it would point at external sources of error such as effects due to the environment or the additional transmons and resonators on the processor (see below).

To fit the pulse parameters, we utilize the Nelder–Mead optimization method introduced in Section 5.3.1. In every iteration, the singlet characterization is evaluated using the current pulse parameters. The objective function used for the optimization is given by the root mean square

$$\sqrt{\frac{1}{3N_\vartheta} \sum_{\vartheta_0,\vartheta_1} \left((E_{01}(\vartheta_0,\vartheta_1) - F_{01}(\vartheta_0,\vartheta_1))^2 + \sum_{l=0,1} (E_l(\vartheta_0,\vartheta_1) - F_l(\vartheta_0,\vartheta_1))^2 \right)}, \qquad (7.10)$$

where $N_\vartheta = 16$ is the number of configurations $(\vartheta_0,\vartheta_1)$ used in the experiment. The 13 fitted pulse parameters are shown in Tab. 7.1. Most of the resulting parameters differ by less than 10% from their respective initial values.

The performance of the singlet characterization using the fitted pulse parameters is shown in Fig. 7.5(c) and (d). We see that applying the new pulses yields much better

agreement between simulation and experimental results. The expectation values E_{01}, E_0, and E_1 deviate equally strongly from the theory. This is especially true for the case shown in Fig. 7.5(c). However, the large oscillation of E_0 in Fig. 7.5(d) (yellow line) is not present in the experimental results (yellow circles); instead, the experimental results rather indicate a constant offset.

We can understand the cause of this large oscillation in the simulation results by investigating the corresponding pulses. As shown in Fig. 7.4, the parameters ϑ_0 and ϑ_1 occur in the U1 gates. These gates are implemented by changing the phase of the following pulses for the H gates. Specifically, they change the phase $\gamma_{ij} \in [0, 2\pi)$ of the microwave pulses defined in Eq. (5.1) (see Section 5.1.1). This change would leave the state of the system invariant if the prepared state was the singlet state given by Eq. (7.3). However, as the pulse parameters have been modified, the logical conclusion is that a slightly different state that is sensitive to these phases has been prepared by the pulses (cf. also the state vector in [Mic2017] obtained from the data produced by the IBM Q processor).

In the experimental data, it is not the oscillation, but rather the almost constant offset of E_0 and E_1 from 0 that stands out. This suggests that another kind of error may be the dominant cause for the deviation from the ideal singlet expectation values.

7.2.2 Effective error model

An obvious alternative source of errors in the device is the environment which is not included in the two-transmon system. To address this hypothesis, a simple method to include effective environment-induced errors is provided by the theory of quantum fault tolerance. In this formalism, errors in the system are modeled by an error channel $\mathcal{E}(\rho)$, which is a linear map on the system's density matrix ρ, i.e.

$$\rho \mapsto \mathcal{E}(\rho) = \sum_\alpha M_\alpha \rho M_\alpha^\dagger, \tag{7.11}$$

where M_α are linear (Kraus) operators defining the error channel (cf. the definition of general quantum operations in Section 2.4). The description is expressed in terms of ρ instead of pure states $|\psi\rangle$ such that non-unitary maps are supported and damping can be modeled. Note that the evolution of a smaller system coupled to a larger system can always be written in the form of Eq. (7.11) if the evolution of the joint system is unitary and the initial state of the joint system is a product state (see Eq. (2.29)).

In this section, we study separate error channels for each qubit. The simplest type of error channels $\mathcal{E}$ considered in the theory of quantum fault tolerance are the so-called depolarizing channel and the amplitude damping channel [Nie2010].

Depolarizing channel

The depolarizing channel $\mathcal{E}_{\text{dep}}$ is defined as

$$\mathcal{E}_{\text{dep}}(\rho) = (1 - p_x - p_y - p_z)\rho + p_x \sigma^x \rho \sigma^x + p_y \sigma^y \rho \sigma^y + p_z \sigma^z \rho \sigma^z, \tag{7.12}$$

where $p_x, p_y, p_z \in [0, 1]$ and $p_x + p_y + p_z < 1$. The interpretation is that at each application of $\mathcal{E}_{\text{dep}}$, a bit flip happens with probability p_x, a phase flip happens with probability p_z,

and a joint bit and phase flip happens with probability p_y. A depolarizing channel is called symmetric if $p_x = p_y = p_z$. It is worth mentioning that a symmetric depolarizing channel for an arbitrary number of qubits n can be written as

$$\mathcal{E}_{\text{dep}}^{(\text{sym},n)}(\rho) = F\rho + (1 - F)\frac{I}{2^n}, \tag{7.13}$$

where the parameter $F \in [0,1]$ can be interpreted as a fidelity. This channel gives rise to a simple mixture of a uniform distribution and the quantum state ρ, which was also considered as a model for the quantum supremacy experiment [Aru2019].

Amplitude damping channel

The amplitude damping channel $\mathcal{E}_{\text{amp}}$ is defined by the following set of Kraus operators in the representation given by Eq. (7.11):

$$M_0 = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \qquad M_1 = \sqrt{p}\begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}, \tag{7.14a}$$

$$M_2 = \sqrt{1-p}\begin{pmatrix} \sqrt{1-\gamma} & 0 \\ 0 & 1 \end{pmatrix}, \qquad M_3 = \sqrt{1-p}\begin{pmatrix} 0 & 0 \\ \sqrt{\gamma} & 0 \end{pmatrix}, \tag{7.14b}$$

where $p, \gamma \in [0,1]$. If $p \neq 1$, the channel is sometimes called *generalized* amplitude damping channel as it can also excite the qubit. In general, the parameter p determines if the energy exchange with the environment rather causes a decay ($p \approx 1$) or an excitation ($p \approx 0$) of the qubit, and the parameter γ is the corresponding rate (per unit time).

Application of the effective error model

We extend the circuit given in Fig. 7.4 with a depolarizing channel after every step that would correspond to a new pulse, and an amplitude damping channel at the end of the circuit. A motivation to have the latter only at the end is given by the fact that energy exchange with the environment is most likely to occur during the measurement process, where information leaves the system and enters the environment (cf. [Jac2014; Gir2014]). However, we also experimented with other approaches that produced only slightly worse results. The new circuit to model effects from the environment in this simple way is shown in Fig. 7.6.

The result of fitting the error-channel parameters p_x, p_y, p_z, p, and γ for each qubit to the experimentally observed data is shown in Fig. 7.7. We find that the effective model can describe the observations very well. In particular, it includes the systematic zig-zag pattern observed for E_{01} in the case $\vartheta_0 = \vartheta_1$ shown in Fig. 7.7(b). We remark that a single bit-flip channel with $p_x \neq 0$ for qubit 0 suffices to model this behavior as well as the reduced amplitude of E_{01} shown in Fig. 7.7(a). However, only the amplitude damping channel can model the constant offset of E_0 and E_1 from 0.

The channel parameters resulting from the fit are given in Tab. 7.2. They can be interpreted in the way that the control qubit 0 is more prone to bit-flip errors (0.7%), while the target qubit 1 is rather susceptible to phase errors (0.3%). This is reasonable
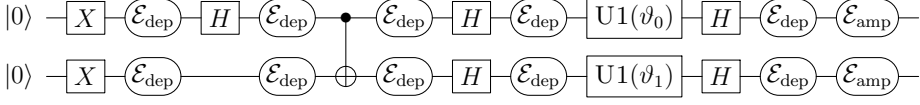
Figure 7.6: Circuit diagram for experiments on the singlet state, extended by effective error channels to test if they can model the experimentally observed deviation from the theoretical result (see Fig. 7.4). The depolarizing channel $\mathcal{E}_{\mathrm{dep}}$ is defined in Eq. (7.12) and is inserted after every step that would correspond to a pulse in an implementation. The amplitude damping channel $\mathcal{E}_{\mathrm{amp}}$ is defined in Eqs. (7.14a) and (7.14b).
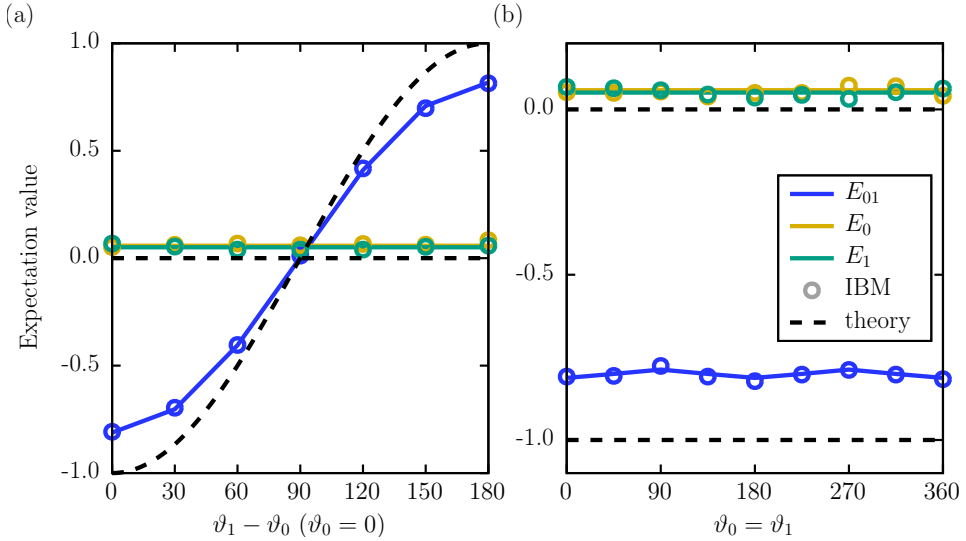

Figure 7.7: Same as Fig. 7.5, except that E_{01}, E_0, and E_1 are not obtained from the transmon simulation, but from an ideal quantum computer simulator such as JUQCS [DeR2019a; Wil2020a] combined with the error channels defined in Eqs. (7.12)–(7.14b). The corresponding circuit is shown in Fig. 7.6. The parameters of the error channels are given in Tab. 7.2.

Table 7.2: Resulting parameters (p_x, p_y, p_z) of the depolarizing channel $\mathcal{E}_{\mathrm{dep}}$ defined in Eq. (7.12) and (p, γ) of the amplitude damping channel $\mathcal{E}_{\mathrm{amp}}$ defined in Eqs. (7.14a) and (7.14b). The channels are added to the circuit as shown in Fig. 7.6 with the same parameters at each step, but different parameters for different qubits.

Qubit	p_x	p_y	p_z	p	γ
0	0.007	0.001	0.000	0.818	0.092
1	0.001	0.001	0.003	0.980	0.053

since for the target qubit, the phase is the most sensitive quantity for the operation of the CNOT gate because a phase shift on the target qubit affects the pulses on both qubits (see the occurrence of ϑ_T for both qubits in the echoed CR scheme in Fig. 5.3(c)). We also see that the interaction with the environment mainly causes qubit relaxation ($p \approx 1$). The corresponding decay rates $\gamma = 0.092$ ($\gamma = 0.053$) for the control (target) qubit are roughly compatible with the qubit relaxation times T_1 of the order of $40\,\mu\mathrm{s}$: for a single application of $\mathcal{E}_{\mathrm{amp}}$ over the course of the circuit in Fig. 7.6, whose execution time T_{exec} is a few microseconds, we have $T_{\mathrm{exec}}/\gamma \approx T_1$. The fact that the decay rate γ of the control qubit is larger than that of the target qubit makes sense as the control qubit of the CNOT gate is driven much more strongly (cf. Fig. 5.1(c)), so it is more susceptible to energy exchange with the environment.

In conclusion, there are errors in the device that can be very well described by simple error channels, such as the depolarizing channel and the amplitude damping channel. As such simple error channels belong to the error model that is addressed by the theory of quantum error correction, this suggests that error correction could work reasonably well for the transmon architecture (see the following section). However, note also that there are more difficult, correlated errors in the device that are very well described by the five-transmon simulation model (see Fig. 7.3(a)). It would be interesting to see whether the five-transmon model can also describe the experiments studied in this section better than the two-transmon model. Initial evidence for this idea is given by the results shown in Fig. 6.2. We leave a detailed analysis of the singlet-state characterization with the five-transmon model for future work.

7.3 Testing quantum fault tolerance

Both simulations and experiments presented in the previous sections suggest that controlling gate-based quantum computers to high accuracy is considerably difficult. For the simulations, this is true even though the model studied in this work is inherently quantum mechanical (see Section 3.2). The same observation is supported by many other experiments implementing the gate-based quantum computer model [She2016a; Gam2017; Nei2018; Mic2017; Aru2019].

The most prominent, long-term solution to this problem is proposed by the theory of quantum error correction and fault tolerance [Sho1996; DiV1996; Got1998b; Cam2017]. The basic idea is that additional *physical* qubits are used to encode a smaller number of so-called *logical* qubits. The logical qubits are designed to be tolerant to errors within certain mathematical models.

Simple models consider discrete, uncorrelated errors such as spontaneous bit or phase flips (see Eq. (7.12)), whereas more sophisticated models consider non-Markovian errors in a general Hamiltonian framework [Ter2005; Ali2006; Ali2007; Aha2008; Ng2009]. In these models, so-called *threshold theorems* are derived. They state that, if the error rates are below a certain threshold, arbitrarily long quantum computation is possible by using a suitable fault-tolerant protocol. As with any mathematical model, however, it is a priori unclear if its predictions hold in practice. Specifically, it is unclear whether such a fault-tolerant protocol systematically improves the results in an actual experiment.

To address this question, we test a full fault-tolerant protocol explicitly designed for small experiments [Got2016], using both the five-transmon, six-resonator model defined in Section 3.4.5, and the 16-qubit processor `ibmqx5` on the IBM Q Experience [IBM2018a]. The protocol is based on the four-qubit code [Leu1997; Vai1996; Gra1997], which has recently also been studied in other experiments [Lin2017; Vui2018; Tak2017; Har2019]. We find that the protocol systematically improves the results in the presence of the inherent control and measurement errors of the studied transmon systems. Part of the work presented in this section has been published in [Wil2018b].

7.3.1 Fault-tolerant protocol

A fault-tolerant protocol considers the full procedure of (1) preparing a certain initial state, (2) applying a few quantum gates to the state, and (3) measuring the result to obtain a distribution of bit strings. For steps (1) and (2), the protocol provides a *circuit encoding* to translate each *bare circuit* (or *logical circuit*) into an *encoded circuit* that requires a larger number of physical qubits. Here, the term "circuit" explicitly includes the gates used to prepare the initial state. For step (3), the fault-tolerant protocol states how the measured bit strings, obtained by measuring the larger number of physical qubits, are to be interpreted to obtain a distribution for the logical qubits. In particular, we consider bare and encoded circuits for two logical qubits.

Definition of the circuit encoding

The protocol under investigation encodes two logical qubits in four physical qubits $q_1 q_2 q_3 q_4$ and an additional so-called *ancillary* qubit q_0. The code can detect an arbitrary single-qubit error [Got2016]. Such a code is typically expressed using the notation $[[4, 2, 2]]$, where a code of the form $[[n_{\text{phy}}, n_{\text{log}}, d]]$ means that the code uses n_{phy} physical qubits to encode the state of n_{log} logical qubits, and the distance d includes information about the number of errors that the code can detect or correct [Nie2010].

The logical two-qubit states of the $[[4, 2, 2]]$ code are defined as

$$\overline{|00\rangle} = (|0000\rangle + |1111\rangle)/\sqrt{2}, \tag{7.15a}$$

$$\overline{|01\rangle} = (|1100\rangle + |0011\rangle)/\sqrt{2}, \tag{7.15b}$$

$$\overline{|10\rangle} = (|1010\rangle + |0101\rangle)/\sqrt{2}, \tag{7.15c}$$

$$\overline{|11\rangle} = (|0110\rangle + |1001\rangle)/\sqrt{2}. \tag{7.15d}$$

By linear combination, one can derive the encoded versions of other logical two-qubit states, e.g.,

$$\overline{|0+\rangle} = (\overline{|00\rangle} + \overline{|01\rangle})/\sqrt{2} = (|0000\rangle + |1100\rangle + |0011\rangle + |1111\rangle)/2, \tag{7.16a}$$

$$\overline{|\Phi^+\rangle} = (\overline{|00\rangle} + \overline{|11\rangle})/\sqrt{2} = (|0000\rangle + |0110\rangle + |1001\rangle + |1111\rangle)/2. \tag{7.16b}$$

Using the definition of the logical two-qubit basis states given by Eqs. (7.15a)–(7.15d), we can derive expressions for logical gates in this code. For instance, a logical bit flip on

Table 7.3: List of the three initial states considered in this experiment. For each state, we list the bare and encoded versions (see Eqs. (7.15a)–(7.16b)) of their preparation circuits. The preparation of the encoded version of $|00\rangle$ requires an additional ancilla qubit. Elementary gates are defined in Tab. B.1 in Appendix B.

Table 7.4: List of the bare and encoded gate elements used to construct circuits to test the fault-tolerant protocol (see Eqs. (7.17a)–(7.17f)). Elementary gates are defined in Tab. B.1 in Appendix B.

qubit 1, denoted by $\overline{\text{X1}}$, has to map $\overline{|0j\rangle}$ to $\overline{|1j\rangle}$ and $\overline{|1j\rangle}$ to $\overline{|0j\rangle}$ for $j = 0, 1$. This is accomplished by flipping the physical qubits q_1 and q_3 in Eqs. (7.15a)–(7.15d). Therefore, we have $\overline{\text{X1}} = X_1 X_3$. In the same way, we find $\overline{\text{X2}} = X_1 X_2$. Similar expressions can be derived for the logical phase flips, namely $\overline{\text{Z1}} = Z_1 Z_2$ and $\overline{\text{Z2}} = Z_1 Z_3$.

In addition to these single-qubit gates, we consider two particular two-qubit gates. The first is a Hadamard gate on both logical qubits followed by swapping the qubits. This gate is denoted by HHS. On a bare two-qubit state, such a transformation can be implemented by the gate sequence $\text{HHS} = \text{CNOT}_{12}\text{CNOT}_{21}\text{CNOT}_{12}H_1H_2$. In the code space defined by Eqs. (7.15a)–(7.15d), this is accomplished by $\overline{\text{HHS}} = H_1H_2H_3H_4$. Finally, we consider the controlled-phase gate CZ. This is an entangling gate with the matrix representation $\text{diag}(1, 1, 1, -1)$. A bare implementation of this gate is given by $\text{CZ} = H_2\text{CNOT}_{12}H_2$. On the encoded states defined by Eqs. (7.15a)–(7.15d), we reach the same effect with the gate sequence $\overline{\text{CZ}} = Z_2Z_3S_1S_2S_3S_4$. In summary, we consider six logical gates whose encoded versions are given by

$$\overline{\text{X1}} = X_1 X_3, \tag{7.17a}$$

$$\overline{\text{X2}} = X_1 X_2, \tag{7.17b}$$

$$\overline{\text{Z1}} = Z_1 Z_2, \tag{7.17c}$$

$$\overline{\text{Z2}} = Z_1 Z_3, \tag{7.17d}$$

$$\overline{\text{HHS}} = H_1H_2H_3H_4, \tag{7.17e}$$

$$\overline{\text{CZ}} = Z_2Z_3S_1S_2S_3S_4, \tag{7.17f}$$

where the definitions of the elementary gates on the right-hand side are given in Tab. B.1 in Appendix B.

The goal of this section is to compare bare versions of a two-qubit circuit with their corresponding encoded versions. We consider circuits composed of one of the initial states $|00\rangle$, $|0+\rangle$, and $|\Phi^+\rangle$, and an arbitrary combination of gates from the gate set $\{\text{X1}, \text{X2}, \text{Z1}, \text{Z2}, \text{HHS}, \text{CZ}\}$. Note that this gate set is not universal, in the sense that not all quantum algorithms can be encoded. The requirement of universality was dropped in favor of having a full fault-tolerant protocol (including the state preparation) that is applicable to small experiments [Got2016].

In Tab. 7.3 and Tab. 7.4, we list all gate sequences used to construct both bare and encoded versions of the circuits. The labels q_i refer to the qubit labels that are used in the experiment below. Note that all circuits need to be expressed in terms of the gates that are supported by the transmon architecture. Therefore, the bare version of the HHS gate contains three CNOT gates (see Tab. 7.4). Similarly, the encoded version of the state preparation of $|00\rangle$ requires five CNOT gates and one additional ancillary qubit (see Tab. 7.3).

The motivation for the ancillary qubit is given in [Got2016] and is based on the simple model of discrete, uncorrelated single-qubit errors such as spontaneous bit or phase flips. An error within this model can be detected by the ancilla qubit being in state $|1\rangle$ or the resulting four-qubit state having an odd parity (i.e., an odd number of 1's), since each four-qubit state in Eqs. (7.15a)–(7.15d) has an even parity.

Evaluation

We consider a two-qubit circuit made from an initial state given in Tab. 7.3 and a sequence of gates from Tab. 7.4. In what follows, we define $p_{j_0 j_1}^{(\text{id})}$ to be the probability distribution of two-bit strings $j_0 j_1$ as computed by an ideal gate-based quantum computer (see Chapter 2). This distribution is compared to the distributions $p_{j_0 j_1}^{(\text{bare})}$ and $p_{j_0 j_1}^{(\text{enc})}$, which are obtained from the evaluation of the bare and encoded version of the circuit as follows.

The evaluation of the bare version of the circuit using the two qubits $q_3 q_4$ is straightforward: In the experiment, we execute the circuit a certain number of times (typically 8192) and count the number $n_{j_0 j_1}$ of measured bit strings $j_0 j_1$. We then obtain a distribution of relative frequencies $p_{j_0 j_1}^{(\text{bare})} = n_{j_0 j_1}/8192$. In the simulation, we directly obtain the probabilities $p_{j_0 j_1}^{(\text{bare})} = |\langle m_3 = j_0, m_4 = j_1 | \Psi \rangle|^2$ from the state vector $|\Psi\rangle$ given by Eq. (3.23).

For the encoded version of the circuit, the distribution $p_{j_0 j_1}^{(\text{enc})}$ is constructed by evaluating the data as dictated by the fault-tolerant protocol: Since the encoded circuit involves five physical qubits, a measurement of all qubits at the end of the circuit produces five bits. If the bit corresponding to the ancilla q_0 is 1, or if the four-bit string corresponding to $q_1 q_2 q_3 q_4$ has an odd parity (i.e., one or three 1's), the run is discarded. Otherwise, we can map the resulting four-bit string to a logical two-qubit state $\overline{|j_0 j_1\rangle}$ using Eqs. (7.15a)–(7.15d). By counting all these mapped two-bit strings, we generate the frequency distribution $p_{j_0 j_1}^{(\text{enc})}$.

The step of systematically discarding some of the measured bit strings is called *postselection* procedure. The corresponding ratio of bit strings that are not discarded is called the postselection ratio r. Note that the essential idea of fault-tolerant protocols based on postselection is that the postselection procedure is systematic. In other words, it is not required to know the ideal, theoretical result of the quantum circuit in order to perform the postselection procedure. Instead, the protocol provides a fixed set of rules (such as "q_0 is 1" or "$q_1 q_2 q_3 q_4$ has an odd parity") that can be checked for the measured bit string to see whether it should be discarded. See Tab. 7.5 below for an example application of the postselection procedure.

To compare the performance of the bare and encoded circuits, we compare the resulting distributions $p_{j_0 j_1}^{(\text{bare/enc})}$ to the ideal distribution $p_{j_0 j_1}^{(\text{id})}$ by means of the statistical distance (cf. Eq. (6.22)),

$$D_{\text{bare/enc}} = \frac{1}{2} \sum_{j_0 j_1} \left| p_{j_0 j_1}^{(\text{bare/enc})} - p_{j_0 j_1}^{(\text{id})} \right|. \tag{7.18}$$

To distinguish between simulation and experimental results, we use the notations $D_{\text{bare/enc}}^{(\text{sim})}$ and $D_{\text{bare/enc}}^{(\text{exp})}$, respectively. Note that from the transmon simulator, we can directly obtain the probabilities $p_{j_0 j_1}^{(\text{bare/enc})}$ from the distribution defined by the state vector $|\Psi\rangle$, so the intermediate step of sampling and counting the outcomes is omitted.

7.3.2 Test systems and circuits

To test the fault-tolerant protocol, the performance of bare and encoded circuits needs to be compared for a representative set of circuits. We generate such a representative set

by applying the procedure defined in [Got2016] using the maximum logical circuit length $T = 10$, the repetition parameter $\mathrm{RP} = 6$, and the periodicity $P = 3$. We obtain a total of 465 circuits, composed of 155 logical gate sequences from Tab. 7.4 for each of the three initial states given in Tab. 7.3. A list of all circuits is given in Listing 7.1, sorted by the total time required for the simulation of all underlying quantum gate pulses.

Listing 7.1 Definition of all $465 = 3 \times 155$ circuits used to test the fault-tolerant protocol, generated according to the procedure specified in [Got2016]. Each of the 155 lines consists of an ID (red), a sequence of logical gates from Tab. 7.4 (blue), and a placeholder for the three initial states $|\mathbf{i}\rangle \in \{|00\rangle, |0+\rangle, |\Phi^+\rangle\}$ (see Tab. 7.3). The circuits are sorted in ascending order by the total time $T_{\mathrm{bare}}^{|00\rangle} \in [0, 10\,\mu\mathrm{s}]$ required for simulating the time evolution of all pulses for the bare version corresponding to the initial state $|00\rangle$ (cf. Section 5.4).

```
0   |i>                                      52  X2  Z1  X2  Z1  X2  Z1  X2  Z1  |i>         104  Z2  X1  Z2  X2  CZ  HHS  CZ  |i>
1   Z2  |i>                                   53  Z1  Z1  X2  X2  Z1  Z1  X2  X2  |i>         105  HHS  CZ  X2  CZ  |i>
2   Z2  Z2  |i>                               54  X2  X1  CZ  X1  |i>                         106  X1  Z2  HHS  CZ  CZ  |i>
3   Z2  Z2  Z2  |i>                           55  X1  X1  X1  X1  X1  |i>                     107  HHS  X2  Z2  CZ  CZ  |i>
4   Z2  Z2  Z2  Z2  |i>                       56  X2  Z1  X2  Z1  X2  Z1  X2  Z1  X2  Z1  |i>  108  Z2  CZ  X1  HHS  X1  CZ  |i>
5   Z2  Z2  Z2  Z2  Z2  |i>                   57  Z2  CZ  X2  X2  X1  Z1  |i>                 109  Z2  X1  Z1  Z1  CZ  Z1  X2  Z1  HHS  CZ  |i>
6   Z2  Z2  Z2  Z2  Z2  Z2  |i>               58  Z1  X2  Z2  CZ  X2  X1  X1  |i>             110  CZ  CZ  CZ  CZ  CZ  |i>
7   Z2  Z2  Z2  Z2  Z2  Z2  Z2  |i>           59  CZ  X1  X2  Z1  Z1  X1  X1  Z1  Z1  Z2  |i>  111  X2  CZ  HHS  X2  CZ  Z2  |i>
8   Z2  Z2  Z2  Z2  Z2  Z2  Z2  Z2  |i>       60  X1  X1  X1  X1  X1  X1  |i>                 112  X2  X1  CZ  HHS  CZ  Z1  Z1  X1  X2  Z2  |i>
9   Z2  Z2  Z2  Z2  Z2  Z2  Z2  Z2  Z2  |i>   61  X2  X2  X2  X2  X2  X2  |i>                 113  HHS  HHS  X2  |i>
10  Z2  Z2  Z2  Z2  Z2  Z2  Z2  Z2  Z2  Z2  |i>  62  CZ  CZ  |i>                             114  Z1  HHS  HHS  X2  |i>
11  Z1  Z2  |i>                               63  X1  X1  X1  X1  X1  X1  X1  |i>             115  X1  HHS  X1  HHS  |i>
12  Z1  Z2  Z1  Z2  |i>                       64  CZ  CZ  X1  |i>                             116  X2  Z2  HHS  X2  Z2  HHS  |i>
13  Z1  Z2  Z1  Z2  Z1  Z2  |i>               65  X1  CZ  Z2  X1  CZ  Z2  |i>                 117  X2  HHS  Z2  Z2  X2  HHS  Z2  Z2  |i>
14  Z1  Z2  Z1  Z2  Z1  Z2  Z1  Z2  |i>       66  CZ  Z2  Z1  Z2  X1  CZ  X2  |i>             118  Z1  Z2  Z1  HHS  X1  Z1  Z2  Z1  HHS  X1  |i>
15  Z1  Z2  Z1  Z2  Z1  Z2  Z1  Z2  Z2  |i>   67  CZ  CZ  X2  |i>                             119  Z1  Z2  X2  Z1  HHS  Z1  Z2  X2  Z1  HHS  |i>
16  Z2  Z1  |i>                               68  X1  X1  X1  X1  X1  X1  X1  X1  |i>         120  CZ  CZ  CZ  CZ  CZ  CZ  |i>
17  Z2  Z1  Z2  Z1  |i>                       69  X2  X2  X2  X2  X2  X2  X2  X2  |i>         121  CZ  CZ  X1  CZ  CZ  X1  CZ  CZ  X1  |i>
18  Z2  Z1  Z2  Z1  Z2  Z1  |i>               70  HHS  |i>                                   122  HHS  X1  Z2  X1  Z2  HHS  X1  Z2  X1  Z2  |i>
19  Z2  Z1  Z2  Z1  Z2  Z1  Z2  Z1  |i>       71  HHS  Z1  |i>                               123  Z2  Z1  HHS  HHS  X2  X1  Z2  CZ  |i>
20  Z2  Z1  Z2  Z1  Z2  Z1  Z2  Z1  Z2  Z1  |i>  72  Z1  Z2  CZ  CZ  X1  X1  X2  X2  |i>      124  CZ  X1  Z2  HHS  X2  X1  Z1  Z1  HHS  |i>
21  Z1  |i>                                   73  X1  X1  X1  X1  X1  X1  X1  X1  X1  |i>     125  CZ  CZ  X2  CZ  CZ  X2  CZ  CZ  X2  |i>
22  X2  Z1  |i>                               74  X1  HHS  |i>                               126  X1  X1  X1  HHS  Z2  X1  X1  X1  HHS  Z2  |i>
23  X1  |i>                                   75  X2  Z2  HHS  |i>                           127  CZ  CZ  CZ  CZ  CZ  CZ  CZ  |i>
24  Z1  Z1  X1  |i>                           76  X2  HHS  Z2  Z2  |i>                       128  HHS  CZ  HHS  CZ  |i>
25  Z2  Z1  X2  Z2  |i>                       77  Z1  Z2  Z1  HHS  X1  |i>                   129  Z1  CZ  Z1  Z2  HHS  Z1  CZ  Z1  Z2  HHS  |i>
26  Z2  X1  Z2  X2  |i>                       78  Z1  Z2  X2  Z1  HHS  |i>                   130  HHS  CZ  X2  HHS  X1  X1  Z1  X1  X2  |i>
27  Z2  X2  Z1  |i>                           79  HHS  X1  Z2  Z2  |i>                       131  HHS  CZ  HHS  X2  CZ  |i>
28  X2  |i>                                   80  Z2  Z2  Z2  X2  Z1  CZ  CZ  Z2  X2  X2  |i>  132  HHS  HHS  HHS  Z1  |i>
29  X1  X1  |i>                               81  X1  X1  X1  X1  X1  X1  X1  X1  X1  X1  |i>  133  X2  Z2  Z1  HHS  Z1  HHS  HHS  |i>
30  X2  X2  |i>                               82  X2  X2  X2  X2  X2  X2  X2  X2  X2  X2  |i>  134  CZ  HHS  X1  Z2  X2  X2  X2  Z2  HHS  CZ  |i>
31  X2  Z2  X2  Z1  |i>                       83  CZ  CZ  CZ  |i>                             135  CZ  CZ  CZ  CZ  CZ  CZ  CZ  CZ  |i>
32  Z1  Z1  X1  Z1  Z1  X1  |i>               84  HHS  X2  X2  Z2  Z2  X1  X1  |i>           136  X1  HHS  X1  HHS  X1  HHS  |i>
33  Z1  X2  X2  |i>                           85  X1  CZ  Z2  X1  CZ  Z2  X1  CZ  Z2  |i>     137  X2  CZ  HHS  X2  Z2  HHS  X2  Z2  HHS  |i>
34  Z2  Z1  X2  Z2  Z2  Z2  Z1  X2  Z2  |i>   86  HHS  X1  Z2  X1  Z2  |i>                   138  Z1  X2  HHS  CZ  X1  HHS  CZ  Z2  X1  X1  |i>
35  Z2  X1  Z2  X2  Z2  X1  Z2  X2  |i>       87  Z1  Z2  X1  Z1  HHS  X2  |i>               139  CZ  CZ  CZ  CZ  CZ  CZ  CZ  CZ  CZ  |i>
36  Z1  Z1  X2  X2  |i>                       88  HHS  X1  X2  X1  X2  Z2  Z1  X1  |i>       140  X1  X1  HHS  Z2  HHS  HHS  X2  Z2  CZ  |i>
37  X1  X2  X1  |i>                           89  X1  X1  X1  HHS  Z2  |i>                   141  X1  CZ  HHS  CZ  HHS  Z1  CZ  CZ  X2  |i>
38  Z1  X2  Z1  X2  X1  Z1  |i>               90  HHS  CZ  |i>                               142  X1  Z2  HHS  CZ  CZ  X1  Z2  HHS  CZ  CZ  |i>
39  X1  X1  X1  X2  X2  Z2  |i>               91  Z1  CZ  Z1  Z2  HHS  |i>                   143  CZ  Z2  HHS  Z2  HHS  CZ  Z2  HHS  |i>
40  Z2  Z1  X2  Z1  X2  Z1  |i>               92  HHS  Z2  CZ  Z1  |i>                       144  HHS  CZ  X2  CZ  HHS  CZ  X2  CZ  |i>
41  Z1  Z1  X1  Z1  Z1  X1  Z1  X1  |i>       93  X2  Z2  HHS  CZ  |i>                       145  CZ  CZ  CZ  CZ  CZ  CZ  CZ  CZ  CZ  CZ  |i>
42  X1  X1  X1  |i>                           94  Z1  HHS  CZ  X2  Z2  |i>                   146  HHS  CZ  HHS  CZ  HHS  CZ  |i>
43  X1  X2  X1  X2  X1  |i>                   95  X2  Z2  Z1  HHS  CZ  |i>                   147  HHS  HHS  X2  HHS  HHS  X2  |i>
44  CZ  |i>                                   96  CZ  X2  CZ  CZ  X2  X2  Z2  Z2  |i>       148  Z1  HHS  HHS  X2  Z1  HHS  HHS  X2  |i>
45  CZ  Z2  |i>                               97  CZ  CZ  CZ  CZ  |i>                       149  X2  CZ  HHS  HHS  HHS  Z2  CZ  CZ  Z1  |i>
46  Z2  Z1  CZ  |i>                           98  Z1  X2  Z1  HHS  CZ  Z2  Z2  Z2  X2  |i>   150  X1  HHS  X1  HHS  X1  HHS  X1  HHS  |i>
47  CZ  Z2  Z1  Z2  |i>                       99  Z2  HHS  CZ  X2  X2  Z1  Z2  X1  X1  |i>   151  HHS  CZ  HHS  CZ  HHS  CZ  HHS  CZ  |i>
48  X1  CZ  Z2  |i>                           100  CZ  CZ  X1  CZ  CZ  X1  |i>               152  X1  HHS  X1  HHS  X1  HHS  X1  HHS  X1  HHS  |i>
49  Z2  Z1  X1  Z2  CZ  |i>                   101  X2  HHS  Z1  X1  X2  CZ  X2  |i>          153  HHS  HHS  X2  HHS  HHS  X2  HHS  HHS  X2  |i>
50  X1  X1  X1  X1  |i>                       102  CZ  CZ  HHS  |i>                          154  HHS  CZ  HHS  CZ  HHS  CZ  HHS  CZ  HHS  CZ  |i>
51  X2  X2  X2  X2  |i>                       103  CZ  CZ  X2  CZ  CZ  X2  |i>
```

We test the fault-tolerant protocol using both the transmon simulator and a real quantum processor on the IBM Q Experience [IBM2016]. For the simulation, we use the large five-transmon system defined in Section 3.4.5. The topology is sketched in Fig. 3.5. The qubit labels $q_0 q_1 q_2 q_3 q_4$ given in this figure are the same that are used to define the circuit components in Tab. 7.3 and Tab. 7.4. Note that the resonator r_5 from q_4 to q_0 has been

added to the simulation model to also permit the implementation of the state preparation of the encoded state $\overline{|00\rangle}$ given in Tab. 7.3. All simulations were performed on the supercomputers JURECA [Jül2018] and JUWELS [Jül2019].

For the real quantum processor, we use five qubits from the 16-qubit processor `ibmqx5` [IBM2018a]. The qubit mapping with respect to the circuit elements shown in Tab. 7.3 and Tab. 7.4 is $q_0 q_1 q_2 q_3 q_4 \mapsto Q_4 Q_3 Q_2 Q_{15} Q_{14}$, where $q_0 = Q_4$ is the ancillary qubit. This five-qubit subset of the 16-qubit device is also indicated in the topology graph of the simulation model in Fig. 3.5. Note that the resonator r_5 in this figure does not exist in the real device. Therefore, the encoded version of the initial state $|00\rangle$ in Tab. 7.3 cannot be performed on the real device. This means that only 5/6 of all experiments defined below can be executed on the processor (the 6 coming from 3 initial states times 2 circuit versions). The idea is that, if simulation and experiment agree for these 5/6 of all experiments, an extrapolation of the simulation results for the remaining 1/6 of the experiments may give an estimate for the performance on a potentially new quantum processor on which the additional connection between the two qubits would exist.

7.3.3 Results

For each of the 465 circuits given in Listing 7.1, we execute the corresponding bare and encoded versions. This defines a total of 930 experiments that are both simulated and run on the `ibmqx5` processor. We evaluate the statistical distances $D_{\text{bare/enc}}$ (see Eq. (7.18)) between the obtained distributions and the ideal probability distributions. All results are summarized in Fig. 7.8.

Figure 7.8(a) shows the statistical distances $D_{\text{bare/enc}}^{(\text{sim})}$ obtained from the transmon simulation model specified in Section 3.4.5 using the pulses with frequency tuning defined in Tab. F.5 and Tab. F.6 (results for pulses without frequency tuning are given in [Wil2018b]). As indicated in this figure, the longest time evolution takes approximately $10\,\mu\text{s}$. We see that for most of the 465 circuits, the encoded version performs better than the bare version, especially for longer time evolutions.

However, a particular exception stands out: In the left panel of Fig. 7.8(a), three encoded circuits (shown as green crosses) with circuit IDs $(40, 52, 56)$ and pulse durations $T_{\text{bare}}^{|00\rangle} \in (0.48\,\mu\text{s}, 0.64\,\mu\text{s}, 0.80\,\mu\text{s})$ have a significantly higher statistical distance than the corresponding bare circuits. Listing 7.1 shows that these circuits correspond to $(\text{X2 Z1})^w |00\rangle$ where $w \in (3, 4, 5)$. For the bare version, such a circuit basically resembles the repeated application of X^π pulses studied in Section 6.2. The reason for this is that the Z_1 gate needs no pulse but only affects the VZ phase of qubit 1 (cf. Section 5.1.1). Therefore, from the results shown in Fig. 6.1(a), we can expect a reasonably good performance for the bare version. The encoded version, however, requires the complicated state preparation of $\overline{|00\rangle}$ (see Tab. 7.3). Furthermore, each $\overline{\text{X2}} = X_1 X_2$ gate requires two X^π pulses (see Tab. 7.3), whose phases γ change by π because of the intermediate $\overline{\text{Z1}} = Z_1 Z_2$ gates (cf. Section 5.1.1). The fact that we do not see this effect for the encoded circuit $\overline{\text{X2}}^4 \overline{|00\rangle}$ with ID 51 implies that it is indeed the different phases that cause the deviation. Apparently, this leads to a significant phase error for the encoded circuit that cannot be detected by postselection anymore. The most severe case with ID 56 can also be seen
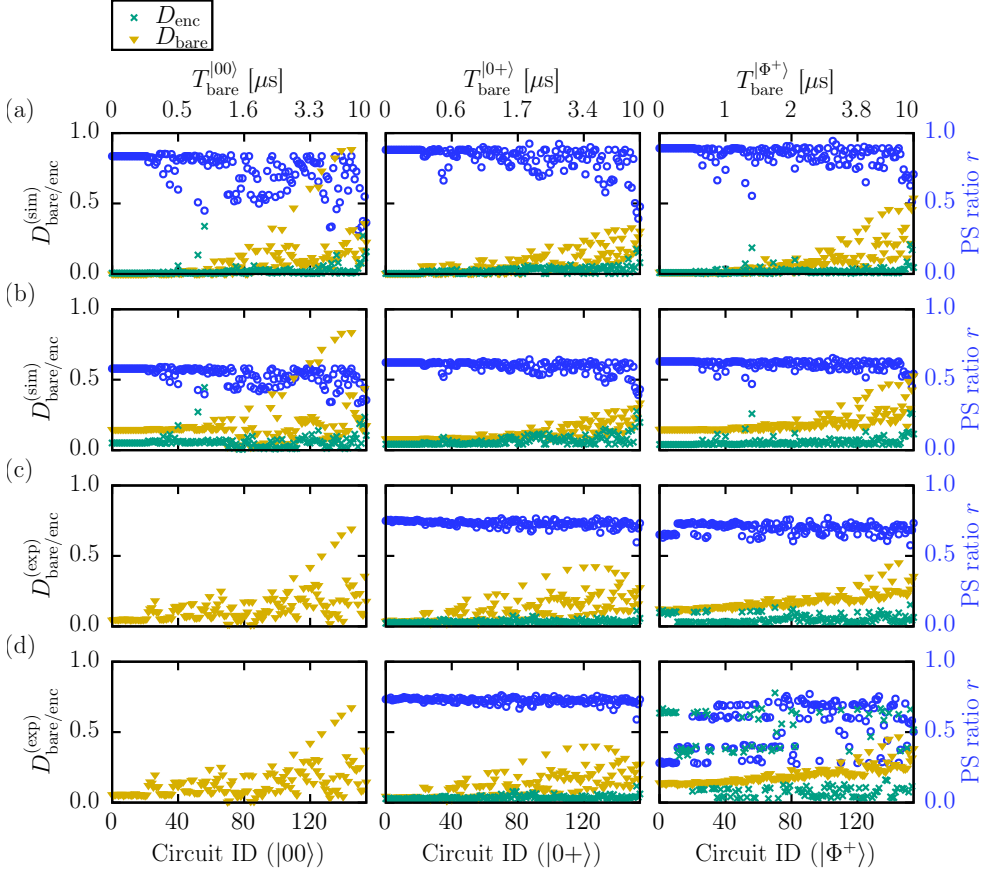
Figure 7.8: Test of the fault-tolerant protocol using (a) the transmon simulation model defined in Section 3.4.5; (b) the same model with an additional measurement error of $\epsilon = 0.08$ (cf. Eq. (7.19)); (c) the `ibmqx5` [IBM2018a] on April 19, 2018; (d) the `ibmqx5` on April 20, 2018. Shown are the statistical distances D_{bare} for the bare circuits (yellow triangles) and D_{enc} for the encoded circuits (green crosses) as defined in Eq. (7.18), as well as the corresponding postselection ratios r (blue circles). The three panels in each row correspond to the three initial states in Tab. 7.3 (the left panels in (c) and (d) only contain data for the bare circuits because the encoded circuits for the initial state $|00\rangle$ cannot be run on the `ibmqx5`). The circuit IDs on the bottom axis are defined in Listing 7.1. They are sorted by the total time $T_{\text{bare}}^{|00\rangle}$ required to simulate all pulses of the bare circuits for the initial state $|00\rangle$ (some representative time scales are indicated on the top axis).

for the initial state $\overline{|\Phi^+\rangle}$ in the right panel of Fig. 6.1(a). Still, it is remarkable that the encoded circuits perform so much better in all the other cases, even though the state preparations in Tab. 7.3 (especially that for $\overline{|00\rangle}$) are much more involved.

In the simulations, we obtain the distributions $p_{j_0 j_1}^{\text{(bare/enc)}}$ for Eq. (7.18) directly from the state vector $|\Psi\rangle$ given by Eq. (3.23). In particular, this means that no measurement error is included in the simulations. To understand the effect of measurement errors on the fault-tolerant scheme, we consider the simple model that for each qubit, a 0 (1) is mistakenly counted as a 1 (0) with probability ϵ. For a distribution p_J of n-bit strings $J = j_0 j_1 \cdots j_{n-1}$, this model is implemented by the transformation

$$ p_J \mapsto \sum_{J'=0}^{2^n - 1} p_{J'} \epsilon^{\Delta(J, J')} (1 - \epsilon)^{n - \Delta(J, J')}, \tag{7.19} $$

where $\Delta(J, J')$ is the Hamming distance between J and J', i.e., the minimum number of bits that must be flipped to achieve $J = J'$. We have $n = 2$ in the bare case and $n = 5$ in the encoded case. Technically, a measurement error according to Eq. (7.19) can be interpreted as a depolarizing channel (see Eq. (7.12)) immediately before the measurement.

Figure 7.8(b) shows the effect of adding a measurement error of $\epsilon = 0.08$ to the simulation results from Fig. 7.8(a). The overall effect is that the statistical distances $D_{\text{bare/enc}}^{\text{(sim)}}$ increase, but those of the bare circuits increase more than those of the encoded circuits. Thus, measurement errors of this sort can be detected and mitigated using the fault-tolerant scheme, which is also reflected by an overall decrease in the postselection ratios in Fig. 7.8(b). However, note that the particular family of encoded circuits $(\overline{X2}\,\overline{Z1})^w$ discussed above still stands out by performing worse than the corresponding unencoded circuits.

In Fig. 7.8(c), we present results obtained from running the fault-tolerance test on the `ibmqx5` [IBM2018a] on April 19, 2018. For the bare versions, all 465 circuits given in Listing 7.1 can be executed. For the encoded versions, 155 out of the 465 circuits (corresponding to the initial state $\overline{|00\rangle}$) cannot be executed, since the resonator r_5 in Fig. 3.5 does not exist in the device. Therefore, the left panel of Fig. 7.8(c) only shows results for the bare circuits.

We see that using the fault-tolerant protocol systematically improves the results. Furthermore, we find good qualitative agreement between experiment and simulation. Extrapolating the simulation results suggests that encoded circuits for the initial state $\overline{|00\rangle}$ would perform similarly well if the device were extended by the appropriate connection.

We repeated the experiments on the `ibmqx5` several times. Most of the time, the results were of the type shown in Fig. 7.8(c). However, the fault-tolerance test was not successful every time we ran the experiment. One such result from April 20, 2018 is shown in Fig. 7.8(d). As one can see, many of the encoded circuits for the initial state $\overline{|\Phi^+\rangle}$ (right panel of Fig. 7.8(d)) have unusually high statistical distances and low postselection ratios. We examined the corresponding calibration parameters and found that qubits Q_{14} and Q_{15} had higher readout errors around 12%. Typically, the reported readout errors are between 4% and 10%. However, this alone cannot explain the high error rates observed in the experiment.

Table 7.5: Evaluation of the data for bit strings produced by the 16-qubit processor `ibmqx5` [IBM2018a] for the encoded circuit Z2^4 $|\Phi^+\rangle$. The circuit has ID 4 in Listing 7.1 and the result corresponds to one of the leftmost green crosses on the right panels in Fig. 7.8(c) and (d). Shown are the 20 most frequent outcomes for the experiment on April 19 (left) and April 20 (right), sorted by relative frequency. To demonstrate the postselection procedure, discarded bit strings are highlighted in red, with the reason given in the third column. Bit strings that are not discarded are highlighted in blue. The qubit mapping on the device is $q_0q_1q_2q_3q_4 \mapsto Q_4Q_3Q_2Q_{15}Q_{14}$, where $q_0 = Q_4$ is the ancillary qubit (see also Fig. 3.5). As this data reveals, a hardware fault apparently caused a bit flip of qubit Q_2 or Q_{15} with very high probability on April 20 such that the four most frequent outcomes are erroneously discarded.

April 19, 2018			April 20, 2018		
Outcome	Frequency	Counted as	Outcome	Frequency	Counted as
0 0000	0.163	$\overline{\|00\rangle}$	0 0100	0.167	(odd parity)
0 1001	0.161	$\overline{\|11\rangle}$	0 1011	0.166	(odd parity)
0 0110	0.130	$\overline{\|11\rangle}$	0 0010	0.160	(odd parity)
0 1111	0.123	$\overline{\|00\rangle}$	0 1101	0.153	(odd parity)
0 0010	0.061	(odd parity)	0 1100	0.054	$\overline{\|01\rangle}$
0 1011	0.053	(odd parity)	0 1010	0.051	$\overline{\|10\rangle}$
0 0100	0.050	(odd parity)	0 0011	0.043	$\overline{\|01\rangle}$
0 1101	0.049	(odd parity)	0 0000	0.043	$\overline{\|00\rangle}$
0 1000	0.037	(odd parity)	0 0101	0.038	$\overline{\|10\rangle}$
0 0001	0.025	(odd parity)	0 1001	0.037	$\overline{\|11\rangle}$
0 1110	0.023	(odd parity)	0 1000	0.014	(odd parity)
0 1010	0.021	$\overline{\|10\rangle}$	0 0110	0.011	$\overline{\|11\rangle}$
0 1100	0.019	$\overline{\|01\rangle}$	0 0001	0.010	(odd parity)
0 0111	0.018	(odd parity)	1 0010	0.009	(wrong ancilla)
0 0011	0.012	$\overline{\|01\rangle}$	0 1111	0.009	$\overline{\|00\rangle}$
0 0101	0.012	$\overline{\|10\rangle}$	1 0100	0.006	(wrong ancilla)
1 0000	0.008	(wrong ancilla)	1 1011	0.006	(wrong ancilla)
1 1001	0.006	(wrong ancilla)	1 1101	0.005	(wrong ancilla)
1 1111	0.006	(wrong ancilla)	0 1110	0.004	(odd parity)
1 0110	0.006	(wrong ancilla)	0 0111	0.002	(odd parity)

Therefore, we take a closer look at the data. Table 7.5 shows results obtained for the encoded circuit $Z2^4 |\Phi^+\rangle$ corresponding to ID 4 in Listing 7.1. The table shows the most frequent five-bit strings obtained on April 19 (left) and April 20 (right). After evaluating the data using the postselection procedure illustrated in the table, we obtain the frequency distributions $p^{(\text{enc})}_{j_0 j_1}$, which are used to evaluate the statistical distance $D^{(\text{exp})}_{\text{enc}}$ given by Eq. (7.18). On April 19, we obtained $D^{(\text{exp})}_{\text{enc}} \approx 0.10$ (shown as one of the first green crosses in the right panel of Fig. 7.8(c)). On April 20, however, we found $D^{(\text{exp})}_{\text{enc}} \approx 0.65$, which corresponds to the much worse result for ID 4 in the right panel of Fig. 7.8(d).

Table 7.5 reveals the reason for this: on April 20, the four most frequent bit strings were discarded due to an odd parity. In other words, they cannot be assigned to one of the logical basis states given by Eqs. (7.15a)–(7.15d). If one of the central two bits corresponding to qubit Q_2 and Q_{15} were flipped, however, the result would be almost the same as on April 19. It seems that a hardware fault caused a bit flip of qubit Q_2 or Q_{15} with very high probability on April 20. Apparently, this systematic error cannot be corrected by the fault-tolerant protocol. We remark that, although reproducible, this problem occurred only in a minority of all runs on the IBM Q processor.

In conclusion, we find that the fault-tolerant protocol provides a systematic procedure to improve the results by encoding a logical state redundantly in a larger number of physical qubits. This is remarkable because the errors in the transmon simulation model and in the real device are not at all guaranteed to be of the simple type assumed in the design of the protocol (see, for instance, Fig. 7.3(a)). Especially, the long circuit used to encode the initial state $|00\rangle$ (see Tab. 7.3) might have seemed unlikely to improve the results. However, we also see that all encoded gate elements used in the scheme do not require two-qubit gates (see Tab. 7.4). Instead, these more error-prone gates (cf. Figs. 6.1 and 6.2) occur only in the initial state preparation of a circuit. This design feature may provide an alternative explanation why the particular protocol under investigation can improve the results. Furthermore, in [Wil2018b], we also studied the performance of the fault-tolerant protocol in the presence of an environment (see also the model studied in Section 4.3). These results suggest that the scheme does not provide significant improvements when the errors are dominated by decoherence, and a related conclusion was drawn for existing stabilizer codes in [Nau2018]. Nevertheless, as long as the errors in quantum information processors are dominated by the inherent control and measurement errors in transmon systems, our results suggest that the performance of a quantum computer can be systematically improved by using an appropriate fault-tolerant protocol.

7.4 Conclusions

The purpose of the first experiment studied in this chapter was to observe crosstalk effects, predicted by the transmon simulator, in an IBM Q processor. We designed a class of circuits for this purpose (see Fig. 7.1), inspired by the state-dependent frequency shifts observed in previous simulations (see Fig. 4.11 and the fifth row in Tab. 6.6). The five-transmon simulation results and experimental results from the `ibmqx4` [IBM2018b] showed almost perfect agreement for time evolutions up to several microseconds (see Fig. 7.3). We also observed that these effects are generic for the transmon architecture in the sense

that they did not depend on the exact values of the device parameters. Moreover, the simulation results suggested that the effects can be systematically reduced or enhanced, which was also confirmed by the experiment (see Fig. 7.3(b)). This observation suggests that the correlated crosstalk errors inherently included in the full dynamics of the simulated transmon system (see Fig. 7.2) are a very good model for the errors in the real processor.

For a set of experiments on the singlet state, we addressed the question if the errors previously observed for a five-transmon processor [Mic2017; Wil2017] can be described purely in terms of miscalibrated pulses for a simulated two-transmon system. To a certain extent, this was possible (see Fig. 7.5), but the two-transmon model does not seem to be capable of describing all deviations from the ideal result. It would be interesting to see if a more extensive analysis using the five-transmon simulation is more appropriate, as indicated by previous results given in Figs. 6.2 and 7.3. However, in this case, we found that simple error channels such as a depolarizing channel and an amplitude damping channel offer a much simpler way of describing the observed results (see Fig. 7.7).

Finally, we tested a full protocol from the theory of quantum fault tolerance, motivated by the observation that simple error channels could describe the errors seen in the singlet-state experiment (cf. Fig. 7.7). The extensive test comprised a total of 930 quantum circuits that were both simulated using the large five-transmon system defined in Section 3.4.5 and run on the processor `ibmqx5` [IBM2018a]. An analysis of the experimental results revealed a systematic hardware fault in the processor during some runs (see Tab. 7.5). Most of the time, however, we observed that the fault-tolerant protocol provides a systematic procedure to improve the results. As this was true for both simulations and experiments (see Fig. 7.8(a)–(c)), we conclude that the fault-tolerant protocol systematically improves the quantum computer's performance if the errors are due to the intrinsic control and measurement errors present in transmon systems.

Given the general qualitative and often also quantitative agreement between transmon simulation and experiment, we can conclude that the IBM Q processors have been engineered very carefully to implement the quantum theoretical model of superconducting transmon systems. It is interesting to see if further development of the processors can also bring them sufficiently close to an implementation of the computational model of an ideal gate-based quantum computer.

Chapter 8

Discussion and conclusion

The goal of this project was to develop a transmon simulator that utilizes the resources of digital supercomputers to study the emerging technology of transmon quantum computers. NISQ devices of this architecture are currently built by several companies such as IBM [IBM2016], Google [Aru2019], and Rigetti Computing [Rig2017]. We designed and implemented a simulation algorithm that computes the real-time dynamics of a system of transmons and couplers by solving the time-dependent Schrödinger equation (TDSE) for a generic model Hamiltonian representing the quantum computing hardware.

The model features an arbitrary number of transmons and resonators, as well as time-dependent pulses used to implement quantum gates on the qubits. Furthermore, we described a way to simulate electromagnetic environments with the model, together with a systematic procedure to extract suitable model parameters from experiments or electromagnetic solvers. The numerical algorithm used to solve the TDSE is unconditionally stable [DeR1987] and can be used to obtain the dynamics for several hundred microseconds on a sub-picosecond scale. In principle, the size of the model is only limited by the available computational resources on the supercomputer. In this work, we presented results for the simulation of up to 16 transmons and resonators, described by more than four billion complex coefficients.

We used simulations of free time evolutions to benchmark the simulation algorithm and found excellent weak and strong scaling on the supercomputer JURECA [Jül2018]. By studying the algorithm's accuracy, we demonstrated that recently proven error bounds for the product-formula algorithm [Wil2020d] are tight.

The simulation approach inherently includes effects beyond the ideal gate-based quantum computer model, such as leakage to higher, non-computational states, crosstalk between the transmons, entanglement between transmons and resonators, and control errors through imperfect pulses applied to the system. All of these effects are known to be limiting factors in current transmon architectures [Woo2018; The2018; Aru2019].

Additionally, we performed experiments on real transmon devices available on the IBM Q Experience [IBM2016]. As the size of these NISQ devices is in the range of what can be simulated on the supercomputers JURECA [Jül2018] and JUWELS [Jül2019], it was possible to relate the simulation results directly to data obtained from real quantum processors.

Relation to perturbative results, master equations, and experiments

We investigated known perturbative results for the model and found that, while they provide a simple, effective description of the system, they develop a drift in time that makes them unsuitable for pulse optimizations. By simulating the dynamics of a transmon system coupled to a bath of harmonic oscillators, we observed a transition from TDSE-based approaches to a Lindblad master equation, with the result that the latter can be an adequate, effective description under certain conditions. Finally, we characterized the resonator-mediated exchange interaction between transmons and found that it is well described by an appropriate ZZ interaction on a two-qubit subspace. We demonstrated that the strength of this interaction can be accurately determined from simulated time evolutions or experiments such as gate set tomography.

Characterizing and predicting the performance of optimized quantum gates

By studying an optimization procedure for the pulses used to implement quantum gates, we found that the Nelder–Mead method is a suitable candidate to obtain pulse parameters that have error rates of the same magnitude as those reported in experiments. In fact, the pulses used for the transmon simulator were often found to perform better in actual applications. This was especially true for two variants (CR1 and CR4) of the echoed cross-resonance pulse (CR2) that is routinely used to implement the two-qubit CNOT gate [IBM2016].

We proved two statements to relate the average gate fidelity [Nie2002] to more sophisticated gate metrics such as the diamond distance [Kit1997]. Unlike previous results, the relations also apply to trace-decreasing quantum operations, which are relevant for transmon systems where leakage is an important limitation. We found that the gate metrics provide useful information about the amount of leakage and the accuracy of a single application of a quantum gate pulse.

However, none of these gate metrics was found to be suitable for predicting the performance of repeated pulse applications in actual quantum algorithms. We observed several cases with poor gate metrics and exceptional performance, as well as almost ideal gate metrics but bad performance in practical applications (see also [Wil2017; McK2019]). Only the diamond distance turned out to provide an upper bound on the observable statistical distance, and even in this case, the diamond distance also needed to be evaluated for repeated pulse applications.

As an alternative to the common gate metrics, we extensively studied the approach of gate set tomography to characterize quantum gates. Gate set tomography requires only the observable relative frequencies from experiments, and we found that especially the completely positive trace-preserving estimates have an exceptional predictive power for the performance of quantum gates in actual applications.

Crosstalk, effective error models, and quantum fault tolerance

We studied a family of quantum circuits designed to probe crosstalk between transmon qubits. We found that the time evolution predicted by the transmon simulator was in excellent agreement with experimental results obtained from the processor `ibmqx4` [IBM2018b], suggesting that the simulation model provides a very good description of the errors present in the device.

For a class of quantum circuits designed to characterize the singlet state, we found that the deviations from the ideal result are most likely not caused by miscalibrated pulses. However, the deviations could be well described in terms of simple error channels from the theory of quantum fault tolerance.

This motivated us to perform an extensive test of a fault-tolerant protocol [Got2016]. Although the protocol cannot correct decoherence errors [Wil2018b], both simulations and experiments on the `ibmqx5` [IBM2018a] suggest that the protocol systematically improves the quantum computer's performance in the presence of control and measurement errors characteristic of the transmon architecture.

Outlook

Many scenarios studied in this work reflect the empirical observation that engineering accurate gate-based quantum computers is remarkably difficult. It is important that detailed simulations such as those presented in this thesis are continuously carried out to understand limitations and find potential ways to overcome systematic errors. There are many interesting paths along which the present work can be continued:

- **Environment simulations:** In its current form, the model Hamiltonian defined in Section 3.2 supports a bath of harmonic oscillators. This is a bosonic bath that can be used to model electromagnetic environments using the procedure described in Section 3.5. It would be interesting to extend the formalism to multiple transmons by following [Nig2012; Ans2019]. Also, the bath simulations can be extended as described in Section 4.3.3. In particular, it would be interesting to analyze the consequences of replacing the bosonic bath with a fermionic bath or a spin bath as studied in [Wil2018b; Wil2020d; Wil2020e].

- **Pulse optimization:** The problem of finding optimal pulses to implement quantum gates is an area of active research. With the success of advanced machine learning techniques such as deep reinforcement learning (see Section 5.5), it would be interesting to find and investigate new pulses using these approaches as done in [Niu2019; An2019]. Also, IBM has recently introduced OpenPulse [McK2018], by which new pulses studied with the transmon simulator (such as the CR variants in Fig. 5.1) can directly be executed on the device. It would be compelling to use this interface to compare pulses for the transmon simulator directly with experiments on the real device.

- **Modeling experiments:** For some of the experiments studied in this project, minor differences between simulation and experiment were still observable. Future work could go into understanding these differences in detail. For instance, the tiny oscillations on the black curve in Fig. 7.3(a) appear to be smoothed out in the experiment, which may be understood by adding an environment to the simulation in the sense of Section 3.5 or Fig. 4.9. Also, it would be interesting to see if a five-transmon simulation for the singlet experiment can describe the deviations better than the two-transmon simulation (see Fig. 7.5). A completely different approach would be to study classical models of quantum information devices (see [Bla2016; Iva2018]) and to understand in which respect their descriptive power differs from the quantum theoretical model used in this work.

- **Extending the simulation algorithm:** A promising direction would be to extend the simulation model with time-dependent magnetic fluxes to include flux-tunable transmons. This would require modifications to the numerical algorithm described in Section 3.3 and enable the simulation of transmon systems controlled by high-speed flux lines [Rot2017] as pursued by the European FET Flagship project OpenSuperQ [Ope2018]. Research efforts in this direction are currently underway.

Appendices

Appendix A

Visualization of quantum gate implementations

In Fig. A.1, we show a screenshot of the scene rendered by `visualizer` (see Section 3.3.3) using the `C++` engine Irrlicht [Geb2005]. The scene is generated from time-evolution data of the five-qubit transmon system used to study repeated gate applications in Section 6.2. It shows the final state after applying two successive X gates on each of the qubits. Although the gate sequence should technically compose an identity operation, one can see that the Bloch vectors are not perfectly straight, despite excellent gate metrics of the X gate pulses (especially for $X_3^{\pi/2}$, see Tab. 6.2).

Figure A.2 shows a selection of visualizations resulting from the simulation of a quantum circuit with random gates from the standard gate set (cf. Appendix B) at three different points in time using a five-transmon simulation (first row) and the ideal quantum computer simulator JUQCS [DeR2019a; Wil2020a] (second row). The particular circuit reveals how crosstalk results in state-dependent frequencies, inducing phase errors that can dramatically increase the error in the system's output distribution in certain cases (see also Fig. 7.3).
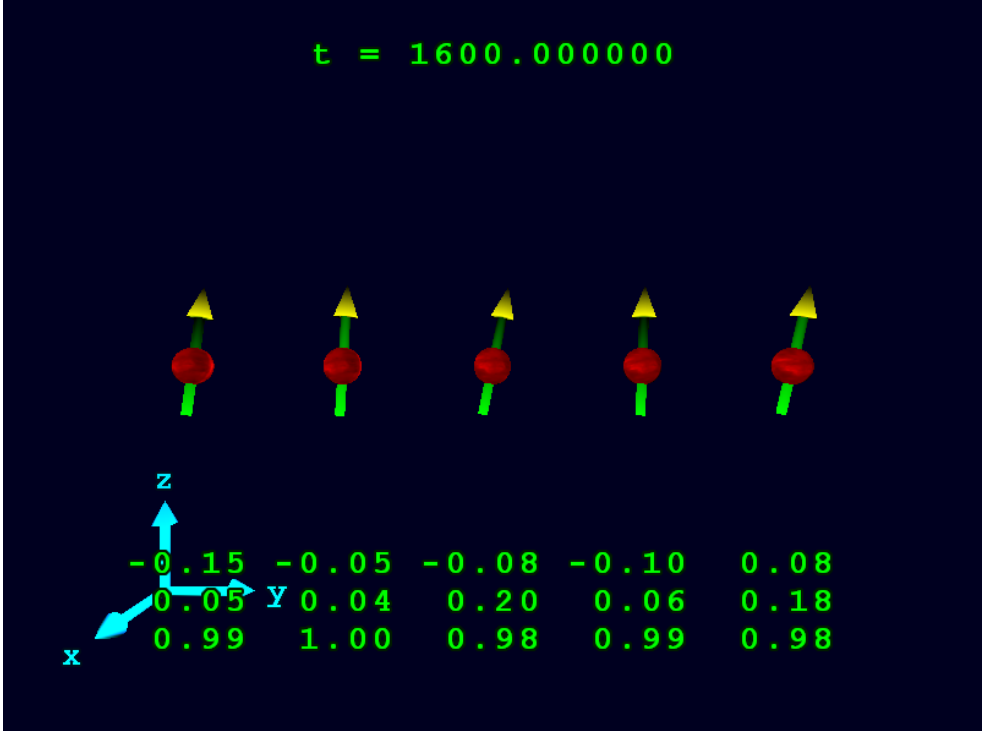
Figure A.1: Example of the scene rendered by `visualizer`. The screenshot shows the Bloch vectors $\vec{r}_i(t)$ (see Eq. (3.38)) of five transmon qubits after the successive application of two X gates (implemented as four $X^{\pi/2}$ gates) on each of the qubits. The simulated system is the small five-transmon system defined in Section 3.4.4. The corresponding gate metrics are given in Tab. 6.2. The Bloch vectors corresponding to qubit $i = 0, \ldots, 4$ are shown from left to right. As the pulse for an $X^{\pi/2}$ gate for this system takes $80\,\text{ns}$ (see Tab. F.3), the depicted final time is $t = 1600\,\text{ns}$. Note that the Bloch vectors at the end of the time evolution are not perfectly straight. In particular, the second and the fourth Bloch vector are similarly tilted even though the corresponding gate metrics (see $X_1^{\pi/2}$ and $X_3^{\pi/2}$ in Tab. 6.2) are of very different quality.

Figure A.2: Selection of scenes rendered by visualizer. The screenshot shows the Bloch vectors $\vec{r}_i(t)$ (see Eq. (3.38)) of five qubits at the indicated times t during the application of a random sequence of quantum gates. The top three images correspond to the time evolution of a five-qubit transmon system where the gates are implemented by pulses. The bottom three images show the exact Bloch vectors predicted by the ideal pen-and-paper quantum computer model described in Chapter 2 and simulated using JUQCS [DeR2019a; Wil2020a]. The rotating frames are different such that different directions of the arrows in the xy plane may not indicate an error. The projection on the z axis, however, indicates errors such that measuring the first qubit shown in the two images on the right would lead to an error of over 50%.

Appendix B

Elementary gate set used for the simulation

In Tab. B.1, we summarize the set of single-qubit gates and their representations in terms of the elementary single-qubit gates U1, U2, and U3 defined in Eqs. (2.15a)–(2.15c). The two-qubit gate used to make the gate set universal [DiV2000] is the CNOT gate defined in Eq. (2.18).

Since the topic of this thesis is the simulation of transmon qubit systems, and the first platform offering access to transmon qubit processors was the IBM Q Experience [IBM2016], the gate set implemented for the simulation is inspired by the gates available on the IBM Q Experience [Cro2017]. Additionally, the gate set includes gates implemented by JUQCS, which was used to compare the simulated gates to their equivalent on an ideal universal quantum computer (see [DeR2007; DeR2019a; Wil2020a] for more information on JUQCS). At the time of writing, JUQCS was used to simulate some of the larger circuits for Google's quantum supremacy experiment [Aru2019], as well as universal quantum circuits with up to 48 qubits, setting the largest simulation of a universal gate-based quantum computer to date [DeR2019a]. In this context, *universal* means that arbitrary circuits can be simulated. If, however, only a specific subset of circuits with a small number of entangling gates need to be simulated, a much larger number of qubits is possible [Ped2017; Boi2017; Che2018a; Che2018b; Mar2018; Vil2020; Wil2020a].

Table B.1: Summary of all single-qubit gates used in this work, including their circuit symbols and matrix representations. The gates listed in the first column and the alternative representations listed in the second column are only equivalent up to a global phase. The relations to the U gates defined in Eqs. (2.15a)–(2.15c) are given in the last column if the gates are implemented in this way. For any gate G, there is also the inverse operation $G^\dagger$ given by the Hermitian conjugate of the corresponding matrix.

Gate	Alternatives	Symbol	Matrix	In U gates
I	Identity, Idle	$-\boxed{I}-$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	—
$U1(\lambda)$	U1 gate	$-\boxed{U1(\lambda)}-$	$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\lambda} \end{pmatrix}$	—
$U2(\phi, \lambda)$	U2 gate	$-\boxed{U2(\phi,\lambda)}-$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -e^{i\lambda} \\ e^{i\phi} & e^{i(\phi+\lambda)} \end{pmatrix}$	—
$U3(\theta, \phi, \lambda)$	U3 gate	$-\boxed{U3(\theta,\phi,\lambda)}-$	$\begin{pmatrix} \cos\frac{\theta}{2} & -e^{i\lambda}\sin\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} & e^{i(\phi+\lambda)}\cos\frac{\theta}{2} \end{pmatrix}$	—
X	Bit flip	$-\boxed{X}-$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$U3(\pi, 0, \pi)$
Y	Bit&Phase flip	$-\boxed{Y}-$	$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$	$U3(\pi, \frac{\pi}{2}, \frac{\pi}{2})$
Z	Phase flip	$-\boxed{Z}-$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$U1(\pi)$
H	Hadamard	$-\boxed{H}-$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$	$U2(0, \pi)$
S	$Z^{\pi/2}, R^z(\frac{\pi}{2})$	$-\boxed{S}-$	$\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$	$U1(\frac{\pi}{2})$
T	$Z^{\pi/4}, R^z(\frac{\pi}{4})$	$-\boxed{T}-$	$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$	$U1(\frac{\pi}{4})$
$+X$	$X^{-\pi/2}, R^x(-\frac{\pi}{2})$	$-\boxed{+X}-$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$	$U2(\frac{\pi}{2}, -\frac{\pi}{2})$
$-X$	$X^{\pi/2}, R^x(\frac{\pi}{2})$	$-\boxed{-X}-$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}$	$U2(-\frac{\pi}{2}, \frac{\pi}{2})$
$+Y$	$Y^{-\pi/2}, R^y(-\frac{\pi}{2})$	$-\boxed{+Y}-$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$	$U2(-\pi, \pi)$
$-Y$	$Y^{\pi/2}, R^y(\frac{\pi}{2})$	$-\boxed{-Y}-$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$	$U2(0, 0)$
$R(k)$	Phase gate	$-\boxed{R(k)}-$	$\begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{pmatrix}$	$U1(\frac{2\pi}{2^k})$

Appendix C

The reason for linear and unitary transformations in quantum theory

The computational model of a gate-based quantum computer described in Chapter 2 is largely based on linear and unitary transformations to describe transitions between states. In fact, the use of unitary matrices can be seen as the characteristic difference between quantum computers and digital or probabilistic machines (which would use Boolean or stochastic matrices, respectively) [Ber1997]. An interesting question is therefore: Where did the use of unitary transformations in quantum theory come from?

To address this question, we review some of the mathematical arguments that have been given during the development of quantum theory for the use of linear and unitary transformations. Note that none of these arguments can prove that Nature has to be described by linear equations and unitary maps; they only illustrate why humans have developed quantum theory on the basis of linear and unitary transformations.

Historically, the question why quantum theory should be linear and unitary was first posed when the TDSE given by Eq. (1.1) was introduced by Schrödinger [Sch1926a; Sch1926b]. As the emergence of his equation was considered "ad hoc" by some researchers, they started a search for arguments why transformations of a quantum state $|\Psi\rangle$ should be linear and unitary. The first such argument was given by Wigner [Wig1931; Wig1959] and later formulated more rigorously by Lomont and Mendelson [Lom1963] and Bargmann [Bar1964]. It has become widely known as *Wigner's theorem*.

C.1 Wigner's theorem

The argument behind Wigner's theorem starts from the assumption that in any experiment, one can only ever observe transition probabilities of the form $|\langle\phi|\psi\rangle|^2$ between quantum states $|\phi\rangle$ and $|\psi\rangle$. Since the laws of Nature are believed to be invariant under space-time symmetry operations, the observable transition probabilities shall be conserved. Mathematically, this means that any transformation U between quantum states shall conserve the absolute value of the inner product between complex vectors, i.e.,

$$\forall |\phi\rangle, |\psi\rangle: \quad |\langle\phi|\psi\rangle| = |\langle\phi'|\psi'\rangle|, \tag{C.1}$$

where $|\phi'\rangle = U(|\phi\rangle)$ and $|\psi'\rangle = U(|\psi\rangle)$. Wigner's theorem essentially proves that among all conceivable transformations U, only *linear and unitary* or *antilinear and antiunitary* operators are compatible with Eq. (C.1). The hardest part of the proof is to show that U must be either linear or antilinear.

Another equivalent formulation of Wigner's theorem often seen in the earlier literature is stated in terms of equivalence classes $\boldsymbol{\psi} := \{c|\psi\rangle : c \in \mathbb{C}, |c| = 1\}$ (so-called *rays*) [Lom1963; Bar1964]. The motivation for this is the idea that two complex vectors $|\psi\rangle$ and $|\psi'\rangle$ describe the same physical state if they differ only by a complex phase. Therefore, only the equivalence class $\boldsymbol{\psi}$ is understood to represent the actual physical state. Wigner's

condition then takes the form

$$\forall \boldsymbol{\phi}, \boldsymbol{\psi} : \quad (\boldsymbol{\phi}, \boldsymbol{\psi}) = (\mathbf{T}(\boldsymbol{\phi}), \mathbf{T}(\boldsymbol{\psi})), \tag{C.2}$$

where $\mathbf{T}$ denotes the given symmetry transformation. The square of the inner product $(\boldsymbol{\psi}, \boldsymbol{\phi})$ in this space can directly be interpreted as the transition probability between $\boldsymbol{\psi}$ and $\boldsymbol{\phi}$. To obtain the more abstract formulation given in Eq. (C.1), the ray expressions given in Eq. (C.2) are defined as $\mathcal{T}(\boldsymbol{\psi}) := U(|\psi\rangle)$ and $(\boldsymbol{\psi}, \boldsymbol{\phi}) := |\langle \psi | \phi \rangle|$, where the complex vectors $|\psi\rangle \in \boldsymbol{\psi}$ and $|\phi\rangle \in \boldsymbol{\phi}$ are the representatives chosen for the computation. This ray formulation of Wigner's theorem comes at the cost of having to prove that the ray expressions are well-defined, in the sense that the results do not depend on the representatives.

There is also a more recent formulation of Wigner's condition that can be seen as combining the mathematical convenience of Eq. (C.1) and the interpretational character of Eq. (C.2). It is based on the density matrix representations of the pure states $\rho_\psi = |\psi\rangle\langle\psi|$ and $\rho_\phi = |\phi\rangle\langle\phi|$. These objects are rank-1 projectors and have the advantage that they do not entail the ambiguity of a complex phase factor in representing the same physical state. In this formulation, Wigner's condition reads

$$\forall \rho_\psi, \rho_\phi : \quad \text{Tr}\, \rho_\psi \rho_\phi = \text{Tr}\, f(\rho_\psi) f(\rho_\phi), \tag{C.3}$$

where the symmetry transformation is denoted by the function f (cf. the generalized Born rule in Eq. (6.23)). In this form, Wigner's theorem has been extended to non-bijective transformations [Geh2014]. Instead of *linear and unitary* or *antilinear and antiunitary* transformations, it then states that the function f has to be implemented by a *linear or antilinear isometry* W such that $f(\rho_\psi) = W \rho_\psi W^\dagger$.

Sketch of the proof

In what follows, we briefly sketch the main ideas of the proof given by Gehér [Geh2014]. To keep it simple, we consider bijective transformations and focus on the formulation of the condition stated in Eq. (C.1).

Let $\{|i\rangle\}$ denote an orthonormal basis (ONB) of the Hilbert space $\mathcal{H}$. Note that, in the context of quantum computing, we almost always deal with finite-dimensional Hilbert spaces such that $i \in \{1, \ldots, N\}$ where $N = \dim \mathcal{H}$ (see below for more information on the infinite-dimensional case).

The proof first defines a set of N vectors $\{|i'\rangle\}$ as the images of this basis under U, i.e., $|i'\rangle := U(|i\rangle)$. One then constructs a linear or antilinear map $\widetilde{U} : \mathcal{H} \to \mathcal{H}$ from these images such that $\widetilde{U} |i\rangle := |i'\rangle$. Using Eq. (C.1), one can show that $\{|i'\rangle\}$ is also an ONB of $\mathcal{H}$. By studying the action of U on the states $|i\rangle - |i+1\rangle$ and $|i\rangle + e^{i\pi/2} |i+1\rangle$, one can show that the given transformation U and the constructed $\widetilde{U}$ coincide, and that there are only two options, namely the linear, unitary and the antilinear, antiunitary case. It is instructive to study both cases and their relevance in quantum theory separately.

The linear, unitary case

In the linear case, we have

$$U(a \ket{\phi} + b \ket{\psi}) = aU(\ket{\phi}) + bU(\ket{\psi}) \tag{C.4}$$

for all states $\ket{\phi}$ and $\ket{\psi}$ and all $a, b \in \mathbb{C}$. Given that U is linear, it is quite straightforward to prove that U has to be unitary. We give a simple proof of this statement.

Since U is linear, it can be represented by a matrix with matrix elements

$$\langle i|U^\dagger U|j \rangle = |\langle i|U^\dagger U|j \rangle|e^{i\varphi_{ij}} = |\langle i|j \rangle|e^{i\varphi_{ij}} = \delta_{ij}e^{i\varphi_{ij}}, \tag{C.5}$$

where $\varphi_{ij} \in [0, 2\pi)$ and we have used Eq. (C.1) to eliminate $U^\dagger U$. In the case $i \neq j$, we have $\langle i|U^\dagger U|j \rangle = 0$, so the phase φ_{ij} is irrelevant. In the case $i = j$, the left-hand side of Eq. (C.5) reads $\langle i|U^\dagger U|i \rangle = \|U \ket{i}\|^2 > 0$, so the phase factor $e^{i\varphi_{ii}} = 1$. Hence we have $\langle i|U^\dagger U|j \rangle = \delta_{ij}$ for all i, j. A similar argument yields $\langle i|UU^\dagger|j \rangle = \delta_{ij}$, so we have $U^\dagger U = UU^\dagger = \mathbb{1}$, which by definition means that U is unitary.

Given that transformations of the physical state $\ket{\Psi}$ can be implemented by unitary operators U, i.e., $\ket{\Psi} \mapsto U \ket{\Psi}$, a derivation of the TDSE given in Eq. (1.1) is very simple: Each unitary operator can be expressed as the exponential $\exp(K)$ of a skew-Hermitian operator K. Writing $K = -iHt$ where H is a Hermitian operator, we have $\ket{\Psi} \mapsto \exp(-iHt) \ket{\Psi}$, or equivalently, Eq. (1.1).

The antilinear, antiunitary case

An antilinear operator is defined by the relation

$$U(a \ket{\phi} + b \ket{\psi}) = a^*U(\ket{\phi}) + b^*U(\ket{\psi}) \tag{C.6}$$

for all states $\ket{\phi}$ and $\ket{\psi}$ and all $a, b \in \mathbb{C}$. There are two things to keep in mind about antilinear operators: First, the conventional Dirac notation may cause problems, in the sense that in general $(\langle i| U) \ket{j} \neq \langle i| (U \ket{j})$. This means that the common expression $\langle i|U|j \rangle$ must be read with caution. Second, the *complex conjugation* operator, which is the simplest example of an antilinear operator, is dependent on the basis in terms of which it is defined. In particular, two antilinear operators U and V defined by $U(\sum_j a_j \ket{j}) = \sum_j a_j^* \ket{j}$ and $V(\sum_j a_j \ket{\tilde{j}}) = \sum_j a_j^* \ket{\tilde{j}}$ with a slightly different basis $\ket{\tilde{j}} = i \ket{j}$ are not the same.

Wigner's theorem states that in the antilinear case, the operator U also needs to be antiunitary. This means that, in terms of the notation used in Eq. (C.1), the operator U has to satisfy $\langle \phi'|\psi' \rangle = \langle \phi|\psi \rangle^*$.

In general, antilinear operators are far less common in quantum theory than linear operators [Bal1998]. One reason for this is that we often work with transformations that depend on a continuous parameter t and shall satisfy the group relation $U(t_1)U(t_2) = U(t_1 + t_2)$. Since the product of two antilinear operators is always linear, such a transformation cannot be implemented by an antilinear operator.

Additional literature

In the past decades, many alternative proofs of Wigner's theorem have been published (see e.g. [Győ2004; Geh2014; Bar2017]). They all explore different routes for the proof in the general case of an infinite-dimensional Hilbert space $\mathcal{H}$, where the mathematical apparatus of functional analysis is utilized [Neu1955; Gus2003]. Some of the proofs require *separability* of $\mathcal{H}$ [Sim2008] (i.e., the basis $\{|i\rangle\}$ is still countable such that $i \in \mathbb{N}$) or differentiability of the transformation [Mou2013]. Gehér's more recent proof can deal with non-separable Hilbert spaces and non-bijective transformations [Geh2014]. In addition to this, generalizations of the proof to several other algebraic structures have been considered [Mol1998; Geh2017]. The diversity of mathematical arguments to prove Wigner's theorem is interesting. However, all these proofs start from the conservation of the absolute value of inner products expressed in Eq. (C.1), meaning that the observable transition probabilities shall be conserved. They provide no conceptual alternative to approach the question where unitary transformations in quantum theory come from. Therefore, we now look at some approaches that do provide conceptual alternatives.

C.2 Alternative approaches

In this section, we list some alternative approaches to the question about the use of linearity and unitarity in quantum theory. Each of these provides separate insights into the topic.

The physically compelling approach

An approach with a conceptual alternative to Wigner's theorem was explored by Jordan [Jor1962; Jor2009]. His goal was to derive Wigner's original condition in Eq. (C.1) and the density matrix formulation in Eq. (C.3) from a physically more compelling starting point. One such starting point is the condition that the state of the system does not depend on anything outside the system, but still allows for a description as part of a larger system. Jordan shows that this condition implies that transformations between quantum states must be linear [Jor2006]. Once this is established, one can derive Wigner's condition given by Eq. (C.1) [Jor1962]. The same arguments given above then yield the unitarity of the transformations and, correspondingly, the TDSE given by Eq. (1.1).

The necessity approach

A deeper insight into the use of unitary operators was presented by Landé in 1969 [Lan1969], who explored the question "Why do the probabilities interfere by way of a matrix product law for the probability amplitudes?" He shows that transition probabilities $P(A \rightarrow B)$ and $P(B \rightarrow C)$ can be related, in a triangular form, to a transition probability $P(A \rightarrow C)$ only by means of unitary matrices. There is no mysterious, fundamental requirement for unitarity. Rather, the use of unitary operators in quantum theory is a necessity due to the physicist's ambition to express such a triangular relation in a convenient mathematical form.

The consistency approach

An interesting, fundamentally different approach has been described by Caticha [Cat1998]. He demonstrates that the *only consistent way* to manipulate probability amplitudes is by means of linear, unitary transformations. The idea goes back to Cox's work on the consistent use of probabilities and their connection to logic [Cox1946; Cox1961]. The fundamental character of probability theory as extended logic has been comprehensively presented by Jaynes [Jay2003].

The information-based approach

A more recent approach advocated by Fuchs is based on the concept of *information.* [Fuc2001; Fuc2002]. The premise is that a quantum state $|\Psi\rangle$ does not represent an objective entity that exists in Nature. Rather, $|\Psi\rangle$ and the probability amplitudes that it contains are only the concise representation of our subjective information about Nature. The subjectivity, in particular, implies that quantum theory does not need interpretations [Fuc2000], and also that the time evolution of a state vector $|\Psi\rangle$ does not represent the real time evolution of a physical system, but rather the evolution of our personal state of knowledge about the system [Fuc2000]. This point of view is inspired by the Bayesian interpretation of probabilities (see also [Jay2003]). As soon as probability amplitudes are seen as representing information, their consistent evolution under linear and unitary transformations follows [Har2001; Sch2003]. In fact, in a simplified context, some argue that the difference between probability theory and quantum theory is that, for the former, the 1-norm of vectors is conserved (so they are transformed by stochastic matrices), while for the latter, the 2-norm is conserved, so they are transformed by unitary matrices [Aar2013].

The logical inference approach

A very conclusive approach is the logical inference (LI) approach [DeR2014; DeR2016]. It starts from the actual data that is obtained in experiments, i.e., individually observed events. LI then tries to infer the *most robust description* of this data, given that the experiments are reproducible. This approach first yields a nonlinear optimization problem for the description of the data. By reformulating this optimization problem, one can then derive the equations of quantum theory, such as the TDSE given by Eq. (1.1) [DeR2014], the Klein-Gordon equation [Don2016], the Pauli equation [DeR2015b], or other well-known equations for quantum mechanical key experiments [DeR2015a]. In the LI approach, the linear and unitary character of the equations follows from a reformulation of the optimization problem. Similar to Caticha's approach mentioned above, the ideas behind LI are based on inductive reasoning in the presence of uncertainty [Cox1946; Cox1961; Tri1969; Jay2003].

C.3 General remarks

It is worth mentioning that also nonlinear and non-unitary transformations are sometimes used in the context of quantum theory. Nonlinear expressions for the state vector $|\Psi\rangle$ are used as an effective tool to address complex systems. For instance, in density functional theory, a linear multi-particle Schrödinger equation is effectively replaced by a nonlinear single-particle Schrödinger equation [Eng2011]. A simple example for a non-unitary map is the evolution of a subsystem, which is typically described by a quantum operation (see Section 2.4). Also, the measurement process in quantum theory is typically described in terms of projections which are non-unitary maps (see, for instance, the description of POVMs in the context of Eq. (6.23)).

Regarding the mathematical arguments presented above, one should keep in mind that there can never be a way to formally prove the need of linear and unitary transformations in descriptions of Nature. The only thing that we can prove is that within a certain theory such as quantum theory, under certain assumptions, the mathematical objects of the theory evolve under linear and unitary transformations. But in principle, any experiment might inspire a new theory capable of better describing the observed data without requiring the concept of linear and unitary operators.

Appendix D

Implementations of the four-component transformations V and $V^\dagger$

This appendix includes `C++` code examples for each of the three alternative implementations of the four-component transformations discussed in Section 3.3.1 and benchmarked in Section 4.1.2. They constitute the V and $V^\dagger$ operations in the second-order Suzuki-Trotter product-formula Eq. (4.5). V and $V^\dagger$ consist of a tensor product of 4×4 matrices (see Eq. (3.31)), which require four-component updates of the form of Eq. (3.34) for all $0 \le i < N_{\text{Tr}}$ and $0 \le r < N_{\text{Res}}$. The central loops of the three different implementations for the transformation V are given in Listings D.1–D.3. A description of the variables and function names used in the code listings is given in Tab. D.1.

Listing D.1 Implementation 0: Complete single loop with branches

```
1   uint64_t inc = 1;   // inc = 0b 00..00 01 00..00
2   for(int i = NTr-1; i >= 0; --i, inc <<= 2) {
3       uint64_t mask = 0b11 * inc;   // mask = 0b 00..00 11 00..00
4       #pragma omp for
5       for(uint64_t KM = 0; KM < dim; ++KM)
6           if((KM & mask) == 0)
7               mul4x4(&Vn[16*i], &psi[KM], inc);
8   }
9   for(int r = NRes-1; r >= 0; --r, inc <<= 2) {
10      uint64_t mask = 0b11 * inc;   // mask = 0b 00..00 11 00..00
11      #pragma omp for
12      for(uint64_t KM = 0; KM < dim; ++KM)
13          if((KM & mask) == 0)
14              mul4x4(&Va[16*r], &psi[KM], inc);
15  }
```

Listing D.2 Implementation 1: Reduced single loop with bitwise operations

```
1   uint64_t inc = 1;   // inc = 0b 00..00 01 00..00
2   for(int i = NTr-1; i >= 0; --i, inc <<= 2) {
3       uint64_t rmask = inc - 1;            // rmask = 0b 00..00 00 11..11
4       uint64_t lmask = dim - 1 - rmask;    // lmask = 0b 11..11 11 00..00
5       #pragma omp for
6       for(uint64_t KMred = 0; KMred < dim/4; ++KMred) {
7           // KM = 0b **..** 00 **..**
8           uint64_t KM = ((KMred & lmask) << 2) | (KMred & rmask);
9           mul4x4(&Vn[16*i], &psi[KM], inc);
10      }
11  }
12  for(int r = NRes-1; r >= 0; --r, inc <<= 2) {
13      uint64_t rmask = inc - 1;            // rmask = 0b 00..00 00 11..11
14      uint64_t lmask = dim - 1 - rmask;    // lmask = 0b 11..11 11 00..00
15      #pragma omp for
16      for(uint64_t KMred = 0; KMred < dim/4; ++KMred) {
17          // KM = 0b **..** 00 **..**
18          uint64_t KM = ((KMred & lmask) << 2) | (KMred & rmask);
19          mul4x4(&Va[16*r], &psi[KM], inc);
20      }
21  }
```

Listing D.3 Implementation 2: Reduced nested loops

```
1   uint64_t inc = 1;   // inc = 0b 00..00 01 00..00
2   for(int i = NTr-1; i >= 0; --i, inc <<= 2) {
3       uint64_t incl = inc << 2;  // incl = 0b 00..01 00 00..00
4       #pragma omp for collapse(2)
5       for(uint64_t K = 0; K < dim; K += incl)  // K = 0b **..** 00 00..00
6           for(uint64_t M = 0; M < inc; ++M)     // M = 0b 00..00 00 **..**
7               mul4x4(&Vn[16*i], &psi[K | M], inc);
8   }
9   for(int r = NRes-1; r >= 0; --r, inc <<= 2) {
10      uint64_t incl = inc << 2; // incl = 0b 00..01 00 00..00
11      #pragma omp for collapse(2)
12      for(uint64_t K = 0; K < dim; K += incl)  // K = 0b **..** 00 00..00
13          for(uint64_t M = 0; M < inc; ++M)     // M = 0b 00..00 00 **..**
14              mul4x4(&Va[16*r], &psi[K | M], inc);
15  }
```

Table D.1: Description of the identifiers used in Listings D.1–D.3.

Identifier	Description
NTr	Number of transmons N_{Tr}
NRes	Number of resonators N_{Res}
dim	Dimension of the Hilbert space $(\dim(\mathcal{H}) = 4^{N_{\mathrm{Tr}}+N_{\mathrm{Res}}}$, see Eq. (3.17))
psi	Coefficients $\psi_{\mathtt{KM}}(t)$ of the state vector $\Psi(t)$ (see Eq. (3.25))
Vn	Matrix $V_i^{(n)}$ including the eigenstates of the charge operator in the transmon basis (see Eq. (3.32b))
Va	Matrix $V_r^{(a)}$ including the eigenstates of the operator $\hat{a}_r + \hat{a}_r^\dagger$ in the Fock basis (see Eq. (3.32a))
inc	Integer indicating the current position of the 4×4 transformation given in Eq. (3.34). It is left-shifted by two bits after each iteration of the outer loops.
i	Transmon index of the current 4×4 transformation, corresponding to the two bits indicated by inc
r	Resonator index of the current 4×4 transformation, corresponding to the two bits indicated by inc
K, M, KM	(Parts of) the index KM given by Eq. (3.24)
KMred	Reduced index consisting of KM without the two bits indicated by inc
mask	Bit mask with the two bits indicated by inc set
rmask	Bit mask for the bits on the right of the position indicated by inc
lmask	Bit mask for the bits on the left of the position indicated by inc (including the two bits in the middle)
incl	Increment for the K part of the index KM given by Eq. (3.24)
mul4x4	A vectorized complex 4×4 transformation implementing Eq. (3.34). The inner loops including this operation are independent and can be parallelized. This has been indicated in the code listings using a simple OpenMP directive.

Appendix E

Error bounds for observables

The accuracy of the product-formula algorithm (see Section 3.3.1) can be controlled by means of rigorous error bounds for the solution of the TDSE [DeR1987; Huy1990] (see Eqs. (4.2a) and (4.2b)). However, the bounds apply to the full state vector, so they may be impractical if we are only interested in the expectation value of a certain observable.

Therefore, we tested two general error bounds for expectation values of observables in Section 4.1.1 (see Eqs. (4.3a) and (4.3b)). The bounds are given by

$$|\langle\psi|A|\psi\rangle - \langle\phi|A|\phi\rangle| \leq 2\sqrt{\Delta}\,\|A\|_2, \tag{E.1a}$$

$$|\langle\psi|A|\psi\rangle - \langle\phi|A|\phi\rangle| \leq 2\sqrt{\Delta}\sqrt{\mathrm{Var}_\psi(A)}\,|\langle\psi|\phi\rangle| + 2\Delta\,\|A\|_2, \tag{E.1b}$$

where A is an observable (i.e., a Hermitian operator), $|\psi\rangle$ and $|\phi\rangle$ are pure states, $\Delta = 1 - |\langle\psi|\phi\rangle|^2$ is the distinguishability between $|\psi\rangle$ and $|\phi\rangle$, $\|A\|_2$ denotes the spectral norm (largest singular value) of A, and $\mathrm{Var}_\psi(A) = \langle A^2\rangle - \langle A\rangle^2$ is the variance of A with respect to the state $|\psi\rangle$.

The second bound was shown to be tight (see Fig. 4.2(a)), and a general proof for its validity is given in [Wil2020d]. For the first bound, we give a short and elementary proof in this appendix.

First, note that for any Hermitian operator B with eigenvectors $|b\rangle$, eigenvalues b, and singular values $|b|$, we have

$$|\mathrm{Tr}\,BA| = \left|\sum_b b\underbrace{\langle b|A|b\rangle}_{\leq\|A\|_2}\right| \leq \left(\sum_b|b|\right)\|A\|_2 = \|B\|_{\mathrm{Tr}}\|A\|_2, \tag{E.2}$$

where $\|B\|_{\mathrm{Tr}} = \sum_b|b|$ is the trace norm of B. We remark that Eq. (E.2) is a special case of Hölder's inequality for Schatten norms, which states that $|\mathrm{Tr}\,X^\dagger Y| \leq \|\sigma(X)\|_p\|\sigma(Y)\|_{p*}$, where X and Y are operators, $\sigma(X)$ and $\sigma(Y)$ denote vectors of their respective singular values, $\|v\|_p = (\sum_i|v_i|^p)^{1/p}$ is the p-norm of a vector v, and $p, p^* \in [1,\infty]$ are chosen so that $1/p + 1/p^* = 1$ [Wat2018]. In this notation, $p = 1$ and $p^* = \infty$ correspond to the trace norm and the spectral norm, respectively.

Applying this result to Eq. (E.1a), we have

$$|\langle\psi|A|\psi\rangle - \langle\phi|A|\phi\rangle| = |\mathrm{Tr}[(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|)A]| \leq \|(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|)\|_{\mathrm{Tr}}\|A\|_2. \tag{E.3}$$

To find the trace norm of $|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$, we make use of Eq. (6.9) for $\alpha = \beta = 1$ (the proof is given in Section 6.1.2):

$$\|(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|)\|_{\mathrm{Tr}} = 2\sqrt{1 - |\langle\psi|\phi\rangle|^2}. \tag{E.4}$$

Recognizing the distinguishability $\Delta = 1 - |\langle\psi|\phi\rangle|^2$, we obtain

$$|\langle\psi|A|\psi\rangle - \langle\phi|A|\phi\rangle| \leq 2\sqrt{\Delta}\,\|A\|_2, \tag{E.5}$$

which proves Eq. (E.1a).

Appendix F

Pulse parameters for quantum gates

For reference, we list all pulse parameters used for the various multi-transmon systems under investigation (see Section 3.4). The pulse parameters are the result of the pulse optimization procedure discussed in Chapter 5 (see Sections 5.1 and 5.2 for the particular meaning of the single-qubit and two-qubit pulse parameters). The performance of the individual gates listed in this appendix is discussed in detail in Chapter 6.

Table F.1: Parameters of the optimized single-qubit GD pulses defined in Eq. (5.9) for the two-transmon system (see Section 3.4.3). For the compilation process (cf. Section 5.4), these parameters are specified in the form of Listing 5.1.

Pulse name	f [GHz]	T_X [ns]	Ω_X	β_X [ns]
xpih-0	5.3463	83	0.002221	0.2309
xpih-1	5.1167	83	0.002269	0.2891
xpi-0	5.3463	83	0.004444	0.2193
xpi-1	5.1167	83	0.004538	0.2239

Table F.2: Parameters of the optimized two-qubit pulses defined in Section 5.2 for the two-transmon system (see Section 3.4.3). We tested three kinds of CR pulses (cf. Fig. 5.1): CR1 (see the text below Eq. (5.17)), CR2 (see the text below Eq. (5.19)), and CR4. The pulse name indicates the control qubit i_C and the target qubit i_T in the form cnot-i_C-i_T. The duration of all single-qubit GD pulses included in these schemes is $(T_X^\pi)_C = (T_X^{\pi/2})_T = 83$ ns. For the compilation process (cf. Section 5.4), these parameters are specified in the form of Listing 5.1.

CR1 Pulse name	f_{i_T} [GHz]	T_{CR} [ns]	Ω_{CR}	Ω_{Cancel}	ϕ_{CR}	ϕ_{Cancel}	ϕ_C	ϕ_T
cnot-0-1	5.1166	41.865	0.0793	0.00618	0.54	0.00	-2.10	0.04
cnot-1-0	5.3464	128.193	0.0940	-0.00162	-2.89	1.72	3.25	1.40

CR2 Pulse name	f_{i_C} [GHz]	f_{i_T} [GHz]	T_{CR} [ns]	Ω_{CR}	$(\Omega_X^\pi)_C$	$(\beta_X^\pi)_C$ [ns]	$(\Omega_X^{\pi/2})_T$	$(\beta_X^{\pi/2})_T$ [ns]
cnot-0-1	5.3463	5.1167	102.9746	0.01111	0.004444	0.2193	0.002269	0.2891
cnot-1-0	5.1167	5.3463	71.5580	0.07058	0.004538	0.2239	0.002221	0.2309

CR4 Pulse name	f_{i_C} [GHz]	f_{i_T} [GHz]	T_{CR} [ns]	Ω_{CR}	GD pulse parameters from Tab. F.1		
cnot-0-1	5.3463	5.1167	50.2385	0.01018			
cnot-1-0	5.1167	5.3463	30.1557	0.06934	$(\Omega_X^\pi, \beta_X^\pi)_C$ and $(\Omega_X^{\pi/2}, \beta_X^{\pi/2}, \Omega_X^\pi, \beta_X^\pi)_T$		

Table F.3: Parameters of the single-qubit GD pulses defined in Eq. (5.9) for the small five-transmon system (see Section 3.4.4). The parameters have been obtained from the pulse optimization procedure described in Section 5.3. The label `withf` indicates pulses with frequency tuning such that the tuned frequency f may differ slightly from the qubit frequency (cf. Tab. 3.5).

Pulse name	f [GHz]	T_X [ns]	Ω_X	β_X [ns]
xpih-0-withf	5.34697	80	0.00231	0.246
xpih-1-withf	5.30232	80	0.00232	0.220
xpih-2-withf	5.11345	80	0.00236	0.218
xpih-3-withf	5.22506	80	0.00233	0.232
xpih-4-withf	5.07065	80	0.00237	0.228

Table F.4: Parameters of the two-qubit CR2 pulses defined in Section 5.2 (see the text below Eq. (5.19)) for the small five-transmon system (see Section 3.4.4). The parameters have been obtained from the pulse optimization procedure described in Section 5.3. The pulse name indicates the control qubit i_C and the target qubit i_T in the form `cnot-`i_C`-`i_T. The label `withf` indicates pulses with frequency tuning such that the tuned frequencies f_{i_C} and f_{i_T} may differ slightly from the qubit frequencies (cf. Tab. 3.5). The duration of all single-qubit GD pulses included in the CR2 pulse is always $(T_X^\pi)_C = (T_X^{\pi/2})_T = 80\,\mathrm{ns}$.

Pulse name	f_{i_C} [GHz]	f_{i_T} [GHz]	T_{CR} [ns]	Ω_{CR}	$(\Omega_X^\pi)_C$	$(\beta_X^\pi)_C$ [ns]	$(\Omega_X^{\pi/2})_T$	$(\beta_X^{\pi/2})_T$ [ns]
cnot-0-2-withf	5.34697	5.11345	100.341	0.0113	0.00463	0.250	0.00236	0.218
cnot-1-2-withf	5.30232	5.11345	121.308	0.0103	0.00465	0.230	0.00236	0.218
cnot-3-2-withf	5.22506	5.11345	88.442	0.0114	0.00469	0.240	0.00236	0.218
cnot-4-2-withf	5.07065	5.11345	48.632	0.0114	0.00502	0.223	0.00236	0.218

Table F.5: Parameters of the single-qubit GD pulses defined in Eq. (5.9) for the large five-transmon system (see Section 3.4.5). The parameters have been obtained from the pulse optimization procedure described in Section 5.3. The corresponding optimization process is visualized in Fig. 5.6. The label `withf` indicates that the drive frequency f has also been optimized such that it may differ slightly from the qubit frequency (cf. Tab. 3.7).

Pulse name	f [GHz]	T_X [ns]	Ω_X	β_X [ns]
xpih-0	4.97154	80	0.00238	1.335
xpih-1	5.07063	80	0.00236	-1.904
xpih-2	5.26657	80	0.00233	-2.165
xpih-3	5.10145	80	0.00236	0.498
xpih-4	4.86036	80	0.00241	2.276
xpih-0-withf	4.97164	80	0.00239	0.239
xpih-1-withf	5.07043	80	0.00236	0.238
xpih-2-withf	5.26634	80	0.00233	0.229
xpih-3-withf	5.10147	80	0.00236	0.232
xpih-4-withf	4.86055	80	0.00241	0.236

Table F.6: Parameters of the two-qubit CR2 pulses defined in Section 5.2 (see the text below Eq. (5.19)) for the large five-transmon system (see Section 3.4.5). The parameters have been obtained from the pulse optimization procedure described in Section 5.3. The corresponding optimization process is visualized in Fig. 5.7. The pulse name indicates the control qubit i_C and the target qubit i_T in the form `cnot-`i_C`-`i_T. The label `withf` indicates pulses with frequency tuning such that the tuned frequencies f_{i_C} and f_{i_T} may differ slightly from the qubit frequencies (cf. Tab. 3.7). The duration of all single-qubit GD pulses included in the CR2 pulse is always $(T_X^\pi)_C = (T_X^{\pi/2})_T = 80$ ns.

Pulse name	f_{i_C} [GHz]	f_{i_T} [GHz]	T_{CR} [ns]	Ω_{CR}	$(\Omega_X^\pi)_C$	$(\beta_X^\pi)_C$ [ns]	$(\Omega_X^{\pi/2})_T$	$(\beta_X^{\pi/2})_T$ [ns]
cnot-1-0	5.07063	4.97154	76.955	0.0097	0.00461	0.640	0.00238	1.335
cnot-1-4	5.07063	4.86036	64.161	0.0183	0.00476	-0.148	0.00241	2.276
cnot-2-1	5.26657	5.07063	33.398	0.0235	0.00465	-0.036	0.00236	-1.904
cnot-3-2	5.10145	5.26657	242.064	0.0111	0.00471	0.508	0.00233	-2.165
cnot-3-4	5.10145	4.86036	33.247	0.0290	0.00465	0.640	0.00241	2.276
cnot-4-0	4.86036	4.97154	105.151	0.0210	0.00449	-1.511	0.00238	1.335
cnot-1-0-withf	5.07043	4.97164	73.538	0.0101	0.00477	0.798	0.00239	0.239
cnot-1-4-withf	5.07043	4.86055	109.439	0.0114	0.00472	0.502	0.00241	0.236
cnot-2-1-withf	5.26634	5.07043	82.077	0.0111	0.00463	0.661	0.00236	0.238
cnot-3-2-withf	5.10147	5.26634	58.763	0.0429	0.00480	-0.198	0.00233	0.229
cnot-3-4-withf	5.10147	4.86055	85.294	0.0118	0.00474	0.247	0.00241	0.236
cnot-4-0-withf	4.86055	4.97164	98.599	0.0239	0.00483	0.115	0.00239	0.239

Appendix G

Average fidelity of trace-decreasing quantum operations

In this appendix, we provide two separate proofs for a generalized version of an explicit relation between the average fidelity F_{avg} (defined in terms of an integral over random states) and the entanglement fidelity F_{ent} (typically accessible in closed-form),

$$F_{\text{avg}}(\mathcal{E}) = \frac{dF_{\text{ent}}(\mathcal{E}) + \text{Tr}\,\mathcal{E}(\mathbb{1}/d)}{d+1}. \tag{G.1}$$

Here, d is the dimension of the Hilbert space (typically 2^n for an n-qubit system), and $\mathcal{E}(\rho) = \sum_\alpha E_\alpha \rho E_\alpha^\dagger$ is a completely positive map which is not necessarily trace-preserving, i.e., $\sum_\alpha E_\alpha^\dagger E_\alpha \leq \mathbb{1}$ (cf. Section 2.4). In the special case that $\mathcal{E}$ is trace-preserving, we have $\text{Tr}\,\mathcal{E}(\mathbb{1}/d) = 1$ such that Eq. (G.1) reduces to the well-known expression given in [Hor1999].

After stating some preliminary definitions to settle the notation, we give both an algebraic proof using methods from quantum information theory and an elementary, analytic proof by direct calculation.

G.1 Preliminaries

The *fidelity* between two quantum states ρ and σ is defined as [Joz1994] (see also [Fuc1995])

$$F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_{\text{Tr}}^2 = \left(\text{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2, \tag{G.2}$$

where we used the definition of the trace norm $\|X\|_{\text{Tr}} = \text{Tr}\sqrt{X^\dagger X}$ (sum of the singular values) of X. If one of the states is pure, e.g. $\rho = |\psi\rangle\langle\psi|$, the fidelity simplifies to the overlap $\langle\psi|\sigma|\psi\rangle$.

The *average fidelity* of a quantum operation $\mathcal{E}$ is defined by averaging the fidelity $F(|\psi\rangle\langle\psi|, \mathcal{E}(|\psi\rangle\langle\psi|))$ over random pure states $|\psi\rangle$,

$$F_{\text{avg}}(\mathcal{E}) = \int \text{d}|\psi\rangle\, \langle\psi|\mathcal{E}(\psi)|\psi\rangle. \tag{G.3}$$

The integral is taken over pure states $|\psi\rangle$ whose $2d$ real coefficients are distributed uniformly on the surface of a $2d$-dimensional unit sphere. For simplicity, we use the notation that $\psi = |\psi\rangle\langle\psi|$ if the meaning is clear from the context.

The *entanglement fidelity* of $\mathcal{E}$ is defined as the fidelity $F(\Phi, (\mathcal{E} \otimes \mathbb{1})(\Phi))$, where $|\Phi\rangle = \sum_j |jj\rangle/\sqrt{d}$ is the maximally entangled state on an extended Hilbert space that is also of dimension d. We have

$$F_{\text{ent}}(\mathcal{E}) = \langle\Phi|(\mathcal{E} \otimes \mathbb{1})(\Phi)|\Phi\rangle = \sum_\alpha \frac{|\text{Tr}\,E_\alpha|^2}{d^2}. \tag{G.4}$$

G.2 Quantum information theoretic proof

The following proof is based on the algebraic proof given in [Nie2002], generalized to non-trace-preserving quantum operations and thus extending the work in [Gil2005].

We consider the so-called *twirled* quantum operation

$$\mathcal{E}_T(\rho) = \int \mathrm{d}U \, U^\dagger \mathcal{E}(U\rho U^\dagger)U, \tag{G.5}$$

where the integral is over the Haar measure on the group of unitary matrices [Spe2012]. This operation leaves both the average fidelity given by Eq. (G.3) and the entanglement fidelity given by Eq. (G.4) invariant, which can be shown by substitution [Nie2002]. Therefore, we have

$$F_{\mathrm{avg}}(\mathcal{E}) = F_{\mathrm{avg}}(\mathcal{E}_T), \tag{G.6a}$$

$$F_{\mathrm{ent}}(\mathcal{E}) = F_{\mathrm{ent}}(\mathcal{E}_T). \tag{G.6b}$$

Furthermore, for any ρ and any unitary operator V, we have

$$V\mathcal{E}_T(\rho)V^\dagger = \int \mathrm{d}U \, VU^\dagger \mathcal{E}(U\rho U^\dagger)UV^\dagger = \mathcal{E}_T(V\rho V^\dagger), \tag{G.7}$$

which can also be shown by substitution, i.e., $W = UV^\dagger$. Let $\rho = P$ be a rank-1 projector and $Q = \mathbb{1} - P$ its orthogonal complement. We define the space S_P (S_Q) as the space onto which P (Q) projects. For any block-diagonal unitary matrix $V = V_P + V_Q$, where V_P (V_Q) is only non-zero on S_P (S_Q), we have $VPV^\dagger = V_P V_P^\dagger = \mathbb{1}_{S_P} = P$ and thus $V\mathcal{E}_T(P)V^\dagger = \mathcal{E}_T(P)$. Since this holds for any block-diagonal unitary matrix of the form $V = V_P + V_Q$, $\mathcal{E}_T(P)$ must also be block-diagonal and each block must be proportional to the identity. The identity on S_P is $\mathbb{1}_{S_P} = P$, and the identity on S_Q is $\mathbb{1}_{S_Q} = Q = \mathbb{1} - P$, so we have

$$\mathcal{E}_T(P) = \alpha P + \beta(\mathbb{1} - P) \tag{G.8}$$

for some α and β. Using Eq. (G.7) again for an arbitrary unitary operator V transforms P in this equation into any other rank-1 projector $P' = VPV^\dagger$, so we see that α and β are the same for each P.

Writing an arbitrary $\rho = \sum_i \rho_i P_i$ as a sum of rank-1 projectors P_i with unit trace, linearity of $\mathcal{E}_T$ yields that for any ρ, $\mathcal{E}_T(\rho) = \alpha\rho + \beta(\mathbb{1} - \rho)$. By replacing $\beta = p/d$ and $\alpha = t - p + p/d$, we obtain

$$\mathcal{E}_T(\rho) = (t - p)\rho + p\mathbb{1}/d, \tag{G.9}$$

where also the parameters p and t are independent of ρ.

The difference to the trace-preserving case considered in [Nie2002] is now that in general, $\mathrm{Tr}\,\mathcal{E}_T(\rho) = t \neq 1$. Consequently, Eq. (G.9) does not represent a depolarizing quantum operation anymore. Note that care must be taken with the identity symbol $\mathbb{1}$ in Eq. (G.9) since it only applies when the map $\mathcal{E}_T$ is restricted to density matrices with $\mathrm{Tr}\,\rho = 1$. If

trace-decreasing quantum operations are applied to operators with trace less than 1 (e.g. if they are chained, $\mathcal{E}_T(\mathcal{E}_T(\rho))$), the correct expression is $\mathcal{E}_T(\rho) = (t - p)\rho + p\mathbb{1}\operatorname{Tr}\rho/d$. To determine the trace parameter t, we evaluate Eqs. (G.5) and (G.9) for $\rho = \mathbb{1}/d$:

$$t = \operatorname{Tr}\mathcal{E}_T(\mathbb{1}/d) = \operatorname{Tr}\int \mathrm{d}U\, U^\dagger\mathcal{E}(UU^\dagger)U/d = \operatorname{Tr}\mathcal{E}(\mathbb{1}/d). \qquad (\text{G.10})$$

Using the simple form of Eq. (G.9), we can evaluate the fidelities in Eqs. (G.6a) and (G.6b) directly. For the average fidelity, we find

$$F_{\mathrm{avg}}(\mathcal{E}) = \int \mathrm{d}|\psi\rangle\,\langle\psi|\mathcal{E}(\psi)|\psi\rangle = \operatorname{Tr}\mathcal{E}(\mathbb{1}/d) - p + \frac{p}{d}, \qquad (\text{G.11})$$

and the entanglement fidelity becomes

$$F_{\mathrm{ent}}(\mathcal{E}) = \langle\Phi|(\mathcal{E}_T \otimes \mathbb{1})(\Phi)|\Phi\rangle = \operatorname{Tr}\mathcal{E}(\mathbb{1}/d) - p + \frac{p}{d^2}. \qquad (\text{G.12})$$

Solving Eq. (G.12) for p and inserting the result into Eq. (G.11) yields the desired relation given by Eq. (G.1).

G.3 Analytic proof

In this version of the proof, we directly evaluate the integral for the average gate fidelity given by Eq. (G.3). First, by expanding the pure state $|\psi\rangle = \sum_i c_i\,|i\rangle$ with $c_i \in \mathbb{C}$, we obtain

$$F_{\mathrm{avg}}(\mathcal{E}) = \sum_\alpha\sum_{ijkl}\langle i|E_\alpha|j\rangle\,\langle k|E_\alpha^\dagger|l\rangle\int\mathrm{d}|\psi\rangle\,c_i^*c_jc_k^*c_l. \qquad (\text{G.13})$$

The integral at the end of this expression can be computed in the following way:

$$\int\mathrm{d}|\psi\rangle\,c_i^*c_jc_k^*c_l = \frac{\int \mathrm{d}a_1\mathrm{d}b_1\cdots\mathrm{d}a_d\mathrm{d}b_d\,\delta(\sum_j(a_j^2 + b_j^2) - 1)c_i^*c_jc_k^*c_l}{\int \mathrm{d}a_1\mathrm{d}b_1\cdots\mathrm{d}a_d\mathrm{d}b_d\,\delta(\sum_j(a_j^2 + b_j^2) - 1)}, \qquad (\text{G.14})$$

where we used the fact that the space of pure states of dimension d is characterized by $c_j = a_j + ib_j$ for $a_j, b_j \in \mathbb{R}$ with $\sum_j(a_j^2 + b_j^2) = 1$. The first thing to note is that the integral is non-zero only if $i = j$ and $k = l$ or $i = l$ and $j = k$, since otherwise the integrand in the numerator is an odd function integrated over a symmetric interval. Hence, the required integrals are $\int\mathrm{d}|\psi\rangle\,|c_i|^2|c_j|^2$ and $\int\mathrm{d}|\psi\rangle\,|c_i|^4$. These integrals have been evaluated in [Ham2000] and [Jin2020]. We do not repeat the full calculation here but, for the sake of reference, we outline three common strategies used to compute such integrals.

One way to evaluate Eq. (G.14) is to use spherical coordinates, i.e., polar coordinates for (a_i, b_i) and (a_j, b_j), and hyperspherical coordinates for the remaining coefficients. Since the integrands only depend on the radii in these coordinates, they reduce to single integrals over these radii multiplied by the surface of the respective spheres. See [Ham2000; Jin2020] for more information.

Another way to compute the integrals is to make use of the representation of the δ-function $\delta(x) = \int dt\, e^{itx}/2\pi$ and closing the contour of integration in the complex plane with $\text{Im}\, t = \varepsilon > 0$. See the supplementary material of [Boi2018] for an example using this approach.

A third option to obtain the result is particularly convenient for numerical work: The coefficients of a random pure state $|\psi\rangle$ can be generated from the normal distribution [Ben2006]

$$p(a_1, b_1, \ldots, a_d, b_d) = \frac{1}{(2\pi)^d} e^{-(a_1^2 + b_1^2 + \cdots + a_d^2 + b_d^2)/2}. \tag{G.15}$$

This is a construction that has a long history (see Muller's method in [Mul1959]) and is commonly used in random matrix theory [Ede2005]. With this strategy, we set $c_j = (a_j + ib_j)/\sum_j (a_j^2 + b_j^2)$ for $a_j, b_j \in \mathbb{R}$, and the evaluation of the integral amounts to

$$\int d|\psi\rangle\, f(c_1, \ldots, c_d) = \int da_1 db_1 \cdots da_d db_d\, p(a_1, b_1, \ldots, a_d, b_d) f(c_1, \ldots, c_d). \tag{G.16}$$

As shown in [Jin2020], this expression basically reduces to a set of Gaussian integrals.

Independent of the strategies used, one obtains

$$\int d|\psi\rangle\, c_i^* c_j c_k^* c_l = \frac{1}{d(d+1)} (\delta_{ij}\delta_{kl} + \delta_{il}\delta_{jk}). \tag{G.17}$$

Inserting this expression into Eq. (G.13), we immediately find

$$F_{\text{avg}}(\mathcal{E}) = \sum_\alpha \frac{|\text{Tr}\, E_\alpha|^2 + \text{Tr}\, E_\alpha^\dagger E_\alpha}{d(d+1)}, \tag{G.18}$$

which is equivalent to the desired relation given by Eq. (G.1) after substituting $F_{\text{ent}}(\mathcal{E}) = \sum_\alpha |\text{Tr}\, E_\alpha|^2/d^2$ (see Eq. (G.4)) and $\text{Tr}\,\mathcal{E}(\mathbb{1}) = \sum_\alpha \text{Tr}\, E_\alpha^\dagger E_\alpha$. We remark that the result given by Eq. (G.18) can also be found in [Ped2007].

Appendix H

Diamond distance between unitary quantum operations

In Section 6.1.2, we obtained an expression for the diamond distance $\eta_\diamond$ between two quantum operations $\mathcal{G}_{id}(\rho) = U\rho U^\dagger$ and $\mathcal{G}_{ac}(\rho) = M\rho M^\dagger$, where U is a unitary matrix representing an ideal quantum gate, and M denotes a matrix describing the actual implementation. In this appendix, we evaluate this expression for the case that both M and U are unitary. The result provides an explicit proof for the statements given in [Aha1998] and [Joh2009] and illustrates the construction.

The expression in Eq. (6.11) reads

$$\eta_\diamond = \frac{1}{2}\|M \cdot M^\dagger - U \cdot U^\dagger\|_\diamond = \frac{1}{2}\max_{|x\rangle}\sqrt{(\langle x|W^\dagger W \otimes \mathbb{1}|x\rangle + 1)^2 - 4|\langle x|W \otimes \mathbb{1}|x\rangle|^2}, \quad \text{(H.1)}$$

where $W = MU^\dagger$. If both M and U are unitary, we have $W^\dagger W = \mathbb{1}$ and therefore $\langle x|W^\dagger W \otimes \mathbb{1}|x\rangle = 1$.

The first step is to diagonalize W so that $V^{-1}WV = \Lambda$, where $\Lambda = \text{diag}(\lambda_i)$ denotes the eigenvalues of W. Since W is unitary, it is a normal matrix, so also V is unitary and preserves the norm. Thus, we can substitute $|x\rangle \mapsto V|x\rangle$ in the maximization to obtain

$$\eta_\diamond = \max_{|x\rangle}\sqrt{1 - |\langle x|\Lambda \otimes \mathbb{1}|x\rangle|^2}. \quad \text{(H.2)}$$

Obviously, the maximum is attained when $|\langle x|\Lambda \otimes \mathbb{1}|x\rangle|^2$ is minimal. We now expand $|x\rangle = \sum_{ij} x_{ij}|i\rangle \otimes |j\rangle$ with $i, j = 0, \ldots, N-1$ and $\sum_{ij}|x_{ij}|^2 = 1$ (we can also think of $|i\rangle$ as the eigenbasis of W; but the important thing is that V is unitary since otherwise, in general, $\sum_{ij}|x_{ij}|^2 \neq 1$). Thus we have

$$\eta_\diamond = \sqrt{1 - \min_{\substack{x_{ij} \in \mathbb{R} \\ \sum|x_{ij}|^2=1}}\left|\sum_{ij}|x_{ij}|^2\lambda_i\right|^2} = \sqrt{1 - \min_{\substack{p_i \geq 0 \\ \sum p_i=1}}\left|\sum_i p_i\lambda_i\right|^2}, \quad \text{(H.3)}$$

where $p_i = \sum_j |x_{ij}|^2 \geq 0$ with $\sum_i p_i = 1$. The minimization is now over all convex combinations of the eigenvalues λ_i, i.e., all points in the set

$$\mathcal{C} = \left\{\sum_i p_i\lambda_i \ : \ p_i \geq 0 \text{ and } \sum_i p_i = 1\right\}. \quad \text{(H.4)}$$

By definition, this set is the convex hull of all λ_i, i.e., a polygon whose vertices are given by λ_i. Since W is unitary, all eigenvalues satisfy $|\lambda_i| = 1$, so the vertices λ_i of the polygon lie on the complex unit circle (see Fig. H.1). The quantity $|\sum_i p_i\lambda_i|$ to be minimized represents the distance from the point $\sum_i p_i\lambda_i \in \mathcal{C}$ to the origin.

There are now two possible cases: Either the origin is inside the polygon ($0 \in \mathcal{C}$), or it is outside. If $0 \in \mathcal{C}$ as shown in Fig. H.1(a), there exists a set $\{p_i\}$ such that $\sum_i p_i\lambda_i = 0$, and we have $\eta_\diamond = 1$.

194

(a)

(b)
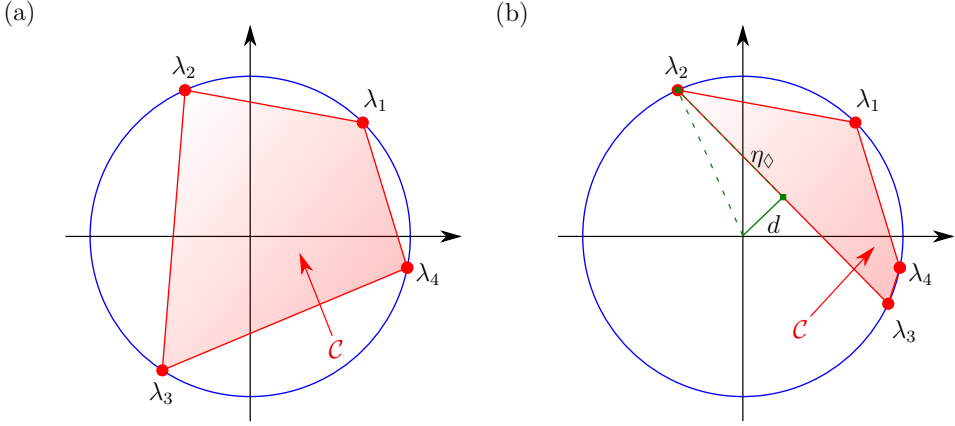
Figure H.1: Illustration of the diamond distance between unitary quantum operations. The red points represent the eigenvalues λ_i. They lie on the unit circle in the complex plane. The polygon is the convex hull $\mathcal{C}$ of these points (see Eq. (H.4)). The two possible cases are: (a) $0 \in \mathcal{C}$, i.e., the origin is inside the polygon such that $\eta_\Diamond = 1$; (b) the origin is outside the polygon such that $\eta_\Diamond = \sqrt{1 - d^2}$, where d is the distance between the origin and the polygon.

In the other case shown in Fig. H.1(b), the point in $\mathcal{C}$ with the closest distance $d = \min|\sum_i p_i \lambda_i|$ to the origin lies in the middle of the line between λ_j and λ_k, for which $|\lambda_j - \lambda_k|$ is maximal. Inserting d into Eq. (H.3), we find that the diamond distance is $\eta_\Diamond = \sqrt{1 - d^2}$, which is the result stated without proof in [Aha1998].

Furthermore, this result can be related to the result given in [Joh2009] by noting that the hypotenuse of the right triangle indicated in Fig. H.1(b) is 1. Hence we have that $\eta_\Diamond = \sqrt{1 - d^2}$ is half the line between λ_j and λ_k, i.e., $\eta_\Diamond = \max_{jk}|\lambda_j - \lambda_k|/2$. This means that $2\eta_\Diamond$ is the diameter of the smallest closed disc enclosing all eigenvalues of W, as stated in [Joh2009].

In summary, we have

$$\eta_\Diamond = \begin{cases} 1 & (0 \in \mathcal{C}) \\ \frac{1}{2} \max_{jk}|\lambda_j - \lambda_k| & \text{(otherwise)} \end{cases}. \tag{H.5}$$

A practical way to check for $0 \in \mathcal{C}$ is to compute all arguments $\varphi_i = \arg(\lambda_i) \in [0, 2\pi)$ and list them in increasing order $0 \leq \varphi_0 \leq \cdots \leq \varphi_{N-1} < 2\pi$. If all N successive differences $\Delta\varphi_i = (\varphi_{i+1} - \varphi_i) \bmod [0, 2\pi)$ (including the case $\varphi_N \equiv \varphi_0$) are in $[0, \pi]$, the convex polygon $\mathcal{C}$ includes the origin.

Note that in the single-qubit case, W has only two eigenvalues such that the convex hull $\mathcal{C}$ is a straight line. In this case, the formula for the second case in Eq. (H.5) is always valid such that $\eta_\Diamond = |\lambda_1 - \lambda_2|/2$.

Appendix I

Proof of a diamond-distance bound for trace-decreasing operations

For trace-preserving quantum operations $\mathcal{E}$, the best known lower bound for the diamond distance in terms of the average fidelity was proven in [Wal2014]. In this appendix, we derive a new bound that also applies to non-trace-preserving quantum operations (cf. Section 2.4). Furthermore, it reduces to the original bound in the trace-preserving case, thereby generalizing the previous result.

The known lower bound for a completely positive trace-preserving map $\mathcal{E}$ reads [Wal2014; Kue2016]

$$\frac{d+1}{d}(1 - F_{\mathrm{avg}}(\mathcal{E})) \leq \eta_\diamond. \tag{I.1}$$

Here, d is the dimension of the Hilbert space, $F_{\mathrm{avg}}(\mathcal{E})$ is the average fidelity defined in Eq. (G.3), and $\eta_\diamond = \|\mathcal{E} - \mathbb{1}\|_\diamond / 2$ is the diamond distance defined in Section 6.1.2. Equation (I.1) is only valid under the assumption that $\mathcal{E}$ is trace-preserving.

In what follows, let $\mathcal{E}$ be a completely positive quantum operation that is not necessarily trace-preserving. We start from the definition of the diamond norm given by Eq. (6.5) for $\mathcal{T} = \mathcal{E} - \mathbb{1}$,

$$\eta_\diamond = \frac{1}{2} \sup_{X \neq 0} \frac{\|((\mathcal{E} - \mathbb{1}) \otimes \mathbb{1})(X)\|_{\mathrm{Tr}}}{\|X\|_{\mathrm{Tr}}}, \tag{I.2}$$

where $\|X\|_{\mathrm{Tr}} = \mathrm{Tr}\,|X|$ with $|X| = \sqrt{X^\dagger X}$ denotes the trace norm (sum of the singular values). Because of the supremum in the definition of $\eta_\diamond$, we have

$$\eta_\diamond \geq \frac{1}{2}\|((\mathcal{E} - \mathbb{1}) \otimes \mathbb{1})(\Phi)\|_{\mathrm{Tr}} = \frac{1}{2}\|J(\mathcal{E}) - \Phi\|_{\mathrm{Tr}}, \tag{I.3}$$

where $\Phi = |\Phi\rangle\langle\Phi|$ with $|\Phi\rangle = \sum_j |jj\rangle / \sqrt{d}$ is the maximally entangled state, and we used the definition of the Choi matrix $J(\mathcal{E}) = (\mathcal{E} \otimes \mathbb{1})(\Phi)$. Note that the order of tensor-product factors in this definition of the Choi matrix is reversed (as compared to Eq. (2.25)) such that it complies with the definition of the diamond norm. The result is independent of the order.

Using the definition of the trace norm given above and inserting $\mathbb{1} = \Phi + \mathbb{1} - \Phi$, we obtain

$$\frac{1}{2}\|J(\mathcal{E}) - \Phi\|_{\mathrm{Tr}} = \frac{1}{2}\Big(\mathrm{Tr}\,\Phi|J(\mathcal{E}) - \Phi| + \mathrm{Tr}\,(\mathbb{1} - \Phi)|J(\mathcal{E}) - \Phi|\Big). \tag{I.4}$$

Since the matrix $J(\mathcal{E}) - \Phi$ is Hermitian, its singular values are the absolute values of its eigenvalues. For any positive semidefinite matrix A and any Hermitian matrix B with singular values $|b|$, eigenvalues b, and eigenvectors $|b\rangle$, we have

$$\mathrm{Tr}\,A|B| = \sum_b |b|\,\mathrm{Tr}\,A\,|b\rangle\langle b| \geq \sum_b b\,\mathrm{Tr}\,A\,|b\rangle\langle b| = \mathrm{Tr}\,AB, \tag{I.5}$$

and similarly, $\operatorname{Tr} A|B| \geq \operatorname{Tr} A(-B)$. Since both Φ and $\mathbb{1} - \Phi$ are projectors, their eigenvalues are 1 and 0, so they are positive semidefinite. Thus, Eq. (I.4) is bounded by

$$\frac{1}{2}\|J(\mathcal{E}) - \Phi\|_{\operatorname{Tr}} \geq \frac{1}{2}\Big(\operatorname{Tr} \Phi(\Phi - J(\mathcal{E})) + \operatorname{Tr}(\mathbb{1} - \Phi)(J(\mathcal{E}) - \Phi)\Big) \qquad \text{(I.6a)}$$

$$= \operatorname{Tr} \Phi(\Phi - J(\mathcal{E})) + \frac{1}{2}\operatorname{Tr}(J(\mathcal{E}) - \Phi) \qquad \text{(I.6b)}$$

$$= 1 - \langle\Phi|J(\mathcal{E})|\Phi\rangle + \frac{1}{2}(\operatorname{Tr} J(\mathcal{E}) - 1). \qquad \text{(I.6c)}$$

Note that for this step, the proof in [Wal2014] made use of the Fuchs-van de Graaf inequality $\|J(\mathcal{E}) - \Phi\|_{\operatorname{Tr}}/2 \geq 1 - \langle\Phi|J(\mathcal{E})|\Phi\rangle$ [Fuc1999], which only works if $\operatorname{Tr} J(\mathcal{E}) = 1$, i.e., if $\mathcal{E}$ is trace-preserving. The step from Eq. (I.4) to Eq. (I.6a), however, is elementary and also works in the non-trace-preserving case.

By identifying the term $\langle\Phi|J(\mathcal{E})|\Phi\rangle$ in Eq. (I.6c) as the entanglement fidelity $F_{\text{ent}}(\mathcal{E})$ defined in Eq. (G.4), we obtain

$$\eta_{\Diamond} \geq 1 - F_{\text{ent}}(\mathcal{E}) + \frac{1}{2}(\operatorname{Tr} J(\mathcal{E}) - 1). \qquad \text{(I.7)}$$

After using $\operatorname{Tr} J(\mathcal{E}) = \operatorname{Tr} \mathcal{E}(\mathbb{1}/d)$ and inserting the relation between entanglement fidelity and average fidelity derived in Appendix G, i.e., $F_{\text{ent}}(\mathcal{E}) = (d+1)F_{\text{avg}}(\mathcal{E})/d - \operatorname{Tr} \mathcal{E}(\mathbb{1}/d)/d$, we finally obtain

$$\eta_{\Diamond} \geq \frac{d+1}{d}(1 - F_{\text{avg}}(\mathcal{E})) - \frac{d+2}{2d}(1 - \operatorname{Tr} \mathcal{E}(\mathbb{1}/d)). \qquad \text{(I.8)}$$

The first term in this expression is the result given in Eq. (I.1), referred to as $\eta_{\Diamond}^{\text{Pauli}}$ in Eq. (6.17a) and [San2016]. The second term is a new contribution that represents the correction required for trace-decreasing quantum operations.

Appendix J

Gate decompositions and effective Hamiltonians

In this appendix, we outline a method to relate an arbitrary quantum operation $\mathcal{G}$ (see Section 2.4) to an effective Hamiltonian H. The method was used to obtain the axis-angle decompositions of the two-qubit GST results studied in Section 6.3 (see Tab. 6.6), which are much more straightforward to interpret than the corresponding Pauli transfer matrices shown in Fig. 6.4. A similar method to obtain such decompositions is implemented by the `pyGSTi` package [Nie2018].

A quantum operation $\mathcal{G}$ can be written in Kraus form (see Eq. (2.24)),

$$\mathcal{G}(\rho) = \sum_{\alpha=1}^{R} E_\alpha \rho E_\alpha^\dagger, \tag{J.1}$$

where E_α are the Kraus operators and R is the Kraus rank of $\mathcal{G}$. The aim of the method is to find a "Hamiltonian" H that approximately generates the evolution described by $\mathcal{G}$ according to

$$\mathcal{G}(\rho) \approx e^{-iH} \rho\, e^{iH}. \tag{J.2}$$

We explicitly left out a symbol for the time in the expression for the generator. Typically, one would rather write $H = \widehat{H}t$, where $\widehat{H}$ is the Hamiltonian and t is the time, to express the characteristic structure of time evolutions in quantum theory. However, since $\mathcal{G}$ describes only one discrete evolution $\rho \mapsto \mathcal{G}(\rho)$, there is no notion of time such that a separation into individual components $\widehat{H}$ and t is arbitrary. For convenience, we still refer to the symbol H as the Hamiltonian.

The relation given in Eq. (J.2) can only be exact if the Kraus rank R in Eq. (J.1) is 1 and the operator E_1 is unitary. This means that the action of $\mathcal{G}$ can be written as $\mathcal{G}(\rho) = U\rho U^\dagger$, where U is a unitary matrix. But even if this is not the case, the following method can be used to produce an approximation to $\mathcal{G}$ in terms of a Hamiltonian, which may be simpler to understand or provide insights into potential errors. In the context of quantum gate optimization, this may provide information on how to improve the gate's implementation.

If $\mathcal{G}$ is not unitary but has a Kraus rank R of almost one (meaning that all but one singular value of the Choi matrix $J(\mathcal{G})$ given by Eq. (2.25) are close to zero), the method yields an effective Hamiltonian that approximates the quantum operation. This Hamiltonian may have non-Hermitian components to model non-unitary or non-trace-preserving quantum operations.

Note that a Hamiltonian generating $\mathcal{G}$ according to Eq. (J.2) is not unique, since the complex matrix exponential is a many-to-one function. One way to see this is that by adding a multiple of 2π to any eigenvalue of H, the matrix exponential e^{-iH} does not change. In what follows, we aim for an expression for H in terms of Pauli matrices that can be readily interpreted as a rotation by a certain angle around an axis specified by H, so this ambiguity can be understood in the context of rotations.

J.1 The matrix logarithm

We start with the matrix representation G of the superoperator $\mathcal{G}$ defined in Eq. (2.27) (the Pauli transfer matrix). We denote the correspondence between the matrix representation G and the map $\mathcal{G}$ by

$$G|\rho\rangle\!\rangle \quad \leftrightarrow \quad \mathcal{G}(\rho), \tag{J.3}$$

where $|\rho\rangle\!\rangle$ denotes a vector representation of the density matrix ρ, obtained by expanding $\rho = \sum_i \rho_i \widehat{P}_i$ in the normalized Pauli basis $\widehat{P}_i$ (see Eq. (2.26)).

If $\mathcal{G}$ preserves Hermiticity, the matrix G is a real $d \times d$ matrix, where $d = N^2 = 4^n$ for an n-qubit system. We define the matrix logarithm of $G = e^L$ as

$$L = \log G = V \log(\Lambda) V^{-1}, \tag{J.4}$$

where $G = V \Lambda V^{-1}$ denotes the eigendecomposition of G (see [Mey2009b] for a common alternative definition of the matrix logarithm). Thus, $\Lambda = \mathrm{diag}(\lambda_0, \ldots, \lambda_{d-1})$ contains the eigenvalues of G, the columns of $V = (v_0, \ldots, v_{d-1})$ contain the right eigenvectors, and the columns of $(V^{-1})^\dagger = (w_0, \ldots, w_{d-1})$ contain the left eigenvectors. In this definition, $\log \lambda_i = \log|\lambda_i| + i \arg \lambda_i$ denotes the principal logarithm defined by $\arg \lambda_i \in (-\pi, \pi)$. If all $\lambda_i \in \mathbb{C} \setminus (-\infty, 0]$, the principal matrix logarithm defined by Eq. (J.4) is unique. If some $\lambda_i = 0$, the matrix G is singular and the matrix logarithm does not exist.

If the matrix logarithm L is real, there is a Hermiticity-preserving map $\mathcal{L}$ such that

$$G|\rho\rangle\!\rangle = e^L|\rho\rangle\!\rangle \quad \leftrightarrow \quad \mathcal{G}(\rho) = e^{\mathcal{L}}\rho. \tag{J.5}$$

Unfortunately, some ideal quantum gates are special in the sense that their Pauli transfer matrices have eigenvalues -1. This is the case for the CNOT gate defined in Eq. (2.18), for which the Pauli transfer matrix corresponding to the map $\rho \mapsto \mathrm{CNOT}\, \rho\, \mathrm{CNOT}^\dagger$ has eigenvalues $\lambda_i = \pm 1$. In this case, however, one can still find a real matrix logarithm since the negative eigenvalues occur in pairs [Cul1966]. It is constructed by choosing, for each pair $\lambda_i = \lambda_j = -1$, two conjugate branches of the logarithm $\log \lambda_{i/j} \leftarrow \pm i\pi$. Similarly, the corresponding real eigenvectors v_i and v_j in the degenerate subspace have to be replaced by conjugate pairs, i.e., $v_i \leftarrow (v_i + iv_j)/\sqrt{2}$ and $v_j \leftarrow (v_i - iv_j)/\sqrt{2}$.

J.2 Extracting the Hamiltonian

The map $\mathcal{L}$ in Eq. (J.5) is typically called the Lindblad operator or *Lindbladian*. The goal is to approximate $\mathcal{L}$ by a map of the form $\mathcal{L}^H(\rho) = -i[H, \rho]$. If this is possible, the correspondence given by Eq. (J.5) becomes

$$e^L|\rho\rangle\!\rangle \quad \leftrightarrow \quad e^{\mathcal{L}}\rho \approx e^{-i[H, \cdot]}\rho = e^{-iH}\rho\, e^{iH}. \tag{J.6}$$

We consider a Hamiltonian H expressed in the Pauli basis (see Eq. (2.26)),

$$H = \sum_{k=0}^{d-1} h_k P_k/2. \tag{J.7}$$

The real coefficients h_k are explicitly defined with respect to $P_k/2$ instead of the normalized basis elements $\widehat{P}_k$. The reason is that in this way, the action of e^{-iH} can be interpreted as a rotation about an axis specified by $\widehat{h} = h/|h|$, where the angle of rotation is given by $\varphi = |h| = \sqrt{\sum_k h_k^2}$ (cf. Eqs. (6.31a) and (6.31b)).

Using the form of the Hamiltonian given by Eq. (J.7), we evaluate the Pauli transfer matrix L^H corresponding to the map $\mathcal{L}^H = -i[H, \cdot]$. Its matrix elements are (cf. Eq. (2.27))

$$L_{ij}^H = \frac{1}{N}\operatorname{Tr} P_i(-i[H, P_j]) = \sum_k h_k \frac{\operatorname{Tr} P_i[P_k, P_j]}{2Ni} = \sum_k h_k s_{kji}, \tag{J.8}$$

where $s_{kji} = \operatorname{Tr}([P_k, P_j]P_i)/2Ni$. Note that, as each P_i is an n-fold tensor product of Pauli matrices, evaluating s_{kji} analytically may be cumbersome. However, s_{kji} can be easily evaluated with computer algebra systems such as Mathematica [Wol2019]. One finds $s_{kji} \in \{1, 0, -1\}$ (see also [Rig2009], in which s_{kji} is called the n-qubit super-commutator). The typical structure of the matrix L^H for $n \geq 1$ qubits is

$$L^H = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & -h_3 & h_2 & \cdots \\ 0 & h_3 & 0 & -h_1 & \cdots \\ 0 & -h_2 & h_1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{J.9}$$

The upper left 4×4 block corresponds to $n = 1$ qubit, the upper left 16×16 block corresponds to $n = 2$ qubits, and so on.

The map $\mathcal{G}$ can be approximately generated by the Hamiltonian given in Eq. (J.7) if the matrix logarithm $L = \log G$ given by Eq. (J.4) has the form of the matrix L^H in Eq. (J.9). We construct a candidate Hamiltonian by projecting L onto this form, such that

$$h_k = \begin{cases} \sum_k L_{ij}s_{kji}/\sum_k |s_{kji}| & (\sum_k |s_{kji}| \neq 0) \\ 0 & (\text{otherwise}) \end{cases} \tag{J.10}$$

for all $k = 0, \ldots, d - 1$. If the decomposition error defined by

$$\gamma = \|G - e^{L^H}\|_F \tag{J.11}$$

is much smaller than 1, we accept the Hamiltonian generator. This was the case for almost all experiments studied in this thesis.

However, since the matrix logarithm given by Eq. (J.4) is not unique, and because the correspondence in Eq. (J.6) is only approximate, there may of course be other Hamiltonians generating the evolution. One way to proceed is to optimize the entries of L towards the intended target gate's matrix logarithm $L^U = \log G^U$, where G^U is the Pauli transfer matrix of the map $\mathcal{G}^U = U \cdot U^\dagger$ and U is the target gate. This can be done by optimizing a joint objective function including both $\|G - e^{L^H}\|$ and $\|L^U - L^H\|$. This option is implemented by the `pyGSTi` package [Nie2018] and was also used to obtain

the decomposition of the CNOT gate reported in Tab. 6.6. The objective function was $\|G - e^{L^H}\|_1 + 10\|L^U - L^H\|_F^2$, where $\|A\|_1 = \sum_{ij}|A_{ij}|$, using the L-BFGS-B algorithm [Zhu1997; Mor2011]. However, for two qubits, L already contains $d \times d = 256$ real numbers, so this approach becomes impractical for more qubits.

Another option is to optimize the d coefficients h_k defined by Eq. (J.7), starting from initial values given by Eq. (J.10) and using the objective function defined in Eq. (J.11). This approach directly relates the Hamiltonian H to the map $\mathcal{G}$ without using an explicit matrix logarithm, so the approach is closer in spirit to the relation anticipated by Eq. (J.2). The advantage is that this approach does not rely on the matrix logarithm as an intermediate step. Furthermore, it only requires an optimization of d real numbers instead of d^2. A similar approach was implemented in [Wil2016].

A third option would be to resolve the ambiguity of the matrix logarithm by adding integer multiples of $2\pi i v_j w_j^\dagger$ to L. This corresponds to changing the eigenvalues $\lambda_j \mapsto \lambda_j + 2\pi i$, which leaves e^L invariant but may lead to imaginary parts in the Hamiltonian. If the resulting Hamiltonian indeed describes the original map $\mathcal{G}$, as reflected by a low decomposition error γ (see Eq. (J.11)), this procedure may provide an effective qubit model for decay in the system under investigation.

Bibliography

[Aar2008]   S. AARONSON, "The Limits of Quantum Computers", *Sci. Am.* **298**, 62 (2008) [cit. on p. 6].

[Aar2013]   S. AARONSON, *Quantum Computing Since Democritus*, Cambridge University Press, 2013 [cit. on p. 182].

[Aar2015]   S. AARONSON, "Read the fine print", *Nat. Phys.* **11**, 291 (2015) [cit. on p. 2].

[Aha1997]   D. AHARONOV AND M. BEN-OR, "Fault-Tolerant Quantum Computation with Constant Error", *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97, 176 (1997) [cit. on p. 2].

[Aha1998]   D. AHARONOV, A. KITAEV, AND N. NISAN, "Quantum Circuits with Mixed States", *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, 20 (1998) [cit. on pp. 112, 113, 194, 195].

[Aha2008]   D. AHARONOV AND M. BEN-OR, "Fault-Tolerant Quantum Computation with Constant Error Rate", *SIAM J. Comput.* **38**, 1207 (2008) [cit. on pp. 111, 155].

[Ale2020]   T. ALEXANDER, N. KANAZAWA, D. EGGER, L. CAPELLUTO, C. WOOD, A. JAVADI-ABHARI, AND D. MCKAY, "Qiskit Pulse: Programming Quantum Computers Through the Cloud with Pulses", arXiv:2004.06755 (2020) [cit. on pp. 90, 125].

[Ali2006]   P. ALIFERIS, D. GOTTESMAN, AND J. PRESKILL, "Quantum Accuracy Threshold for Concatenated Distance-3 Codes", *Quantum Inf. Comput.* **6**, 97 (2006) [cit. on pp. 111, 155].

[Ali2007]   P. ALIFERIS AND B. TERHAL, "Fault-tolerant Quantum Computation for Local Leakage Faults", *Quantum Inf. Comput.* **7**, 139 (2007) [cit. on pp. 2, 111, 155].

[An2019]   Z. AN AND D. ZHOU, "Deep reinforcement learning for quantum gate control", *EPL* **126**, 60002 (2019) [cit. on pp. 107, 169].

[Ans2019]   M. ANSARI, "Superconducting qubits beyond the dispersive regime", *Phys. Rev. B* **100**, 024509 (2019) [cit. on pp. 48, 169].

[Aru2019]   F. Arute, K. Arya, R. Babbush, D. Bacon, J. Bardin, R. Barends, R. Biswas, S. Boixo, F. Brandao, D. 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. Harrigan, M. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. Humble, S. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Niu, E. Ostby, A. Petukhov, J. Platt, C. Quintana, E. Rieffel, P. Roushan, N. Rubin, D. Sank, K. Satzinger, V. Smelyanskiy, K. Sung, M. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. Yao, P. Yeh, A. Zalcman, H. Neven, and J. Martinis, "Quantum supremacy using a programmable superconducting processor", *Nature* **574**, 505 (2019) [cit. on pp. 2, 3, 25, 27, 153, 155, 167, 176].

[Ata1940]   J. Atanasoff, "Computing Machine for the Solution of large Systems of Linear Algebraic Equations", grant proposal (reprinted in [Ran1973]), 1940 [cit. on p. 1].

[Aud2018]   C. Audet and C. Tribes, "Mesh-based Nelder–Mead algorithm for inequality constrained optimization", *Comput. Optim. Appl.* **71**, 331 (2018) [cit. on p. 99].

[Bab1837]   C. Babbage, "On the Mathematical Powers of the Calculating Engine", unpublished manuscript (reprinted in full in [Ran1973]), 1837 [cit. on p. 1].

[Bab2018]   R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. Chan, "Low-Depth Quantum Simulation of Materials", *Phys. Rev. X* **8**, 011044 (2018) [cit. on p. 2].

[Bal1998]   L. Ballentine, *Quantum Mechanics: A Modern Development*, World Scientific, 1998 [cit. on pp. 1, 24, 136, 180].

[Bar1957]   J. Bardeen, L. Cooper, and J. Schrieffer, "Theory of Superconductivity", *Phys. Rev.* **108**, 1175 (1957) [cit. on p. 24].

[Bar1964]   V. Bargmann, "Note on Wigner's Theorem on Symmetry Operations", *J. Math. Phys.* **5**, 862 (1964) [cit. on p. 178].

[Bar1995]   A. Barenco, C. Bennett, R. Cleve, D. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, "Elementary gates for quantum computation", *Phys. Rev. A* **52**, 3457 (1995) [cit. on p. 2].

[Bar2007]   S. Barnett and J. Vaccaro, *The Quantum Phase Operator: A Review*, CRC Press, 2007 [cit. on p. 25].

[Bar2013] R. BARENDS, J. KELLY, A. MEGRANT, D. SANK, E. JEFFREY, Y. CHEN, Y. YIN, B. CHIARO, J. MUTUS, C. NEILL, P. O'MALLEY, P. ROUSHAN, J. WENNER, T. WHITE, A. CLELAND, AND J. MARTINIS, "Coherent Josephson Qubit Suitable for Scalable Quantum Integrated Circuits", *Phys. Rev. Lett.* **111**, 080502 (2013) [cit. on p. 27].

[Bar2014] R. BARENDS, J. KELLY, A. MEGRANT, A. VEITIA, D. SANK, E. JEFFREY, T. WHITE, J. MUTUS, A. FOWLER, B. CAMPBELL, Y. CHEN, Z. CHEN, B. CHIARO, A. DUNSWORTH, C. NEILL, P. O'MALLEY, P. ROUSHAN, A. VAINSENCHER, J. WENNER, A. KOROTKOV, A. CLELAND, AND J. MARTINIS, "Superconducting quantum circuits at the surface code threshold for fault tolerance", *Nature* **508**, 500 (2014) [cit. on pp. 115, 116].

[Bar2017] J. BARVÍNEK AND J. HAMHALTER, "Linear algebraic proof of Wigner theorem and its consequences", *Mathematica Slovaca* **67**, 371 (2017) [cit. on p. 181].

[Bel1964] J. BELL, "On the Einstein Podolsky Rosen paradox", *Physics* **1**, 195 (1964) [cit. on p. 10].

[Bel2004] J. BELL AND A. ASPECT, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, Cambridge University Press, 2004 [cit. on p. 10].

[Ben1973] C. BENNETT, "Logical Reversibility of Computation", *IBM J. Res. Dev.* **17**, 525 (1973) [cit. on p. 12].

[Ben1980] P. BENIOFF, "The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines", *J. Stat. Phys.* **22**, 563 (1980) [cit. on p. 2].

[Ben2006] I. BENGTSSON AND K. ŻYCZKOWSKI, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, Cambridge University Press, 2006 [cit. on pp. 10, 111, 116, 193].

[Ben2010] A. BEN-AROYA AND A. TA-SHMA, "On the Complexity of Approximating the Diamond Norm", *Quantum Inf. Comput.* **10**, 77 (2010) [cit. on p. 112].

[Ber1997] E. BERNSTEIN AND U. VAZIRANI, "Quantum Complexity Theory", *SIAM J. Comput.* **26**, 11 (1997) [cit. on p. 178].

[Bil2015] P.-M. BILLANGEON, J. TSAI, AND Y. NAKAMURA, "Circuit-QED-based scalable architectures for quantum information processing with superconducting qubits", *Phys. Rev. B* **91**, 094517 (2015) [cit. on pp. 80, 82, 147].

[Bla2004] A. BLAIS, R.-S. HUANG, A. WALLRAFF, S. GIRVIN, AND R. SCHOELKOPF, "Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation", *Phys. Rev. A* **69**, 062320 (2004) [cit. on pp. 2, 25, 68, 80].

[Bla2016]   J. BLACKBURN, M. CIRILLO, AND N. GRØNBECH-JENSEN, "A survey of classical and quantum interpretations of experiments on Josephson junctions at very low temperatures", *Phys. Rep.* **611**, 1 (2016) [cit. on pp. 23, 170].

[Bla2017]   J. BLACKBURN, M. CIRILLO, AND N. GRØNBECH-JENSEN, "Investigation of low temperature quantum crossover in Josephson junctions", *J. Appl. Phys.* **122**, 133904 (2017) [cit. on p. 23].

[Blu2013]   R. BLUME-KOHOUT, J. GAMBLE, E. NIELSEN, J. MIZRAHI, J. STERK, AND P. MAUNZ, "Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit", arXiv:1310.4492 (2013) [cit. on pp. 125, 127, 129].

[Blu2017]   R. BLUME-KOHOUT, J. GAMBLE, E. NIELSEN, K. RUDINGER, J. MIZRAHI, K. FORTIER, AND P. MAUNZ, "Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography", *Nat. Commun.* **8**, 14485 (2017) [cit. on pp. 89, 125, 126, 128, 129].

[Bog1958]   N. BOGOLJUBOV, "On a new method in the theory of superconductivity", *Nuovo Cim* **7**, 794 (1958) [cit. on p. 56].

[Boi2017]   S. BOIXO, S. ISAKOV, V. SMELYANSKIY, AND H. NEVEN, "Simulation of low-depth quantum circuits as complex undirected graphical models", arXiv:1712.05384 (2017) [cit. on p. 176].

[Boi2018]   S. BOIXO, S. ISAKOV, V. SMELYANSKIY, R. BABBUSH, N. DING, Z. JIANG, M. BREMNER, J. MARTINIS, AND H. NEVE, "Characterizing quantum supremacy in near-term devices", *Nat. Phys.* **14**, 595 (2018) [cit. on pp. 2, 193].

[Bon2010]   I. BONGIOANNI, L. SANSONI, F. SCIARRINO, G. VALLONE, AND P. MATALONI, "Experimental quantum process tomography of non-trace-preserving maps", *Phys. Rev. A* **82**, 042307 (2010) [cit. on p. 111].

[Boo1847]   G. BOOLE, *The Mathematical Analysis of Logic: Being an Essay Towards a Calculus of Deductive Reasoning*, Philosophical Library (reprinted by Cambridge University Press), 1847 [cit. on pp. 1, 6].

[Bou1998]   V. BOUCHIAT, D. VION, P. JOYEZ, D. ESTEVE, AND M. DEVORET, "Quantum coherence with a single Cooper pair", *Phys. Scr.* **1998**, 165 (1998) [cit. on p. 25].

[Bou2012]   J. BOURASSA, F. BEAUDOIN, J. GAMBETTA, AND A. BLAIS, "Josephson-junction-embedded transmission-line resonators: From Kerr medium to in-line transmon", *Phys. Rev. A* **86**, 013814 (2012) [cit. on p. 48].

[Boy2004]   S. BOYD AND L. VANDENBERGHE, *Convex Optimization*, Cambridge University Press, 2004 [cit. on p. 113].

[Bra2011]   S. BRAVYI, D. DIVINCENZO, AND D. LOSS, "Schrieffer-Wolff transformation for quantum many-body systems", *Ann. Phys. (N.Y.)* **326**, 2793 (2011) [cit. on p. 69].

[Bre2007]   H. BREUER AND F. PETRUCCIONE, *The Theory of Open Quantum Systems*, Oxford University Press, 2007 [cit. on pp. 19, 75].

[Bru2004]   H. BRUUS AND K. FLENSBERG, *Many-Body Quantum Theory in Condensed Matter Physics: An Introduction*, Oxford University Press, 2004 [cit. on p. 32].

[Bul2016]   C. BULTINK, M. ROL, T. O'BRIEN, X. FU, B. DIKKEN, C. DICKEL, R. VERMEULEN, J. DE STERKE, A. BRUNO, R. SCHOUTEN, AND L. DICARLO, "Active Resonator Reset in the Nonlinear Dispersive Regime of Circuit QED", *Phys. Rev. Applied* **6**, 034008 (2016) [cit. on p. 72].

[Cal2018]   S. CALDWELL, N. DIDIER, C. RYAN, E. SETE, A. HUDSON, P. KARALEKAS, R. MANENTI, M. DA SILVA, R. SINCLAIR, E. ACALA, N. ALIDOUST, J. ANGELES, A. BESTWICK, M. BLOCK, B. BLOOM, A. BRADLEY, C. BUI, L. CAPELLUTO, R. CHILCOTT, J. CORDOVA, G. CROSSMAN, M. CURTIS, S. DESHPANDE, T. BOUAYADI, D. GIRSHOVICH, S. HONG, K. KUANG, M. LENIHAN, T. MANNING, A. MARCHENKOV, J. MARSHALL, R. MAYDRA, Y. MOHAN, W. O'BRIEN, C. OSBORN, J. OTTERBACH, A. PAPAGEORGE, J.-P. PAQUETTE, M. PELSTRING, A. POLLORENO, G. PRAWIROATMODJO, V. RAWAT, M. REAGOR, R. RENZAS, N. RUBIN, D. RUSSELL, M. RUST, D. SCARABELLI, M. SCHEER, M. SELVANAYAGAM, R. SMITH, A. STALEY, M. SUSKA, N. TEZAK, D. THOMPSON, T.-W. TO, M. VAHIDPOUR, N. VODRAHALLI, T. WHYLAND, K. YADAV, W. ZENG, AND C. RIGETTI, "Parametrically Activated Entangling Gates Using Transmon Qubits", *Phys. Rev. Applied* **10**, 034050 (2018) [cit. on p. 27].

[Cam2017]   E. CAMPBELL, B. TERHAL, AND C. VUILLOT, "Roads towards fault-tolerant universal quantum computation", *Nature* **549**, 172 (2017) [cit. on p. 155].

[Car1968]   P. CARRUTHERS AND M. NIETO, "Phase and Angle Variables in Quantum Mechanics", *Rev. Mod. Phys.* **40**, 411 (1968) [cit. on p. 25].

[Car2008]   H. CARTERET, D. TERNO, AND K. ŻYCZKOWSKI, "Dynamics beyond completely positive maps: Some properties and applications", *Phys. Rev. A* **77**, 042113 (2008) [cit. on p. 20].

[Cat1998]   A. CATICHA, "Consistency, amplitudes, and probabilities in quantum theory", *Phys. Rev. A* **57**, 1572 (1998) [cit. on p. 182].

[Cat2011]   G. CATELANI, J. KOCH, L. FRUNZIO, R. SCHOELKOPF, M. DEVORET, AND L. GLAZMAN, "Quasiparticle Relaxation of Superconducting Qubits in the Presence of Flux", *Phys. Rev. Lett.* **106**, 077002 (2011) [cit. on p. 79].

[Cat2012]   G. CATELANI, S. NIGG, S. GIRVIN, R. SCHOELKOPF, AND L. GLAZMAN, "Decoherence of superconducting qubits caused by quasiparticle tunneling", *Phys. Rev. B* **86**, 184514 (2012) [cit. on p. 79].

[Che2016]   Z. Chen, J. Kelly, C. Quintana, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. Fowler, E. Lucero, E. Jeffrey, A. Megrant, J. Mutus, M. Neeley, C. Neill, P. O'Malley, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. White, A. Korotkov, and J. Martinis, "Measuring and Suppressing Quantum State Leakage in a Superconducting Qubit", *Phys. Rev. Lett.* **116**, 020501 (2016) [cit. on pp. 39, 85].

[Che2018a]   J. Chen, F. Zhang, C. Huang, M. Newman, and Y. Shi, "Classical Simulation of Intermediate-Size Quantum Circuits", arXiv:1805.01450 (2018) [cit. on p. 176].

[Che2018b]   Z.-Y. Chen, Q. Zhou, C. Xue, X. Yang, G.-C. Guo, and G.-P. Guo, "64-qubit quantum circuit simulation", *Sci. Bull.* **63**, 964 (2018) [cit. on p. 176].

[Chi2018a]   N.-H. Chia, H.-H. Lin, and C. Wang, "Quantum-inspired sublinear classical algorithms for solving low-rank linear systems", arXiv:1811.04852 (2018) [cit. on p. 5].

[Chi2018b]   Chinese Academy of Sciences, *Quantum Computing Cloud Platform*, 2018, `http://quantumcomputer.ac.cn/index.html` [cit. on pp. 2, 3].

[Cho1975]   M.-D. Choi, "Completely positive linear maps on complex matrices", *Linear Algebra Appl.* **10**, 285 (1975) [cit. on p. 17].

[Cho2010]   J. Chow, L. DiCarlo, J. Gambetta, F. Motzoi, L. Frunzio, S. Girvin, and R. Schoelkopf, "Optimized driving of superconducting artificial atoms for improved single-qubit gates", *Phys. Rev. A* **82**, 040305 (2010) [cit. on p. 88].

[Cho2011]   J. Chow, A. Córcoles, J. Gambetta, C. Rigetti, B. Johnson, J. Smolin, J. Rozen, G. Keefe, M. Rothwell, M. Ketchen, and M. Steffen, "Simple All-Microwave Entangling Gate for Fixed-Frequency Superconducting Qubits", *Phys. Rev. Lett.* **107**, 080502 (2011) [cit. on p. 89].

[Cho2012]   J. Chow, J. Gambetta, A. Córcoles, S. Merkel, J. Smolin, C. Rigetti, S. Poletto, G. Keefe, M. Rothwell, J. Rozen, M. Ketchen, and M. Steffen, "Universal Quantum Gate Set Approaching Fault-Tolerant Thresholds with Superconducting Qubits", *Phys. Rev. Lett.* **109**, 060501 (2012) [cit. on pp. 25, 115].

[Chr2017]   D. Chruściński and S. Pascazio, "A Brief History of the GKLS Equation", *Open Syst. Inf. Dyn* **24**, 1740001 (2017) [cit. on p. 75].

[Chu1997]   I. Chuang and M. Nielsen, "Prescription for experimental determination of the dynamics of a quantum black box", *J. Mod. Opt.* **44**, 2455 (1997) [cit. on p. 127].

[Cir1995]   J. Cirac and P. Zoller, "Quantum Computations with Cold Trapped Ions", *Phys. Rev. Lett.* **74**, 4091 (1995) [cit. on p. 2].

[Cla1988] J. Clarke, A. Cleland, M. Devoret, D. Esteve, and J. Martinis, "Quantum Mechanics of a Macroscopic Variable: The Phase Difference of a Josephson Junction", *Science* **239**, 992 (1988) [cit. on p. 22].

[Col1978] J. Colpa, "Diagonalization of the quadratic boson hamiltonian", *Physica A* **93**, 327 (1978) [cit. on p. 56].

[Coo1965] J. Cooley and J. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series", *Math. Comput.* **19**, 297 (1965) [cit. on p. 16].

[Cop2017] B. Copeland, "The Modern History of Computing", in: *The Stanford Encyclopedia of Philosophy*, ed. by E. Zalta (2017) [cit. on p. 1].

[Cór2013] A. Córcoles, J. Gambetta, J. Chow, J. Smolin, M. Ware, J. Strand, B. Plourde, and M. Steffen, "Process verification of two-qubit quantum gates by randomized benchmarking", *Phys. Rev. A* **87**, 030301 (2013) [cit. on pp. 89, 94].

[Cox1946] R. Cox, "Probability, Frequency and Reasonable Expectation", *Am. J. Phys.* **14**, 1 (1946) [cit. on p. 182].

[Cox1961] R. Cox, *The Algebra of Probable Inference*, Johns Hopkins University Press, 1961 [cit. on p. 182].

[Cro2017] A. Cross, L. Bishop, J. Smolin, and J. Gambetta, "Open Quantum Assembly Language", arXiv:1707.03429 (2017) [cit. on pp. 13, 32, 87–89, 105, 176].

[Cul1966] W. Culver, "On the existence and uniqueness of the real logarithm of a matrix", *Proc. Amer. Math. Soc.* **17**, 1146 (1966) [cit. on p. 199].

[Dem2008] J. Demmel, O. Marques, B. Parlett, and C. Vömel, "Performance and Accuracy of LAPACK's Symmetric Tridiagonal Eigensolvers", *SIAM J. Sci. Comput.* **30**, 1508 (2008) [cit. on p. 61].

[DeR1987] H. De Raedt, "Product formula algorithms for solving the time dependent Schrödinger equation", *Comput. Phys. Rep.* **7**, 1 (1987) [cit. on pp. 3, 33, 34, 38, 39, 60, 62, 167, 186].

[DeR2000] H. De Raedt, A. Hams, K. Michielsen, and K. De Raedt, "Quantum Computer Emulator", *Comput. Phys. Commun.* **132**, 1 (2000) [cit. on pp. 3, 33].

[DeR2002] H. De Raedt, K. Michielsen, A. Hams, S. Miyashita, and K. Saito, "Quantum spin dynamics as a model for quantum computer operation", *Eur. Phys. J. B* **27**, 15 (2002) [cit. on p. 33].

[DeR2005] K. De Raedt, H. De Raedt, and K. Michielsen, "Deterministic event-based simulation of quantum phenomena", *Comput. Phys. Commun.* **171**, 19 (2005) [cit. on p. 10].

Bibliography

[DeR2006] H. DE RAEDT AND K. MICHIELSEN, "Computational Methods for Simulating Quantum Computers", *Quantum and Molecular Computing, Quantum Simulations, Handbook of Theoretical and Computational Nanotechnology* **3**, 1, ed. by M. Rieth and W. Schommers (2006) [cit. on pp. 3, 33].

[DeR2007] K. DE RAEDT, K. MICHIELSEN, H. DE RAEDT, B. TRIEU, G. ARNOLD, M. RICHTER, TH. LIPPERT, H. WATANABE, AND N. ITO, "Massively parallel quantum computer simulator", *Comput. Phys. Commun.* **176**, 121 (2007) [cit. on pp. 3, 176].

[DeR2014] H. DE RAEDT, M. KATSNELSON, AND K. MICHIELSEN, "Quantum theory as the most robust description of reproducible experiments", *Ann. Phys. (N.Y.)* **347**, 45 (2014) [cit. on p. 182].

[DeR2015a] H. DE RAEDT, M. KATSNELSON, H. DONKER, AND K. MICHIELSEN, "Quantum theory as a description of robust experiments: application to Stern-Gerlach and Einstein-Podolsky-Rosen-Bohm experiments", *Proc. SPIE* **9570**, 957002 (2015) [cit. on p. 182].

[DeR2015b] H. DE RAEDT, M. KATSNELSON, H. DONKER, AND K. MICHIELSEN, "Quantum theory as a description of robust experiments: Derivation of the Pauli equation", *Ann. Phys. (N.Y.)* **359**, 166 (2015) [cit. on p. 182].

[DeR2016] H. DE RAEDT, M. KATSNELSON, AND K. MICHIELSEN, "Quantum theory as plausible reasoning applied to data obtained by robust experiments", *Philos. Trans. R. Soc. A* **374**, 20150233 (2016) [cit. on p. 182].

[DeR2017] H. DE RAEDT, F. JIN, M. KATSNELSON, AND K. MICHIELSEN, "Relaxation, thermalization, and Markovian dynamics of two spins coupled to a spin bath", *Phys. Rev. E* **96**, 053306 (2017) [cit. on pp. 75, 77].

[DeR2019a] H. DE RAEDT, F. JIN, D. WILLSCH, M. WILLSCH, N. YOSHIOKA, N. ITO, S. YUAN, AND K. MICHIELSEN, "Massively parallel quantum computer simulator, eleven years later", *Comput. Phys. Commun.* **237**, 47 (2019) [cit. on pp. 3, 32, 105, 154, 173, 175, 176].

[DeR2019b] H. DE RAEDT, M. KATSNELSON, D. WILLSCH, AND K. MICHIELSEN, "Separation of conditions as a prerequisite for quantum theory", *Ann. Phys. (N.Y.)* **403**, 112 (2019) [cit. on pp. 128, 129].

[DeR2020] H. DE RAEDT, M. JATTANA, D. WILLSCH, M. WILLSCH, F. JIN, AND K. MICHIELSEN, "Discrete-Event Simulation of an Extended Einstein-Podolsky-Rosen-Bohm Experiment", *Front. Phys.* **8**, 160 (2020) [cit. on p. 10].

[Deu1985] D. DEUTSCH, "Quantum theory, the Church-Turing principle and the universal quantum computer", *Proc. R. Soc. Lond. A* **400**, 97 (1985) [cit. on p. 2].

[Deu1995] D. DEUTSCH, A. BARENCO, AND A. EKERT, "Universality in Quantum Computation", *Proc. R. Soc. Lond. A* **449**, 669 (1995) [cit. on pp. 2, 14].

[Dev1985] M. DEVORET, J. MARTINIS, AND J. CLARKE, "Measurements of Macroscopic Quantum Tunneling out of the Zero-Voltage State of a Current-Biased Josephson Junction", *Phys. Rev. Lett.* **55**, 1908 (1985) [cit. on p. 22].

[Dev1997] M. DEVORET, "Quantum Fluctuations in Electrical Circuits", *Fluctuations Quantiques/Quantum Fluctuations: Les Houches Session LXIII*, 351, ed. by S. Reynaud, E. Giacobino, and J. Zinn-Justin (1997) [cit. on pp. 22, 68].

[Did2018] N. DIDIER, E. SETE, M. DA SILVA, AND C. RIGETTI, "Analytical modeling of parametrically modulated transmon qubits", *Phys. Rev. A* **97**, 022330 (2018) [cit. on pp. 27, 29].

[Dir1927] P. DIRAC AND N. BOHR, "The quantum theory of the emission and absorption of radiation", *Proc. R. Soc. Lond. A* **114**, 243 (1927) [cit. on p. 25].

[DiV1995] D. DIVINCENZO, "Two-bit gates are universal for quantum computation", *Phys. Rev. A* **51**, 1015 (1995) [cit. on pp. 2, 14].

[DiV1996] D. DIVINCENZO AND P. SHOR, "Fault-Tolerant Error Correction with Efficient Quantum Codes", *Phys. Rev. Lett.* **77**, 3260 (1996) [cit. on p. 155].

[DiV2000] D. DIVINCENZO, D. BACON, J. KEMPE, G. BURKARD, AND K. WHALEY, "Universal quantum computation with the exchange interaction", *Nature* **408**, 339 (2000) [cit. on pp. 14, 176].

[Dom2016] J. DOMINY AND D. LIDAR, "Beyond complete positivity", *Quantum Inf. Process.* **15**, 1349 (2016) [cit. on p. 20].

[Don2016] H. DONKER, M. KATSNELSON, H. DE RAEDT, AND K. MICHIELSEN, "Logical inference approach to relativistic quantum mechanics: Derivation of the Klein-Gordon equation", *Ann. Phys. (N.Y.)* **372**, 74 (2016) [cit. on p. 182].

[Dop2009] F. DOPICO AND C. JOHNSON, "Parametrization of the Matrix Symplectic Group and Applications", *SIAM J. Matrix Anal. Appl.* **31**, 650 (2009) [cit. on p. 57].

[DWa2019] D-WAVE, *Leap*, 2019, https://www.dwavesys.com/take-leap [cit. on p. 2].

[Ede2005] A. EDELMAN AND N. RAO, "Random matrix theory", *Acta Numer.* **14**, 233 (2005) [cit. on p. 193].

[Eke1996] A. EKERT AND R. JOZSA, "Quantum computation and Shor's factoring algorithm", *Rev. Mod. Phys.* **68**, 733 (1996) [cit. on p. 2].

[Eme2005] J. EMERSON, R. ALICKI, AND K. ŻYCZKOWSKI, "Scalable noise estimation with random unitary operators", *J. Opt. B: Quantum Semiclassical Opt.* **7**, S347 (2005) [cit. on p. 111].

[Eng2011] E. ENGEL AND R. DREIZLER, *Density Functional Theory: An Advanced Course*, Springer-Verlag Berlin Heidelberg, 2011 [cit. on p. 183].

[Far2000]   E. FARHI, J. GOLDSTONE, S. GUTMANN, AND M. SIPSER, "Quantum Computation by Adiabatic Evolution", arXiv:quant-ph/0001106 (2000) [cit. on p. 2].

[Fel2009]   L. FELSEN, M. MONGIARDO, AND P. RUSSER, *Electromagnetic Field Computation by Network Methods*, Springer-Verlag Berlin Heidelberg, 2009 [cit. on p. 51].

[Fey1982]   R. FEYNMAN, "Simulating physics with computers", *Int. J. Theor. Phys.* **21**, 467 (1982) [cit. on p. 2].

[Fin1994]   A. FINNILA, M. GOMEZ, C. SEBENIK, C. STENSON, AND J. DOLL, "Quantum annealing: A new method for minimizing multidimensional functions", *Chem. Phys. Lett.* **219**, 343 (1994) [cit. on p. 2].

[Fos1924]   R. FOSTER, "A reactance theorem", *Bell Syst. Tech. J.* **3**, 259 (1924) [cit. on pp. 48, 50].

[Fox2006]   M. FOX, *Quantum Optics: An Introduction*, Oxford University Press, 2006 [cit. on pp. 24, 69, 79].

[Fre1982]   E. FREDKIN AND T. TOFFOLI, "Conservative logic", *Int. J. Theor. Phys.* **21**, 219 (1982) [cit. on p. 12].

[Fuc1995]   C. FUCHS, "Distinguishability and Accessible Information in Quantum Theory", PhD thesis, University of New Mexico, 1995 [cit. on p. 190].

[Fuc1999]   C. FUCHS AND J. VAN DE GRAAF, "Cryptographic distinguishability measures for quantum-mechanical states", *IEEE Trans. Inf. Theory* **45**, 1216 (1999) [cit. on p. 197].

[Fuc2000]   C. FUCHS AND A. PERES, "Quantum Theory Needs No 'Interpretation'", *Phys. Today* **53**, 70 (2000) [cit. on p. 182].

[Fuc2001]   C. FUCHS, "Quantum Foundations in the Light of Quantum Information", arXiv:quant-ph/0106166 (2001) [cit. on pp. 127, 182].

[Fuc2002]   C. FUCHS, "Quantum Mechanics as Quantum Information (and only a little more)", arXiv:quant-ph/0205039 (2002) [cit. on pp. 127, 182].

[Gam2011]   J. GAMBETTA, F. MOTZOI, S. MERKEL, AND F. WILHELM, "Analytic control methods for high-fidelity unitary operations in a weakly nonlinear oscillator", *Phys. Rev. A* **83**, 012308 (2011) [cit. on p. 88].

[Gam2013]   J. GAMBETTA, "Control of Superconducting Qubits", *Quantum Information Processing: Lecture Notes of the 44th IFF Spring School, Schriften des Forschungszentrums Jülich, Reihe Schlüsseltechnologien / Key Technologies*, Vol. **52**, ed. by D. DiVincenzo (2013) [cit. on pp. 29, 68, 69, 80, 82, 86, 88, 101].

[Gam2017]   J. GAMBETTA, J. CHOW, AND M. STEFFEN, "Building logical qubits in a superconducting quantum computing system", *npj Quantum Inf.* **3**, 2 (2017) [cit. on pp. 116, 155].

[Gar2004] D. GARANIN AND R. SCHILLING, "Quantum nonlinear spin switching model", *Phys. Rev. B* **69**, 104412 (2004) [cit. on p. 44].

[Geb2005] N. GEBHARDT, T. ALTEN, A. CELIS, AND J. GOEWERT, `http://irrlicht.sourceforge.net/`, 2005 [cit. on pp. 42, 173].

[Geh2014] G. GEHÉR, "An elementary proof for the non-bijective version of Wigner's theorem", *Phys. Lett. A* **378**, 2054 (2014) [cit. on pp. 179, 181].

[Geh2017] G. GEHÉR, "Wigner's theorem on Grassmann spaces", *J. Funct. Anal.* **273**, 2994 (2017) [cit. on p. 181].

[Ger2005] C. GERRY AND P. KNIGHT, *Introductory Quantum Optics*, Cambridge University Press, 2005 [cit. on p. 70].

[Gil2005] A. GILCHRIST, N. LANGFORD, AND M. NIELSEN, "Distance measures to compare real and ideal quantum processes", *Phys. Rev. A* **71**, 062310 (2005) [cit. on pp. 112, 191].

[Gil2018] A. GILYÉN, S. LLOYD, AND E. TANG, "Quantum-inspired low-rank stochastic regression with logarithmic dependence on the dimension", arXiv: 1811.04909 (2018) [cit. on p. 5].

[Gir2014] S. GIRVIN, "Circuit QED: superconducting qubits coupled to microwave photons", in: *Quantum Machines: Measurement and Control of Engineered Quantum Systems: Lecture Notes of the Les Houches Summer School*, Oxford University Press, 2014 [cit. on p. 153].

[Gla1963] R. GLAUBER, "Coherent and Incoherent States of the Radiation Field", *Phys. Rev.* **131**, 2766 (1963) [cit. on p. 79].

[Goo2016] I. GOODFELLOW, Y. BENGIO, AND A. COURVILLE, *Deep Learning*, MIT Press, 2016 [cit. on pp. 107, 129].

[Gor1976] V. GORINI, A. KOSSAKOWSKI, AND E. SUDARSHAN, "Completely positive dynamical semigroups of N-level systems", *J. Math. Phys.* **17**, 821 (1976) [cit. on p. 74].

[Got1998a] D. GOTTESMAN, "The Heisenberg representation of quantum computers", *Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, 32, ed. by S. Corney, R. Delbourgo, and P. Jarvis (1998) [cit. on p. 18].

[Got1998b] D. GOTTESMAN, "Theory of fault-tolerant quantum computation", *Phys. Rev. A* **57**, 127 (1998) [cit. on p. 155].

[Got2016] D. GOTTESMAN, "Quantum fault tolerance in small experiments", arXiv: 1610.03507 (2016) [cit. on pp. 156, 158, 160, 169].

[Gov2020] L. GOVIA, G. RIBEILL, D. RISTÈ, M. WARE, AND H. KROVI, "Bootstrapping quantum process tomography via a perturbative ansatz", *Nat. Commun.* **11**, 1084 (2020) [cit. on p. 138].

[Gra1997] M. GRASSL, T. BETH, AND T. PELLIZZARI, "Codes for the quantum erasure channel", *Phys. Rev. A* **56**, 33 (1997) [cit. on p. 156].

[Gre2015]   D. GREENBAUM, "Introduction to Quantum Gate Set Tomography", arXiv: 1509.02921 (2015) [cit. on pp. 125, 129].

[Grø2010]   N. GRØNBECH-JENSEN, J. MARCHESE, M. CIRILLO, AND J. BLACK-BURN, "Tomography and Entanglement in Coupled Josephson Junction Qubits", *Phys. Rev. Lett.* **105**, 010501 (2010) [cit. on p. 23].

[Gro2012]   P. DE GROOT, S. ASHHAB, A. LUPAŞCU, L. DICARLO, F. NORI, C. HARMANS, AND J. MOOIJ, "Selective darkening of degenerate transitions for implementing quantum controlled-NOT gates", *New J. Phys.* **14**, 073038 (2012) [cit. on p. 89].

[Gus2003]   S. GUSTAFSON AND I. SIGAL, *Mathematical Concepts of Quantum Mechanics*, Springer-Verlag Berlin Heidelberg, 2003 [cit. on p. 181].

[Győ2004]   M. GYŐRY, "A new proof of Wigner's theorem", *Rep. Math. Phys.* **54**, 159 (2004) [cit. on p. 181].

[Ham2000]   A. HAMS AND H. DE RAEDT, "Fast algorithm for finding the eigenvalue distribution of very large matrices", *Phys. Rev. E* **62**, 4365 (2000) [cit. on p. 192].

[Har2001]   L. HARDY, "Quantum Theory From Five Reasonable Axioms", arXiv: quant-ph/0101012 (2001) [cit. on p. 182].

[Har2009]   A. HARROW, A. HASSIDIM, AND S. LLOYD, "Quantum Algorithm for Linear Systems of Equations", *Phys. Rev. Lett.* **103**, 150502 (2009) [cit. on p. 2].

[Har2010]   R. HARRIS, M. JOHNSON, T. LANTING, A. BERKLEY, J. JOHANSSON, P. BUNYK, E. TOLKACHEVA, E. LADIZINSKY, N. LADIZINSKY, T. OH, F. CIOATA, I. PERMINOV, P. SPEAR, C. ENDERUD, C. RICH, S. UCHAIKIN, M. THOM, E. CHAPPLE, J. WANG, B. WILSON, M. AMIN, N. DICKSON, K. KARIMI, B. MACREADY, C. TRUNCIK, AND G. ROSE, "Experimental investigation of an eight-qubit unit cell in a superconducting optimization processor", *Phys. Rev. B* **82**, 024511 (2010) [cit. on p. 2].

[Har2019]   R. HARPER AND S. FLAMMIA, "Fault-Tolerant Logical Gates in the IBM Quantum Experience", *Phys. Rev. Lett.* **122**, 080504 (2019) [cit. on p. 156].

[Hel2019]   J. HELSEN, F. BATTISTEL, AND B. TERHAL, "Spectral quantum tomography", *npj Quantum Inf.* **5**, 74 (2019) [cit. on p. 127].

[Hen2010]   J. HENDRICKX AND A. OLSHEVSKY, "Matrix p-Norms Are NP-Hard to Approximate If $p \neq 1, 2, \infty$", *SIAM J. Matrix Anal. Appl.* **31**, 2802 (2010) [cit. on p. 111].

[Hor1996]   M. HORODECKI, P. HORODECKI, AND R. HORODECKI, "Separability of mixed states: necessary and sufficient conditions", *Phys. Lett. A* **223**, 1 (1996) [cit. on p. 10].

[Hor1999]   M. HORODECKI, P. HORODECKI, AND R. HORODECKI, "General teleportation channel, singlet fraction, and quasidistillation", *Phys. Rev. A* **60**, 1888 (1999) [cit. on pp. 110, 111, 190].

[Huy1990]   J. Huyghebaert and H. De Raedt, "Product formula methods for time-dependent Schrödinger problems", *J. Phys. A: Math. Gen.* **23**, 5777 (1990) [cit. on pp. 33, 60, 62, 186].

[IBM2016]   IBM Q team, *Quantum Experience*, 2016, `https://quantum-computing.ibm.com/` [cit. on pp. 2–4, 13, 25, 27, 44–46, 89, 94, 109, 123, 141, 160, 167, 168, 176].

[IBM2018a]   IBM Q team, *IBM Q 16 Rueschlikon backend specification V1.1.0*, 2018, `https://github.com/Qiskit/ibmq-device-information/blob/master/backends/rueschlikon/V1/version_log.md` [cit. on pp. 48, 88, 141, 156, 161–164, 166, 169].

[IBM2018b]   IBM Q team, *IBM Q 5 Tenerife backend specification V1.1.0*, 2018, `https://github.com/Qiskit/ibmq-device-information/blob/master/backends/tenerife/V1/version_log.md` [cit. on pp. 40, 141, 142, 145, 147, 165, 169].

[Iva2018]   O. Ivakhnenko, S. Shevchenko, and F. Nori, "Simulating quantum dynamical phenomena using classical oscillators: Landau-Zener-Stückelberg-Majorana interferometry, latching modulation, and motional averaging", *Sci. Rep.* **8**, 12218 (2018) [cit. on pp. 23, 170].

[Iye2018]   P. Iyer and D. Poulin, "A small quantum computer is needed to optimize fault-tolerant protocols", *Quantum Sci. Technol.* **3**, 030504 (2018) [cit. on p. 125].

[Jac1999]   J. Jackson, *Classical electrodynamics*, Wiley, 1999 [cit. on p. 22].

[Jac2014]   K. Jacobs, *Quantum Measurement Theory and its Applications*, Cambridge University Press, 2014 [cit. on p. 153].

[Jay1963]   E. Jaynes and F. Cummings, "Comparison of quantum and semiclassical radiation theories with application to the beam maser", *Proc. IEEE* **51**, 89 (1963) [cit. on pp. 68, 70].

[Jay1996]   E. Jaynes, "Probability in Quantum Theory", *Complexity, Entropy And The Physics Of Information*, ed. by W. Zurek (1996) [cit. on p. 23].

[Jay2003]   E. Jaynes and G. Bretthorst, *Probability Theory: The Logic of Science*, Cambridge University Press, 2003 [cit. on pp. 1, 182].

[Jin2013]   F. Jin, K. Michielsen, M. Novotny, S. Miyashita, S. Yuan, and H. De Raedt, "Quantum decoherence scaling with bath size: Importance of dynamics, connectivity, and randomness", *Phys. Rev. A* **87**, 022117 (2013) [cit. on p. 74].

[Jin2020]   F. Jin, D. Willsch, M. Willsch, H. Lagemann, K. Michielsen, and H. De Raedt, "Random state technology", in preparation, 2020 [cit. on pp. 192, 193].

[Joh2009]   N. Johnston, D. Kribs, and V. Paulsen, "Computing Stabilized Norms for Quantum Operations via the Theory of Completely Bounded Maps", *Quantum Inf. Comput.* **9**, 0016 (2009) [cit. on pp. 112–114, 194, 195].

[Joh2011]   M. JOHNSON, M. AMIN, S. GILDERT, T. LANTING, F. HAMZE, N. DICK-
             SON, R. HARRIS, A. BERKLEY, J. JOHANSSON, P. BUNYK, E. CHAPPLE,
             C. ENDERUD, J. HILTON, K. KARIMI, E. LADIZINSKY, N. LADIZINSKY,
             T. OH, I. PERMINOV, C. RICH, M. THOM, E. TOLKACHEVA, C. TRUN-
             CIK, S. UCHAIKIN, J. WANG, B. WILSON, AND G. ROSE, "Quantum
             annealing with manufactured spins", *Nature* **473**, 194 (2011) [cit. on p. 2].

[Joh2012]   J. JOHANSSON, P. NATION, AND F. NORI, "QuTiP: An open-source
             Python framework for the dynamics of open quantum systems", *Comput.
             Phys. Commun.* **183**, 1760 (2012) [cit. on pp. 92, 144].

[Joh2013]   J. JOHANSSON, P. NATION, AND F. NORI, "QuTiP 2: A Python frame-
             work for the dynamics of open quantum systems", *Comput. Phys. Commun.*
             184, 1234 (2013) [cit. on pp. 92, 144].

[Jor1962]   T. JORDAN, M. PINSKY, AND E. SUDARSHAN, "Dynamical Mappings of
             Density Operators in Quantum Mechanics. II. Time Dependent Mappings",
             J. Math. Phys. **3**, 848 (1962) [cit. on p. 181].

[Jor2006]   T. JORDAN, "Assumptions that imply quantum dynamics is linear", *Phys.
             Rev. A* **73**, 022101 (2006) [cit. on p. 181].

[Jor2009]   T. JORDAN, "Why quantum dynamics is linear", *J. Phys: Conf. Ser.* **196**,
             012010 (2009) [cit. on p. 181].

[Jos1962]   B. JOSEPHSON, "Possible new effects in superconductive tunnelling", *Phys.
             Lett.* **1**, 251 (1962) [cit. on p. 22].

[Joz1994]   R. JOZSA, "Fidelity for Mixed Quantum States", *J. Mod. Opt.* **41**, 2315
             (1994) [cit. on p. 190].

[Joz2001]   R. JOZSA, "Quantum factoring, discrete logarithms, and the hidden sub-
             group problem", *Comput. Sci. Eng.* **3**, 34 (2001) [cit. on p. 16].

[Jül2018]   JÜLICH SUPERCOMPUTING CENTRE, "JURECA: Modular supercomputer
             at Jülich Supercomputing Centre", *J. of Large-Scale Res. Facil.* **4**, A132
             (2018) [cit. on pp. 3, 47, 60, 63, 66, 74, 83, 102, 103, 124, 161, 167].

[Jül2019]   JÜLICH SUPERCOMPUTING CENTRE, "JUWELS: Modular Tier-0/1 Super-
             computer at the Jülich Supercomputing Centre", *J. of Large-Scale Res.
             Facil.* **5**, A135 (2019) [cit. on pp. 3, 74, 161, 167].

[Kad1998]   T. KADOWAKI AND H. NISHIMORI, "Quantum annealing in the transverse
             Ising model", *Phys. Rev. E* **58**, 5355 (1998) [cit. on p. 2].

[Kad2016]   A. KADIN AND S. KAPLAN, "Proposed Experiments to Test the Founda-
             tions of Quantum Computing", vixra:1607.0105 (2016) [cit. on p. 23].

[Kel2014]   J. KELLY, R. BARENDS, B. CAMPBELL, Y. CHEN, Z. CHEN, B. CHIARO,
             A. DUNSWORTH, A. FOWLER, I.-C. HOI, E. JEFFREY, A. MEGRANT,
             J. MUTUS, C. NEILL, P. O'MALLEY, C. QUINTANA, P. ROUSHAN, D.
             SANK, A. VAINSENCHER, J. WENNER, T. WHITE, A. CLELAND, AND J.
             MARTINIS, "Optimal Quantum Control Using Randomized Benchmarking",
             Phys. Rev. Lett. **112**, 240504 (2014) [cit. on p. 116].

[Ker2016]   I. KERENIDIS AND A. PRAKASH, "Quantum Recommendation Systems", arXiv:1603.08675 (2016) [cit. on p. 5].

[Kho2009]   K. KHODJASTEH AND L. VIOLA, "Dynamical quantum error correction of unitary operations with bounded controls", *Phys. Rev. A* **80**, 032314 (2009) [cit. on p. 89].

[Khr2010]   A. KHRENNIKOV, *Ubiquitous Quantum Structure: From Psychology to Finance*, Springer-Verlag Berlin Heidelberg, 2010 [cit. on p. 1].

[Kit1997]   A. KITAEV, "Quantum computations: algorithms and error correction", *Russ. Math. Surveys* **52**, 1191 (1997) [cit. on pp. 109, 111, 112, 168].

[Kiv1990]   S. KIVELSON AND D. ROKHSAR, "Bogoliubov quasiparticles, spinons, and spin-charge decoupling in superconductors", *Phys. Rev. B* **41**, 11693 (1990) [cit. on p. 79].

[Kni2008]   E. KNILL, D. LEIBFRIED, R. REICHLE, J. BRITTON, R. BLAKESTAD, J. JOST, C. LANGER, R. OZERI, S. SEIDELIN, AND D. WINELAND, "Randomized benchmarking of quantum gates", *Phys. Rev. A* **77**, 012307 (2008) [cit. on p. 111].

[Koc2007]   J. KOCH, T. YU, J. GAMBETTA, A. HOUCK, D. SCHUSTER, J. MAJER, A. BLAIS, M. DEVORET, S. GIRVIN, AND R. SCHOELKOPF, "Charge-insensitive qubit design derived from the Cooper pair box", *Phys. Rev. A* **76**, 042319 (2007) [cit. on pp. 3, 25, 27, 29, 41, 68, 69, 74].

[Kos1972]   A. KOSSAKOWSKI, "On quantum statistical mechanics of non-Hamiltonian systems", *Rep. Math. Phys.* **3**, 247 (1972) [cit. on p. 75].

[Kra1971]   K. KRAUS, "General state changes in quantum theory", *Ann. Phys. (N.Y.)* **64**, 311 (1971) [cit. on p. 17].

[Ku2020]   J. KU, X. XU, M. BRINK, D. MCKAY, J. HERTZBERG, M. ANSARI, AND B. PLOURDE, "Suppression of Unwanted ZZ Interactions in a Hybrid Two-Qubit System", arXiv:2003.02775 (2020) [cit. on p. 80].

[Kue2016]   R. KUENG, D. LONG, A. DOHERTY, AND S. FLAMMIA, "Comparing Experiments to the Fault-Tolerance Threshold", *Phys. Rev. Lett.* **117**, 170502 (2016) [cit. on pp. 115, 196].

[Küh2019]   M. KÜHN, S. ZANKER, P. DEGLMANN, M. MARTHALER, AND H. WEISS, "Accuracy and Resource Estimations for Quantum Chemistry on a Near-Term Quantum Computer", *J. Chem. Theory Comput.* **15**, 4764 (2019) [cit. on p. 2].

[Lan1969]   A. LANDÉ, "Quantum Fact and Fiction III", *Am. J. Phys.* **37**, 541 (1969) [cit. on p. 181].

[LaR2019]   R. LAROSE, "Overview and Comparison of Gate Level Quantum Software Platforms", *Quantum* **3**, 130 (2019) [cit. on p. 2].

[Lei1996]   D. LEIBFRIED, D. MEEKHOF, B. KING, C. MONROE, W. ITANO, AND
             D. WINELAND, "Experimental Determination of the Motional Quantum
             State of a Trapped Atom", *Phys. Rev. Lett.* **77**, 4281 (1996) [cit. on p. 127].

[Leu1997]   D. LEUNG, M. NIELSEN, I. CHUANG, AND Y. YAMAMOTO, "Approximate
             quantum error correction can lead to better codes", *Phys. Rev. A* **56**, 2567
             (1997) [cit. on p. 156].

[Li2008]   J. LI, K. CHALAPAT, AND G. PARAOANU, "Entanglement of supercon-
             ducting qubits via microwave fields: Classical and quantum regimes", *Phys.
             Rev. B* **78**, 064503 (2008) [cit. on p. 80].

[Lie1888]   S. LIE AND F. ENGEL, *Theorie der Transformationsgruppen*, Teubner, 1888
             [cit. on p. 33].

[Lin1976]   G. LINDBLAD, "On the generators of quantum dynamical semigroups",
             Commun. Math. Phys. **48**, 119 (1976) [cit. on p. 74].

[Lin2017]   N. LINKE, M. GUTIERREZ, K. LANDSMAN, C. FIGGATT, S. DEBNATH,
             K. BROWN, AND C. MONROE, "Fault-tolerant quantum error detection",
             Sci. Adv. **3**, e1701074 (2017) [cit. on p. 156].

[Lin2019]   J. LIN, B. BUONACORSI, R. LAFLAMME, AND J. WALLMAN, "On the
             freedom in representing quantum operations", *New J. Phys.* **21**, 023006
             (2019) [cit. on pp. 111, 129].

[Lom1963]   J. LOMONT AND P. MENDELSON, "The Wigner Unitarity-Antiunitarity
             Theorem", *Ann. Math.* **78**, 548 (1963) [cit. on p. 178].

[Los1992]   D. LOSS AND K. MULLEN, "Commutation relations for periodic opera-
             tors", *J. Phys. A: Math. Gen.* **25**, L235 (1992) [cit. on p. 25].

[Lue2004]   M. LUERSEN, R. LE RICHE, AND F. GUYON, "A constrained, globalized,
             and bounded Nelder–Mead method for engineering optimization", *Struct.
             Multidiscip. Optim.* **27**, 43 (2004) [cit. on p. 99].

[Lui1999]   A. LUIS AND L. SÁNCHEZ-SOTO, "Complete Characterization of Arbitrary
             Quantum Measurement Processes", *Phys. Rev. Lett.* **83**, 3573 (1999) [cit.
             on p. 127].

[Mag2012]   E. MAGESAN, J. GAMBETTA, AND J. EMERSON, "Characterizing quan-
             tum gates via randomized benchmarking", *Phys. Rev. A* **85**, 042311 (2012)
             [cit. on p. 111].

[Mag2020]   E. MAGESAN AND J. GAMBETTA, "Effective Hamiltonian models of the
             cross-resonance gate", *Phys. Rev. A* **101**, 052308 (2020) [cit. on pp. 82, 89,
             96].

[Maj2007]   J. MAJER, J. CHOW, J. GAMBETTA, J. KOCH, B. JOHNSON, J. SCHREIER,
             L. FRUNZIO, D. SCHUSTER, A. HOUCK, A. WALLRAFF, A. BLAIS, M.
             DEVORET, S. GIRVIN, AND R. SCHOELKOPF, "Coupling superconducting
             qubits via a cavity bus", *Nature* **449**, 443 (2007) [cit. on p. 80].

[Mal2020]   M. Malekakhlagh, E. Magesan, and D. McKay, "First-principles analysis of cross-resonance gate operation", arXiv:2005.00133 (2020) [cit. on pp. 89, 96].

[Mar1987]   J. Martinis, M. Devoret, and J. Clarke, "Experimental tests for the quantum behavior of a macroscopic degree of freedom: The phase difference across a Josephson junction", *Phys. Rev. B* **35**, 4682 (1987) [cit. on p. 22].

[Mar2007]   J. Marchese, M. Cirillo, and N. Grønbech-Jensen, "Classical analogs for Rabi-oscillations, Ramsey-fringes, and spin-echo in Josephson junctions", *Eur. Phys. J. Spec. Top.* **147**, 333 (2007) [cit. on p. 23].

[Mar2018]   I. Markov, A. Fatima, S. Isakov, and S. Boixo, "Quantum Supremacy Is Both Closer and Farther than It Appears", arXiv:1807.10749 (2018) [cit. on p. 176].

[McK2017]   D. McKay, C. Wood, S. Sheldon, J. Chow, and J. Gambetta, "Efficient Z gates for quantum computing", *Phys. Rev. A* **96**, 022330 (2017) [cit. on pp. 85, 87, 117].

[McK2018]   D. McKay, T. Alexander, L. Bello, M. Biercuk, L. Bishop, J. Chen, J. Chow, A. Córcoles, D. Egger, S. Filipp, J. Gomez, M. Hush, A. Javadi-Abhari, D. Moreda, P. Nation, B. Paulovicks, E. Winston, C. Wood, J. Wootton, and J. Gambetta, "Qiskit Backend Specifications for OpenQASM and OpenPulse Experiments", arXiv: 1809.03452 (2018) [cit. on pp. 107, 169].

[McK2019]   D. McKay, S. Sheldon, J. Smolin, J. Chow, and J. Gambetta, "Three-Qubit Randomized Benchmarking", *Phys. Rev. Lett.* **122**, 200502 (2019) [cit. on p. 168].

[Mey2009a] K. Meyer, G. Hall, and D. Offin, "Symplectic Transformations", in: *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, Springer New York, 2009 [cit. on p. 57].

[Mey2009b] K. Meyer, G. Hall, and D. Offin, "Topics in Linear Theory", in: *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, Springer New York, 2009 [cit. on p. 199].

[Mic2005]   K. Michielsen, K. De Raedt, and H. De Raedt, "Simulation of Quantum Computation: A Deterministic Event-Based Approach", *J. Comput. Theor. Nanosci.* **2**, 227 (2005) [cit. on p. 23].

[Mic2017]   K. Michielsen, M. Nocon, D. Willsch, F. Jin, Th. Lippert, and H. De Raedt, "Benchmarking gate-based quantum computers", *Comput. Phys. Commun.* **220**, 44 (2017) [cit. on pp. 141, 149, 150, 152, 155, 166].

[Mit2019]   S. Mittal, "A survey of techniques for dynamic branch prediction", *Concurrency and Computation: Practice and Experience* **31**, e4666 (2019) [cit. on p. 36].

[Mni2015]   V. Mnih, K. Kavukcuoglu, D. Silver, A. Rusu, J. Veness, M. Bellemare, A. Graves, M. Riedmiller, A. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning", *Nature* **518**, 529 (2015) [cit. on p. 107].

[Mol1998]   L. Molnár, "An algebraic approach to Wigner's unitary-antiunitary theorem", *J. Austral. Math. Soc.* **65**, 354 (1998) [cit. on p. 181].

[Mon1995]   C. Monroe, D. Meekhof, B. King, W. Itano, and D. Wineland, "Demonstration of a Fundamental Quantum Logic Gate", *Phys. Rev. Lett.* **75**, 4714 (1995) [cit. on p. 2].

[Mor2011]   J. Morales and J. Nocedal, "Remark on "Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound Constrained Optimization"", *ACM Trans. Math. Softw.* **38**, 7 (2011) [cit. on pp. 99, 201].

[Mot2009]   F. Motzoi, J. Gambetta, P. Rebentrost, and F. Wilhelm, "Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits", *Phys. Rev. Lett.* **103**, 110501 (2009) [cit. on pp. 88, 117].

[Mou2013]   A. Mouchet, "An alternative proof of Wigner theorem on quantum transformations based on elementary complex analysis", *Phys. Lett. A* **377**, 2709 (2013) [cit. on p. 181].

[Mul1959]   M. Muller, "A Note on a Method for Generating Points Uniformly on N-dimensional Spheres", *Commun. ACM* **2**, 19 (1959) [cit. on p. 193].

[Mül2009]   C. Müller, A. Shnirman, and Y. Makhlin, "Relaxation of Josephson qubits due to strong coupling to two-level systems", *Phys. Rev. B* **80**, 134517 (2009) [cit. on p. 79].

[Mül2019]   C. Müller, J. Cole, and J. Lisenfeld, "Towards understanding two-level-systems in amorphous solids: insights from quantum circuits", *Rep. Prog. Phys.* **82**, 124501 (2019) [cit. on p. 79].

[Nat2019]   National Academies of Sciences, Engineering, and Medicine, *Quantum Computing: Progress and Prospects*, The National Academies Press, 2019 [cit. on pp. 2, 25].

[Nau2018]   H. Naus and R. Versluis, "Consequences of unitary evolution of coupled qubit-resonator systems for stabilizing circuits in surface codes", arXiv: 1811.09832 (2018) [cit. on p. 165].

[Nei2018]   C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, R. Barends, B. Burkett, Y. Chen, Z. Chen, A. Fowler, B. Foxen, M. Giustina, R. Graff, E. Jeffrey, T. Huang, J. Kelly, P. Klimov, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. White, H. Neven, and J. Martinis,

"A blueprint for demonstrating quantum supremacy with superconducting qubits", *Science* **360**, 195 (2018) [cit. on pp. 27, 155].

[Nel1965]   J. NELDER AND R. MEAD, "A Simplex Method for Function Minimization", *Comput. J.* **7**, 308 (1965) [cit. on p. 99].

[Neu1955]   J. VON NEUMANN, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955 [cit. on pp. 1, 181].

[Ng2009]   H. NG AND J. PRESKILL, "Fault-tolerant quantum computation versus Gaussian noise", *Phys. Rev. A* **79**, 032318 (2009) [cit. on pp. 111, 155].

[Nie2002]   M. NIELSEN, "A simple formula for the average gate fidelity of a quantum dynamical operation", *Phys. Lett. A* **303**, 249 (2002) [cit. on pp. 109–111, 168, 191].

[Nie2010]   M. NIELSEN AND I. CHUANG, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, Cambridge University Press, 2010 [cit. on pp. 1, 2, 5, 15, 16, 19, 20, 127, 148, 152, 156].

[Nie2017]   E. NIELSEN, K. RUDINGER, R. BLUME-KOHOUT, A. BESTWICK, B. BLOOM, M. BLOCK, S. CALDWELL, M. CURTIS, A. HUDSON, J.-L. ORGIAZZI, A. PAPAGEORGE, A. POLLORENO, M. REAGOR, N. RUBIN, M. SCHEER, M. SELVANAYAGAM, E. SETE, R. SINCLAIR, R. SMITH, M. VAHIDPOUR, M. VILLIERS, W. ZENG, AND C. RIGETTI, "Efficient gate set tomography on a multi-qubit superconducting processor", *APS Meeting Abstracts*, L46.008 (2017) [cit. on p. 138].

[Nie2018]   E. NIELSEN, R. BLUME-KOHOUT, K. RUDINGER, T. PROCTOR, L. SALDYT, J. GROSS, T. SCHOLTEN, AND D. NADLINGER, *pyGSTio/pyGSTi: Version 0.9.4.4*, 2018, https://doi.org/10.5281/zenodo.1209246 [cit. on pp. 126, 130, 131, 198, 200].

[Nig2012]   S. NIGG, H. PAIK, B. VLASTAKIS, G. KIRCHMAIR, S. SHANKAR, L. FRUNZIO, M. DEVORET, R. SCHOELKOPF, AND S. GIRVIN, "Black-Box Superconducting Circuit Quantization", *Phys. Rev. Lett.* **108**, 240502 (2012) [cit. on pp. 42, 48, 50–52, 54, 169].

[Niu2019]   M. NIU, S. BOIXO, V. SMELYANSKIY, AND H. NEVEN, "Universal quantum control through deep reinforcement learning", *npj Quantum Inf.* **5**, 33 (2019) [cit. on pp. 107, 169].

[Ope2018]   OPENSUPERQ, 2018, https://opensuperq.eu/ [cit. on p. 170].

[Orú2019]   R. ORÚS, S. MUGEL, AND E. LIZASO, "Quantum computing for finance: Overview and prospects", *Rev. Phys.* **4**, 100028 (2019) [cit. on p. 2].

[Pai2011]   H. PAIK, D. SCHUSTER, L. BISHOP, G. KIRCHMAIR, G. CATELANI, A. SEARS, B. JOHNSON, M. REAGOR, L. FRUNZIO, L. GLAZMAN, S. GIRVIN, M. DEVORET, AND R. SCHOELKOPF, "Observation of High Coherence in Josephson Junction Qubits Measured in a Three-Dimensional Circuit QED Architecture", *Phys. Rev. Lett.* **107**, 240501 (2011) [cit. on p. 25].

[Pal2017] P. Palittapongarnpim, P. Wittek, E. Zahedinejad, S. Vedaie, and B. Sanders, "Learning in Quantum Control: High-Dimensional Global Optimization for Noisy Quantum Dynamics", *Neurocomputing* **268**, 116 (2017) [cit. on p. 107].

[Par2006] G. Paraoanu, "Microwave-induced coupling of superconducting qubits", *Phys. Rev. B* **74**, 140504 (2006) [cit. on p. 89].

[Pau2003] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge University Press, 2003 [cit. on p. 112].

[Ped2007] L. Pedersen, N. Møller, and K. Mølmer, "Fidelity of quantum operations", *Phys. Lett. A* **367**, 47 (2007) [cit. on pp. 111, 193].

[Ped2017] E. Pednault, J. Gunnels, G. Nannicini, L. Horesh, T. Magerlein, E. Solomonik, and R. Wisnieff, "Breaking the 49-Qubit Barrier in the Simulation of Quantum Circuits", arXiv:1710.05867 (2017) [cit. on p. 176].

[Per2019] A. Perdomo-Ortiz, A. Feldman, A. Ozaeta, S. Isakov, Z. Zhu, B. O'Gorman, H. Katzgraber, A. Diedrich, H. Neven, J. de Kleer, B. Lackey, and R. Biswas, "Readiness of Quantum Optimization Machines for Industrial Applications", *Phys. Rev. Applied* **12**, 014004 (2019) [cit. on p. 2].

[Pom2020] J. Pommerening and D. DiVincenzo, "What is measured when a qubit measurement is performed on a multi-qubit chip?", arXiv:2001.11756 (2020) [cit. on p. 69].

[Pop2014] I. Pop, K. Geerlings, G. Catelani, R. Schoelkopf, L. Glazman, and M. Devoret, "Coherent suppression of electromagnetic dissipation due to superconducting quasiparticles", *Nature* **508**, 369 (2014) [cit. on p. 79].

[Poy1997] J. Poyatos, J. Cirac, and P. Zoller, "Complete Characterization of a Quantum Process: The Two-Bit Quantum Gate", *Phys. Rev. Lett.* **78**, 390 (1997) [cit. on p. 127].

[Pre2007] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Cambridge University Press, 2007 [cit. on pp. 40, 41, 99–101].

[Pre2012] J. Preskill, "Quantum computing and the entanglement frontier", arXiv:1203.5813 (2012) [cit. on p. 2].

[Pre2018] J. Preskill, "Quantum Computing in the NISQ era and beyond", *Quantum* **2**, 79 (2018) [cit. on p. 2].

[Pro2017] T. Proctor, K. Rudinger, K. Young, M. Sarovar, and R. Blume-Kohout, "What Randomized Benchmarking Actually Measures", *Phys. Rev. Lett.* **119**, 130502 (2017) [cit. on pp. 111, 116, 125].

[Pud2012]   K. PUDENZ AND D. LIDAR, "Quantum adiabatic machine learning", *Quantum Inf. Process.* **12**, 2027 (2012) [cit. on p. 2].

[Qi2019]   J. QI AND H. NG, "Comparing the randomized benchmarking figure with the average infidelity of a quantum gate-set", *Int. J. Quantum Inf* **17**, 1950031 (2019) [cit. on p. 111].

[Rai2001]   J. RAIMOND, M. BRUNE, AND S. HAROCHE, "Manipulating quantum entanglement with atoms and photons in a cavity", *Rev. Mod. Phys.* **73**, 565 (2001) [cit. on p. 2].

[Ran1973]   B. RANDELL, *The Origins of Digital Computers: Selected Papers*, Springer-Verlag Berlin Heidelberg, 1973 [cit. on pp. 1, 204, 229].

[Rea2018]   M. REAGOR, C. OSBORN, N. TEZAK, A. STALEY, G. PRAWIROAT-MODJO, M. SCHEER, N. ALIDOUST, E. SETE, N. DIDIER, M. DA SILVA, E. ACALA, J. ANGELES, A. BESTWICK, M. BLOCK, B. BLOOM, A. BRADLEY, C. BUI, S. CALDWELL, L. CAPELLUTO, R. CHILCOTT, J. CORDOVA, G. CROSSMAN, M. CURTIS, S. DESHPANDE, T. EL BOUAYADI, D. GIRSHOVICH, S. HONG, A. HUDSON, P. KARALEKAS, K. KUANG, M. LENIHAN, R. MANENTI, T. MANNING, J. MARSHALL, Y. MOHAN, W. O'BRIEN, J. OTTERBACH, A. PAPAGEORGE, J.-P. PAQUETTE, M. PELSTRING, A. POLLORENO, V. RAWAT, C. RYAN, R. RENZAS, N. RUBIN, D. RUSSEL, M. RUST, D. SCARABELLI, M. SELVANAYAGAM, R. SINCLAIR, R. SMITH, M. SUSKA, T.-W. TO, M. VAHIDPOUR, N. VODRAHALLI, T. WHYLAND, K. YADAV, W. ZENG, AND C. RIGETTI, "Demonstration of universal parametric entangling gates on a multi-qubit lattice", *Sci. Adv.* **4**, eaao3603 (2018) [cit. on p. 27].

[Ric2013]   S. RICHER, "Perturbative analysis of two-qubit gates on transmon qubits", Master's thesis, RWTH Aachen University, 2013, `https://www.quantuminfo.physik.rwth-aachen.de/global/show_document.asp?id=aaaaaaaaaajiobd` [cit. on pp. 69, 80, 131].

[Rie2019]   D. RIEGER AND I. POP, personal communication, 2019 [cit. on pp. 43, 75, 77, 79].

[Rig2009]   C. RIGETTI, "Quantum Gates for Superconducting Qubits", PhD thesis, Yale University, 2009 [cit. on pp. 11, 200].

[Rig2010]   C. RIGETTI AND M. DEVORET, "Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies", *Phys. Rev. B* **81**, 134507 (2010) [cit. on p. 89].

[Rig2012]   C. RIGETTI, J. GAMBETTA, S. POLETTO, B. PLOURDE, J. CHOW, A. CÓRCOLES, J. SMOLIN, S. MERKEL, J. ROZEN, G. KEEFE, M. ROTHWELL, M. KETCHEN, AND M. STEFFEN, "Superconducting qubit in a waveguide cavity with a coherence time approaching 0.1 ms", *Phys. Rev. B* **86**, 100506 (2012) [cit. on p. 25].

[Rig2017]   Rigetti Computing, 2017, https://www.rigetti.com/ [cit. on pp. 2, 3, 25, 167].

[Riv1978]   R. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public-key Cryptosystems", *Commun. ACM* **21**, 120 (1978) [cit. on p. 2].

[Rot2017]   M. Roth, M. Ganzhorn, N. Moll, S. Filipp, G. Salis, and S. Schmidt, "Analysis of a parametrically driven exchange-type gate and a two-photon excitation gate between superconducting qubits", *Phys. Rev. A* **96**, 062323 (2017) [cit. on p. 170].

[San2016]   Y. Sanders, J. Wallman, and B. Sanders, "Bounding quantum gate error rate based on reported average fidelity", *New J. Phys.* **18**, 012002 (2016) [cit. on pp. 111, 112, 115, 117, 123, 125, 197].

[Sch1926a]   E. Schrödinger, "Quantisierung als Eigenwertproblem", *Ann. Phys. (Berl.)* **384**, 361 (1926) [cit. on p. 178].

[Sch1926b]   E. Schrödinger, "Quantisierung als Eigenwertproblem", *Ann. Phys. (Berl.)* **384**, 489 (1926) [cit. on p. 178].

[Sch1966]   J. Schrieffer and P. Wolff, "Relation between the Anderson and Kondo Hamiltonians", *Phys. Rev.* **149**, 491 (1966) [cit. on p. 69].

[Sch2003]   R. Schack, "Quantum Theory from Four of Hardy's Axioms", *Found. Phys.* **33**, 1461 (2003) [cit. on p. 182].

[Sek2014]   Y. Sekino and S. Ishizaka, "Quantum-information division and an optimal uncorrelated channel", *Phys. Rev. A* **89**, 034304 (2014) [cit. on p. 111].

[She2016a]   S. Sheldon, L. Bishop, E. Magesan, S. Filipp, J. Chow, and J. Gambetta, "Characterizing errors on qubit operations via iterative randomized benchmarking", *Phys. Rev. A* **93**, 012301 (2016) [cit. on pp. 116, 155].

[She2016b]   S. Sheldon, E. Magesan, J. Chow, and J. Gambetta, "Procedure for systematically tuning up cross-talk in the cross-resonance gate", *Phys. Rev. A* **93**, 060302 (2016) [cit. on pp. 89, 96, 98, 116, 133].

[Sho1994]   P. Shor, "Algorithms for quantum computation: discrete logarithms and factoring", *Proceedings 35th Annual Symposium on Foundations of Computer Science*, 124 (1994) [cit. on pp. 2, 16].

[Sho1996]   P. Shor, "Fault-tolerant quantum computation", *Proceedings of 37th Conference on Foundations of Computer Science*, 56 (1996) [cit. on pp. 2, 155].

[Sho1997]   P. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer", *SIAM J. Comput.* **26**, 1484 (1997) [cit. on p. 2].

[Sil2016]   D. Silver, A. Huang, C. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of Go with deep neural networks and tree search", *Nature* **529**, 484 (2016) [cit. on p. 107].

[Sim2008]   R. Simon, N. Mukunda, S. Chaturvedi, and V. Srinivasan, "Two elementary proofs of the Wigner theorem on symmetry in quantum mechanics", *Phys. Lett. A* **372**, 6847 (2008) [cit. on p. 181].

[Smi1981]   J. Smith, "A Study of Branch Prediction Strategies", *Proceedings of the 8th Annual Symposium on Computer Architecture, ISCA '81*, 135 (1981) [cit. on p. 36].

[Smi1993]   D. Smithey, M. Beck, M. Raymer, and A. Faridani, "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum", *Phys. Rev. Lett.* **70**, 1244 (1993) [cit. on p. 127].

[Sol2014]   F. Solgun, D. Abraham, and D. DiVincenzo, "Blackbox quantization of superconducting circuits using exact impedance synthesis", *Phys. Rev. B* **90**, 134504 (2014) [cit. on pp. 48, 50].

[Sol2015]   F. Solgun, "Analysis and Synthesis of Multi-Qubit, Multi-Mode Quantum Devices", PhD thesis, RWTH Aachen University, 2015 [cit. on pp. 48, 50].

[Sol2016]   F. Solgun and J. Gambetta, personal communication, 2016 [cit. on pp. 44, 147].

[Sol2019]   F. Solgun, D. DiVincenzo, and J. Gambetta, "Simple Impedance Response Formulas for the Dispersive Interaction Rates in the Effective Hamiltonians of Low Anharmonicity Superconducting Qubits", *IEEE Trans. Microwave Theory Tech.* **67**, 928 (2019) [cit. on pp. 48, 80].

[Son2019]   C. Song, J. Cui, H. Wang, J. Hao, H. Feng, and Y. Li, "Quantum computation with universal error mitigation on a superconducting quantum processor", *Sci. Adv.* **5**, (2019) [cit. on p. 138].

[Spe2012]   C. Spengler, M. Huber, and B. Hiesmayr, "Composite parameterization and Haar measure for all unitary and special unitary groups", *J. Math. Phys.* **53**, 013501 (2012) [cit. on p. 191].

[Sti1955]   W. Stinespring, "Positive Functions on C*-Algebras", *Proc. Amer. Math. Soc.* **6**, 211 (1955) [cit. on p. 17].

[Sur2015]   B. Suri, Z. Keane, L. Bishop, S. Novikov, F. Wellstood, and B. Palmer, "Nonlinear microwave photon occupancy of a driven resonator strongly coupled to a transmon qubit", *Phys. Rev. A* **92**, 063801 (2015) [cit. on pp. 72, 75, 79].

[Suz1976]   M. Suzuki, "Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems", *Commun. Math. Phys.* **51**, 83 (1976) [cit. on p. 33].

[Suz1985]   M. Suzuki, "Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics", *J. Math. Phys.* **26**, 601 (1985) [cit. on p. 33].

[Tak2016]   M. Takita, A. Córcoles, E. Magesan, B. Abdo, M. Brink, A. Cross, J. Chow, and J. Gambetta, "Demonstration of Weight-Four Parity Measurements in the Surface Code Architecture", *Phys. Rev. Lett.* **117**, 210505 (2016) [cit. on p. 95].

[Tak2017]   M. Takita, A. Cross, A. Córcoles, J. Chow, and J. Gambetta, "Experimental Demonstration of Fault-Tolerant State Preparation with Superconducting Qubits", *Phys. Rev. Lett.* **119**, 180501 (2017) [cit. on pp. 44, 89, 94, 95, 118, 122, 125, 156].

[Tan2018]   E. Tang, "Quantum-inspired classical algorithms for principal component analysis and supervised clustering", arXiv:1811.00414 (2018) [cit. on p. 5].

[Tan2019]   E. Tang, "A Quantum-Inspired Classical Algorithm for Recommendation Systems", *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019*, 217 (2019) [cit. on p. 5].

[Ter2000]   B. Terhal, "Bell inequalities and the separability criterion", *Phys. Lett. A* **271**, 319 (2000) [cit. on p. 10].

[Ter2005]   B. Terhal and G. Burkard, "Fault-tolerant quantum computation for local non-Markovian noise", *Phys. Rev. A* **71**, 012336 (2005) [cit. on pp. 111, 155].

[The2018]   L. Theis, F. Motzoi, S. Machnes, and F. Wilhelm, "Counteracting systems of diabaticities using DRAG controls: The status after 10 years(a)", *EPL* **123**, 60001 (2018) [cit. on pp. 88, 167].

[Tri1969]   M. Tribus, T. Irvine, and J. Hartnett, *Rational Descriptions, Decisions and Designs: Pergamon Unified Engineering Series*, Pergamon Press, 1969 [cit. on p. 182].

[Tri2019]   V. Tripathi, M. Khezri, and A. Korotkov, "Operation and intrinsic error budget of a two-qubit cross-resonance gate", *Phys. Rev. A* **100**, 012301 (2019) [cit. on p. 89].

[Tro1959]   H. Trotter, "On the Product of Semi-Groups of Operators", *Proc. Amer. Math. Soc.* **10**, 545 (1959) [cit. on p. 33].

[Tur1937]   A. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem", *Proc. London Math. Soc. Ser. 2* **42**, 230 (1937) [cit. on p. 1].

[Vai1996]   L. Vaidman, L. Goldenberg, and S. Wiesner, "Error prevention scheme with four particles", *Phys. Rev. A* **54**, R1745 (1996) [cit. on p. 156].

[Val1958]   J. VALATIN, "Comments on the theory of superconductivity", *Nuovo Cim* **7**, 843 (1958) [cit. on p. 56].

[Vil2020]   B. VILLALONGA, D. LYAKH, S. BOIXO, H. NEVEN, T. HUMBLE, R. BISWAS, E. RIEFFEL, A. HO, AND S. MANDRÀ, "Establishing the quantum supremacy frontier with a 281 Pflop/s simulation", *Quantum Sci. Technol.* **5**, 034003 (2020) [cit. on p. 176].

[Vio2002]   D. VION, A. AASSIME, A. COTTET, P. JOYEZ, H. POTHIER, C. URBINA, D. ESTEVE, AND M. DEVORET, "Manipulating the Quantum State of an Electrical Circuit", *Science* **296**, 886 (2002) [cit. on pp. 2, 25].

[Voo2017]   U. VOOL AND M. DEVORET, "Introduction to quantum electromagnetic circuits", *Int. J. Circ. Theor. Appl.* **45**, 897 (2017) [cit. on pp. 22, 23, 25].

[Vui2018]   C. VUILLOT, "Is error detection helpful on IBM 5Q chips?", *Quantum Inf. Comput.* **18**, 0949 (2018) [cit. on p. 156].

[Wal2014]   J. WALLMAN AND S. FLAMMIA, "Randomized benchmarking with confidence", *New J. Phys.* **16**, 103032 (2014) [cit. on pp. 111, 115, 196, 197].

[Wal2015a] J. WALLMAN, "Bounding experimental quantum error rates relative to fault-tolerant thresholds", arXiv:1511.00727 (2015) [cit. on p. 115].

[Wal2015b] J. WALLMAN, C. GRANADE, R. HARPER, AND S. FLAMMIA, "Estimating the coherence of noise", *New J. Phys.* **17**, 113020 (2015) [cit. on pp. 109, 116].

[Wal2016]   J. WALLMAN, M. BARNHILL, AND J. EMERSON, "Robust characterization of leakage errors", *New J. Phys.* **18**, 043021 (2016) [cit. on p. 116].

[Wat2009]   J. WATROUS, "Semidefinite Programs for Completely Bounded Norms", *Theory Comput.* **5**, 217 (2009) [cit. on p. 112].

[Wat2013]   J. WATROUS, "Simpler semidefinite programs for completely bounded norms", *Chicago J. Theoret. Comput. Sci.* **2013**, 8 (2013) [cit. on p. 112].

[Wat2018]   J. WATROUS, *The Theory of Quantum Information*, Cambridge University Press, 2018 [cit. on pp. 5, 111–113, 186].

[Wen2017]   G. WENDIN, "Quantum information processing with superconducting circuits: a review", *Rep. Prog. Phys.* **80**, 106001 (2017) [cit. on pp. 2, 21].

[Wie2014]   C.-R. WIE, "Bloch sphere model for two-qubit pure states", arXiv:1403.8069 (2014) [cit. on p. 11].

[Wig1931]   E. WIGNER, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*, Vieweg+Teubner Verlag, 1931 [cit. on p. 178].

[Wig1959]   E. WIGNER, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*, Academic Press, 1959 [cit. on p. 178].

[Wil2016]   D. WILLSCH, "Simulation of quantum computer hardware based on superconducting circuits", Master's thesis, RWTH Aachen University, 2016, `http://hdl.handle.net/2128/21812` [cit. on pp. 25, 28, 31, 33, 35, 42, 68, 86, 88, 95, 98, 131, 133, 201].

[Wil2017]   D. WILLSCH, M. NOCON, F. JIN, H. DE RAEDT, AND K. MICHIELSEN, "Gate-error analysis in simulations of quantum computers with transmon qubits", *Phys. Rev. A* **96**, 062302 (2017) [cit. on pp. 4, 33, 39, 44, 80, 85, 90, 99, 109, 141, 149, 166, 168, 233].

[Wil2018a]   D. WILLSCH AND M. WILLSCH, *A modular implementation of the Nelder–Mead optimization method*, 2018, https://github.com/dixr/neldermead/ [cit. on p. 100].

[Wil2018b]   D. WILLSCH, M. WILLSCH, F. JIN, H. DE RAEDT, AND K. MICHIELSEN, "Testing quantum fault tolerance on small systems", *Phys. Rev. A* **98**, 052348 (2018) [cit. on pp. 4, 33, 47, 120, 141, 156, 161, 165, 169, 233].

[Wil2020a]   D. WILLSCH, H. LAGEMANN, M. WILLSCH, F. JIN, H. DE RAEDT, AND K. MICHIELSEN, "Benchmarking Supercomputers with the Jülich Universal Quantum Computer Simulator", *NIC Symposium 2020, Publication Series of the John von Neumann Institute for Computing (NIC) NIC Series* **50**, 255, ed. by M. Müller, K. Binder, and A. Trautmann (2020) [cit. on pp. 3, 154, 173, 175, 176].

[Wil2020b]   D. WILLSCH, M. WILLSCH, H. DE RAEDT, AND K. MICHIELSEN, "Support vector machines on the D-Wave quantum annealer", *Comput. Phys. Commun.* **248**, 107006 (2020) [cit. on p. 2].

[Wil2020c]   M. WILLSCH, D. WILLSCH, K. MICHIELSEN, F. JIN, T. DENKMAYR, S. SPONAR, Y. HASEGAWA, AND H. DE RAEDT, "Long-Time Correlations in Single-Neutron Interferometry Data", *J. Phys. Soc. Jpn.* **89**, 064005 (2020) [cit. on p. 128].

[Wil2020d]   M. WILLSCH, "Study of quantum annealing by simulating the time evolution of flux qubits", PhD thesis, RWTH Aachen University, 2020 [cit. on pp. 33, 34, 38, 62, 79, 83, 131, 167, 169, 186].

[Wil2020e]   M. WILLSCH, D. WILLSCH, F. JIN, H. DE RAEDT, AND K. MICHIELSEN, "Real-time simulation of flux qubits used for quantum annealing", *Phys. Rev. A* **101**, 012327 (2020) [cit. on pp. 33, 79, 169].

[Wil2020f]   M. WILLSCH, D. WILLSCH, K. MICHIELSEN, AND H. DE RAEDT, "Discrete-Event Simulation of Quantum Walks", *Front. Phys.* **8**, 145 (2020) [cit. on p. 10].

[Wol2019]   WOLFRAM RESEARCH, INC, *Mathematica, Version 12.0*, 2019, http://reference.wolfram.com/language/ [cit. on p. 200].

[Woo2015]   C. WOOD, J. BIAMONTE, AND D. CORY, "Tensor networks and graphical calculus for open quantum systems", *Quantum Inf. Comput.* **15**, 0579 (2015) [cit. on p. 18].

[Woo2018]   C. WOOD AND J. GAMBETTA, "Quantification and characterization of leakage errors", *Phys. Rev. A* **97**, 032306 (2018) [cit. on pp. 39, 85, 167].

[Xin2017]   T. Xin, D. Lu, J. Klassen, N. Yu, Z. Ji, J. Chen, X. Ma, G. Long, B. Zeng, and R. Laflamme, "Quantum State Tomography via Reduced Density Matrices", *Phys. Rev. Lett.* **118**, 020401 (2017) [cit. on p. 127].

[Yan2008]   N. Yanofsky and M. Mannucci, *Quantum Computing for Computer Scientists*, Cambridge University Press, 2008 [cit. on p. 5].

[Zar2006]   V. Zarikian, "Alternating-projection algorithms for operator-theoretic calculations", *Linear Algebra Appl.* **419**, 710 (2006) [cit. on p. 112].

[Zha2016]   P. Zhao, H. De Raedt, S. Miyashita, F. Jin, and K. Michielsen, "Dynamics of open quantum spin systems: An assessment of the quantum master equation approach", *Phys. Rev. E* **94**, 022126 (2016) [cit. on p. 75].

[Zhu1997]   C. Zhu, R. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-scale Bound-constrained Optimization", *ACM Trans. Math. Softw.* **23**, 550 (1997) [cit. on pp. 99, 201].

[Zus1936]   K. Zuse, "Method for Automatic Execution of Calculations with the Aid of Computers", patent application (reprinted in full in [Ran1973]), 1936 [cit. on p. 1].

List of publications

(1) K. MICHIELSEN, M. NOCON, D. WILLSCH, F. JIN, TH. LIPPERT, H. DE RAEDT,
"Benchmarking gate-based quantum computers",
Comput. Phys. Commun. **220**, 44 (2017)

(2) D. WILLSCH, M. NOCON, F. JIN, H. DE RAEDT, K. MICHIELSEN,
"Gate-error analysis in simulations of quantum computers with transmon qubits",
Phys. Rev. A **96**, 062302 (2017)

(3) D. WILLSCH, M. WILLSCH, F. JIN, H. DE RAEDT, K. MICHIELSEN,
"Testing quantum fault tolerance on small systems",
Phys. Rev. A **98**, 052348 (2018)

(4) H. DE RAEDT, F. JIN, D. WILLSCH, M. WILLSCH, N. YOSHIOKA, N. ITO,
S. YUAN, K. MICHIELSEN,
"Massively parallel quantum computer simulator, eleven years later",
Comput. Phys. Commun. **237**, 47 (2019)

(5) H. DE RAEDT, M. KATSNELSON, D. WILLSCH, K. MICHIELSEN,
"Separation of conditions as a prerequisite for quantum theory",
Ann. Phys. (N. Y.) **403**, 112 (2019)

(6) M. WILLSCH, D. WILLSCH, F. JIN, H. DE RAEDT, K. MICHIELSEN,
"Real-time simulation of flux qubits used for quantum annealing",
Phys. Rev. A **101**, 012327 (2020)

(7) D. WILLSCH, H. LAGEMANN, M. WILLSCH, F. JIN, H. DE RAEDT, K. MICHIELSEN,
"Benchmarking Supercomputers with the Jülich Universal Quantum Computer Simulator",
NIC Symposium 2020, Publication Series of the John von Neumann Institute for Computing (NIC) NIC Series **50**, 255 (2020)

(8) D. WILLSCH, M. WILLSCH, H. DE RAEDT, K. MICHIELSEN,
"Support vector machines on the D-Wave quantum annealer",
Comput. Phys. Commun. **248**, 107006 (2020)

(9) M. WILLSCH, D. WILLSCH, K. MICHIELSEN, H. DE RAEDT,
"Discrete-Event Simulation of Quantum Walks",
Front. Phys. **8**, 145 (2020)

(10) H. DE RAEDT, M. JATTANA, D. WILLSCH, M. WILLSCH, F. JIN, K. MICHIELSEN,
"Discrete-Event Simulation of an Extended Einstein-Podolsky-Rosen-Bohm Experiment",
Front. Phys. **8**, 160 (2020)

(11) M. WILLSCH, D. WILLSCH, K. MICHIELSEN, F. JIN, T. DENKMAYR, S. SPONAR, Y. HASEGAWA, H. DE RAEDT,
     "Long-time correlations in single-neutron interferometry data",
     J. Phys. Soc. Jpn. **89**, 064005 (2020)

(12) M. WILLSCH, D. WILLSCH, F. JIN, H. DE RAEDT, K. MICHIELSEN,
     "Benchmarking the Quantum Approximate Optimization Algorithm",
     Quantum Inf. Process. **19**, 197 (2020)

Eidesstattliche Erklärung

Ich, Dennis Willsch, erkläre hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt;

2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;

3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;

4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;

5. Alle wesentlichen Quellen von Unterstützung wurden benannt;

6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;

7. Teile dieser Arbeit wurden zuvor veröffentlicht und zwar in:

[Wil2017]   D. WILLSCH, M. NOCON, F. JIN, H. DE RAEDT, K. MICHIELSEN, "Gate-error analysis in simulations of quantum computers with transmon qubits", *Phys. Rev. A* **96**, 062302 (2017)

[Wil2018b] D. WILLSCH, M. WILLSCH, F. JIN, H. DE RAEDT, K. MICHIELSEN, "Testing quantum fault tolerance on small systems", *Phys. Rev. A* **98**, 052348 (2018)

Datum und Unterschrift

Acknowledgments

There are many people to whom I am deeply grateful for supporting me during the past three years (and also before that, of course).

Kristel, first and foremost, I am immensely happy that you have given me the opportunity to do my PhD in your group. I really appreciate the time you have taken to guide me through this project, despite your full schedule of over 80 hours every week. I am truly grateful to you and Hans, you are great scientists and I am genuinely looking forward to working together with you in the future SDL at JSC so that we can make something out of all the interesting, unfinished projects we have started recently. And Hans, besides simulating time-dependent tantrum systems in D-dimensional Dilbert spaces (where clarity is an issue, after all), I have not forgotten that we still need to meet at the *voormalig station Raeren*; we will definitely make it this year (you only need to let me know a day in advance so that me and my grandfather's old bike have a chance to make it in time). Toedeloe.

David, thank you very much for proofreading my thesis and for the time you have taken for discussions with me, especially during the final stage, and for always getting back to me with additional thoughts by email. I am excited to put the results of our latest discussion on modeling electromagnetic environments to work.

Thomas, thanks a lot for finding the time to join my presentation on the problems of quantum theory and for providing valuable ideas for discussion. I very much appreciate you and Kristel offering me a position in the young quantum computing SDL that is about to start at JSC.

Seiji, thank you very much for the ideas you have given me during our regular discussions that have taken place almost every, like, year. I recall that your suggestions have helped me find an efficient scheme to manage all four-level transformations on the supercomputer. Your comments were also very helpful in finding the QUBO formulation of the SVM experiments that we did. The last time we met was just a few weeks ago — I hope your train was still coming, after 400 years.

Special thanks go to my wife Madita. I truly appreciate that you spent so much time proofreading my thesis, even though we both had less than a week left until the final deadline, and for spotting all those silly little errors that everyone else overlooked. I thank you for the countless great moments that we have shared. Living, working, and laughing with you makes me really happy.

Marco and Berni, our *Sauerbratengruppe* and *Trauzeugen*, thanks a lot for numerous bright and challenging discussions (in both space and time) and countless great moments spent sometimes in far off places and sometimes in local pubs around the corner. Thanks also to Nicolas for that gorgeous piece of software (I don't think you will ever read this — but Marco and Berni will know what it's about).

My sincere thanks go to my family, especially my parents, grandparents, and the parents of my wife, for supporting us in every possible way, always trying to get us back to life when we were deeply lost in physics and work. Kelli, thank you very much for reading my thesis and for always showing a keen interest in my work. Chris, thanks a lot for reading and helping me with the most crucial parts, as usual :D

I would like to dedicate this work to the memory of my grandfather Willi, who always, after some persuasion by grandma and me, enjoyed playing yet another round of Skat with us. I thank you for being the first to supply me with challenging games and complicated books on maths and physics (*un ahle Verzällcher us uns kölsche Sproch*) when I wasn't even able to read (let alone understand) half of the words. I have always admired you.

Finally, I would like to thank all my friends, especially Andi, Felix, Felix, Janni, Jojo, Kevin, Mel, Philipp, Philipp, Philipp, Philipp, and all the others who have more diverse names ;) Thanks for the time we spent together!

JÜLICH
Forschungszentrum