Optimizing Robust Quantum Gates in Open Quantum Systems

DISSERTATION

zur Erlangung des akademischen Grades
DOKTOR DER NATURWISSENSCHAFTEN
(Dr. rer. Nat.)
im Fach Physik

eingereicht am
Fachbereich Mathematik und Naturwissenschaften
Universität Kassel

von

Dipl.-Phys. Michael Goerz

Betreuerin: Prof. Dr. Christiane P. Koch

Zweitgutachter: Prof. Dr. Frank Wilhelm-Mauch

Eingereicht am: 30. Januar 2015
Disputation am: 04. Mai 2015

Optimizing Robust Quantum Gates in Open Quantum Systems

DISSERTATION

in partial fulfillment of the requirements
for the degree of
DOCTOR RERUM NATURALIUM
(Dr. rer. Nat.)
in the field of Physics

submitted to the
Fachbereich Mathematik und Naturwissenschaften
Universität Kassel

by

Dipl.-Phys. Michael Goerz

Advisor: Prof. Dr. Christiane P. Koch

May 2015

Optimizing Robust Quantum Gates in Open Quantum Systems

ZUSAMMENFASSUNG

Wir befinden uns an der Schwelle einer Revolution zu einer Quantentechnologie, die nicht nur auf der passiven Nutzung von Quanteneffekten, sondern auf ihrer aktiven Kontrolle beruht. An vorderster Front beinhaltet dies die Realisierung eines Quantencomputers. Das Kodieren von Informationen in Quantenzuständen als "Qubits" erlaubt es, Verschränkung und Quantensuperposition zu nutzen, um Rechnungen durchzuführen, die auf einem klassischen Computer unpraktikabel sind. Eine zentrale Schwierigkeit ist es dabei, Dekohärenz zu vermeiden – der Verlust von Quanteneigenschaften aufgrund ungewollter Wechselwirkung mit der Umgebung. Diese Arbeit thematisiert die Realisierung verschränkender Zwei-Qubit-Gatter, die sowohl gegenüber Dekohärenz als auch klassischen Störeinflüssen robust sind. Sie behandelt dabei drei Aspekte: die Nutzung effizienter numerischer Methoden zur Simulation und optimaler Kontrolle offener und geschlossener Quantensysteme, die Rolle fortgeschrittener Optimierungsfunktionale zur Begünstigung von Robustheit, sowie die Anwendung dieser Techniken auf zwei führende Umsetzungen von Quantencomputern, gefangene Atome und supraleitende Schaltkreise.

Nach einem Überblick über die theoretischen und numerischen Grundlagen beginnt der zentrale Teil dieser Arbeit mit der Idee einer Ensembleoptimierung, um Robustheit sowohl gegenüber klassischen Fluktuationen als auch Dekohärenz zu erreichen. Für das Beispiel eines kontrollierten Phasengatters auf gefangenen Rydberg-Atomen wird gezeigt, dass Gatter erreichbar sind, die um mindestens eine Größenordnung robuster sind als der beste bekannte analytische Ansatz. Darüber hinaus bleibt diese Robustheit selbst dann erhalten, wenn die Gatterdauer signifikant gegenüber der kurzmöglichsten Dauer des analytischen Gatters verkürzt wird.

Supraleitenden Schaltkreise sind eine besonders vielversprechende Architektur zur Implementierung eines Quantencomputers. Ihre Flexibilität wird durch Optimierungen sowohl für diagonale als auch nicht-diagonale Gatter gezeigt. Um Robustheit gegenüber Dekohärenz zu gewährleisten, ist es essentiell, das Gatter in so kurzer Zeit wie möglich zu realisieren. Das Erreichen dieses Ziels wird durch die Optimierung hin zu einem beliebigen perfekten Verschränker erleichtert, basierend auf einer geometrischen Theorie der Zwei-Qubit-Gatter. Für das Beispiel supraleitender Qubits wird gezeigt, dass dieser Ansatz zu kürzeren Gatterzeiten, höheren Fidelitäten, sowie schnellerer Konvergenz führt, im Vergleich zur Optimierung hin zu vorausbestimmten, festen Zwei-Qubit-Gattern.

Eine Optimierung im Liouville-Raum zur sauberen Berücksichtigung von Dekohärenzeffekten ist mit erheblichen numerischen Herausforderungen verbunden, da die Dimension im Vergleich zum Hilbert-Raum quadratisch wächst. Es kann allerdings gezeigt werden, dass es für ein unitäres Optimierungsziel ausreichend ist, höchstens drei Zustände anstelle der vollen Basis des Liouville-Raums zu propagieren. Sowohl für das Beispiel gefangener Rydberg-Atome also auch für supraleitende Qubits wird die erfolgreiche Optimierung von Quantengattern gezeigt, mit einem numerischen Aufwand, der weit unterhalb der bisher angenommenen Untergrenze liegt. Insgesamt zeigen die Ergebnisse dieser Arbeit zu einem umfassenden Gerüsts zur Optimierung robuster Quantengatter, und bereiten den Weg für die mögliche Realisierung eines Quantencomputers.

Optimizing Robust Quantum Gates in Open Quantum Systems

ABSTRACT

We are currently at the cusp of a revolution in quantum technology that relies not just on the passive use of quantum effects, but on their active control. At the forefront of this revolution is the implementation of a quantum computer. Encoding information in quantum states as "qubits" allows to use entanglement and quantum superposition to perform calculations that are infeasible on classical computers. The fundamental challenge in the realization of quantum computers is to avoid decoherence – the loss of quantum properties – due to unwanted interaction with the environment. This thesis addresses the problem of implementing entangling two-qubit quantum gates that are robust with respect to both decoherence and classical noise. It covers three aspects: the use of efficient numerical tools for the simulation and optimal control of open and closed quantum systems, the role of advanced optimization functionals in facilitating robustness, and the application of these techniques to two of the leading implementations of quantum computation, trapped atoms and superconducting circuits.

After a review of the theoretical and numerical foundations, the central part of the thesis starts with the idea of using ensemble optimization to achieve robustness with respect to both classical fluctuations in the system parameters, and decoherence. For the example of a controlled phasegate implemented with trapped Rydberg atoms, this approach is demonstrated to yield a gate that is at least one order of magnitude more robust than the best known analytic scheme. Moreover this robustness is maintained even for gate durations significantly shorter than those obtained in the analytic scheme.

Superconducting circuits are a particularly promising architecture for the implementation of a quantum computer. Their flexibility is demonstrated by performing optimizations for both diagonal and non-diagonal quantum gates. In order to achieve robustness with respect to decoherence, it is essential to implement quantum gates in the shortest possible amount of time. This may be facilitated by using an optimization functional that targets an arbitrary perfect entangler, based on a geometric theory of two-qubit gates. For the example of superconducting qubits, it is shown that this approach leads to significantly shorter gate durations, higher fidelities, and faster convergence than the optimization towards specific two-qubit gates.

Performing optimization in Liouville space in order to properly take into account decoherence poses significant numerical challenges, as the dimension scales quadratically compared to Hilbert space. However, it can be shown that for a unitary target, the optimization only requires propagation of at most three states, instead of a full basis of Liouville space. Both for the example of trapped Rydberg atoms, and for superconducting qubits, the successful optimization of quantum gates is demonstrated, at a significantly reduced numerical cost than was previously thought possible. Together, the results of this thesis point towards a comprehensive framework for the optimization of robust quantum gates, paving the way for the future realization of quantum computers.

Contents

To my parents

Acknowledgments

First and foremost, I would like to thank my advisor, Christiane Koch, for her constant support and guidance. She has been truly exceptional in her dedication and patience, giving me the opportunity to work on a wide range of fascinating problems in quantum computing and optimal control, and to follow my passion for computing. Also, she has been providing me the chance, through several extended research visits and workshops, e.g. the KITP program on Control of Complex Quantum Systems, to exchange ideas and collaborate with a great number of wonderful people.

I would like to thank Birgitta Whaley for letting me spend a combined few months in her group at Berkeley. Her guidance has also been invaluable. Many of the results presented in this thesis have originated from work done as part of this collaboration. Felix Motzoi has taught me much about superconducting qubits. Jon Aytac and especially Eli Halperin spent much effort in analyzing the analytical schemes for Rydberg gates.

My thanks also goes to all the past and present members of the group in Kassel, Daniel Reich, Giulia Gualdi, Wojciech Skomorowski, Michał Tomza, Martin Berglund, and Esteban Goetz. Daniel and Giulia in particular have contributed much to my work. Daniel's deep understanding of the mathematics of optimal control has been the basis of the results in chapter 7. Giulia has taught me much about open quantum systems. As part of their Bachelor theses, Daniel Basilewitsch and Lutz Marder helped in some of the calculations for the transmon system, and in the implementation of the Newton propagator.

On a personal level, I would like to thank Cathy Kudlick for being a source of inspiration and comfort. Anton Haase was incredibly helpful in preparing for the defense of this thesis. Lastly, and especially, I thank my parents, Christoph Goerz and Lieve Gevaert-Goerz, to whom this thesis is dedicated, for their endless support.

1

Introduction

Quantum mechanics has been one of the most fundamental and significant achievements of modern physics. While to this day, the theory can seem perplexing and at odds with the categories of everyday experience, it has withstood any test put to it, and indeed allowed unprecedented insights into the fundamentals of nature. Many of the philosophical questions that quantum mechanics raises remain open. Is quantum mechanics just a mathematical tool that happens to yield accurate predictions, or are wave functions in fact “real” objects? Can it be true that particles are everywhere at once, until a measurement pins them down? Is it acceptable to have a theory of nature that involves blind chance at a fundamental level? Can the gap between the quantum and the classical ever be bridged? All of these questions continue to be explored by both philosophers and scientists.

At the same time, the theory itself has matured, and many scientists and engineers have gone beyond such fundamental concerns and moved on to explore its applications, sparking a technological revolution. Remarkably, while the knowledge of quantum mechanics seems utterly irrelevant to the average person on the street, it is at the core of an immeasurable number of today’s technologies. In fact, one would be hard-pressed to find any technological advance of the last decades that does not in some way rely on our understanding of quantum mechanics. Without it, we would have no lasers, no MRI scanners, no modern chemistry, and maybe most importantly, no modern electronics and thus no information technology.

Presently, we are at the cusp of a second wave of quantum technology that is based not just on the passive understanding of quantum effects, but on the active control and manipulation of quantum systems [1]. This new "quantum engineering" has considerable challenges, but also holds the promise of unprecedented new possibilities; for example the control of chemical reactions with shaped laser pulses [2], or advances in renewable energy, where a new generation of solar power cells could mimic the process of photosynthesis [3, 4]. Probably the most far-reaching example, and the main focus of this thesis is the field of quantum information processing, where logical operations are performed by manipulating the quantum system on which the information is encoded [5].

1.1 Coherent Control of Quantum Systems

A common thread through the newly developing quantum technologies is the concept of *coherent control*. The control problem can be formulated as follows: Given a quantum system in a well-defined initial state, and a Hamiltonian

$$\hat{\mathsf{H}} = \hat{\mathsf{H}}_0 + \hat{\mathsf{H}}_c[u(t)] \tag{1.1}$$

that includes a control parameter $u(t)$, which choice of $u(t)$ ensures that the system evolves to some desired target state, or implements a desired process? Instead of a single control $u(t)$, we might also have a set of multiple control parameters $\{u_i(t)\}$.

For example, the interaction of an electronic or nuclear spin with a magnetic field $\vec{B}(t)$ is described by the Hamiltonian

$$\hat{\mathsf{H}} = -\gamma \left(\hat{\mathsf{S}}_x B_x(t) + \hat{\mathsf{S}}_y B_y(t) + \hat{\mathsf{S}}_z B_z(t) \right) , \tag{1.2}$$

where γ is the gyromagnetic ratio, $\hat{\mathsf{S}}_x = \frac{\hbar}{2}\hat{\sigma}_x$, $\hat{\mathsf{S}}_y = \frac{\hbar}{2}\hat{\sigma}_y$, and $\hat{\mathsf{S}}_z = \frac{\hbar}{2}\hat{\sigma}_z$ are the operators measuring the spin in the three spatial directions, proportional to the Pauli matrices

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \hat{\sigma}_x = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} , \quad \hat{\sigma}_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.3}$$

The components of the magnetic field $B_x(t)$, $B_y(t)$, $B_z(t)$ are the control parameters, and we may ask which magnetic field will bring an arbitrary initial state $|\Psi_0(t=0)\rangle$ to an arbitrary target state $|\Psi^{\text{tgt}}(t=T)\rangle$ at some final time T. On a larger scale, such a control of spin systems forms the basis of next-generation medical imaging technology or spectroscopy of complex molecules, among other applications. In the same way, we can ask which laser field will drive the electronic states of atoms to bind them into ultra-cold molecules [6], which voltages to the electrodes of an ion-trap will transport an ion over a large distance [7], or which variation of the confinement potential of a Bose-Einstein condensate will split the wave function [8].

In all cases, the solution to the control problem relies on the well-defined phase relation between the basis states of the Hilbert space, exploiting the interference between multiple pathways. This is what defines the control as *coherent* and thus puts the control problem entirely in the quantum domain.

The control problem is not limited to simple state-to-state transformations, but can equally apply to the implementation of a desired unitary evolution $\hat{\mathsf{O}}$ on an N-dimensional Hilbert space. In this case, we start from a set of basis states $\{\vert\phi_1\rangle, \dots, \vert\phi_N\rangle\}$ and attempt to find a single control that maps each of the basis states $\vert\phi_j\rangle$ to $\hat{\mathsf{O}}\,\vert\phi_j\rangle$:

$$\forall j: \quad \hat{\mathsf{U}}(T, 0; u(t))\,\vert\phi_j\rangle \overset{!}{=} \hat{\mathsf{O}}\,\vert\phi_j\rangle \tag{1.4}$$

where $\hat{\mathsf{U}}(T, 0; u(t))$ is the time evolution operator induced by the Hamiltonian $\hat{\mathsf{H}}[u(t)]$ from $t = 0$ to $t = T$. Note that the control $u(t)$ is the same for each of the basis states. If $\{\vert\phi_j\rangle\}$ is a complete basis, then any state $\vert\Psi\rangle$ can be expanded as $\vert\Psi\rangle = \sum_j \alpha_j\,\vert\phi_j\rangle$, and thus we have implemented $\hat{\mathsf{U}}\,\vert\Psi\rangle = \hat{\mathsf{O}}\,\vert\Psi\rangle$ for an arbitrary state. This type of control problem is especially relevant to quantum computation where $\hat{\mathsf{O}}$ is a *quantum gate*.

We can go further and formulate the control problem on a more mathematical basis by defining a functional $J_T\left(\{\vert\Psi_j(T)\rangle\}\right)$ that becomes minimal if (and only if) the control implements the desired target. For example, for a state-to-state transition, one possible choice is

$$\begin{aligned} J_{T,\text{ss}}(\vert\Psi(T)\rangle) &= 1 - \left\vert\left\langle \Psi(t=0) \,\middle\vert\, \hat{\mathsf{U}}^\dagger(T, 0; u(t)) \,\middle\vert\, \Psi^{\text{tgt}}(t=T) \right\rangle\right\vert^2 \\ &= 1 - \left\vert\left\langle \Psi(T) \,\middle\vert\, \Psi^{\text{tgt}}(T) \right\rangle\right\vert^2 , \end{aligned} \tag{1.5}$$

where $\hat{\mathsf{U}}(T, 0; u(t))$ is again the time evolution operator; $\hat{\mathsf{U}}^\dagger$ acts to the left in Eq. (1.5), as $\vert\Psi(t)\rangle = \hat{\mathsf{U}}(t, 0; u(t))\,\vert\Psi(t=0)\rangle$. The task has now become an *optimization problem* of finding the $u(t)$ that minimizes J_T. The time evolution operator places an implicit constraint on the optimization, in that the state $\vert\Psi(t)\rangle$ must be a solution to the proper equation of motion, usually the Schrödinger equation:

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\,\vert\Psi(t)\rangle = \hat{\mathsf{H}}[u(t)\,\vert\Psi(t)\rangle\ . \tag{1.6}$$

The methods for solving this constrained optimization problem constitutes the framework of *optimal control*.

In simple cases, the solution to the control problem can be derived analytically. In the above example of the two-level system of a single spin, we can find a solution simply by solving the Schrödinger equation.

We write the initial state $|\Psi\rangle$ in its Feynman-Vernon-Hellwarth (FVH) representation [9, 10], a three component vector $\vec{r}$ obtained from projecting the density operator $\hat{\rho} \equiv |\Psi\rangle\langle\Psi|$ onto the three Pauli matrices

$$r_i = \mathrm{tr}\left[\hat{\sigma}_i \hat{\rho}\right] . \tag{1.7}$$

The pulse is also represented as a vector $\vec{\Omega}$ with the three components $B_x(t)$, $B_y(t)$, $B_z(t)$, and the FVH-vector now simply precesses around $\vec{\Omega}$,

$$\dot{\vec{r}} = \vec{\Omega}(t) \times \vec{r}. \tag{1.8}$$

The optimal solution for a state transfer from $\vec{r}_0(t=0)$ to $\vec{r}_{\mathrm{tgt}}(t=T)$ is easy to see: choose $\vec{\Omega}$ such that it points to the center of the geodesic circle connecting $\vec{r}_0$ and $\vec{r}_{\mathrm{tgt}}$, and switch the magnetic field to an arbitrary constant amplitude until the target state is reached.

More generally, analytical solutions are usually based on a direct application of Pontryagin's maximum principle [11], a generalization of the Euler-Lagrange equation familiar from the variational calculus of classical mechanics. For spin systems, a number of non-trivial control problems have been solved using a technique named *geometric control*, such as the implementation of quantum gates [12], dissipative state-to-state transfer [13], and control under inhomogeneous magnetic fields [14].

Usually, an analytic solution to the control problem is only possible as long as the system or the optimization functional is reasonably “simple”. One problem is that the analytic solutions are often of *bang-bang* type, i.e., they switch instantaneously between zero and some constant amplitude. This type of control would not be compatible with constraints on the smoothness or spectral width of the control, reflecting that such controls may not be realistic. For example, a laser field cannot be switched instantaneously.

Furthermore, analytical solutions are usually restricted to Hilbert spaces of very small dimension. A standard approach is to try to derive a reduced effective model of a physical system that is valid in a certain parameter regime. For example, in a two-photon transition between two levels via one or more intermediary levels, the intermediate levels can be *adiabatically eliminated* if both driving fields are far off-resonant from the intermediary levels and if the pulse shapes are slowly varying. Another example where the dynamics can be effectively limited to a subset of the full system is population-transfer in a Λ-shaped three-level system via the popular STIRAP mechanism [15]. In both of these examples, the approximations only hold under severe restrictions, specifically that the dynamics must be *adiabatic*, i.e., that the system is always in an instantaneous eigenstate of the time-dependent Hamiltonian. This is only true if the controls vary on a sufficiently slow time scale, which is in direct contradiction to the requirement to implement the target process in as little time as possible. In chapter 4, we will see the limitations of such techniques in the context of a quantum gate optimization.

To circumvent the limitations of analytical optimal control, we turn to *numerical optimal control* which takes a different approach. Starting from a set of sub-optimal “guesses”, the control fields are iteratively improved: The equation of motion is solved numerically, allowing to evaluate the functional, and to modify each $u_i(t)$ such that the value of the optimization functional decreases. This new optimized set of controls is then used as the guess for another iteration. Eventually, assuming there are no local minima in the optimization landscape, the procedure will converge towards an optimal solution.

Numerical optimal control theory (OCT) holds the promise of providing an extremely versatile tool for designing the controls necessary in a new generation of quantum technology, going beyond simple models that can be approached analytically, and instead meeting the demands imposed by real-life implementations. Fundamentally, the method is only limited by the available computational resources. This does not imply that OCT can provide a complete black-box solution. The efficient numerical simulation of the system dynamics relies on an suitable model that identifies the relevant degrees of freedom. Moreover, the iterative nature of the control scheme makes it necessary to start from a reasonably well-chosen guess pulse in order to converge quickly. Designing such a guess pulse still requires a thorough understanding of least some possible control mechanisms. Lastly, the optimized pulses resulting from OCT can be extremely complex, sometimes making their physical implementation a challenge. From a theorist’s perspective, understanding the mechanisms employed by the pulse can be a non-trivial task, but it provides the chance for an important interplay between analytical and numerical solutions. Ideally, entirely new control mechanisms can be identified, or OCT can find the most suitable out of several possible strategies, and the gained knowledge can go into the design of better guess pulses, or even new analytic schemes. Thus, OCT does not substitute for a deep understanding of the physics of a given quantum system, but it augments the framework of coherent control by an important and powerful tool.

The methods of optimal control, both analytic and numerical, originate from extensive work in mathematics and engineering [16, 11, 17, 18, 19]. In the context of quantum systems, these ideas were first applied to the control of molecular interactions [20, 2], and later to the control of spins in nuclear magnetic resonance [21, 22, 23] as well as to a wide range of other quantum systems [24]. Some of the central algorithms used in numerical optimal control of quantum systems are presented in chapter 3.

1.2 Quantum Computation

Quantum computation [5] is a particularly far-reaching and exciting type of quantum technology. As microchips become smaller and smaller, following Moore's law [25], they will reach sizes at which quantum effects are dominant. At this point, in order to achieve further improvements in computing power, a paradigm shift is necessary, where quantum effects are no longer seen as a perturbation to classical electronic circuits, but are actively involved in the computational process.

The motivation for quantum computers is more fundamental than that, however. The "weirdness" of quantum mechanics promises radically more powerful ways of processing information. A key observation is that quantum mechanics cannot efficiently be simulated on a classical computer [26], giving the motivation for quantum computing was to use one quantum system to simulate the behavior of another. This is possible because the theoretical framework of quantum mechanics is entirely abstract from the diverse range of systems it models. Therefore, a quantum system that is easily engineered and controlled in the lab could stand in for another system that is less accessible. The far-reached applications of a quantum simulator typically include open problems in solid states physics, such as high-temperature superconductivity or quantum phase transitions [27, 28], but extend to other fields such as high-energy physics or cosmology as well [29].

The reason that a classical computer cannot efficiently simulate a large quantum system originates from the feature that distinguishes quantum states from classical states: *quantum superposition* and *entanglement*. Whereas a register of n classical bits encodes exactly one of 2^n possible values composed of n binary digits b_i, $\text{reg} = |b_0 \dots b_n|$, the quantum version of this register is described by a quantum state

$$|\Psi\rangle = \sum_{i=0}^{2^n-1} a_i \left(\left| q_0^{(i)} \right\rangle \otimes \cdots \otimes \left| q_{n-1}^{(i)} \right\rangle \right) , \tag{1.9}$$

with $q_j^{(i)} \in \{0, 1\}$, i.e., a simultaneous superposition of all the 2^n eigenstates of the n-quantum-bit-register, with complex coefficients a_i. The amount of information required to describe the state of a quantum system grows exponentially with the size of the system, but grows only linearly for a classical system. This raises the question whether more generally, a quantum system could also implement a *universal* computer that would be fundamentally more powerful than classical computers.

In theoretical computer science, the question of computability is addressed in the model of the Turing machine [30] as a universal computer, i.e., one that can compute any computable function. As an important prerequisite to the possibility of a universal quantum computer, a quantum version of such a Turing machine has been shown [31] to be at least as powerful as its classical counterpart, and to solve at least some problems with greater efficiency [32].

In practice, the concept of a universal quantum computer takes its inspiration from the standard digital computer, implemented as a network of electronic logical gates on a microchip. Such a computer can perform any manipulation of bits in its memory. Similarly, a gate-based model of a quantum computer, reviewed in chapter 2 is implemented as a network of quantum gates that perform any operation on the quantum bits in its memory. Compared to classical circuits, there are some caveats imposed by the laws of quantum mechanics [5]: the quantum network has to be *unitary*, i.e., fully reversible and the copying of a quantum state is not allowed [33].

At first glance, quantum superposition might seem to make it obvious that a quantum computer could solve problems at exponentially greater efficiency than a classical computer. In a quantum analogue of a classical circuit, one could simply prepare the quantum superposition of all possible inputs, and the circuit would simultaneously calculate all possible results, in a form of *quantum parallelism*. Unfortunately, the resulting quantum state has to be measured in order to read it out, collapsing the superposition of results to one result at random. Designing quantum algorithms in which some of the computational power resulting from quantum parallelism survives the measurement process requires great ingenuity. The most striking case in which such an algorithm has been devised is Shor's method of factoring numbers [34]. Given an integer that is the product of two large prime numbers, the best classical algorithm to find the two factors runs in near exponential time [35], whereas Shor's algorithm can perform the same task in polynomial time on a quantum computer.

In general, the understanding of all the ways quantum algorithms could outperform classical algorithms is incomplete. An underlying theme, however, is that it can be possible to obtain some *global* property, i.e., a property shared by *all* states in the superposition. The question then becomes how to cleverly manipulate the input of the quantum computer such that the resulting output has some interesting global property, which will survive the measurement. This is exactly what Shor's algorithm exploits. It turns out [36] that finding the prime factors of a large integer N can be mapped to finding the period of the function

$$f(x) = a^x \mod N \tag{1.10}$$

for random integers $a < N$. The period is found using a quantum Fourier transform that maps every eigenstate $|i\rangle$ of an n-dimensional Hilbert space as

$$|i\rangle \to \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \exp\left[2\pi \mathrm{i}\frac{ij}{n}\right] |j\rangle \ . \tag{1.11}$$

This is analogous to the classical Fourier transform, cf. appendix E; as in the classical case, each value in the spectrum gives a global property of the original function. In the case of the quantum Fourier transform, a measurement has the greatest likelihood for yielding the dominant frequency component of $f(x)$, and thus ultimately solves the factoring problem.

The factoring of large integers is a highly relevant problem, because the security of one of the most widely used encryption systems, the RSA algorithm [37], relies on it. Having a machine that could factor large integers efficiently would instantly break all such encryption.

Beyond the quantum Fourier transform and the algorithms derived from it, there is currently one other class of quantum algorithms that are fundamentally faster than their classical counterparts. The core problem is finding one specific value from a list of N non-sorted entries. In the classical case, there is no better solution than to look at all N entries of the list, stopping when the target element is found. Quantum mechanically, Grover's quantum search algorithm [38] can perform the same task in only $\sqrt{N}$ steps.

While Grover's algorithm provides a less impressive advantage over the classical algorithm, its applications are much farther reaching. There is a large class of "hard" problems where there is no way to find a solution in polynomial time, but it is easy to verify that a given solution is the correct one. In complexity theory, such problems belong to the category "NP" [39]. These include many famous graph-theoretical optimization problems, such as versions of the traveling salesman problem [40] or the graph coloring problem [41]. In the worst case, an NP problem can be solved by enumerating all N possible solutions and selecting the desired one, where N depends exponentially on the size of the problem. Using Grover's algorithm in such a case would give a very significant speedup $N \to \sqrt{N}$, even though the solution remains exponential in principle.

The underlying idea of Grover's algorithm is that since a solution can easily be identified, it is possible to implement an *oracle* quantum gate that selectively flips the phase of the eigenstate associated with the correct solution. Once the eigenstate has been "tagged" in this way in a superposition of all eigenstates, the *amplitude amplification* technique [42, 43] is employed, which increases the relative amplitude of the tagged state. This amplification has to be repeated on the order of $\sqrt{N}$ time before the target state dominates over all other eigenstates, and can be measured with high likelihood. Essentially, the quantum search algorithm finds a needle in a haystack by increasingly growing the needle and shrinking the haystack, until the needle can be easily picked out by a blind grab.

There are a wide variety of quantum systems that have been considered to implement such a universal quantum computer [5], from nuclear spins [44], to quantum dots [45], to trapped ions [46] to nitrogen-vacancy centers [47] – to name just a few examples. No implementation to date has overcome all the technical challenges required to implement a working, universal, large-scale quantum computer. In this thesis, two systems in particular are used as examples, trapped Rydberg atoms, reviewed in section 4.1 of chapter 4, and superconducting circuits, reviewed in chapter 5. The latter are particularly promising candidates for quantum computing, since they share many of their techniques with present day classical computer chips and can be engineered with great versatility.

There are some aspects of using quantum systems to store and process information that go beyond the ideas of universal quantum computers. With the cryptographic applications of the most prominent quantum algorithms threatening the security of existing encryption schemes, quantum mechanics has also provided a possible answer, in the form of *quantum communication* and *quantum cryptography* [48, 49, and references therein]. Information is stored in the quantum states of a photon, which can easily be transmitted through optical fibers. Since unknown quantum states cannot be copied [33], and any measurement collapses the quantum state, it is impossible to wiretap the optical fiber without the eavesdropping being detected. Thus, quantum communication provides a fundamentally secure channel of communication. Once a secure channel is established, it can be used to exchange cryptographic keys, providing the basis for an unbreakable encryption system [50], although in practical implementations, loopholes may still exist [51]. Such *quantum key-distribution networks* have been successfully implemented and are commercially available [52, 53, 54].

There has also been considerable interest in quantum computers that do not fit into the standard gate model, i.e., "special-purpose" quantum computers. One approach is that of *adiabatic quantum computing* [55]. The idea is to engineer a complex Hamiltonian $\hat{\mathsf{H}}(T)$ whose (unknown) ground state encodes the result of a computation. The system is then initialized to the known ground state of a simpler Hamiltonian $\hat{\mathsf{H}}(0)$. Then, the simple Hamiltonian is slowly ("adiabatically") transformed into the complex result Hamiltonian. According to the adiabatic theorem, the system will remain in the ground state of the evolving $\hat{\mathsf{H}}(t)$ at all times. In this way, the ground state of the complex Hamiltonian, and thus the solution to the computation problem is found. Performed fully coherently, at zero temperature, adiabatic quantum computing can be shown to be mathematically equivalent to the standard gate model [56]. A "messier" version at non-zero temperature, in which the qubits are strongly coupled to their environment throughout, but still maintain some quantum coherence is known as *quantum annealing*. A quantum processor based on quantum annealing has been built and marketed [57] by the Canadian company DWave, and is to date the only commercially available quantum information processing device (outside of quantum communication networks). However, the DWave processor has been subject to severe criticism [58, 59, 60], questioning whether it provides any speedup over classical computers. Nonetheless, applications of the architecture to machine learning [61, 62] are currently being explored.

1.3 Decoherence

The standard description of quantum mechanics [63] considers a quantum system to be in complete isolation. This is almost never a realistic assumption. Any quantum system will have some remaining interaction with its environment; this is especially true in a control context, where simply the fact that the system can be influenced from the outside indicates that it cannot be completely isolated. For example, the same dipole moment that allows an atom to be controlled with a laser beam also couples that atom to possibly unwanted stray photons, or even the vacuum mode of the electromagnetic field, causing spontaneous emission. Such unwanted interaction with the environment is the source of *decoherence* [64], i.e., the loss of “quantumness”. More precisely, the fixed phase relation between the eigenstates of the Hilbert space is lost. Since coherent control relies on exploiting the interference of exactly these phases, decoherence is fundamentally detrimental to the objective. For a technology built specifically on quantum features, decoherence is an obvious challenge.

Mathematically, the effects of decoherence are well-described. Instead of a state vector $|\Psi\rangle$ in Hilbert space, the system is now modeled as a density matrix $\hat{\rho}$ in Liouville space. The density matrix formalism, presented in chapter 2, allows to represent both *pure states* that are equivalent to the state vectors in Hilbert space, but also *mixed states* that describe the system when decoherence has taken place. In a closed system, the direct equivalent of the Schrödinger equation (1.6) is the Liouville-von Neumann equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\rho}_{\text{closed}} = \left[\hat{\mathsf{H}}, \hat{\rho}_{\text{closed}}\right] . \tag{1.12}$$

For an open system, the decoherence is expressed as a dissipator $\mathcal{D}$, which occurs as an additional term in the Liouville-von Neumann equation.

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\rho}_{\text{open}} = \left[\hat{\mathsf{H}}, \hat{\rho}_{\text{open}}\right] + \mathcal{D}\left[\hat{\rho}_{\text{open}}\right] . \tag{1.13}$$

In principle, the methods of OCT still apply to the open quantum system, except that the control functional to be minimized now has to be expressed in Liouville space. For example, the Liouville space equivalent to Eq. (1.5) is

$$J_{T,\mathrm{ss}}(\hat{\rho}) = 1 - \left| \mathrm{tr} \left[\hat{\rho}^{\dagger}(T) \hat{\rho}^{\mathrm{tgt}} \right] \right|^{2} . \tag{1.14}$$

In the presence of decoherence, the promise of optimal control is only reinforced. One reason is that in Liouville space, it is even harder than in Hilbert space to devise analytical control schemes. Thus, numerical tools are often the only way to effectively tackle the problem. Examples for the use of numerical optimal control of systems that are subject to decoherence include laser cooling of the internal degrees of freedom in molecules, maximizing coherences [65], distilling out the influence of noise [66], guiding the dynamics of a quantum state [67], and the mitigation of Markovian dephasing noise on single qubits [68].

For the realization of large scale quantum computing, it is crucial to limit or circumvent the effects of decoherence. The phase relation between the states in the quantum computer is at the heart of both the algorithms discussed in section 1.2. In Shor's algorithm, the underlying quantum Fourier transform identifies global properties of the wave function, inherently embodied in phase information. In Grover's algorithm, the correct solution to the search is tagged by a phase flip, allowing it to be amplified. Thus, the loss of phase information destroys all chance for true quantum computation, and the speedup it promises over classical computers. Optimal control provides a tool for addressing the issue of decoherence, allowing to implement quantum gates with sufficiently high fidelity that *quantum error correction* can guarantee fault-tolerant quantum computing.

Taking decoherence into account explicitly in the optimization allows to actively search for solutions where its effects are minimized. Generally, not all levels suffer equally from decoherence, and optimal control can restrict the solution to those levels that are least affected. In chapter 4, we consider a case where there is significant spontaneous decay from only one of the levels, and see how optimal control circumvents the effects of decoherence by not populating this level. Lastly, an obvious way to counter decoherence, is to find solutions for the control problem that act on a faster time scale than the dissipation, so that the desired process is implemented before the decoherence becomes too noticeable. In this context, it becomes especially important that numerical optimal control can find the shortest possible controls, reaching what has been called the quantum speed limit [69, 70, 71, 72, 73].

1.4 Organization of the Thesis

The work presented in this thesis explores possibilities for using numerical optimal control to implement robust two-qubit gates in open quantum systems. Chapter 2 gives an overview of the theoretical foundations, outlining the concepts of quantum computing, followed by an introduction to the description of open quantum systems. Chapter 3 discusses the necessary numerical tools for the representation and dynamical simulation of both open and closed quantum systems, as well as algorithms for optimal control.

Chapter 4 presents a first example for the application of these techniques to the implementation of a quantum gate for trapped neutral atoms. Robustness with respect both to dissipation and fluctuations in technical parameters is achieved by performing the optimization in Liouville space, and by optimizing for an ensemble of Hamiltonians that samples over the relevant range of fluctuations. By including the desire for robustness explicitly in the optimization functional, optimal control can find solutions that meet the desired objectives.

One of the most promising architectures for quantum computation is based on superconducting circuits, reviewed in chapter 5. The qubits in this implementation allow for unparalleled flexibility, where parameters and interactions can be engineered over a wide range. Especially with the development of the *transmon* qubit, decoherence times have been pushed to a regime where large scale quantum computing appears within reach. Entangling gates on two transmons can be implemented by coupling both qubits to a shared transmission line resonator, which mediates an effective coupling between them. We discuss the derivation of an effective model for the qubit-qubit interaction, which is the basis for some of the results presented in chapters 6 and 7. Furthermore, we review some of the basic gate mechanisms that have been demonstrated experimentally, before exploring the use of an off-resonant driving field to induce Stark shifts in the levels of the logical subspace, realizing a holonomic phasegate. Finally, we show optimization results for a CPHASE, CNOT, and the holonomic gate in the full qubit-qubit-cavity model.

Chapter 6 illustrates the power of optimal control to exploit any freedom allowed by the optimization functional. Instead of optimizing for a specific quantum gate implemented on superconducting qubits, the target is an arbitrary *perfect entangler*, i.e., a gate that entangles some initially separable states. Not restricting the optimization unnecessarily holds the promise of allowing optimal control to identify the gates least affected by decoherence, or the gates that can be implemented in the shortest possible time, in the hope the decoherence will only become relevant on longer time scales.

The quadratic scaling of Liouville space compared to Hilbert space provides significant challenges for the simulation and optimization of open quantum systems. Chapter 7 illustrates the possibility of a significant reduction in the resources required for optimization. Under the assumption that the optimization target is unitary (e.g., a quantum gate), it is not necessary to consider a full basis of Liouville space consisting of d^2 matrices, where d is the dimension of the underlying Hilbert space. Instead, at most 3 states need to be included in the optimization. This is illustrated for both trapped neutral atoms and superconducting circuits. Finally, chapter 8 summarizes and gives an outlook.

2

Quantum Information in Open Quantum Systems

The central focus of this thesis is the numerical realization of quantum gates that are robust with respect to dissipation. While the numerical tools are presented in chapter 3, here, a concise overview of the theoretical foundations is given. We touch upon two sub-fields of quantum theory. The first is quantum computing, i.e., the theory of how information can be encoded and processed in quantum systems. After introducing the basic terminology of quantum mechanics in section 2.1 and defining quantum bits (qubits) in section 2.2, we discuss single-qubit and two-qubit gates in sections 2.3–2.4. The second field is the theory of open quantum systems, i.e., of quantum systems undergoing dissipation. The state of such a system can no longer be described as a vector in Hilbert space. Instead, the *density matrix formalism* reviewed in section 2.5 is used. The time-evolution within this formalism is presented in section 2.6, before discussing some of the techniques for countering decoherence in section 2.7.

The material presented here is not a complete review of either topic by far; the intention is to define the central terms, introduce the notation used throughout this thesis, and to provide a point of reference for the following chapters. The description of the fundamental concepts of quantum mechanics follow the book by Sakurai and Napolitano [63], whereas the discussion on quantum computation and open quantum systems is mostly based on the textbooks by Nielsen and Chuang [5], and Breuer and Petruccione [64], respectively.

2.1 The Postulates of Quantum Mechanics

The fundamental postulate of quantum mechanics is that the state of a system is completely described by a wave function $|\Psi\rangle$, element of a Hilbert space, and that every physical quantity, or *observable*, is associated with a Hermitian operator $\hat{\mathsf{A}}$. We also define the dual $\langle\Psi|$ that allows to formulate the inner product $(\cdot,\cdot) \to \mathbb{C}$. For two states $|\Psi\rangle$, $|\phi\rangle$, their inner product, or *overlap*, is written as $\langle\Psi\,|\,\phi\rangle$. All states are normalized as $\||\Psi\rangle\| = \sqrt{\langle\Psi\,|\,\Psi\rangle} = 1$.

Quantum mechanics is inherently probabilistic; knowledge of a quantum state $|\Psi\rangle$ does not in general determine a unique value for an observable, but only predicts the *expectation value*

$$\left\langle\hat{\mathsf{A}}\right\rangle = \left\langle\Psi\,\middle|\,\hat{\mathsf{A}}\,\middle|\,\Psi\right\rangle = \Big(|\Psi\rangle\,,\hat{\mathsf{A}}\,|\Psi\rangle\Big)\;. \tag{2.1}$$

Since the operator $\hat{\mathsf{A}}$ is Hermitian, its expectation value is real, as required for a physically meaningful quantity. Of particular interest is the Hamiltonian, whose expectation value is the total energy E of the system, as expressed by the time-independent Schrödinger equation,

$$\hat{\mathsf{H}}\,|\Psi\rangle = E\,|\Psi\rangle\;. \tag{2.2}$$

Eq. (2.2) takes the form of an eigenvalue problem, where $|\Psi\rangle$ is the eigenstate associated with the eigenvalue E. For bound states, i.e., those that are localized by a potential to a finite region of space, the spectrum of the Hamiltonian is discrete. That is, Eq. (2.2) can be fulfilled only for a set of N eigenstates,

$$\hat{\mathsf{H}}\,|\phi_n\rangle = E_n\,|\phi_n\rangle\;,\quad n = 1\ldots N\,. \tag{2.3}$$

The discrete, or “quantized” energy levels are what gives quantum theory its name.

The set of eigenstates $\{|\phi_n\rangle\}$ is *complete*, i.e., any state $|\Psi\rangle$ can be expanded as

$$|\Psi\rangle = \sum_{n=1}^{N} a_n\,|\phi_n\rangle\;,\quad \sum_{n=1}^{N} |a_n|^2 = 1\,, \tag{2.4}$$

with complex coefficients a_n. Equivalently, introducing the *dyadic product* $\hat{\mathsf{P}}_{i,j} = |\phi_i\rangle\langle\phi_j|$, defined as

$$\hat{\mathsf{P}}_{i,j}\,|\Psi\rangle = (|\phi_i\rangle\langle\phi_j|)\,|\Psi\rangle = \langle\phi_j\,|\,\Psi\rangle\,|\phi_i\rangle\;, \tag{2.5}$$

we can write the completeness relation

$$\sum_{n=1}^{N} |\phi_n\rangle\langle\phi_n| = \mathbb{1}\,. \tag{2.6}$$

$\hat{\mathsf{P}}_n = |\phi_n\rangle\langle\phi_n|$ is also called the *projector* onto the eigenstate state $|\phi_n\rangle$.

By definition, eigenstates are also linearly independent and thus pair-wise orthogonal, $\langle\phi_i\,|\,\phi_j\rangle = \delta_{ij}$, where δ_{ij} is the Kronecker-Delta. The set of eigenstates $\{|\phi_1\rangle \dots |\phi_N\rangle\}$ is therefore a complete orthonormal basis of the Hilbert space of dimension N. Note that the basis is not unique; any unitary transformation $\hat{\mathsf{U}}$, with $\hat{\mathsf{U}}\hat{\mathsf{U}}^\dagger = \mathbb{1}$, will generate a “rotated” basis.

Having expanded an arbitrary state $|\Psi\rangle$ in the eigenbasis of an operator in Eq. (2.4), we define the *measurement* process according to the Copenhagen interpretation: The “projective” measurement of an operator $\hat{\mathsf{A}}$ in a given basis instantaneously “collapses” the state $|\Psi\rangle$ to one of the operator’s eigenstates $|\phi_n\rangle$. The probability of obtaining any specific eigenstate is given by the absolute-square of the expansion coefficient a_n. This motivates the normalization of the coefficients as $\sum |a_n|^2 = 1$, since the sum of all probabilities must be 1. The result of the measurement (and any subsequent measurement in the same basis) is the eigenvalue associated with $|\phi_n\rangle$.

For a finite and discrete Hilbert space as described above, it is useful to represent states as complex vectors with N components, and the operators as $N \times N$ matrices, relative to a fixed basis $\{|\phi_n\rangle\}$. The entries in the vector representation of $|\Psi\rangle$ are the expansion coefficients a_n of Eq. (2.4). The entries A_{ij} of the operator-matrix are given by

$$A_{ij} = \left\langle \phi_i \,\middle|\, \hat{\mathsf{A}} \,\middle|\, \phi_j \right\rangle . \tag{2.7}$$

2.2 Storing Information in Quantum Systems

Having a discrete and finite Hilbert space raises the possibility of using it to encode information digitally. In classical computing, it has proven most practical to work in a binary representation, using the logical values 0 and 1 (implemented as low and high voltage in an electrical circuit). A single logical value is called a *bit* (binary digit). In analogy, we consider a 2-dimensional Hilbert space and label its basis states $|0\rangle$ and $|1\rangle$. An arbitrary state of this system, called *quantum bit* (qubit), is

$$|\Psi\rangle_{1q} = a_0\,|0\rangle + a_1\,|1\rangle\ , \quad |a_0|^2 + |a_1|^2 = 1\,. \tag{2.8}$$

Whereas the classical bit is in a state of either 0 or 1, the qubit can be in any superposition of $|0\rangle$ and $|1\rangle$.

The theory of quantum mechanics is universal in the sense that it describes *any* quantum system on an abstract level with the same formalism. In practice, the systems can be extremely diverse, from nuclear spins to the electronic states of atoms and molecules, quantum dots, Bose-Einstein-condensates, nano-mechanical devices, photons (free or in optical cavities), or superconducting circuits. Mathematically, these only differ in the structure of their Hamiltonian. From an experimentalist's points of view, of course, they are vastly different. All of these systems are in principle candidates for the implementation of a quantum computer. However, in order to be practically useful for quantum computation, a quantum system must fulfill the five DiVincenzo criteria [74]:

1. a scalable physical system with well characterized qubits,
2. the ability to initialize the state of the qubits,
3. long relevant decoherence times,
4. a "universal" set of quantum gates, and
5. a qubit-specific measurement capability.

Already the first requirement puts severe restrictions on a candidate system. While spin-$\frac{1}{2}$ particles have two possible states (spin-up and spin-down) and thus provide a "natural qubit", this is not the case in general. Most systems have a Hilbert space of dimension larger than 2. In this case, two of the levels must be well-separated from the remainder of the Hilbert space, forming a *logical subspace.* An example for a system that is *not* suitable in this regard is the quantum harmonic oscillator, which has all equidistant energy levels. Thus, no two levels can be addressed separately from the others and could provide a qubit. On the other hand, if there is sufficient anharmonicity in the levels, the system can provide a well-defined qubit. In many cases, the additional levels can be exploited as ancillas to implement quantum information tasks, as we will see in the examples in chapters 4–7. However, if the entire population does not eventually return to the logical subspace, this introduces an error into the quantum operation.

Scalability means that the system must *coherently* support a large number of qubits (at least several hundred). Mathematically, the n-qubit Hilbert space is constructed from the tensor product of the n single-qubit Hilbert spaces,

$$
\begin{aligned}
\left|\Psi\right\rangle_{nq} &= \sum_{b_i=0,1} a_{b_1 b_2 \dots b_n} \left|b_1\right\rangle \otimes \left|b_2\right\rangle \otimes \dots \otimes \left|b_n\right\rangle \\
&= \sum_{b_i=0,1} a_{b_1 b_2 \dots b_n} \left|b_1 b_2 \dots b_n\right\rangle \;,
\end{aligned}
\tag{2.9}
$$

where $\left|b_i\right\rangle$ represents the two basis states $\left|0\right\rangle$, $\left|1\right\rangle$ of the i'th qubit. For example, for a two-qubit state,

$$
\left|\Psi\right\rangle_{2q} = a_{00}\left|00\right\rangle + a_{01}\left|01\right\rangle + a_{10}\left|10\right\rangle + a_{11}\left|11\right\rangle \;. \tag{2.10}
$$

In general, it is *not* possible to write

$$\begin{aligned} \left|\Psi\right\rangle_{2q} &= \left(a_0 \left|0\right\rangle + a_1 \left|1\right\rangle\right) \otimes \left(b_0 \left|0\right\rangle + b_1 \left|1\right\rangle\right) \\ &= \underbrace{a_0 b_0}_{a_{00}} \left|00\right\rangle + \underbrace{a_0 b_1}_{a_{01}} \left|01\right\rangle + \underbrace{a_1 b_0}_{a_{10}} \left|10\right\rangle + \underbrace{a_1 b_1}_{a_{11}} \left|11\right\rangle \,. \end{aligned} \tag{2.11}$$

If Eq. (2.11) holds, $\left|\Psi\right\rangle_{2q}$ is said to be *separable.* Otherwise, $\left|\Psi\right\rangle_{2q}$ is entangled. The entanglement of a given two-qubit state can be quantified through the *concurrence* [75]

$$C\left(\left|\Psi\right\rangle_{2q}\right) = \left|\left\langle \Psi_{2q} \middle| \hat{\sigma}_y \otimes \hat{\sigma}_y \middle| \Psi^*_{2q} \right\rangle\right| \,, \tag{2.12}$$

where $\hat{\sigma}_y$ is the Pauli-y matrix, see Eq. (2.17) below, and $\left|\Psi^*\right\rangle_{2q}$ is the complex conjugate of $\left|\Psi\right\rangle_{2q}$ (in a given vector representation). The concurrence takes values in the range $[0, 1]$, where a state for which $C\left(\left|\Psi\right\rangle_{2q}\right) = 1$ is called *maximally entangled.* The standard example for maximally entangled states are the *Bell states*

$$\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}\left(\left|00\right\rangle + \left|11\right\rangle\right) \,, \qquad \left|\Phi^-\right\rangle = \frac{1}{\sqrt{2}}\left(\left|00\right\rangle - \left|11\right\rangle\right) \,, \tag{2.13}$$

$$\left|\Psi^+\right\rangle = \frac{1}{\sqrt{2}}\left(\left|01\right\rangle + \left|10\right\rangle\right) \,, \qquad \left|\Psi^-\right\rangle = \frac{1}{\sqrt{2}}\left(\left|00\right\rangle - \left|11\right\rangle\right) \,. \tag{2.14}$$

Even though DiVincenzo's criterion 1 requires the system to scale to a large number of qubits, we have focused here on the description of two-qubit systems. The reason for this is that any operation on more than two qubits can be expressed as a series of operations where in each operation, at most two of the qubits interact [76].

Criteria 2 and 5 concern the input-output of a quantum computer. For most implementations, this is less of a fundamental concern than a (possibly difficult) technical challenge. It can be useful to employ the methods of quantum control discussed in chapter 1 to bring the system into a desired initial state. Measurement and detection are primarily an experimentalist's task, although theory can certainly contribute relevant aspects. For example, the ideas on which the results of chapter 7 build have been applied to significantly reduce the number of measurements required for the characterization of unitary processes [77, 78]. Also, for superconducting qubits, optimal control has been applied to the measurement process [79].

On the other hand, the implementation of quantum gates, i.e., a unitary transformation on all the eigenstates of the logical subspace (criterion 4), is where numerical optimization techniques shine, and it is the central focus of the remainder of the thesis. Lastly, the long decoherence times required by criterion 3 are a significant inherent challenge. The effects of decoherence are discussed in sections 2.5–2.7 below; implementing quantum gates in the presence of decoherence is the focus of chapters 4 and 7.

2.3 Gate-Based Quantum Computing

2.3.1 Single-Qubit Gates

While the *encoding* of information only relies on the existence of quantized energy levels in quantum systems, its *processing* implies time-dependence. From a given input state at time zero, a quantum computer applies a series of logical transformations, eventually resulting in a state at time T. Physically, the input state undergoes the time evolution of the system's Hamiltonian. For a system of qubits, we call the resulting time evolution operator $\hat{\mathsf{U}}(T, 0)$ a quantum gate U. For a static Hamiltonian, the time evolution operator

$$\hat{\mathsf{U}}(t, 0) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}t} \tag{2.15}$$

is the solution of the time-dependent Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\left|\Psi\right\rangle = \hat{\mathsf{H}}\left|\Psi\right\rangle \,. \tag{2.16}$$

A quantum gate operating on a single qubit can be any unitary 2 × 2 matrix. Their mathematical properties are best understood with the help of the three Pauli matrices

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \,, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \,, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \,. \tag{2.17}$$

First, the Pauli matrices are important single-qubit quantum gates themselves. In this context, they are also written as X, Y, and Z, respectively. The X gate flips a qubit (analogously to a classical logical NOT operation on a bit), whereas Z creates a π phase shift between $|0\rangle$ and $|1\rangle$.

Even more importantly, though, we can use the Pauli matrices (together with the identity $\mathbb{1}$) to describe any single-qubit Hamiltonian

$$\hat{\mathsf{H}}_{1q} = c_0\mathbb{1} + c_1\hat{\sigma}_x + c_2\hat{\sigma}_y + c_3\hat{\sigma}_z \tag{2.18}$$

with coefficients $c_i \in \mathbb{R}$. The gate U induced by the Hamiltonian $\hat{\mathsf{H}}_{1q}$ acting for a duration t is

$$U = \hat{\mathsf{U}}[t, 0] = \mathrm{e}^{-\mathrm{i}c_0 t}\, \mathrm{e}^{-\mathrm{i}(c_1\hat{\sigma}_x + c_2\hat{\sigma}_y t + c_3\hat{\sigma}_z)t} \,. \tag{2.19}$$

From

$$X = \mathrm{e}^{-\mathrm{i}\frac{\pi}{4}(\hat{\sigma}_x - \mathbb{1})} \,, \tag{2.20}$$

and equivalently for $\hat{\sigma}_y$ and $\hat{\sigma}_z$, we see that the Pauli matrices are their own generators. The contribution by the identity, defined by c_0, only induces a global phase, which is physically irrelevant. Beyond the Pauli matrices,

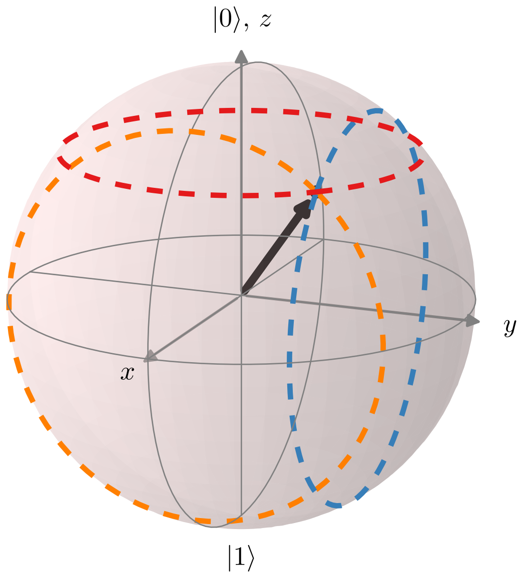


Figure 2.1: Bloch sphere with Bloch vector for the qubit state $|\Psi\rangle = \frac{3}{11}|0\rangle + \frac{1+\mathrm{i}}{11}|1\rangle$. The precession induced by a Hamiltonian proportional to $\hat{\sigma}_x$, $\hat{\sigma}_y$, and $\hat{\sigma}_z$ is indicated by the orange, blue and red circles, respectively

frequently encountered single-qubit gates are the Hadamard gate,

$$H = \frac{1}{\sqrt{2}}(X + Y) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} , \tag{2.21}$$

which brings the logical eigenstates into superposition, and the single-qubit phase gate as a generalization of the Z gate,

$$S_\phi = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\phi} \end{pmatrix} , \tag{2.22}$$

generating a relative phase between $|0\rangle$ and $|1\rangle$.

An arbitrary quantum state $|\Psi\rangle$ can be associated with the *density matrix*

$$\hat{\rho} = |\Psi\rangle\langle\Psi| , \tag{2.23}$$

i.e., the projector onto that specific state. If we expand this matrix into the three Pauli matrices, we obtain a vector $\vec{r} \in \mathbb{R}^3$ that represents the state up to a global phase. This vector is called the *Bloch vector*. We have already encountered the Bloch vector in chapter 1 under the name “Feynman-Vernon-Hellwarth representation”. Since for a properly normalized qubit state, $|\vec{r}| = 1$, the Bloch vector is on the unit sphere, called *Bloch sphere* in this context. Like in chapter 1, $\vec{r}$ precesses around the vector (c_1, c_2, c_3) representing the Hamiltonian, obtained from Eq. (2.19). Specifically, Eq. (2.19) corresponds to the solution of Eq. (1.8). The Bloch sphere is illustrated in Fig 2.1 for an arbitrary qubit state. A term proportional to $\hat{\sigma}_x$ induces a rotation around the x-axis, shown in orange, and correspondingly $\hat{\sigma}_y$ and $\hat{\sigma}_z$ around the y- and z-axes (blue and red, respectively).

2.3.2 Two-Qubit Gates

For a composite system of multiple qubits, the total Hilbert space is the tensor product of the single-qubit Hilbert spaces. Since for a single-qubit state the Pauli matrices (including the identity) form a complete basis according to Eq. (2.18), for a two-qubit Hilbert space, this means that any Hamiltonian is of the form

$$\hat{\mathsf{H}}_{2q} = \sum_{i,j} c_{ij} \hat{\sigma}_i \otimes \hat{\sigma}_j \,, \tag{2.24}$$

generated by the sixteen operators $\hat{\sigma}_i \otimes \hat{\sigma}_j$, for $\hat{\sigma}_{i,j} \in \{\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z, \mathbb{1}\}$, with the expansion coefficients

$$c_{ij} = \frac{1}{4} \operatorname{tr}\left[(\hat{\sigma}_i \otimes \hat{\sigma}_j)^\dagger \hat{\mathsf{H}}_{2q}\right] . \tag{2.25}$$

This is analogous to Eq. (2.10) for states. If there is no interaction between the two qubits, the Hamiltonian for the two-qubit system is

$$\hat{\mathsf{H}}_0 = \hat{\mathsf{H}}_{1q} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathsf{H}}_{1q} \,. \tag{2.26}$$

In matrix representation, for two operators of dimension N, the tensor product is defined as

$$\hat{\mathsf{A}} \otimes \hat{\mathsf{B}} = \begin{pmatrix} a_{11}\hat{\mathsf{B}} & \dots & a_{1N}\hat{\mathsf{B}} \\ \vdots & \ddots & \vdots \\ a_{N1}\hat{\mathsf{B}} & \dots & a_{NN}\hat{\mathsf{B}} \end{pmatrix} . \tag{2.27}$$

The "local" Hamiltonian $\hat{\mathsf{H}}_0$ consists only of terms proportional to the 6 operators $\hat{\sigma}_x^{(1)}$, $\hat{\sigma}_y^{(1)}$, $\hat{\sigma}_z^{(1)}$, $\hat{\sigma}_x^{(2)}$, $\hat{\sigma}_y^{(2)}$, $\hat{\sigma}_z^{(2)}$, and $\mathbb{1}$, where $\hat{\sigma}_i^{(1)} = \hat{\sigma}_i \otimes \mathbb{1}$ acts only on the first qubit, and equivalently $\hat{\sigma}_i^{(2)}$ acts only on the second qubit. Starting from a non-entangled state, evolution under $\hat{\mathsf{H}}_0$ does not yield an entangled state. In order to generate entanglement, there needs to exist a physical interaction between the two qubits, reflected in an interaction Hamiltonian $\hat{\mathsf{H}}_I$,

$$\hat{\mathsf{H}}_{2q} = \hat{\mathsf{H}}_{1q} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathsf{H}}_{1q} + \hat{\mathsf{H}}_I \,, \tag{2.28}$$

where $\hat{\mathsf{H}}_I$ is spanned by the remaining 9 terms $\hat{\sigma}_i \otimes \hat{\sigma}_j$, $i, j = x, y, z$. From hereon, we abbreviate $\hat{\sigma}_i \otimes \hat{\sigma}_j$ as $\hat{\sigma}_i \hat{\sigma}_j$ in the context of two-qubit Hilbert spaces.

Two very common interactions found in physical systems are $\hat{\sigma}_z\hat{\sigma}_z$ and $\hat{\sigma}_x\hat{\sigma}_x + \hat{\sigma}_y\hat{\sigma}_y$. The former means that the interaction between two qubits induces a relative shift on one or more of the energy levels, resulting in a diagonal gate

$$D = \mathrm{diag}\left\{\mathrm{e}^{\mathrm{i}\phi_{00}}, \mathrm{e}^{\mathrm{i}\phi_{01}}, \mathrm{e}^{\mathrm{i}\phi_{10}}, \mathrm{e}^{\mathrm{i}\phi_{11}}\right\}\,, \tag{2.29}$$

where $\phi_i = \Delta E_i t$ is a phase induced by the shift ΔE_i of the level i relative to the evolution under $\hat{\mathsf{H}}_0$. Defining the non-local phase

$$\gamma = \phi_{00} - \phi_{01} - \phi_{10} + \phi_{11}\,, \tag{2.30}$$

this gate induces the entanglement [80, 73]

$$C(D) = \left|\sin\frac{\gamma}{2}\right|\,. \tag{2.31}$$

In general, the entanglement generated by a gate $\hat{\mathsf{U}}$ is defined as the maximum concurrence according to Eq. (2.12) that the state $\hat{\mathsf{U}}\vert\Psi\rangle$ can have, assuming the input state $\vert\Psi\rangle$ is separable [81].

For $\gamma = \pi$, the gate is a perfect entangler, i.e. $C(D)$ takes the maximum value 1 and there exists a separable state that is mapped to a maximally entangled state. The canonical form in this case is the Controlled-Phase (CPHASE) gate,

$$\mathrm{CPHASE} = \mathrm{diag}\{1, 1, 1, -1\}\,. \tag{2.32}$$

In the taxonomy of two-qubit gates, the CPHASE is one in the large class of "controlled-unitary" gates. In this type of gate, the first qubit acts as a "control" qubit. If the control qubit is in the state $\vert 1\rangle$, the single-qubit gate U is performed on the second "target" qubit. If the control qubit is in the state $\vert 0\rangle$, the target qubit remains unchanged. In the case of the CPHASE, U is the single-qubit phase gate $Z = \hat{\sigma}_z$. Probably the most well-known controlled-unitary is the CNOT gate [82],

$$\mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\,. \tag{2.33}$$

In this case, U is the logical-not operation implemented by $X = \hat{\sigma}_x$.

To understand the $\hat{\sigma}_x\hat{\sigma}_x + \hat{\sigma}_y\hat{\sigma}_y$ interaction, we introduce

$$\hat{\sigma}_+ = \hat{\sigma}_x + \mathrm{i}\hat{\sigma}_y\,, \tag{2.34}$$

$$\hat{\sigma}_- = \hat{\sigma}_x - \mathrm{i}\hat{\sigma}_y\,. \tag{2.35}$$

These are the raising and lowering operators on the qubit. The interaction is then written as

$$\hat{\sigma}_x \hat{\sigma}_x + \hat{\sigma}_y \hat{\sigma}_y = \frac{1}{2} \left(\hat{\sigma}_+ \hat{\sigma}_- + \hat{\sigma}_- \hat{\sigma}_+ \right) . \tag{2.36}$$

We see that this can be interpreted as an *exchange interaction*, where an excitation moves from one qubit to the other. This interaction induces the iSWAP gate

$$\text{iSWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{2.37}$$

It is equivalent two consecutive CNOT operations, where the target qubit for the second CNOT is the control qubit from the first CNOT. Therefore, iSWAP is also called the Double-CNOT (DCNOT) gate. A summary of the most relevant two-qubit gates and the Hamiltonians inducing them is given in appendix D.

In a physical implementation not based on actual spins we often have to include additional levels beyond $\vert 0 \rangle$ and $\vert 1 \rangle$ in the description of the system. For a time evolution $\hat{\mathsf{U}}(T, 0)$, the implemented quantum gate is

$$U = \hat{\mathsf{P}} \hat{\mathsf{U}}(T, 0) \, , \tag{2.38}$$

where $\hat{\mathsf{P}}$ is the projector onto the logical subspace,

$$\hat{\mathsf{P}} = \vert 0 \rangle \langle 0 \vert + \vert 1 \rangle \langle 1 \vert \ , \tag{2.39}$$

and equivalently for a two-qubit system. The gate U will only be unitary if at time T, all population returns to the logical subspace. This does not imply that the population must also remain in the logical subspace during the entire duration of the dynamics. In fact, it may be necessary to leave the logical subspace in order to implement a quantum gate. An example is the implementation of a two-qubit gate using trapped Rydberg atoms, as is the focus of chapter 4. The logical subspace is defined by two low-lying electronic states of the atoms. However, the atoms can be excited to a Rydberg state $\vert r \rangle$, i.e., a state with large principal quantum number. When both atoms are in $\vert r \rangle$, they feel a dipole-dipole interaction, shifting the $\vert rr \rangle$ level and thus creating the necessary entanglement for a quantum gate. While the Hamiltonian can no longer be expressed in terms of Pauli matrices, the intuition for the implementation of gates remains: the simple shift of the $\vert rr \rangle$ level can be mapped onto the logical subspace as a $\hat{\sigma}_z \hat{\sigma}_z$ interaction and allows the implementation of a CPHASE.

For superconducting qubits, discussed in chapter 5, the logical subspace consists of the two lowest levels of an anharmonic ladder. The Hamiltonian includes the term $\hat{\mathsf{b}}_1^\dagger \hat{\mathsf{b}}_2 + \hat{\mathsf{b}}_1 \hat{\mathsf{b}}_2^\dagger$, where $\hat{\mathsf{b}}_{1,2}$ is the lowering operator for the first and second (multilevel) qubit and works analogously to the interaction in Eq. (2.36), allowing for the implementation of an iSWAP gate, among others.

2.3.3 Controllability

So far, we have only considered a static interaction Hamiltonian. The situation becomes more interesting if the Hamiltonian includes one or more time-dependent controls $u_i(t)$,

$$\hat{\mathsf{H}} = \hat{\mathsf{H}}_0 + \sum_{i=1}^{m} u_i(t) \hat{\mathsf{H}}_i \,. \tag{2.40}$$

If we restrict ourselves to the logical subspace such that each $\hat{\mathsf{H}}_i$ takes the form of Eq. (2.24), understanding which gates are induced by Eq. (2.40) relies on the theory of dynamical Lie algebras and groups [83]. In practical terms, the procedure for determining which gates can be implemented in a given system, is as follows:

1. Start from the set of independent terms in the Hamiltonian,

$$\mathfrak{L} = \{\hat{\mathsf{H}}_0, \hat{\mathsf{H}}_1, \dots \hat{\mathsf{H}}_m\} \,. \tag{2.41}$$

2. Calculate all commutators of the elements of $\mathfrak{L}$. Extend $\mathfrak{L}$ by those commutators that are linearly independent from the existing elements.

3. Repeat the procedure, building nested commutators, until it yields no further new elements.

The reachable gates are those in the span of the dynamical Lie algebra,

$$\mathrm{e}^{-\mathrm{i}\mathfrak{L}t} \equiv \left\{ \prod_{\hat{\mathsf{A}}_i \in \mathfrak{L}} \mathrm{e}^{-\mathrm{i}\hat{\mathsf{A}}_i t} \right\} \,. \tag{2.42}$$

If the dimension of the dynamical Lie algebra is full, i.e., it contains 16 linearly independent elements for a two-qubit Hamiltonian, the system is *fully controllable*, i.e., every quantum gate can be implemented in principle. An example for such a controllability analysis is given in section 6.2 of chapter 6.

From a control perspective, it is essential to have a measure for whether a unitary evolution $\hat{\mathsf{U}}(T,0)$ implements a specific quantum gate $\hat{\mathsf{O}}$ at time

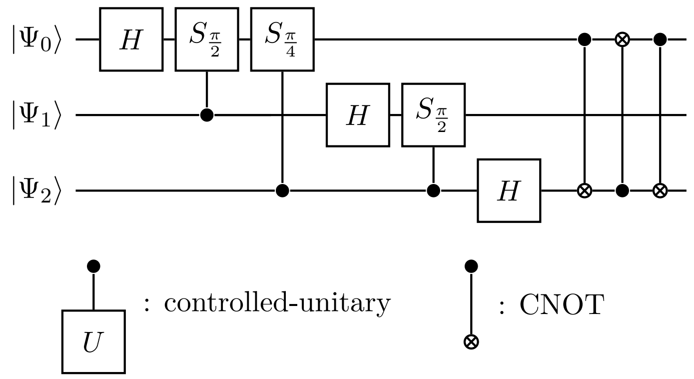


Figure 2.2: Exemplary circuit diagram, showing the implementation of a quantum Fourier transform on three qubits.

T. The gate fidelity in this case is defined as

$$F(\hat{\mathsf{U}}(T,0)) = \frac{1}{N^2}\left|\mathrm{tr}\left[\hat{\mathsf{O}}^\dagger \hat{\mathsf{P}}\hat{\mathsf{U}}(T,0)\right]\right|^2 , \tag{2.43}$$

where $\hat{\mathsf{P}}$ is the projector to the logical subspace. The square-modulus ensures that the gate $\hat{\mathsf{O}}$ only has to be implemented up to a global phase. The gate fidelity is not necessarily the best functional to be used for the purpose of optimization, as will be discussed in chapter 3.

2.3.4 Universal Sets of Gates

Any practical quantum computer will operate on a large number of qubits. In principle, a quantum algorithm could be described as a *black box* unitary transformation acting on the entire Hilbert space $\mathcal{H}^N$ for N qubits. However, this approach would hardly be practical. The logical circuits in a classical computer are decomposed into small logical elements. Indeed, it can be shown that any logical function can be implemented using only the classical NAND gate, taking two bits as input and returning *true* unless both of the inputs are *true*. This greatly simplifies the production of integrated circuits, as only a single logical element needs to be repeated over and over again on the circuit board. In a gate-based quantum computer, we follow the same approach [82]. Unlike in the classical case, there is no single gate that can be used to implement an arbitrary quantum transformation.

However, it has been shown that any quantum transformation can be decomposed into the set of all single-qubit gates, together with one two-qubit quantum gate, most commonly CNOT [84]. This is the motivation to restricting any discussion about the implementation of gates to single- and two-qubit gates. The decomposition of a larger gate into single- and two-qubit gates is commonly in the form of a quantum circuit diagram.

An example is shown in Fig 2.2, giving a decomposition for a quantum-Fourier transform on three qubits. In many physical implementations, single-qubit gates are relatively easy to implement, since they require only local manipulation of a qubit. In contrast, the implementation of an entangling two-qubit gate is usually non-trivial.

Instead of requiring arbitrary single-qubit gates, one can also find a finite set of single-qubit gates that, together with an entangling gate such as CNOT, can implement an arbitrary unitary up to a predefined precision. There are several such possible sets; one example consists of the Hadamard gate H, the phase gate $S_{\frac{\pi}{4}}$, and the CNOT gate. With H and $S_{\frac{\pi}{4}}$ alone, any single-qubit gate can be approximated to arbitrary precision, and in fact with relatively good efficiency: in order to approximate a single-qubit gate with an error of ε, on the order of $O\left[\log^4(1/\varepsilon)\right]$ operations are required [85].

The question of efficiency arises also for the choice of the two-qubit gate included in the universal set. While the use of CNOT is standard, it has been shown that nearly *any* non-local two-qubit gate is universal in combination with single-qubit gates [86, 87], albeit not necessarily with high efficiency. Using the CNOT gate and single-qubit operations, any other two-qubit gate can be implemented using at most three applications of CNOT [88]. However, the

$$\text{B-GATE} = \begin{pmatrix} \cos\frac{\pi}{8} & 0 & 0 & \mathrm{i}\sin\frac{\pi}{8} \\ 0 & \cos\frac{3\pi}{8} & \mathrm{i}\sin\frac{3\pi}{8} & 0 \\ 0 & \mathrm{i}\sin\frac{3\pi}{8} & \cos\frac{3\pi}{8} & 0 \\ \mathrm{i}\sin\frac{\pi}{8} & 0 & 0 & \cos\frac{\pi}{8} \end{pmatrix} \tag{2.44}$$

has been shown [89] to require only two applications for universality.

2.4 Two-Qubit Gates in the Weyl Chamber

Even more important than theoretical considerations about which single-qubit and two-qubit gates are good candidates for efficient universal quantum computing, is the question which gates can easily and with high fidelity be implemented in a given physical implementation. Also, it might be beneficial to include a small number of additional gates for the implementation of a specific algorithm, rather than insisting on a standard set. It is not always obvious which gates are easiest to implement for a given physical system. For the Rydberg Hamiltonian discussed in chapter 4, only diagonal gates are possible, and thus neither CNOT nor the B-GATE could be included in a universal set on this platform.

The theory of local invariants [90, 91] provides a way to address this issue. Under the assumption that single-qubit gates are easily available, we

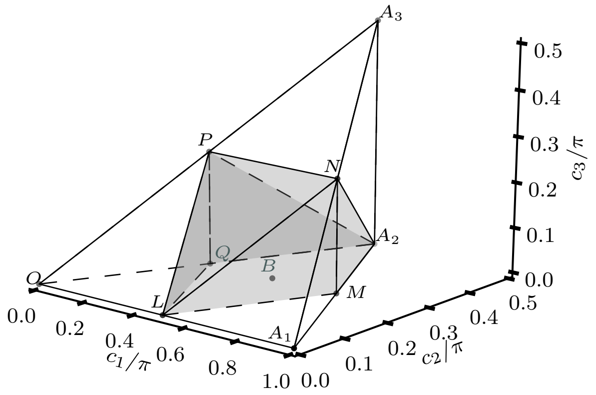


Figure 2.3: Weyl chamber of two-qubit gates. The labeled vertices and midpoints correspond to the equivalence class of prominent gates, with O and A_2 for the identity (i.e., all local gates), L for [CNOT/CPHASE], Q and M for [$\sqrt{\text{iSWAP}}$], P and N for two different [$\sqrt{\text{SWAP}}$], A_2 for [iSWAP/DCNOT], B for [B-GATE] and A_3 for [SWAP]. The shaded polyhedron at the center contains all gates that are perfect entanglers, cf. appendix D.

characterize a quantum gate only by its nonlocal component. For example, a CNOT gate can easily be obtained from a CPHASE using only two additional Hadamard gates,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \left(\mathbb{1} \otimes \hat{\mathsf{H}}\right) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \left(\mathbb{1} \otimes \hat{\mathsf{H}}\right) . \tag{2.45}$$

In this sense, CPHASE and CNOT are *locally equivalent*, i.e., they only differ by local operations. All gates fulfilling this property are considered to be in a single local equivalence class. We denote the equivalence class as e.g. [CNOT].

The notion of the local equivalence class is based on the Cartan decomposition of the general dynamical Lie group of the two-qubit Hilbert space, which yields that any two-qubit gate can be written (up to a global phase) as [91]

$$\hat{\mathsf{U}} = \hat{\mathsf{k}}_1 \exp\left[\frac{\mathrm{i}}{2}\left(c_1 \hat{\sigma}_x \hat{\sigma}_x + c_2 \hat{\sigma}_y \hat{\sigma}_y + c_3 \hat{\sigma}_z \hat{\sigma}_z\right)\right] \hat{\mathsf{k}}_2 \,, \tag{2.46}$$

for real coefficients (c_1, c_2, c_3), where $\hat{\mathsf{k}}_1$ and $\hat{\mathsf{k}}_2$ contain only local operations. Every point in $\mathbb{R}^3$ given by the coordinates (c_1, c_2, c_3) represents a local equivalence class. Removing symmetries in the coefficients results in the geometric shape shown in Fig. 2.3, named the *Weyl chamber*. There is one remaining mirror symmetry through the L–A_2 line on the ground surface of the Weyl chamber, such that e.g. O and A_1, and Q and M are associated with the same local equivalence class. Otherwise, every point in the Weyl chamber uniquely corresponds to a set of gates that differ only by local transformations. The c_1 axis describes the diagonal gates in Eq. (2.29), with $c_1 = \gamma$ – or, more generally, all controlled-unitaries. The gates at any point except O, A_2 (local gates), and A_3 (gates equivalent to the SWAP gate) create non-zero entanglement. Most two-qubit gates, those indicated by the shaded polyhedron in Fig. 2.3 are perfect entanglers [92]. Based on the Weyl chamber coordinates (c_1, c_2, c_3), the *local invariants* (g_1, g_2, g_3) can be derived as

$$\begin{aligned} g_1 &= \frac{1}{4}\big[\cos(2c_1) + \cos(2c_2) + \cos(2c_3) + \\ &\qquad + \cos(2c_1)\cos(2c_2)\cos(2c_3)\big]\,, && (2.47a)\\ g_2 &= \frac{1}{4}\sin(2c_1)\sin(2c_2)\sin(2c_3)\;, && (2.47b)\\ g_3 &= \cos(2c_1) + \cos(2c_2) + \cos(2c_3)\;, && (2.47c) \end{aligned}$$

uniquely identifying a gate's local equivalence class. That is, two gates with the same values (g_1, g_2, g_3) differ only by single-qubit operations.

The local invariants have the additional benefit that they can be calculated analytically directly from the gate $\hat{\mathsf{U}}$, whereas the procedure for determining the Weyl chamber coordinates [93] requires numerical diagonalization and branch-selection of the complex logarithm. As an alternative to Eq. (2.47), the local invariants may also be written as

$$\begin{aligned} g_1 &= \frac{1}{16}\mathfrak{Re}\left\{\mathrm{tr}^2(\hat{\mathfrak{m}})\right\}\;, && (2.48a)\\ g_2 &= \frac{1}{16}\mathfrak{Im}\left\{\mathrm{tr}^2(\hat{\mathfrak{m}})\right\}\;, && (2.48b)\\ g_3 &= \frac{1}{4}\left[\mathrm{tr}^2(\hat{\mathfrak{m}}) - \mathrm{tr}\left(\hat{\mathfrak{m}}^2\right)\right]\;, && (2.48c) \end{aligned}$$

where

$$\hat{\mathfrak{m}} = \hat{\mathsf{U}}_B^T \hat{\mathsf{U}}_B\,, \qquad (2.49)$$

and

$$\hat{\mathsf{U}}_B = \hat{\mathsf{Q}}\hat{\mathsf{U}}\hat{\mathsf{Q}}^\dagger \qquad (2.50)$$

is the gate written in the *magic Bell basis*,

$$\hat{\mathsf{Q}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & \mathrm{i} \\ 0 & \mathrm{i} & 1 & 0 \\ 0 & \mathrm{i} & -1 & 0 \\ 1 & 0 & 0 & -\mathrm{i} \end{pmatrix}. \qquad (2.51)$$

2.5 Density Matrix Formalism

In the context of the Bloch sphere, we have already introduced the *pure state density operator*,

$$\hat{\rho} = |\Psi\rangle\langle\Psi| \tag{2.52}$$

for an arbitrary quantum state $|\Psi\rangle$. When $|\Psi\rangle$ is written as a Hilbert space vector with respect to a fixed basis $\{|\phi_i\rangle\}$, we obtain the *density matrix* representation in that basis, with

$$\rho_{ij} = \langle\phi_i\,|\,\hat{\rho}\,|\,\phi_j\rangle\ . \tag{2.53}$$

The density matrix is completely equivalent to the state space vector in Hilbert space. Via a simple application of the chain rule in the time-dependent Schrödinger equation (2.16), the equation of motion of the density matrix is derived as the Liouville-von Neumann equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\rho} = \left[\hat{\mathsf{H}}, \hat{\rho}\right]\ . \tag{2.54}$$

Furthermore, the expectation value of any Hermitian Hilbert space operator $\hat{\mathsf{A}}$ is

$$\left\langle\hat{\mathsf{A}}\right\rangle = \operatorname{tr}\left[\hat{\mathsf{A}}\hat{\rho}\right]\ . \tag{2.55}$$

The measurement process exposes the probabilistic aspect of quantum mechanics. For example, a measurement in the canonical basis of the state $|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ of a two-level system yields the statistical ensemble $\{(\frac{1}{2}, |0\rangle), (\frac{1}{2}, |1\rangle)\}$. That is, after the measurement the system is in the state $|0\rangle$ *or* in the state $|1\rangle$, each with 50% probability. The concept of the density matrix can be extended to describe the state of the system after the measurement. For a statistical ensemble of N quantum states $|\Psi_i\rangle$, $i \in [1, N]$, each occurring with probability p_i, the *mixed state density matrix* is defined as

$$\hat{\rho} = \sum_{i=1}^{N} p_i\,|\Psi_i\rangle\langle\Psi_i|\ . \tag{2.56}$$

Given an arbitrary density matrix $\hat{\rho}$, the expectation value for the result of a population measurement with regard to a given state $|\Psi\rangle$ is

$$p = \langle\Psi\,|\,\hat{\rho}\,|\,\Psi\rangle\ . \tag{2.57}$$

This probability for finding the system in the state $|\Psi\rangle$ now combines both the *inherent* quantum mechanical probability due to the projective measurement, and as well as the simple (classical) lack of knowledge of the quantum state, e.g. through previous non-recorded measurements of the system.

The generalized density matrix has the following properties:

1. $\hat{\rho}$ is Hermitian and positive-semidefinite.
2. $\operatorname{tr}[\hat{\rho}] = 1$, i.e., the total population is normalized to 1
3. $\frac{1}{N} \leq \operatorname{tr}\left[\hat{\rho}^2\right] \leq 1$, where N is the dimension of the Hilbert space. A purity of $\frac{1}{N}$ is obtained for the completely mixed state $\hat{\rho} = \frac{1}{N}\mathbb{1}$, whereas a purity of 1 is obtained for any pure state $\hat{\rho} = |\Psi\rangle\langle\Psi|$.

With respect to a given basis $\{|\phi_i\rangle\}$, the diagonal elements of the density matrix give the population in each of the basis states. The off-diagonal elements ρ_{ij} are called *coherences*; they express the phase relationship between the basis states $|\phi_i\rangle$ and $|\phi_j\rangle$, assuming both states have non-zero population. When a coherence is zero, this indicates an *incoherent* superposition of those two basis states.

Mathematically, the density matrices are the elements of *Liouville space.* The inner product of Liouville space is defined via the Hilbert-Schmidt product

$$(\hat{\rho}_1, \hat{\rho}_2) = \operatorname{tr}[\hat{\rho}_1^\dagger \hat{\rho}_2] = \operatorname{tr}[\hat{\rho}_1 \hat{\rho}_2]\,. \tag{2.58}$$

Thus, the proper mathematical (“induced”) norm of a density matrix is

$$\|\hat{\rho}\| = \sqrt{\operatorname{tr}\hat{\rho}^2}\,. \tag{2.59}$$

Since $\|\hat{\rho}\|^2$ is identical to the purity, only pure states have norm one. This sometimes causes confusion, as it is common to consider $\operatorname{tr}[\hat{\rho}] = 1$ as the “normalization condition” for a density matrix. However, this refers to the normalization of probabilities, not the norm of the state. A basis of Liouville space is given by all the dyadic products $\{|\phi_i\rangle\langle\phi_j|\}$ for the eigenstates $\{|\phi_i\rangle\}$ of the associated Hilbert space. The dimension of Liouville space is therefore N^2 if N is the dimension of the Hilbert space. Note, however, that the matrices $|\phi_i\rangle\langle\phi_j|$ for $i \neq j$ are not themselves density matrices, since they are not Hermitian.

For a composite system in the Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ where N_A and N_B is the dimension of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, an arbitrary density matrix takes the form

$$\hat{\rho} = \sum_{i,i'=1}^{N_A} \sum_{j,j'=1}^{N_B} \rho_{ij,i'j'} \, |ij\rangle\langle i'j'| \,, \tag{2.60}$$

where $\{|i\rangle\}$ and $\{|j\rangle\}$ are a basis of $\mathcal{H}_A$ and $\mathcal{H}_B$. The density matrix formalism allows to describe the state of only subsystem $\mathcal{H}_A$ as the reduced

$$\begin{tikzcd}
\hat{\rho}_{SE}(0) = \hat{\rho}_S(0) \otimes \hat{\rho}_E \arrow[r, "\text{unitary evolution}"] \arrow[d, "\mathrm{tr}_E"'] & \hat{\rho}_{SE}(t) = \hat{\mathsf{U}}(t,0)\, \hat{\rho}_{SE}(0)\, \hat{\mathsf{U}}^{\dagger}(t,0) \arrow[d, "\mathrm{tr}_E"'] \\
\hat{\rho}_S(0) \arrow[r, "\text{dynamical map}"] & \hat{\rho}_S(t) = \mathcal{E}(t,0)\hat{\rho}_S(0)
\end{tikzcd}$$

Figure 2.4: Diagram showing how the dynamical map $\mathcal{E}(t,0)$ is obtained by tracing out the environment from the unitary evolution of the total state. Adapted from [64].

density matrix

$$\hat{\rho}_A = \mathrm{tr}_B\left[\hat{\rho}\right] = \sum_{i,i'=1}^{N_A} \underbrace{\left[\sum_{j,j'=1}^{N_B} \sum_{j''=1}^{N_B} \rho_{ij,i'j'} \left\langle j'' \,\middle|\, j \right\rangle \left\langle j' \,\middle|\, j'' \right\rangle\right]}_{(\hat{\rho}_A)_{i,i'}} \left|i\right\rangle\left\langle i'\right| . \tag{2.61}$$

The result of this partial trace is a state where all knowledge of subsystem B has been erased. The resulting density matrix $\hat{\rho}_A$ will have purity one only if the subsystems A and B in $\hat{\rho}$ were not entangled.

Typically, we are interested in some relatively small quantum system (a qubit, for example), which is not completely isolated from its environment. While the qubit and the environment are initially in the separable state $\hat{\rho}_S \otimes \hat{\rho}_E$ (otherwise, the qubit would not be considered well-defined as required by the DiVincenzo criteria), the non-zero interaction between them will lead to entanglement between system and environment at later times t. We consider the qubit in this case to be an *open quantum system*.

In principle, we could simply extend the model to include the environment. In practice, this is usually not feasible due to the large or infinite number of degrees of freedom of the environment. However, at least formally, the time evolution in this composite Hilbert space is unitary with Eq. (2.54), and the state of the qubit could be obtained by taking a partial trace over the environment, giving the equation of motion

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \hat{\rho}_S = \mathrm{tr}_E\left[\hat{\mathsf{H}}, \hat{\rho}\right] . \tag{2.62}$$

The solution of Eq. (2.62) is called the *dynamical map*,

$$\hat{\rho}_S(t) = \mathcal{E}(t,0)\hat{\rho}_S(0) = \mathrm{tr}_E\left[\hat{\mathsf{U}}(t,0)\, \hat{\rho}_{SE}(0)\, \hat{\mathsf{U}}^{\dagger}(t,0)\right] . \tag{2.63}$$

The relationship between the unitary evolution in the composite system-environment Hilbert space and the dynamical map is illustrated in Fig. 2.4.

Since tracing out the environment can bring the system to a mixed state, the interaction with the environment is the source of decoherence on the system.

2.6 Master Equation in Lindblad Form

While the approach of including the environment in the total Hilbert space is formally correct, an accurate model often has to include so many degrees of freedom of the environment that an exact treatment of the full system-environment-dynamics is intractable. Therefore, we must find an *effective* description of the system dynamics by itself, i.e., determine an explicit expression for the dynamical map $\mathcal{E}(t,0)$ that only includes operators acting on $\mathcal{H}_S$.

2.6.1 Kraus Operator Representation

From Eq. (2.63), expanding the state of the environment as

$$\hat{\rho}_E = \sum_{\beta=1}^{N_E} \lambda_\beta \left|\phi_\beta\right\rangle \left\langle\phi_\beta\right| \,, \tag{2.64}$$

where $\{|\phi_\beta\rangle\}$ is a complete basis of $\mathcal{H}_E$, and writing out the partial trace yields

$$\begin{aligned} \hat{\rho}_S(t) &= \sum_\alpha \left\langle \phi_\alpha \middle| \hat{\mathsf{U}}(t,0) \left(\hat{\rho}_S(0) \otimes \sum_\beta \lambda_\beta \left|\phi_\beta\right\rangle \left\langle\phi_\beta\right| \right) \hat{\mathsf{U}}^\dagger(t,0) \middle| \phi_\alpha \right\rangle \\ &= \sum_{\alpha\beta} \underbrace{\sqrt{\lambda_\beta} \left\langle \phi_\alpha \middle| \hat{\mathsf{U}}(t,0) \middle| \phi_\beta \right\rangle}_{\hat{\mathsf{K}}_i(t)} \hat{\rho}_S(0) \underbrace{\sqrt{\lambda_\beta}^{\,*} \left\langle \phi_\beta \middle| \hat{\mathsf{U}}(t,0)^\dagger \middle| \phi_\alpha \right\rangle}_{\hat{\mathsf{K}}_i(t)^\dagger} \,. \end{aligned} \tag{2.65}$$

The resulting expression

$$\mathcal{E}(t,0)\hat{\rho}(0) = \sum_{i=1}^{N_E^2} \hat{\mathsf{K}}_i(t) \hat{\rho}_S \hat{\mathsf{K}}_i^\dagger(t) \tag{2.66}$$

for the dynamical map is a *Kraus decomposition*, the operators $\hat{\mathsf{K}}_i$ acting on the elements of $\mathcal{H}_S$ are the *Kraus operators*. The only assumption has been that system and environment are initially separable. If the Kraus operators fulfill the condition $\sum_i \hat{\mathsf{K}}_i^\dagger \hat{\mathsf{K}} = \mathbb{1}$, the resulting $\hat{\rho}(t)$ has trace 1, as required for a proper density matrix. The specific Kraus operators defined in Eq. (2.65) are not the only choice of Kraus operators describing the dynamical map. Any unitary transformation of the Kraus operators yields another set of Kraus operators [5]. The minimum number $1 \leq K \leq N_S^2$ of non-zero Kraus operators that can be found this way is called the *Kraus rank* of the dynamical map.

2.6.2 The Quantum Dynamical Semigroup

If the action of the dynamical map is independent from any previous evolution of the total Hilbert space, a dynamical map where the time interval t is split into t_1 and t_2 can be written as

$$\mathcal{E}(t_1 + t_2, 0) = \mathcal{E}(t_2, 0)\mathcal{E}(t_1, 0) \tag{2.67}$$

This independence property is a strong assumption. Whether it is justified depends very much on the specific structure of the environment and the system-environment interaction. An immediate implication is that the dynamical map is "contracting", i.e. the distance (distinguishability) between two initial state never increases. A general dynamical map does not have an inverse if it induces decoherence. Together with Eq. (2.67), this can be shown to imply [64] that $\{\mathcal{E}(t, 0)\}$ has the mathematical structure of a "continuous one-parameter semi-group".

It follows from the semi-group structure [64] that the dynamical map can be written in the form

$$\mathcal{E}(t, 0) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\mathcal{L}t} \,, \tag{2.68}$$

where the factor $\frac{-\mathrm{i}}{\hbar}$ has been factored out of $\mathcal{L}$ in order to yield the same structure as the time evolution operator in Hilbert space in Eq. (2.15). The equation of motion for the state of the system is then

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\rho}_S(t) = \mathcal{L}\left[\hat{\rho}_S(t)\right] \equiv \lim_{t\to 0}\frac{1}{t}\left(\mathcal{E}(t, 0)\hat{\rho}_S - \hat{\rho}_S\right) \,. \tag{2.69}$$

We choose a complete basis $\{\hat{\mathsf{F}}_\mathrm{i}\}$ of N_S^2 operators of dimension N_S, acting on the elements of $\mathcal{H}_S$, such that $\mathrm{tr}\left[\hat{\mathsf{F}}_{N_S^2}\right] = 1$ and $\mathrm{tr}\left[\hat{\mathsf{F}}_i\right] = 0$ for $i < N_S^2$. These operators can be expanded in the set of Kraus operators, Eq. (2.66). After some algebraic manipulation [64], the equation of motion can be shown to take the form

$$\mathcal{L}\left[\hat{\rho}_S\right] = \left[\hat{\mathsf{H}}, \hat{\rho}_S\right] + \mathrm{i}\sum_{i,j=1}^{N_S^2-1} a_{ij}\left(\hat{\mathsf{F}}_i\hat{\rho}_S\hat{\mathsf{F}}_j^\dagger - \frac{1}{2}\left\{\hat{\mathsf{F}}_j^\dagger\hat{\mathsf{F}}_i, \hat{\rho}_S\right\}\right) \,, \tag{2.70}$$

where $\{\cdot, \cdot\}$ denotes the anti-commutator, and $\hat{\mathsf{H}}$ is a particular Hermitian operator constructed from $\{\hat{\mathsf{F}}_i\}$. As a last step, one can once more rotate the operator basis $\{\hat{\mathsf{F}}_i\}$ to a new set $\{\hat{\mathsf{A}}_k\}$ such that the coefficient matrix a_{ij}

in Eq. (2.70) is diagonalized with eigenvalues γ_k, and arrive at the master equation in Lindblad form [94, 95],

$$\mathcal{L}\left[\hat{\rho}_S\right] = \left[\hat{\mathsf{H}}, \hat{\rho}_S\right] + \mathrm{i}\sum_k \gamma_k \left(\hat{\mathsf{A}}_k \hat{\rho}_S \hat{\mathsf{A}}_k^\dagger - \frac{1}{2}\left\{\hat{\mathsf{A}}_k^\dagger \hat{\mathsf{A}}_k, \hat{\rho}_S\right\}\right) . \tag{2.71}$$

The operators $\{\hat{\mathsf{A}}_k\}$ are the *Lindblad operators* describing the dissipative process. In comparison with the unitary Liouville-von-Neumann equation (2.54), we see that we have obtained an additional *dissipator* $\mathcal{D}$ that captures the non-unitary effects of the system-environment interaction,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\rho}_S = \mathcal{L}\left[\hat{\rho}_S\right] = \left[\hat{\mathsf{H}}, \hat{\rho}_S\right] + \mathrm{i}\mathcal{D}\left[\hat{\rho}_S\right] . \tag{2.72}$$

Each Lindblad operator $\hat{\mathsf{A}}_k$ can be interpreted as a *decoherence channel*, and the associated γ_k is the *rate* at which this channel acts. It is common to absorb γ_k in the Lindblad operators, and to write the dissipator as either

$$\mathcal{D}\left[\hat{\rho}\right] = \sum_k \hat{\mathsf{A}}_{\gamma_k} \hat{\rho}_S \hat{\mathsf{A}}_{\gamma_k}^\dagger - \frac{1}{2}\left(\hat{\mathsf{A}}_{\gamma_k}^\dagger \hat{\mathsf{A}}_{\gamma_k} \hat{\rho}_S + \hat{\rho}_S \hat{\mathsf{A}}_{\gamma_k}^\dagger \hat{\mathsf{A}}_{\gamma_k}\right) \tag{2.73}$$

with $\hat{\mathsf{A}}_{\gamma_k} = \sqrt{\gamma_k}\hat{\mathsf{A}}_k$, or

$$\mathcal{D}\left[\hat{\rho}\right] = \sum_k 2\hat{\mathsf{A}}_{\frac{\gamma_k}{2}} \hat{\rho}_S \hat{\mathsf{A}}_{\frac{\gamma_k}{2}}^\dagger - \hat{\mathsf{A}}_{\frac{\gamma_k}{2}}^\dagger \hat{\mathsf{A}}_{\frac{\gamma_k}{2}} \hat{\rho}_S - \hat{\rho}_S \hat{\mathsf{A}}_{\frac{\gamma_k}{2}}^\dagger \hat{\mathsf{A}}_{\frac{\gamma_k}{2}} \tag{2.74}$$

with $\hat{\mathsf{A}}_{\frac{\gamma_k}{2}} = \sqrt{\frac{\gamma_k}{2}}\hat{\mathsf{A}}_k$.

2.6.3 Decay and Dephasing

While the derivation of the master equation from the semi-group properties of the dynamical map yields the general structure of the dissipator, it does not provide any practical way of obtaining the Lindblad operators $\hat{\mathsf{A}}_i$. However, we can adopt a phenomenological perspective to determine the $\hat{\mathsf{A}}_i$ that are associated with decoherence channels. In a system with the eigenbasis $\{|i\rangle\}$, we consider the dissipator in Eq. (2.73) and two fundamental decoherence processes:

- $\hat{\mathsf{A}}_{\gamma_1} = \sqrt{\gamma_1}\,|0\rangle\langle 1|$. This is a *jump operator* from level $|1\rangle$ to $|0\rangle$, representing a decrease in the energy of the system.
- $\hat{\mathsf{A}}_{\gamma_2^*} = \sqrt{2\gamma_2^*}\,|1\rangle\langle 1|$. The time evolution under this projector does not change the energy of the system, but induces a phase on the level $|1\rangle$.

Inserting $\hat{\mathsf{A}}_{\gamma_1}$ into Eq. (2.71) with $\hat{\mathsf{H}} = 0$ and solving the resulting set of coupled differential equations for all entries of $\hat{\rho}(t) = \hat{\rho}_S(t)$ yields

$$\rho_{00}(t) = \rho_{00}(0) + \left(1 - \mathrm{e}^{-\gamma_1 t}\right) \rho_{11}(t)\,, \tag{2.75a}$$

$$\rho_{11}(t) = \rho_{11}(0)\; \mathrm{e}^{-\gamma_1 t}\,, \tag{2.75b}$$

$$\rho_{1,j\neq 1}(t) = \rho_{1j}(0)\; \mathrm{e}^{-\frac{\gamma_1}{2} t}\,, \tag{2.75c}$$

$$\rho_{i\neq 1,1}(t) = \rho_{i1}(0)\; \mathrm{e}^{-\frac{\gamma_1}{2} t}\,, \tag{2.75d}$$

with all other entries remaining constant. Thus, the channel results in an exponential decay of population from level $|1\rangle$ to level $|0\rangle$ with a decay rate of γ_1. This *incoherent* population transfer is associated with the loss of phase information, described by Eqs. (2.75c, 2.75d), the decay-induced dephasing.

Similarly, for $\hat{\mathsf{A}}_{\gamma_2}$, we obtain

$$\rho_{1,j\neq 1}(t) = \rho_{1j}(0)\; \mathrm{e}^{-\gamma_2 t}\,, \tag{2.76a}$$

$$\rho_{i\neq 1,1}(t) = \rho_{i1}(0)\; \mathrm{e}^{-\gamma_2 t}\,. \tag{2.76b}$$

i.e., an exponential decrease in all coherences related to level $|1\rangle$, without any effect in the populations. We therefore label this dissipative channel *pure dephasing.*

For the example of a two-level system starting in the initial state $|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)$, the time evolution for the combination of both of these channels reads

$$\hat{\rho}(t) = \frac{1}{2}\begin{pmatrix} 2 - \mathrm{e}^{-\gamma_1 t} & \mathrm{e}^{-\left(\frac{\gamma_1}{2}+\gamma_2\right)t} \\ \mathrm{e}^{-\left(\frac{\gamma_1}{2}+\gamma_2\right)t} & \mathrm{e}^{-\gamma_1 t} \end{pmatrix}\,. \tag{2.77}$$

Borrowing from the terminology of nuclear magnetic resonance (NMR), we introduce the longitudinal and transversal relaxation times T_1 and T_2 as the characteristic time of the exponential decline of the population and coherences, respectively [96],

$$T_1 = \frac{1}{\gamma_1}, \qquad T_2 = \frac{1}{\frac{\gamma_1}{2} + \gamma_2}\,. \tag{2.78}$$

For a spin in the Bloch sphere, see Fig 2.1, T_1 is measured from the decay of the component of the spin in the direction of the magnetic field (in the direction of the z-axis), whereas T_2 is measured as the decay of the spin-component perpendicular to the magnetic field. With the pure-dephasing time

$$T_2^* = \frac{1}{\gamma_2}\,, \tag{2.79}$$

the relationship

$$\frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{T_2^*} \tag{2.80}$$

holds. Even outside of NMR, it is extremely common to find the decoherence of a quantum system characterized in terms of T_1 and T_2^* times.

In the context of this thesis, the terms “decoherence” and “dissipation” are used interchangeably to refer to any process that lowers the purity of a quantum state. Especially in NMR, a distinction between “decoherence” and “dissipation” is sometimes made, where the term “dissipation” indicates that energy is exchanged with environment, whereas “decoherence” strictly refers to the loss of phase information.

2.6.4 A Microscopic View

The master equation in Lindblad form has been derived using only the mathematical property of the dynamical semigroup defined in Eq. (2.67). What does the requirement that the dynamical maps over two time intervals are independent mean physically? Is there a general method for obtaining the Lindblad operators? These questions can be answered by a rigorous derivation of the equation of motion from a microscopic model of the system and the environment. We consider the full Hamiltonian $\hat{\mathsf{H}}$ acting on the elements of $\mathcal{H}_{SE}$,

$$\hat{\mathsf{H}} = \hat{\mathsf{H}}_S + \hat{\mathsf{H}}_E + \hat{\mathsf{H}}_{SE}\,, \tag{2.81}$$

where $\hat{\mathsf{H}}_S$ is the Hamiltonian of only the system, $\hat{\mathsf{H}}_E$ is the Hamiltonian of only the environment, and $\hat{\mathsf{H}}_{SE}$ models the system-environment interaction. Based on the definition of the dynamical map in Eq. (2.63) and a series of approximations, a master equation of Lindblad form (2.71) for $\hat{\rho}_S(t)$ can be obtained [64, 97]. A review of the steps typically taken in that derivation illuminates the underlying physical assumptions of the master equation in Lindblad form:

1. *Separability.* The system and the environment must initially be in a separable state, $\hat{\rho}(0) = \hat{\rho}_S(0) \otimes \hat{\rho}_E(0)$.

2. *Weak system-environment coupling.* A weak coupling allows for a perturbative approach, where the Liouville-von Neumann equation (2.54) is integrated once and inserted into Eq. (2.62) to give the first order equation (in the interaction picture)

$$\frac{\partial}{\partial t}\hat{\rho}_S(t) = -\int_0^t \mathrm{tr}_E\left[\hat{\mathsf{H}}_{SE}(t), \left[\hat{\mathsf{H}}_{SE}(t'), \hat{\rho}_{SE}(t')\right]\right] \mathrm{d}t'\,. \tag{2.82}$$

3. *Born approximation.* As a further consequence of the weak system-environment coupling, and if the environment is much larger than the system, the environment will remain in its original state. The total

state is separable at all times,

$$\hat{\rho}_{SE}(t) \approx \hat{\rho}_S(t) \otimes \hat{\rho}_E \,, \qquad \hat{\rho}_E \equiv \hat{\rho}_E(0) \,. \tag{2.83}$$

4. *Time locality.* Under the assumption that the environment retains no memory, it is valid to say that past states have no influence on the dynamics and set $\hat{\rho}_S(t') = \hat{\rho}_S(t)$ in Eq. (2.82). The equation of motion obtained thereby is known as the *Redfield equation.* It is possible to derive a master equation of a very similar form to Eq. (2.71), but with time-dependent and possibly negative decay rates [97].

5. *Markov approximation.* Correlations in the environment are assumed to decay much faster than the time scale on which the system evolves. Mathematically, this allows to take the limit of the integration to infinity. This is a further “no memory” effect. In combination, all of the above approximations yield the *Born-Markov* master equation,

   $$\frac{\partial}{\partial t}\hat{\rho}_S(t) = -\int_0^{\infty} \mathrm{tr}_E \left[\hat{\mathsf{H}}_{SE}(t), \left[\hat{\mathsf{H}}_{SE}(t-s), \hat{\rho}_S(t) \otimes \hat{\rho}_E\right]\right] \mathrm{d}s \,, \tag{2.84}$$

   with $s \equiv t - t'$.

6. *Secular approximation.* In order to arrive from Eq. (2.84) to the master equation in Lindblad form, fast-rotating terms in the interaction frame [64] must be negligible, in the sense of the rotating wave approximation outlined in appendix B.

A system for which the Born-Markov approximation is particularly well fulfilled is an atom interacting with a quantized electromagnetic field [64, 98]. For the specific example of a two-level system, the system and environment are modeled as [97]

$$\hat{\mathsf{H}}_S = \frac{\hbar\omega_0}{2}\hat{\sigma}_z \,, \qquad \hat{\mathsf{H}}_E = \sum_k \hbar\omega_k \hat{\mathsf{b}}_k^\dagger \hat{\mathsf{b}}_k \,, \tag{2.85}$$

and the interaction between system and environment as

$$\hat{\mathsf{H}}_{SE} = \sum_k \left(g_k \hat{\mathsf{b}}_k + g_k \hat{\mathsf{b}}_k^\dagger\right)\left(\hat{\sigma}_+ + \hat{\sigma}_-\right) \,, \tag{2.86}$$

where $\hat{\sigma}_z$, $\hat{\sigma}_+$, and $\hat{\sigma}_-$ are the Pauli matrices defined in Eq. (2.17) and Eq. (2.35), ω_0 is the energy spacing in the two-level system, ω_k is the energy of mode k of the electromagnetic field, and g_k is the coupling strength between the atom and the environment. The derivation of the master equation in Lindblad form, using the approximations listed above, yields [64]

$$\begin{aligned}\frac{\partial}{\partial t}\hat{\rho}_S(t) = \gamma_0(M+1)\left(\hat{\sigma}_-\hat{\rho}_S(t)\hat{\sigma}_+ - \frac{1}{2}\hat{\sigma}_+\hat{\sigma}_-\hat{\rho}_S(t) - \frac{1}{2}\hat{\rho}_S(t)\hat{\sigma}_+\hat{\sigma}_-\right) + \\ + \gamma_0 M\left(\hat{\sigma}_+\hat{\rho}_S(t)\hat{\sigma}_- - \frac{1}{2}\hat{\sigma}_-\hat{\sigma}_+\hat{\rho}_S(t) - \frac{1}{2}\hat{\rho}_S(t)\hat{\sigma}_-\hat{\sigma}_+\right),\end{aligned} \tag{2.87}$$

with the decoherence rate γ_0 and M being the expectation value for the number of photons in the field. At non-zero temperature, $M > 0$, and Eq. (2.87) describes the processes of spontaneous emission, stimulated emission, and absorption. At zero temperature, $M = 0$, only spontaneous emission remains, taking a form identical to the phenomenological decay in section 2.6.3.

All dissipation in the following chapters will be assumed to be Markovian, allowing the dynamics to be modeled by Eq. (2.71). This does not imply that non-Markovian system-bath interactions do not also commonly occur. Markovianity implies that information that has passed from the system to the environment is lost, whereas in a non-Markovian setting, there may be a backflow of information [99, 100], opening up additional potential for control. However, the treatment of non-Markovian dynamics is beyond the scope of this thesis; we will briefly return to it in the outlook in chapter 8.

2.7 Implementing Quantum Gates in Open Quantum Systems

For some quantum control tasks, the presence of decoherence is not detrimental and can even be exploited. For example, for a number of biological processes such as photosynthesis and the "chemical compass" in the retina of birds, the role of quantum effects have been demonstrated, with some evidence that these processes are assisted by noise [101]. Another typical example is cooling, where dissipation is used to remove entropy from the system to the environment [102]. A similar concept is employed for the so-called "quantum governor" [66], where the system is stabilized by transferring fluctuations to the environment.

For the implementation of quantum gates, however, the presence of (Markovian) decoherence is always pernicious, since by definition, the desired process is unitary. The goal is therefore to avoid dissipative processes as much as possible. If suitable symmetries are present in the system, it may be possible to find *decoherence free subspaces* [103] in which a qubit can be encoded; an approach that has been verified experimentally [104]. In the context of NMR, *dynamical decoupling* [105] is a popular method, where a series of short pulses are applied to the system. The implementation of fast gates is an obvious requirement in open quantum systems; by operating at the *quantum speed limit* [69, 106, 70, 71, 72], the effects of decoherence are minimized. This makes the use of numerical methods of optimal control especially important, as they are able to reach this limit [107, 73]. If the error due to decoherence in the gate implementation can be kept below a certain threshold of 10^{-3}–10^{-4}, depending on the system, quantum error correction can be applied [108, and references therein]. The basic idea of quantum error correction is to encode the information of one *logical* qubit in a system of multiple *physical* qubits.

In an open quantum system, the gate fidelity defined in Eq. (2.43) is no longer a well-defined quantity, since instead of a time evolution operator $\hat{\mathsf{U}}(T,0)$, the dynamics is now described by the dynamical map $\mathcal{E}(T,0)\hat{\rho}$. The *average fidelity* [109]

$$F_{\text{avg}} = \int \langle\Psi|\hat{\mathsf{O}}^{\dagger}\mathcal{E}(T,0)[|\Psi\rangle\langle\Psi|]\hat{\mathsf{O}}|\Psi\rangle \, \mathrm{d}\Psi \,, \tag{2.88}$$

is a measure for how much the dynamical map corresponds to a desired unitary gate $\hat{\mathsf{O}}$. It is defined via the Haar measure that averages over all possible states. In practice, the gate fidelity can be evaluated as [110]

$$\begin{split} F_{\text{avg}} = \frac{1}{N(N+1)} \sum_{i,j=1}^{N} \Bigg(&\langle\varphi_i|\hat{\mathsf{O}}^{\dagger}\mathcal{E}(T,0)\,[|\varphi_i\rangle\langle\varphi_j|]\,\hat{\mathsf{O}}|\varphi_j\rangle + \\ &+ \operatorname{tr}\left[\hat{\mathsf{O}}\,|\varphi_i\rangle\langle\varphi_i|\,\hat{\mathsf{O}}^{\dagger}\mathcal{E}(T,0)\,[|\varphi_j\rangle\langle\varphi_j|]\right] \Bigg)\,, \end{split} \tag{2.89}$$

where N is the dimension of the Hilbert space, and $\{|\varphi_i\rangle\}$ are the logical basis states. Since the average fidelity is more general than the gate fidelity in Eq. (2.43), it can also be used in the unitary case, where the dynamical map is given by

$$\mathcal{E}_U(T,0)\,[|\Psi\rangle\langle\Psi|] = \hat{\mathsf{U}}(T,0)\,|\Psi\rangle\langle\Psi|\,\hat{\mathsf{U}}^{\dagger}(T,0)\,. \tag{2.90}$$

Specifically for a two-qubit gate with unitary evolution, the average gate fidelity can also be written as

$$F_{\text{avg}} = \frac{1}{20}\left(\left| \operatorname{tr}\left[\hat{\mathsf{O}}^{\dagger}\hat{\mathsf{U}}\right] \right|^2 + \operatorname{tr}\left[\hat{\mathsf{O}}^{\dagger}\hat{\mathsf{U}}\hat{\mathsf{U}}^{\dagger}\hat{\mathsf{O}}\right] \right). \tag{2.91}$$

In the following chapters, we will consider several functionals that may be used to to steer the optimization of a quantum gate. While these functionals have advantages over Eq. (2.88) for the purpose of optimization, they are not easily comparable. The average fidelity therefore serves as a well-defined and experimentally accessible benchmark that can be used to characterize the result of an optimization, independent of the optimization functional.

3

Numerical Methods in Quantum Information Processing

The design of robust quantum gates under realistic conditions provides considerable numerical challenges. Even with a focus on two-qubit gates, it is generally not sufficient to model the system in the four-dimensional logical subspace spanned by the states $\{\vert 00\rangle, \vert 01\rangle, \vert 10\rangle, \vert 11\rangle\}$. For an accurate description of the underlying physical system, further degrees of freedom need to be taken into account. For example, for the case of trapped Rydberg atoms in chapter 4, the gate mechanism makes use of a Rydberg level $\vert r\rangle$ and an additional intermediate level $\vert i\rangle$, outside of the logical subspace, in order to generate the entanglement necessary for a two-qubit gate. Therefore, the minimum dimension of the total Hilbert space grows to 16. More generally, for trapped atoms, it may be necessary to also include their vibrational degree of freedom, which may further enlarge the Hilbert space by a factor of several hundred [73, 80]. For superconducting qubits discussed in chapters 5–7, the qubits are weakly anharmonic ladders, such that the population of higher levels cannot be neglected. Furthermore, the interaction between the qubit is via a transmission line resonator, which can receive significant excitation. With n_q levels per qubit and n_c levels for the resonator having to be taken into account, the total dimension of the Hilbert space $n_q^2 n_c$ is generally at least several hundred.

The problem is exacerbated when taking into account dissipation, such that the system has to be modeled as a density matrix in Liouville space that scales quadratically relative to the underlying Hilbert space. Hilbert space dimensions of several hundreds are numerically very manageable. The corresponding Liouville space dimension however, approaching 10^5, would push the boundaries of typically available computational resources. Therefore, it is vital to employ efficient data structures and algorithms that are problem-specific to the simulation of quantum dynamics and control, using highly optimized implementations. In the context of the work presented in this thesis, considerable effort has been spent to implement algorithms for the efficient representation, simulation, and optimization of both closed and open quantum systems, as part of the QDYN Fortran 90 library. This chapters reviews some of the key techniques that have been implemented and used in obtaining the results in chapters 4–7.

3.1 Numerical Representation of Quantum Systems

At the heart of any numerical treatment lies the ability to model the physical system in an optimal mathematical representation, and to translate this mathematical model into efficient data structures. It is important to realize how intimately efficient algorithms depend on the underlying data structures, which must therefore be designed with great care.

The first step in modeling a physical system numerically is to discretize it. In practice, this means choosing a suitable representation for the Hamiltonian (i.e., a set of basis functions). Naturally, the number of degrees of freedom should be as small as possible, while still providing an accurate description.

If only bound states of the Hamiltonian up to some maximum energy are relevant, an energy representation $\left\langle \phi_i \middle| \hat{\mathsf{H}} \middle| \phi_j \right\rangle$ is an obvious choice, where $\{|\phi_1\rangle, \dots, |\phi_N\rangle\}$ is the set of the first N eigenstates of the drift Hamiltonian. The resulting matrix representation is almost always very sparse. Consider a typical Hamiltonian of the form

$$\hat{\mathsf{H}}(t) = \hat{\mathsf{H}}_0 + \epsilon(t)\hat{\mathsf{H}}_1 \,. \tag{3.1}$$

The drift Hamiltonian $\hat{\mathsf{H}}_0$ is entirely diagonal, since $\hat{\mathsf{H}}(t)$ is represented in the eigenstates of $\hat{\mathsf{H}}_0$, but also the control Hamiltonian $\hat{\mathsf{H}}_1$ will generally have a sparse structure. For example if $\hat{\mathsf{H}}_1$ is the dipole operator that describes the interaction of an atom with an electromagnetic field, see appendix A, the non-zero elements in $\hat{\mathsf{H}}_1$ are determined by the selection rules. Sparsity is further increased in a composite system,

$$\hat{\mathsf{H}} = \hat{\mathsf{H}}_1 \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathsf{H}}_1 + \hat{\mathsf{H}}_I \,, \tag{3.2}$$

where $\hat{\mathsf{H}}_1$ is the Hamiltonian for each of the two subsystems, assuming that the interaction Hamiltonian $\hat{\mathsf{H}}_I$ is also sparse. Data structures for storing sparse matrices are well-established [111, 112], and are at least partly available in standard numerical libraries such as LAPACK [113]. Where these existing data structures are too general, it is advisable to implement data structures that reflect the sparsity of the representation to the greatest extent possible. The possibility to efficiently apply a Hamiltonian to a wave function is key to any further numerical methods of time propagation or optimal control.

The same sparsity argument applies equally in Liouville space. Given a master equation (2.71) in Lindblad form, the commutator $[\hat{\mathsf{H}}, \hat{\rho}]$ can be efficiently calculated if $\hat{\mathsf{H}}$ is stored in a sparse matrix format. Likewise, the Lindblad operators $\{\hat{\mathsf{A}}_i\}$ or alternatively the entire dissipator $\mathcal{L}_D$ can be stored as a sparse matrix.

While the energy representation is very straightforward, it is neither the only nor generally the most efficient choice. In the standard example of the harmonic oscillator [63],

$$\hat{\mathsf{H}}_{ho} = \hbar\omega\left(\hat{\mathsf{a}}^\dagger\hat{\mathsf{a}} + \frac{1}{2}\right)\,, \tag{3.3}$$

$\hat{\mathsf{a}}$ can be written in either energy representation (truncated to N eigenstates) or coordinate representation,

$$\hat{\mathsf{a}} = \sum_{n=0}^{N} \sqrt{n}\,|n-1\rangle\langle n| = \sqrt{\frac{m\,\hbar\omega}{2}}\left(\hat{\mathsf{x}} + \frac{\mathrm{i}\hat{\mathsf{p}}}{m\,\hbar\omega}\right)\,, \tag{3.4}$$

leading to

$$\hat{\mathsf{H}}_{ho} = \sum_{n=0}^{N} \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle\langle n| = \frac{\hat{\mathsf{p}}^2}{2m} + \frac{m\omega^2}{2}\hat{\mathsf{x}}^2\,. \tag{3.5}$$

While the energy representation is already discrete and finite, by choice of the truncated set of energy basis functions $\{|0\rangle, \dots, |N\rangle\}$, the coordinate representation must still be discretized for numerical use. That is, we must also select a finite set of basis functions in which to represent the Hamiltonian. The Fourier method [114, 9] defines a *pair* of such sets. One set consists of the "spectral" basis of plane waves, $|n\rangle = |\mathrm{e}^{\mathrm{i}k_n x}\rangle$, the other of a basis of associated "pseudo-spectral" functions [9, 80]. In practical terms, the representation of a state in the pseudo-spectral basis is simply the vector $(\Psi(x_i), \dots, \Psi(x_N))$ for x_i on an equidistant grid. The operator $\hat{\mathsf{x}}^2$ in this basis is the diagonal matrix $\hat{\mathsf{x}}^2 = \sum_n x_n^2\,|n\rangle\langle n|$, whereas $\hat{\mathsf{p}}^2$ is a dense matrix. The representation in the plane wave basis is the Fourier transform of $\Psi(x_i)$, and consequently $\hat{\mathsf{p}}^2$ is diagonal, whereas $\hat{\mathsf{x}}^2$ is dense. This suggests to split the Hamiltonians into two terms, the kinetic operator $\hat{\mathsf{T}} = \hat{\mathsf{p}}^2/(2m)$ and the harmonic potential $\hat{\mathsf{V}} = \frac{1}{2}m\omega^2\hat{\mathsf{x}}^2$ and to store each in the representation in which it is sparse (i.e. diagonal). In the application of the total Hamiltonian, the state vector or density matrix must then be converted from one representation to the other, which can be done efficiently using the fast-Fourier-transform (FFT), cf. appendix E. The necessary number of grid points in x and p is completely determined by the phase space volume $[x_{\text{min}}, x_{\text{max}}] \times [p_{\text{min}}, p_{\text{max}}]$ in which the dynamics take place, see appendix E. The Fourier representation fills this volume with the minimal number of equally distributed points.

In the example of the harmonic oscillator, the Fourier representation only becomes useful if the system is driven with a term $\epsilon(t)(\hat{\mathsf{a}} + \hat{\mathsf{a}}^\dagger)$. In this case, the energy representation becomes banded instead of diagonal, whereas the Fourier representation remains diagonal. Which representation is more efficient then depends on the maximum excitation; for $n > 32$, the Fourier grid is generally more efficient. However, in the description of molecular dynamics, where the method was developed [114, 115], the Fourier representation is the default choice. There, the Born-Oppenheimer approximation yields energy surfaces $\hat{\mathsf{V}}_{BO}(R)$ for each of the electronic degrees of freedom of a diatomic molecule. For two electronic surfaces, the Hamiltonian might be written as e.g.

$$\hat{\mathsf{H}} = \begin{pmatrix} \hat{\mathsf{T}}(k) + \hat{\mathsf{V}}_1(R) & \mu(R)\epsilon(t) \\ \hat{\mu}(R)\epsilon(t) & \hat{\mathsf{T}}(k) + \hat{\mathsf{V}}_2(R) \end{pmatrix} . \tag{3.6}$$

With the spatial degree of freedom represented on a Fourier grid with N_R points, $\hat{\mathsf{H}}$ is a sparse matrix of dimension $2N_R$, and can be applied efficiently by storing only the vectors $V_1(R)$, $V_2(R)$, and $\mu(R)$ in the pseudo-spectral representation and the vector $T(k) \propto k^2$ for the kinetic operator in the plane wave representation. For a more realistic example, see Ref. [6].

A common situation for such molecular systems is that the phase space is used very unevenly, requiring an excessive number of points to sample it evenly in the Fourier representation. This issue can be addressed by using a mapped grid [116, 117, 118, 119, 120], where a coordinate transformation is used to mold the used phase space volume into a square.

In some dynamical problems such as the transport of trapped ions, the wave packet is within in a small region of phase space at every instant, but moves over large distances (and values of velocity) over time. This situation is efficiently modeled with a moving grid [121, 7], where the phase space volume $[x_{\text{min}}(t), x_{\text{min}}(t) + \Delta x] \times [p_{\text{min}}(t), p_{\text{min}}(t) + \Delta p]$ is moved dynamically along with the wave packet.

The pseudo-spectral basis functions associated with the Fourier grid are not the only example of a *collocation* method, where the expansion coefficients of a state are the values of the wave functions $\Psi(x_j)$ for points x_j. However, the Fourier grid is the only choice for which the x_j are equidistant. Generally, the spectral basis functions should be chosen such that they diagonalize the kinetic operator; e.g. for spherical symmetry with the magnetic quantum number $m = 0$, these would be Legendre polynomials. The general use of spectral methods has an extensive body of work behind it [122] and is central to the efficient numerical treatment of molecular dynamics.

3.2 Simulation of Quantum Dynamics

Once the application of the Hamiltonian or Liouvillian to a state has been realized, the next step is to simulate the dynamics of the quantum system. This usually means solving the time-dependent Schrödinger equation (2.16) or the Liouville-von Neumann equation (2.72).

There are two possible approaches to obtaining a solution. The first is to simply take the equation of motion and apply one of the generic numerical methods for solving ordinary differential equations (ODEs), like one of the Runge-Kutta (RK) methods [123, 124]. This approach has the benefit that ODE solvers such as RK45 are readily available in numerical libraries. They are also very flexible with respect to the equation of motion, as long is it is reasonably well-behaved [123]. For example, instead of the Schrödinger equation, the non-linear Gross-Pitaevskii-equation describing a Bose-Einstein condensate [125], can easily be solved. However, general ODE solvers will lack in numerical efficiency and, most importantly, accuracy.

The alternative approach is to solve the equation of motion analytically, and then to evaluate that solution numerically. In this way, results of arbitrary precision can be obtained, with the obvious caveat that the propagation scheme will be specific to a particular equation of motion and its solution.

The Schrödinger equation for a time-independent Hamiltonian has the solution

$$|\Psi(T)\rangle = \hat{\mathsf{U}}(T,0)\,|\Psi(0)\rangle = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}\,T}\,|\Psi(0)\rangle \;, \tag{3.7}$$

For time-dependent Hamiltonians, we approximate $\hat{\mathsf{H}}(t)$ as piecewise constant on a time grid with time step dt. Then, the total time evolution operator is the product of the time evolution operators at each time step,

$$\hat{\mathsf{U}}[T,0] = \prod_{i=1}^{nt-1} \underbrace{\hat{\mathsf{U}}(t_i + dt, t_i)}_{\equiv \hat{\mathsf{U}}_i} = \prod_{i=1}^{nt-1} \exp\left[-\frac{\mathrm{i}}{\hbar}\underbrace{\hat{\mathsf{H}}\left(t_i + \frac{dt}{2}\right)}_{\equiv \hat{\mathsf{H}}_i} dt\right]. \tag{3.8}$$

The time step dt must be chosen sufficiently small that this is a good approximation; in practice, convergence is checked by verifying that the numerical results remain stable within a desired precision when dt is decreased. The number of necessary time steps can be reduced significantly if a rotating wave approximation (RWA) is justified. As shown in appendix B, the RWA allows to eliminate fast oscillations in the pulse, leaving only a slowly varying shape.

A naive way of evaluating $\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}_i\,dt}$ is to diagonalize the Hamiltonian and use the eigendecomposition

$$\hat{\mathsf{H}}_i = \hat{\mathsf{W}}_i \begin{pmatrix} \lambda_1^{(i)} & & \\ & \ddots & \\ & & \lambda_N^{(i)} \end{pmatrix} \hat{\mathsf{W}}_i^\dagger , \tag{3.9}$$

where $\hat{\mathsf{W}}_i$ contains the eigenvectors of $\hat{\mathsf{H}}_i$ as columns, to write the propagation step as

$$\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}_i dt} \left|\Psi\right\rangle = \hat{\mathsf{W}}_i \begin{pmatrix} \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\lambda_1^{(i)}dt} & & \\ & \ddots & \\ & & \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\lambda_N^{(i)}dt} \end{pmatrix} \hat{\mathsf{W}}_i^\dagger \left|\Psi\right\rangle . \tag{3.10}$$

This “exact” exponentiation is suitable for Hilbert spaces of trivially small dimension < 10. Since diagonalization scales as N^3 with the dimension N of the matrix [126], full diagonalization at every time step quickly becomes numerically infeasible. Moreover, $\hat{\mathsf{H}}$ and $\hat{\mathsf{W}}$ must be constructed as dense matrices.

For a system of non-trivial size, $\exp[-\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}dt]$ is evaluated by expanding the exponential in a polynomial series,

$$\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}\,dt} \left|\Psi\right\rangle \approx \sum_{n=0}^{N-1} a_n P_n(\hat{\mathsf{H}}) \left|\Psi\right\rangle . \tag{3.11}$$

where $P_n(\hat{\mathsf{H}})$ is a polynomial of degree n and $\{a_n\}$ are the expansion coefficients. Applying $P_n(\hat{\mathsf{H}})$ to $\left|\Psi\right\rangle$ then simply means repeated applications of $\hat{\mathsf{H}}$. For this reason, an efficient propagation relies on the proper use of sparsity in storing the Hamiltonian and spectral methods such as the Fourier grid.

The idea of evaluating the exponential as a polynomial series is already presupposed by the very definition of the exponential of an operator,

$$\exp[\hat{\mathsf{A}}] \equiv \sum_{n=0}^{\infty} \frac{1}{n!} \hat{\mathsf{A}}^n . \tag{3.12}$$

However, this expansion converges particularly slowly and is numerically unstable [127]. Thus, it is not suitable for time propagation. Instead, a polynomial basis must be chosen such that Eq. (3.11) converges quickly and can be truncated as early as possible.

3.2.1 Chebychev Propagation

For a function $f(x)$ with $x \in [-1, 1]$, it can be shown [128] that the fastest converging polynomial series is the expansion in Chebychev polynomials

$$P_n(x) = \cos(n\theta); \qquad \theta = \arccos(x)\,. \tag{3.13}$$

The function $f(x)$ must be sampled at N discrete points $\{x_k\}$ that are either the roots or the extrema of the N'th Chebychev polynomial; the expansion coefficients are

$$a_n = \frac{2 - \delta_{n0}}{N} \sum_{k=0}^{N-1} f(x_k) P_n(x_k)\,. \tag{3.14}$$

When using the Chebychev expansion for propagation, the requirement that the argument of $f(x)$ must be real translates into $\hat{\mathsf{H}}$ being Hermitian. Secondly, to account for the requirement that $x \in [-1, 1]$, the Hamiltonian must be normalized as [114, 9, 129]

$$\hat{\mathsf{H}}_{\text{norm}} = 2\frac{\hat{\mathsf{H}} - E_{\text{min}}\,\mathbb{1}}{\Delta} - \mathbb{1}\,, \tag{3.15}$$

where $\Delta = E_{\text{max}} - E_{\text{min}}$ is the spectral radius and E_{max} and E_{min} are the smallest and largest eigenvalue.

For $f(\hat{\mathsf{H}}) = \exp\left[-\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}\,dt\right]$, the series converges for N being a small multiple of $\lfloor\alpha\rfloor$ with $\alpha = \frac{\Delta}{2}\,dt$. In this case, the expansion coefficients can be calculated analytically as [127]

$$a_n = (2 - \delta_{n0})\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\left(\frac{\Delta}{2} + E_{\text{min}}\right)dt} J_k(\alpha)\,, \tag{3.16}$$

where $J_k(\alpha)$ is the Bessel function of first kind. In order to calculate the propagated state,

$$|\Psi\rangle = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}\,dt}\,|\Psi_0\rangle = \sum_n a_n \underbrace{P_n(-\mathrm{i}\hat{\mathsf{H}}_{\text{norm}})\,|\Psi_0\rangle}_{\equiv|\Phi_n\rangle}\,, \tag{3.17}$$

the series is truncated as soon as $|a_k|$ is below machine precision. Since both $\hat{\mathsf{H}}_{\text{norm}}$ and $|\Psi_0\rangle$ are normalized, this guarantees that the entire term in the series is below machine precision as well. Eq. (3.17) is evaluated using the recursive relationship of the Chebychev polynomials [114, 9, 129],

$$|\Phi_0\rangle = |\Psi_0\rangle\;, \tag{3.18}$$

$$|\Phi_1\rangle = -\mathrm{i}\hat{\mathsf{H}}_{\text{norm}}\,|\Phi_0\rangle\;, \tag{3.19}$$

$$|\Phi_n\rangle = -2\mathrm{i}\hat{\mathsf{H}}_{\text{norm}}\,|\Phi_{n-1}\rangle + |\Phi_{n-2}\rangle\;. \tag{3.20}$$

The full algorithm is summarized in appendix F as Algorithm 2.

The propagator is stable as long as the spectrum of the Hamiltonian is in the range $[E_{\text{min}}, E_{\text{min}} + \Delta]$ that is used to obtain the Chebychev coefficients. Therefore, for propagation on a time grid with a Hamiltonian of the form $\hat{\mathsf{H}} = \hat{\mathsf{H}}_0 + \epsilon(t)\hat{\mathsf{H}}_1$, the Chebychev coefficients can be calculated once and then re-used for every propagation step. A good heuristic in this case is to choose a spectral radius that includes the spectrum of both $\hat{\mathsf{H}}_{\text{max}} = \hat{\mathsf{H}}_0 + \max_t \epsilon(t)\hat{\mathsf{H}}_1$, and $\hat{\mathsf{H}}_{\text{min}} = \hat{\mathsf{H}}_0 + \min_t \epsilon(t)\hat{\mathsf{H}}_1$.

Still, it is essential to have a good approximation for the spectral range of a given Hamiltonian. Specifically for molecular dynamics, estimating the spectral range from the extrema of the potentials and the kinetic energy gives sufficiently accurate results [127]. More generally, a very good approximation of the minimum and maximum eigenvalue of an operator can be obtained using the *Arnoldi* method. Starting from a random state $|\Psi_0\rangle$, the *Krylov space* is built by repeatedly applying $\hat{\mathsf{H}}$ and orthonormalizing the obtained states [130, 115]. The projection of the Hamiltonian into this Krylov space yields a Hessenberg matrix, whose maximum and minimum eigenvalues converge towards the maximum and minimum eigenvalues of the Hamiltonian. The algorithm is detailed in appendix F. For each iteration j, the spectral range is calculated from the eigenvalues obtained in line 10 of Algorithm 5. The iteration continues until the result is converged to some predefined precision.

The Chebychev method can also be applied to other equations of motion, as long as an analytical solution can be derived and expanded in Chebychev polynomials. The expansion coefficients in this case cannot be expressed in Bessel functions, but must be derived using a cosine transform, as outlined in appendix E. In this way, a Chebychev propagator for the inhomogeneous Schrödinger equation has been derived [129].

For a truly time-dependent Hamiltonian that is not well-approximated as piecewise constant the formal solution of the Schrödinger equation (2.16) is $|\Psi(t)\rangle = \hat{\mathsf{U}}(t,0)|\Psi(0)\rangle$, with

$$\hat{\mathsf{U}}(t,0) = \mathcal{T} \exp\left[-\frac{\mathrm{i}}{\hbar}\int_0^t \hat{\mathsf{H}}(t')\,\mathrm{d}t'\right], \tag{3.21}$$

where $\mathcal{T}$ represents the time ordering operator. This can be rewritten as the solution to an inhomogeneous Schrödinger equation [131], and thus also be evaluated using the inhomogeneous Chebychev propagator.

In principle, the solution to the Liouville-von Neumann equation (2.72) takes the same form as the time evolution operator in Hilbert space, i.e.,

$$\hat{\rho}(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\mathcal{L}t}\hat{\rho}(0)\,. \tag{3.22}$$

However, unlike the Hamiltonian, the Liouvillian is generally not a Hermitian operator, except when there is no dissipator, i.e. $\mathcal{L}[\hat{\rho}] = [\hat{\mathsf{H}}, \hat{\rho}]$. In this case, the spectrum of $\mathcal{L}$ is in the interval $[-\Delta, +\Delta]$, where Δ is the spectral radius of $\hat{\mathsf{H}}$.

3.2.2 Newton Propagation

In the general case of a dissipative Liouvillian, instead of a Chebychev expansion, an expansion into Newton polynomials can be used. For a general function $f(z)$ with $z \in \mathbb{C}$, the expansion in Newton polynomials $R_n(z)$ reads

$$f(z) \approx \sum_{n=0}^{N-1} a_n R_n(z)\,, \quad R_n(z) = \prod_{j=0}^{n-1} (z - z_j)\ , \tag{3.23}$$

for a set of sampling points $\{z_j\}$ at which the interpolation is exact. The coefficients are defined as the *divided differences* [132],

$$a_0 = f(z_0)\,, \tag{3.24a}$$

$$a_1 = f(z_1) - f(z_0)\,, \tag{3.24b}$$

$$a_n = \frac{f(z_n) - \sum_{j=0}^{n-1} a_j \prod_{k=0}^{j-1} (z_n - z_k)}{\prod_{j=0}^{n-1} (z_n - z_j)}\,. \tag{3.24c}$$

For solving the Liouville-von Neumann equation, $f(z) = \mathrm{e}^{-\mathrm{i} z\, dt}$, where the argument z is $\mathcal{L}/\hbar$. Thus, the propagation is written as

$$\hat{\rho} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\mathcal{L}\, dt}\, \hat{\rho}_0 \approx \underbrace{\sum_{n=0}^{N-1} a_n \left(\mathcal{L} - z_n \mathbb{1}\right)}_{\equiv p_{N-1}(\mathcal{L})} \hat{\rho}_0\,, \tag{3.25}$$

where the polynomial is evaluated through repeated application of Eq. (2.72).

The central issue for obtaining a fast-converging series is a proper choice of the sampling points $\{z_j\}$. The fastest convergence results from using the complex eigenvalues of $\mathcal{L}$ [133]. However, the exact eigenvalues of the Liouvillian are not readily available. More generally, arbitrary points from the spectral domain of $\mathcal{L}$ can be used as sampling points.

A widely used method is to estimate the spectral domain and to encircle it with a rectangle or ellipse [134, 132, 135]. Then, a large number of expansion coefficients are calculated from sampling points on that boundary. The same coefficients are used for the propagation of all Liouvillians on the time grid under the assumption that they all fit into the same encirclement. The series is truncated as soon as convergence is reached. This is similar to the method employed for the Chebychev propagator, where a set of coefficients is calculated once and then used for the propagation of any

Hamiltonian that is within the same spectral range.

A middle path between the exact eigenvalues of $\mathcal{L}$ and the crude encirclement of the spectral domain is the use of the Krylov method to obtain approximate eigenvalues. The same method was already employed to estimate the spectral radius for the Chebychev propagator. The Arnoldi algorithm 5 in appendix F for $\hat{A} = \mathcal{L}$ and using $\vec{v} = \hat{\rho}$ as a starting vector yields a set of approximate eigenvalues of $\mathcal{L}$, as well as a Hessenberg matrix $\hat{H}$ that is the projection of $\mathcal{L}$ into the Krylov subspace, and the set of Arnoldi vectors that span that subspace. Instead of using $\mathcal{L}$ as the argument of the polynomial p_{N-1}, the Hessenberg matrix may be used. If $\hat{V}_N$ is the transformation matrix between the full Liouville space and the reduced Krylov space, consisting of the Arnoldi vectors as columns, the propagation is evaluated using

$$\mathcal{L}\hat{\rho} \approx \hat{V}_N \, p_{m-1}(\hat{H}) \, \hat{V}_N^{\dagger} \hat{\rho}_0 \,. \tag{3.26}$$

Assuming N is much smaller than the full dimension of the Liouville space, most of the numerical effort is in the Arnoldi algorithm, in constructing the Krylov space.

However, even for moderate values of N (typically on the order of 100), the Arnoldi algorithm can require prohibitive amounts of memory. This is because a full set of N Arnoldi vectors, each of the dimension of the Liouville space, need to be stored. To counter this problem, an iterative scheme has been developed [136]. Instead of performing the Arnoldi algorithm to a high order N, until convergence is reached in the propagation, we stop at some small order $m < 10$. This gives a first approximation to the propagated density matrix,

$$\hat{\rho}^{(1)} = p^{(0)}_{m-1}(\mathcal{L})\hat{\rho}_0 = \sum_{n=0}^{m-1} a_n R_n(\mathcal{L})\hat{\rho}_0 \,. \tag{3.27}$$

The idea is now to iteratively add remaining terms to the Newton series in chunks of size m, retaining all coefficients and sampling points, but restarting the Arnoldi procedure in every iteration.

Adding the next m terms to Eq. (3.27) yields

$$\begin{aligned} \hat{\rho}^{(2)} &= \hat{\rho}^{(1)} + \sum_{n=m}^{2m-1} a_n R_n(\mathcal{L})\hat{\rho}_0 \\ &= \hat{\rho}^{(1)} + \underbrace{\left(\sum_{n=0}^{m-1} a_{m+n} R^{(1)}_n(\mathcal{L})\right)}_{\equiv p^{(1)}_{m-1}} \underbrace{\left(R^{(0)}_m(\mathcal{L})\hat{\rho}_0\right)}_{\equiv \hat{\sigma}^{(1)}} , \end{aligned} \tag{3.28}$$

with

$$R_n^{(0)}(\mathcal{L}) = \prod_{j=0}^{n-1}(\mathcal{L} - z_j\mathbb{1}), \qquad R_n^{(1)}(\mathcal{L}) = \prod_{j=0}^{n-1}(\mathcal{L} - z_{n+j}\mathbb{1})\,. \tag{3.29}$$

That is, the terms in $R_n(\mathcal{L})$ already known from the calculation of $\hat{\rho}^{(1)}$ have been pulled out, and yield a new "starting vector" $\hat{\sigma}^{(1)}$, which is the argument to a Newton series of only m new terms. The new sampling points on which the $R_n^{(1)}$ are evaluated are obtained by applying the Arnoldi procedure to $\hat{\sigma}^{(1)}$. The Newton coefficients continue recursively from the previous restart. The third iteration yields

$$\hat{\rho}^{(3)} = \hat{\rho}^{(2)} + \underbrace{\left(\sum_{n=0}^{m-1} a_{2m+n} R_n^{(2)}(\mathcal{L})\right)}_{\equiv p_{m-1}^{(2)}} \underbrace{\left(R_m^{(1)}(\mathcal{L})\hat{\sigma}_1\right)}_{\equiv \hat{\sigma}^{(2)}}\,. \tag{3.30}$$

The Newton propagator continues, adding the m terms evaluating

$$p_{m-1}^{(s)}(\mathcal{L})\hat{\sigma}^{(s)} = \sum_{n=0}^{m-1} a_{sm+n} \prod_{k=0}^{n-1} \left(\mathcal{L} - z_{sm+k}\mathbb{1}\right)\hat{\sigma}^{(s)} \tag{3.31}$$

with

$$\hat{\sigma}^{(s)} = p_{m-1}^{(s-1)}\hat{\sigma}^{(s-1)} \tag{3.32}$$

at every restart iteration. The complete algorithm is listed in appendix F.

In the implementation of the algorithm, there are two details that need to be taken into account for numerical stability. First, the denominator of the divided differences in Eq. (3.24) may become extremely small if consecutive sampling points are close to each other. This can be addressed by reordering the points such that the denominator in the divided differences is maximized. This process is called *Leja ordering* [137]. The reverse problem that the sampling points are too far apart, causing an underflow in the calculation of coefficients can be avoided by normalizing the Liouvillian as

$$\tilde{\mathcal{L}} = \frac{1}{\rho}\left(\mathcal{L} - c\right)\,, \tag{3.33}$$

where c is an estimate for the center of the spectrum of $\mathcal{L}$, and the eigenvalues are roughly contained in a radius ρ around c. These values can be estimated from the sampling points obtained in the first iteration of the Newton propagator. The normalization of the Liouvillian is in some sense similar to the normalization of the Hamiltonian in the Chebychev propagator, but it is crucial there since the Chebychev polynomials are only defined in the domain $[-1, 1]$. For the Newton propagator, the normalization is only for numerical stability.

The primary use for the Newton propagator is to solve the Liouville-von Neumann equation with a dissipative term. However, it can also be used to solve the Schrödinger equation with a non-Hermitian Hamiltonian. This is sometimes used as an ad-hoc model for spontaneous decay, avoiding the significant overhead incurred by modeling the system properly in Liouville space. As seen in section 2.6.3 of chapter 2, spontaneous decay from a level $|n\rangle$ results in the population of that level decreasing proportionally to $\mathrm{e}^{-\gamma_1 t}$. To obtain the same decay behavior in Hilbert space, the complex amplitude a_n of the level $|n\rangle$ must decay at a rate of $\frac{\gamma_1}{2}$, since the population of $|n\rangle$ is given by $|a_n|^2$. This is achieved by adding a non-Hermitian term to the Hamiltonian, resulting in

$$\hat{\mathsf{H}}_\gamma = \hat{\mathsf{H}} - \mathrm{i}\frac{\gamma_1}{2} |n\rangle \langle n| \,. \tag{3.34}$$

Propagation with such a Hamiltonian does not conserve the norm of $|\Psi\rangle$; the population that has decayed simply vanishes. Specifically, it is not added to the state to which $|n\rangle$ decays, since the incoherent superposition of states cannot be modeled in Hilbert space. Nonetheless, a non-Hermitian Hamiltonian can be a useful way to determine e.g. how much the fidelity of a quantum operation is affected if spontaneous decay is taken into account.

3.3 Optimization Methods

3.3.1 Optimization of Two-Qubit Quantum Gates

Building upon the ability to simulate the dynamics of a quantum system, the final step is to apply the methods of numerical optimal control (OCT), as introduced in chapter 1. The focus of this thesis is the realization of two-qubit quantum gates $\hat{\mathsf{O}}$.

For a quantum gate to be successfully implemented at time T, all basis states of the two-qubit logical subspace must evolve according to

$$\hat{\mathsf{U}}(0,T) |k\rangle \overset{!}{=} \mathrm{e}^{\mathrm{i}\phi} \hat{\mathsf{O}} |k\rangle \,, \qquad |k\rangle \in \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\} \,, \tag{3.35}$$

with an arbitrary global phase ϕ. Likewise, in Liouville space,

$$\mathcal{E}(T,0)\hat{\rho}_{ij} \overset{!}{=} \hat{\mathsf{O}} \hat{\rho}_{ij} \hat{\mathsf{O}}^\dagger \,, \qquad \hat{\rho}_{ij} = |i\rangle \langle j| \,, \tag{3.36}$$

where $\mathcal{E}(T,0)\hat{\rho}$ denotes the dynamical map.

The most straightforward way to express this in a functional is in terms of the complex overlaps,

$$\tau_k = \left\langle k \,\middle|\, \hat{\mathsf{O}}^\dagger \hat{\mathsf{U}}(T,0) \,\middle|\, k \right\rangle \tag{3.37}$$

in Hilbert space, or

$$\tau_k = \operatorname{tr}\left[\left(\hat{\mathsf{O}}^\dagger \hat{\rho}_k^\dagger \hat{\mathsf{O}}\right) \mathcal{E}(T,0)\hat{\rho}_k\right] \tag{3.38}$$

in Liouville space. Two commonly used possibilities for obtaining a real-valued functional are [138]

$$J_{T,\text{sm}} = 1 - \frac{1}{N^2}\left|\sum_{k=1}^{N}\tau_k\right|^2 = 1 - \frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\tau_l^*\tau_k\,, \tag{3.39}$$

$$J_{T,\text{re}} = 1 - \frac{1}{N}\mathfrak{Re}\left[\sum_{k=1}^{N}\tau_k\right] = 1 - \frac{1}{N}\sum_{k=1}^{N}\frac{1}{2}\left(\tau_k + \tau_k^*\right)\,, \tag{3.40}$$

where $N = 4$ in Hilbert space and $N = 16$ in Liouville space. $J_{T,\text{sm}}$ takes its minimum value of 0 if Eq. (3.35) is fulfilled with an arbitrary phase ϕ, whereas $J_{T,\text{re}}$ is 0 only for a global phase of $\phi = 0$. Note that in Hilbert space,

$$\sum_{k=1}^{N}\tau_k = \operatorname{tr}[\hat{\mathsf{O}}^\dagger\hat{\mathsf{U}}] \tag{3.41}$$

is the Hilbert-Schmidt overlap of the target gate with the implemented gate. In Liouville space, $J_{T,\text{sm}}$ and $J_{T,\text{re}}$ are equivalent, as global phases cannot be expressed; $J_{T,\text{re}}$ is preferred due to its simpler form.

It is important to note that $J_{T,\text{re}}$ or $J_{T,\text{sm}}$ are used as optimization functionals only. While in Hilbert space, $J_{T,\text{sm}}$ still has a well-defined meaning (the probability that a state $\hat{\mathsf{U}}\vert\Psi\rangle$ is obtained in a measurement of $\hat{\mathsf{O}}\vert\Psi\rangle$), this is no longer the case in Liouville space. In general, a physically meaningful figure of merit of the success of implementing a quantum gate is through the average gate fidelity F_{avg} defined in Eq. (2.88). For optimization purposes, we have the freedom to choose an alternative figure of merit J_T with a simpler structure, as long as both J_T and the "physical" gate error $1 - F_{\text{avg}}$ reach their minimum value for an optimal pulse, and only return to F_{avg} afterwards to compare results.

The Hamiltonian that induces $\hat{\mathsf{U}}(T,0)$ or $\mathcal{E}(T,0)$ takes the form $\hat{\mathsf{H}} = \hat{\mathsf{H}}_d + \hat{\mathsf{H}}_c[\epsilon(t)]$, with the drift Hamiltonian $\hat{\mathsf{H}}_d$ and a control Hamiltonian $\hat{\mathsf{H}}_c[\epsilon(t)]$ that depends on a driving field $\epsilon(t)$, parametrized through an arbitrary number of independent control parameters ϵ_j. In most cases considered here, the pulse is approximated as piecewise constant on a time grid, and the control parameters are the values of $\epsilon_j \equiv \epsilon(t_j + \frac{dt}{2})$. A more constrained parametrization with fewer control parameters would be provided through predefined analytical shapes, e.g. for a Gauss-shaped pulse, the ϵ_j would be the amplitude, duration, and central frequency; for a spectral decomposition, the ϵ_j would be the amplitudes for the frequencies ω_j.

The task is to minimize a functional such as $J_{T,\text{sm}}$ or $J_{T,\text{re}}$, in an iterative procedure: in each OCT iteration i, we find an update to the pulse $\epsilon^{(i)}(t)$ such that the updated pulse $\epsilon^{(i+1)}(t)$ yields an improved value for the functional. The optimization loop continues until J reaches a value that is smaller than some predefined limit, or until the value of J shows no significant improvement.

There are two basic categories of algorithms for finding updates for the control parameters that improve the value of the functional. The first are "gradient-free", employing only evaluations of the functional. Most prominently, this includes the downhill-simplex algorithm discussed in section 3.3.2 below. Gradient-free optimization algorithms are extremely versatile and easy to apply, but they also tend to converge very slowly, particularly for a large number of control parameters. Moreover, they are prone to running into local minima, although more advanced *global* methods such as genetic algorithms [139] and swarm search optimization [140] also exist. Gradient-free optimization methods can easily be incorporated into experimental setups as a *closed-loop* control, where a measurement determines the figure of merit and drives a variation of the control parameters for the next iteration.

The second category are gradient-based methods. Including information about how the optimization functional varies with changes in the controls greatly speeds up convergence. However, it requires to derive analytical expressions for the gradient, and then additional numerical resources for evaluating that gradient. A *concurrent* scheme like the gradient ascent discussed in section 3.3.3 varies each control parameter individually according to the derivative of the optimization functional with respect to that parameter. In contrast, Krotov's method presented in section 3.3.4 takes the time-continuous $\epsilon(t)$ as a whole. Subsequent discretization to a time grid yields a *sequential* scheme in which the update for ϵ_j takes into account the updates at earlier times. While not applicable to arbitrary pulse parametrizations, Krotov's method guarantees monotonic convergence. Gradient-based methods are typically used as an *open-loop* control, where the entire optimization is performed based on a numerical simulation, before taking the final optimized set of controls to an experimental implementation. Open- and closed-loop control may also be combined in a hybrid scheme [141] that brings together the advantages of both methods for experimental applications.

For simplicity, we have only considered a single control Hamiltonian and driving field here; in general, there may be multiple control Hamiltonians and driving fields. In this case, each driving field is parametrized independently.

3.3.2 Downhill Simplex Optimization

The downhill simplex, or Nelder-Mead algorithm [142] is a particularly simple, gradient-free optimization method that is very effective if there is only a handful of N optimization parameters. The idea is to construct a simplex polytope consisting of $N+1$ points in the N-dimensional parameter space. The point with the largest value of the functional is then replaced by reflection on the remainder of the polytope, followed by some contraction and expansion steps, yielding a new point with an improved value of the functional. Intuitively, the simplex "rolls" down the optimization landscape.

The fact that the algorithm only relies on the *evaluation* of the optimization functional makes it extremely versatile, giving a black box method for the optimization of arbitrary figures of merit. It is well-suited as a pre-optimization for finding pulses of simple analytical forms that may then be optimized further on using a gradient-based method on a time grid. In particular, the pulse duration T can easily be included as a control parameter for a simplex method. In contrast, T must be fixed when optimizing on a time grid.

Non-gradient methods may be used for controls that are inherently coarse-grained, due to the limitations of experimental setups, e.g. the limited number of pixels in early femtosecond pulse shapers [143]. Some control problems, specifically in quantum many-body systems [144] have been demonstrated to have solutions that contain only a small number of frequency components. The CRAB algorithm [145] (Chopped RAndom Basis) has been developed for this class of problems; it uses a pre-specified number of N frequency components, but chooses those frequencies at random. That is, the control pulse is parametrized as

$$\epsilon_{\text{CRAB}}(t) = S(t) \sum_{n=1}^{N} \left(a_n \sin(\omega_n t) + b_n \cos(\omega_n t)\right) \,, \tag{3.42}$$

with a pulse shape $S(t)$. The frequencies are chosen as

$$\omega_n = \frac{2\pi n}{T}(1 + r_n)\,; \qquad r_n \in [-0.5, 0.5]\,, \tag{3.43}$$

where T is the pulse duration and the r_n are random numbers. The coefficients a_n and b_n are optimized using the downhill-simplex algorithm, with many optimizations running in parallel, using different randomized frequencies. The CRAB method can be useful since it limits the effective number of control parameters without assuming a simple analytical pulse shape, but it fails for control problems whose solution require a large number of spectral components.

3.3.3 Gradient Ascent

The GRAPE algorithm [22] (Gradient Ascent Pulse Engineering) considers the gradient $\frac{\partial J}{\partial \epsilon_j}$ with respect to any control parameter and then updates that control parameter according to

$$\epsilon_j^{(i+1)} = \epsilon_j^{(i)} - \alpha \frac{\partial J}{\partial \epsilon_j}\,, \tag{3.44}$$

using a suitable step width α.

For functionals depending on the overlap between propagated and expected state, and assuming a time-grid parametrization of the pulse, the gradient of Eq. (3.37) becomes

$$\begin{aligned} \frac{\partial \tau_k}{\partial \epsilon_j} &= \frac{\partial}{\partial \epsilon_j} \left\langle k \middle| \hat{\mathsf{O}}^\dagger \hat{\mathsf{U}}_{nt-1} \dots \hat{\mathsf{U}}_j \dots \hat{\mathsf{U}}_1 \middle| k \right\rangle \\ &= \left\langle k \middle| \hat{\mathsf{O}}^\dagger\, \hat{\mathsf{U}}_{nt-1} \dots \hat{\mathsf{U}}_{j+1} \frac{\partial \hat{\mathsf{U}}_j}{\partial \epsilon_j} \hat{\mathsf{U}}_{j-1} \dots \hat{\mathsf{U}}_1 \middle| k \right\rangle \\ &= \left\langle \chi_k(t_{j+1}) \middle| \frac{\partial \hat{\mathsf{U}}_j}{\partial \epsilon_j} \middle| \phi_k(t_j) \right\rangle\,, \end{aligned} \tag{3.45}$$

where $|\phi_k(t)\rangle$ is the forward-propagated basis state $|k\rangle$, and $|\chi(t)\rangle$ is the backward-propagated target state $\hat{\mathsf{O}}|k\rangle$. The numerical effort in calculating the gradient compared to a simple evaluation of the functional is therefore an additional backward propagation. Moreover, the states of either the backward or the forward propagation at every point in time need to be stored in order to calculate the gradient. The derivative of the j'th time evolution operator is given by

$$\frac{\partial \hat{\mathsf{U}}_j}{\partial \epsilon_j} = \frac{\partial}{\partial \epsilon_j} e^{-\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}_j dt} \equiv \sum_{n=1}^{\infty} \frac{(-\mathrm{i}\, dt/\hbar)^n}{n!} \sum_{k=0}^{n-1} \hat{\mathsf{H}}_j^k \left(\frac{\partial \hat{\mathsf{H}}_j}{\partial \epsilon_j} \right) \hat{\mathsf{H}}_j^{n-k-1}\,. \tag{3.46}$$

In Liouville space, $\hat{\mathsf{H}}$ is replaced by $\mathcal{L}$. For the two functionals $J_{T,\text{sm}}$ and $J_{T,\text{re}}$ of Eq. (3.39) and Eq. (3.40), the total gradient is

$$\frac{\partial J_{T,\text{sm}}}{\partial \epsilon_j} = -\frac{1}{N^2} \sum_{k=1}^{N} \sum_{l=1}^{N} \left[\frac{\partial \tau_l^*}{\partial \epsilon_j} \tau_k + \tau_l^* \frac{\partial \tau_k}{\partial \epsilon_j} \right] = -\frac{2}{N} \mathfrak{Re} \sum_{k=1}^{N} \sum_{l=1}^{N} \tau_l^* \frac{\partial \tau_k}{\partial \epsilon_j}\,, \tag{3.47}$$

$$\frac{\partial J_{T,\text{re}}}{\partial \epsilon_j} = -\frac{1}{N} \sum_{k=1}^{N} \frac{1}{2} \left(\frac{\partial \tau_k}{\partial \epsilon_j} + \frac{\partial \tau_k^*}{\partial \epsilon_j} \right) = -\frac{1}{N} \mathfrak{Re} \sum_{k=1}^{N} \frac{\partial \tau_k}{\partial \epsilon_j}\,. \tag{3.48}$$

Using only the gradient of the functional to steer the optimization will generally not yield sufficiently fast convergence; the situation could be improved by using Newton's method, taking into account the second derivative, i.e., the Hessian

$$H_{jj'} = \frac{\partial^2 J}{\partial \epsilon_j \partial \epsilon_{j'}}. \tag{3.49}$$

However, calculating the Hessian is generally prohibitively expensive. Therefore, quasi-Newton methods are employed [146]. The idea is to estimate the Hessian using only gradient information. There are several approaches of how to achieve this, the most popular one [17] is the BFGS method [147, 148, 149, 150]. A memory-efficient version of this method that also allows to define bounds on the control, named LBFGS-B [151] is widely available as a Fortran library [152].

It is crucial for the application of the LBFGS-B method that the gradient is calculated to full machine precision. The Taylor series in Eq. (3.46) would usually have to be evaluated to very high order. A more efficient and stable method for calculating the gradient is provided by the observation that [153]

$$\begin{aligned} \exp\left[-\frac{\mathrm{i}}{\hbar}\begin{pmatrix} \hat{\mathsf{H}}_j & \frac{\partial \hat{\mathsf{H}}_j}{\partial \epsilon_j} \\ 0 & \hat{\mathsf{H}}_j \end{pmatrix} dt\right] \begin{pmatrix} 0 \\ \vert\Psi\rangle \end{pmatrix} &= \begin{pmatrix} \hat{\mathsf{U}}_j & \frac{\partial \hat{\mathsf{U}}_j}{\partial \epsilon_j} \\ 0 & \hat{\mathsf{U}}_j \end{pmatrix} \begin{pmatrix} 0 \\ \vert\Psi\rangle \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial \hat{\mathsf{U}}_j}{\partial \epsilon_j} \vert\Psi\rangle \\ \hat{\mathsf{U}}_j \vert\Psi\rangle \end{pmatrix}. \end{aligned} \tag{3.50}$$

Eq. (3.50) can be efficiently evaluated by applying the Newton propagator of appendix F.

3.3.4 Krotov's Method

For time-continuous controls, Krotov's method [154] considers a functional of the form

$$J[\epsilon^{(i)}(t)] = J_T(\{\phi_k^{(i)}(T)\}) + \int_0^T g_a[\epsilon^{(i)}(t)]\,\mathrm{d}t + \int_0^T g_b[\{\phi_k^{(i)}(t)\}]\,\mathrm{d}t\,. \tag{3.51}$$

In addition to the final time functional, e.g. $J_{T,\mathrm{sm}}$ or $J_{T,\mathrm{re}}$ defined in Eq. (3.39) and Eq. (3.40), *running costs* that depend on the control field and the propagated states at each point in time are also included. As before, for a gate optimization, the $\{\vert\phi_k^{(i)}(t)\rangle\}$ are the basis states $\{\vert k\rangle\}$ propagated under the pulse $\epsilon^{(i)}(t)$ at the current OCT iteration i.

Krotov's method uses an auxiliary functional to disentangle the interdependence of the states and the field, allowing to find an updated $\epsilon^{(i+1)}(t)$ such that $J[\epsilon^{(i+1)}] < J[\epsilon^{(i)}]$ is guaranteed. The derivation, see Ref. [155], yields the condition

$$\begin{aligned}\left.\frac{\partial g_a}{\partial \epsilon}\right|_{\epsilon^{(i+1)}(t)} = 2\mathfrak{Im}\Bigg[&\sum_{k=1}^{N}\left\langle \chi_k^{(i)}(t)\right|\left(\left.\frac{\partial \hat{\mathsf{H}}}{\partial \epsilon}\right|_{\substack{\phi^{(i+1)}(t)\\ \epsilon^{(i+1)}(t)}}\right)\left|\phi_k^{(i+1)}(t)\right\rangle + \\ &+\frac{1}{2}\sigma(t)\left\langle \Delta\phi_k(t)\right|\left(\left.\frac{\partial \hat{\mathsf{H}}}{\partial \epsilon}\right|_{\substack{\phi^{(i+1)}(t)\\ \epsilon^{(i+1)}(t)}}\right)\left|\phi_k^{(i+1)}(t)\right\rangle\Bigg]\,,\end{aligned} \tag{3.52}$$

with

$$|\Delta\phi_k(t)\rangle \equiv |\phi_k^{(i+1)}(t)\rangle - |\phi_k^{(i)}(t)\rangle\,. \tag{3.53}$$

Assuming the equation of motion for the forward propagation of $|\phi_k(0)\rangle = |k\rangle$ is written as

$$\frac{\partial}{\partial t}\left|\phi_k^{(i+1)}(t)\right\rangle = -\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}^{(i+1)}\left|\phi_k^{(i+1)}(t)\right\rangle\,, \tag{3.54}$$

the co-states $|\chi_k\rangle$ are backward-propagated under the old pulse as

$$\frac{\partial}{\partial t}\left|\chi_k^{(i)}(t)\right\rangle = -\frac{\mathrm{i}}{\hbar}\hat{\mathsf{H}}^{\dagger\,(i)}\left|\chi_k^{(i)}(t)\right\rangle + \left.\frac{\partial g_b}{\partial\left\langle\phi_k\right|}\right|_{\phi^{(i)}(t)}\,, \tag{3.55}$$

with the boundary condition

$$\left|\chi_k^{(i)}(T)\right\rangle = -\left.\frac{\partial J_T}{\partial\left\langle\phi_k\right|}\right|_{\phi_k^{(i)}(T)}\,. \tag{3.56}$$

In Eq. (3.52), $\sigma(t)$ is a scalar function that must be properly chosen to ensure monotonic convergence. In many cases, it is sufficient to set $\sigma(t) \equiv 0$, in particular if the equation of motion is linear ($\hat{\mathsf{H}}$ does not depend on $|\phi_k(t)\rangle$), the functional J_T is convex, and no state-dependent constraints are used ($g_b \equiv 0$). Even for some types of state-dependent constraints $\sigma(t)$ may be set to zero, specifically for keeping the population in an allowed subspace [156]. However, a state-dependent constraint adds an inhomogeneity to the equation of motion for $|\chi_k(t)\rangle$.

Where $\sigma(t) \neq 0$ is required, it can be determined numerically as shown in Ref. [155]. In chapter 6, final-time functionals that depend higher than quadratically on the states are considered, while the equation of motion remains the linear Schrödinger equation. In this case,

$$\sigma(t) \equiv -\max\left(\varepsilon_A, 2A + \varepsilon_A\right)\,, \tag{3.57}$$

where ε_A is a small non-negative number that can be used to enforce strict inequality in the second order optimality condition. The optimal value for

A in each iteration can be determined numerically as [155]

$$A = \frac{2\sum_{k=1}^{N} \mathfrak{Re}\left[\langle \chi_k(T) \,|\, \Delta\phi_k(T)\rangle\right] + \Delta J_T}{\sum_{k=1}^{N} \left|\Delta\phi_k(T)\right|^2}\,, \tag{3.58}$$

with

$$\Delta J_T \equiv J_T(\{\phi_k^{(i+1)}(T)\}) - J_T(\{\phi_k^{(i)}(T)\})\,. \tag{3.59}$$

In order to obtain an explicit equation for $\epsilon^{(i+1)}(t)$, a state-dependent running cost g_a must be used, and usually takes the form

$$g_a[\epsilon(t)] = \frac{\lambda_a}{S(t)} \left(\epsilon(t) - \epsilon^{\text{ref}}(t)\right)^2\,, \tag{3.60}$$

with a scaling parameter λ_a and a shape function $S(t) \in [0,1]$. When ϵ^{ref} is set to the optimized field $\epsilon^{(i)}$ from the previous iteration,

$$g_a[\epsilon^{(i+1)}(t)] = \frac{\lambda_a}{S(t)} \left(\Delta\epsilon(t)\right)^2\,, \quad \Delta\epsilon(t) \equiv \epsilon^{(i+1)}(t) - \epsilon^{(i)}(t)\,, \tag{3.61}$$

and for $\sigma(t) \equiv 0$, the explicit first-order Krotov update equation is obtained [157, 138],

$$\Delta\epsilon(t) = \frac{S(t)}{\lambda_a} \mathfrak{Im} \left[\sum_{k=1}^{N} \left\langle \chi_k^{(i)}(t) \middle| \left(\left.\frac{\partial \hat{\mathsf{H}}}{\partial \epsilon}\right|_{\substack{\phi^{(i+1)}(t)\\ \epsilon^{(i+1)}(t)}} \right) \middle| \phi_k^{(i+1)}(t) \right\rangle \right]\,. \tag{3.62}$$

If $S(t) \in [0,1]$ is chosen as a function that smoothly goes to zero at $t = 0$ and $t = T$, then the update will be suppressed there, and thus a smooth switch-on and switch-off can be maintained. The scaling factor λ_a controls the overall magnitude of the pulse update. Values that are too large will change $\epsilon^{(i)}(t)$ by only a small amount, causing slow convergence.

If the reference field $\epsilon^{\text{ref}}(t)$ is set to zero, such that Eq. (3.60) is a penalty on the pulse fluence, the update equation turns into a simple replacement, where $\epsilon^{(i+1)}(t)$ is directly given by the right hand side of Eq. (3.62). However, the lack of an explicit dependence on the previous field leads to numerical instability. A better approach for penalizing large pulse amplitudes is to use the state-dependent running cost

$$g_a[\epsilon^{(i+1)}(t)] = \frac{\lambda_a}{S(t)} \left(\epsilon^{(i+1)}(t) - \epsilon^{(i)}(t)\right)^2 + \frac{\lambda_\epsilon}{S(t)} \left(\epsilon^{(i+1)}(t)\right)^2\,, \tag{3.63}$$

which leads to

$$\epsilon^{(i+1)}(t) = \frac{\lambda_a}{\lambda_a + \lambda_\epsilon} \epsilon^{(i)}(t) + \frac{S(t)}{\lambda_a + \lambda_\epsilon} \frac{\delta\epsilon(t)}{2}\,, \tag{3.64}$$

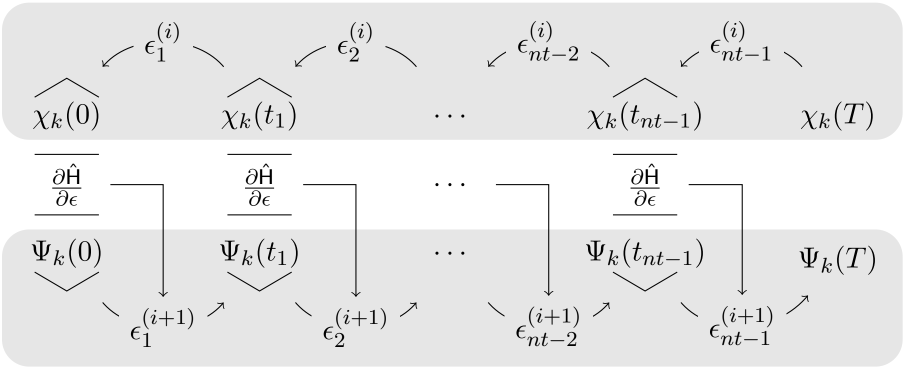


Figure 3.1: Sequential update scheme in Krotov's method on a time grid.

where $\delta\epsilon(t)$ is the right hand side of Eq. (3.52). It is also possible to derive update equations for choices of g_a that include spectral constraints [158, 159].

The functional J_T enters the first-order update equation only in the boundary condition for the backward propagated co-state, Eq. (3.56). For the standard functionals defined in Eq. (3.39) and Eq. (3.40), this evaluates to

$$-\left.\frac{\partial J_{T,\text{sm}}}{\partial \langle \phi_k |}\right|_{\phi_k^{(i)}(T)} = \left(\frac{1}{N^2}\sum_{l=1}^{N}\tau_l\right)\hat{\mathsf{O}}\,|k\rangle\ , \tag{3.65}$$

$$-\left.\frac{\partial J_{T,\text{re}}}{\partial \langle \phi_k |}\right|_{\phi_k^{(i)}(T)} = \frac{1}{2N}\hat{\mathsf{O}}\,|k\rangle\ . \tag{3.66}$$

If $\hat{\mathsf{H}}$ depends more than linearly on the field, the derivative

$$\left.\frac{\partial \hat{\mathsf{H}}}{\partial \epsilon}\right|_{\substack{\phi^{(i+1)}(t)\\ \epsilon^{(i+1)}(t)}}$$

yields an explicit dependence on $\epsilon^{(i+1)}(t)$ on the right hand side of Eq. (3.62). In this case, the usual approach is to enforce $\epsilon^{(i+1)}(t) \approx \epsilon^{(i)}(t)$ with a large value of λ_a. Alternatively, $\Delta\epsilon(t)$ may be determined in a self-consistent loop. This is especially relevant if instead of $\epsilon(t)$, a parametrization $\epsilon(u(t))$ is used, where $u(t)$ is the optimized control field. For example, $\epsilon(t) = u^2(t)$ is used to ensure that $\epsilon(t) > 0$, and

$$\epsilon(t) = \frac{\epsilon_{\text{max}} - \epsilon_{\text{min}}}{2}\tanh(u(t)) + \frac{\epsilon_{\text{max}} + \epsilon_{\text{min}}}{2} \tag{3.67}$$

keeps $\epsilon(t)$ bounded between ϵ_{min} and ϵ_{max} [160].

Discretization to a time grid yields the numerical scheme shown in Fig. 3.1, and resolves the seeming contradiction that the calculation of $\epsilon^{(i+1)}(t)$ requires knowledge of the states $|\Psi_k^{(i+1)}(t)\rangle$ propagated under $\epsilon^{(i+1)}(t)$. The scheme starts with $|\chi_k(T)\rangle$ obtained from Eq. (3.56), which is backward-propagated under Eq. (3.55). All backward-propagated states $|\chi(t)\rangle$ must be stored. The first pulse value is updated according to Eq. (3.62), using $|\chi_k(0)\rangle$ and the known initial state $|\Psi_k(0)\rangle = |k\rangle$. Then, $|\Psi_k(0)\rangle$ is forward-propagated by one time step under Eq. (3.54) using the updated pulse value. The updates proceed sequentially, until the final forward-propagated state $|\Psi_k(T)\rangle$ is reached. For numerical stability, it is useful to define the normalized

$$|\Psi_k^{\mathrm{bw}}(T)\rangle = \frac{1}{\|\chi_k\|} |\chi_k(T)\rangle \tag{3.68}$$

and then later multiply again with $\|\chi_k\|$ when calculating the pulse update.

At first glance, there is a striking similarity between the Krotov update formula (3.62) and the gradient in Eq. (3.45), except that for GRAPE/LBFGS-B *both* backward- and forward-propagation are performed with the same pulse $\epsilon^{(i)}(t)$. However, a closer look shows the key difference between GRAPE/LBFGS-B and Krotov's method: whereas the former is inherently discrete, the latter gives a continuous update equation that is discretized only afterwards. In fact the monotonic convergence of Krotov's method is only guaranteed in the continuous limit; a coarse time step must be compensated by larger values of λ_a, slowing down convergence. Generally, choosing λ_a too small will lead to numerical instabilities and unphysical features in the optimized pulse. A lower limit for λ_a can be determined from the requirement that the change $\Delta\epsilon(t)$ should be at most on the same order of magnitude as the guess pulse $\epsilon^{(i)}(t)$ for that iteration. The Cauchy-Schwarz inequality applied to the update equation (3.62) yields

$$\|\Delta\epsilon(t)\|_\infty \leq \frac{\|S(t)\|}{\lambda_a} \sum_k \|\chi_k\|_\infty \|\psi_k\|_\infty \left\| \frac{\partial \hat{\mathsf{H}}}{\partial \epsilon} \right\|_\infty \overset{!}{\leq} \left\| \epsilon^{(i)}(t) \right\|_\infty . \tag{3.69}$$

Since $S(t) \in [0,1]$ and $|\psi_k\rangle$ is normalized, the condition for λ_a becomes

$$\lambda_a \geq \frac{1}{\max |\epsilon^{(i)}(t)|} \left[\sum_k \|\chi_k\|_\infty \right] \left\| \frac{\partial \hat{\mathsf{H}}}{\partial \epsilon} \right\|_\infty . \tag{3.70}$$

From a practical point of view, the best strategy is to start the optimization with a comparatively large value of λ_a, and after a few iterations lower λ_a as far as possible without introducing numerical instabilities. The value of λ_a may be adjusted dynamically with the rate of convergence. Generally, the optimal choice of λ_a requires some trial and error. Inspired by gradient ascent, it has been proposed to employ quasi-Newton methods to determine the Krotov step width λ_a dynamically [161].

When using the rotating wave approximation (RWA), cf. appendix B, it is important to remember that the target transformation $\hat{\mathsf{O}}$ is usually defined in the lab frame, not in the rotating frame. This is relevant for the construction of $\vert\chi_k(T)\rangle$. The easiest approach is to transform the result of the forward propagation $\vert\phi_k(T)\rangle$ from the rotating frame to the lab frame, then construct $\vert\chi_k(T)\rangle$ for the next OCT iteration, and transform $\vert\chi_k(T)\rangle$ back to the rotating frame, before starting the backward-propagation for the next OCT iteration. When the RWA is used, the control-pulses are complex-valued. In this case, the Krotov update equation is valid for both the real and the imaginary part independently. That is, in the update for the real part of the pulse, all derivatives are also taken with respect to only the real part, and likewise for the imaginary part.

The control equations have been written in the notation of Hilbert space. However, they are equally valid for a gate optimization in Liouville space, by replacing states with density matrices, $\hat{\mathsf{H}}$ with $\mathcal{L}$, and inner products with Hilbert-Schmidt products. An explicit formulation is given in chapters 4 and 7.

3.3.5 Choosing an Optimization Method

Whether to use a gradient-free optimization method, gradient ascent, or Krotov's method depends on the size of the problem (both the Hilbert space dimension and the number of control parameters), the requirements on the control pulse, and the optimization functional. Gradient-free methods should be used if propagation is extremely cheap (small Hilbert space dimension), the number of independent control parameters is relatively small, or the functional is of a form that does not allow to calculate gradients.

Gradient ascent should be used if the control parameters are discrete, such as on a coarse-grained time grid, and the derivative of J with respect to each control parameter is known. Moreover, evaluation of the gradient must be numerically feasible. Parametrization on a time grid, as shown in section 3.3.3, directly yields an efficient scheme. Gradients for many other parametrizations (e.g. an expansion in Fourier coefficients) can be derived from that time grid gradient [162].

Krotov's method should be used if the control is near-continuous, and if the derivative of J_T with respect to the states, Eq. (3.56), can be calculated. Sometimes, a functional must be rewritten in an alternative form before this is possible, as illustrated in chapter 6. When these conditions are met, Krotov's method gives excellent convergence, although it is often observed to slow down when getting close to the minimum of J. Since quasi-Newton gradient ascent does not show such a slow-down, it can be beneficial to switch from Krotov's method to LBFGS-B in the final stage of the optimization.

4

Robustness through Ensemble Optimization

The numerical design of robust quantum gates must address two issues. The first is the unwanted interaction with the environment, inducing decoherence. In order to take decoherence effects into account, the system must be modeled using the formalism of Liouville space, as reviewed in chapter 2, and include the dominant decoherence channels in the equations of motion. Applying the optimal control techniques (OCT) of chapter 3 allows to find pathways that avoid these dominant channels and still yield unitary quantum gates with high fidelity. Beyond that, however, there are also imperfections, uncertainties, and fluctuations in the classical parameters of the experimental setup: variations in the control, or extraneous electromagnetic fields. These can be modeled as variations of the energies and coupling strengths in the Hamiltonian. The challenge of real-world quantum computing is to implement quantum gates that are robust with respect to *both* decoherence and macroscopic fluctuations.

This chapter, adapted from Ref. [163], illustrates an approach for obtaining such robust gates, for the example of a CPHASE gate on trapped Rydberg atoms: we apply Krotov's method, section 3.3.4 of chapter 3, in a Liouville space formulation, and include classical fluctuations through an ensemble approach, obtaining an optimized control field that induces the desired gate not just under the ideal Hamiltonian, but also Hamiltonians with varied parameters. Similar approaches have been used in the context of NMR spectroscopy [164, 165].

4.1 Implementation of Quantum Gates with Rydberg Atoms

Rydberg states of trapped neutral atoms provide an attractive platform for the realization of quantum information processing. The qubit is encoded in two hyperfine states of the electronic ground state; two atoms in these states are effectively non-interacting and well-isolated from the environment. Using a focused laser beam, the atoms may be excited to a Rydberg state with high principal quantum number. When both atoms are in the Rydberg state, they feel a strong dipole-dipole interaction [166]. Thus, by exciting to the Rydberg state, the interaction between two relatively distant atoms may be switched on and off, generating entanglement and allowing the implementation of a quantum gate [167, 168].

Arrays of trapped neutral atoms have been realized in optical lattices with single site occupancy [169, 170], optical tweezers [171], or on atom chips [172]. Atoms can be kept at a well-defined relative distance and be addressed individually [173, 174, 175, 176]. Therefore, they have the potential to provide a scalable architecture for quantum computing, if fast, high-fidelity quantum gates can be implemented.

The dipole-dipole interaction between two atoms in the Rydberg state, shifting the energy of the two-atom state $|rr\rangle$ by the interaction energy u, results in a blockade effect [177, 171]. If u is significantly larger than the energy of a laser tuned to the transition $|0\rangle \rightarrow |r\rangle$ of the atom, the laser will be far off-resonant for reaching $|rr\rangle$. Thus, exciting one atom to $|r\rangle$ blocks the second atom from being excited. If the laser energy is much larger than u, the blockade may be broken, as the transition to $|rr\rangle$ is now only weakly detuned.

Both in the blockade and the non-blockade regime, there are proposals for the implementation of quantum gates [178]. Outside the blockade regime, single-site addressability of the atoms is not required. Exciting the atoms to the $|rr\rangle$ state for a duration of T, a two-qubit phase of $\gamma = uT$ is generated. However, excitation to the $|rr\rangle$ state can induce atomic motion, limiting the fidelity. Using optimal control, the motional degree of freedom can be restored; it has been shown that errors on the order of 10^{-3} are attainable [73, 179].

Here, we focus on the blockade regime and assume that the atoms can be individually addressed. A CPHASE gate can then be implemented by the scheme shown in Fig. 4.1 [178]. First, a π-pulse is applied only to the left atom, transferring population from $|0\rangle$ to $|r\rangle$, with a phase factor of $\mathrm{e}^{\mathrm{i}\frac{\pi}{2}}$, see appendix C. The $|1\rangle$ level is completely isolated and is thus not affected by the pulse. The second step is a 2π-pulse on the right atom. Because of the Rydberg blockade, the $|r0\rangle$ and $|r1\rangle$ state are unaffected, whereas

$$\begin{array}{ccccccc}
 & \pi\text{-pulse} & & 2\pi\text{-pulse} & & \pi\text{-pulse} & \\
 & \text{(left)} & & \text{(right)} & & \text{(left)} & \\
\vert 00\rangle & \rightarrow & \mathrm{i}\,\vert r0\rangle & \rightarrow & \mathrm{i}\,\vert r0\rangle & \rightarrow & -\vert 00\rangle \\
\vert 10\rangle & \rightarrow & \vert 10\rangle & \rightarrow & -\vert 10\rangle & \rightarrow & -\vert 10\rangle \\
\vert 01\rangle & \rightarrow & \mathrm{i}\,\vert r1\rangle & \rightarrow & \mathrm{i}\,\vert r1\rangle & \rightarrow & -\vert 01\rangle \\
\vert 11\rangle & \rightarrow & \vert 11\rangle & \rightarrow & \vert 11\rangle & \rightarrow & \vert 11\rangle
\end{array}$$

Figure 4.1: Scheme for the implementation of a CPHASE gate with individually addressable atoms in the Rydberg blockade regime

$\vert 10\rangle \rightarrow \mathrm{i}\vert 1r\rangle \rightarrow -\vert 10\rangle$. Lastly, another π-pulse on the left atom completes the gate.

While conceptually, the available schemes provide a clear recipe for the implementation of quantum gates, their experimental realization holds considerable challenges. While the dipole-dipole interaction between two atoms in the Rydberg states potentially allows for fast gates, there are a number of intrinsic and technical sources of error that can restrict both the achieved fidelity and the speed of operation.

The transition $\vert 0\rangle \rightarrow \vert r\rangle$ is not directly accessible by available laser frequencies. Therefore, it must be implemented as a two-photon transition via an intermediary level. This complicates the implementation of the necessary π- and 2π-pulses for the gate scheme in Fig. 4.1 considerably. Loss from the intermediary level has been identified as one of the primary technical challenges in the implementation of a gate, together with fluctuations in pulse amplitude, and fluctuations in the Rydberg level, which is highly sensitive to external electromagnetic fields [180].

In section 4.3, we consider several analytic pulse sequences that implement the gate scheme via an intermediary level, realizing effective Rabi-pulses either via adiabatic elimination [181] or a STIRAP process [182], and systematically analyze their robustness with respect to the primary fluctuating parameters. Even in the best case, we find that gate fidelities drop by several percent over a realistic range of fluctuations. Moreover, gate durations of at least 800 ns are required. In related existing proposals to mitigate the effects of intrinsic errors in Rydberg gates using adiabatic passage techniques [183], the resulting pulse sequences typically also result in relatively long gate times of μs or longer [184, 185].

We therefore turn to optimal control, using Krotov's method presented in chapter 3, applied to the density matrix formalism in order to take into account spontaneous emission from the intermediary level. Secondly, we obtain robustness with respect to fluctuations in pulse amplitude, expressed through the dipole moment, and Rydberg energies, expressed through the two-photon detuning. The results of the optimal control yield systematically

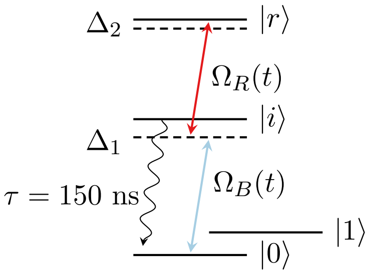


Figure 4.2: Level scheme for a single atom.

higher gate fidelities than all analytic approaches, showing improvement of an order of magnitude to reach gate errors of order 10^{-4} for equivalent gate times. Most importantly, optimal control can deliver gate performance that is also extremely robust with respect to experimental fluctuations, with the gate error staying below or at the order of 10^{-3} even for large fluctuations, i.e. below the quantum correction limit [186]. Lastly, optimal control also allows to address the issue of gate duration. By systematically decreasing the gate time until the optimization no longer yields a gate of sufficiently high fidelity [107], a "quantum speed limit" for performing the gate [187] can be determined. Compared to the analytical solutions, using optimal control we can significantly shorten the total gate duration, to $\sim$100 ns, without loss in either robustness or fidelity.

4.2 Model

We consider two cesium atoms trapped in an optical lattice with single-site addressability. The qubit states are encoded in hyperfine levels of the ground state, $|0\rangle = |6\,^2S_{1/2}, F = 3\rangle$, $|1\rangle = |6\,^2S_{1/2}, F = 4\rangle$.

For technical reasons, the Rydberg level, here $|r\rangle = |50D_{3/2}\rangle$, is accessed by a two-photon transition via an intermediate state [171, 167], $|i\rangle = |7P_{3/2}\rangle$. In the basis $\{|0\rangle, |1\rangle, |i\rangle, |r\rangle\}$, the Hamiltonian for a single atom, using a two-photon rotating-wave approximation, see appendix B, reads

$$\hat{\mathsf{H}}_{1\mathrm{q}} = \begin{pmatrix} 0 & 0 & \Omega_B(t) & 0 \\ 0 & E_1 & 0 & 0 \\ \Omega_B(t) & 0 & \Delta_1 & \Omega_R(t) \\ 0 & 0 & \Omega_R(t) & \Delta_2 \end{pmatrix}, \tag{4.1}$$

where $\Omega_B(t), \Omega_R(t)$ are the Rabi frequencies of the 'blue' and 'red' pulses, $\Omega_\alpha = \frac{1}{2}\mu_{ij}^{\alpha} E(t)$, cf. Fig. 4.2, and Δ_1, Δ_2 are the one-photon and two-photon detunings. The two atoms are kept at a distance of 5 µm such that their interaction is negligible except when both atoms are in the Rydberg state.

Single-Photon Detuning	Δ_1	=	1.273	GHz
Two-Photon Detuning	Δ_2	=	0	
Qubit Energy	E_1	=	9.100	GHz
Interaction Energy	u	=	57.26	MHz
Intermediate State Lifetime	τ	=	150.0	ns

Table 4.1: Rydberg atom system parameters

The Hamiltonian for the two atoms, including their Rydberg interaction, is written as

$$\hat{\mathsf{H}}_{2q} = \hat{\mathsf{H}}_{1q} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathsf{H}}_{1q} - u \left|rr\right\rangle \left\langle rr\right| , \tag{4.2}$$

with interaction energy u. The parameters are summarized in Table 4.1. Rabi frequencies of $\Omega_B = 171.5\,\mathrm{MHz}$ and $\Omega_R = 148.4\,\mathrm{MHz}$ have been implemented for this system and values up to ~250 MHz are expected to be experimentally feasible [188]. Note that we can restrict our model to a single m_J-sublevel of the Rydberg state (e.g., $m_J = 5/2$) by choice of laser polarization, such that the shift due to the interaction u is insensitive to magnetic field [166] and may then be ignored. In order clarify the physics of the alternate excitation path to the excited state, we have approximated the intermediate state as a single level. In practice, since the alternate excitation path is approximately resonant with the intermediate state, it would be desirable to also include details of the intermediate state hyperfine structure in the control optimization. This might affect the optimal laser detuning but will not cause any qualitative changes in the results obtained here.

The intermediate level undergoes spontaneous decay to the ground state. Thus, the full dynamics must be described by a master equation in Lindblad form,

$$\frac{\partial}{\partial t}\hat{\rho}(t) = -\frac{\mathrm{i}}{\hbar}[\hat{\mathsf{H}}_{2q}(t), \hat{\rho}(t)] + \mathcal{L}_D(\hat{\rho}(t)) . \tag{4.3}$$

$\mathcal{L}_D$ is the dissipator for the spontaneous decay of each atom,

$$\mathcal{L}_D(\hat{\rho}) = \frac{1}{\hbar\tau} \sum_{i=1,2} \left(\hat{\mathsf{A}}_i \hat{\rho} \hat{\mathsf{A}}_i^\dagger - \frac{1}{2} \left\{ \hat{\mathsf{A}}_i^\dagger \hat{\mathsf{A}}_i, \hat{\rho} \right\} \right) , \tag{4.4}$$

with $\hat{\mathsf{A}}_1 = \left|0\right\rangle \left\langle i\right| \otimes \mathbb{1}$, $\hat{\mathsf{A}}_2 = \mathbb{1} \otimes \left|0\right\rangle \left\langle i\right|$, and τ the lifetime of state $\left|i\right\rangle$.

Resonant excitation of both atoms to the Rydberg state leads to an acceleration of the atoms towards each other due to the dependence of the Rydberg interaction strength on interatomic separation [178]. The minimum gate duration is then determined either by the inverse of the interaction, u, or by the period of the atomic motion in the trap [73]. The gate duration may be limited further by the inverse of the experimentally realizable Rabi frequencies.

We consider here the Rydberg blockade regime which avoids resonant excitation into $|rr\rangle$. It corresponds to

$$\Omega_{\text{eff}} \ll u, \tag{4.5}$$

where [189]

$$\Omega_{\text{eff}} \approx \frac{\Omega_R \Omega_B}{2\Delta_1}\,. \tag{4.6}$$

The original proposal of the Rydberg gate [178] in this regime requires the atoms to be individually addressable, and employs the pulse sequence shown in Fig. 4.1. If the qubits are initially in $|00\rangle$, a non-local phase is accumulated during the middle pulse because of the detuning of level $|rr\rangle$ due to the interaction, u, and we can thus execute a CPHASE gate. This is in principle feasible with the experimental setup of Ref. [173].

We quantify success in terms of the gate error defined as $1 - F_{\text{avg}}$, where F_{avg} is the average gate fidelity, Eq. (2.91). The fidelity is calculated with respect to the target CPHASE gate, for $\hat{\mathsf{U}}$, the projection of the time evolution operator onto the logical subspace ($\hat{\mathsf{U}}$ is unitary if and only if there is no loss from that subspace at final time T).

4.3 Analytic pulse sequences

When a resonant two-photon transition is employed via an intermediate level, the two-level system $\{|0\rangle, |r\rangle\}$ for one atom in the original proposal [178] is replaced by $\{|0\rangle, |i\rangle, |r\rangle\}$. The π and 2π population flips can then be realized either with two simultaneous pulses, namely Ω_B connecting $|0\rangle$ and $|i\rangle$ and Ω_R connecting $|i\rangle$ and $|r\rangle$; or via a STIRAP process [182], where Ω_R acts as a “Stokes” pulse, preceding but overlapping Ω_B, the “pump” pulse. Both methods may be combined in a mixed scheme, where a STIRAP sequence is used for the π flip acting on the left atom, while the 2π flip on the right atom is realized using simultaneous pulses. The pulse schemes should avoid putting any population in the intermediary state; if this condition is satisfied there are no relevant dissipative effects and the dynamics can be described in Hilbert space, using the time-dependent Schrödinger equation in place of Eq. (4.3). The following sections discuss the merits and drawbacks of all three approaches, and numerically analyze the robustness with respect to pulse timing, fluctuations of the Rydberg level, and fluctuations of the pulse amplitude. We find that the mixed scheme is the most robust of these analytic approaches, due to its selective use of STIRAP on the control qubit only. However, all STIRAP based schemes are found to require either large pulse amplitudes or exceedingly long pulse times.

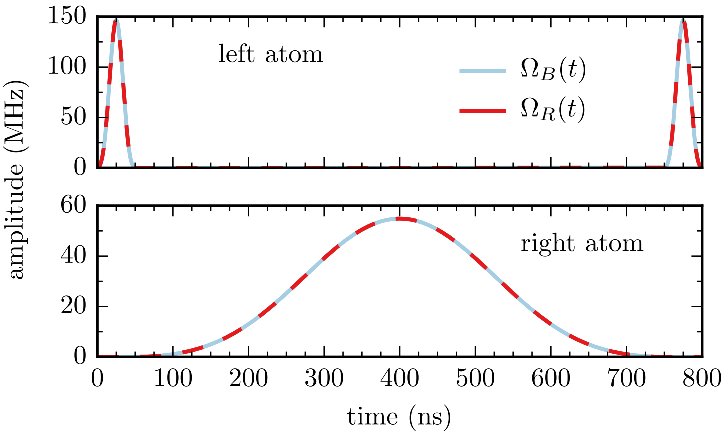


Figure 4.3: Three sequential Blackman pulse pairs implementing a CPHASE gate.

4.3.1 Sequence of three simultaneous pulse pairs

We first consider the realization of all population transfers using simultaneous pulse pairs. The pulses are of Blackman shape,

$$\Omega(t) = \frac{E_0}{2}\left(1 - a - \cos\left(\frac{2\pi t}{T}\right) + a\cos\left(\frac{4\pi t}{T}\right)\right)\,, \tag{4.7}$$

with $a = 0.16$ and E_0 the peak amplitude. This pulse shape is essentially identical to a Gaussian centered at $T/2$ with a width of $\sigma = T/6$, but, unlike the Gaussian, is exactly zero at $t = 0$ and $t = T$. Other pulse shapes are possible.

A pulse sequence that realizes the two π-flips on the left atom and one 2π-flip on the right atom is shown in Fig. 4.3. Due to the large single photon detuning of 1.3 GHz, the intermediate level can be adiabatically eliminated. This places a restriction on the pulse amplitude,

$$\Omega_B, \Omega_R \ll \Delta_1\,. \tag{4.8}$$

The 2π pulse is more stringently restricted by the blockade condition in Eq. (4.5). With the pulse duration being inversely proportional to the pulse amplitude, both effects result in a quantum speed limit.

Quantitatively, the limitations are illustrated in Fig. 4.4 which shows the gate error (black solid line) vs. duration of the middle 2π pulse, using a duration of 50 ns for the initial and final π pulse. The breakdown of adiabatic elimination becomes apparent in the peak population of the $|0i\rangle$ state (blue short-dashed line), whereas a breaking of the Rydberg blockade is observed in the peak population in the $|rr\rangle$ state (orange long-dashed line). Gate errors below 10^{-3} are only reached for pulse durations of $\geq$800 ns. The gate time is dominated by the central 2π pulse, which must be sufficiently

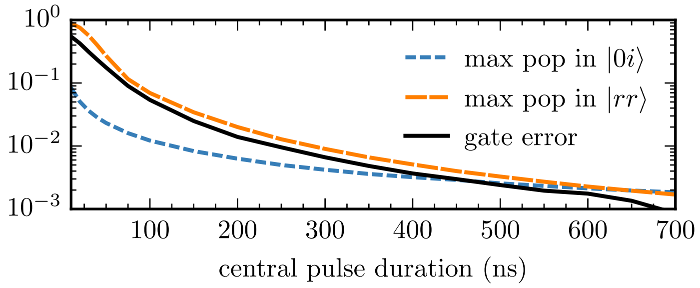


Figure 4.4: Quantum speed limit for the Rydberg gate using simultaneous Blackman pulse pairs. The time window is only that of the center 2π pulse in the scheme. As a measure of the breakdown of the Rydberg blockade, the maximum population in the $|rr\rangle$ state during that pulse is shown, as well as the maximum population in the $|01\rangle$ state, as a measure of the breakdown of the adiabatic elimination of the intermediate level. Finally, we show the total gate error obtained when combining the center 2π pulse of the given duration with two 50 ns π pulses on the left atom.

weak to not break the Rydberg blockade. Already, the pulse amplitude is remarkably close to the interaction energy, pushing the limits of condition (4.5). Note that the choice of identical peak Rabi frequencies for the red and blue laser, $\Omega_{B,\max} = \Omega_{R,\max}$, is the only ratio possible to guarantee complete population inversion in a three-level system using simultaneous pulses when the intermediate level is adiabatically eliminated [189].

Population and phase dynamics obtained with simultaneous red and blue pulses are shown in Fig. 4.5. As described in section 4.2, the population undergoes a π Rabi cycle on the left atom, followed by a 2π pulse on the right atom, followed by a π pulse on the left atom, cf. Fig. 4.5 (a,b,c). The intermediate levels (red long-dashed lines) receive almost no population. Thus, for this time scale, spontaneous decay is not an issue. As can be seen from Fig. 4.5 (f), the non-local phase is accumulated in the $|00\rangle$ state entirely during the central 2π pulse. Although the Rydberg blockade is not broken, and the population remains in $|r0\rangle$, the state accumulates an additional phase due to the detuned pulse driving the transition out of $|r0\rangle$. This additional phase is critical for the success of the gate.

4.3.2 Sequence of STIRAP pulse pairs

STIRAP is a popular scheme to achieve population transfer in three-level systems, avoiding population in the intermediate level at all times [190]. It is based on adiabatically following a dynamic dark state that does not contain an $|i\rangle$-component. In our setup, the scheme for transferring population from $|0\rangle$ to $|r\rangle$ is realized by first switching on the red laser, acting as a “Stokes” pulse, followed by the blue laser, acting as the “pump” pulse. The

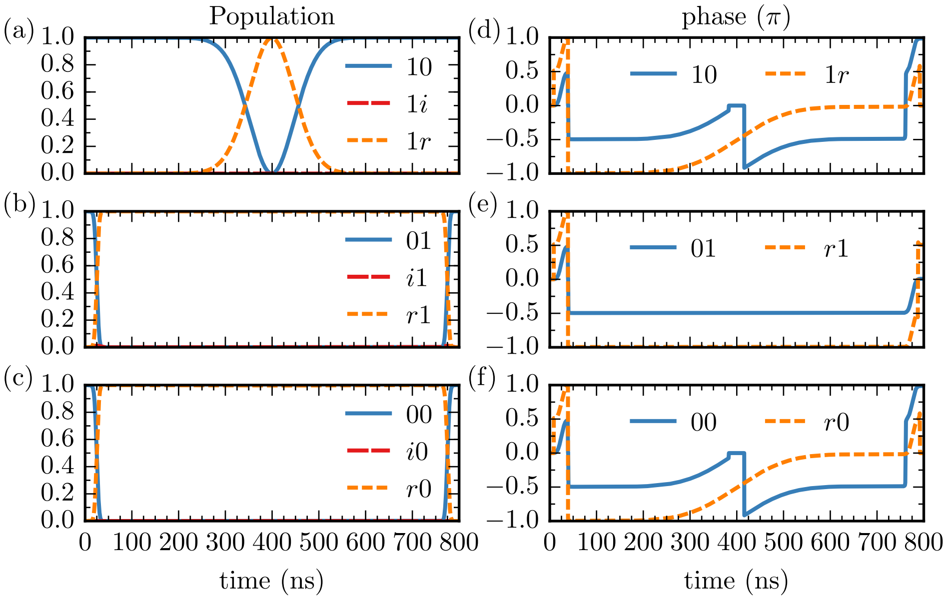


Figure 4.5: Population and phase dynamics using the simultaneous pulses shown in Fig. 4.3. Since the population in the intermediary states $|1i\rangle$, $|i1\rangle$, $|i0\rangle$ are effectively zero throughout, there are not included in the phase dynamics, panels (d-f).

two pulses must overlap, but the process is robust with respect to the laser amplitude and the exact overlap of the pulses, as long as the condition for adiabatic following, roughly given by [190]

$$\Omega_B \Delta\tau, \Omega_R \Delta\tau \gg 10 \tag{4.9}$$

is met, where $\Delta\tau$ is the time for which the pulses overlap. Thus, for short pulses, large amplitudes are required. However, for a Rydberg gate, the blockade condition, Eq. (4.5), also needs to be fulfilled, limiting the maximum Rabi frequency. Therefore STIRAP can only employ comparatively long pulses for the center 2π Rabi flip on the right atom.

In order to quantify the violation of the blockade condition, we define the “blockade efficiency” as

$$B = \max(P_{1r}) - \frac{1}{2}P_{1r}(T) - \left(\max(P_{rr}) - \frac{1}{2}P_{rr}(T)\right), \tag{4.10}$$

where T is the total time of the pulse sequence and P_{1r} and P_{rr} are the population in $|1r\rangle$ and $|rr\rangle$, respectively. B takes values between zero and one, with one corresponding to a perfect blockade. Both maximum and final-time populations appear in B because in order to have full Rabi cycling, the Rydberg level must be fully populated (giving a maximum population of one) and then fully depopulated (giving a final population

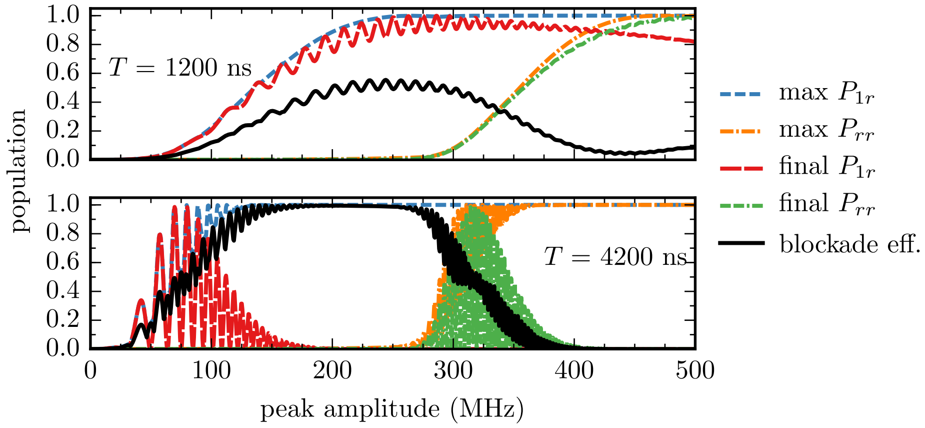


Figure 4.6: Breakdown of the Rydberg blockade for STIRAP: Only long gate durations allow for amplitudes that are sufficiently large to ensure adiabaticity in STIRAP while being small enough not to break the Rydberg blockade (lower panel).The amplitude of the two central pulse pairs are systematically scanned while the amplitude of the first and last pulse pairs are kept constant.

of zero), i.e., considering only the maximum population does not allow for distinguishing between π and 2π pulses. We only obtain $B = 1$ when the population completes a 2π cycle through $|1r\rangle$ whenever the system begins in $|10\rangle$, but never reaches $|rr\rangle$ whenever the system begins in $|00\rangle$. The blockade condition, Eq. (4.5), depends on the peak amplitude of the pulses whereas the adiabaticity condition, Eq. (4.9), depends on the pulses' complete Rabi angle. For short central 2π pulses the Rabi angle will not be sufficiently large to satisfy the adiabaticity condition without requiring a peak amplitude so high that it will break the blockade. This is illustrated in Fig. 4.6 (top), where for small amplitudes both the maximum and final $|1r\rangle$ populations rise together: the Rabi angle is less than π. Then, as the final $|1r\rangle$ population begins to fall such that the adiabaticity condition of STIRAP is better fulfilled, the blockade is broken, causing the drop in the blockade efficiency, concurrent with a rise in both the maximum and final $|rr\rangle$ populations. In Fig. 4.6 (bottom), the maximum and final $|1r\rangle$ populations rise together, but $|1r\rangle$ is now fully depopulated, thus achieving full Rabi cycling, before breaking the blockade. This corresponds to the area where $B \approx 1$ seen in the graph. We do not see a rise in the maximum and final $|rr\rangle$ population until high peak amplitudes.

A corresponding sequence of STIRAP pulse pairs, using short pulses on the left atom and long pulses on the right atom, is shown in Fig. 4.7. In principle, the pulses on the left atom can be made arbitrarily short, at the expense of extremely large field amplitudes. Taking into account realistic restrictions on the available laser power, the gate time will generally become prohibitively large.

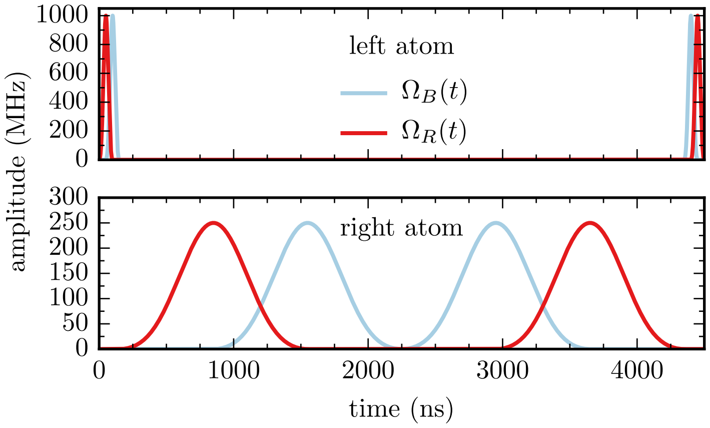


Figure 4.7: A sequence of STIRAP pulse pairs to implement the Rydberg CPHASE gate: While the pulses acting on the left atom can be made very short (limited effectively by the power of the driving laser), the pulses acting on the right atom need to be sufficiently long to avoid breaking the Rydberg blockade.

Additionally, since the STIRAP pulses are so robust to two-photon detuning [191], STIRAP will, to some extent, resolve the non-resonant Rydberg levels that are not explicitly considered, leading to unwanted population dynamics. Even if this population transfer to extraneous levels is invertible, it will lead to undesired phase accumulation as the higher and lower energy levels rotate with different frequencies than the rotating frame.

4.3.3 Mixed scheme: STIRAP-π-pulses and simultaneous 2π-pulses

The primary drawbacks of the simultaneous pulses are the unwanted population in the intermediate level for the pulses acting on the left atom and a relatively large sensitivity of the pulses to variations in pulse area. On the other hand, the primary drawback of STIRAP is the breakdown of the Rydberg blockade, which results from employing an extremely long pulse acting on the right atom. This issue, however, is not present when using STIRAP for the pulses acting on the left atom. We therefore investigate a mixed scheme, consisting of STIRAP pulses to drive the π rotations on the left atom and simultaneous pulses to drive the 2π rotation on the right atom, cf. Fig. 4.8. By doing so we use each method where it is most effective. Furthermore, the pulses on the left and right atom can be overlapped without any appreciable loss in fidelity. This is because the pulses acting on the right atom only drive significant population transfer during the central third of the pulses. As long as the left atom is populated by the time the amplitude of the pulses acting on the right atom becomes

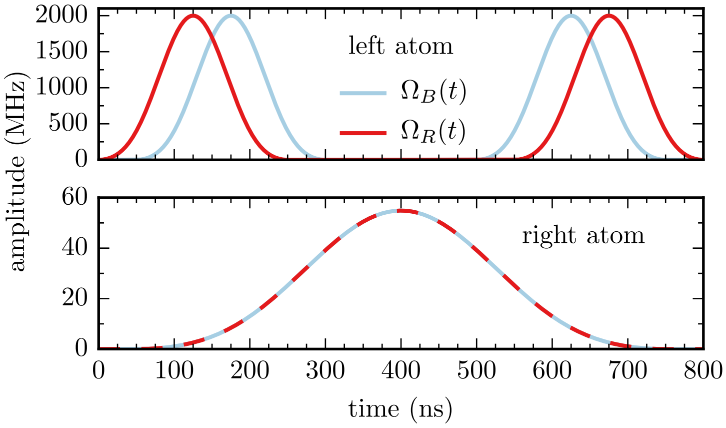


Figure 4.8: Mixed scheme: STIRAP pulse pairs for robust population transfer on the left atom, and simultaneous pulses for the 2π rotation of the right atom.

significant, the blockade is still effective. The two STIRAP pulses acting on the left atom, that bookend the central pulses acting on the right atom, are moved in towards the center. In fact the pulses can be compressed quite far: By overlapping the STIRAP pulses with the central pulses for 250 ns (see Fig. 4.8), the gate duration can be reduced from 1300 ns to 800 ns. The gate duration in the mixed scheme is limited by the laser power available for driving the left atom.

4.3.4 Robustness

For all three variants of pulse sequences, the gate fidelity in an actual experiment will be compromised by noise and experimental inaccuracies. In the following, we consider three main sources of errors: inaccuracies in timing between the pulses acting on the left and right qubit, inaccuracies in pulse amplitudes, and fluctuations of the Rydberg level due to, e.g., the presence of DC electric fields [160]. The latter results in a non-zero two-photon detuning. To analyze the robustness with respect to all of these fluctuations, we determine the expectation value of the gate fidelity under the assumption that the timing offset, the transition dipole, and the two-photon detuning differ from the optimal values by Δ_{time}, Δ_{Ω}, and Δ_{ryd} drawn from a Gaussian distribution centered at 0 of width σ_{time}, σ_{Ω}, and σ_{ryd}, respectively. For the pulse amplitudes, the variation is given in percent of the original amplitudes. The expectation value of the avergage gate fidelity is given by

$$\tilde{F}(\sigma_x) = \int \frac{1}{\sqrt{2\pi\sigma_x^2}} \mathrm{e}^{-\frac{\Delta x}{2\sigma_x^2}} F_{\text{avg}}(\Delta x)\,\mathrm{d}x, \qquad (4.11)$$

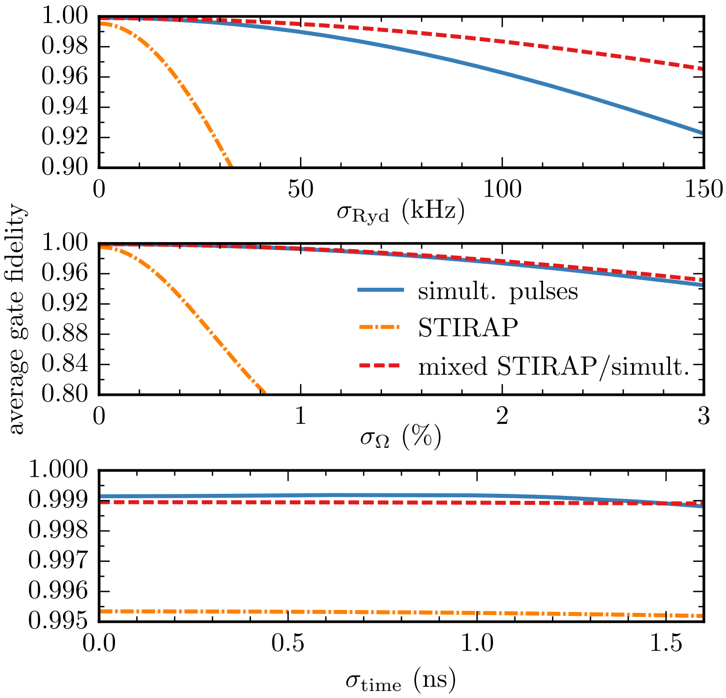


Figure 4.9: Robustness of the Rydberg gate with respect to Rydberg level fluctuations (top), amplitude fluctuations (middle), and fluctuations in the relative timing between pulses acting on the left and right atom. All fluctuations are drawn from a Gaussian distribution of width σ_{ryd}, σ_{Ω}, and σ_{time}, respectively. Note that the (%) in the middle panel refers to the percent by which the each pulse was uniformly scaled down. For the Rydberg level, $\sigma_{\mathrm{ryd}} = 150\,\mathrm{kHz}$ represents a 0.3% variation of $u = 57\,\mathrm{MHz}$.

with $\sigma_x = \sigma_{\mathrm{time}}, \sigma_{\Omega}, \sigma_{\mathrm{ryd}}$, and $\Delta x = \Delta_{\mathrm{time}}, \Delta_{\Omega}, \Delta_{\mathrm{ryd}}$, and F_{avg} given by Eq. (2.91). Sampling over 1000 variations of each parameter allows to evaluate the integral in Eq. (4.11) numerically.

Figure 4.9 shows the resulting expectation value of the gate fidelity vs. standard deviation of the fluctuations in pulse timings, pulse amplitudes, and energy of the Rydberg level. The gate is found to be very robust with respect to pulse timings and fairly robust with respect to amplitudes: only errors of more than a few nanoseconds in timing and several per cent in amplitude reduce the gate fidelity appreciably. A larger sensitivity is found with respect to the position of the Rydberg level: Fluctuations on the order of 1% of the interaction energy $u = 57\,\mathrm{MHz}$ reduce the gate fidelity to around 0.5 even for the most robust scheme, and even those on the order of 0.1% of u reduce the fidelity appreciably, cf. top panel of Fig. 4.9. This is not surprising, since a 'wrong' energy of the Rydberg level leads to a non-zero two-photon detuning, Δ_2, and thus affects both the population transfer for the left atom and the non-local phase accumulated during the pulse acting on the right atom. This additional phase is by assumption unknown and thus cannot be accounted for. Depending on the choice of the Rydberg level, the fluctuations of the level energy may be suppressed down to 100 kHz or less [192]. Gate fidelities of about 0.98 are then within reach, as shown in the upper panel of Fig. 4.9.

Though all the schemes behave similarly with respect to variations in timing, there are significant differences in each scheme's robustness to fluctuations in pulse amplitude and Rydberg level energy. For inaccuracies in pulse amplitude, cf. Fig. 4.9 (middle), the fidelity achieved with STIRAP pulses is far more susceptible to small variations than both other schemes. This is due to the additional phase accumulated for STIRAP during the central pulse acting on the right atom, caused by undesired population entering $|ri\rangle$, cf. section 4.3.2. Additional corrections would be required to compensate undesired phase evolution [185]. The mixed scheme performs slightly better than the simultaneous scheme in this respect, as the robust STIRAP pulses acting on the left atom can achieve efficient population transfer at a wide variety of amplitudes. With respect to fluctuations in the energy of the Rydberg level, in Fig 4.9 (top) the longer a given scheme populates $|r0\rangle$, the less robust that scheme is. When the population is in the detuned $|r0\rangle$ state, it accumulates an undesired phase, and this, not the loss in population transfer efficiency, is the primary reason for the drop in fidelity. The longer a scheme remains in $|r0\rangle$, the longer it takes to accumulate this additional phase. The mixed scheme, which overlaps the pulses acting on the left and right atom and thus populates $|r0\rangle$ for the shortest time possible, is the most robust to fluctuations in the Rydberg level energy. This is followed by the simultaneous scheme, which fully populates $|r0\rangle$ for 700 ns, and finally the STIRAP scheme, which fully populates $|r0\rangle$ for 4200 ns. Counterintuitively, then, the schemes actually are less robust with respect to variations in Rydberg level energy the longer they become.

4.4 Optimal control

The use of optimal control theory (OCT) allows to obtain non-analytic pulses that are not bound by conditions of adiabaticity, and that can realize gate times at the quantum speed limit [73, 193, 160]. Here, we extend the application of optimal control to increase the robustness of the pulses with respect to fluctuations in amplitude and the energy of the Rydberg level due to external fields. This is achieved by requiring the gate fidelity, Eq. (2.88), to be close to one not only for the ideal Hamiltonian $\hat{\mathsf{H}}_0$, Eq. (4.2), but also for an ensemble of perturbed Hamiltonians $\{\hat{\mathsf{H}}_n\}$, $n = [1, N-1]$ that sample the relevant parameter space of variations. Unlike in the analytical pulse schemes, the optimized control pulses will not consist of sub-pulses, but will be completely overlapping. Therefore, an analysis of the robustness with respect to pulse timing is not meaningful in this context.

4.4.1 Control Equations

Since population in the intermediate state $|i\rangle$ should be avoided, we now perform the optimization in Liouville space, and include spontaneous emission for the intermediate level with a lifetime of 150 ns [194, 195]. The optimization functional to be minimized reads

$$J = 1 - \frac{1}{16N} \sum_{n=0}^{N-1} \sum_{k=1}^{16} \mathfrak{Re} \left\{ \operatorname{tr} \left[\hat{\mathsf{O}} \hat{\rho}_k(0) \hat{\mathsf{O}}^\dagger \hat{\rho}_{k,n}(T) \right] \right\} - \sum_{j=1}^{4} \lambda_j \int_0^T \frac{(\Delta\Omega_j(t))^2}{S(t)} \, \mathrm{d}t \,, \tag{4.12}$$

with

$$\Delta\Omega_j(t) = \Omega_j^{(i+1)}(t) - \Omega_j^{(i)}(t) \,. \tag{4.13}$$

in iteration OCT iteration i. The first part of J is a final time cost that measures the Hilbert-Schmidt overlap of the propagated states $\hat{\rho}_{k,n}(T)$ with the target states $\hat{\mathsf{O}}\hat{\rho}_k(0)\hat{\mathsf{O}}^\dagger$, where $\hat{\mathsf{O}}$ is the CPHASE gate, up to a trivial global phase due to the natural time evolution of the $|1\rangle$ state. It corresponds to Eq. (3.40) in Liouville space, over all states in the basis, and for each ensemble. The set of $\hat{\rho}_k$ are the canonical basis elements of the two-qubit Liouville space, $\{|i\rangle\langle j|\}\ \forall i,j \in \{00, 01, 10, 11\}$. The state $\hat{\rho}_{k,n}(T)$ is the state $\hat{\rho}_k(0)$ propagated under the n'th ensemble Hamiltonian $\hat{\mathsf{H}}_n$. In order for a robust gate to be successfully implemented, the overlap must become maximal for all of the N ensemble members.

The second term corresponds to Eq. (3.61), but for the four controls $\Omega_j(t)$: the fields of the red and blue lasers for the left and right atom, respectively. $S(t)$ is a shape function for $\Delta\Omega_j$ that maintains smooth switch-on and switch-off of the pulses. The gate duration T is fixed for the optimization, but can be systematically varied in order to determine the quantum speed limit. As shown in chapter 7, for numerical efficiency, the full basis of 16 states can be replaced by just two density matrices specifically tailored to the optimization problem, exploiting the fact that we optimize for a diagonal unitarity and not a general dynamical map [196]. The time dependent states $\hat{\rho}_{k,n}(t)$ are determined by the equation of motion,

$$\frac{\partial}{\partial t} \hat{\rho}_{k,n}(t) = -\frac{\mathrm{i}}{\hbar} [\hat{\mathsf{H}}_n(t), \hat{\rho}_{k,n}(t)] + \mathcal{L}_D(\hat{\rho}_{k,n}(t)) \tag{4.14}$$

with $\hat{\rho}_{k,n}(t=0) = \hat{\rho}_k(0)$, and $\mathcal{L}_D$ according to Eq. (4.4).

For our choice of functional, the linear version of Krotov's method [138, 155] is sufficient. The update equation for each control is given by Eq. (3.62), but written out in Liouville space. The inner product is new given by the

Hilbert-Schmidt overlap. Since the dissipator does not depend on the control,

$$\frac{\partial \mathcal{L}}{\partial \Omega_j} \hat{\rho}_{k,n}^{(i+1)}(t) = \left[\frac{\partial \hat{\mathsf{H}}_n}{\partial \Omega_j}, \hat{\rho}_{k,n}^{(i+1)}(t) \right] . \tag{4.15}$$

Thus, the update equation for each pulse reads

$$\Delta \Omega_j(t) = \frac{S(t)}{\lambda_j} \sum_{n=0}^{N-1} \sum_{k=1}^{16} \mathfrak{Im} \left\{ \operatorname{tr} \left(-\mathrm{i}\, \hat{\sigma}_{k,n}^{(i)}(t) \left[\frac{\partial \hat{\mathsf{H}}_n}{\partial \Omega_j}, \hat{\rho}_{k,n}^{(i+1)}(t) \right] \right) \right\} , \tag{4.16}$$

with the $\hat{\sigma}_{k,n}^{(i)}(t)$ being a set of co-states backwards propagated with the pulse from the previous iteration, cf. (3.55),

$$\frac{\partial \hat{\sigma}_{k,n}(t)}{\partial t} = -\frac{\mathrm{i}}{\hbar} [\hat{\mathsf{H}}_n(t), \hat{\sigma}_{k,n}(t)] - \mathcal{L}_D(\hat{\sigma}_{i,n}(t)) . \tag{4.17}$$

Note the change of sign for the dissipator compared to Eq. (4.14). This reflects the propagation under the adjoint equation of motion, expressed as $\hat{\mathsf{H}}^\dagger$ in Eq. (3.55). The 'initial' condition for the propagation of $\hat{\sigma}$ is determined by the final time objective, cf. Eq.(3.56), resulting in

$$\hat{\sigma}_{k,n}(t = T) = \hat{\mathsf{O}} \hat{\rho}_k(0) \hat{\mathsf{O}}^\dagger . \tag{4.18}$$

This corresponds to Eq. (3.66), with the normalization prefactor absorbed in λ_j. The states $\hat{\rho}_{k,n}^{(i+1)}(t)$ in Eq. (4.16) are forward propagated using the pulse of the current iteration, according to Eq. (4.14). In the case of the rotating wave approximation where the $\Omega_j(t)$ are complex, Eq. (4.16) is valid for both the real and the imaginary part of the pulse.

4.4.2 Optimized Rydberg Gates

In order to optimize for robustness with respect to both amplitude fluctuations and fluctuations of the Rydberg level, we choose an ensemble of $N = 24$ Hamiltonians, evenly sampling the values of Δ_{ryd} between $\pm 300\,\text{kHz}$ and variations of the dipole coupling strength between $\pm 5\%$. The resulting pulses and their spectra are shown in Fig. 4.10, in the two-color rotating frame. In the spectra, the central frequency of zero corresponds to a laser frequency of the blue pulse that is detuned by Δ_1 with respect to the $|0\rangle \rightarrow |i\rangle$ transition. For the red pulse, it indicates the frequency for which there is a two-photon resonance with the $|0\rangle \rightarrow |r\rangle$ transition. The guess pulses from which the optimization started are indicated in orange; they are inspired by the analytic scheme of the previous section, consisting of two π pulses on the left atom and simultaneously one 2π pulse on the right atom. The gate duration was set to $T = 800\,\text{ns}$, matching the shortest gate duration obtained for the analytic schemes in the previous section. The choice of the guess pulse is arbitrary in principle, but has significant impact on the convergence speed and the characteristics of the optimized pulse. Indeed, the optimized pulse shapes still roughly follow the shapes of the guess pulses. However, especially for the left atom, there are fast oscillations present in the optimized pulse shapes. These correspond to a second laser frequency. As can be seen from the spectra shown in Fig. 4.10 (c), this second frequency is at $+\Delta_1$ for the blue pulse and at $-\Delta_1$ for the red pulse. This allows a pathway interference to be set up in the energy level space, as follows. Pulses at these additional frequencies are now two-photon resonant with the $|0\rangle \rightarrow |r\rangle$ transition and thereby introduce a second excitation pathway whose interference with the primary pathway can be exploited as a control mechanism. The optimization procedure takes advantage of this mechanism to enhance the overall gate fidelity by utilizing the interference between excitation amplitude contributions from these distinct paths connecting the atomic ground and Rydberg state. We note that the blue side peak is smaller as a result of the smaller amplitude of the corresponding laser. Indeed, in the spectra of the pulses acting on the right atom, cf. Fig. 4.10 (d), the second frequency is mostly absent, except for the very beginning and end of the red pulse.

The population induced by the optimized pulses with the ideal Hamiltonian $\hat{\mathsf{H}}_0$ is shown in Fig. 4.11. Even though the optimized pulses have frequency components that are resonant with the $|0\rangle \rightarrow |i\rangle$ transition, the intermediate level is now never significantly populated, due to destructive interference along the two pathways in energy space. Suppression of the intermediate state population may be aided by the STIRAP-like feature of the optimized pulse shape, in Fig. 4.10 (a) and (b), where the red laser (counter-intuitively) precedes the blue laser in the initial depopulation of the $|0\rangle$ level of the left atom, and follows it in the final repopulation. Furthermore, the population of the $|01\rangle$ state stays remarkably constant, despite the rather large amplitudes of the laser fields in Fig. 4.10 (a). Again, this is due to the interfering multiple pathways. In contrast, the dynamics of the $|10\rangle$ state is much more straightforward, on account of the absence of the second laser frequency. The pulse consists effectively of a single 2π pulse, although not with full population transfer. The Rydberg blockade is almost fully maintained, cf. the lack of population in the $|rr\rangle$ state in the bottom panel of Fig. 4.11. Also, the right atom in the time evolution of the $|00\rangle$ state is almost unaffected by the pulse on the right atom, resulting in very similar population dynamics for the $|00\rangle$ and $|01\rangle$ states.

Optimal control also holds the promise of finding pulses approaching the quantum speed limit. With numerical optimization, we can find solutions with gate durations far below $T = 800\,\mathrm{ns}$ required for the analytic schemes,

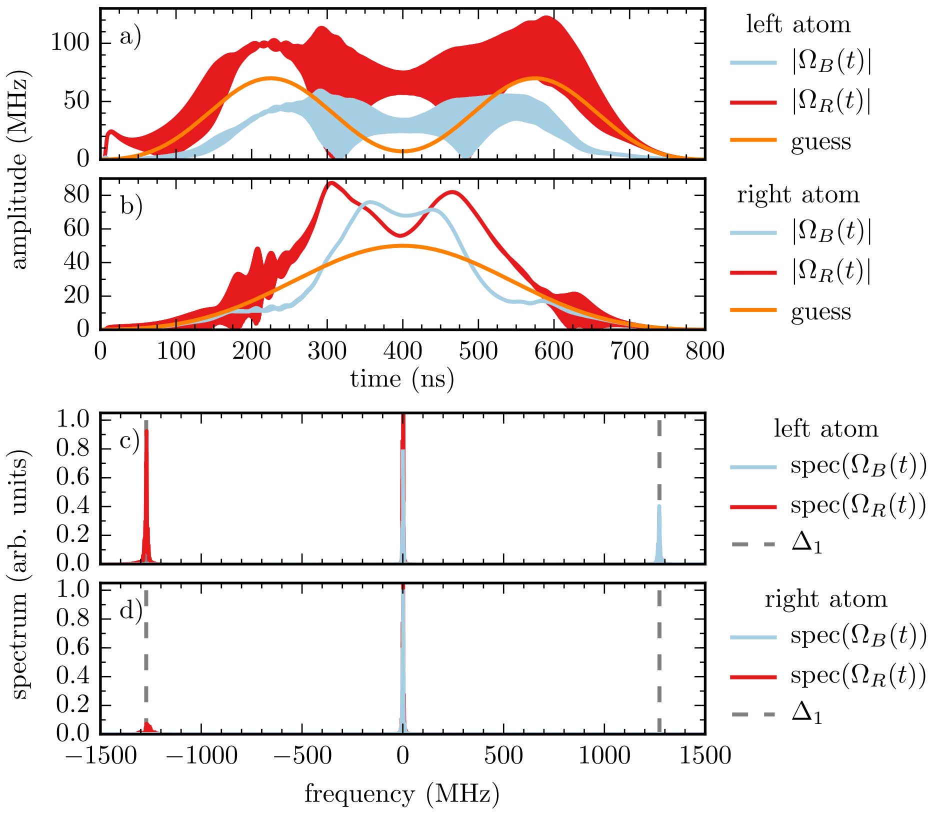


Figure 4.10: Amplitudes and spectra of pulses optimized with respect to variations in both two-photon detuning and pulse amplitude, for a gate duration of $T = 800\,$ns. The central peaks in the spectra are truncated to emphasize the side-peaks, reaching 2.0 (red) and 0.8 (blue) in panel (c), and 1.4 (red) and 1.0 (blue) in panel (d). The frequencies matching $\pm\Delta_1$ are indicated by vertical dashed gray lines.

although very short pulses may require unfeasibly large pulse amplitudes. The pulses and spectra resulting from an optimization for $T = 100\,$ns are shown in Fig. 4.12. The pulses are optimized for robustness, using the same ensemble of Hamiltonians as for the $T = 800\,$ns pulses. The pulse shapes again follow the features of the guess pulse, and are only slightly more complex than those for 800 ns in Fig. 4.10, and should be well within reach of established nanosecond pulse shaping techniques [197].

The spectra in Fig. 4.12 (c) and (d) reveal that a similar pathway interference mechanism as seen in the optimized $T = 800\,$ns pulse sequences operates, realized by the additional frequencies at $\pm\Delta_1$. The most significant difference from Fig. 4.10 is that now the additional frequencies are present for both the left and the right atom throughout the entire gate duration. As a result of the shorter time window, the peaks in the spectrum are broadened and the pulse amplitudes are now significantly higher. Generally,

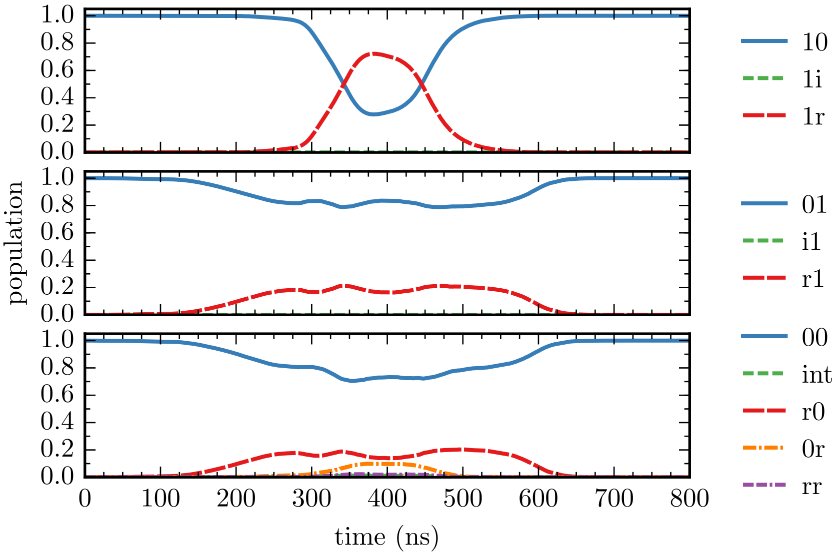


Figure 4.11: Dynamics under the pulses optimized with respect to fluctuations in both the Rydberg level and pulse amplitudes, as shown in Fig. 4.10. The intermediate population in the bottom panel ("int") is integrated over the states $|0i\rangle$, $|i0\rangle$, $|ii\rangle$, $|ir\rangle$, and $|ri\rangle$. The shown dynamics implement the desired CPHASE gate up to a gate error of 1.04×10^{-4}.

the optimization becomes harder for shorter pulse durations and the available control mechanism must now be used more efficiently: This rationalizes the presence of the second laser frequency throughout all pulses.

The population dynamics in Fig. 4.13 show some significant differences from the dynamics shown in Fig. 4.11, as a result of the increase in laser amplitude. Most importantly, the Rydberg blockade is now broken, resulting in a significant population of the $|rr\rangle$ state, cf. the purple curve in the bottom panel. This nicely illustrates the power of OCT; while the analytic schemes rely on maintaining the blockade regime, the optimization has no such restrictions, and will exploit any pathways available in the time evolution generated by the two-qubit Hamiltonian. There is some minor population in the intermediate states during the propagation of the $|00\rangle$ state, cf. the blue line in the bottom panel of Fig. 4.13. However, since the dynamics result from an optimization that explicitly took into account the spontaneous decay from the intermediate level, we are guaranteed that the population in this level is below a threshold value that would affect the gate fidelity.

In Fig. 4.14, we compare the effect of fluctuations due to electric fields and pulse amplitude fluctuations on the gate fidelity for the pulses obtained with OCT, cf. Figs. 4.10 and 4.12, to that for the most robust gates achieved with the analytic schemes, i.e., the mixed scheme employing STIRAP for the pulses on the left atom, and simultaneous pulses for the right atom, cf.

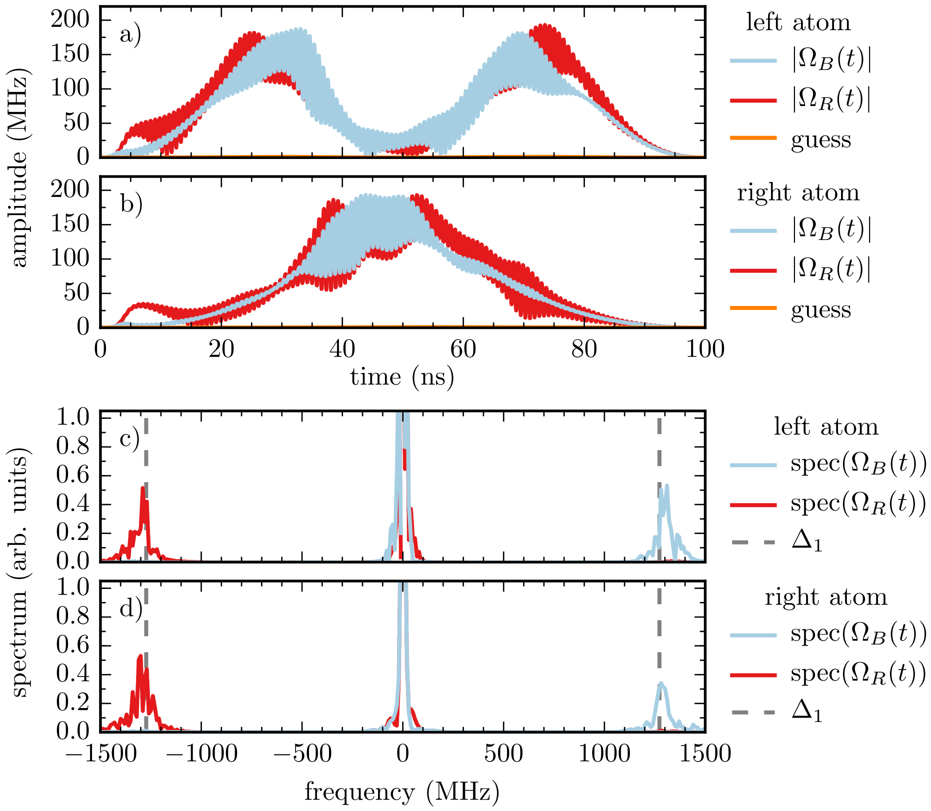


Figure 4.12: Amplitudes and spectra of pulses optimized with respect to variations in both two-photon detuning and pulse amplitude, for a gate duration of $T = 100\,\text{ns}$. The spectra are drawn on the same scale as in Fig. 4.10, with the central peaks in panel (c) reaching 4.5 (blue) and 3.0 (red), and 4.5 for both pulses in panel (d).

Fig. 4.9. The optimized pulses are significantly more robust with respect to both sources of error by at least an order of magnitude, with the gate fidelity staying above 99.9% even for large variations, whereas for the analytic pulses, it drops below 97% for fluctuations of the Rydberg level (top panel) and 95% for amplitude fluctuations (bottom panel). Note that in contrast to the analytic mixed scheme, the optimized pulses do not require unfeasibly large pulse amplitudes. In contrast, the scheme using only simultaneous pulses but more realistic pulse amplitudes would be even more sensitive – particularly to fluctuations of the Rydberg level (cf. the drop to 92% gate fidelity in the top panel of Fig. 4.9). The price for this additional robustness offered by the numerically optimized pulses is a slightly more complex pulse shape and the presence of a second frequency.

It is important to note that the solutions provided by OCT are not unique; the pulses obtained depend on the guess pulses, the exact choice of optimization functional, and on arbitrary scaling parameters such as the

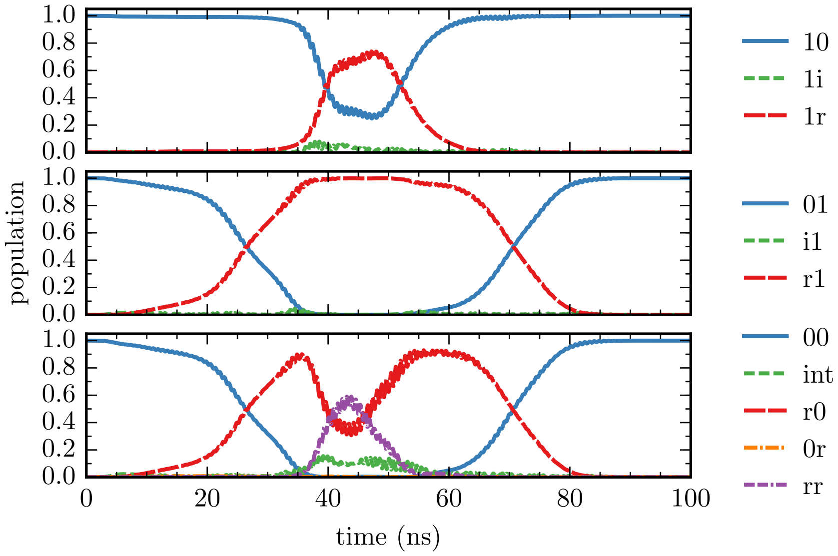


Figure 4.13: Dynamics under the optimized pulses shown in Fig. 4.12. The gate error is 1.92×10^{-4}.

λ_j in Eq. (4.12). By tuning these parameters carefully, the optimization may be steered towards desired pulse features. It is also possible to add additional constraints to the optimization functional in order to preselect optimization pathways [158]. For example, the $|rr\rangle$ state could be defined as a forbidden subspace in order to enforce the blockade regime. This could be desirable, since the breaking of the Rydberg blockage can excite the atoms vibrationally, resulting in a possible increase of gate duration due to the additional time required to restore the original state of motion. The vibrational motion was not taken into account in the our model. It could be included explicitly in the optimization to determine the speed limit outside of the blockage regime [73]. One could also include spectral constraints to impose a pre-specified pulse bandwidth or suppress undesired frequency components [158, 159]. Optimizing to extremely high fidelities often leads to very large pulse amplitudes or complex pulse features that are undesirable from an experimental point of view. Thus, it is usually best to stop the optimization as soon as the reached fidelities are “good enough”, as was done for the optimized pulses shown as solid blue and yellow lines in Fig. 4.14. In principle, however, pulses of much higher fidelity and robustness than those shown here can be found. This is illustrated by the dotted green line in Fig. 4.14, which shows the result of a further optimization of the pulse for $T = 800\,$ns. While these pulses achieve a gate fidelity well above that required for fault tolerant quantum computation [198, 199, 186], the resulting highly optimized pulses have unfeasibly large pulse amplitudes of 1100 MHz and 330 MHz for the blue and red laser, respectively.

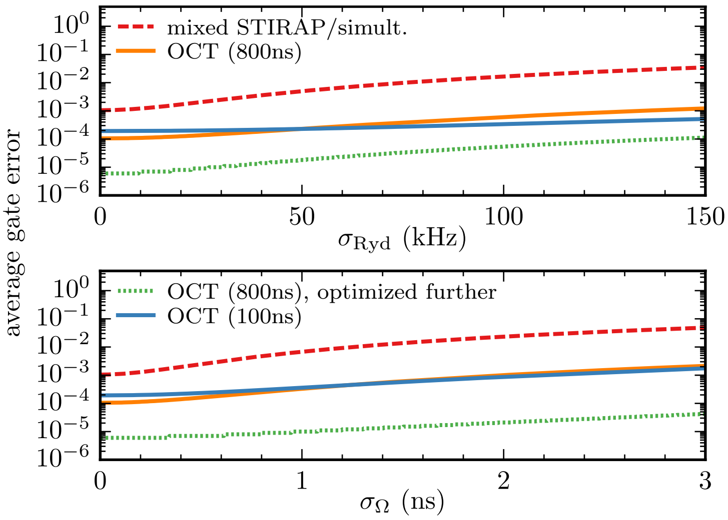


Figure 4.14: Expectation value of the gate error in the presence of fluctuations in the $|rr\rangle$ state due to DC electric fields (top), and pulse amplitude fluctuations (bottom). The red dashed curve shows the most robust analytical pulse, cf. the red dashed curve in Fig. 4.9. The solid yellow and blue lines are for the optimized pulses shown in Figs. 4.10, 4.12, respectively. The dotted green line is for a further optimized pulse at $T = 800\,\text{ns}$, without any consideration of limits on the pulse amplitude or complexity. Note that both panels show the robustness for same set of pulses, i.e. the pulses were optimized with respect to *both* variations in the two-photon detuning and the pulse amplitude.

The results presented here illustrate the power of optimal control theory as a tool for finding pulses that are robust both with respect to decoherence and to experimental fluctuations. For realistic noise levels, we were able to generate pulses that yield gate errors well below 10^{-3}, with errors below 10^{-5} being reached when no limits are placed on pulse amplitudes. Optimized pulse sequences are not only more robust but can also be of much shorter duration, significantly shorter than those obtained from the best analytical schemes. This is a considerable advantage, allowing to perform a larger number of gates before decoherence effects become significant. The technique of optimizing over an ensemble of perturbed Hamiltonians in order to achieve robust pulses is not limited to the example of Rydberg gates, but is generally applicable to other systems and optimizations problems.

5

Quantum Gates with Superconducting Transmon Qubits

Superconducting circuits have emerged as one of the most promising environments for quantum information processing. Unlike most other implementations (e.g. trapped atoms like the Rydberg gate discussed in chapter 4), they are macroscopic, in the sense that the relevant degrees of freedom are the quantities of electrical circuits, such as charge, current, or flux. However, unlike in classical electrical circuits, these variables are now fully quantized and described by a wave function. The quantum behavior results from the electrons forming Cooper pairs and condensing to a *collective* quantum state as the system is cooled below some (material-dependent) critical temperature T_c [200].

From a technical perspective, superconducting qubits have the significant advantage of building upon fabrication techniques used in standard integrated circuits. A superconducting layer of usually aluminum or niobium is layered on a silicon wafer [201, 202]. The circuit elements are then patterned using the standard tools of optical or electron-beam lithography and chemical etching: a photo-sensitive "resin" material is added to the chip and then illuminated with UV light or an electron beam through a mask. Then, the developed resin is removed and the exposed superconducting material is etched away, or alternatively, electrical leads are deposited in the resin gaps [203]. For superconducting circuits relevant to quantum computing, one of the central components is the *Josephson junction*, described in section 5.1. It consists of two layers of superconducting material separated by an insulating layer. This insulating layer is generally created by oxidation of the superconducting material [201, 202]. Using these well-established production techniques, a quantum computer based on a superconducting architecture could readily be produced industrially, provided that the more fundamental challenges in building a large-scale quantum processor can be met, i.e., implementing universal quantum gates and scaling the circuit to a significant number of qubits while maintaining coherence.

This chapter reviews the fundamentals of superconducting qubits, with emphasis on the *transmon* design. Two transmons can be coupled via a shared transmission line, allowing for the exchange of virtual photons. Using perturbation theory, an effective model can be derived in which the qubit-cavity coupling has been eliminated. This motivates the two-transmon Hamiltonian used in chapters 6 and 7. We discuss some of the gate mechanisms that have been used to implement two-qubit gates using transmons and explore possibilities for optimal control.

5.1 The Josephson Junction

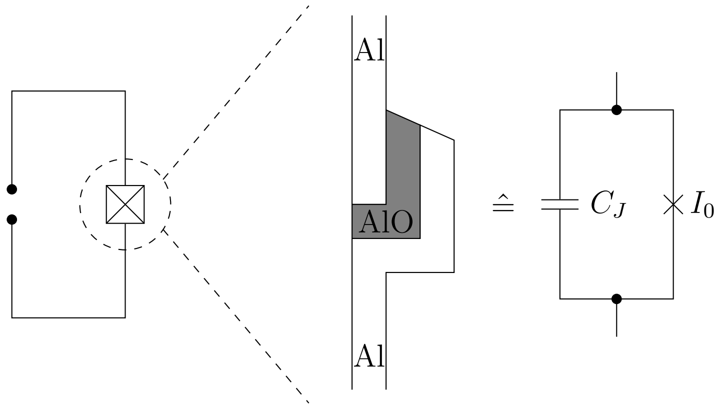


Figure 5.1: Josephson junction. The junction consists of two superconducting layers of e.g. aluminum, separated by an insulating layer, e.g. aluminum oxide. The junction combines the properties of a capacitor, characterized by C_J or the charging energy $E_C \propto \frac{1}{C_J}$, and an inductive tunneling current, characterized by I_0 or the Josephson energy $E_J \propto I_0$.

The Josephson junction consists of two superconducting leads separated by a thin insulating layer, as shown in Fig. 5.1. This structure acts as a capacitor with capacitance C_j. As known from classical electrodynamics, if one of the leads (capacitor plates) holds a charge of Q, the energy stored in the capacitor is

$$E_{\text{capacitor}} = \frac{1}{2}\frac{Q^2}{C_J}\,. \tag{5.1}$$

junction parameters	
I_0	critical current
C_j	junction capacitance
Q_r	offset charge
E_c	charging energy, $E_c = \frac{1}{2}\frac{(2e)^2}{C_J}$
E_J	Josephson energy, $E_J = \phi_0 I_0$
L_{J0}	effective inductance, $L_{J0} = \frac{\phi_0}{I_0}$
dynamic quantities	
$\Phi(t)$	branch flux of junction, see Eq. (5.3); $\phi(t) \in \mathbb{R}$
$I(t)$	Josephson current, see Eq. (5.2); $I(t) \in \mathbb{R}$
$\delta(t)$	"phase", $\delta(t) = \frac{2\pi\Phi(t)}{\Phi_0} = \frac{\Phi(t)}{\phi_0}$; $\delta \in \mathbb{R}$
$\theta(t)$	condensate phase diff., $\theta(t) = \delta(t) \bmod 2\pi$; $\theta(t) \in [0, 2\pi)$
$L_J(\delta)$	Josephson inductance, $L_J(\delta) = L_{J0}\cos^{-1}\delta$
$n(t)$	number of tunneled cooper pairs; $n(t) \in \mathbb{N}$
fundamental constants	
Φ_0	flux quantum, $\Phi_0 = \frac{h}{2e}$
ϕ_0	reduced flux quantum, $\phi_0 = \frac{\Phi_0}{2\pi}$

Table 5.1: Summary of quantities and constants used for the characterization of a Josephson junction. The first group contains static parameters that depend on the geometric properties of the junction at production. An exception is E_J which can also be made tunable by splitting the junction in two and running an external flux through the resulting loop. The second group are the dynamic quantities from which the tunneling energy and the capacitive energy of the junction derive, and which allow to formulate the Hamiltonian. The "phase" $\delta(t)$ is more properly known as the "gauge-invariant phase difference" across the junction.

In addition to the capacitive effect, Cooper pairs of electrons can tunnel through the insulating layer (owing to the quantum mechanical nature of the circuit), resulting in a *Josephson current* that is described by [204]

$$I(t) = I_0 \sin \delta(t) = I_0 \sin \frac{2\pi\Phi(t)}{\Phi_0} = I_0 \sin \frac{\Phi(t)}{\phi_0}\,, \tag{5.2}$$

where I_0 is the maximum current the junction can support, and $\delta(t)$ is the phase difference of the wave function on each side of the junction, which is directly related to the branch flux [205]

$$\Phi(t) = \int_{-\infty}^{t} V(t')\,\mathrm{d}t' \tag{5.3}$$

of the junction element, where $V(t)$ is the voltage across the junction. All the quantities typically used to describe the junction are listed in Table 5.1.

If the junction holds an initial charge difference ("offset charge") of Q_r and n Cooper pairs of electrons (charge $2e$) tunnel through the junction, the capacitive energy of the junction according to Eq. (5.1) is

$$E_{jj}^{\text{capacitive}} = E_c\left(n - \frac{Q_r}{2e}\right)^2, \qquad E_c = \frac{1}{2}\frac{(2e)^2}{C_J}\,. \tag{5.4}$$

The offset charge $Q_r/2e$ will generally be much larger than n, and cannot easily be controlled during the production process. The energy due to the tunneling current is

$$E_{jj}^{\text{inductive}} = \int\limits_{-\infty}^{t} I(t')V(t')\,\mathrm{d}t' = \int\limits_{0}^{\Phi(t)} I_0 \sin\left(\frac{\Phi}{\phi_0}\right)\mathrm{d}\Phi = -E_J \cos\delta\,, \tag{5.5}$$

with $E_J = I_0\phi_0$, $\delta = \frac{\Phi}{\phi_0}$, cf. Table 5.1, and $V(t) = \frac{\mathrm{d}\Phi(t)}{\mathrm{d}t}$, cf. Eq. (5.3). Together, both terms yield the Hamiltonian

$$\hat{\mathsf{H}}_{jj} = E_c\left(\hat{\mathsf{n}} - \frac{Q_r}{2e}\right)^2 - E_J \cos\hat{\theta}\,. \tag{5.6}$$

where $\hat{\mathsf{n}}$ and $\hat{\theta}$ are quantum variables obeying the canonical commutator relationship $[\hat{\theta}, \hat{\mathsf{n}}] = \mathrm{i}\hbar$.

5.2 The Cooper Pair Box

The Hamiltonian in Eq. (5.6) is dominated by the value of Q_r, which takes a random value during the production process and is subject to significant noise fluctuations [206]. This makes the naked Josephson junction ill-suited

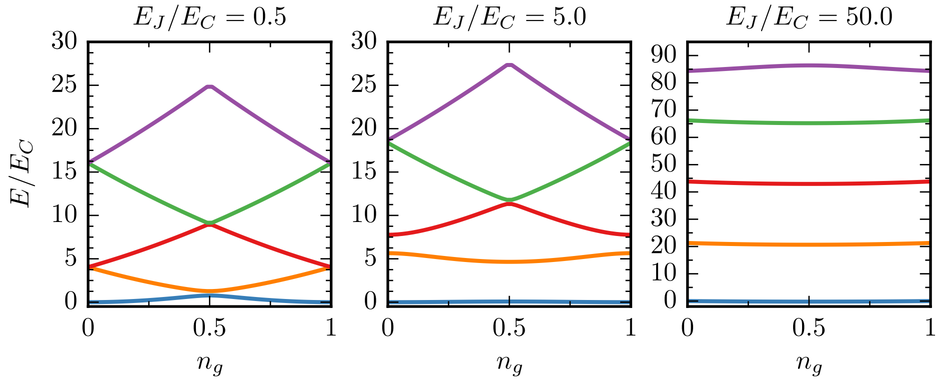


Figure 5.2: Lowest five eigenvalues of the charge qubit Hamiltonian, obtained from diagonalization of Eq. (5.9), depending on the offset charge n_g, for different ratios of the qubit's Josephson and charging energies, cf. [212].

as a qubit. To obtain a well-defined system, there are three "traditional" superconducting qubit designs that add further elements to the Josephson junction circuit [207]: the Cooper pair box ("charge qubit") the flux qubit, and the phase qubit.

The most straightforward of these is the charge qubit, obtained by biasing the Josephson junction with a voltage source and placing it in series with a capacitance C_g [208, 209]. The form of the resulting Hamiltonian is identical to the naked Josephson junction in Eq. (5.6), but the offset charge, respectively the charging energy is now a well-defined quantity, determined by C_g and the bias voltage U:

$$\hat{\mathsf{H}}_{CPB} = 4E_C\,(\hat{\mathsf{n}} - n_g)^2 - E_J \cos\hat{\theta}\,, \tag{5.7}$$

with

$$4E_C = \frac{(2e)^2}{2\,(C_J + C_g)}\,, \qquad n_g = \frac{Q_r}{2e} + \frac{C_g U}{2e}\,. \tag{5.8}$$

E_J can be made tunable by splitting the Josephson junction and running a magnetic flux through the resulting loop [210]. Eq. (5.7) can be written entirely in the charge basis [211],

$$\hat{\mathsf{H}}_{CPB} = 4E_C\,(\hat{\mathsf{n}} - n_g)^2 - \frac{E_J}{2}\sum_n \left(|n\rangle\langle n+1| + |n+1\rangle\langle n|\right)\,. \tag{5.9}$$

The qubit is encoded in the lowest two eigenstates of $\hat{\mathsf{H}}_{CPB}$. The eigenvalues depend parametrically on the offset charge n_g; this dependence is named the *charge dispersion*. The separation of qubit levels as well as the charge dispersion depend on the ratio E_J/E_C, as shown in Fig. 5.2. Fluctuations in Q_r (and thus n_g) lead to dephasing of the qubit. Therefore, the qubit is usually operated at the *charge degeneracy point* $n_g = 0.5$, where the charge

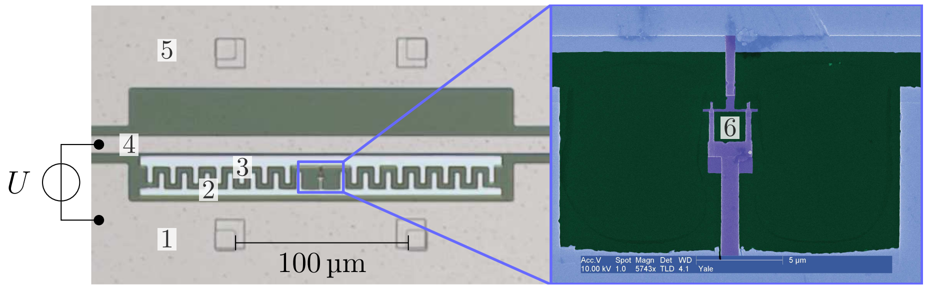


Figure 5.3: Optical image of transmon inside of a resonator, adapted from [211]. The circuit consists of five superconducting islands: the lower ground plane (1), the lower transmon island (2), the upper transmon island (3), the coplanar wave guide center pin (4), and the upper ground plane (5). The zoom shows the split Josephson junction (6). The islands (4) and (1) are connected with a bias voltage.

dispersion is zero to first order. At $n_g = 0.5$, small values of E_J/E_C result in the lowest two levels being separated by approximately E_J from each other, and by approximately $8E_C$ from the next higher level, making the charge qubit a two-level system in very good approximation. Higher values of E_J/E_C lead to less sensitivity with respect to charge noise, as the curves flatten out, but also decrease the anharmonicity of the levels. Both in order to maintain a well-defined two-level system, and for technical reasons related to the read-out of the qubit [213], $E_J/E_C < 1$ in the traditional charge qubit.

5.3 The Transmon Qubit

5.3.1 A Charge-Insensitive Cooper Pair Box

The transmon qubit [212] extends the charge qubit by shunting it with an additional capacitance C_B and placing it in a coplanar wave guide resonator. We refer to this resonator also as the transmission line, or simply the cavity. All components are shown in Fig. 5.3. The capacitance C_B is between the islands (2) and (3). The capacitance C_g is between island (3) and (4). A full circuit analysis [212, 211, 214] shows that the Hamiltonian takes the same form as that of the charge qubit, Eq. (5.7). However, C_B increases the denominator of the charging energy, Eq. (5.8), lowering E_C and thus increasing the ratio E_J/E_C to values of several tens or hundreds. The design of the transmon exploits that with increasing E_J/E_C, the charge dispersion decreases exponentially, while the anharmonicity only decreases algebraically [213]. Thus, one obtains a weakly anharmonic ladder that is robust against charge noise, cf. the right panel of Fig. 5.2. Choosing $E_J \gg E_C$ limits θ in Eq. (5.7) to small values, allowing to neglect n_g and to expand the cosine potential in a Taylor series as $\cos\theta \approx 1 - \frac{\theta^2}{2} + \frac{\theta^4}{24}$. Rewriting the charge and phase states in terms of ladder operators $\hat{\mathsf{b}}$, $\hat{\mathsf{b}}^\dagger$,

$$\hat{\mathsf{n}} = \frac{\mathrm{i}}{2}\left(\frac{E_J}{2E_C}\right)^{\frac{1}{4}}\left(\hat{\mathsf{b}}^\dagger - \hat{\mathsf{b}}\right) \qquad \hat{\theta} = \left(\frac{2E_C}{E_J}\right)^{\frac{1}{4}}\left(\hat{\mathsf{b}}^\dagger + \hat{\mathsf{b}}\right) \tag{5.10}$$

results in the *Duffing oscillator* (up to a constant)

$$\hat{\mathsf{H}}_T = \sqrt{8E_C E_J}\,\hat{\mathsf{b}}^\dagger\hat{\mathsf{b}} - \frac{E_C}{12}\left(\hat{\mathsf{b}}^\dagger + \hat{\mathsf{b}}\right)^4 . \tag{5.11}$$

Leading order perturbation theory brings this to [212, 215]

$$\hat{\mathsf{H}}_T = \omega_q \hat{\mathsf{b}}^\dagger\hat{\mathsf{b}} + \frac{\alpha}{2}\hat{\mathsf{b}}^\dagger\hat{\mathsf{b}}^\dagger\hat{\mathsf{b}}\hat{\mathsf{b}}\,, \tag{5.12}$$

with the qubit frequency $\omega_q \approx \sqrt{8E_J E_C} - E_C$ and the anharmonicity $\alpha = -E_C$.

Besides contributing to the charging energy, the waveguide cavity is also central for control and readout of the transmon. The interaction between the qubit with the ladder operator $\hat{\mathsf{b}}$, $\hat{\mathsf{b}}^\dagger$ and the cavity with ladder operators $\hat{\mathsf{a}}$, $\hat{\mathsf{a}}^\dagger$ is described by [216]

$$H_I = g\left(\hat{\mathsf{b}}^\dagger + \hat{\mathsf{b}}\right)\left(\hat{\mathsf{a}}^\dagger + \hat{\mathsf{a}}\right) , \tag{5.13}$$

where g is the interaction strength. The transmon system is operated in the *strong coupling regime* where g is much larger than the qubit and cavity decay rates γ and κ. This allows excitations to be exchanged between the qubit and the cavity. If $g \ll \omega_q, \omega_c$, a rotating wave approximation (RWA) yields a Jaynes-Cummings Hamiltonian, in which excitations are preserved [217],

$$H_I = g\left(\hat{\mathsf{b}}^\dagger\hat{\mathsf{a}} + \hat{\mathsf{b}}\hat{\mathsf{a}}^\dagger\right) , \tag{5.14}$$

with

$$\gamma, \kappa \ll g \ll \omega_q, \omega_c\,. \tag{5.15}$$

The total Hamiltonian of the transmon and the transmission line in the RWA is

$$\hat{\mathsf{H}}_T = \omega_q \hat{\mathsf{b}}^\dagger\hat{\mathsf{b}} + \frac{\alpha}{2}\hat{\mathsf{b}}^\dagger\hat{\mathsf{b}}^\dagger\hat{\mathsf{b}}\hat{\mathsf{b}} + g(\hat{\mathsf{b}}^\dagger\hat{\mathsf{a}} + \hat{\mathsf{b}}\hat{\mathsf{a}}^\dagger) + \omega_c\hat{\mathsf{a}}^\dagger\hat{\mathsf{a}}\,. \tag{5.16}$$

It is also possible to consider the transmon in the *ultra-strong coupling regime*, where the rotating wave approximation breaks down [218]. The eigenstates of Eq. (5.16) mix the bare eigenstates of the qubit and the cavity; the degree of mixing is determined by the ratio of the detuning $\omega_c - \omega_q$ and the coupling strength g. In the *resonant regime*, $\omega_c = \omega_q$, the eigenstates of $\hat{\mathsf{H}}_T$ are equal superpositions of qubit and cavity eigenstates. This regime is used for measurement and characterization of the transmon device [219].

For the implementation of quantum gates, the qubit is usually tuned to the *dispersive regime*,

$$|\omega_c - \omega_q| \gg g \,. \tag{5.17}$$

The eigenstates are now very close to the bare qubit and cavity eigenstates. The slight mixing shifts the energy of qubit and the cavity by a small amount χ, see section 5.4.

The parameters of the qubit and the coupling to the cavity can be engineered with great flexibility. Qubit frequencies can be readily set anywhere from 2 to 15 GHz, and couplings can range from a few kilohertz to nearly 1 GHz [220]. Moreover, when implemented with a split Josephson junction as depicted in Fig. 5.3, the qubit frequencies can be dynamically tuned by applying a magnetic flux through the junction loop.

5.3.2 Coupled Transmon Qubits

Connecting two transmon qubits to the same transmission line [221] creates an effective coupling between them. Each qubit is characterized by its own frequency, anharmonicity, and coupling strength to the cavity, ω_1, α_1, g_1, respectively ω_2, ω_2, α_2 in Hamiltonian (5.16). Together, the Hamiltonian for the two qubits and the shared transmission line with frequency ω_c reads

$$\begin{aligned}\hat{\mathsf{H}}_{2T} = \sum_{q=1,2} &\left[\omega_q \hat{\mathsf{b}}_q^\dagger \hat{\mathsf{b}}_q + \frac{\alpha_q}{2} \hat{\mathsf{b}}_q^\dagger \hat{\mathsf{b}}_q^\dagger \hat{\mathsf{b}}_q \hat{\mathsf{b}}_q + g_q (\hat{\mathsf{b}}_q^\dagger \hat{\mathsf{a}} + \hat{\mathsf{b}}_q \hat{\mathsf{a}}^\dagger)\right] + \\ &+ \omega_c \hat{\mathsf{a}}^\dagger \hat{\mathsf{a}} + \epsilon^*(t) \hat{\mathsf{a}} + \epsilon(t) \hat{\mathsf{a}}^\dagger \,.\end{aligned} \tag{5.18}$$

We have included a driving term with the external field $\epsilon(t)$ on the transmission line. The field $\epsilon(t)$ is taken as complex to account for the possibility of a rotating wave approximation. As described in appendix B, this shifts ω_1, ω_2, ω_c by the frequency of the lab frame $\epsilon(t)$ and allows to describe the field $\epsilon(t)$ by its shape only.

The interaction with the cavity allows for the exchange of *virtual photons* [221], where an excitation moves from qubit (1) to qubit (2) without populating the cavity. Driving $\epsilon(t)$ at the qubit frequency results in a cavity-mediated excitation of the qubit. Thus, the shared transmission line resonator allows for extensive control over the two-transmon system.

In addition to the coupling via the transmission line, the qubits could also be coupled directly, either capacitively or inductively [214], adding a static interaction term $J(\hat{\mathsf{b}}_1^\dagger \hat{\mathsf{b}}_2 + \hat{\mathsf{b}}_1 \hat{\mathsf{b}}_2^\dagger)$ to Eq. (5.18), where J is the strength of the interaction. In this case, the interaction must be activated by dynamically tuning the qubit frequencies into resonance, see section 5.5.

5.4 Effective Description of the Transmon in the Dispersive Regime

Understanding how the cavity mediates an effective interaction between the qubits is possible by partially diagonalizing the Hamiltonian (5.18) to remove the qubit-cavity coupling. First, Eq. (5.18) is rewritten in the energy basis,

$$\hat{\mathsf{b}} = \sum_i \sqrt{i}\,|i-1\rangle\langle i| \,, \tag{5.19}$$

to read

$$\begin{aligned}\hat{\mathsf{H}}_{2T} = &\underbrace{\sum_{q=1,2}\sum_i \omega_i^{(q)}\hat{\Pi}_i^{(q)} + \omega_c\hat{\mathsf{a}}^\dagger\hat{\mathsf{a}}}_{\hat{\mathsf{H}}_0} + \underbrace{\sum_{q=1,2}\sum_i g_i^{(q)}\left(\hat{\mathsf{C}}_i^{+\,(q)}\hat{\mathsf{a}} + \hat{\mathsf{C}}_i^{-\,(q)}\hat{\mathsf{a}}^\dagger\right)}_{\hat{\mathsf{H}}_I} + \\ &+ \epsilon^*(t)\hat{\mathsf{a}} + \epsilon(t)\hat{\mathsf{a}}^\dagger \,,\end{aligned} \tag{5.20}$$

with the qubit energy levels

$$\omega_i^{(1,2)} = i\omega_{1,2} + \frac{\alpha_{1,2}}{2}(i^2 - i)\,, \tag{5.21}$$

the qubit-cavity coupling

$$g_i^{(1,2)} = \sqrt{i}\,g_{1,2}\,, \tag{5.22}$$

and the projectors and jump operators

$$\hat{\Pi}_i^{(1,2)} = |i\rangle\langle i|_{1,2}\,, \quad \hat{\mathsf{C}}_i^{+\,(1,2)} = |i\rangle\langle i-1|_{1,2}\,, \quad \hat{\mathsf{C}}_i^{-\,(1,2)} = |i-1\rangle\langle i|_{1,2}\,. \tag{5.23}$$

The unitary transformation that decouples qubits and cavity takes the form

$$\hat{\mathsf{H}}' = e^{\hat{\mathsf{S}}}\hat{\mathsf{H}}e^{-\hat{\mathsf{S}}}\,. \tag{5.24}$$

It is known as a Schrieffer-Wolff-transformation, with

$$\hat{\mathsf{S}} = -\sum_{q=1,2}\sum_i \frac{g_i^{(q)}}{\omega_i^{(q)} - \omega_{i-1}^{(1)} - \omega_c}\left(\hat{\mathsf{a}}^\dagger\hat{\mathsf{C}}_i^{-\,(1)} - \hat{\mathsf{a}}\hat{\mathsf{C}}_i^{+\,(1)}\right)\,. \tag{5.25}$$

$\hat{\mathsf{S}}$ is constructed such that $[\hat{\mathsf{S}}, \hat{\mathsf{H}}_0] = -\hat{\mathsf{H}}_I$, cf. Ref. [222]. Therefore, to first order, it removes the unwanted interaction term $\hat{\mathsf{H}}_I$. Eq. (5.24) is evaluated using the Hadamard-Lemma,

$$\hat{\mathsf{H}}' = e^{\hat{\mathsf{S}}}\hat{\mathsf{H}}e^{-\hat{\mathsf{S}}} = \hat{\mathsf{H}} + [\hat{\mathsf{S}}, \hat{\mathsf{H}}] + \frac{1}{2}[\hat{\mathsf{S}}, [\hat{\mathsf{S}}, \hat{\mathsf{H}}]] + \ldots \tag{5.26}$$

After some algebra [223, 216], one obtains the transformed Hamiltonian

$$
\begin{aligned}
\hat{\mathsf{H}}' &= \omega_c \hat{\mathsf{a}}\hat{\mathsf{a}}^\dagger + \epsilon^*(t)\hat{\mathsf{a}} + \epsilon(t)\hat{\mathsf{a}}^\dagger + \sum_{q=1,2}\sum_{i=0}^{N_q-1} (\chi_i^{(q)} - \chi_{i+1}^{(q)})\hat{\mathsf{a}}^\dagger\hat{\mathsf{a}}\hat{\Pi}_i^{(q)} + \\
&+ \sum_{q=1,2}\sum_{i=0}^{N_q-1}\left[(\omega_i^{(q)} + \chi_i^{(q)})\hat{\Pi}_i^{(q)} + g_i^{\mathrm{eff}\,(q)}\epsilon(t)(\hat{\mathsf{C}}_i^{+\,(q)} + \hat{\mathsf{C}}_i^{-\,(q)})\right] + \\
&+ \sum_{ij} J_{ij}^{\mathrm{eff}}(\hat{\mathsf{C}}_i^{-\,(1)}\hat{\mathsf{C}}_j^{+\,(2)} + \hat{\mathsf{C}}_i^{+\,(1)}\hat{\mathsf{C}}_j^{-\,(2)}),
\end{aligned}
\tag{5.27}
$$

truncated to N_q transmon levels, with the "effective" quantities

$$
g_i^{\mathrm{eff}\,(1,2)} = \frac{g_i^{(1,2)}}{(\omega_i^{(1,2)} - \omega_{i-1}^{(1,2)} - \omega_c)}\,, \tag{5.28a}
$$

$$
J_{ij}^{\mathrm{eff}} = \frac{1}{2} g_i^{\mathrm{eff}\,(1)} g_j^{(2)} + \frac{1}{2} g_j^{\mathrm{eff}\,(2)} g_i^{(1)}\,, \tag{5.28b}
$$

$$
\chi_i^{(1,2)} = \frac{(g_i^{(1,2)})^2}{(\omega_i^{(1,2)} - \omega_{i-1}^{(1,2)} - \omega_c)}\,. \tag{5.28c}
$$

All terms in which $g_i^{\mathrm{eff}\,(1,2)}$ appears in higher than quadratic order have been neglected. The qubit levels obtain a *Lamb-shift* χ_i. Moreover, the term proportional to $\hat{\mathsf{a}}^\dagger\hat{\mathsf{a}}\hat{\Pi}_i$ can be interpreted either as a qubit-dependent shift of the cavity levels, or a further *ac Stark shift* of the qubit levels depending on the number of photons in the cavity. The mediated interaction between the two qubits is given by J_{ij}^{eff}. Lastly, we see that $\epsilon(t)$ can drive the qubit transitions with a reduced transition strength g_i^{eff}. The Schrieffer-Wolff-transformation is a perturbative treatment into orders of g^{eff}, which must be small, corresponding to the dispersive condition (5.17). Thus, Eq. (5.24) is also known as the dispersive transformation, and is closely related to the method of *adiabatic elimination* [224].

Assuming that the frequency of $\epsilon(t)$ is far detuned from ω_c, such that the cavity is never populated, ($\langle\hat{\mathsf{a}}^\dagger\hat{\mathsf{a}}\rangle = 0$), it can be fully integrated out and one obtains an effective Hamiltonian for the reduced system,

$$
\begin{aligned}
\hat{\mathsf{H}}_{\mathrm{red}} &= \sum_{q=1,2}\sum_{i=0}^{N_q-1}\left[(\omega_i^{(q)} + \chi_i^{(q)})\hat{\Pi}_i^{(q)} + g_i^{\mathrm{eff}\,(q)}\epsilon(t)(\hat{\mathsf{C}}_i^{+\,(q)} + \hat{\mathsf{C}}_i^{-\,(q)})\right] + \\
&+ \sum_{ij} J_{ij}^{\mathrm{eff}}\left(\hat{\mathsf{C}}_i^{-\,(1)}\hat{\mathsf{C}}_j^{+\,(2)} + \hat{\mathsf{C}}_i^{+\,(1)}\hat{\mathsf{C}}_j^{-\,(2)}\right).
\end{aligned}
\tag{5.29}
$$

The derivation of effective models using Schrieffer-Wolff transformations, perturbation theory, and adiabatic elimination theory is an extensive field of study. Some of the methods are reviewed in Ref. [216].

Depending on the parameter regime, often a series of additional Schrieffer-Wolff transformations on Eq. (5.27) is employed to further reduce the degrees of freedom, e.g. to decouple well-separated subspaces [222]. The validity of the dispersive and any further transformation to a given order must be verified for each choice of parameters.

Approximately, Eq. (5.29) takes the form of an effective Jaynes-Cummings Hamiltonian on the qubits [222]

$$\begin{aligned} \hat{\mathsf{H}}_{\text{red}} \approx \sum_{q=1,2} \left(\left(\omega_q' - \frac{\alpha_q}{2} \right) \hat{\mathsf{b}}_q^\dagger \hat{\mathsf{b}}_q + \frac{\alpha_q}{2} \left(\hat{\mathsf{b}}_q^\dagger \hat{\mathsf{b}}_q \right)^2 \right) + \\ + J^{\text{eff}} \left(\hat{\mathsf{b}}_1^\dagger \hat{\mathsf{b}}_2 + \hat{\mathsf{b}}_1 \hat{\mathsf{b}}_2^\dagger \right) + \epsilon(t) \left(\hat{\mathsf{b}}_1 + \hat{\mathsf{b}}_1^\dagger + \lambda \hat{\mathsf{b}}_2 + \lambda \hat{\mathsf{b}}_2^\dagger \right) , \end{aligned} \tag{5.30}$$

where the Lamb shifts have been absorbed in ω_q' and λ expresses the relative effective driving strength between the two qubits. While Eq. (5.30) matches the level-dependence of the effective quantities in Eq. (5.28) only approximately, it is still useful in practical applications. It is valid for weak anharmonicities and low qubit excitation [216]. Moreover, deviations from the exact theory may be compensated by measuring the effective parameters experimentally, instead of deriving them from first principles. The Hamiltonian is used as a simplified model for transmon qubits in chapters 6 and 7.

5.5 Gate Mechanisms for the Transmon

A variety of approaches have been used to implement entangling two-qubit gates on transmon qubits. We give a brief overview over some of the techniques that have been realized experimentally.

The conceptually most straightforward method is the *Direct Resonant iSWAP (DRI) Gate* [225]. The transmons are coupled statically with a capacitor, yielding an interaction on the order of 4 MHz. Since there is no drive on the transmission line resonator during the implementation of the gate, the cavity can be integrated out and the transmon truncated to two levels. The effective Hamiltonian therefore reads simply as

$$\hat{\mathsf{H}} = -\frac{\omega_1}{2} - \frac{\omega_2}{2} + g\, \hat{\sigma}_y \hat{\sigma}_y \,. \tag{5.31}$$

The interaction is "activated" by tuning the qubit frequencies into resonance on a time scale $\frac{1}{g} \ll \tau \ll \frac{1}{\omega_{1,2}}$, using a magnetic flux through the split Josephson junction. After a duration of $T = \frac{1}{8g}$, it generates a $\sqrt{\text{iSWAP}}$ up to single qubit transformations.

The *higher-level resonance-induced dynamical CPHASE (DP) gate* [226] explicitly exploits the effective interaction mediated by the cavity. It models the system with a generalized form of the Jaynes-Cummings-Hamiltonian of Eq. (5.18) as

$$\hat{\mathsf{H}} = \omega_c \hat{\mathsf{a}}^\dagger \hat{\mathsf{a}} + \sum_{q=1,2} \left[\sum_{i=0}^{N_q-1} \omega_i^{(q)} \left|i\right\rangle \left\langle i\right|_q + \left(\hat{\mathsf{a}} + \hat{\mathsf{a}}^\dagger\right) \sum_{i,j=0}^{N_q-1} g_{jk}^{(q)} \left|i\right\rangle \left\langle j\right|_q \right] , \quad (5.32)$$

with all anharmonicities absorbed in $\omega_i^{(q)}$ and with a coupling matrix $g_{jk}^{(q)}$. The qubit frequencies again are flux-tunable. For the parameter choices in Ref. [226], $\omega_c = 6.9\,\mathrm{GHz}$ and $g_1 = 199\,\mathrm{MHz}$, $g_2 = 183\,\mathrm{MHz}$, the logical level $|02\rangle$ decreases more rapidly than level $|11\rangle$ when tuning the qubit frequencies. For a specific flux value, these two levels would become degenerate. However, the cavity also mediates a coupling $|11\rangle \leftrightarrow |02\rangle$, cf. Eq. (5.30), causing an avoided crossing at the degeneracy point. The resulting large shift in $|11\rangle$ induces a relative phase and thus implements a CPHASE gate on a time scale of 30 ns, without applying a field to the waveguide resonator. However, this neglects the time required to tune the qubits into the interaction regime. As the avoided crossing is approached, the tuning speed must be sufficiently slow to remain adiabatic, inversely proportional to the splitting between $|11\rangle$ and $|02\rangle$.

Using the flux-tunability of the qubit frequency has the disadvantage of reduced coherence times due to flux noise and the risk of unwanted interactions during the tuning process [227]. An alternative is the use of shaped microwave pulses on the transmission line resonator.

The *resonator-sideband-induced (RSI) gate* [228] employs sideband transitions [229] common in trapped-ion implementations of quantum computing [46, 230]. Unlike the DRI or DP gates, the system is actively driven with pulses on the transmission line. The basic idea is to entangle the qubit state with the first excitation of the cavity, and then to transfer this entanglement onto the two-qubit state. The generation of entanglement can be understood by considering the initial state $|000\rangle$, where the quantum numbers indicate the excitation of the first qubit, second qubit, and the cavity. A series of pulses is applied, consisting of

1. a π pulse (see appendix C) on the second qubit,
2. a $\frac{\pi}{2}$ pulse on the sideband of the first qubit, and
3. a π-pulse on the sideband transition of the second qubit.

The sideband transition excites or de-excites both the qubit and the cavity simultaneously via a two-photon transition. This transforms the qubit-qubit-cavity state as (neglecting global phases)

$$|000\rangle \to |010\rangle \to \frac{1}{\sqrt{2}} \left(|010\rangle + |111\rangle\right) \to \frac{1}{\sqrt{2}} \left(|010\rangle + \mathrm{i}\,|100\rangle\right) .$$

In total, it brings the initial separable two-qubit state to an entangled Bell state,

$$|00\rangle \to \frac{1}{\sqrt{2}} \left(|01\rangle + \mathrm{i}\,|10\rangle\right).$$

Note that in the last step, there exists no right-qubit blue sideband transition for the first term $|010\rangle$, which enables the transfer of the entanglement to the qubits. The duration of the full pulse sequence for the parameters in Ref. [228] is $T = 150\,\mathrm{ns}$. As described in Ref. [231], the mechanism can be extended to implement a CNOT gate. Generally, many of the gate mechanisms for trapped ions can be translated to superconducting qubits, with the cavity taking the place of the vibrational excitation of the ions.

Two further gate implementations for transmon qubits exploit specific resonance conditions with levels outside of the logical subspace. Both of these start from an effective model, Eq. (5.30), where the qubit-cavity interaction has been transformed into an effective qubit-qubit interaction, and an effective driving term for the qubit transitions. Thus, we describe the state of the system in the basis $|ij\rangle$, where i and j are the excitation quantum numbers of the left and right qubit, respectively. While in general, the effective interaction is weak, building resonances into the higher qubit levels can significantly enhance the generation of entanglement. In the case of the *Bell-Rabi (BR) Gate* [222], the transmon is engineered such that the qubit-qubit detuning is close to the anharmonicity of one of the qubits, $\omega_1 - \omega_2 \approx \alpha_1$. This brings the levels $|11\rangle$ and $|20\rangle$ into resonance. Driving at a frequency slightly detuned by δ from the center between both qubits,

$$\omega_d = \frac{1}{2}\left(\omega_1 + \omega_2\right) - \delta\,, \tag{5.33}$$

the levels $|00\rangle$, $|11\rangle$, $|21\rangle$ are close to zero in the rotating frame at ω_d, and far detuned from the remaining levels. Using a Schrieffer-Wolff transformation similar to the one discussed in section 5.4, the coupling between this "low-energy manifold" and the remaining Hilbert space can be eliminated. In the frame defined by the transformation, it can be shown [222] that there is a resulting two-photon transition between the state $|00\rangle$ and $|11\rangle$. With a specific choice of the drive-detuning δ, this implements a gate that is equivalent to $\sqrt{\mathrm{iSWAP}}$.

The *Microwave-Activated CPHASE (MAP) Gate* [227] uses an alternative resonance condition. Here, the qubits are engineered such that $|12\rangle$ and $|03\rangle$ align, which translates to the condition

$$\omega_1 - \omega_2 = 2\alpha_2\,. \tag{5.34}$$

The effective coupling between these two levels shifts them, thus lifting the degeneracy of the two transitions $|01\rangle \leftrightarrow |02\rangle$ and $|11\rangle \leftrightarrow |12\rangle$. The system is driven with a detuning of δ from these two transitions,

$$\omega_d = \omega_2 + \alpha_2 - \delta\,. \tag{5.35}$$

With the resonance condition in Eq. (5.34), ω_d takes the same form as Eq. (5.33). In the rotating frame with ω_d, the combined Hamiltonian (drift Hamiltonian and drive Hamiltonian for a field with amplitude Ω) can be fully diagonalized. The field-dressed level $|11\rangle$ obtains a Stark shift. The shift induces a CPHASE gate, with the total interaction energy

$$\zeta \approx \zeta_0 + \frac{\Omega^2}{2\delta}\zeta_2\,, \tag{5.36}$$

where ζ_0 is the static interaction and the term proportional to ζ_2 is the field-induced interaction, both defined in Ref. [227]. For the parameters chosen there, a fully entangling gate is reached after 510 ns. A detailed discussion of the effective Hamiltonian and the gate mechanism for both the BR and MAP gate may be found in Ref. [216].

5.6 A Holonomic Phasegate

We might wonder whether we can generalize the idea of using Stark shifts for the implementation of gates to include the cavity. Instead of inducing the shift in the MAP gate by driving an off-resonant transition to higher qubit levels, it has been proposed to use a transition to a cavity level, implementing a *Driven Resonator-Induced CPHASE (RIP) Gate* [227, 232]. As an alternative picture, we could think of the resonance induced dynamical phase (DP) gate discussed in section 5.5, but instead of flux-tuning the qubit frequencies into resonance, the necessary shifts would be obtained by driving the system with an off-resonant microwave field near the cavity frequency.

We consider the full Hamiltonian of two transmon qubits coupled via a shared transmission line (cavity), Eq. (5.18), that is driven at a frequency ω_d detuned by a few tens of MHz from the cavity frequency ω_c,

$$\epsilon(t) = \epsilon_0 S(t)\cos(\omega_d t), \quad S(t) \in [0,1]\,. \tag{5.37}$$

The Hamiltonian can be transformed to a rotating frame at ω_d, see appendix B, with the basis transformation

$$\hat{\mathsf{U}}_{\text{RWA}}(t) = \sum_{ijn} \mathrm{e}^{\mathrm{i}(i+j+n)\omega_d t} \left|ijn\right\rangle \left\langle ijn\right|\,, \tag{5.38}$$

where i, j, n are the quantum numbers for the excitation of the left qubit, right qubit, and the cavity, respectively. Assuming Eq. (5.37), the field in the rotating wave approximation (RWA) is

$$\Omega(t) = \frac{1}{2}\epsilon_0 S(t)\,. \tag{5.39}$$

Going to the RWA allows to obtain the field-dressed energy levels by diagonalization of $\hat{\mathsf{H}}_{\text{RWA}}(t)$ for a fixed value of the field amplitude ϵ_0. A two-qubit interaction

$$E_\gamma(\epsilon_0) = E_{00}(\epsilon_0) - E_{01}(\epsilon_0) - E_{10}(\epsilon_0) + E_{11}(\epsilon_0) \tag{5.40}$$

results if the energy levels that define the logical subspace are shifted non-linearly. Keeping the field at amplitude ϵ_0 for duration T implements a diagonal gate

$$D = \operatorname{diag}\left\{\mathrm{e}^{\mathrm{i}\phi_{00}}, \mathrm{e}^{\mathrm{i}\phi_{01}}, \mathrm{e}^{\mathrm{i}\phi_{10}}, \mathrm{e}^{\mathrm{i}\phi_{11}}\right\}\,, \tag{5.41}$$

where the $\phi_{00} = \mathrm{e}^{\mathrm{i}E_{00}(\epsilon_0)T}$, and equivalently for ϕ_{01}, ϕ_{10}, ϕ_{11}. As discussed in chapter 2, this gate is a CPHASE gate with the entangling phase

$$\gamma(\epsilon_0, T) = E_\gamma(\epsilon_0)T\,, \tag{5.42}$$

cf. Eq. (2.30), up to single-qubit operations.

However, the implementation of the gate assumes that the population is in the field-dressed logical eigenstates at every point in time. This can be ensured by switching the pulse shape $S(t)$ on and off adiabatically: at $t = 0$, the pulse is off and the system is in one of the logical eigenstates, e.g. $|00\rangle$. The pulse is then slowly switched on, reaching its peak value ϵ_0 at $t = T/2$, after which it is slowly switched off again, reaching zero at $t = T$. If the switch-on and switch-off is sufficiently slow, then according to the adiabatic theorem, the system will remain in the field-dressed logical eigenstate with energy e.g. $E_{00}(\Omega(t))$ at every point in time. Specifically, the derivative of $\Omega(t)$ must be sufficiently small to avoid jumping over any avoided crossing, as described by Landau-Zener's formula [233]. At final time T, the system returns to its original state, but with a geometric Berry phase [234], e.g.

$$\phi_{00} = \int_0^T E_{00}(\Omega(t))\,\mathrm{d}t\,. \tag{5.43}$$

Thus, following the terminology of Ref. [235], we refer to this realization of a diagonal gate as a *holonomic phasegate*.

5.6.1 Entanglement Creation

The effective entangling energy obtained at final time T depends not only on the peak amplitude of the pulse, but also on the choice of the system

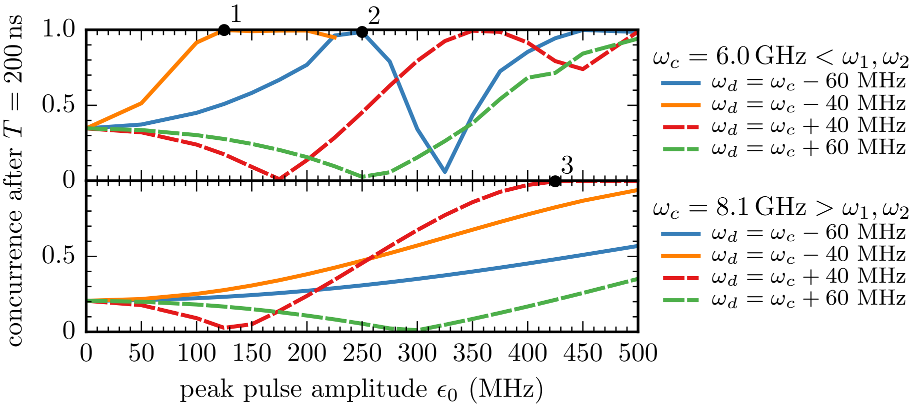


Figure 5.4: Entanglement (concurrence) generated after $T = 200\,\mathrm{ns}$, over the peak amplitude of the pulse, for different choices of pulse driving frequency ω_d. The top and bottom panels show results for a cavity frequency of 860 MHz left and right of the qubits, respectively. The remaining parameters are listed in Table 5.2. The gates implemented at the select points labeled 1–3 are analyzed in Table 5.3.

cavity frequency	ω_c	6.00, 8.11	GHz
left qubit frequency	ω_1	6.86	GHz
right qubit frequency	ω_2	7.25	GHz
qubit anharmonicity	α_1, α_2	-300	MHz
qubit-cavity coupling	g_1, g_2	70	MHz

Table 5.2: Parameters for the transmon Hamiltonian Eq. (5.18) for a holonomic phase gate.

point	$1 - C$	pop. loss	$\langle i, j \rangle_{\max}$	$\langle n \rangle_{\max}$	n_q	n_c
1	1.56×10^{-3}	1.70×10^{-2}	1.55	7.70	10	25
2	1.43×10^{-2}	9.89×10^{-3}	3.04	12.92	11	30
3	3.19×10^{-3}	3.60×10^{-3}	1.04	48.46	6	70

Table 5.3: Properties of gates implemented for the parameters at labeled points in Fig. 5.4. For each point, the entanglement error $1 - C$ is given, cf. Fig. 5.4 for the values of the concurrence C. Furthermore, the loss of population from the logical subspace at final time T, the peak expectation value $\langle i, j \rangle_{\max}$ at $t = \frac{T}{2}$ for the excitation of either one of the qubits, and the peak expectation value $\langle n \rangle_{\max}$ for the excitation of the cavity are listed. Lastly, n_q and n_c are the number of qubit and cavity levels that must be taken into account to reach numerical convergence.

parameters and the driving frequency of the microwave pulse. This is explored in Fig. 5.4, showing the entanglement generated by a Blackman pulse of duration $T = 200\,\text{ns}$ with varying peak amplitude ϵ_0. The Blackman shape is nearly identical to a Gaussian centered around $\frac{T}{2}$ with width $\sigma = \frac{T}{6}$, but is exactly zero at $t = 0$ and $t = T$, see Eq. (4.7). The qubit parameters are listed in Table 5.2. Even in the field-free case, $\epsilon_0 = 0$, the static coupling between qubit and cavity induces a Lamb-shift in the qubit levels. To first order, these shifts are described by $\chi_i^{(1,2)}$ in Eq. (5.29). Considering only the levels of the logical subspace, the cavity-mediated interaction energy is $\zeta \equiv E_\gamma(\epsilon = 0)$ with E_γ defined in Eq. (5.40). At final time T, this results in the entanglement

$$C(\zeta, T) = \sin\left[\frac{1}{2}\zeta T\right], \tag{5.44}$$

representing the y-intercept in Fig. 5.4. Note that for the DP gate, the qubit parameters are flux-tuned to ensure $C(\zeta, T) = 1$. Here, the qubit parameters are not tuned, but the interaction due to the pulse-induced Stark shifts (from the combination of all terms proportional to $\hat{\mathsf{a}}$, $\hat{\mathsf{a}}^\dagger$ in Eq. (5.29), and higher order) yields the entanglement for the realization of the gate. For a two-qubit gate, it must add to the always-on entanglement due to ζ, whereas for single qubit gates, it must cancel ζ. The field-free parameters in Table 5.2 are chosen such that the entanglement created due to ζ at time T is less than 0.5, to ensure that the realization of single-qubit gates is not harder than that of two-qubit gates.

The magnitude of the Stark shift is proportional to $|\epsilon_0|^2$, i.e., it depends quadratically on the field. This reflected in the parabolic shape of the curves in Fig. 5.4.

Generally, the magnitude and sign of the Lamb and Stark shifts depends on the magnitude and sign of the qubit-qubit detuning, the qubit-cavity detuning, and the cavity-drive detuning. Instead of an exact theoretical analysis, based on Eq. (5.29) and higher order terms (a procedure that quickly becomes too tedious to perform by hand), we can obtain empirical insight into how these quantities affect the generation of entanglement from Fig. 5.4.

The two panels consider the situation where the cavity frequency is left or right of the qubits, while the qubits are at fixed frequencies ω_1 and ω_2, see Table 5.2. In both cases, the absolute value of the qubit-cavity detuning is the same, but the sign of the detuning changes. This leads to a change in the always-on interaction ζ, resulting in less entanglement in the field-free case when $\omega_c > \omega_{1,2}$.

Furthermore, when higher levels of the transmon are taken into account, the Hamiltonians for the two choices of ω_c are in fact *not* symmetric with respect to the qubits: With a negative anharmonicity, higher qubit transitions are closer to the cavity frequency if $\omega_c < \omega_{1,2}$ and increasingly farther detuned if $\omega_c > \omega_{1,2}$. Since avoided crossings with higher levels can be relevant for the resulting shifts in the levels of the logical subspace, this can have considerable impact on the implementation of the gate.

The choice of the drive frequency w_d as either smaller or larger than the cavity frequency determines whether the Lamb and Stark shifts interfere constructively or destructively; driving left of cavity (solid blue and orange lines) results in constructive interference, whereas driving right of the cavity (dashed red and green lines) results in destructive interference. This is independent of the sign of the qubit-cavity detuning, as the same behavior is observed in both panels.

In the case of constructive interference, we observe that the rate at which entanglement is created is proportional to ζ , from the greater slope of the solid blue and orange curve in the top panel, compared to the bottom panel. This may be expected from Eq. (5.29) by the observation that both the Lamb shift and the Stark shift contain $\chi_i^{(q)}$ as the leading term. Interestingly, in the case of destructive interference ($\omega_d > \omega_c$) for a driving frequency close to the cavity (dashed red curve), we do not observe the same behavior; this curve is simply shifted down in the bottom panel. Presumably, this choice of parameters causes higher order terms in the Stark shift to cancel the leading order term.

The implication of this is that for small values of ζ, the fastest way to realize an entangling gate (respectively, using the smallest pulse amplitude) may be to choose a drive frequency such that the entanglement from the Lamb and the Stark shift interfere destructively. For small pulse amplitudes, respectively short durations, the same drive frequency could be used to implement both single qubit gates ($\epsilon_0 = 125\,\text{MHz}$ for the red curve in the bottom panel), while stronger drives, respectively longer durations, would allow the realization of an entangling gate (point 3 at $\epsilon_0 = 425\,\text{MHz}$ in the bottom panel). In contrast, for the top panel one would implement entangling gates at points labeled 1 or 2, but switch the drive frequency to $\omega_d > \omega_c$ for single-qubit gates.

Setting the drive frequency closer to that of the cavity increases the Stark shift significantly, cf. (5.36), as shown by the greater slope of the orange and red curves relative to the blue and green curves, in both panels. However, getting too close to the cavity will drive up cavity population significantly. Also, it will cause non-adiabatic effects. That is, the dynamics might jump over at least some of the avoided crossings. These are hinted at by kinks in the curves of Fig. 5.4. In more detail, we consider the numbered points in Fig. 5.4. The properties of the dynamics at these points are listed in Table 5.3.

The population dynamics of the state $|00\rangle$ for the parameters at point

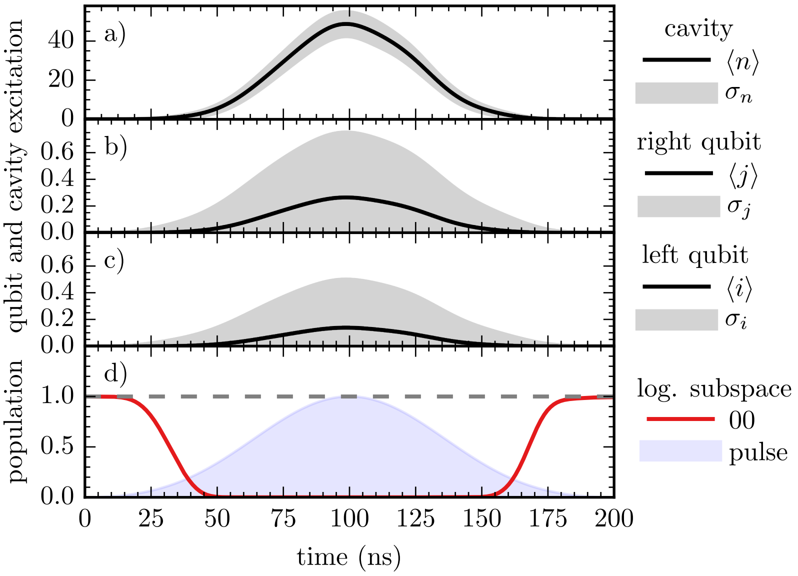


Figure 5.5: Exemplary population dynamics for an adiabatic pulse. The dynamics shown are for the propagation of the logical eigenstate $|00\rangle$ for the parameters corresponding to point (3) in the bottom panel of Fig. 5.4. Panels (a-c) show the expectation value of the cavity (a), the right qubit (b), and the left qubit (c), in the bare basis. The standard deviation for each expectation value is shown in gray. In panel (d), population in the (non-field-dressed) logical eigenstate $|00\rangle$. The pulse shaped, normalized to the peak amplitude $\epsilon_0 = 425\,\text{MHz}$, is shown in light blue. The dashed gray line indicates the value 1.0.

3 are shown in Fig. 5.5. The dynamics are exemplary for an adiabatic evolution. Note that excitations are the projection of the field-dressed states on the non-dressed basis. This shows intuitively the effect of the off-resonant drive: it shifts the qubit and cavity wavepackets out of their equilibrium position proportionally to the square of the field amplitude shown in light blue in panel (d). Most importantly, the adiabaticity of the dynamics result in the population, panel (d), returning exactly to the initial state at final time T, up to a Berry phase. Non-adiabatically jumping over an avoided crossing would be reflected in the plot of the qubit and cavity excitation by kinks and asymmetries; the excitation would no longer smoothly follow the pulse shape. As a consequence, the system would generally no longer return to the initial state. The non-adiabaticity can be measured by the loss from the logical subspace,

$$p_{\text{loss}} = 1 - \frac{1}{4}\,\text{tr}\left[\tilde{\mathsf{U}}^{\dagger}\tilde{\mathsf{U}}\right]\,, \tag{5.45}$$

where $\tilde{\mathsf{U}}$ is the projection of the full time evolution operator onto the logical subspace. For the three numbered points in Fig. 5.4, the value of p_{loss} is listed in Table 5.3. Points 1 and 2 show a significantly larger loss than point

3. Accordingly, for point 3, there are no discernible nonadiabatic effects visible in the population dynamics, while for points 1 and 2, the propagation of at least one of the logical basis states has some visible deviation from the ideal dynamics.

The shift of the qubit and cavity excitation is also inverse proportional to the detuning of the drive from the respective frequency. Thus, in Fig. 5.5, the cavity, panel (a), is strongly excited, whereas the right qubit, panel (b), is only slightly excited, as it is far off-resonant. The left qubit, panel (c), is even further detuned from the drive, and it thus receives even less excitation. The data for the peak expectation value for the qubit and cavity excitation in Table 5.3 illustrates this further. For points 1 and 2, $\omega_1, \omega_2 < \omega_d < \omega_c$, and both qubit and cavity population increase as the pulse amplitude increases from $\epsilon_0 = 125\,\text{MHz}$ for point 1 to $\epsilon_0 = 250\,\text{MHz}$ for point 2. However, since for point 3, the cavity drive is much farther detuned from the qubits, $\omega_1, \omega_2 < \omega_c < \omega_d$, the qubit excitation is much lower, even though the pulse amplitude increases to $\epsilon_0 = 425\,\text{MHZ}$.

Another observation based on the data of Table 5.3 is that for points 1 and 2, the number of qubit and cavity levels that need to be taken into account in the dynamics is higher than the excitation would imply. This illustrates that higher levels of the transmon can be relevant to the realization of the gate, even without being populated. The same situation occurs in the DP gate mechanism in section 5.5, where the $|02\rangle$ level causes the shift in the $|11\rangle$ level, without ever being populated. When the cavity and drive frequency are to the right of the qubits, far less qubit levels need to be taken into account.

Beyond the parameters explored in Table 5.2 and Fig. 5.4, increasing the qubit-cavity detuning has the effect of reducing the entanglement generated by the Stark shift. Decreasing it to values significantly smaller than those used in Table 5.2 leads to strongly non-adiabatic dynamics. Similarly, lowering the value of the qubit-cavity coupling g reduces the generation of entanglement, but enlarging it leads to undesirably large values of ζ.

The empirical observations on the parameter-dependence of the entanglement creation are essential for a good choice of parameters for the realization of a holonomic phasegate, and moreover they are relevant for guiding a more rigorous analytical derivation. Both points 1 and 3 in Fig. 5.4 are worthwhile for further consideration. Placing the cavity frequency to the right of the qubits has the significant benefit of providing a lower value of ζ, allowing to implement single and two-qubit gates with the same drive frequency, showing more adiabatic behavior despite stronger pulse amplitudes, and exciting the qubit levels far less. However, the strong pulses imply a large cavity population. Placing the cavity left of the qubits allows to implement the gate with far less pulse amplitude and cavity excitation, but runs the risk of population loss from the logical subspace due to non-adiabatic effects.

5.6.2 Simplex Optimization of a Holonomic Gate

Since the points in Fig. 5.4 only sample a discrete set of parameter choices, none of the gates considered there are exact perfect entanglers, cf. the entanglement error in Table 5.4. Furthermore, the gate duration has been fixed at $T = 200\,\text{ns}$. In order to find exact perfect entanglers and also to minimize the gate duration, we can use a simplex optimization as discussed in chapter 3 to optimize the gate duration and the peak amplitude of the Blackman shape, for the two most promising types of parameters of Fig. 5.4, $\omega_d < \omega_c = 6.0\,\text{GHz} < \omega_1, \omega_2$ and $\omega_d > \omega_c = 8.11\,\text{GHz} > \omega_1, \omega_2$. These correspond to the most promising parameter sets from Fig. 5.4: the orange and blue curve in the top panel, and the green and red curve in the bottom panel. We also explore the option of using smaller pulse-cavity detunings.

The figure of merit for the optimization must encode three requirements:

1. The obtained gate must be a perfect entangler, as measured by the non-local phase γ defined in Eq. (5.42), which should take the value π.
2. The gate duration T should be as short as possible.
3. The dynamics should be adiabatic, as measured by population loss from the logical subspace, Eq. (5.45).

This allows to write the functional

$$F(\epsilon_0, T) = (1 - p_{\text{loss}})^n \frac{|\gamma/\pi|}{T^{1/m}}, \qquad n = 10, \quad m = 1, 2\,, \tag{5.46}$$

which takes its maximum value if all three requirements are fulfilled. To weight the three conditions, the exponents n and m are used. Adiabaticity is the most difficult to achieve, and is weighted most strongly by setting $n = 10$. For the gate duration, $m = 2$ has proved the most effective choice, although $m = 1$ can be used to put more emphasis on a shorter gate duration.

The results of the simplex optimization are shown in Table 5.4. The obtained gates are comparable to those obtained from systematic variation in Fig. 5.4 and Table 5.3. Nearly all of the gates obtained by the simplex optimization are now exact perfect entanglers.

For $\omega_c = 6.00\,\text{GHz}$ and $\omega_d = \omega_c - 40\,\text{MHz}$, the optimization yielded considerable improvement of the quality of the gate, compared to point 1 in Table 5.3. The gate duration is slightly shorter, while population loss has decreased and entanglement has increased. For $\omega_c = 8.11\,\text{GHz}$ the results are less positive; for $\omega_d = \omega_c + 40\,\text{MHz}$, there is in fact a larger loss of population than for point 3 in Table 5.3 (although the gate is now a perfect entangler). This illustrates a shortcoming of the simplex algorithm, which is sensitive to local traps. A slight change in parameters can cause

δ [MHz]	T [ns]	ϵ_0 [MHz]	pop. loss	$\langle i,j\rangle_{\max}$	$\langle n\rangle_{\max}$	n_q	n_c
$\omega_d < \omega_c = 6.00\,\mathrm{GHz} < \omega_1, \omega_2$							
-20	423	19	1.32×10^{-1}	2.34	21.75	12	50
-30	214	44	4.53×10^{-3}	1.15	4.81	9	40
-40	165	121	7.50×10^{-3}	1.36	7.99	10	40
-40*	146	127	1.09×10^{-2}	1.40	8.21	10	60
-50	311	144	5.47×10^{-3}	1.15	5.20	5	40
$\omega_d > \omega_c = 8.11\,\mathrm{GHz} > \omega_1, \omega_2$							
20	291	67	2.30×10^{-2}	1.01	9.92	5	40
30	308	174	9.69×10^{-3}	1.02	18.24	5	50
40	220	372	1.10×10^{-2}	1.05	36.31	6	70
40*	125	471	5.94×10^{-2}	1.10	64.82	6	100
50	168	670	1.27×10^{-2}	1.09	65.64	6	110

Table 5.4: Simplex optimization of pulse time T and peak pulse amplitude ϵ_0 for the realization of a holonomic phasegate, for different pulse-cavity detunings $\delta \equiv \omega_d - \omega_c$ and for two different cavity frequencies (left and right of the qubits, cf. the two panels in Fig. 5.4). All parameters are listed in Table 5.2. The figure of merit for the optimization is Eq. (5.46) with $n = 10$ and $m = 2$. We list the loss of population from the logical subspace at final time T, the peak expectation values $\langle i,j\rangle_{\max}$, $\langle n\rangle_{\max}$ of the qubit and cavity excitation, and the number n_q, n_c of qubit and cavity levels that must be included for numerical convergence. In the optimizations labeled by an asterisk, the figure of merit is taken with $m = 1$, putting more emphasis on short gate durations.

a sudden non-adiabatic loss of population, producing a local minimum for the optimization functional.

The results for varying values of $\delta = \omega_d - \omega_c$ confirm that when placing the drive frequency closer to the cavity, lower pulse amplitudes produce higher entanglement, but also, it is easier to lose population from the logical subspace. For $\delta = -20\,\mathrm{MHz}$, the gate fails completely: there is large excitation of the qubit, 13% loss of population, and the concurrence only reaches 0.84. This explains the somewhat counter-intuitive long gate durations for both small and large values of δ. For small values, the loss of adiabaticity can be countered by making the pulse longer, and thus change more slowly. For large values, it is more difficult to obtain entanglement, so that the pulse needs to be active for a longer duration before a concurrence of 1.0 is reached. For $\delta = -50\,\mathrm{MHz}$, $1 - C = 8 \times 10^{-4}$. The best result is obtained for $\delta = -40\,\mathrm{MHz}$ (shortest gate duration) or $\delta = -30\,\mathrm{MHz}$ (least loss of population). For $\omega_c = 8.11\,\mathrm{GHz}$, moving the drive frequency closer to the cavity is less detrimental; a perfect entangler is still achieved, albeit

cavity frequency	ω_c	8.3	GHz
left qubit frequency	ω_1	6.5	GHz
right qubit frequency	ω_2	6.6	GHz
qubit anharmonicity	α_1, α_2	-300	MHz
qubit-cavity coupling	g_1, g_2	100	MHz

Table 5.5: Parameters for the transmon Hamiltonian Eq. (5.18) [236]

at a relatively high loss of population.

Shorter gates can be obtained by putting more emphasis on the minimization of T, as was done for the two optimizations labeled with an asterisk. However, this is at the cost of larger pulse amplitudes and increased loss of population. Also for $\omega_c = 8.11\,\text{GHz}$ and $\delta = 40\,\text{GHz}$, the gate falls short of being an exact perfect entangler, $1 - C = 1 \times 10^{-5}$.

Lastly, the results in Table 5.4 confirm the observations of Fig. 5.4 and Table 5.3, that choosing $\omega_c > \omega_1, \omega_2$ requires far stronger pulses, causing very large cavity excitation. On the other hand, far less qubit levels need to be taken into account.

5.7 Optimization of Transmon Quantum Gates with Krotov's Method

5.7.1 Optimization for CPHASE and CNOT

In order to explore the potential for more advanced optimal control of two-qubit quantum gates for two transmons with a shared transmission line, we use Krotov's method as presented in chapter 3 to optimize the full system Hamiltonian in Eq. (5.18) for the parameters given in Table 5.5. The qubit parameters are assumed fixed, and the control is entirely through the microwave field on the transmission line resonator. This distinguishes the approach from recent applications of similar control techniques to flux-tunable qubits where the control parameter was the qubit-cavity detuning [237].

The optimization functional to be minimized is

$$J = 1 - \frac{1}{16} \left| \sum_{i=1,4} \left\langle \phi_k \,\middle|\, \hat{\mathsf{O}}^\dagger \hat{\mathsf{U}}(T, 0; \epsilon) \,\middle|\, \phi_k \right\rangle \right|^2 + \frac{\lambda_a}{S(t)} \int_0^T (\Delta\epsilon(t))^2 \, \mathrm{d}t \,, \quad (5.47)$$

where $\{|\phi_k\rangle\}$ are the states $\{|000\rangle, |010\rangle, |100\rangle, |110\rangle\}$ that span the logical two-qubit subspace. The three quantum numbers indicate the excitation of the first qubit, second qubit, and the cavity. $\hat{\mathsf{U}}(T, 0; \epsilon)$ is the time evolution operator for the propagation under the pulse $\epsilon(t)$, $\hat{\mathsf{O}}$ is the target quantum gate. We optimize both for $\hat{\mathsf{O}} = \text{CPHASE}$ and $\hat{\mathsf{O}} = \text{CNOT}$. The shape function $S(t)$ ensures a smooth switch-on and switch-off, λ_a is an arbitrary scaling parameter. Starting from a guess pulse driving at the frequency of the second qubit and small amplitude, Krotov's method iteratively updates the pulse according to

$$\Delta\epsilon(t) = \frac{S(t)}{\lambda_a} \mathfrak{Im} \left[\sum_{k=1}^{4} \left\langle \chi_k^{(i)}(t) \middle| \left(\hat{\mathsf{a}} + \hat{\mathsf{a}}^\dagger \right) \middle| \phi_k^{(i+1)}(t) \right\rangle \right] , \tag{5.48}$$

cf. Eq. (3.62), in order to minimize Eq. (5.47). The states $\{|\phi_k\rangle\}$ are forward propagated with the updated pulse $\epsilon^{(i+1)}(t)$, while the states $\{|\chi_k\rangle\}$ are backward-propagated with the pulse $\epsilon^{(i)}(t)$ of the current iteration, with the boundary condition given by Eq. (3.65).

The results of the optimization for a CPHASE gate with pulse durations between 100 ns and 1000 ns are shown in Fig. 5.6. The gate error $1 - F_{\text{avg}}$ with F_{avg} according to Eq. (2.88) is shown in blue in panel (d). The gate error is limited to values $> 1 \times 10^{-3}$, i.e., slightly above the quantum error correction limit. For gate durations ≥ 200 ns, the gate error is entirely due to loss of population from the logical subspace at time T, evaluated according to Eq. (5.45). It is shown as the black dashed line in panel (d) of Fig. 5.6. For $T = 100$ ns, the gate error is not simply due to the inability to bring return the population to the logical subspace: for short gate durations, there is not sufficient time to exploit the cavity-mediated qubit interaction to generate the necessary entanglement for the desired gate. Panels (b) and (c) show the peak expectation value of the qubit and cavity population during the propagation of any of the logical basis states. For the qubit, it shows that population is mostly within the first three levels of the transmon. The system would *not* accurately be described by truncating it to two levels, however. In fact, the number of levels that must be included in the simulation for full convergence is a multiple of the mean excitation number. The cavity is strongly driven to large excitations, which is the cause for the limited gate fidelity. The gate error in panel (d) and the peak cavity excitation in panel (b) are strongly correlated. In cases where the cavity population reaches large values during the implementation of the gate, it becomes extremely hard to bring the cavity to the ground state at final time T. This results in the loss of population from the logical subspace, and limits the gate fidelity. The large cavity population also makes the gate sensitive to spontaneous decay of the cavity levels. The yellow and red curve in panel (d) show the gate error if the cavity decays weakly, with a lifetime of $\tau = 100\,\mu\text{s}$, and more strongly, with a lifetime of $\tau = 25\,\mu\text{s}$, exploring the realistic range for current transmission line resonators. In both cases, the gate error increases to values of 1×10^{-2} or worse.

The system of two coupled transmons can implement not only a CPHASE

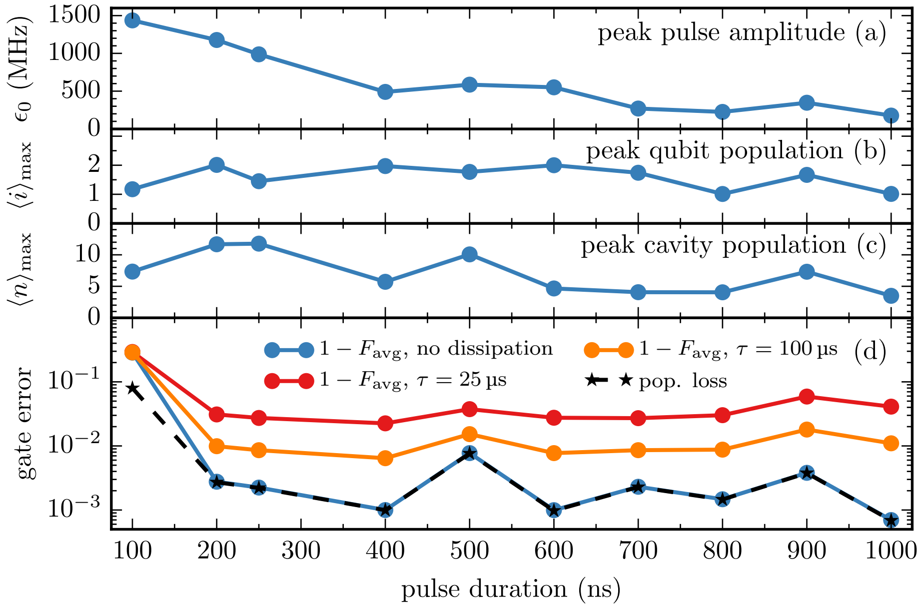


Figure 5.6: Results for the unconstrained optimization of a CPHASE gate for different gate durations. The panels show the peak amplitude of the optimized pulse (a), the peak expectation value of the qubit number operator (b) and the cavity number operator (c). Panel (d) shows the gate error without dissipation, and with spontaneous decay of the cavity for a lifetime of $\tau = 100\,\mu\text{s}$ and $\tau = 25\,\mu\text{s}$. For the non-dissipative case, the loss of population from the logical subspace at final time T is shown.

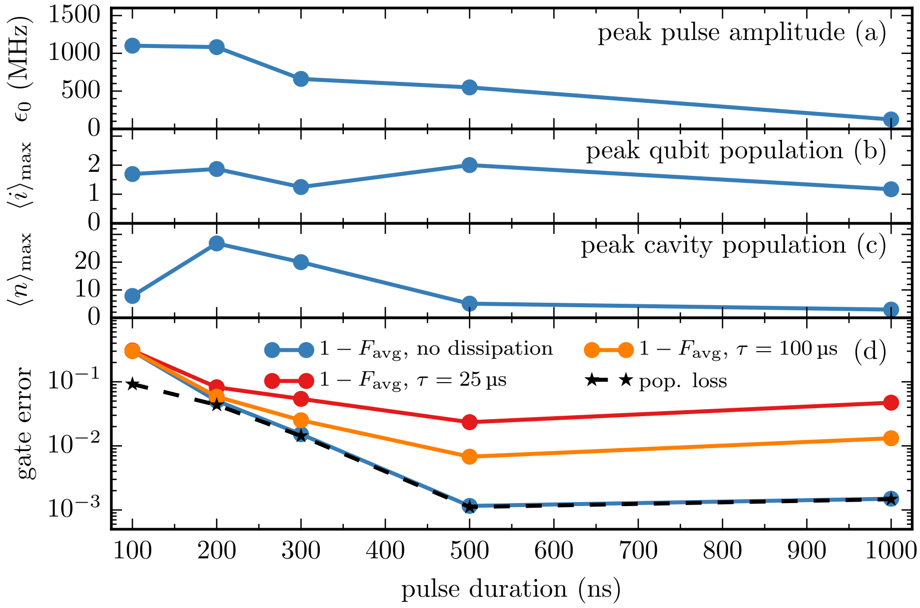


Figure 5.7: Results for the unconstrained optimization of a CNOT gate for different gate durations, cf. Fig. 5.6.

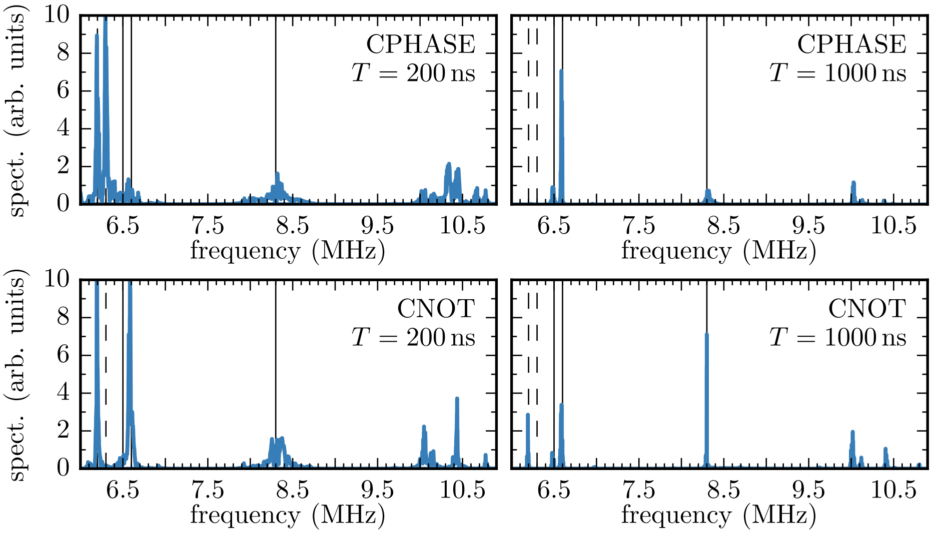


Figure 5.8: Spectra of select pulses resulting from the optimization for CPHASE (top) and CNOT (bottom), for a gate duration of $T = 200\,\text{ns}$ (left) and $T = 1000\,\text{ns}$ (right), cf. Fig 5.6 and Fig. 5.7. The bare qubit and cavity frequencies are indicated by the solid black vertical lines, and the qubit transitions $|1\rangle \rightarrow |2\rangle$ by the dashed lines.

gate, but but a wide range of other entangling quantum gates. In Fig. 5.7, the results for an optimization towards a CNOT gate for a few gate durations is shown. The pulse amplitudes in panel (a) are similar to those obtained in the CPHASE optimization. Also, the reachable gate error is on the order of 10^{-3}. The qubit excitation again shows significant population of the third level of the anharmonic transmon ladder. For short gate durations, the peak cavity population is significantly higher than for CPHASE, resulting in considerably worse fidelities at 200 ns and 300 ns. A quantum speed limit, where the gate error is not determined just by the loss of population from the logical subspace alone is again found for $T < 200\,\text{ns}$.

While the optimization yields complicated dynamics that make a detailed understanding of the gate mechanisms far from trivial, it is instructive to consider the spectra of the optimized pulses. For the optimized pulses of CPHASE and CNOT, for each gate duration of $T = 200\,\text{ns}$ and $T = 1000\,\text{ns}$, the spectra are shown in Fig. 5.8. Intuitively, for $T = 200\,\text{ns}$, the implementation of the gate is much harder than for $T = 1000\,\text{ns}$ and requires greater pulse amplitude. Consequently, there are more spectral components and broader peaks in the spectra for the short gate durations. For CPHASE at $T = 200\,\text{ns}$, the spectrum is dominated by the frequencies corresponding to the anharmonic transition $|1\rangle \rightarrow |2\rangle$ for the both left and right qubit. This matches the peak qubit excitation of $\langle i \rangle = 2$ in panel (b) of Fig. 5.6. The two qubit frequencies and the cavity frequencies are also present, together with high-frequency sideband transitions. For example, the first two of these transitions at 10.0 GHz and 10.1 GHz correspond to the transitions

$$\left|1\right\rangle_q \otimes \left|0\right\rangle_c \;\longleftrightarrow\; \left|0\right\rangle_q \otimes \left|2\right\rangle_c\,, \qquad q = 1, 2 \tag{5.49}$$

that transfers two cavity excitations into one qubit excitation of the left or right qubit, or vice versa. For $T = 1000\,\text{ns}$, the spectrum for the CPHASE gate is relatively close to the spectrum of the original guess pulse, driving mainly on the frequency of the second qubit. However, the cavity frequencies and the sideband transition of Eq. (5.49) are also present. The absence of the frequency driving the $\left|1\right\rangle \leftrightarrow \left|2\right\rangle$ transition of the qubits corresponds to the peak qubit population of $\langle i \rangle = 1$ for $T = 1000\,\text{ns}$ in panel (c) of Fig. 5.6. The optimized pulses for a CNOT gate are roughly similar to those for CPHASE. However, the combination of driving the right qubit frequency and the left anharmonic transition appear relevant, as they are present both for short and long gate duration. Counterintuitively, the significant drive of the cavity frequency for CNOT at $T = 1000\,\text{ns}$ does *not* correspond to a large peak cavity population in panel (c) of Fig. 5.7. This illustrates that the dynamics of the optimized pulses are very non-trivial and that the high-frequency sideband frequencies play a significant role. For example, it is conceivable that the transition in Eq. (5.49) acts in conjunction with the cavity drive to manipulate the qubit state. Lastly, the spectra also illustrate the Stark shift of the cavity level, respectively the Lamb shifts on the qubit levels that is induced by the qubit-cavity interaction: the qubit frequencies are shifted very slightly to the left of the bare frequency, whereas the cavity frequency is very slightly shifted right.

5.7.2 Optimization for a Holonomic Phasegate

Given the limited success of an unconstrained direct optimization for a CPHASE or a CNOT gate, we may attempt to set up the optimization for a specific gate mechanism. Therefore, we return to the holonomic phasegate and attempt to use Krotov's method to improve upon the results of section 5.6. The first term in Eq. (5.47) must be replaced with a final time functional that reflects the objectives of the holonomic mechanism. While the functional in Eq. (5.46) used in the simplex optimization is not suitable for Krotov's method, since it would be extremely hard to evaluate the boundary condition for the backward-propagated states, Eq. (3.56), it is relatively straightforward to derive a more suitable functional. The two conditions that we consider are:

1. The gate should be diagonal; every one of the logical eigenstates

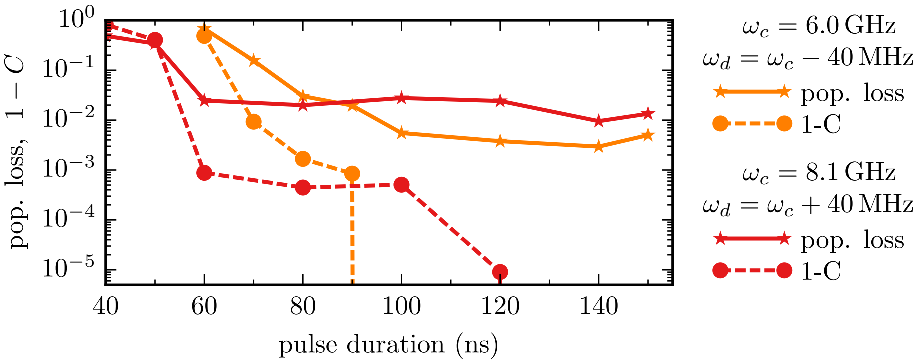


Figure 5.9: Success of optimization for a holonomic phasegate using Krotov's method for two parameter sets, cf. Fig. 5.4. All remaining parameters are given in Table 5.2. The success is evaluated in terms of the loss of population from the logical subspace at final time T, and the concurrence C of the obtained gate $\hat{\mathsf{U}}$.

should be mapped onto itself, up to a Berry phase.

2. The Berry phases should yield maximum entanglement, $\gamma = \pi$, according to Eq. (2.30).

The two conditions enter as two terms in the final time functional,

$$J_T = J_{ss} + w_\gamma J_\gamma \,, \tag{5.50}$$

with a factor w_γ to weight their relative importance. With

$$\tau_{00} = \left\langle 00 \middle| \hat{\mathsf{U}} \middle| 00 \right\rangle \,, \tag{5.51}$$

and equivalently for $|01\rangle$, $|10\rangle$, and $|11\rangle$, the first term takes the form

$$J_{ss} = 4 - |\tau_{00}|^2 - |\tau_{01}|^2 - |\tau_{10}|^2 - |\tau_{11}|^2 \,, \tag{5.52}$$

i.e. four simultaneous state-to-state transitions [138]. As shown in appendix H, the second term can be written as

$$J_\gamma = 2 + \tau_{00}\tau_{01}^*\tau_{10}^*\tau_{11} + \tau_{00}^*\tau_{01}\tau_{10}\tau_{11}^* \,. \tag{5.53}$$

The boundary condition for the backward propagation may also be found in appendix H.

The results of an optimization for the two parameter sets corresponding to the solid orange curve in the top panel and the dashed red curve in the bottom panel of Fig. 5.4, used also in the simplex optimization in section 5.6.2, is shown in Fig. 5.9. The gate durations are varied between 40 ns and 150 ns, the approximate shortest gate duration obtained in the simplex optimization. A fully entangling gate can be achieved for $T > 90\,\mathrm{ns}$

($\omega_c = 6.0\,\mathrm{GHz}$), respectively $T > 120\,\mathrm{ns}$ ($\omega_c = 8.1\,\mathrm{GHz}$). However, there is considerable loss of population from the logical subspace, comparable to the results obtained by simplex optimization, cf. 5.4. As observed both in section 5.6.2 and section 5.7.1, it is the inability to return the state entirely to the logical subspace that fundamentally limits the optimization success to values above the error correction threshold.

The population dynamics for $T = 120\,\mathrm{ns}$ and $\omega_c = 6.0$, cf. the respective point on the solid red line in Fig. 5.9 is shown in Fig. 5.10 and should be compared to Fig. 5.5. The dynamics are considerably non-adiabatic, showing direct driving of both the qubits and the cavity. However, the highly symmetric shape of the pulse, the population dynamics and the excitation of both qubits and cavity correspond to what one would hope to achieve by applying optimal control to the holonomic gate. Since adiabaticity cannot be achieved on the given time-scale, optimal control corrects the non-adiabatic effects by ensuring that for a symmetric pulse, any jumps over an avoided crossing that happen during the switch-on also happen during the switch-off. However, the optimized pulse is only partially successful in this endeavor, as evidenced by the non-zero excitation of the left qubit at final time T, cf. panel (c).

The spectrum of the optimized pulse in panel (d) of Fig. 5.10 is shown in the top panel of Fig. 5.11, together with the spectrum of the optimized pulse for $\omega_c = 8.1\,\mathrm{GHz}$, also at $T = 120\,\mathrm{ns}$ (bottom panel). In both cases, the optimized pulse does not only contain the central frequency for the off-resonant drive at ω_d, but also frequencies that resonantly drive the cavity and both qubits. In the top panel, these are shifted significantly from their bare values, whereas for the bottom panel, the shifts are not discernible. This corresponds to the results of Fig. 5.4. The mirrored frequencies are due to the pulse being complex-valued in the RWA and correspond to the time-dependence of the complex phase.

For all the optimizations presented in this section, the CPHASE, CNOT, and the holonomic gate, the fidelities of the optimized gates are limited by loss of population from the logical subspace. To achieve high fidelity gates, more advanced approaches will be required. First, it appears that the Krotov-gradient is not sufficiently sensitive to the small but still significant loss of population. Switching the optimization to a method that uses second order information as soon as convergence stagnates, such as the GRAPE/LBFGS method presented in chapter 3 provides a possible solution. More fundamentally, however, it might be more promising to avoid the problematic and unwanted excitation of high qubit and cavity levels altogether, by placing constraints on the optimization. These could take the form of state-dependent constraints [156] to avoid high excitations, or using spectral constraints [158, 159] to enforce simpler spectra in the optimized

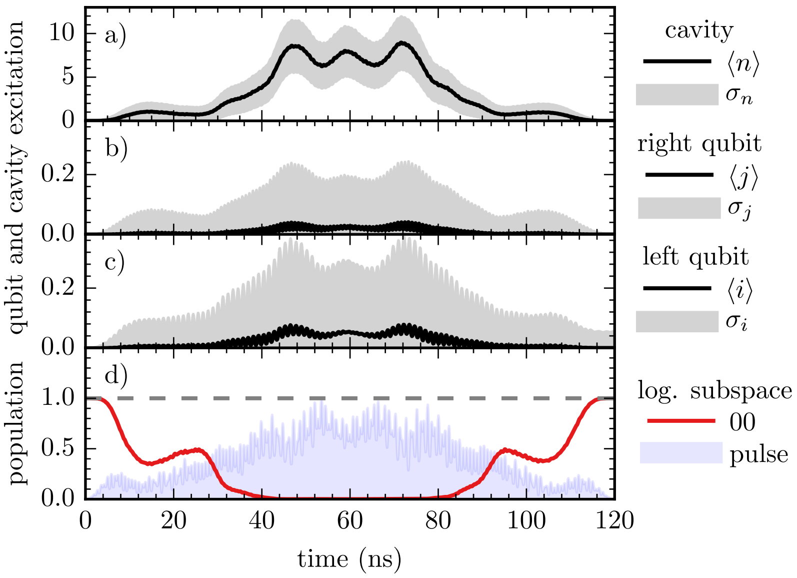


Figure 5.10: Population dynamics for logical eigenstate $|00\rangle$ under the optimized pulse for $T = 120\,\text{ns}$ and $\omega_c = 6.0\,\text{GHZ}$, cf. the corresponding point in the solid red line of Fig. 5.9. The figure follows the same conventions as Fig. 5.5.

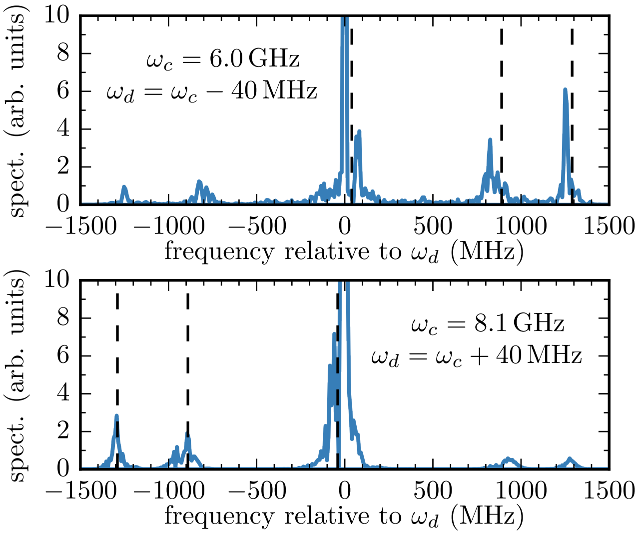


Figure 5.11: Spectra of pulses for $T = 120\,\text{ns}$. The top panel corresponds to the pulse shown in panel (d) of Fig. 5.10, for $\omega_d < \omega_c = 6.0\,\text{GHZ} < \omega_1, \omega_2$ (cf. yellow curves in Fig. 5.9). The bottom panel shows the spectrum of the optimized pulse for $\omega_d > \omega_c = 8.1\,\text{GHZ} > \omega_1, \omega_2$ (cf. red curves in Fig. 5.9). The spectrum is shown in the rotating frame at ω_d, corresponding to zero in the spectrum. The bare frequencies of the qubit and the cavity are indicated by vertical dashed lines. The central peaks extend to a value ≈ 40.

pulses or to remove e.g. cavity driving from the optimization of the CPHASE and CNOT, and resonant driving from the optimization of the qubit. A further option is to reduce the complexity of the model by considering an effective model of the cavity-mediated qubit-qubit interaction, as discussed in section 5.4. This is the approach we take in chapters 6 and 7. However, we have found that an unconstrained optimization in the effective model will generally cause the correspondence between the full and the effective model to break down. Again, this issue might be addressed by placing constraints on the optimization.

Limiting the excitation of the qubits and the cavity will also improve the robustness of the implemented gates with respect to decoherence, cf. Figs. 5.6 and 5.7, as the lower levels are least affected by spontaneous decay and dephasing. Going further, once the optimization of the transmon qubits has been brought to yield fidelities below the quantum error correction limit, the techniques of chapter 4 and the following chapters 6 and 7 may be used to address the ultimate goal of designing quantum gates for the transmon that are as robust as possible against the effects of decoherence.

6

Optimization for a Perfect Entangler

The implementation of a universal quantum computer requires at least one entangling two-qubit gate [86, 87], together with arbitrary single-qubit gates. However, as discussed in chapter 2, this need not be one of the "standard" gates, such as CNOT. The construct of the Weyl chamber, see section 2.4 in chapter 2, shows that for any two-qubit gate, there is an infinite number of equivalent quantum gates that differ only by additional local operations. For example the CPHASE gate is equivalent in this sense to CNOT, see Eq. (2.45).

In an optimization context, this has far-reaching consequences. Not every Hamiltonian allows for the implementation of arbitrary two-qubit gates. For example, the trapped Rydberg atoms discussed in chapter 4 can only implement diagonal gates, with a fixed global phase. The optimization target must be chosen to properly reflect this. Reachability is not the only consideration, however. Equally important is the question whether one two-qubit gate may be "easier" to implement than some other (possibly equivalent) two-qubit gate, in terms of the minimum required pulse duration, the required pulse energy, or the difficulty of keeping the population in the logical subspace in a multi-level system. For example, a Hamiltonian containing among other terms an exchange interaction $(\hat{\sigma}_+\hat{\sigma}_- + \hat{\sigma}_-\hat{\sigma}_+)$ will likely find it easier to implement the $\sqrt{\text{iSWAP}}$ gate than the locally equivalent M-gate. The latter is generated by a term $(3\hat{\sigma}_x\hat{\sigma}_x + \hat{\sigma}_y\hat{\sigma}_y)$, cf. Table D.1, that would be part of the Lie algebra only indirectly. Likewise, the implementation of iSWAP would likely require more time than $\sqrt{\text{iSWAP}}$, even though both gates generate the same amount of entanglement.

Using an optimization functional such as Eq. (3.39) or Eq. (3.40) that targets a specific quantum gate is therefore unnecessarily restrictive. A far better approach is to formulate a more general optimization functional, and allow optimal control to select the specific gate that most effectively fits the objective. The geometric theory of local invariants has been combined with optimal control [238, 193], using a functional that minimizes the Cartesian distance to a target point in the Weyl chamber (c_1, c_2, c_3), see Eq. (2.46), respectively a set of local invariants (g_1, g_2, g_3), see Eq. (2.47).

This chapter, adapted from [239, 240], addresses the realization of perfect entanglers, i.e., quantum gates that are capable of transforming some separable states into maximally entangled states. This recognizes the fundamental role of entanglement, as the resource that two-qubit gates contribute to the gate model of universal quantum computing. Moreover, since nearly 85% of gates in SU(4) are perfect entanglers [241, 92], optimizing for an arbitrary perfect entangler gives a much broader target.

Extending the idea of optimization in the Weyl chamber, a functional for an arbitrary perfect entangler has been formulated [239]. As shown in Fig. 2.3, the perfect entanglers form a polyhedron bounded by the planes

$$c_1 + c_2 = \frac{\pi}{2}, \quad c_1 - c_2 = \frac{\pi}{2}, \quad c_2 + c_3 = \frac{\pi}{2}. \tag{6.1}$$

Thus, an optimization could simply minimize the distance to the nearest wall of that polyhedron.

Both the underlying local-invariants functional and the newly developed perfect entanglers functionals are reviewed in section 6.1. An optimization in the Weyl chamber, either towards a specific point or towards a general perfect entangler is most useful for quantum computing implementations that can generate a large number of gates, such as the superconducting circuits presented in chapter 5. In order to obtain some insight in the structure of the dynamics generated by a typical Hamiltonian, section 6.2 shows the controllability for an effective description of two qubits, truncated to two levels. Section 6.3 then applies the method to a specific example of transmon qubits, illustrating the power of an optimization towards a perfect entangler in comparison to the optimization towards a local equivalence class.

While the perfect entanglers functional is tested on a Hilbert space description, the ultimate motivation for using a more general functional is to counter decoherence. First, dissipation might affect the processes implementing different gates differently. In such a situation, a direct optimization might result in very different fidelities. Optimizing for a general perfect entanlger, on the other hand, would automatically yield the gate that is implementable with highest fidelity under dissipation. In a situation where dissipation cannot be circumvented, the only option is to perform the gate on a faster time scale than that of the decoherence. The equation of the quantum speed limit then becomes crucial. Since the speed limit may differ for various gates, the perfect entanglers functional allows to identify the optimal gate in such cases. Thus, even without going to the formalism of Liouville space, the possibility to implement *fast* gates provides an answer to the problem of implementing gates in an open quantum system. The message of this chapter therefore connects to that of chapter 4, in that the key to finding robust quantum gates with optimal control is to encode all relevant requirements, and *only* the relevant requirements in the optimization functional.

6.1 The Perfect Entanglers Functional

6.1.1 Formulation in c-space

The Cartan decomposition

$$\hat{\mathsf{U}} = \hat{\mathsf{k}}_1 \hat{\mathsf{A}}_U \hat{\mathsf{k}}_2 \,, \quad \hat{\mathsf{A}}_u = \exp\left[\frac{\mathrm{i}}{2}\sum_{j=1}^{3} c_j \hat{\sigma}_j \hat{\sigma}_j\right] \,, \tag{6.2}$$

cf. Eq. (2.46), allows to determine the “true” two-qubit part $\hat{\mathsf{A}}_U$ for an arbitrary two-qubit gate $\hat{\mathsf{U}}$. The “local” components $\hat{\mathsf{k}}_1, \hat{\mathsf{k}}_2$ can be written in terms of single-qubit gates

$$\hat{\mathsf{k}}_i = \hat{\mathsf{U}}_i^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathsf{U}}_i^{(2)} \,, \quad \hat{\mathsf{U}}_i^{(1,2)} \in \mathrm{SU}(2) \,, \tag{6.3}$$

whereas $\hat{\mathsf{A}}_U$ allows for no such decomposition. The entanglement power of $\hat{\mathsf{U}}$ is provided only by the non-local $\hat{\mathsf{A}}_U$.

Therefore, the optimization for a general perfect entangler starts from the prerequisite of eliminating all non-local operations from the figure of merit. Replacing the target gate $\hat{\mathsf{O}}$ and achieved gate $\hat{\mathsf{U}}$ in the functional (3.40) by only their non-local components $\hat{\mathsf{A}}_O$, $\hat{\mathsf{A}}_U$ yields an optimization functional $\varepsilon_{\text{lec}} = 1 - F_{\text{lec}}$ with

$$F_{\text{lec}} = \frac{1}{4}\mathfrak{Re}\left(\mathrm{tr}\left[\hat{\mathsf{A}}_O^\dagger \hat{\mathsf{A}}_U\right]\right) \,. \tag{6.4}$$

The functional ε_{lec} takes its minimum value of zero if and only if $\hat{\mathsf{O}}$ and $\hat{\mathsf{U}}$ have the same non-local component, as determined by the Weyl chamber coordinates (c_1, c_2, c_3). Rewriting Eq. (6.4) explicitly in these coordinates yields [239]

$$F_{\text{lec}} = \cos\frac{\Delta c_1}{2}\cos\frac{\Delta c_2}{2}\cos\frac{\Delta c_3}{2} \,, \qquad \Delta c_i \equiv c_{i,O} - c_{i,U} \,. \tag{6.5}$$

Thus, the optimization can be interpreted as minimizing the geometric distance $(\Delta c_1, \Delta c_2, \Delta c_3)$ to a target point in the Weyl chamber.

This may now be extended further to obtain a functional that optimizes for a general perfect entangler, by minimizing the distance not to a specific point in the Weyl chamber, but to the polyhedron of perfect entanglers. For any gate $\hat{\mathsf{U}}$, we identify the closest of the planes bounding the perfect entanglers, Eq. (6.1). Then, the distance of $(c_{1,U}, c_{2,U}, c_{3,U})$ and its projection onto that plane is minimized, resulting in [239]

$$F_{\text{PE}}(\hat{\mathsf{U}}) = \begin{cases} \cos^2 \frac{c_{U,1}+c_{U,2}-\frac{\pi}{2}}{4} & c_1 + c_2 \leq \frac{\pi}{2} \\ \cos^2 \frac{c_{U,2}+c_{U,3}-\frac{\pi}{2}}{4} & c_2 + c_3 \geq \frac{\pi}{2} \\ \cos^2 \frac{c_{U,1}-c_{U,2}-\frac{\pi}{2}}{4} & c_1 - c_2 \geq \frac{\pi}{2} \\ 1 & \text{otherwise.} \end{cases} \tag{6.6}$$

Both F_{lec} and F_{PE} take values in $[0, 1]$ and can thus be interpreted as fidelities.

Generally, the logical two-qubit subspace is embedded in a larger Hilbert space, such that while the dynamics in the total Hilbert space are unitary, the dynamics in the subspace may not be. In this case, a closest unitary $\hat{\mathsf{U}}$ can be derived from the non-unitary (projected) gate $\tilde{\mathsf{U}}$: If $\tilde{\mathsf{U}}$ has the singular value decomposition

$$\tilde{\mathsf{U}} = \hat{\mathsf{V}}\,\hat{\Sigma}\,\hat{\mathsf{W}}^{\dagger}\,, \tag{6.7}$$

then the closest unitary is

$$\hat{\mathsf{U}} = \arg\min_{u} \left\Vert \tilde{\mathsf{U}} - u \right\Vert = \hat{\mathsf{V}}\,\hat{\mathsf{W}}^{\dagger}\,. \tag{6.8}$$

The distance between $\hat{\mathsf{U}}$ and $\tilde{\mathsf{U}}$ is a measure of unitarity. The total local-equivalence-class and perfect-entangler fidelities then become

$$F_{\text{lec}}(\tilde{\mathsf{U}}) = F_{\text{lec}}(\hat{\mathsf{U}}) - \left\Vert \tilde{\mathsf{U}} - \hat{\mathsf{U}} \right\Vert\,, \tag{6.9}$$

$$F_{\text{PE}}(\tilde{\mathsf{U}}) = F_{\text{PE}}(\hat{\mathsf{U}}) - \left\Vert \tilde{\mathsf{U}} - \hat{\mathsf{U}} \right\Vert\,. \tag{6.10}$$

The optimization goal is to find that $\tilde{\mathsf{U}} = \hat{\mathsf{U}}$ which minimizes $1 - F_{\text{lec}}$ or $1 - F_{\text{PE}}$.

6.1.2 Formulation in g-space

The formulation of the local equivalence class and perfect entanglers functionals in terms of the Weyl chamber coordinates (c_1, c_2, c_3), Eq. (6.9) and Eq. (6.10), have the disadvantage that there is no way to evaluate them

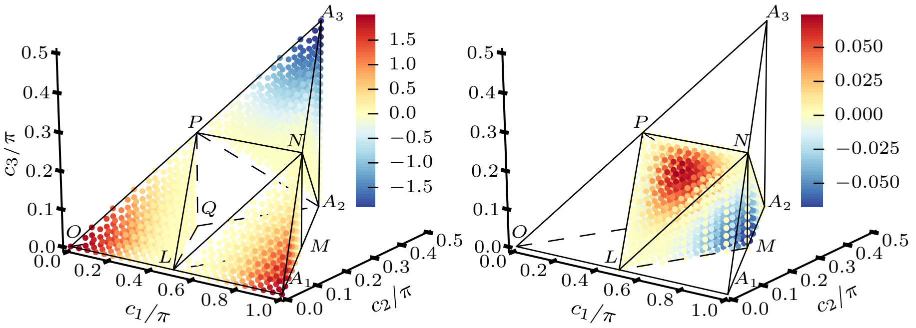


Figure 6.1: Value for the perfect-entanglers J_{PE} as defined in Eq. (6.12) for sampling points in the Weyl chamber. On the left, values for points in the regions W_0, W_0^*, and W_1 (outside of the perfect entanglers polyhedron). The values are in the range $-2 \leq J_{\mathrm{PE}} \leq 2$. On the right, values $-0.07 \lesssim J_{\mathrm{PE}} \lesssim 0.07$ inside the perfect-entanglers polyhedron. The functional takes slightly positive values in the center top region, and slightly negative values near the outer bottom regions, behind both the M point and the Q point (not visible).

analytically for a given gate $\hat{\mathsf{U}}$. This is because calculation of the Weyl coordinates themselves relies on numerical diagonalization of $\hat{\mathsf{U}}$, as discussed in chapter 2. Therefore, use of the functionals F_{lec} and F_{PE} is restricted to optimization methods that are not gradient-based, such as the CRAB algorithm presented in section 3.3.2 of chapter 3. An application of the functional using this algorithm can be found in Ref. [240].

For the other optimal control approaches discussed in chapter 3, the gradient of the optimization functional needs to be evaluated. We must therefore use an equivalent functional, based not on the Weyl space coordinates (c_1, c_2, c_3), but on the local invariants (g_1, g_2, g_3), see Eq. (2.48), which depend analytically on $\hat{\mathsf{U}}$ and therefore can be differentiated either with respect to the field (GRAPE) or with respect to the states (Krotov).

For the local invariants, an appropriate functional can be postulated as [238, 193]

$$J_{\mathrm{LI}}(\hat{\mathsf{U}}) = (\Delta g_1)^2 + (\Delta g_2)^2 + (\Delta g_3)^2 \,, \tag{6.11}$$

where Δg_i is the Euclidean distance between local invariant g_i of the obtained unitary $\hat{\mathsf{U}}$ and the optimal gate $\hat{\mathsf{O}}$. Eq. (6.11) takes its minimum value zero if and only if the gates $\hat{\mathsf{U}}$ and $\hat{\mathsf{O}}$ are identical up to local transformations. This makes it slightly more general than Eq. (6.5), as there are some gates that are locally equivalent but have different Weyl-chamber coordinates, such as the points Q and M in Fig. 2.3.

In order to derive a functional based on the local invariants (g_1, g_2, g_3) for an arbitrary perfect entangler, Eq. (2.47) is used to rewrite Eq. (6.6), resulting in [239]

$$J_{\mathrm{PE}}(\hat{\mathsf{U}}) = g_3\sqrt{g_1^2 + g_2^2} - g_1 \,. \tag{6.12}$$

Eq. (6.12) takes the value zero at the boundary to the perfect entanglers. The value of J_{PE} for gates throughout the Weyl chamber is shown in Fig. 6.1. It is important to note that J_{PE} can also take negative values. Typically, an optimization will start from a guess pulse that generates little entanglement, locally equivalent to a gate close to the identity. In the Weyl chamber, these are defined as the regions W_0 (delimited by O, L, Q, and P) and W_0^* (delimited by A_1, L, M, and N). In these regions, the functional goes smoothly from its maximum value of 2 to 0 on the boundary of the perfect entanglers. The sign of the functional is reversed if the optimization were to start from near the SWAP gate at A_3. This region is denoted as W_2, delimited by A_3, A_2, P, N. Typically, this is not a problem. Thinking of entanglement as a resource that is generated by an interaction acting over a certain period of time, continuing to apply the same interaction after the entanglement has reached its maximum value of 1 will now disentangle the entangled states. For example, a

$$\hat{\sigma}_x \hat{\sigma}_x + \hat{\sigma}_y \hat{\sigma}_y + \hat{\sigma}_z \hat{\sigma}_z$$

interaction will generate gates along the $O - P - A_3$ line (see appendix D), building up entanglement as it goes towards the $\sqrt{\text{SWAP}}$ perfect entangler and then removing it again until the non-entangling SWAP gate is reached. Thus, guess pulses would be expected to start in W_2 only for extremely simple systems; when this occurs, the sign of the functional must be reversed.

Inside the polyhedron of perfect entanglers, J_{PE} can also take non-zero values; it is zero on the planes L–N–A_2 and L–P–A_2, as well as on all boundaries of the polyhedron. Thus, if an optimization is allowed to continue after passing through the perfect-entanglers boundary, it will be pushed towards the center of the sub-polyhedron delimited by the point L, M, A_2, and N on the right; or L, Q, A_2 and P on the left.

Like before, we must take into account non-unitarity due to projection onto the logical subspace. Just as for the value of the functional, F_{lec} and F_{PE}, the expression $\left\Vert \tilde{\mathsf{U}} - \hat{\mathsf{U}} \right\Vert$ used to evaluate unitarity in section 6.1.1 cannot easily be differentiated. As an alternative, we minimize the loss of population from the logical subspace,

$$p_{\text{loss}} = \frac{1}{4} \operatorname{tr} \left[\tilde{\mathsf{U}}^\dagger \tilde{\mathsf{U}} \right] \tag{6.13}$$

as an alternative measure of unitarity.

In total, the optimization functional for the local invariants and perfect entanglers then reads

$$J_{\mathrm{LI}}(\tilde{\mathsf{U}}) = w J_{\mathrm{LI}}(\hat{\mathsf{U}}) + (w-1)\left(1 - \frac{1}{4}\operatorname{tr}\left[\tilde{\mathsf{U}}^\dagger \tilde{\mathsf{U}}\right]\right)\,, \tag{6.14}$$

$$J_{\mathrm{PE}}(\tilde{\mathsf{U}}) = w J_{\mathrm{PE}}(\hat{\mathsf{U}}) + (w-1)\left(1 - \frac{1}{4}\operatorname{tr}\left[\tilde{\mathsf{U}}^\dagger \tilde{\mathsf{U}}\right]\right)\,. \tag{6.15}$$

In order to weight the relative importance of the Weyl-chamber optimization and the unitarity, the factor $w \in [0,1]$ is used. This factor can be adaptively changed during the optimization in order to improve convergence.

6.2 Controllability of Superconducting Qubits

Optimization towards an arbitrary perfect entangler is most meaningful if the system dynamics allows the polyhedron of perfect entanglers to be approached from more than one direction or, more generally, for optimization paths in the Weyl chamber that explore more than one dimension. We therefore investigate the corresponding requirements on a generic two-qubit Hamiltonian,

$$\begin{aligned}\hat{\mathsf{H}}[u_1(t), u_2(t)] = \sum_{\alpha=1,2} \frac{\omega_\alpha}{2}\hat{\sigma}_z^{(\alpha)} + u_1(t)\left(\hat{\sigma}_x^{(1)} + \lambda\hat{\sigma}_x^{(2)}\right) + \\ + u_2(t)\left(\hat{\sigma}_x^{(1)}\hat{\sigma}_x^{(2)} + \hat{\sigma}_y^{(1)}\hat{\sigma}_y^{(2)}\right)\,.\end{aligned} \tag{6.16}$$

Here, $\hat{\sigma}_i^{(\alpha)}$ is the i'th Pauli operator acting on the α'th qubit of transition frequency ω_α, $u_1(t)$ the single-qubit control field, where λ describes how strongly $u_1(t)$ couples to the second qubit relative to the first one, and $u_2(t)$ is the two-qubit interaction control field.

The Hamiltonian in Eq. (6.16) is of the form typical for an effective description of superconducting qubits, truncated to two levels. In principle, truncation of the Hamiltonian can have significant influence on controllability. For example, a two-level system is fully controllable, whereas and infinite harmonic oscillator is not. For superconducting qubits, however, we do not expect controllability to be enhanced by truncation. Therefore, an analysis still provides valuable insight into how the Hamiltonian acts in the Weyl chamber.

We analyze the solutions to the differential equation

$$\frac{\partial}{\partial t}\hat{\mathsf{U}}(t) = -\mathrm{i}\hat{\mathsf{H}}\left[u(t)\right]\hat{\mathsf{U}}\left(t\right)\,, \quad \hat{\mathsf{U}}(0) = \mathbb{1}\,, \tag{6.17}$$

for the unitary transformations $\hat{\mathsf{U}}$ generated by the Hamiltonian (6.16). The reachable set of unitary transformations for a Hamiltonian is given in terms of the corresponding dynamical Lie algebra, see section 2.3.3 in chapter 2.

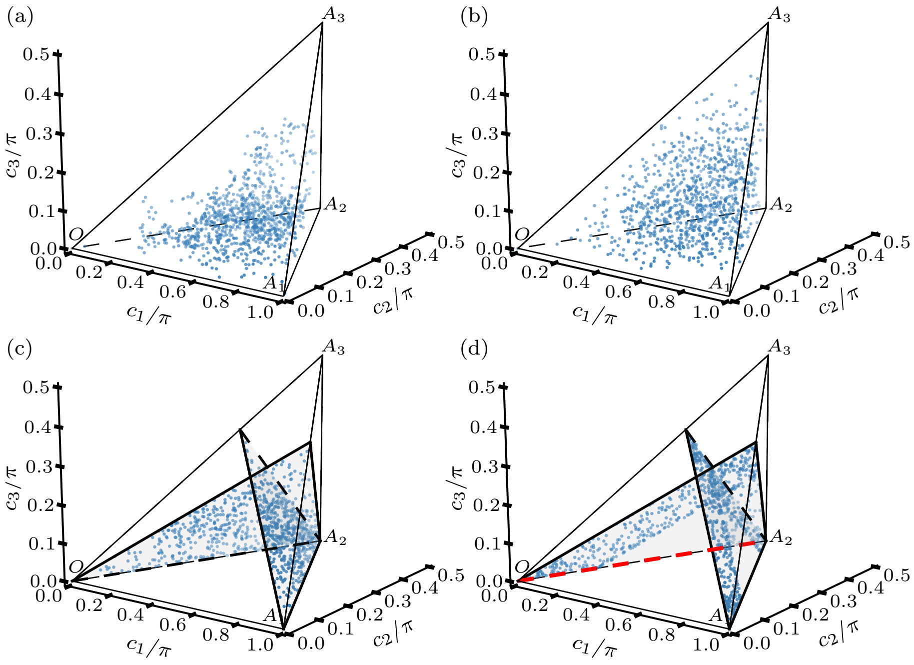


Figure 6.2: Sampling of reachable points in the Weyl chamber, obtained by solving Eq. (6.17) for the Hamiltonian (6.16) with $\lambda = 1$. In panel (a), $u_1(t)$ is a random pulse $\in [0, 1]$, $u_2(t) \equiv 10^{-3}$, and $\omega_1 = 1.0 \neq \omega_2 = 1.1$. In panel (b), in addition $\omega_1 = \omega_2 = 1.0$. In panel (c), $u_1(t)$ and $u_2(t)$ are both independently random $\in [0, 1]$, and $\omega_1 = \omega_2 = 0$. In panel (d), in addition $u_1(t) \equiv u_2(t)$. The red line in panel (d) is obtained for $u_1(t) \equiv 0$ and $u_2(t)$ random $\in [0, 1]$ ($\omega_1 = \omega_2 = 0$).

It can be generated by taking the terms in (6.16) as a basis (neglecting orthonormalization for simplicity),

$$\hat{\sigma}_z^{(1)}\,,\ \hat{\sigma}_z^{(2)}\,,\ \hat{\sigma}_x^{(1)} + \lambda\hat{\sigma}_x^{(2)}\,,\ \hat{\sigma}_x^{(1)}\hat{\sigma}_x^{(2)} + \hat{\sigma}_y^{(1)}\hat{\sigma}_y^{(2)}\,,$$

and constructing the repeated Lie brackets of these operators. This quickly yields all 15 canonical basis operators of $SU(4)$, consisting of the single-qubit operators $\hat{\sigma}_x^{(1)}$, $\hat{\sigma}_x^{(2)}$, $\hat{\sigma}_y^{(1)}$, $\hat{\sigma}_y^{(2)}$, $\hat{\sigma}_z^{(1)}$, and $\hat{\sigma}_z^{(2)}$, as well as the entangling operators $\hat{\sigma}_x^{(1)}\hat{\sigma}_y^{(2)}$, $\hat{\sigma}_y^{(1)}\hat{\sigma}_x^{(2)}$, $\hat{\sigma}_y^{(1)}\hat{\sigma}_z^{(2)}$, $\hat{\sigma}_z^{(1)}\hat{\sigma}_y^{(2)}$, $\hat{\sigma}_x^{(1)}\hat{\sigma}_z^{(2)}$, $\hat{\sigma}_z^{(1)}\hat{\sigma}_x^{(2)}$, $\hat{\sigma}_x^{(1)}\hat{\sigma}_x^{(2)}$, $\hat{\sigma}_y^{(1)}\hat{\sigma}_y^{(2)}$, and $\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}$. Hence the system is completely controllable, and any point in the Weyl chamber can be reached.

The complete controllability can be verified numerically, by solving Eq. (6.17) for a random sequence of pulse values. The gates obtained within 1000 steps are shown in panel (a) of Fig. 6.2, and demonstrate full controllability, since there are points in all regions of the Weyl chamber. Continuing the procedure to infinity would eventually fill the entire chamber. Neither setting $u_2(t)$ constant nor choosing $\lambda = 0$ places any restrictions on the controllability – indeed it is sufficient if either the single qubit terms or the interaction term is controllable. While the controllability in this example was analyzed for arbitrary values of the parameters, the form of the Hamiltonian and the ratio between $\omega_{1,2}$ and u_2 fits the description of superconducting transmon qubits, with qubit energies in the GHz range and static qubit-qubit-coupling in the MHz range.

Introducing symmetries in the Hamiltonian (6.16) reduces the controllability. First, we consider a situation in which the two qubits operate at the same frequency $\omega_1 = \omega_2$. In this case, the dynamic Lie algebra consists of only 9 instead of 15 operators. Consequently, not every two-qubit gate can be implemented. However, the nine operators include $\hat{\sigma}_x^{(1)}\hat{\sigma}_x^{(2)}$, $\hat{\sigma}_y^{(1)}\hat{\sigma}_y^{(2)}$, $\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}$, which are sufficient to reach every point in the Weyl chamber, cf. Eq. (2.46). This is illustrated in panel (b) of Fig. 6.2. Despite the reduced controllability, the Weyl chamber is more evenly filled after the same 1000 propagation steps as in panel (a). This counterintuitive finding is due to the lower dimension of the random walk, with no resources being "wasted" on the missing six single-qubit directions.

The set of gates that can be implemented with Hamiltonian (6.16) is more severely restricted if both qubits are completely degenerate, $\omega_1 = \omega_2 = 0$. This is typical for superconducting charge qubits operated at the "charge degeneracy point". Without any drift term, the Lie algebra consists of only four generators, $\hat{\sigma}_z^{(1)}\hat{\sigma}_y^{(2)} + \hat{\sigma}_y^{(1)}\hat{\sigma}_z^{(2)}$ and $\hat{\sigma}_y^{(1)}\hat{\sigma}_y^{(2)} - \hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}$ in addition to the two original terms. The implications for controllability in the Weyl chamber are not immediately obvious since three generators can be sufficient to obtain full Weyl chamber controllability. The easiest approach is to perform a numerical analysis, the results of which are shown in panel (c) of Fig. 6.2. Two independent randomized pulses $u_1(t)$ and $u_2(t)$ were used. The reachable points lie on a plane, which due to the reflection symmetries appears as two triangular branches, indicated by the shaded triangles, O–$(\frac{2\pi}{3}, \frac{\pi}{3}, \frac{\pi}{3})$–$A_2$ and A_1–$(\frac{\pi}{3}, \frac{\pi}{3}, \frac{\pi}{3})$–$A_2$. Note that almost none of the common two-qubit gates are included in this set.

If only a single pulse is available to drive both the single-qubit and two-qubit terms, $u_1(t) \equiv u_2(t)$, and the qubits are degenerate, $\omega_1 = \omega_2 = 0$, there is a single generator for the dynamics. This situation is shown in panel (d) of Fig. 6.2. Although there is only a single generator for the dynamics, a two-dimensional subset of the Weyl chamber can be reached. However, the subset is no longer the full plane as it is for two independent pulses, panel (c). Without single-qubit control, the center of the plane is not longer reachable. It is important to remember that while a single generator yields points on a line in the Weyl chamber (not necessarily a straight one), it can still fill an arbitrary subset of the Weyl chamber, due to reflections at the boundaries. A similar example, restricted to the ground plane of the Weyl chamber, has been analyzed in Ref. [91].

left qubit frequency	ω_1	4.380	GHz
right qubit frequency	ω_2	4.614	GHz
left qubit anharmonicity	α_1	-210	MHz
right qubit anharmonicity	α_2	-215	MHz
effective qubit-qubit coupling	J^{eff}	-3.0	MHz
relative coupling strength	λ	1.03	

Table 6.1: Parameters for the transmon Hamiltonian Eq. (6.19)

.

Lastly, if there is no control over the individual qubits at all, $u_1(t) \equiv 0$, the only remaining generator is $\hat{\sigma}_x^{(1)}\hat{\sigma}_x^{(2)} + \hat{\sigma}_y^{(1)}\hat{\sigma}_y^{(2)}$. This corresponds to the straight line O–A_2 in the Weyl chamber, shown in red in panel (d) of Fig. 6.2. The line is reflected back onto itself at the A_2 point. Thus, in this case only a truly one-dimensional subset of reachable gates in the Weyl chamber can be realized.

For a Hamiltonian that allows for a one-dimensional search-space only, optimal control calculations with a functional targeting all perfect entanglers will not yield results better than direct gate optimization. In the above example, the perfect entangler that can be generated with the shortest gate duration is the $\sqrt{\text{iSWAP}}$ at the point Q.

In contrast, for Hamiltonians allowing for two or three search directions in the Weyl chamber, especially panels (a) and (b) in Fig. 6.2, the polyhedron of perfect entanglers may be approached from several different angles. Optimization with a functional targeting all perfect entanglers is then non-trivial. In such a search, the optimized solution will depend on additional constraints in the functional and the initial guess field.

6.3 Optimization of Transmon Qubits

6.3.1 Model

Having verified that the Hamiltonian in Eq. (6.16) allows for full controllability, we now extend the discussion to a more realistic description of two transmons, i.e., anharmonic multi-level systems whose interaction is mediated by a cavity. This is described by a generalized Jaynes-Cummings Hamiltonian, see chapter 5. The energy of each transmon qubit transition given by ω_1, ω_2 for the first (“left”) and second (“right”) transmon respectively. Following a Duffing oscillator model, higher levels show an anharmonicity of α_1, respectively α_2. Each qubit couples to the cavity with coupling strength g_1, g_2. In the dispersive limit $|\omega_i - \omega_r| \gg |gi|$, $i = 1, 2$, with ω_r the cavity frequency, the cavity can be eliminated and an effective two-transmon Hamiltonian is obtained. The coupling between each transmon and the cavity turns into an effective qubit-qubit coupling

$$J^{\text{eff}} \approx \frac{g_1 g_2}{(\omega_1 - \omega_r)} + \frac{g_1 g_2}{(\omega_2 - \omega_r)} \,. \tag{6.18}$$

In most current setups, $J^{\text{eff}} \ll |\omega_2 - \omega_1|$, and the two-transmon Hamiltonian can be approximated as [222]

$$\begin{aligned} \hat{\mathsf{H}}_{2T} \approx & \sum_{i=1,2} \left(\left(\omega_i + \frac{\alpha_i}{2} \right) \hat{\mathsf{b}}_i^\dagger \hat{\mathsf{b}}_i - \frac{\alpha_i}{2} \left(\hat{\mathsf{b}}_i^\dagger \hat{\mathsf{b}}_i \right)^2 \right) + \\ & + J^{\text{eff}} \left(\hat{\mathsf{b}}_1^\dagger \hat{\mathsf{b}}_2 + \hat{\mathsf{b}}_1 \hat{\mathsf{b}}_2^\dagger \right) + \Omega(t) \left(\hat{\mathsf{b}}_1 + \hat{\mathsf{b}}_1^\dagger + \lambda \hat{\mathsf{b}}_2 + \lambda \hat{\mathsf{b}}_2^\dagger \right) \,, \end{aligned} \tag{6.19}$$

cf. Eq. (5.30), where $\Omega(t)$ is the driving field in the rotating wave approximation, see Eq. (5.39). Typical parameters are given in Table 6.1. A two-level truncation of this Hamiltonian corresponds to Eq. (6.16).

6.3.2 Krotov's Method

The optimization is performed using Krotov's method, section 3.3.4 of chapter 3, for both the local invariants functional (6.14) and the perfect entanglers functional (6.15). No dissipation is taken into account. The basis states for the forward propagation are the Bell states

$$|\phi_1\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle - \mathrm{i}|11\rangle \right) \,, \qquad |\phi_2\rangle = -\frac{1}{\sqrt{2}} \left(\mathrm{i}|01\rangle - |10\rangle \right) \,, \tag{6.20}$$

$$|\phi_3\rangle = -\frac{1}{\sqrt{2}} \left(\mathrm{i}|01\rangle + |10\rangle \right) \,, \qquad |\phi_4\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle + \mathrm{i}|11\rangle \right) \,, \tag{6.21}$$

as required for the calculation of the local invariants, cf. Eq.(2.51). The equations of motion are given by the standard Schrödinger equation, see Eq. (3.54) and Eq. (3.55).

The boundary conditions for the backward propagation,

$$\left| \chi_k^{(i)}(T) \right\rangle = - \left. \frac{\partial J_T}{\partial \langle \phi_k |} \right|_{\phi_k^{(i)}(T)} \,, \tag{6.22}$$

cf. Eq. (3.56), are straightforward to calculate, although tedious. For J_{LI}, they can be found in Ref.[238]. For J_{PE}, they are listed in appendix G.

Both J_{LI} and J_{PE} are highly non-convex functionals. Therefore, they require the second order contribution in the Krotov update equation. With

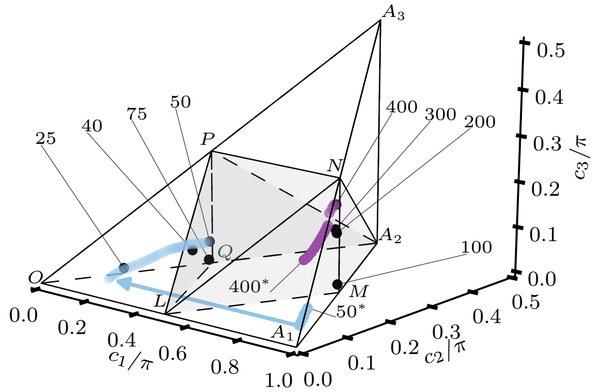


Figure 6.3: Optimized gates in the Weyl chamber, for two transmon qubits, optimized with Krotov's method for the perfect-entangler (PE) functional in Eq. (6.15). The point at which each optimization enters the PE polyhedron, or the end point of the optimization if no PE can be obtained, is shown by a black dot and labeled with the gate duration. The entire optimization paths for $T = 50\,\text{ns}$ and $T = 400\,\text{ns}$ are shown in light blue and dark purple, respectively, with the starting points labeled by 50^* and 400^*.

the standard constraint (3.61), this update equation is

$$\begin{aligned}\Delta\Omega(t) = \frac{S(t)}{\lambda_a}\mathfrak{Im}\Bigg[&\sum_{k=1}^{N}\left\langle \chi_k^{(i)}(t)\middle|\left(\frac{\partial\hat{\mathsf{H}}}{\partial\Omega}\bigg|_{\substack{\phi^{(i+1)}(t)\\ \Omega^{(i+1)}(t)}}\right)\middle|\phi_k^{(i+1)}(t)\right\rangle + \\ &+\frac{1}{2}\sigma(t)\left\langle \Delta\phi_k(t)\middle|\left(\frac{\partial\hat{\mathsf{H}}}{\partial\Omega}\bigg|_{\substack{\phi^{(i+1)}(t)\\ \Omega^{(i+1)}(t)}}\right)\middle|\phi_k^{(i+1)}(t)\right\rangle\Bigg]\,,\end{aligned} \tag{6.23}$$

cf. Eq. (3.52), with $\sigma(t)$ given by Eq. (3.57). The optimization is carried out for different gate durations between 25 ns and 400 ns, using a sine-squared pulse of 35 MHz peak amplitude as the guess pulse $\Omega^{(0)}(t)$.

6.3.3 Optimization Results

Figure 6.3 shows the results of the optimization in the Weyl chamber. The point at which each optimization enters the perfect entanglers polyhedron is indicated by a black dot and labeled with the gate duration. For $T < 50\,\text{ns}$, no perfect entangler can be reached – defining heuristically the quantum speed limit (QSL) for this transformation. In order to illustrate how the optimization proceeds, the optimization paths for $T = 50\,\text{ns}$, i.e., the gate at the QSL, and a high-fidelity gate ($T = 400\,\text{ns}$) are traced in light blue

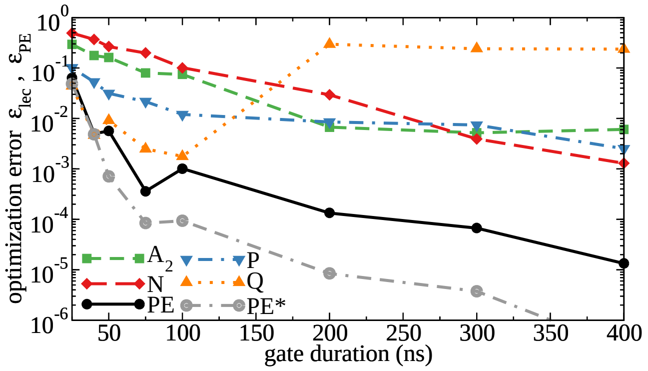


Figure 6.4: Comparison of optimization success for the PE functional compared to the local invariants (LI) functional for several points in the Weyl chamber. The optimization success using Krotov's method is measured in c-space, although the optimization functionals are defined in g-space (see text for details). For the LI-optimization, the results are fully converged. For the PE-optimization, the results are converged to a relative change below 10^{-2} (black solid curve) and 10^{-3} (gray dash-dash-dotted curve).

and dark purple, respectively. Both optimizations start in the W_0^* region (near the A_1 point). The gate obtained with the guess pulse for $T = 50\,$ns is significantly farther away from the surface of the polyhedron of PE than that for the guess pulse with $T = 400\,$ns. Optimization for $T = 400\,$ns therefore moves directly towards the W_0^* surface of the PE polyhedron, whereas the optimization for $T = 50$ ns enters the ground plane and emerges in the W_0 region, before finally reaching the W_0 surface of the polyhedron of perfect entanglers. The jump from W_0^* to W_0 is indicated by the light blue arrow. We find the optimization to enter W_0 from W_0^* for durations $< 100\,$ns, whereas for longer gate duration the optimizations stay within W_0^* entirely. The different optimization paths are a result of the competition between the two objectives – to reach a perfect entangler, and to implement a gate that is unitary in the logical subspace (the points shown in Fig. 6.3 are the Weyl chamber coordinates of the unitary U closest to the actual time evolution $\tilde{U}$). The latter objective is more difficult to realize for shorter gate durations, resulting in a more indirect approach to the polyhedron of perfect entanglers than one might expect when considering that objective alone.

It is instructive to compare the optimization success of the perfect entangler functional, Eq. (6.15) with the optimization under the local-invariants functional, Eq. (6.14) for a few select points of the Weyl chamber. This is shown in Fig. 6.4. While the optimization was driven by the g-space formulation of the functionals, the fidelities F_{lec} and F_{PE} define a more intuitive figure of merit for the analysis. As defined in Eq. (6.9) and Eq. (6.10), the fidelity is reduced by non-unitarity.

The results of Fig. 6.4 show how for different gate durations, different gates are easiest to reach. In agreement with the results of Fig. 6.3, for durations < 50 ns, the jump in the optimization error indicates a speed limit. For short gate durations, 50 ns $\leq T \leq$ 100 ns, optimization towards the point Q in the Weyl chamber is most successful. This matches the optimized gates for $T \leq 100\,\text{ns}$ in Fig. 6.3 being near the Q point. Also correspondingly, the longer gate durations end near the N point. The failure to reach the point Q for longer durations is due to the symmetry structure of the Weyl chamber. Namely, for the ground plane of the chamber, there is a mirror axis defined by the line through L and A_2, where mirrored points are in the same local equivalence class. Both the Q-point and the M point have local invariants of $g_1 = \frac{1}{4}, g_2 = 0, g_3 = 1$. Since the optimization was performed in g-space, these two points are not distinguishable; indeed, for long gate durations, the Q-optimization successfully reached the M point.

In comparison with the local invariants optimization, the perfect entanglers functional shows excellent performance. It automatically identifies the optimal gate for a given gate duration and reaches significantly better fidelities. This is due to the fact that the desired entangling power of U can usually be obtained in just a few tens of iterations of the algorithm, and the remainder of the optimization then focuses on improving the unitarity of the obtained gate $\tilde{U}$. Most strikingly, we find that for the optimization towards a specific local equivalence class, the convergence rate becomes extremely small as the optimum is approached. All the results shown in Fig. 6.4 are converged to a relative change below 10^{-4}, such that no measurable improvement can be expected within a reasonable number of iterations. While in principle (due to the full controllability of the system), the direct optimizations should yield arbitrarily small gate errors, as long as the gate duration is above the quantum speed limit, in practice this depends on numerical parameters such as the weight λ_a in Krotov's method and may take a extremely large number of iterations or stagnate, as we observe here. The perfect entangler optimization shows remarkable robustness with respect to this issue. We observed very little slow-down in convergence. The black curve in Fig. 6.4 for the PE-optimization already yields a significantly smaller optimization error than any of the LI-optimizations, but is only converged to a relative change of 10^{-2}. Even the gray dash-dash-dotted curve, labeled PE*, is only converged to a relative change of 10^{-3}, and thus the optimization would still yield considerably better results if it were to be continued.

The values of the optimization error in Fig. 6.4 of 10^{-3} or 10^{-2} should not be understood to indicate a gate error above the quantum error correction threshold. Whereas the optimization error relates only to a figure of merit used for optimization, the relevant physical quantity that would be

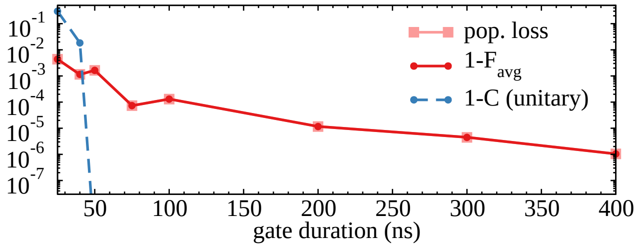


Figure 6.5: Analysis of the sources of error for PE of 2 transmon qubits: Population loss from the logical subspace (light red squares), concurrence error of the closest unitary gate Û in the logical subspace (blue circles), and average gate error, $\varepsilon_{avg} = 1 - F_{\text{avg}}$, with which Û is implemented (red circles).

determined in an experiment is the average gate fidelity, Eq. (2.88).

Figure 6.5 shows the generated entanglement as measured by the concurrence and the average gate error, $\varepsilon_{avg} = 1 - F_{avg}$, together with the population loss from the logical subspace. Ô is taken to be the unitary that is closest to the projection of the realized operation from the full Hilbert space onto the logical subspace. For $T > 50\,\text{ns}$, the gate errors are at or below 10^{-4}. For shorter gate durations, insufficient entanglement is generated, cf. blue dashed curve in Fig. 6.5. Once T is sufficiently large to generate the desired entanglement, the only source of error is loss of population from the logical subspace, shown in light red in Fig. 6.5. This loss does not depend on the choice of the weight w in Eq. (6.15). When the gate duration is increased, optimization yields gates that are exponentially more unitary, as indicated by the linear decrease of the average gate error in our semi-log plot. The difficulty to ensure unitarity on the logical subspace is typical for weakly anharmonic ladders, as found in superconducting transmon or phase qubits. Optimal control can be successfully employed to tackle the problem of ensuring unitarity in the logical subspace, in addition to generating entanglement.

7

Efficient Optimization of Unitaries in Liouville Space

The preceding chapters 4 and 6 have illustrated the power of optimal control theory (OCT) for the robust implementation of quantum gates. Both dissipation and classical noise can be taken into account explicitly in the functional, allowing OCT to determine the pathways that result in the best possible fidelities under the present physical constraints. Using more general functionals to give the optimization more freedom can greatly enhance the chances of success.

Many implementations of quantum computing show continuing advances in maintaining coherence. For superconducting qubits, coherence times have increased from 2 ns for the first Cooper pair boxes [209] to ~0.1 ms for 3D transmons [242], an improvement over five orders of magnitude. In light of this, it may be possible to forgo a treatment in Liouville space, using OCT together with the perfect-entanglers functional to bring two-qubit gate durations to the quantum speed limit. Ideally, the quantum gate can be realized on a time scale where decoherence is not relevant. However, as long as at least some channels of dissipation act on the time scale of the gate implementation, a realistic description will generally require to modeling the dynamics in Liouville space. Moreover, in a non-Markovian setting, it may be possible to exploit strong interactions with parts of the environment to aid the implementation of a quantum gate [243].

Modeling quantum dynamics in Liouville space raises the question of numerical efficiency. Compared to Hilbert space, where states are described as a complex vector of dimension N, the corresponding density matrix has dimension N^2. Moreover, optimization for a quantum gate using the standard functionals presented in chapter 3 requires to propagate a full basis of the logical subspace of dimension d. Again, this scales quadratically in Liouville space. For gradient-based optimization schemes such as GRAPE or Krotov's method, it is necessary to store the entire time evolution of each required state, see Fig. 3.1. This can quickly push the boundaries of available numerical resources. For example, optimizing a two-qubit gate, $d = 4$, on a Hilbert space of modest dimension $N = 500$ (e.g., two transmon qubits with 5 levels coupled to a cavity of 20 levels) requires 4 megabyte storage for a single state. For a time discretization of only 1000 steps, at least 16000 states (1000×16 matrices spanning the logical Liouville subspace) must be kept in memory, requiring 64 gigabytes of memory.

This chapter, adapted from Ref. [196], illustrates that it is *not* necessary to propagate a full basis of Liouville space when optimizing for a unitary operation. This insight builds on results from quantum tomography [77, 78, 244] concerning the minimum resources required to characterize the implementation of quantum gates. In essence, the direct extension of functionals such as Eq. (3.39) and Eq. (3.40) from Hilbert to Liouville space [66, 245, 246] overlooks the fact that in quantum gate optimization, the target is a unitary operation and not a general dynamical map. Thus, the gate fidelity can be determined without a full reconstruction of the dynamical map, which would indeed require a basis that spans the full Liouville space.

In general, propagation of only three states is sufficient to optimize a quantum gate. Mathematically, this is based on the observation that only two states are necessary to distinguish any two unitaries, irrespective of Hilbert space dimension [77]. A third state is required to enforce that the dynamical map on the optimization subspace is contracting and population conserving.

If the system Hamiltonian only allows for the implementation of diagonal gates, the number of required states reduces to two; this is demonstrated for the example of a Rydberg gate, using the same model as in chapter 4. The general case of non-diagonal quantum gates is illustrated by a $\sqrt{\text{iSWAP}}$ gate on transmon qubits, using the same model as in chapter 6. While the use of three, respectively two states is sufficient in principle, the rate of convergence can be improved by extending the number of states. This corresponds to the observation that while two states represent the minimal set of states required to distinguish any two unitaries, they do not allow to deduce bounds on the gate error [77]. A set of d or $2d$ allows to determine a numerical and analytical bound, respectively. Therefore, depending on the desired gate error, propagation of either $d + 1$ (including the extra state required to enforce unitarity) or $2d$ states is the numerically most efficient choice. Already for a two-qubit gate, this represents a significant reduction in the number of states that need to be propagated, from 16 for the full Liouville space basis to 8 and 5, respectively.

7.1 A Minimal Set of States for the Optimal Control of Unitaries

7.1.1 Unitary Operations in Liouville Space

When optimizing a quantum gate $\hat{\mathsf{O}}$, we require that despite the presence of dissipation, the evolution of a state $\hat{\rho}$ is unitary,

$$\hat{\rho}(T) = \mathcal{E}\hat{\rho}(t=0) \equiv \hat{\mathsf{U}}\hat{\rho}(0)\hat{\mathsf{U}}^{\dagger}\,, \tag{7.1}$$

or at least as unitary as possible. In a second step, we then require that the obtained gate $\hat{\mathsf{U}}$ is identical to the desired gate $\hat{\mathsf{O}}$. Intuitively, Eq. (7.1) is fulfilled if any pure state remains pure under propagation. If we assume that the dynamics are indeed unitary, we may the ask how to determine the implemented gate U by propagating a well-chosen set of states. If we know U to be diagonal,

$$\hat{\mathsf{U}} = \operatorname{diag}\left(\mathrm{e}^{\mathrm{i}\phi_{00}}, \mathrm{e}^{\mathrm{i}\phi_{01}}, \mathrm{e}^{\mathrm{i}\phi_{10}}, \mathrm{e}^{\mathrm{i}\phi_{11}}\right)\,, \tag{7.2}$$

this becomes quite easy, as we only need to determine the four phases to specify the gate. Propagation of the state $\hat{\rho}_{\mathrm{sp}}$ corresponding to the coherent superposition of all d logical basis states,

$$\left(\hat{\rho}_{\mathrm{sp}}(0)\right)_{ij} = \frac{1}{d}\,, \tag{7.3}$$

yields

$$\hat{\mathsf{U}}\hat{\rho}_{\mathrm{sp}}\hat{\mathsf{U}}^{\dagger} = \begin{pmatrix} 1 & \mathrm{e}^{\mathrm{i}(\phi_1-\phi_2)} & \mathrm{e}^{\mathrm{i}(\phi_1-\phi_3)} & \mathrm{e}^{\mathrm{i}(\phi_1-\phi_4)} \\ \mathrm{e}^{\mathrm{i}(\phi_2-\phi_1)} & 1 & \mathrm{e}^{\mathrm{i}(\phi_2-\phi_3)} & \mathrm{e}^{\mathrm{i}(\phi_2-\phi_4)} \\ \mathrm{e}^{\mathrm{i}(\phi_3-\phi_1)} & \mathrm{e}^{\mathrm{i}(\phi_3-\phi_2)} & 1 & \mathrm{e}^{\mathrm{i}(\phi_3-\phi_4)} \\ \mathrm{e}^{\mathrm{i}(\phi_4-\phi_1)} & \mathrm{e}^{\mathrm{i}(\phi_4-\phi_2)} & \mathrm{e}^{\mathrm{i}(\phi_4-\phi_3)} & 1 \end{pmatrix}. \tag{7.4}$$

We may choose $\phi_1 = 0$, which corresponds to a specific global phase; all other phases can then directly be read off from the entries of Eq. (7.4). When $\hat{\mathsf{U}}$ is not diagonal, the result of propagating a mixed state with non-degenerate eigenvalues, e.g.

$$(\hat{\rho}_{\mathrm{mixed}})_{ij} = \frac{2\,(d-i+1)}{d\,(d+1)}\delta_{ij}\,, \tag{7.5}$$

can be diagonalized to obtain the basis $\{|\tilde{\Psi}_k\rangle\}$. We then use that $\hat{\mathsf{U}}$ can be written as [244]

$$\hat{\mathsf{U}} = \sum_k \mathrm{e}^{\mathrm{i}\phi_k} |\tilde{\Psi}_k\rangle\langle\Psi_k| \,, \tag{7.6}$$

where $\{|\Psi_k\rangle\}$ are the canonical basis states. The phases are obtained from the propagation of $\hat{\rho}_{\mathrm{sp}}$ above, as

$$\phi_k = \arg \left\langle \tilde{\Psi}_k \middle| \hat{\rho}_{\mathrm{sp}}(T) \middle| \tilde{\Psi}_1 \right\rangle \,. \tag{7.7}$$

The numerical diagonalization that is required in order to construct the gate means that there is no way to calculate analytical derivatives of $\hat{\mathsf{U}}$. While a direct gate optimization with functionals based on Eq. (3.38) and using Krotov's method is possible, more advanced functionals that depend explicitly only on $\hat{\mathsf{U}}$, such as those discussed in chapter 6, are not available as easily. They require an *analytical* construction of

$$\hat{\mathsf{U}} = \begin{pmatrix} r_{11}\mathrm{e}^{\mathrm{i}\phi_{11}} & r_{12}\mathrm{e}^{\mathrm{i}\phi_{12}} & r_{13}\mathrm{e}^{\mathrm{i}\phi_{13}} & r_{14}\mathrm{e}^{\mathrm{i}\phi_{14}} \\ r_{21}\mathrm{e}^{\mathrm{i}\phi_{21}} & r_{22}\mathrm{e}^{\mathrm{i}\phi_{22}} & r_{23}\mathrm{e}^{\mathrm{i}\phi_{23}} & r_{24}\mathrm{e}^{\mathrm{i}\phi_{24}} \\ r_{31}\mathrm{e}^{\mathrm{i}\phi_{31}} & r_{32}\mathrm{e}^{\mathrm{i}\phi_{32}} & r_{33}\mathrm{e}^{\mathrm{i}\phi_{33}} & r_{34}\mathrm{e}^{\mathrm{i}\phi_{34}} \\ r_{41}\mathrm{e}^{\mathrm{i}\phi_{41}} & r_{42}\mathrm{e}^{\mathrm{i}\phi_{42}} & r_{43}\mathrm{e}^{\mathrm{i}\phi_{43}} & r_{44}\mathrm{e}^{\mathrm{i}\phi_{44}} \end{pmatrix} , \tag{7.8}$$

in the case of a two-qubit gate. Such a construction is provided by propagating the d states that contain a non-zero element only in the first column, e.g. the dyadic products

$$\hat{\rho}_1 = |00\rangle\langle 00| \;, \quad \hat{\rho}_2 = |01\rangle\langle 00| \;, \tag{7.9a}$$

$$\hat{\rho}_3 = |10\rangle\langle 00| \;, \quad \hat{\rho}_4 = |11\rangle\langle 00| \;. \tag{7.9b}$$

The entries (a,b) of the propagated state $\hat{\rho}_k$ are

$$\left(\hat{\mathsf{U}}\hat{\rho}_k\hat{\mathsf{U}}^\dagger\right)_{ab} = r_{b1} r_{ai} \mathrm{e}^{\mathrm{i}(\phi_{ai}-\phi_{b1})} \,. \tag{7.10}$$

The propagation of $\hat{\rho}_1$ alone determines all the entries in the first column of Eq. (7.8) (again with choosing one of the phases as zero). Plugging in the obtained values into the propagation of $\hat{\rho}_2$ then completely determines the second column of Eq. (7.8), and so forth. This gives an explicit construction of $\hat{\mathsf{U}}$, which now in principle allows to calculate the necessary derivatives for functionals like the one for a general perfect entangler, see appendix G.

So far, we have only considered how to obtain the gate $\hat{\mathsf{U}}$ under the assumption that the dynamical map is unitary. However, when not propagating the full basis of Liouville space, it is not guaranteed that the non-unitarity is detectable from the time evolution of propagated states. The non-unitarity might be either due to dissipation, or due to loss from the logical subspace by leakage into higher levels. It can be shown [196, 77] that for a direct gate optimization, in addition to the states in Eq. (7.3) and Eq. (7.5), a third state must be added to guarantee that the optimization is able to detect non-unitarity in the dynamical map. The optimization is successful if, and only if

$$\hat{\rho}_k(T) = \hat{\mathsf{O}}\hat{\rho}_k(0)\hat{\mathsf{O}}^\dagger \tag{7.11}$$

for the three states $\hat{\rho}_k$. The total set of states that is sufficient for a direct gate optimization reads

$$(\hat{\rho}_1(0))_{ij} = \frac{2\,(d-i+1)}{d\,(d+1)}\delta_{ij}\,, \tag{7.12a}$$

$$(\hat{\rho}_2(0))_{ij} = \frac{1}{d}\,, \tag{7.12b}$$

$$(\hat{\rho}_3(0))_{ij} = \frac{1}{d}\delta_{ij}\,, \tag{7.12c}$$

where the matrix elements are given in the optimization subspace, all other elements are zero. For the optimization with a functional that depends explicitly only on $\hat{\mathsf{U}}$, as discussed above, unitarity is guaranteed by the functional [244]

$$J_U = \sum_{k=1}^{3}\sum_{l=1}^{d} \mathrm{tr}^2\left[\hat{\rho}_k^l(T) - \hat{\rho}_k^l\right] \tag{7.13}$$

going to zero for the three states defined in Eq. (7.12).

7.1.2 Optimization Functional

In order to employ optimal control theory, we must define a distance measure J_T between the desired unitary $\hat{\mathsf{O}}$ and the actual evolution, as a final time functional. We generalize Eq. (3.40) to

$$J_T = 1 - \sum_{k=1}^{n} \frac{w_k}{\mathrm{tr}[\hat{\rho}_k^2(0)]}\,\mathfrak{Re}\left\{\mathrm{tr}\left[\hat{\mathsf{O}}\hat{\rho}_k(0)\hat{\mathsf{O}}^\dagger\hat{\rho}_k\,(T)\right]\right\}\,, \tag{7.14}$$

with $n = 3$ and using the states $\hat{\rho}_k(0)$ defined in Eq. (7.12). This is in contrast to Refs. [66, 245, 246], where n was taken to be the Liouville space dimension corresponding to $\hat{\mathsf{O}}$, i.e., $n = 2^{2N}$ for N qubits, and $\hat{\rho}_k$ an orthonormal basis (under the Hilbert-Schmidt product) of Liouville space. The three states are constructed such that the first one fixes a basis, and the corresponding Hilbert-Schmidt product in Eq. (7.14) checks whether the gate is correctly implemented in this basis. It misses errors for gates that are diagonal in the basis, i.e., phase errors [77]. The second state is therefore chosen to detect phase errors with its contribution to Eq. (7.14) [77]. The term for the third state determines whether the dynamical map attained at time T conserves the population within the optimization subspace. In Eq. (7.14), w_k are weights, normalized as $\sum_{k=1}^{n} w_k = 1$, that allow to put more or less importance to the properties checked by each of the three states. In order to evaluate J_T, the time evolved states $\hat{\rho}_k(T)$ need to be obtained by solving the equation of motion describing the open system's evolution for $\hat{\rho}_k$. While in general the dynamics can be non-Markovian, we will restrict ourselves to a Markovian master equation in the examples below. We assume the coherent part to include coupling to an external control, i.e., the Hamiltonian is of the form $\hat{\mathsf{H}}(t) = \hat{\mathsf{H}}_0 + \epsilon(t)\hat{\mathsf{H}}_1$, and generalization to several controls $\epsilon_i(t)$ is straightforward.

The states $\hat{\rho}_1$ and $\hat{\rho}_2$ of Eq. (7.12), while sufficient in principle to distinguish any two unitaries, do not allow for stating bounds on the gate error [77]. Meaningful bounds on the gate error can be obtained numerically by replacing $\hat{\rho}_1$, $\hat{\rho}_2$ by a set of $d+1$ states, whereas analytical bounds can be deduced from $2d$ states [77, 247, 248]. Motivated by this fact, we define two additional sets of states that can be employed in Eq. (7.14). When n in Eq. (7.14) is taken to be equal to $d+1$, the totally mixed state of Eq. (7.12a) is replaced by d pure states,

$$\hat{\rho}_j(0) = |\varphi_j\rangle\langle\varphi_j|\,, \tag{7.15}$$

with $j = 1, \ldots, d$ and $\{|\varphi_j\rangle\}$ the logical basis. $\hat{\rho}_{d+1}(0)$ is simply equal to $\hat{\rho}_2(0)$ of Eq. (7.12b). In this case, Eq. (7.12c) is not required since the $d+1$ pure states are sufficient to enforce the dynamical map on the optimization subspace to be contracting and norm conserving. Similarly, the functional (7.14) employing $n = 2d$ states is constructed by replacing $\hat{\rho}_1(0)$ of Eq. (7.12a) by $\hat{\rho}_j$, $j = 1, \ldots, d$ of Eq. (7.15) and $\hat{\rho}_2(0)$ of Eq. (7.12b) by

$$\hat{\rho}_{d+j}(0) = |\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_j|\,, \tag{7.16}$$

with $j = 1, \ldots, d$, where the states $|\tilde{\varphi}_j\rangle$ form a mutually unbiased basis [249, 250] with respect to the canonical basis $\{|\varphi_j\rangle\}$. For two qubits ($d = 4$), an example for such a basis is given by

$$|\tilde{\varphi}_1\rangle = \frac{1}{2}\left(|00\rangle + |01\rangle + |10\rangle + |11\rangle\right)\,, \tag{7.17a}$$

$$|\tilde{\varphi}_2\rangle = \frac{1}{2}\left(|00\rangle - |01\rangle + |10\rangle - |11\rangle\right)\,, \tag{7.17b}$$

$$|\tilde{\varphi}_3\rangle = \frac{1}{2}\left(|00\rangle + |01\rangle - |10\rangle - |11\rangle\right)\,, \tag{7.17c}$$

$$|\tilde{\varphi}_4\rangle = \frac{1}{2}\left(|00\rangle - |01\rangle - |10\rangle + |11\rangle\right)\,. \tag{7.17d}$$

7.1.3 Optimization with Krotov's Method

In order to obtain an update equation for Krotov's method, a field dependent constraint as in Eq. (3.61) is added to the final time functional, Eq. (7.14). As discussed in chapter 3, more complex additional constraints are conceivable, for example restricting the spectral width of the pulse or confining the accessible state space [159, 158]. The full optimization functional is linear in the states $\hat{\rho}_k(T)$ and does not depend on the states at intermediate times t, so that the linear version of Krotov's method is sufficient to yield a monotonically convergent optimization algorithm. Modeling the dissipative time evolution by a Markovian master equation,

$$\frac{d\hat{\rho}}{dt} = \mathcal{L}(\hat{\rho}) = -i[\hat{\mathsf{H}}(t), \hat{\rho}] + \mathcal{L}_D(\hat{\rho}) \,, \tag{7.18}$$

the control equations then read

$$\frac{\partial \hat{\rho}_k}{dt} = -i[\hat{\mathsf{H}}, \hat{\rho}_k] + \mathcal{L}_D(\hat{\rho}_k) \,, \tag{7.19a}$$

$$\frac{\partial \hat{\sigma}_k}{dt} = -i[\hat{\mathsf{H}}, \hat{\sigma}_k] - \mathcal{L}_D(\hat{\sigma}_k) \quad \text{and} \quad \hat{\sigma}_k(t=T) = \frac{w_k}{\mathrm{tr}[\hat{\rho}_k^2(0)]} \hat{\mathsf{O}} \hat{\rho}_k(0) \hat{\mathsf{O}}^\dagger \,, \tag{7.19b}$$

$$\Delta\epsilon(t) = \frac{S(t)}{\lambda_a} \sum_{k=1}^{n} \mathfrak{Im} \left\{ \mathrm{tr} \left(\hat{\sigma}_k^{\mathrm{old}}(t) \frac{\partial \mathcal{L}\left(\hat{\rho}_k\right)}{\partial \epsilon} \Big|_{\rho_k^{\mathrm{new}}, \epsilon^{\mathrm{new}}} \right) \right\} \tag{7.19c}$$

with $k = 1, 2, 3$ when the initial conditions $\hat{\rho}_k(0)$ of Eq. (7.12) are employed or $k = 1, \dots d^2$ with d the dimension of Hilbert space when a full basis of Liouville space is propagated. In Eq. (7.19c), the states $\hat{\sigma}_k^{\mathrm{old}}$ are backward-propagated with the pulse of the previous iteration ('old'), whereas the states $\hat{\rho}_k^{\mathrm{new}}$ are forward-propagated with the updated pulse ('new'). The derivative with respect to the field is given by the commutator

$$\frac{\partial \mathcal{L}\left(\hat{\rho}\right)}{\partial \epsilon} = -i \left[\frac{\partial \hat{\mathsf{H}}}{\partial \epsilon}, \hat{\rho} \right] \tag{7.20}$$

and has to be evaluated for the 'new' field and the states $\hat{\rho}$ propagated under the 'new' field. For a complex control, which occurs for example when using the rotating wave approximation (RWA), Eq. (7.19c) holds for both the real and the imaginary part of $\epsilon(t)$.

The value of the optimization functional in Eq. (7.14) depends on the number and the specific choice of initial states as well as the choice of weights. It is therefore not suitable to compare the convergence behavior between different sets of states. Instead, we employ the average gate fidelity as defined in Eq. (2.89).

single-photon detuning Δ_1	600 MHz
two-photon detuning Δ_2	0
excitation energy E_1	6.8 GHz
Rabi frequencies Ω_R, Ω_B	300 MHz
interaction energy U	50 MHz
lifetime $\tau = 1/\gamma$	25 ns

Table 7.1: Parameters of the Hamiltonian, Eq. (4.1), for implementing a controlled phasegate with two rubidium atoms.

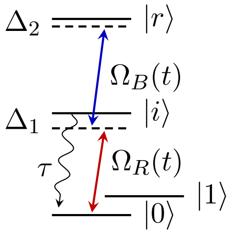


Figure 7.1: Atomic levels for two-photon near-resonant excitation to a Rydberg state.

7.2 Example I: Diagonal gates

It is quite common that a two-qubit Hamiltonian allows only for diagonal gates, such as a controlled phasegate. A prominent example are non-interacting qubit carriers that interact only when excited into an auxiliary state where they accumulate a non-local phase [178]. Neutral trapped atoms with long-range interaction in a Rydberg state, discussed in chapter 4, present a physical implementation of this setting [178, 166]. We consider here the same system as in chapter 4, albeit in a slightly different parameter regime. The parameters correspond to optically trapped rubidium atoms (as opposed to cesium atoms in chapter 4) and are summarized in Table 7.1. As before, the excitation to the Rydberg state proceeds by a near-resonant two-photon process via an intermediary state. As shown in Fig. 7.1, the transition $|0\rangle \rightarrow |i\rangle$ now corresponds to a red laser frequency, whereas $|i\rangle \rightarrow |r\rangle$ is a blue frequency, in contrast to Fig. 4.2. The gate is implemented for non-individually addressable atoms, and outside of the Rydberg blockade regime [178], cf. the discussion in chapter 4. The Hamiltonian and the equation of motion is given by Eqs. (4.1–4.4).

Figure 7.2 shows the gate error of the controlled phasegate versus iteration of the optimization algorithm when using a full basis, i.e., 16 states, or using three, respectively two, states in Eq. (7.14). The minimum number of states in this example is two since the Hamiltonian admits only diagonal gates,

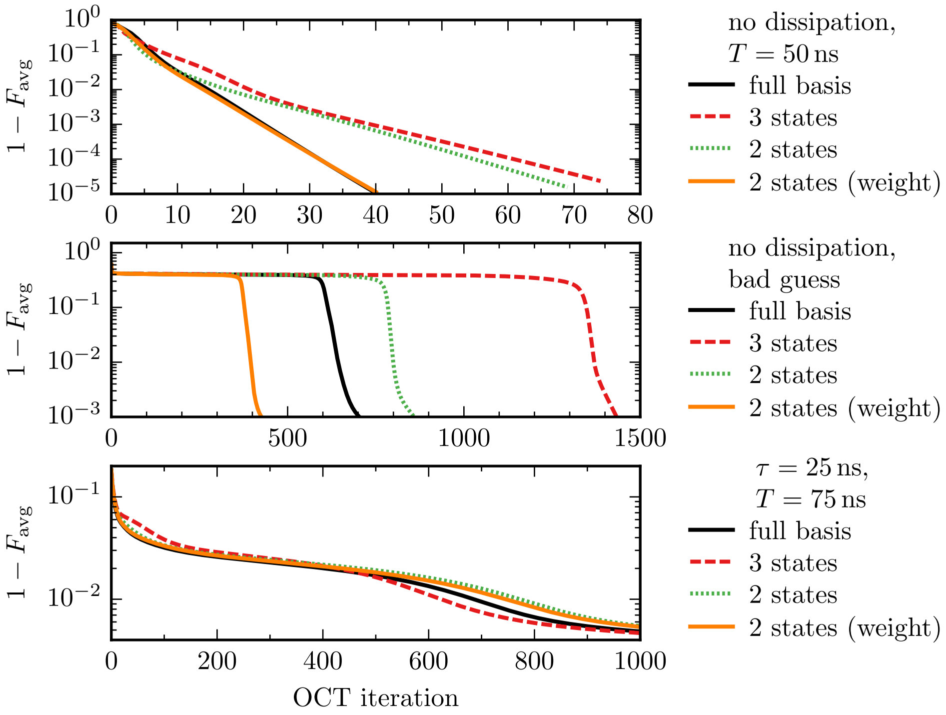


Figure 7.2: Optimizing a controlled phasegate for two trapped neutral atoms that are excited to a Rydberg state. The convergence is shown as the gate error, $1 - F_{\text{avg}}$, over OCT iterations, using the full basis as well as several reduced sets states. The calculations employ equal weights of all states, except for those shown in orange where $w_2/w_3 = 10$. The optimization shown in the bottom panel takes into account spontaneous emission from the intermediate state, with a lifetime of $\tau = 25\,\text{ns}$. The gate duration is $T = 50\,\text{ns}$ for the top and middle panels, and $T = 75\,\text{ns}$ for the bottom panel. The number of iterations and the reached gate error differ significantly in all three situations, cf. the different x- and y-axes scales.

i.e., only phase errors and norm conservation within the logical subspace have to be checked. Therefore, $\hat{\rho}_1$ in Eq. (7.12a) can be omitted, and the two remaining states are $\hat{\rho}_2$ (phase errors) and $\hat{\rho}_3$ (norm conservation) of Eqs. (7.12b, 7.12c). The relative weights w_2 and w_3 in Eq. (7.14) can be modified to emphasize one of the two aspects. Figure 7.2 therefore also compares two states with equal and unequal weights in Eq. (7.14), cf. green dotted and orange solid lines. The fastest convergence was obtained for $w_2/w_3 = 10$. The panels from top to bottom show the optimization without any dissipation, starting from a well-chosen guess pulse; an optimization starting with a bad guess pulse of insufficient fluence; and an optimization taking into account spontaneous decay from the intermediate level. As the main observation, Fig. 7.2 clearly demonstrates that only two states are sufficient to optimize a quantum gate for a Hamiltonian of this kind. The optimization for coherent time evolution (top panel), shows that while the use of three states converges to gate errors as small as those obtained with the full basis, the convergence rate is only about half that of the full basis. This is due to two factors: (i) For the optimization with three states, there is no bound on the distance between the value J_T and the gate error, such that the path in the optimization landscape may be less direct until an asymptotic value is reached. Since without dissipation, there is no limit to the gate error, the convergence of J_T and that of $1 - F_{\text{avg}}$ stay on different trajectories. (ii) The reduced sets of states are constructed specifically to take into account decoherence. In particular, the third state contributes significantly less information that is relevant for reaching the optimization target than the second state. The convergence can be improved dramatically by weighting the three states according to the relevance of the information they carry. In this respect, the use of only two initial states can be seen as choosing $w_1 = 0$. Taking $w_2 > w_3$ addresses the issue of $\hat{\rho}_3$ contributing less to the optimization. Choosing proper weights allows for ensuring the convergence of optimization with a reduced set of states to be as fast as the optimization using the full basis.

The importance of choosing weights appropriate to the optimization problem becomes even more evident when the optimization starts from a bad guess pulse of insufficient fluence, as shown in the center panel of Fig. 7.2. The features observed in Fig. 7.2 are typical: The plateau near the beginning corresponds to the optimization increasing the intensity of the pulse without any significant improvement in the gate error, before converging quickly once the pulse is sufficiently intense. The end of the plateau can be significantly influenced by the choice of weights, cf. solid orange and dotted green curves in the middle panel of Fig. 7.2. Remarkably, the optimal choice of using two properly weighted initial states outperforms the use of the full basis. This might be explained by the fact that each of the three states in the reduced set has a specific physical role to play in the optimization, and this role can be emphasized by choice of the weight. In contrast, all states in the full basis fulfill the same role in the optimization, and thus there is no way in which different weights on individual states would improve the convergence.

One should point out that even in the cases where the use of two or three states shows a slower convergence than that of the full basis, they still outperform the full basis in terms of numerical resources. Since both CPU time and the required memory scale linearly with the number of initial states in the optimization, using only two states compared to 16 has a 1:8 advantage, which more than offsets the factor of two in the convergence rate in the middle panel of Fig. 7.2.

Naturally, without the presence of decoherence, there is no reason to perform the optimization in Liouville space. Therefore, the results shown here only serve to illustrate the general convergence behavior of a reduced set of initial states. The more relevant case of non-coherent dynamics is shown in the bottom panel of Fig. 7.2. The presence of decoherence implies the existence of an asymptotic bound on the gate error. This constraint on the optimization landscape (together with the further constraint that only diagonal gates are reachable) ensures that all sets of reduced states converge at a similar rate, once the asymptotic region is approached. We expect that all choices reach the same asymptotic value; which choice yields the best fidelity after a specific number of iterations cannot be predicted in general. Factoring in all necessary resources, optimization using two states with unequal weights dramatically outperforms optimization using the full basis in this example.

The optimized pulse and spectrum in the case of coherent dynamics is presented in Fig. 7.3. The result shown here is obtained from the optimization using two initial states with unequal weights. However, the pulse is indistinguishable from the one obtained using the full basis, consistent with the identical convergence behavior for the two sets in the upper panel of Fig. 7.2. The optimized pulses only show relatively small amplitude modulations compared to the guess pulse (dotted line). These modulations appear as small side-peaks in the spectrum. In the time interval in which there is a significant pulse amplitude, the complex phase only deviates by about $\frac{\pi}{10}$ from zero. This phase evolution is reflected in the asymmetry of the spectrum for the red and the blue pulse (bottom panel). The spectrum nicely illustrates the mechanism of control: while each spectrum by itself is asymmetric, the red pulse showing negative frequencies, the blue pulse showing positive frequencies, the sum of both pulses is again symmetric, i.e., positive and negative frequencies cancel out. This means that the combination of both pulses is two-photon resonant with the transition $|0\rangle \rightarrow |r\rangle$, providing multiple pathways for the same transition whose interference might be exploited by the optimization.

The population dynamics induced by the optimized pulses are shown in Fig. 7.4. The two-photon resonance of the pulse expresses itself in a direct Rabi cycling (see appendix C) between $|0\rangle$ and $|r\rangle$ on the left qubit in the propagation of $|01\rangle$ (top panel). The population shows roughly a 4π Rabi flip due to the relatively high pulse intensity. The nearly 25% of the population in the intermediate states in the propagation of $|00\rangle$ (bottom panel) is due to the fact that the decay from these levels was not included in the optimization, and thus the optimization algorithm makes no attempt at suppressing population in these states.

For the optimization with dissipation, the optimized pulse and pulse spectrum is shown in Fig. 7.5. The characteristics of the pulses are quite different compared to the coherent case. The red pulse remains close to the

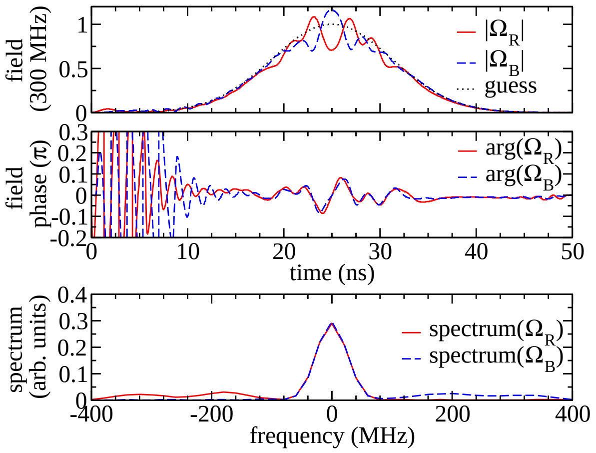


Figure 7.3: The optimized pulses $\Omega_{B,R}(t)$ for the blue and red laser cf. Fig 7.1, resulting from optimization using two states with unequal weights without spontaneous decay (corresponding to the orange solid line in the top panel of Fig. 7.2). The pulse amplitudes are shown in the top panel, the complex phase in the center panel, and the pulse spectrum in the bottom panel. The guess pulse, indicated by the black dotted line in the top panel, is identical for both the red and the blue laser. In the spectrum, frequency 0 corresponds to the carrier frequencies of the laser pulses.

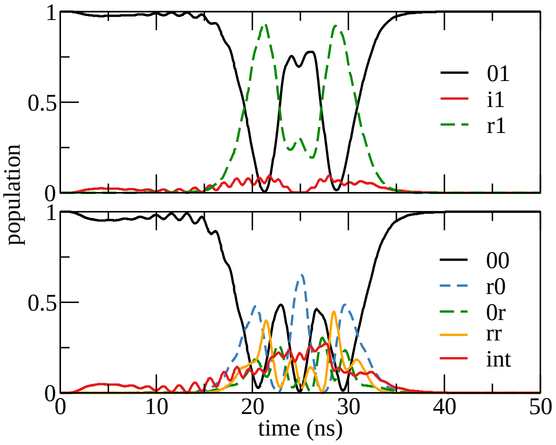


Figure 7.4: Population dynamics under the pulse shown in Fig. 7.3, for the logical basis states $|01\rangle$ (top) and $|00\rangle$ (bottom). The intermediate population (”int”) is integrated over all levels with decay, i.e., $|0i\rangle$, $|i0\rangle$, $|ii\rangle$, $|ir\rangle$, and $|ri\rangle$.

single Gaussian peak of the guess pulse, except for being slightly narrower. The blue pulse has a more complex structure. It is overall broader than the red pulse and consists of three distinctive features: an initial peak that overlaps but precedes the red pulse, followed by some amplitude oscillations in the center of the pulse, and lastly another peak symmetric to the first, thus following the red laser pulse, with some overlap. For both pulses, the complex phase, shown in the center panel, is close to zero when there is significant pulse amplitude. In the spectrum (bottom panel), the overall narrowing and broadening of the red and blue pulse, respectively, is reflected in a broadening and narrowing of the central peak in the spectrum. The amplitude modulations on the blue pulse appear as side-lobes in the spectrum.

The initial and final peak of the blue pulse, together with the red pulse are reminiscent of the counter-intuitive pulse scheme of STIRAP, with the blue laser acting as the "Stokes" pulse and the red laser as "pump". The STIRAP-like behavior appears also in the population dynamics, shown in Fig. 7.6, as a population inversion between level $|0\rangle$ and $|r\rangle$, *without* any population in the intermediate decaying state. The amplitude modulations in the central region of both pulses then induce some additional dynamics, generating the entanglement needed for the gate. Note that the pulse duration for the dissipative process ($T = 75\,\text{ns}$) is longer than that of the coherent process ($T = 50\,\text{ns}$). This is necessary to allow for an adiabatic time evolution that is essential to the STIRAP-like behavior. Overall, the decaying intermediate state population (red lines in Fig. 7.6) is almost completely suppressed, which is in contrast to the optimization not taking into account the dissipation, cf. the red lines in Fig. 7.4. Both Figs. 7.4 and 7.6 show a significant population of the $|rr\rangle$ state. This is not surprising since the parameters of Table 7.1 are not in the regime of the Rydberg blockade [178, 166].

7.3 Example II: Non-diagonal gates

Superconducting qubits represent a physical realization of a quantum processor where the Hamiltonian admits both diagonal and non-diagonal entangling gates. In fact, there exist superconducting architectures that admit several two-qubit gates simultaneously [251, 222]. We consider here the example of two transmon qubits coupled via a shared transmission line resonator. In the dispersive limit, the interaction of each qubit with the resonator leads to an effective coupling J between the two qubits, and the cavity can be

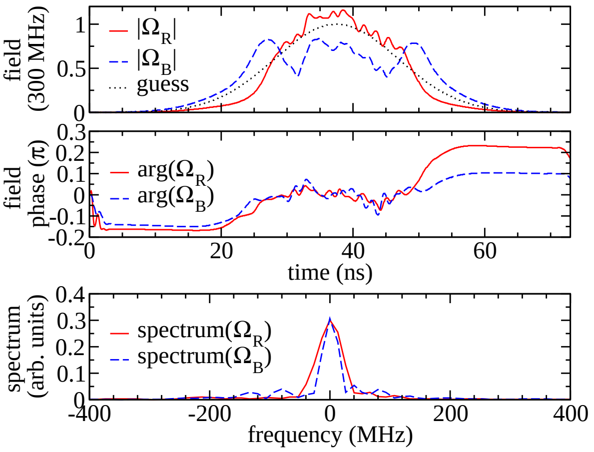


Figure 7.5: The optimized pulses resulting from optimization using two weighted states and including spontaneous decay (orange solid line in Fig. 7.2, bottom panel), using the same conventions as Fig. 7.3.

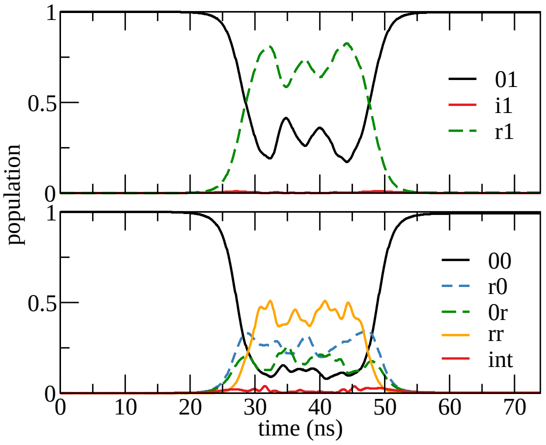


Figure 7.6: Dissipative population dynamics under the pulse shown in Fig. 7.5, for the initial states $\hat{\rho}(0) = |01\rangle\langle 01|$ (top) and $\hat{\rho}(0) = |00\rangle\langle 00|$ (bottom). The intermediate population ("int") is integrated over all levels with decay, i.e. $|0i\rangle$, $|i0\rangle$, $|ii\rangle$, $|ir\rangle$, and $|ri\rangle$.

qubit frequency ω_1	4.3796 GHz
qubit frequency ω_2	4.6137 GHz
drive frequency ω_d	4.4985 GHz
anharmonicity α_1	−239.3 MHz
anharmonicity α_2	−242.8 MHz
effective qubit-qubit coupling J	−2.3 MHz
qubit 1 decay time T_1	38.0 µs
qubit 2 decay time T_1	32.0 µs
qubit 1 dephasing time T_2^*	29.5 µs
qubit 2 dephasing time T_2^*	16.0 µs

Table 7.2: Parameters of the transmon Hamiltonian, Eq. (7.21), and Liouvillian, Eq. (7.22), taken from Ref. [222].

integrated out [222]. The resulting Hamiltonian reads, cf. Eq (5.30),

$$\begin{aligned}\hat{\mathsf{H}} = \sum_{i=1,2} \left(\left(\omega_i - \frac{\alpha_i}{2} \right) \hat{\mathsf{b}}_i^\dagger \hat{\mathsf{b}}_i + \frac{\alpha_i}{2} \left(\hat{\mathsf{b}}_i^\dagger \hat{\mathsf{b}}_i \right)^2 \right) + \\ + J^{\text{eff}} \left(\hat{\mathsf{b}}_1^\dagger \hat{\mathsf{b}}_2 + \hat{\mathsf{b}}_1 \hat{\mathsf{b}}_2^\dagger \right) + \epsilon(t) \left(\hat{\mathsf{b}}_1 + \hat{\mathsf{b}}_1^\dagger + \hat{\mathsf{b}}_2 + \hat{\mathsf{b}}_2^\dagger \right) ,\end{aligned} \tag{7.21}$$

where $\hat{\mathsf{b}}_{1,2}$, $\hat{\mathsf{b}}_{1,2}^\dagger$ are the ladder operators for the first and second qubit, $\omega_{1,2}$ and $\alpha_{1,2}$ represent the frequency and anharmonicity, J is the effective qubit-qubit-interaction, and $\epsilon(t)$ is the driving field, see Eq. (5.37). The two most relevant dissipation channels are energy relaxation and pure dephasing of the qubits, described by the decay rate $\gamma = 1/T_1$ and dephasing rate $\gamma_\phi = 1/T_2^*$ for each qubit. The corresponding dissipator reads

$$\mathcal{L}_D(\hat{\rho}) = \sum_{q=1,2} \left(\gamma_q \sum_{i=1}^{N-1} iD \left[|i-1\rangle\langle i|_q \right] \hat{\rho} + \gamma_{\phi,q} \sum_{i=0}^{N-1} i^2 D \left[|i\rangle\langle i|_q \right] \hat{\rho} \right) , \tag{7.22}$$

with

$$D\left[\hat{\mathsf{A}}\right]\hat{\rho} = \hat{\mathsf{A}}\hat{\rho}\hat{\mathsf{A}}^\dagger - \frac{1}{2}\left(\hat{\mathsf{A}}^\dagger\hat{\mathsf{A}}\hat{\rho} + \hat{\rho}\hat{\mathsf{A}}^\dagger\hat{\mathsf{A}}\right) , \tag{7.23}$$

and each qubit, $q = 1, 2$, truncated at level N. The parameters of the coupled transmon qubits are summarized in Table 7.2. We employ a RWA, centered at the drive frequency ω_d. The pulse is then described by the shape $\Omega(t)$ instead of the fast-oscillating $\epsilon(t)$, see Eq. (5.39).

The Hamiltonian in Eq. (7.21) can generate a large number of entangling two-qubit gates; we find $\sqrt{\text{iSWAP}}$ to be a fast converging non-diagonal

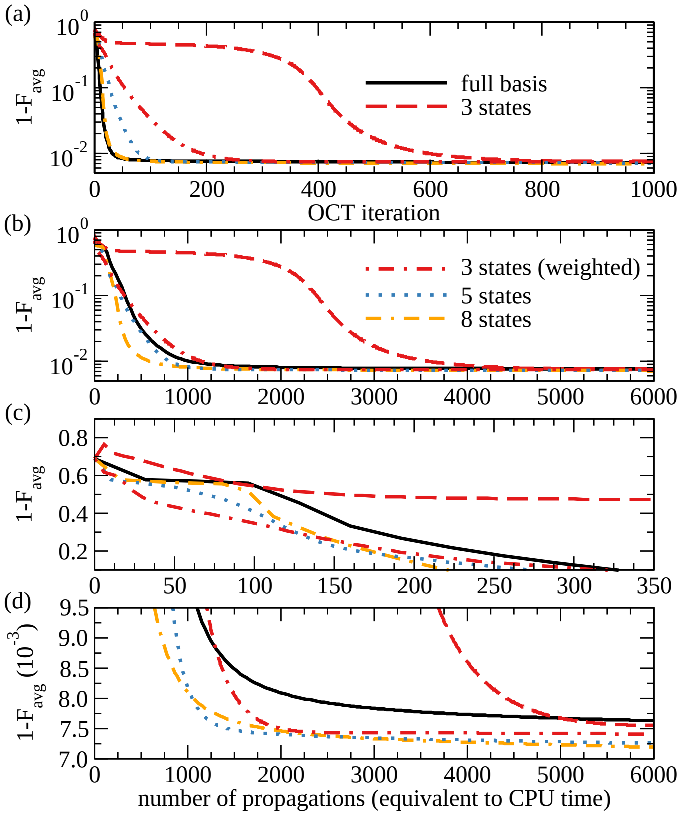


Figure 7.7: Optimizing a $\sqrt{\text{iSWAP}}$ gate for two transmons in the presence of energy relaxation and pure dephasing (with the rates given in Table 7.2): Convergence for five choices of sets of initial states, as described in the text. The gate duration is $T = 400\,\text{ns}$. The panels from top to bottom show the gate error over the number of iterations; the gate error over the number of state propagations, indicative of the required CPU time; a zoom on the initial phase of the optimization; and a zoom on the asymptotic convergence (panels c and d both using a linear scale).

perfect entangler, and thus choose

$$\hat{\mathsf{O}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{\mathrm{i}}{\sqrt{2}} & 0 \\ 0 & \frac{\mathrm{i}}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{7.24}$$

as the optimization target. Figure 7.7 shows the convergence behavior for several choices of initial states: the 16 canonical states of the full basis of Liouville space; the three states given in Eq. (7.12) with equal weight and with $w_1/w_2 = w_1/w_3 = 20$; a set of 5 states consisting of $\hat{\rho}_1$ expanded into four pure states, cf. Eq. (7.15) plus $\hat{\rho}_2$ of Eq. (7.12b); and lastly a set of eight states, cf. Eqs. (7.15) and (7.16), consisting of the expansion of $\hat{\rho}_1$ and the four pure states of a mutually unbiased basis, as explained in section 7.1. As seen in the top panel, all choices show good convergence. A plateau corresponding to a slowing of convergence is observed only for the three states with equal weights. But even in this case, the same asymptotic value for the gate error is obtained as for the other choices, see also Fig. 7.7 (d). The advantage of employing the reduced sets of states in the optimization functional, Eq. (7.14), becomes most apparent in Fig. 7.7 (b) which shows the gate error over the number of state propagations. Since optimization requires two propagations per iteration and state, i.e., the backward and forward propagation in Eq. (7.19), the number of state propagations corresponds directly to the CPU time that is required to obtain a given fidelity. That is, in panels (b-d), the lines are simply rescaled depending on the respective number of states. Figure 7.7 (c) and (d) show a zoom on the same data, once for the initial phase of the optimization and once for the asymptotic behavior. All reduced sets except for the three states with equal weights perform better than the full set during the initial phase. Also, for this specific optimization problem, all reduced sets reach a slightly better asymptotic value than the full set, although we expect that ultimately all curves will converge to the same value. Figure 7.7 suggests that the reduced sets have a significant advantage in reaching a good fidelity with a given amount of resources, especially since in practice, an optimization is usually stopped near the beginning of the asymptotic regime. Indeed, the full set shows an advantage only in the intermediate regime between gate errors of 10 and 1 percent, and only over the sets of three states. The choice of 5 or 8 states outperforms the full set in all cases. One should note that the savings in computational resources due to the use of a reduced set of states also extends to the amount of memory required, which is proportional to the number of states. Since in the optimization algorithm, propagated states over the entire time grid need to be stored, these savings can be very substantial.

For the three states with equal weights the gate error shows a non-monotonic behavior in the upper left corner of Fig. 7.7 (c). This is due to the optimization functional, Eq. (7.14), not being equivalent to the gate error F_{avg}, Eq. (2.89). Specifically, for a set of three states, no bound on the distance between J_T and $1 - F_{\text{avg}}$ can be derived [77]. Thus, the gate error might increase even though J_T decreases. In fact, the behavior of J_T is fully monotonic as expected (data not shown). With an increasing number of states in the chosen set, the value of the optimization functional is more closely connected to the gate fidelity; and for 5 and 8 states numerical, respectively analytical, bounds can be found [77, 247]. For this reason, we expect the sets of 5 and 8 states to show a faster convergence than the 3 states, when measured in OCT iterations, although not necessarily in CPU time. This expectation is confirmed by Fig. 7.7. The weak correspondence between the optimization functional and the gate error for three states is most likely also the reason for the plateau observed for the red dashed line in Fig. 7.7 (a) and (b). However, the use of three states can still be a good choice since weighting the states properly improves the convergence significantly. The weights have to be chosen empirically, but the choice can be guided by physical intuition. The three states are responsible for ensuring that the realized gate is diagonal in the correct basis, that the relative phases match the target once the correct basis has been found, and that the gate is unitary on the logical subspace, respectively. The weights should reflect which of these requirements is most difficult to realize. In the present example this is finding the correct basis in which the gate is diagonal. Therefore the choice of $w_1/w_2 = w_1/w_3 = 20$ gave the best convergence rate. This is in contrast to the optimization of the Rydberg gate in section 7.2, in which the gate was already known to be diagonal, and the first state could be left out of the optimization entirely. Generally, using the set of three states with equal weights is not recommended.

Comparing Fig. 7.7 with the bottom panel of Fig. 7.2 for the Rydberg gate shows that the different choices of basis sets show a slightly wider range of the convergence rate. This can be attributed to the fact that for the Rydberg gate, the optimization landscape is severely constrained since only diagonal gates can be reached. In contrast, the transmon Hamiltonian can generate both diagonal and non-diagonal gates, resulting in a more complex optimization landscape. Different choices of initial states can thus take more strongly varying pathways.

It must be noted that the different pathways through the optimization landscape for different choices of the set of states or different weights result in optimized fields that are generally not identical. For the Rydberg gate, where the dynamics are constrained to be diagonal, pulses with the same fidelities are virtually indistinguishable. This is not the case for the transmon qubits. Here, there are at least slight differences in the pulses obtained asymptotically, and even for pulses with identical fidelities (at crossing points of the curves in Fig. 7.7). This illustrates that the solution to an optimal control problem is in general not unique. Different solutions with identical fidelities can also be obtained e.g. by starting from different guess pulses.

Figure 7.8 shows the optimization of a $\sqrt{\text{iSWAP}}$ gate for two transmons in the case of weak dissipation, where the decay and dephasing times from Table 7.2 have been increased by a factor of 10. A comparison of Fig. 7.8 (a) with Fig. 7.7 (a) shows that the convergence behavior is essentially the same except for the value of the asymptote. We find an asymptotic gate error of approximately 7×10^{-3} with full dissipation, 7×10^{-4} with weak dissipation, and no asymptote without dissipation (data not shown). The

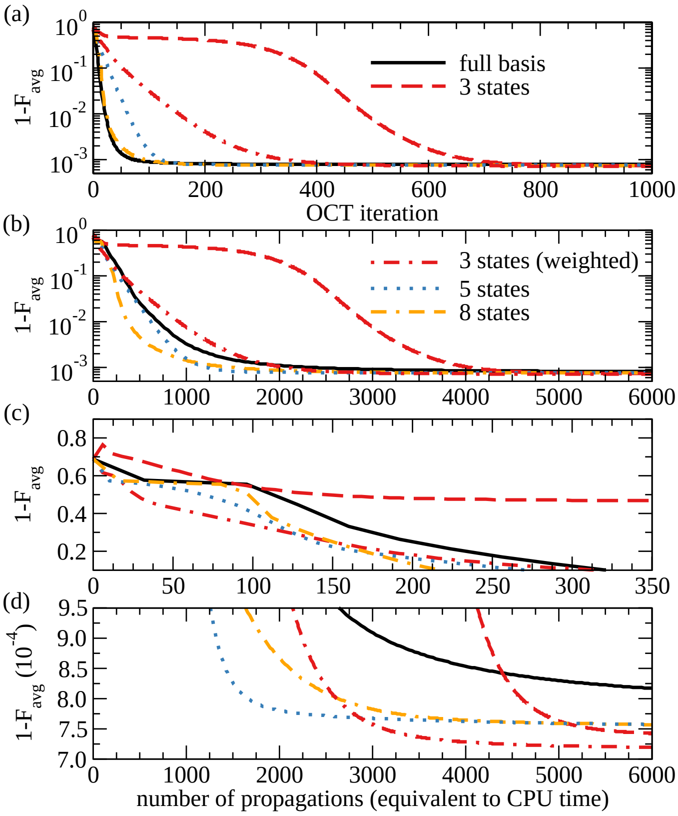


Figure 7.8: Optimizing a $\sqrt{\text{iSWAP}}$ gate for two transmons with weak dissipation, using decay and dephasing times increased by a factor of 10 compared to Fig. 7.7 (with all quantities and labels as defined in Fig. 7.7). The gate duration is $T = 400\,\text{ns}$. The weaker dissipation results in an asymptotic gate error of approximately 7.5×10^{-4} compared to 7.5×10^{-3} in Fig. 7.7, cf. the y-axis scales in both figures.

value of the asymptote is logarithmically proportional to the decay and dephasing rates. This is as expected since the pulse duration is kept constant at 400 ns and the gate fidelity is solely limited by dissipation. Our claim that the dissipation only affects the asymptotic convergence is supported by a comparison of the initial convergence in Figs. 7.7 (c) and 7.8 (c), which remarkably are completely identical. Furthermore, the crossing between the black solid and red dot-dashed lines for the full basis and the three states with unequal weights near 1000 propagations and that between the blue dotted and orange dash-dash-dotted lines for the sets of 5, respectively 8, states near 1300 propagations in Fig. 7.8 (d) can also be seen in Fig. 7.7 (d). There are however some slight differences in the asymptotically reached values, in that the choice of 3 states (with both equal and unequal weights) reaches a slightly smaller gate error than in the case of full dissipation. Again, we expect that ultimately, all curves will converge to the same value.

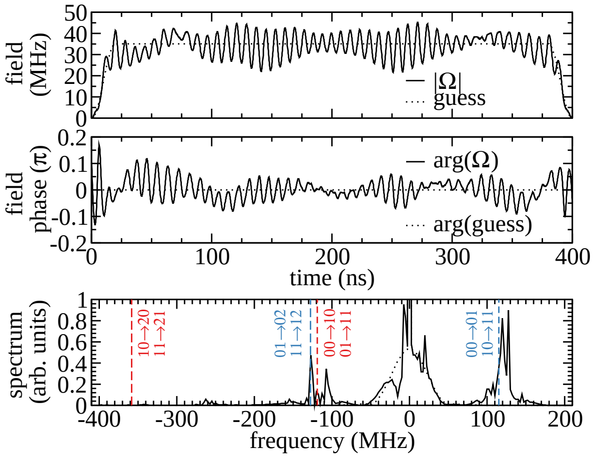


Figure 7.9: Shape and spectrum of an optimized pulse, from optimization with 3 weighted states, with strong dissipation. The panels from top to bottom show the amplitude, complex phase, and spectrum of the optimized pulse $\Omega(t)$. The spectrum is shown in the rotating frame, with zero corresponding to the driving frequency w_d of the field. The transition frequencies from the logical subspace are indicated by vertical dashed lines. These are $\Delta_1 = w_1 - w_d = -118.88\,\text{MHz}$ and $\Delta_1 - \alpha_1 = -358.18\,\text{MHz}$ in red for the left qubit, and $\Delta_2 = w_2 - w_d = 115.20\,\text{MHz}$ and $\Delta_2 - \alpha_2 = -127.58\,\text{MHz}$ in blue for the right qubit. The central peak in the spectrum has been cut off to show the relevant side-peaks, and would extend to a value of approximately 10.0. For all quantities, the values for the guess pulse are shown as a dotted line.

Which set of states reaches the best gate error at a specific point near the beginning of the asymptotic region seems to depend on the slope of the convergence curve as the limit is approached. This can depend on any number of factors including, e.g., the choice of λ_a in Eq. (3.61). Again, empirically, the reduced sets of states show a significant numerical advantage over the full basis also for weak dissipation.

As an example, the optimized pulse obtained using a set of three states with unequal weights, taking into account the full dissipation, is presented in Fig. 7.9, along with the pulse spectrum. The population dynamics that this pulse induces when propagating the logical basis states $\hat{\rho}(t=0) = |01\rangle\langle 01|$ and $\hat{\rho}(t=0) = |11\rangle\langle 11|$ is shown in Fig. 7.10. As can be seen in the top panel of Fig. 7.9, the optimized pulse shows small oscillations around the guess peak amplitude of 35 MHz. The complex phase, shown in the middle panel, stays relatively close to zero, indicating that the optimization employs mainly amplitude modulation. The pulse amplitude is roughly time-symmetric. The pulse spectrum shown in the bottom panel of Fig. 7.9 relates easily to the pulse shape. The strongest frequency component remains the driving frequency of the guess pulse (zero in the spectrum). The small oscillations in the pulse shape are approximately 8 ns apart, corresponding to a frequency of ±125 MHz, which is present in the spectrum. There are peaks with exponentially decaying amplitude in the spectrum at multiples of these values. The width of the central peak is due to the 20 ns switch-on and switch-off time of the pulse, and is unchanged from the guess pulse. The fact that there is not a single, but a double peak around ±125 MHz corresponds the slow beats in the pulse shape. The slight asymmetry of the spectrum is caused by the complex phase of the optimized pulse.

The spectrum of the optimized pulse is very instructive in understanding the population dynamics in Fig. 7.10. The most relevant transition frequencies from the logical subspace are indicated by vertical lines in the spectrum in the lower panel of Fig. 7.9. Clearly, the peaks around 125 MHz are nearly resonant with the excitation of the left and right qubit, and the excitation to level $\vert 2\rangle$ of the right qubit. There is no significant component in the spectrum that could excite to the level $\vert 2\rangle$ of the left qubit. Consequently, in the population dynamics of both the $\vert 01\rangle\langle 01\vert$ and $\vert 11\rangle\langle 11\vert$ state, the right qubit (top panel) leaves the logical subspace (expectation value $\langle j\rangle > 1.0$) to a much more significant extent than the left qubit (middle panel). This behavior is slightly more pronounced for $\vert 11\rangle\langle 11\vert$, which is the only state for which the total subspace population (gray curve in bottom panel) drops below 80% for a significant amount of time. The fact that for all logical basis states, most of the dynamics occurs within the logical subspace is due to the presence of decoherence, where higher levels have faster decay and faster dephasing due to a stronger coupling to the cavity. In an optimization without dissipation (data not shown), the optimized dynamics would generally veer farther outside the logical subspace. Lastly, the population dynamics show the expected behavior for the $\sqrt{\text{iSWAP}}$ gate: the $\vert 01\rangle$ state ends up in a coherent superposition between $\vert 01\rangle$ and $\vert 10\rangle$, whereas $\vert 11\rangle$ returns to its original state at the end of the gate.

In conclusion, we have demonstrated a significant reduction in the computational resources that are required to optimize a quantum gate in the framework of open quantum systems. This opens the door to the numerical realization of robust quantum gates that would otherwise be unfeasible with current numerical resources. An example where this becomes particularly relevant is the ensemble optimization presented in chapter 4. There, we optimized over an ensemble of 20 Liouvillians. This would ordinarily require the propagation of 320 Liouville space matrices, each of which for a gradient-based optimization must then be stored for each point of the time grid. Based on the results of section 7.2, these 320 propagations could be reduced to 40. It is important to note that while we have used a Markovian master equation to model the dissipation, the result that a reduced number of states is sufficient for the optimization of a unitary process is independent of the equation of motion or a specific decoherence model. Specifically, it also allows to explore master equations that go beyond the Markovian regime.

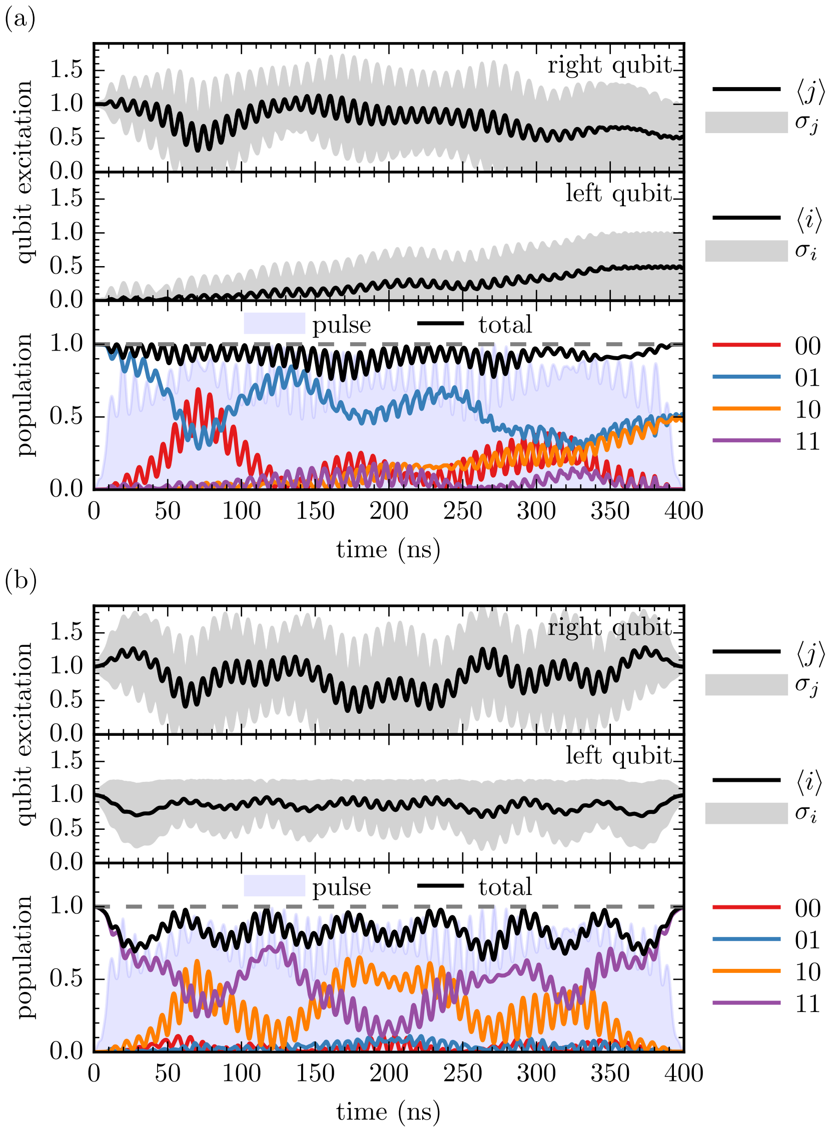


Figure 7.10: Population dynamics for $\hat{\rho}(t=0) = |01\rangle\langle01|$ (a) and $\hat{\rho}(t=0) = |11\rangle\langle11|$ (b) under the pulse shown in Fig. 7.9. For each of the two propagated states, the expectation value of the right qubit excitation quantum number j is shown in the top panel, with the standard deviation in gray, the expectation value for the corresponding quantum number i for the left qubit is shown in the center panel, and the population dynamics for all the logical subspace states is shown in the bottom panel (colored lines), along with the total population in the logical subspace (black line).

8

Summary and Outlook

8.1 Summary and Conclusions

The material presented in this thesis provides a comprehensive framework for the realization of robust two-qubit quantum gates using numerical optimal control. It covers three aspects:

1. the use of efficient numerical tools for the modeling, simulation, and control of both closed and open quantum systems,
2. the facilitation of robustness with respect to decoherence and classical fluctuations of system parameters through advanced control techniques, and
3. the application of these techniques to two of the leading candidates for the implementation of universal quantum computers, trapped atoms and superconducting circuits.

The numerical approach to the design of quantum gates allows to go beyond parameter regimes where the system can accurately be described by simple, analytically solvable models. The flip side of this promise is that the design of *efficient* representations, algorithms, and implementations becomes crucial. Some of the fundamental techniques have been presented in chapter 3. An important part of the work in this thesis has consisted of contributing to their implementation in the QDYN Fortran 90 library. The core aspect of the efficient storage and application of Hamiltonians, respectively Liouvillians is to exploit their sparsity. Especially for systems with spatial degrees of freedom, obtaining sparse operators may require the use of spectral and pseudo-spectral representations. For the propagation of the time-dependent Schrödinger equation, respectively the Liouville von-Neumann equation for open quantum systems, the most efficient method is the expansion of the time evolution operator, respectively the dynamical map in a fast-converging polynomial series. For an Hermitian operator, the fastest converging expansion is in Chebychev polynomials. For a non-Hermitian operator (especially a Liouvillian), an expansion in Newton polynomials must be used instead. A memory efficient variation of this method involves the use of restarted Arnoldi iterations. Both the Chebychev propagator and the Newton propagator with restarted Arnoldi are presented in pseudocode in appendix F. For optimal control, algorithms can be divided into gradient-free optimization methods that are applicable to low-dimension control problems, and gradient-based methods that provide significantly faster convergence, but require additional numerical effort to evaluate the gradient, and place restrictions on the types of functionals that can be optimized. The GRAPE/LBFGS algorithm is suitable for coarse discrete parametrizations of the control, whereas Krotov's method guarantees monotonic convergence for quasi-time-continuous controls. These methods, and Krotov's method in particular, form the basis of the results presented in this thesis.

The realization of robust quantum gates can be achieved by using advanced control techniques. A first example of such a technique is the optimization over an ensemble of Hamiltonians to address the issue of classical fluctuations and uncertainties in the system and/or the control. In chapter 4, this has been used to obtain quantum gates for trapped Rydberg atoms that are robust with respect to variations of the amplitude of the control field, and fluctuations in the Rydberg levels caused by stray electromagnetic fields. With respect to the best available analytical schemes, robustness has been increased by more than one order of magnitude. Moreover, this level of robustness can be maintained as pulses are shortened to approach the quantum speed limit, beyond what is achievable using analytic pulses. Lastly, the optimized pulses successfully avoided the effects of spontaneous decay from an intermediary level for the transition to the Rydberg state.

Superconducting qubits provide one of the most versatile platforms for quantum computing. They may be engineered in an almost arbitrary range of parameters, using widely available production techniques. Recent implementations using the transmon design are approaching decoherence times close to 0.1 ms. This makes transmon qubits one of the most promising contenders for quantum computing. Yet, the implementation of high fidelity entangling gates reaching the quantum error correction limit has not been achieved. In chapter 5, the fundamental concepts and recent advances in the realization of two-qubit gates for transmon qubits are reviewed. We have shown that a holonomic phasegate can be implemented by driving the system with an off-resonant drive near the cavity transition which induces a Stark shift on the levels of the logical subspace. Using Krotov's method, we obtain optimized pulses that successfully implement a CPHASE and a CNOT gate, as well as the holonomic phasegate. However, fidelities remain limited by the loss of population from the logical subspace. Addressing this loss of population in order to achieve high fidelity gates below the quantum error correction limit, and taking into account the dominant sources of decoherence will be the focus of future work.

Using an effective model described in chapter 5 for two transmon qubits, a further advanced control technique that aids the realization of robust gates has been illustrated in chapter 6: Based on the fundamental insight that any perfect entangler together with single qubit operations is sufficient for universal quantum computing, the application of a functional that optimizes for an arbitrary perfect entangler has been demonstrated. Avoiding the use of an overly specific optimization functional, such as the optimization for a "standard" quantum gate such as CNOT or iSWAP, allows the optimal control method to find the entangling gate that is easiest to realize. This becomes especially relevant as constraints are added to the optimization. The optimization functional is based on the geometric theory of two-qubit gates, reviewed in section 2.4 of chapter 2. Any two-qubit gate can be mapped to a point in the Weyl chamber. Gates that map to the same point are identical up to single-qubit operations. The set of perfect entanglers form a polyhedron inside the Weyl chamber. In earlier work [238, 193], the optimization towards a specific point in the Weyl chamber has been demonstrated, allowing the optimization towards a gate up to arbitrary single-qubit operations. The results of chapter 6 have demonstrated that an optimization towards the surface of the polyhedron of perfect entanglers is significantly faster and reaches better fidelities than the optimization towards a specific point in the Weyl chamber, and thus a specific two-qubit gate. There are two aspects under which the optimization towards an arbitrary perfect entangler benefits the implementation of gates that are robust with respect to dissipation. First, the optimization can find the perfect entangler that can be realized in the shortest amount of time, ideally shorter than the relevant dissipation processes. Second, the presence of decoherence effectively places a constraint on the system. For example, dissipation rates increase with higher excitation numbers. Thus, in order to minimize the effects of decoherence, it is beneficial to implement a quantum gate with the least excitation. Optimizing for a general perfect entangler has the potential to automatically identify the quantum gate that is least affected by decoherence. This remains to be demonstrated in future work.

In order to allow the optimization to avoid the effects of decoherence, the dissipation must be explicitly included in the equations of motion. Modeling the dynamics in Liouville space greatly exacerbates the numerical challenges. Instead of Hilbert space vectors of dimension d, every state is now a density matrix of dimension d^2. This both significantly increases the required storage for storing all propagated states, and the CPU time required in propagation, as all matrix-vector operations in Hilbert space become matrix-matrix operations in Liouville space. The optimization of a two-qubit gate in Hilbert space requires the propagation of the four logical basis states $\{\left|00\right\rangle, \left|01\right\rangle, \left|10\right\rangle, \left|11\right\rangle\}$. A naive extension of the standard functionals to Liouville space would require the propagation of the 16 matrices that form the basis of the two-qubit Liouville logical subspace.

Chapter 7 has demonstrated that the propagation of the full Liouville space basis is not necessary for the optimization of a unitary gate, but that a reduced set of states can be employed. In general, the minimum number of states that need to be propagated is 3. In situations where the Hamiltonian only allows for a subset of two-qubit gates to be implemented, this number may reduce further. For example, the Hamiltonian for two trapped Rydberg atoms, chapter 4, only allows for diagonal gates. We have demonstrated that in this case, propagation of only two density matrices is sufficient to optimize for a CPHASE gate. For the general case, we have considered again the effective model for two coupled transmon qubits. There, optimization has been shown to be successful using three states. However, in order to achieve an efficiency comparable to the propagation of the full basis, the three states must be properly weighted. We have shown that convergence can be further improved using a set of 5, respectively 8 states. In all cases, considerable savings both in memory and CPU time have been demonstrated, addressing the issue of numerical efficiency of the optimization of quantum gates in Liouville space.

8.2 Future Work

In order to achieve the ultimate goal of implementing a universal set of quantum gates for superconducting qubits at high fidelity with robustness against decoherence, several issues must yet be addressed. In a full model, Eq. (5.18), fidelities have been limited by the loss of population from the logical subspace caused by high cavity excitation in particular (chapter 5). These excitations necessitate a large Hilbert space dimension $d > 3000$, making optimization in the corresponding Liouville space (dimension d^2) numerically infeasible. Thus, no pulses optimized with respect to dissipation have been obtained.

There are several approaches to improving the fidelities in this case. First, the gradient employed in Krotov's method has been demonstrated to become small when close to the optimum [146], yielding slow asymptotic convergence. An optimization algorithm using second order information, such as LBFGS, can be more effective in this region, allowing to return population fully to the logical subspace. However, an even better approach is to avoid large excitations in the first place. A state-dependent constraint [156] can be added to the optimization functional in order to suppress high cavity population. Alternatively, frequencies driving unwanted transitions can be avoided in the optimized pulses using spectral constraints [158, 159]. To avoid high excitation of the qubit levels, the DRAG technique [252, 252] could be used, adding a second pulse that suppresses excitation out of the logical subspace.

Eliminating large excitations allows to truncate the Hilbert space at significantly smaller dimension, making it feasible to describe the system in Liouville space and to take dissipative effects into account in the optimization explicitly. Decay and dephasing rates for higher levels of the transmon qubit have recently been obtained [253], allowing to model the dynamics as a master equation in Lindblad form. We can apply the full range of techniques presented in this thesis to obtain maximally robust gates. Optimization for a general perfect entangler (chapter 6) allows to identify the gate that can be implemented in the shortest amount of time in order to beat decoherence. Numerical effort is kept minimal by propagating only a reduced set of states (chapter 7). Also, one could consider using ensemble optimization (chapter 4) to obtain further robustness against fluctuations of the system parameters, e.g. the frequency of the transmon qubits.

An alternative possibility to reduce the dimension of the Hilbert space is to employ an effective model that eliminates the cavity from the description of the system. For a simplified effective model, Eq. (5.30), gates of arbitrary fidelity in the non-dissipative case, and gate fidelities limited only by the decoherence in the general case have been attained. However, we have found the pulses obtained from this model not to be transferable to the full model of Eq. (5.18). Even for the properly derived first-order effective Hamiltonian of Eq. (5.29), agreement between the full and effective model is limited to small pulse amplitudes [254]. In order to to obtain optimized pulses that are exchangeable between the effective and the full model, at high fidelity, a rigorous derivation to higher order perturbation theory is necessary. The use of a computer algebra system may be required for this. In this way, full correspondence between the parameters of the effective and the full model can be obtained, and the validity of any approximations made in the derivation ensured. The transformation yielding the effective Hamiltonian can also be applied to the full master equation [255, 256]. Decay of the cavity leads to an additional effective dephasing of the qubit in this case.

Specifically for the holonomic gate considered in chapter 5, a drive that is only slightly detuned from the cavity induces a significant Stark shift in the cavity levels. In the non-dressed basis, this is reflected in a large cavity excitation, see Fig 5.5. Therefore, it is not possible to simply suppress cavity population with a state-dependent constraint. However, in principle, an effective model can be derived through a series of transformations [257]. Most importantly, these include a Polaron transformation

$$D(\alpha) = \exp\left(\alpha \hat{\mathsf{a}}^{\dagger} - \alpha^{*} \hat{\mathsf{a}}\right) \tag{8.1}$$

that describes the displacement of the cavity by α, which is induced by the field $\Omega(t)$. With a properly chosen value for α depending on $\Omega(t)$ the driving term for the cavity is eliminated from the Hamiltonian in the shifted frame [229], allowing to truncate the Hilbert space to a smaller dimension.

One of the fundamental promises of numerical optimal control is the ability to go beyond the regime of simple analytically solvable models. Thus, it would be possible to explore parameter regimes beyond those of the standard Jaynes-Cummings model for the qubit-cavity interaction, Eq. (5.18). For example, in the ultra-strong coupling regime, Eq. (5.15) would be violated, and the interaction Hamiltonian would take the form of Eq. (5.13). A master equation for superconducting qubits in the ultrastrong coupling regime has been derived in Ref. [218]. It might even be feasible to go beyond the Duffing oscillator model, Eq. (5.11), and to model the qubit state as a wave-packet in the cosine potential of Eq. (5.7).

While with the approaches discussed above we can fully expect to obtain high fidelity quantum gates for transmon qubits, including qubit decay and dephasing as well as cavity decay will place limits on the achievable fidelity, cf. chapter 7. In this case, we are limited to using optimal control to implement quantum gates on a shorter time scale than the decoherence. Furthermore, we can minimize – but not eliminate – the effects of decoherence, e.g. by avoiding cavity population. Microscopically, decoherence in superconducting qubits has been linked to dielectric defects as a primary source [258, 259], which may be modeled as two-level systems. While weak coupling leads phenomenologically to decay and dephasing and a master equation of Lindblad form, for strong coupling the dynamics can be non-Markovian [260, 261, 262]. That is, the conditions listed in section 2.6.4 of chapter 2 are violated and the dynamics can no longer be treated with a master equation in Lindblad form. Remarkably, while in the Markovian regime, information that passes from the system to the environment through dissipation is irretrievably lost, non-Markovian dynamics allow for a backflow of information from the environment to the system [99, 100]. This opens up new possibilities for control, and may in fact aid in the implementation of quantum gates. For example, it has recently been shown that if the qubit strongly couples to just a few of the bath spins, which in turn couple weakly to the rest of the environment, the non-Markovianity can be exploited for the realization of quantum gates [243].

Numerical optimal control has previously been applied to non-Markovian quantum systems [263, 264, 265, 266], provided the dynamics can be calculated with sufficient efficiency. One possibility to tackle the simulation of the system's time evolution is by a renormalization approach [267]. It is important to note that none of the techniques presented in this thesis are limited to a specific, e.g. Markovian, dissipation model or a specific equation of motion. They may thus be used to explore a wider range of physics that may lead to entirely new possibilities of quantum control.

A

Interaction between an Atom and a Laser Field

We consider an electromagnetic field with the vector and electrostatic potential

$$\vec{A}(\vec{r},t) = \frac{E_0}{\omega}\vec{e}_z \sin\left(ky-\omega t\right) \,, \tag{A.1}$$

$$\Phi(\vec{r},t) = 0\,. \tag{A.2}$$

This corresponds to a cosine shape electromagnetic field propagating along the y-axis, with the electric field oscillating in z-direction and the magnetic field oscillating in y-direction,

$$\vec{E}(\vec{r},t) = -\frac{\partial}{\partial t}\vec{A}(\vec{r},t) - \vec{\nabla}\Phi(\vec{r},t) = E_0\vec{e}_z\cos(ky-\omega t)\,, \tag{A.3}$$

$$\vec{B}(\vec{r},t) = \vec{\nabla}\times\vec{A}(\vec{r},t) = \frac{\partial}{\partial y}A_z\vec{e}_x = B_0\vec{e}_x\cos(ky-\omega t)\,, \tag{A.4}$$

with $B_0 = \frac{E_0}{\omega}k = \frac{E_0}{c}$, where $k=\frac{\omega}{c}$, c is the speed of light and ω is the laser frequency.

The Hamiltonian for an atom's valence electron at position $\vec{r}$, with electron mass m, electron-charge q, and $\vec{r}$ and $\vec{p}$ now being operators, reads

$$\begin{aligned}\hat{\mathsf{H}} &= \frac{1}{2m}\left[\vec{\mathsf{p}} - q\vec{\mathsf{A}}(\vec{\mathsf{r}},t)\right]^2 + \hat{\mathsf{V}}(\vec{\mathsf{r}}) - \frac{q}{m}\vec{\mathsf{S}}\cdot\vec{\mathsf{B}}(\vec{\mathsf{r}},t) + \vec{\nabla}\hat{\Phi}(\vec{\mathsf{r}},t)\\ &= \hat{\mathsf{H}}_0 - \frac{q}{m}\vec{\mathsf{p}}\cdot\vec{\mathsf{A}} - \frac{q}{m}\vec{\mathsf{S}}\cdot\vec{\mathsf{B}} + \frac{q}{2m}\left[\vec{\mathsf{A}}(\vec{\mathsf{r}},t)\right]^2\,,\end{aligned} \tag{A.5}$$

with the electron's drift Hamiltonian

$$\hat{\mathsf{H}}_0 = \frac{\vec{\mathsf{p}}^2}{2m} + \hat{\mathsf{V}}(\vec{\mathsf{r}})\,, \tag{A.6}$$

and the spin operator $\vec{\mathsf{S}}$ coupling to the magnetic field. The origin of the coordinate system is in the atoms nucleus. Since $\left\|\vec{A}^2\right\| \ll \left\|\vec{A}\right\|$ for realistic laser field amplitudes, we set the last term to zero.

The spatial dependence of the vector potential, $ky = \frac{2\pi y}{\lambda}$, where y is on the order of an atomic radius a_0 and λ is the wavelength of the laser is extremely small. We can therefore Taylor-expand the vector potential as

$$\begin{aligned} \vec{A}(\vec{r},t) &= \frac{E_0}{w}\vec{e}_z \sin(ky - \omega t) \\ &= \frac{E_0}{2\mathrm{i}\omega}\vec{e}_z \left(\mathrm{e}^{\mathrm{i}ky}\mathrm{e}^{-\mathrm{i}\omega t} - \mathrm{e}^{-\mathrm{i}ky}\mathrm{e}^{\mathrm{i}\omega t}\right) \\ &\approx \frac{E_0}{2\mathrm{i}\omega}\vec{e}_z \left((1+\mathrm{i}ky)\mathrm{e}^{-\mathrm{i}\omega t} - (1-\mathrm{i}ky)\mathrm{e}^{\mathrm{i}\omega t}\right) \\ &= \frac{E_0}{w}\vec{e}_z \sin(\omega t) + B_0 y \vec{e}_z \cos(\omega t)\,. \end{aligned} \tag{A.7}$$

Also, since an electron that is localized with a Bohr radius a_0 must have a minimum momentum $\vec{\mathsf{p}}$ such that $\frac{\hbar}{p} \leq a_0$, and $\vec{\mathsf{S}}$ is on the order of $\hbar$, we can show $\left\|\vec{\mathsf{p}}\cdot\vec{\mathsf{A}}\right\| \gg \left\|\vec{\mathsf{S}}\cdot\vec{\mathsf{B}}\right\|$,

$$\frac{\left\|\vec{\mathsf{S}}\cdot\vec{\mathsf{B}}\right\|}{\left\|\vec{\mathsf{p}}\cdot\vec{\mathsf{A}}\right\|} \approx \frac{\hbar k E_0/\omega}{pE_0/\omega} = \frac{\hbar k}{p} < \frac{a_0}{\lambda} \ll 1\,. \tag{A.8}$$

Therefore, we are justified in approximating

$$\vec{B}(\vec{r},t) \approx B_0 \vec{e}_x \cos\omega t\,. \tag{A.9}$$

Inserting this and Eq. (A.7) into Eq. (A.5) yields

$$\hat{\mathsf{H}} \approx \hat{\mathsf{H}}_0 - \frac{q}{m}\frac{E_0}{\omega}\hat{\mathsf{p}}_z \sin(\omega t) - \frac{qB_0}{m}B_0\hat{\mathsf{p}}_z\hat{\mathsf{y}}\cos(\omega t) - \frac{q}{2m}\hat{\mathsf{S}}_x B_0\cos(\omega t)\,. \tag{A.10}$$

Furthermore,

$$\hat{\mathsf{p}}_z\hat{\mathsf{y}} = \frac{1}{2}\left(\hat{\mathsf{p}}_z\hat{\mathsf{y}} - \hat{\mathsf{z}}\hat{\mathsf{p}}_y\right) + \frac{1}{2}\left(\hat{\mathsf{p}}_z\hat{\mathsf{y}} - \hat{\mathsf{z}}\hat{\mathsf{p}}_y\right) = \frac{1}{2}\hat{\mathsf{L}}_x + \frac{1}{2}\left(\hat{\mathsf{p}}_z\hat{\mathsf{y}} - \hat{\mathsf{z}}\hat{\mathsf{p}}_y\right)\,, \tag{A.11}$$

resulting in

$$\begin{aligned} \hat{\mathsf{H}} \approx \hat{\mathsf{H}}_0 &- \frac{q}{m}\frac{E_0}{\omega}\hat{\mathsf{p}}_z \sin(\omega t) - \frac{q}{2mc}E_0\cos(\omega t)\left[\hat{\mathsf{p}}_z\hat{\mathsf{y}} - \hat{\mathsf{z}}\hat{\mathsf{p}}_y\right] + \\ &- \frac{q}{2m}\left(\hat{\mathsf{L}}_x + \hat{\mathsf{S}}_x\right)B_0\cos(\omega t)\,. \end{aligned} \tag{A.12}$$

The three interaction terms are interpreted as follows:

- $$\hat{\mathsf{H}}_{ED} = \frac{q}{m}\frac{E_0}{w}\hat{\mathsf{p}}_z \sin(\omega t) \tag{A.13}$$

  is the *electric dipole* interaction in momentum space. It can be rewritten to its more familiar form in coordinate space

  $$\hat{\mathsf{H}}_{ED} = q\hat{\mathsf{z}}\, E_0 \cos(\omega t) = \hat{\mu} E_0 \cos(\omega t)\,, \tag{A.14}$$

  where $\hat{\mu}$ has been introduced as the dipole operator.

- $$\hat{\mathsf{H}}_{EQ} = -\frac{q}{2mc} E_0 \cos(\omega t) \left[\hat{\mathsf{p}}_z \hat{\mathsf{y}} - \hat{\mathsf{z}}\hat{\mathsf{p}}_y\right] \tag{A.15}$$

  describes the *electric quadrupole* interaction.

- $$\hat{\mathsf{H}}_{MD} = -\frac{q}{2m}\left(\hat{\mathsf{L}}_x + \hat{\mathsf{S}}_x\right) B_0 \cos(\omega t) \tag{A.16}$$

  describes the *magnetic dipole* interaction.

Both the electric quadrupole and the magnetic dipole are negligible compared to the electric dipole. Therefore, in the dipole-approximation the total Hamiltonian becomes

$$\hat{\mathsf{H}} \approx \hat{\mathsf{H}}_0 + \hat{\mu} E(t)\,. \tag{A.17}$$

The dipole approximation results from the assumption that the wavelength of the laser is much larger than the width of the atom, and thus that the spatial dependence of the field can be dropped, allowing to define the z-component of the electric field as

$$E(t) = E_0 \cos(\omega t)\,. \tag{A.18}$$

When Eq. (A.17) is written in the energy representation given by the eigenstates of $\hat{\mathsf{H}}_0$, the selection rules for the dipole transitions are obtained. That is, for certain quantum numbers the corresponding matrix element of $\hat{\mu}$ will vanish. For a Hydrogen atom with eigenstates $|nlm\rangle$, the dipole is zero unless $\Delta l = 1$ and $\Delta m = 0, 1$.

For example, we may consider a Hamiltonian for a sub-system consisting of three levels $|0\rangle$, $|1\rangle$, and $|2\rangle$, with energies E_0, E_1, E_2. The dipole transition $|0\rangle \rightarrow |1\rangle$ and $|2\rangle \rightarrow |3\rangle$ is allowed with a resulting dipole moment of $\mu_{01} = \langle 0 \,|\, \hat{\mu} \,|\, 1\rangle$ and $\mu_{12} = \langle 1 \,|\, \hat{\mu} \,|\, 2\rangle$, respectively, but $|1\rangle \rightarrow |3\rangle$ is forbidden. This Hamiltonian would we written in the energy representation as

$$\hat{\mathsf{H}} = \begin{pmatrix} E_0 & \mu_{01}E(t) & 0 \\ \mu_{01}E(t) & E_1 & \mu_{12}E(t) \\ 0 & \mu_{12}E(t) & E_2 \end{pmatrix}\,, \tag{A.19}$$

the form used for the Hamiltonians e.g. in chapters 4 and 7.

B

The Rotating Wave Approximation

In many cases, we can simplify the Hamiltonian and make it analytically tractable by transforming it from the *lab frame* to the *rotating frame* oscillating at the laser frequency ω_L. The rotating frame is defined by a time-dependent unitary transformation $\hat{\mathsf{U}}(t)$. Every state $|\Psi\rangle$ in the lab frame is transformed to the corresponding state in the rotating frame as

$$|\tilde{\Psi}(t)\rangle = \hat{\mathsf{U}}(t)|\Psi(t)\rangle \tag{B.1}$$

In order to derive the Hamiltonian in the rotating frame, we demand the Schrödinger equation to be fulfilled,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\tilde{\Psi}\rangle = \tilde{\mathsf{H}}|\tilde{\Psi}\rangle\,, \tag{B.2}$$

which leads to

$$\begin{aligned}\mathrm{i}\hbar\frac{\partial}{\partial t}|\tilde{\Psi}\rangle &= \mathrm{i}\hbar\dot{\mathsf{U}}|\Psi\rangle + \hat{\mathsf{U}}\mathrm{i}\hbar\frac{\partial}{\partial t}|\Psi\rangle \\ &= \mathrm{i}\hbar\dot{\mathsf{U}}|\Psi\rangle + \hat{\mathsf{U}}\hat{\mathsf{H}}|\Psi\rangle \\ &= \mathrm{i}\hbar\dot{\mathsf{U}}\hat{\mathsf{U}}^{\dagger}|\tilde{\Psi}\rangle + \hat{\mathsf{U}}\hat{\mathsf{H}}\hat{\mathsf{U}}^{\dagger}|\tilde{\Psi}\rangle \\ &= \left(\mathrm{i}\hbar\dot{\mathsf{U}}\hat{\mathsf{U}}^{\dagger} + \hat{\mathsf{U}}\hat{\mathsf{H}}\hat{\mathsf{U}}^{\dagger}\right)|\tilde{\Psi}\rangle.\end{aligned} \tag{B.3}$$

So, the transformed Hamiltonian is

$$\tilde{\mathsf{H}} = \mathrm{i}\hbar\dot{\mathsf{U}}\hat{\mathsf{U}}^{\dagger} + \hat{\mathsf{U}}\hat{\mathsf{H}}\hat{\mathsf{U}}^{\dagger}. \tag{B.4}$$

For a system of two dipole-coupled levels separated by energy $\hbar\omega_1$, driven by a pulse $\epsilon(t)$, the Hamiltonian reads

$$\hat{\mathsf{H}} = \begin{pmatrix} 0 & \mu\epsilon(t) \\ \mu\epsilon(t) & \hbar\omega_1 \end{pmatrix} , \tag{B.5}$$

where μ is the dipole strength. We consider the pulse

$$\epsilon(t) = S(t)\cos(\omega_L t) = \frac{S(t)}{2}\left(\mathrm{e}^{\mathrm{i}\omega_L t} + \mathrm{e}^{-\mathrm{i}\omega_L t}\right) , \tag{B.6}$$

with frequency ω_L and a slowly varying shape $S(t)$. The rotating frame for this pulse is defined by

$$\hat{\mathsf{U}}(t) = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{i\omega_L t} \end{pmatrix} . \tag{B.7}$$

Applying Eq. (B.4) yields

$$\tilde{\mathsf{H}} = \begin{pmatrix} 0 & \mu\epsilon(t)\mathrm{e}^{-\mathrm{i}\omega_L t} \\ \mu\epsilon(t)\mathrm{e}^{+\mathrm{i}\omega_L t} & \hbar(\omega_1 - \omega_L) \end{pmatrix} . \tag{B.8}$$

The energy level has been shifted down by ω_L, resulting from the term $\mathrm{i}\hbar\dot{\hat{\mathsf{U}}}\hat{\mathsf{U}}^\dagger$, and the couplings obtain a time-dependent phase-factor, due to the term $\hat{\mathsf{U}}\hat{\mathsf{H}}\hat{\mathsf{U}}^\dagger$. Up to this point, the transformation is exact. For the off-diagonal terms, we now find

$$\mu\epsilon(t)\mathrm{e}^{\pm\mathrm{i}\omega_L t} = \frac{\mu}{2}\epsilon(t)\left(1 + \mathrm{e}^{\pm 2\mathrm{i}\omega_L t}\right) \approx \frac{\mu}{2}S(t) \equiv \Omega(t) . \tag{B.9}$$

The approximation is valid for $\omega_L \gg 1$, where the fast oscillations at twice the laser frequency average out, leaving only the slowly varying pulse shape. With the detuning $\Delta \equiv \hbar(\omega_1 - \omega_L)$, the RWA-Hamiltonian is therefore

$$\hat{\mathsf{H}}_{\text{RWA}} = \begin{pmatrix} 0 & \Omega(t) \\ \Omega(t) & \Delta \end{pmatrix} . \tag{B.10}$$

In chapter 4, a cesium atom was considered where the ground state $|0\rangle$ is excited to the Rydberg state $|r\rangle$ by a two-photon transition via the intermediary level $|i\rangle$, see Fig. 4.2. In the lab frame, the Hamiltonian reads

$$\hat{\mathsf{H}} = \begin{pmatrix} 0 & \mu_B\epsilon_B(t) & 0 \\ \mu_B\epsilon_B(t) & \hbar\omega_i & \mu_R\epsilon_R(t) \\ 0 & \mu_R\epsilon_R(t) & \hbar\omega_r \end{pmatrix} , \tag{B.11}$$

with the red and blue laser field

$$\epsilon_{R,B}(t) = S_{R,B}(t)\cos(\omega_{R,B}t). \tag{B.12}$$

The appropriate two-color RWA is defined by

$$\hat{\mathsf{U}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\omega_B t} & 0 \\ 0 & 0 & e^{i(\omega_B+\omega_R)t} \end{pmatrix} . \tag{B.13}$$

The transformed Hamiltonian (with $\mu_{R,B}$ absorbed in $\epsilon_{R,B}$) is

$$\begin{aligned} \tilde{\mathsf{H}} &= \begin{pmatrix} 0 & \epsilon_B(t)e^{-i\omega_B t} & 0 \\ \epsilon_B(t)e^{i\omega_B t} & \hbar(\omega_i - \omega_R) & \epsilon_R(t)e^{-i(\omega_R+\omega_B)t+i\omega_B t} \\ 0 & \epsilon_R(t)e^{i(\omega_R+\omega_B)t-i\omega_B t} & \hbar(\omega_r - \omega_R - \omega_B) \end{pmatrix} \\ &\approx \begin{pmatrix} 0 & \Omega_B(t) & 0 \\ \Omega_B(t) & \Delta_1 & \Omega_R(t) \\ 0 & \Omega_R(t) & \Delta_2 \end{pmatrix} , \end{aligned} \tag{B.14}$$

with

$$\Omega_{R,B}(t) = \frac{1}{2}\mu_{R,B}S_{R,B}(t) \tag{B.15}$$

and the single- and two-photon detuning Δ_1 and Δ_2, respectively. This corresponds to Eq. (4.1).

In general, the appropriate RWA-transformation $\hat{\mathsf{U}}$ can be read off from a level diagram; $\hat{\mathsf{U}}$ is always diagonal with entries $e^{i\omega_i t}$, where ω_i is the amount by which the i'th energy level is shifted down. The shifts result from shortening each transition by the pulse frequency that is to be eliminated [189]. For example, starting from the diagram in Fig 4.2, shortening the blue transition means that both $|i\rangle$ and $|r\rangle$ are shifted down by ω_B. Then, shortening the red transition means that $|r\rangle$ is shifted down by an additional amount ω_R, resulting in Eq. (B.13).

C

Rabi-Cycling in the Two-Level System

We consider a two-level system

$$|\Psi(t)\rangle = a(t)\,|0\rangle + b(t)\,|1\rangle \tag{C.1}$$

with time-dependent complex coefficients $a(t)$ and $b(t)$. Under the rotating wave approximation presented in appendix B, the Hamiltonian takes the form

$$\hat{\mathsf{H}} = \begin{pmatrix} 0 & \frac{1}{2}\Omega_0(t) \\ \frac{1}{2}\Omega_0(t) & \Delta \end{pmatrix}\,. \tag{C.2}$$

where $\Omega_0(t)$ is a slowly varying pulse shape and Δ is the detuning of the central pulse frequency from the $|0\rangle \rightarrow |1\rangle$ transition. We first consider a pulse that is a simple continuous oscillation i.e. a constant $\Omega_0(t) \equiv \Omega_0$ in the RWA.

For the initial conditions $a(0) = 1$ and $b(0) = 0$, the Schrödinger equation then has the solution [9]

$$a(t) = \mathrm{e}^{-\frac{\mathrm{i}}{2}\Delta t}\left(\cos\left(\frac{\Omega t}{2}\right) - \mathrm{i}\frac{\Delta}{\Omega}\sin\left(\frac{\Omega t}{2}\right)\right)\,, \tag{C.3a}$$

$$b(t) = -\mathrm{i}\frac{\Omega_0}{\Omega}\mathrm{e}^{-\frac{\mathrm{i}}{2}\Delta t}\sin\left(\frac{\Omega t}{2}\right)\,, \tag{C.3b}$$

with

$$\Omega = \sqrt{\Delta^2 + \Omega_0^2}\,. \tag{C.4}$$

The population undergoes *Rabi oscillations* between $|0\rangle$ and $|1\rangle$ with a period of $\frac{2\pi}{\Omega}$,

$$|a(t)|^2 = \left(\frac{\Delta}{\Omega}\right)^2 + \left(\frac{\Omega_0}{\Omega}\right)^2 \cos^2\left(\frac{\Omega t}{2}\right)\,, \tag{C.5a}$$

$$|b(t)|^2 = \left(\frac{\Omega_0}{\Omega}\right)^2 \sin^2\left(\frac{\Omega t}{2}\right)\,. \tag{C.5b}$$

In the case of non-zero detuning, $\Delta \neq 0$, the population is only transferred partially, but at higher frequency. In the resonant case, $\Delta = 0$, the Rabi-frequency is $\Omega = \Omega_0$, and there is complete population transfer at $t = \frac{\pi}{\Omega}$. The complete coherent transfer of population from $|0\rangle$ to $|1\rangle$ is therefore called a "π-pulse". However, according to Eq. (C.3), it also induces a phase factor of $\mathrm{e}^{\mathrm{i}\frac{\pi}{2}}$,

$$\pi\text{-pulse:} \qquad |0\rangle \rightarrow \mathrm{i}\,|1\rangle\ .$$

Transferring the population up and down again in a full Rabi cycle, or 2π-pulse, yields a phase factor -1,

$$2\pi\text{-pulse:} \qquad |0\rangle \rightarrow -\,|0\rangle\ .$$

In order to restore the original state with a phase of zero, two full cycles are necessary.

For a time-dependent but slowly varying pulse shape $\Omega(t)$, the Rabi angle Ωt in the argument of the sines and cosines in Eq. (C.3) and Eq. (C.5) generalizes to

$$\Omega t \rightarrow \int_0^t \Omega(t')\,\mathrm{d}t'\,.$$

Thus, the amount of population that is transferred depends only on the pulse area, not the specific shape of the pulse.

D

Overview of Two-Qubit Gates

The CNOT gate is the archetypal two-qubit gate in sets of operations for universal quantum computing; it flips the *target* qubit if the *control* qubit is in state $|1\rangle$.

$$\mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{D.1}$$

The CPHASE gate induces a phase shift of γ on the target qubit if the control qubit is in state $|1\rangle$.

$$\mathrm{CPHASE}_\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \mathrm{e}^{\mathrm{i}\gamma} \end{pmatrix} \tag{D.2}$$

The gate is a perfect entangler for $\gamma = \pi$, where it is locally equivalent to CNOT. Indeed, *all* controlled operators are locally equivalent to a CPHASE_γ [91]. We refer to CPHASE_π simply as CPHASE. In the Weyl chamber, the CPHASE_γ gates are on the line O–A_1.

The SWAP gate exchanges the two qubits. The gates at the A_3 point in the Weyl chamber are the only true two-qubit gates that yield zero entanglement.

$$\text{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{D.3}$$

The $\sqrt{\text{SWAP}}$, located at the point P in the Weyl chamber, however, is a perfect entangler, indicating that a SWAP gate is implemented by first entangling and then disentangling the two qubits.

$$\sqrt{\text{SWAP}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} - \frac{\mathrm{i}}{2} & \frac{1}{2} + \frac{\mathrm{i}}{2} & 0 \\ 0 & \frac{1}{2} + \frac{\mathrm{i}}{2} & \frac{1}{2} - \frac{\mathrm{i}}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{D.4}$$

A secondary square root of SWAP is located at the N point; it is simply the complex conjugate of the principal square root, Eq. (D.4), and we thus label is as $\sqrt{\text{SWAP}}^*$.

The iSWAP gate performs a SWAP, with and additional relative phase shift of π. The gate is also known as DCNOT (Double-CNOT), since is implemented by two consecutive CNOT gates, where the control qubit for the second CNOT is the target qubit of the first CNOT.

$$\text{iSWAP} = \text{DCNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{D.5}$$

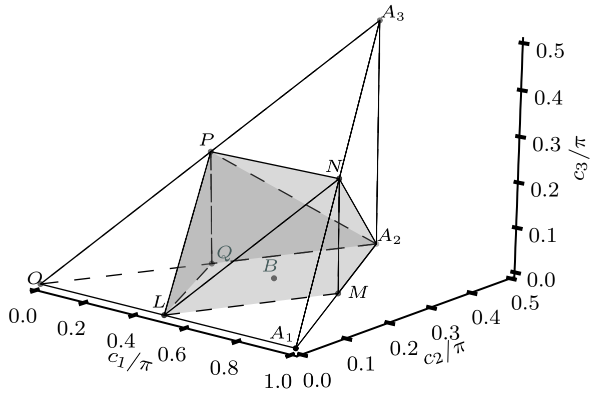
The principal square root of iSWAP is located at the Q point in the Weyl chamber.

$$\sqrt{\text{iSWAP}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{\mathrm{i}}{\sqrt{2}} & 0 \\ 0 & \frac{\mathrm{i}}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{D.6}$$

Note that none of the gates at the M point are square roots of the exact iSWAP, even though their square is still locally equivalent to iSWAP.

The B-GATE is at the center of the perfect entanglers. It has been shown to be extremely efficient for the construction of arbitrary two-qubit gates [89].

$$\text{B-GATE} = \begin{pmatrix} \cos\frac{\pi}{8} & 0 & 0 & \mathrm{i}\sin\frac{\pi}{8} \\ 0 & \cos\frac{3\pi}{8} & \mathrm{i}\sin\frac{3\pi}{8} & 0 \\ 0 & \mathrm{i}\sin\frac{3\pi}{8} & \cos\frac{3\pi}{8} & 0 \\ \mathrm{i}\sin\frac{\pi}{8} & 0 & 0 & \cos\frac{\pi}{8} \end{pmatrix} \tag{D.7}$$



Gate	Hamiltonian	c_1	c_2	c_3	W	g_1	g_2	g_3
$\mathbb{1}$	(single qubit gates)	0	0	0	O	1	0	3
		π	0	0	A_1	1	0	3
CNOT	$\hat{\sigma}_z^{(1)} + \hat{\sigma}_x^{(2)} - \hat{\sigma}_z\hat{\sigma}_x$	$\frac{\pi}{2}$	0	0	L	0	0	1
CPHASE$_\gamma$	$\hat{\sigma}_z^{(1)} + \hat{\sigma}_z^{(2)} - \hat{\sigma}_z\hat{\sigma}_z$	$\frac{\gamma}{2}$	0	0		$g_1(\gamma)$	0	$g_3(\gamma)$
iSWAP DCNOT	$\hat{\sigma}_x\hat{\sigma}_x + \hat{\sigma}_y\hat{\sigma}_y$ $= \frac{1}{2}(\hat{\sigma}_+\hat{\sigma}_- + \hat{\sigma}_-\hat{\sigma}_+)$	$\frac{\pi}{2}$	$\frac{\pi}{2}$	0	A_2	0	0	-1
$\sqrt{\text{iSWAP}}$	$\hat{\sigma}_x\hat{\sigma}_x + \hat{\sigma}_y\hat{\sigma}_y$	$\frac{\pi}{4}$	$\frac{\pi}{4}$	0	Q	$\frac{1}{4}$	0	1
SWAP	$\hat{\sigma}_x\hat{\sigma}_x + \hat{\sigma}_y\hat{\sigma}_y + \hat{\sigma}_z\hat{\sigma}_z$	$\frac{\pi}{2}$	$\frac{\pi}{2}$	$\frac{\pi}{2}$	A_3	-1	0	-3
$\sqrt{\text{SWAP}}$	$\hat{\sigma}_x\hat{\sigma}_x + \hat{\sigma}_y\hat{\sigma}_y + \hat{\sigma}_z\hat{\sigma}_z$	$\frac{\pi}{4}$	$\frac{\pi}{4}$	$\frac{\pi}{4}$	P	0	$\frac{1}{4}$	0
$\sqrt{\text{SWAP}}^*$	$-\hat{\sigma}_x\hat{\sigma}_x - \hat{\sigma}_y\hat{\sigma}_y - \hat{\sigma}_z\hat{\sigma}_z$	$\frac{3\pi}{4}$	$\frac{\pi}{4}$	$\frac{\pi}{4}$	N	0	$-\frac{1}{4}$	0
B-GATE	$2\hat{\sigma}_x\hat{\sigma}_x + \hat{\sigma}_y\hat{\sigma}_y$	$\frac{\pi}{2}$	$\frac{\pi}{4}$	0	B	0	0	0
M-GATE	$3\hat{\sigma}_x\hat{\sigma}_x + \hat{\sigma}_y\hat{\sigma}_y$	$\frac{3\pi}{4}$	$\frac{\pi}{4}$	0	M	$\frac{1}{4}$	0	1

Table D.1: Summary of two-qubit gates at special points in the Weyl chamber (shown at the top, with the polyhedron of perfect entanglers indicated by the shaded area). For each gate, the Hamiltonian generating that gate up to a global phase is given in terms of the Pauli matrices, where $\hat{\sigma}_i^{(1,2)}$ indicates an operator acting only on the first and second qubit, respectively, and $\hat{\sigma}_i\hat{\sigma}_j$ is a shorthand for $\hat{\sigma}_i^{(1)} \otimes \hat{\sigma}_j^{(2)}$. Also, the Weyl coordinates c_1, c_2, c_3, the name of the respective point in the Weyl chamber, and the local invariants g_1, g_2, g_3 are listed. The Weyl chamber coordinates for the controlled phase gate are $g_1(\gamma) = \cos^2\frac{\gamma}{2}$ and $g_3(\gamma) = 1 + 2\cos^2\frac{\gamma}{2}$.

E

Applications of the Fast-Fourier-Transform

The Fast-Fourier-Transform [268] is a versatile numerical tool. It is an efficient implementation of the discrete Fourier transform that maps between a function in a canonical variable and a function in the conjugate variable, e.g. space (x) and wavenumber (k, equal to momentum with $\hbar = 1$), or time (t) and angular frequency (ω). For these two examples, the discrete Fourier transform F of a function f reads as

$$F(\omega_j) = \sum_{n=0}^{N-1} f(t_n) \mathrm{e}^{-\mathrm{i}\frac{2\pi j n}{N}} \approx \sum_{n=0}^{N-1} f(t_n) \mathrm{e}^{-\mathrm{i}\omega_j t_n} \,, \tag{E.1}$$

$$F(k_j) = \sum_{n=0}^{N-1} f(x_n) \mathrm{e}^{-\mathrm{i}\frac{2\pi j n}{N}} \approx \sum_{n=0}^{N-1} f(x_n) \mathrm{e}^{-\mathrm{i}k_j x_n} \,, \tag{E.2}$$

where N is the sampling size, and using the correspondence

$$w_j t_n \approx \frac{2\pi j n}{N} \,, \qquad k_j x_n \approx \frac{2\pi j n}{N} \,. \tag{E.3}$$

This transformation is widely implemented in numerical libraries as $F(\omega_j) \equiv \mathrm{FFT}\, f(t_n)$ with the inverse transform defined by

$$f(t_n) = \frac{1}{N} \mathrm{FFT}^{-1}\, F(\omega_j) = \frac{1}{N} \mathrm{FFT}^{-1}\, \mathrm{FFT}\, f(t_n) \,. \tag{E.4}$$

The FFT scales as $N \log N$ with the sampling size [269].

E.1 The Frequency Grid

When using the output of the FFT routine, e.g. as the spectrum of a time-dependent signal $f(t_n)$, it is important to understand exactly which angular frequency values ω_j the resulting amplitudes $F(\omega_j)$ correspond to. This differs between odd and even N, a detail that is often neglected and can lead to subtle errors in numerical calculations.

If the original signal is of duration $T = t_{N-1} - t_0$, and has a sampling rate $dt = \frac{T}{N-1}$, the result of a call to the FFT routine is an array of N complex numbers, containing the amplitudes for angular frequencies between $-\omega_{\max}$ and $+\omega_{\max}$, with

$$\omega_{\max} = \begin{cases} \frac{\pi}{dt} = \frac{(N-1)\pi}{T} & \text{if } N \text{ even} \\ \frac{N-1}{N}\frac{\pi}{dt} = \frac{N-1}{N}\frac{(N-1)\pi}{T} & \text{if } N \text{ odd} \end{cases} \tag{E.5}$$

The layout of the frequency array also depends on whether N is odd or even. In any case, the frequency array consists of two parts: the first sub-array of length l contains the amplitudes of the positive frequencies, the remaining sub-array of length $N - l$ (running from $l + 1$ to N) contains the amplitudes for the negative frequencies.

- N even.
  For even N, there are $l = N/2$ positive frequencies, and the values correspond to
  $$0, d\omega, \ldots, \omega_{\max} - d\omega,$$
  followed by
  $$-\omega_{\max}, -\omega_{\max} + d\omega, \ldots, -d\omega.$$
- N odd.
  For odd N, there are $l = N/2 + 1$ positive frequencies, and the values correspond to
  $$0, d\omega, \ldots, \omega_{\max},$$
  followed by
  $$-\omega_{\max}, -\omega_{\max} + d\omega, \ldots, -d\omega.$$

The spectral resolution is

$$d\omega = \begin{cases} \frac{2\omega_{\max}}{N} = \frac{2\pi}{T} - \frac{2\pi}{NT} & \text{if } N \text{ even} \\ \frac{2\omega_{\max}}{N-1} = \frac{2\pi}{T} & \text{if } N \text{ odd} \end{cases}. \tag{E.6}$$

With $t_n = n\frac{T}{N}$ and $\omega_j = j\,d\omega$, Eq. (E.3) is recovered. If $f(t) \in \mathbb{R}$, then $F(\omega) = F(-\omega)^*$. For complex signals, on the other hand, the positive and the negative part of the spectrum are *not* equivalent.

E.2 Derivatives and the Kinetic Operator

The FFT can be used to calculate the derivative of a signal $f(x)$ evenly sampled at N points.

$$\frac{\partial}{\partial x} \equiv \frac{1}{N}\,\mathrm{FFT}^{-1}\,(\mathrm{i}k)\,\mathrm{FFT} \tag{E.7}$$

The kinetic operator in one Cartesian dimension is

$$\hat{\mathsf{T}} = \frac{\hat{\mathsf{p}}^2}{2m} = -\frac{\hbar}{2m}\frac{\partial^2}{\partial x^2}\,. \tag{E.8}$$

Using the FFT, this becomes

$$\begin{aligned} \hat{\mathsf{T}} &= -\frac{\hbar^2}{2m}\frac{1}{N}\,\mathrm{FFT}^{-1}\,\mathrm{i}k\,\mathrm{FFT}\,\frac{1}{N}\,\mathrm{FFT}^{-1}\,\mathrm{i}k\,\mathrm{FFT} \\ &= \frac{\hbar^2}{2m}\frac{1}{N}\,\mathrm{FFT}^{-1}\,k^2\,\mathrm{FFT}\,. \end{aligned} \tag{E.9}$$

E.3 Cosine-Transform and Chebychev Coefficients

For $f(x) = e^x$, the Chebychev coefficients can be derived analytically to be proportional to the Bessel functions, see section 3.2.1 in chapter 3. For a general function, however, the coefficients are calculated via a cosine transform [129]. Since the Chebychev polynomials P_n are defined only in the interval $[-1, 1]$, the function $f(x)$ must be rescaled as $f(x) = \tilde{f}(\xi)$ with $\xi \in [-1, 1]$. The coefficients are then calculated as

$$a_n = \frac{2-\delta_{n,0}}{\pi}\int_{-1}^{+1}\frac{\tilde{f}(\xi)P_n(\xi)}{\sqrt{1-\xi^2}}\,\mathrm{d}\xi \approx \frac{2-\delta_{n,0}}{\pi}\sum_{k=0}^{N-1} w_k f_k \cos\left(n\theta_k\right)\,, \tag{E.10}$$

with $f_k \equiv f(x_k) = \tilde{f}(\xi_k)$, $\theta_k \equiv \arccos(\xi_k)$, and weights w_k that depend on the choice of the sampling points x_k. The approximation becomes exact for $N \to \infty$. There are two possible choices of sampling points:

- Gauss-Lobatto-Chebychev grid ("closed interpolation": $\xi_k \in [1, -1]$)

  $$\xi_k = \cos\left(\frac{k\pi}{N-1}\right)\,, \qquad w_k = \left(\frac{1}{2}\right)^{\delta_{n,0}}\left(\frac{1}{2}\right)^{\delta_{n,N-1}}\frac{\pi}{N-1}\,. \tag{E.11}$$

  The ξ_k are the extrema of the Chebychev polynomials, plus endpoints. The resulting formula for the coefficients a_n is defined as a discrete cosine transform of type I (DCT I) and may be implemented via the FFT [270]. The complete calculation of the coefficients is shown in Algorithm 1.

- Gauss-Chebychev (“open interpolation”: $\xi_k \in (1, -1)$)

  $$\xi_k = \cos\left(\frac{\left(k+\frac{1}{2}\right)\pi}{N}\right), \qquad w_k = \frac{\pi}{N}\,. \tag{E.12}$$

  The ξ_k are the roots of the Chebychev polynomials. The resulting formula for the a_n is defined as a discrete cosine transform of type II (DCT II). There is no direct mapping to the FFT in this case.

Algorithm 1 ChebychevCoefficients for expansion of $f(x)$.

Input: $f(x)$ with $x \in [x_{\min}, x_{\max}]$; maximum number $n_{\max}$ of coefficients

Output: Array of Chebychev coefficients $[a_0 \dots a_n]$, $n < n_{\max}$ allowing to approximate $f(x)$ to predefined precision.

```
procedure ChebyCoeffs($f(x)$)
    $\Delta = x_{\max} - x_{\min}$
    $\alpha = \frac{1}{2}\Delta$; $\beta = \alpha + x_{\min}$
    $F_{0:2n_{\max}-3} = 0$                          ▷ allocation to size $2(n_{\max}-1)$
    for $k = 0 : n_{\max} - 1$ do
        $\xi_k = \cos\left(\frac{k\pi}{n_{\max}-1}\right)$                  ▷ $\xi_k = +1, \dots, -1$
        $F_k = f(\alpha\xi_k + \beta)$
    end for
    for $k = 1 : n_{\max} - 2$ do                    ▷ mirror, without endpoints
        $F_{n_{\max}-1+k} = F_{n_{\max}-1-k}$
    end for
    $F = \mathrm{FFT}(F)$
    $F = F/(n_{max} - 1)$
    $F_0 = \frac{1}{2}F_0$
    $F_{n_{\max}-1} = \frac{1}{2}F_{n_{\max}-1}$
    for $i = 0 : n_{\max} - 1$ do
        $a_i = F_i$
        if $|a_i| <$ limit then exit loop with $n = i$
    end for
    return $[a_0, \dots a_n]$
end procedure
```

F

Propagation Algorithms

F.1 Chebychev Propagator

For the evaluation of the time evolution operator $f(\pm\hat{A}\,dt) = \mathrm{e}^{\pm\mathrm{i}\hat{A}\,dt}$, the Chebychev propagation [114, 129] is implemented by Algorithm 2. The operator $\hat{A}$ may be a Hamiltonian, in which case the state vectors $\vec{v}$ are Hilbert space states, or $\hat{A}$ may be a Liouvillian with no dissipators, in which case $\vec{v}$ is a density matrix.

A re-calculation of the Chebychev coefficients in line 4 is necessary only if the spectral radius Δ and minimum eigenvalue E_{min} have changed, otherwise, the coefficients from a previous calculation can be re-used. For a more general function $f(\hat{A})$, instead of ExpChebyCoeffs, the routine ChebyCoeffs defined in Algorithm 1 may be used, together with an appropriate pre-factor in lines 8 and 11. The coefficients $[a_0 \dots a_n]$ are the same for both a forward and a backward propagation. They are real for the exponential function, but may be complex for an arbitrary $f(\hat{A})$.

Algorithm 2 CHEBYCHEV-PROPAGATOR Evaluate $\vec{w} = f(\pm\hat{A}\,dt)\vec{v}$, with $f(\pm\hat{A}\,dt) = \mathrm{e}^{\pm\mathrm{i}\hat{A}\,dt}$.

Input: input vector $\vec{v} \in \mathbb{C}^N$; operator $\hat{A} \in \mathbb{C}^{N\times N}$; time step dt;
Output: Approximation of propagated vector $\vec{w} = e^{-i\hat{A}\,dt}\vec{v} \in \mathbb{C}^N$

```
procedure CHEBY($\vec{v}$, $\hat{A}$, $dt$)
    $\Delta$ = spectral radius of $\hat{A}$
    $E_{\min}$ = minimum eigenvalue of $\hat{A}$
    $[a_0 \dots a_n]$ = EXPCHEBYCOEFFS($\Delta$, $E_{\min}$, $dt$)
    $d = \frac{1}{2}\Delta$; $\beta = d + E_{\min}$
    $\vec{v}_0 = \vec{v}$
    $\vec{w}^{(0)} = a_0\vec{v}_0$
    $\vec{v}_1 = \pm\frac{\mathrm{i}}{d}\left(\hat{A}\vec{v}_0 - \beta\vec{v}_0\right)$
    $\vec{w}^{(1)} = \vec{w}^{(0)} + a_1\vec{v}_1$
    for $i = 2 : n$ do
        $\vec{v}_i = \pm\frac{2\mathrm{i}}{d}\left(\hat{A}\vec{v}_{i-1} - \beta\vec{v}_{i-1}\right) + \vec{v}_{i-2}$
        $\vec{w}^{(i)} = \vec{w}^{(i-1)} + a_i\vec{v}_i$
    end for
    return $e^{\pm\mathrm{i}\,\beta\,dt}\vec{w}^{(n)}$
end procedure
```

Algorithm 3 CHEBYCHEVCOEFFICIENTS for $f(\pm\hat{A}\,dt) = e^{\pm i\hat{A}\,dt}$.

Input: spectral radius Δ of $\hat{A}$; minimum eigenvalue $E_{\min}$ of $\hat{A}$; time step dt
Output: Array of Chebychev coefficients $[a_0 \dots a_n]$ allowing to approximate $f(\hat{A}\,dt)$ to pre-defined precision.

```
procedure EXPCHEBYCOEFFS($\Delta$, $E_{\min}$, $dt$)
    $\alpha = \frac{1}{2}\Delta\,dt$
    $a_0 = J_0(\alpha)$                ▷ 0'th order Bessel-function of first kind
    for $i = 1 : n_{\max} \approx 4\lfloor\alpha\rfloor$ do
        $a_i = 2J_i(\alpha)$           ▷ $i$'th order Bessel-function of first kind
        if $|a_i| <$ limit then exit loop with $n = i$
    end for
    return $[a_0, \dots a_n]$
end procedure
```

F.2 Newton Propagator with Restarted Arnoldi

For the evaluation of the time evolution operator $f(\pm\hat{A}\,dt) = \mathrm{e}^{\pm\mathrm{i}\hat{A}\,dt}$, where $\hat{A}$ has complex eigenvalues, the Newton algorithm using a restarted Arnoldi scheme [136] can be used, as discussed in section 3.2.2 of chapter 3. The operator $\hat{A}$ is typically a dissipative Liouvillian, but may also be a Hamiltonian with non-Hermitian terms. The propagation scheme is implemented by algorithm 4, using the Arnoldi algorithm 5 as a central component.

The algorithm proceeds in iterations, with each iteration adding m new terms to the Newton expansion of the propagator. In each iteration, starting from a vector $\vec{v}_s$, the Arnoldi algorithm is performed to obtain the Arnoldi vectors spanning the m-dimensional Krylov subspace, as well as the Hessenberg matrix $\hat{H}$, i.e. the projection of $\hat{A}$ into the Krylov space. From the eigenvalues of $\hat{H}$ (the "Ritz values"), new sampling points for the Newton polynomials are chosen, in Leja ordering to maximize numerical stability, using algorithm 6. Then, new Newton coefficients are calculated using algorithm 7. Lastly, the coefficients and Leja points are used to calculate the contribution to the propagation result, and the starting vector for the next iteration.

Moreover:

- The check in line 10 catches if $\vec{v}_s$ is an eigenstate of $\hat{A}$.
- For the normalization in line 13, the normalization radius ρ and center c should be calculated in the *first* iteration and the re-used in subsequent iterations. For the set of Leja points $Z_0 = [z_0 : z_{m-1}]$, they are calculated as

$$c = \frac{1}{m} \sum_{j=0}^{m-1} z_j \,, \qquad \rho = \Pi_{j=0}^{m-1} \left| c - z_j \right|^{\frac{1}{m}} \,. \tag{F.1}$$

- The call to ExtendLeja in line 15 adds Leja points $z_{n_s} \dots z_{n_s+m-1}$.
- The call to ExtendNewtonCoeffs in line 16 adds the coefficients $a_{n_s} \dots a_{n_s+m-1}$
- The loop ending in line 22 implements the formula

$$\vec{r}_{m-1} = \Pi_{j=1}^{k} (\hat{H} - z_{n_s+j-1}\mathbb{1})\vec{e}_1 \,, \tag{F.2}$$

$$\vec{p}_{m-1} = \sum_{k=0}^{m-1} a_{n_s+k}\, \vec{r}_k \,. \tag{F.3}$$

Algorithm 4 RestartedNewton: Evaluate $\vec{w} = f(\pm\hat{A}\,dt)\vec{v}$, with $f(\pm\hat{A}\,dt) = \mathrm{e}^{\pm\mathrm{i}\hat{A}\,dt}$.

Input: input vector $\vec{v} \in \mathbb{C}^N$; operator $\hat{A} \in \mathbb{C}^{N\times N}$; time step dt; maximum size m of Hessenberg matrices

Output: Approximation of propagated vector $\vec{w} = \mathrm{e}^{\pm\mathrm{i}\hat{A}\,dt}\vec{v} \in \mathbb{C}^N$

```
procedure RestartedNewton($\vec{v}$, $\hat{A}$, $dt$, $m$)
    $A_0 = \emptyset$; $Z_0 = \emptyset$                          ▷ $A$, $Z$ use zero-based indexing
    $\vec{w}^{(0)} = \vec{0} \in \mathbb{C}^N$
    $\vec{v}_0 = \vec{v} \in \mathbb{C}^N$
    $\beta = \|\vec{v}_0\|$
    $\vec{v}_0 = \vec{v}_0/\beta$
    $s = 0$
    while not converged do                           ▷ Iteration $s \to s+1$
        $U$, $\hat{H}$, $Z$, $m$ = Arnoldi($\hat{A}$, $dt$, $\vec{v}_s$, $m$)
        if $m = 0$, $s = 0$ then
            return $\mathrm{e}^{\pm\mathrm{i}\beta H_{1,1}}\vec{v}_s$
        end if
        Normalize $Z$ with center $c$ and radius $\rho$
        $n_s = |Z_s|$
        $Z_{s+1}$ = ExtendLeja($Z_s, Z, m$)
        $A_{s+1}$ = ExtendNewtonCoeffs($A_s, Z_{s+1}, \rho, c$)
        $\vec{r}_0 = \beta\vec{e}_1 \in \mathbb{C}^{m+1}$                       ▷ $\vec{e}_1$ is unit vector
        $\vec{p}_0 = a_{n_s}\vec{r}_0$
        for k=1:m-1 do
            $\vec{r}_k = (\frac{1}{\rho}\hat{H} - (z_{n_s+k-1}) + \frac{c}{\rho})\vec{r}_{k-1}$
            $\vec{p}_k = \vec{p}_{k-1} + a_{n_s+k}\vec{r}_k$
        end for
        $w_+^{(s)} = \sum_{i=1}^{m} [\vec{p}_{m-1}]_i\, \vec{u}_i$                      ▷ $\vec{u}_i \in U$
        $\vec{w}^{(s+1)} = \vec{w}^{(s)} + \vec{w}_+^{(s)}$
        $\vec{r}_m = \left(\frac{1}{\rho}\hat{H} - \left(z_{n_s+m-1} + \frac{c}{\rho}\right)\mathbb{1}\right)\vec{r}_{m-1}$
        $\beta = \|\vec{r}_m\|$
        $\vec{r}_m = \vec{r}_m/\beta$
        $\vec{v}_{s+1} = \sum_{i=1}^{m+1} [\vec{r}_m]_i\, \vec{u}_i$                      ▷ $\vec{u}_i \in U$
        $s = s + 1$
        converged if $\frac{\|\vec{w}_+^s\|}{\|\vec{w}^s\|}$ < limit
    end while
    return $\vec{w}^{(s)}$
end procedure
```

Algorithm 5 ARNOLDI: Obtain the $m \times m$ Hessenberg matrix for an operator $\hat{A}dt$ by projecting it into the Krylov space starting from a vector $\vec{v}$.

Input: Operator $\hat{A}$, Time step dt; input vector $\vec{v}$; maximum order $m_{\max}$

Output: Array of $m+1$ (extended) Arnoldi vectors, each of the same dimension as $\vec{v}$; Extended $(m+1) \times (m+1)$ Hessenberg matrix $\hat{H}$; Accumulated Ritz values Z; Dimension m of Hessenberg matrix

```
 1: procedure ARNOLDI($\hat{A}$, $dt$, $\vec{v}$, $m_{\max}$)
 2:     $\beta = \|\vec{v}\|$; $\vec{u}_1 = \vec{v}/\beta$; $U = [\vec{u}_1]$; $Z = \emptyset$; $m = m_{\max}$
 3:     $\hat{H}_{1:(m+1),1:(m+1)} = 0$
 4:     for $j = 1 : m_{\max}$ do
 5:         $\vec{u}_{j+1} = \hat{A}\vec{u}_j$
 6:         for $i = 1 : j$ do
 7:             $H_{i,j} = dt \, \langle \vec{u}_i \,|\, \vec{u}_{j+1} \rangle$
 8:             $\vec{u}_{j+1} = \vec{u}_{j+1} - \frac{H_{i,j}}{dt}\vec{u}_i$
 9:         end for
10:         $Z = Z \cup \text{eigenvalues}(\hat{H}_{1:j,1:j})$
11:         $h_{next} = |\vec{u}_{j+1}|$
12:         if $h_{next} \approx 0$: $m = j$, exit loop
13:         $\vec{u}_{j+1} = \vec{u}_{j+1}/h_{next}$
14:         $U = U \cup \vec{u}_{j+1}$
15:         $H_{j+1,j} = h_{next}\, dt$
16:     end for
17:     return $U = [\vec{u}_1 : \vec{u}_{m+1}]$, $\hat{H}_{1:(m+1),1:(m+1)}$, $Z$, $m$
18: end procedure
```

Notes:

- At the end of the loop in line 9, $\hat{H}$ is complete as a $j \times j$ matrix. The loop itself is a Gram-Schmidt orthonormalization of the Arnoldi vectors $\vec{u}_j$.
- The eigenvalues of the Hessenberg matrix in line 10 can be obtained with the QR method, implemented in LAPACK as `ZHSEQR`.

Algorithm 6 EXTENDLEJA: Choose m new points from a set of Ritz values to extend an existing array of Leja points

Input: Array Z_s of n_s existing Leja points; Array Z of new candidate points (Ritz values); Number m of points to pick from Z.
Output: Array Z_{s+1} of $n_s + m$ Leja-ordered points

```
procedure ExtendLeja($Z_s, Z, m$)
    $n_0 = 1$; $Z_{s+1} = Z_s$
    if $Z_s = \emptyset$ then
        $z = \max(\text{abs}(Z))$
        $Z_{s+1} = Z_{s_1} \cup z$; remove $z$ from $Z$
        $n_0 = 2$
    end if
    for $n = n_0 : m$ do
        Select $z_i \in Z$ that maximizes $\Pi_{z_j \in Z_{s+1}} |z_i - z_j|$
        $Z_{s+1} = Z_{s+1} \cup z_i$; remove $z_i$ from $Z$
    end for
    return $Z_{s+1}$
end procedure
```

Algorithm 7 ExtendNewtonCoeffs: Choose m new points from a set of Ritz values to extend an existing array of Leja points

Input: Array $A_s = [a_0 \dots a_{n_s-1}]$ of n_s Newton coefficients from previous iteration; Array $Z_{s+1} = [z_0 \dots z_{n_s-1+m}]$ of Leja points; Normalization radius ρ; Normalization center c

Output: Array $A_{s+1} = [a_0 \dots\dots a_{n_s-1+m}]$ of $n_s + m$ Newton coefficients

```
procedure ExtendNewtonCoeffs($A_s, Z_{s+1}, \rho, c$)
    $A_{s+1} = A_s$; $n_0 = n_s = |A_s|$, $m = |Z_{s+1}| - n_s$;
    Define $f(z) = \mathrm{e}^{\pm\mathrm{i}(\rho z + c)}$
    if $n_s = 0$ then
        $a_0 = f(z_0)$
        $A_{s+1} = A_{s+1} \cup a_0$
        $n_0 = 1$
    end if
    for $k = n_0 : n_s - 1 + m$ do
        $a_k = \frac{f(z_k) - a_0 - \sum_{n=1}^{k-1} a_n \Pi_{j=0}^{n-1}(z_k - z_j)}{\Pi_{j=0}^{k-1}(z_k - z_j)}$
        $A_{s+1} = A_{s+1} \cup a_k$
    end for
    return $A_{s+1}$
end procedure
```

G

Krotov Boundary Condition for the Perfect Entangler Functional

We derive the boundary condition

$$\left|\chi^{(i)}(T)\right\rangle = -\left.\frac{\partial J_T}{\partial \langle\phi_k|}\right|_{\phi_k^{(i)}(T)}, \tag{G.1}$$

cf. Eq. (3.56), for the backwards-propagation of the co-state $|\chi(t)\rangle$ in the update equation (3.62) of Krotov's method, for the perfect-entanglers functional

$$J_{\text{PE}} = g_3\sqrt{g_1^2 + g_2^2} - g_1\,, \tag{G.2}$$

cf. Eq. (6.12).

The forward propagation of the four Bell states

$$|\phi_1\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - \mathrm{i}|11\rangle\right)\,, \qquad |\phi_2\rangle = -\frac{1}{\sqrt{2}}\left(\mathrm{i}|01\rangle - |10\rangle\right)\,, \tag{G.3}$$

$$|\phi_3\rangle = -\frac{1}{\sqrt{2}}\left(\mathrm{i}|01\rangle + |10\rangle\right)\,, \qquad |\phi_4\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + \mathrm{i}|11\rangle\right)\,. \tag{G.4}$$

in the two-qubit logical subspace yields a gate U_B, as a 4×4 matrix. We split this matrix into real and imaginary part

$$A \equiv \mathfrak{Re}[U_B], \qquad B \equiv \mathfrak{Im}[U_B]\,. \tag{G.5}$$

The co-state at final time T can then be written as

$$\left|\chi(T)\right\rangle = \sum_{l=1}^{4} a_{lk}^{*} \left|\phi_l(T)\right\rangle \,, \tag{G.6}$$

with

$$a_{kl} = -\left(\frac{\partial F_{PE}}{\partial A_{kl}} + \mathrm{i}\frac{\partial F_{PE}}{\partial B_{kl}}\right) \,. \tag{G.7}$$

We can rewrite

$$\begin{aligned} \sqrt{g_1^2 + g_2^2} &= \sqrt{\left(\mathfrak{Re}\frac{\mathrm{tr}^2[m]}{16\det U_B}\right)^2 + \left(\mathfrak{Im}\frac{\mathrm{tr}^2[m]}{16\det U_B}\right)^2} \\ &= \frac{1}{16}\left(\mathfrak{Re}^2\left[\mathrm{tr}[m]\right] + \mathfrak{Im}^2\left[\mathrm{tr}[m]\right]\right) \,, \end{aligned} \tag{G.8}$$

with $m \equiv U_B^T U_B$, and having used $\left|\mathrm{tr}^2[m]\right| = \left|\mathrm{tr}[m]\right|^2$.

With Eq.(G.8), and

$$g_3 = \frac{\mathrm{tr}^2[m] - \mathrm{tr}[m^2]}{4\det U_B} \,, \tag{G.9}$$

the perfect-entanglers functional can be rewritten as

$$\begin{aligned} F_{PE} = &\left(\frac{1}{\det U_B}\right)\left(\frac{1}{4}(\mathrm{tr}^2[m] - \mathrm{tr}[m^2])\right)\left(\frac{1}{16}\mathfrak{Re}^2[\mathrm{tr}[m]]\right) + \\ &+ \left(\frac{2}{\det U_B}\right)\left(\frac{1}{4}(\mathrm{tr}^2[m] - \mathrm{tr}[m^2])\right)\left(\frac{1}{16}\mathfrak{Im}^2[\mathrm{tr}[m]]\right)\left(\frac{1}{16}\mathfrak{Re}[\mathrm{tr}^2[m]]\right) \,. \end{aligned} \tag{G.10}$$

The derivatives in Eq. (G.7) can now be evaluated via the chain rule, where the derivatives of the individual terms are

$$\frac{\partial}{\partial A_{ab}}\left(\frac{1}{\det[U_B]}\right) = -\frac{\det[U_B^{(ab)'}]}{\det^2[U_B]} \,, \tag{G.11}$$

$$\frac{\partial}{\partial B_{ab}}\left(\frac{1}{\det[U_B]}\right) = -\mathrm{i}\frac{\det[U_B^{(ab)'}]}{\det^2[U_B]} \,, \tag{G.12}$$

where $U_B^{(ab)'}$ is obtained from U_B by replacing the a'th column by the b'th unit vector, following Leibniz' formula, cf. Ref [238]; furthermore,

$$\begin{aligned}
&\frac{\partial}{\partial A_{ab}}\left(\frac{1}{4}(\mathrm{tr}^2[m]-\mathrm{tr}[m^2])\right)\\
&=\sum_{ki}\Bigg(\Big(A_{ab}A_{ki}A_{ki}-A_{ab}B_{ki}B_{ki}-2B_{ab}A_{ki}B_{ki}\\
&\qquad -A_{ki}A_{ai}A_{kb}+A_{kb}B_{ai}B_{ki}+2B_{kb}A_{ai}B_{ki}\Big)+\\
&\quad +\mathrm{i}\Big(B_{ab}A_{ki}A_{ki}-B_{ab}B_{ki}B_{ki}+2A_{ab}A_{ki}B_{ki}-A_{ai}A_{ki}B_{kb}\\
&\qquad -A_{kb}A_{ki}B_{ai}-A_{kb}A_{ai}B_{ki}+B_{ai}B_{ki}B_{kb}\Big)\Bigg)\,,
\end{aligned} \tag{G.13}$$

$$\begin{aligned}
&\frac{\partial}{\partial B_{ab}}\left(\frac{1}{4}(\mathrm{tr}^2[m]-\mathrm{tr}[m^2])\right)\\
&=\sum_{ki}\Bigg(\Big(B_{ab}B_{ki}B_{ki}-B_{ab}A_{ki}A_{ki}-2A_{ab}A_{ki}B_{ki}\\
&\qquad -B_{ki}B_{ai}B_{kb}+B_{kb}A_{ai}A_{ki}+2A_{kb}B_{ai}A_{ki}\Big)+\\
&\quad +\mathrm{i}\Big(-A_{ab}B_{ki}B_{ki}+A_{ab}A_{ki}A_{ki}-2B_{ab}B_{ki}A_{ki}+B_{ai}B_{ki}A_{kb}\\
&\qquad +B_{kb}B_{ki}A_{ai}+B_{kb}B_{ai}A_{ki}-A_{ai}A_{ki}A_{kb}\Big)\Bigg)\,,
\end{aligned} \tag{G.14}$$

$$\frac{\partial}{\partial A_{ab}}\left(\frac{1}{16}\mathfrak{Re}^2[\mathrm{tr}[m]]\right)=\frac{1}{4}\sum_{ki}\left(A_{ab}A_{ki}A_{ki}-A_{ab}B_{ki}B_{ki}\right)\,, \tag{G.15}$$

$$\frac{\partial}{\partial B_{ab}}\left(\frac{1}{16}\mathfrak{Re}^2[\mathrm{tr}[m]]\right)=\frac{1}{4}\sum_{ki}\left(B_{ab}B_{ki}B_{ki}-B_{ab}A_{ki}A_{ki}\right)\,, \tag{G.16}$$

$$\frac{\partial}{\partial A_{ab}}\left(\frac{1}{16}\mathfrak{Im}^2[\mathrm{tr}[m]]\right)=\frac{1}{2}\sum_{ki}\left(A_{ki}B_{ab}B_{ki}\right)\,, \tag{G.17}$$

$$\frac{\partial}{\partial B_{ab}}\left(\frac{1}{16}\mathfrak{Im}^2[\mathrm{tr}[m]]\right)=\frac{1}{2}\sum_{ki}\left(B_{ki}A_{ab}A_{ki}\right)\,, \tag{G.18}$$

$$\begin{aligned}
&\frac{\partial}{\partial A_{ab}}\left(\frac{1}{16}\mathfrak{Re}[\mathrm{tr}^2[m]]\right)\\
&=\frac{1}{4}\sum_{ki}\Bigg(\Big(A_{ab}A_{ki}A_{ki}-A_{ab}B_{ki}B_{ki}-2B_{ab}A_{ki}B_{ki}\Big)+\\
&\qquad +\mathrm{i}\Big(+B_{ab}A_{ki}A_{ki}-B_{ab}B_{ki}B_{ki}+2A_{ab}A_{ki}B_{ki}\Big)\Bigg)\,,
\end{aligned} \tag{G.19}$$

and finally

$$\begin{aligned}
&\frac{\partial}{\partial B_{ab}}\left(\frac{1}{16}\mathfrak{Re}[\mathrm{tr}^2[m]]\right) \\
&= \frac{1}{4}\sum_{ki}\bigg(\Big(B_{ab}B_{ki}B_{ki} - B_{ab}A_{ki}A_{ki} - 2A_{ab}B_{ki}A_{ki}\Big)+ \\
&\qquad + \mathrm{i}\Big(-A_{ab}B_{ki}B_{ki} + A_{ab}A_{ki}A_{ki} - 2B_{ab}B_{ki}A_{ki}\Big)\bigg)\,.
\end{aligned} \tag{G.20}$$

H

Optimization Functional for a Holonomic Phasegate

We formulate a functional for the holonomic phasegate of chapter 5, i.e. an arbitrary diagonal perfect entangler. The functional can be expressed both in Hilbert space or Liouville space. We derive the boundary condition for the backward propagation in Krotov's method, Eq. (3.56).

H.1 Hilbert Space Functional

The functional is split in two parts, corresponding to the two requirements that the gate must be diagonal, and that it should be a perfect entangler,

$$J_T = J_{ss} + w_\gamma J_\gamma . \tag{H.1}$$

The two terms may be weighted with w_γ.

The first part of the functional states that every logical eigenstate should be mapped onto itself, irrespective of any phase.

$$\begin{aligned} J_{ss} = 4 &- \left|\left\langle 00 \middle| \hat{\mathsf{U}} \middle| 00 \right\rangle\right|^2 - \left|\left\langle 01 \middle| \hat{\mathsf{U}} \middle| 01 \right\rangle\right|^2 + \\ &- \left|\left\langle 10 \middle| \hat{\mathsf{U}} \middle| 10 \right\rangle\right|^2 - \left|\left\langle 11 \middle| \hat{\mathsf{U}} \middle| 11 \right\rangle\right|^2 \end{aligned} \tag{H.2}$$

J_{ss} takes values in $[0, 4]$ and optimizes for an arbitrary diagonal gate.

The second requirement of having a perfect entangler is fulfilled if the non-local phase, cf. Eq. (2.30) reaches a value of π, $\gamma = \phi_{00} - \phi_{10} - \phi_{01} + \phi_{11} = \pi$

where ϕ_{00} is the phase obtained by the state $|00\rangle$ and equivalently for the other states of the logical basis.

We find

$$\left\langle 00 \middle| \hat{\mathsf{U}} \middle| 00 \right\rangle \left\langle 01 \middle| \hat{\mathsf{U}}^\dagger \middle| 01 \right\rangle \left\langle 10 \middle| \hat{\mathsf{U}}^\dagger \middle| 10 \right\rangle \left\langle 11 \middle| \hat{\mathsf{U}} \middle| 11 \right\rangle = e^{i\gamma} . \tag{H.3}$$

Together with

$$2\cos(\gamma) = e^{i\gamma} + e^{-i\gamma} , \tag{H.4}$$

and the condition that $\cos(\gamma) = -1$ for $\gamma = \pi$, this determines the functional J_γ,

$$\begin{aligned} J_\gamma = 2 &+ \left\langle 00 \middle| \hat{\mathsf{U}} \middle| 00 \right\rangle \left\langle 01 \middle| \hat{\mathsf{U}}^\dagger \middle| 01 \right\rangle \left\langle 10 \middle| \hat{\mathsf{U}}^\dagger \middle| 10 \right\rangle \left\langle 11 \middle| \hat{\mathsf{U}} \middle| 11 \right\rangle \\ &+ \left\langle 11 \middle| \hat{\mathsf{U}}^\dagger \middle| 11 \right\rangle \left\langle 10 \middle| \hat{\mathsf{U}} \middle| 10 \right\rangle \left\langle 01 \middle| \hat{\mathsf{U}} \middle| 01 \right\rangle \left\langle 00 \middle| \hat{\mathsf{U}}^\dagger \middle| 00 \right\rangle . \end{aligned} \tag{H.5}$$

J_γ takes values in $[0, 4]$.

In order to derive the boundary condition for the backward propagation of Krotov's method, we use the abbreviation

$$\tau_{00} \equiv \left\langle 00 \middle| \hat{\mathsf{U}} \middle| 00 \right\rangle , \tag{H.6}$$

and equivalently for $|01\rangle$, $|10\rangle$, and $|11\rangle$. The two terms of the functional are then

$$J_{ss} = 4 - \tau_{00}\tau_{00}^* - \tau_{01}\tau_{01}^* - \tau_{10}\tau_{10}^* - \tau_{11}\tau_{11}^* , \tag{H.7a}$$

$$J_\gamma = 2 + \tau_{00}\tau_{01}^*\tau_{10}^*\tau_{11} + \tau_{00}^*\tau_{01}\tau_{10}\tau_{11}^* . \tag{H.7b}$$

In order to calculate the boundary conditions for the backward propagated states,

$$|\chi(T)\rangle_i = -\frac{\partial J_T}{\partial \langle i(T)|}; \quad |i(T)\rangle \equiv \hat{\mathsf{U}} |i\rangle ; \quad i = \{00, 01, 10, 11\} , \tag{H.8}$$

we use

$$\frac{\partial \tau_i}{\partial \langle j(T)|} = 0, \qquad \frac{\partial \tau_i^*}{\partial \langle j(T)|} = \delta_{ij} |i\rangle ; \qquad i, j = \{00, 01, 10, 11\} . \tag{H.9}$$

This yields

$$|\chi(T)\rangle_{00} = (\tau_{00} - w_\gamma \tau_{01}\tau_{10}\tau_{11}^*) |00\rangle , \tag{H.10a}$$

$$|\chi(T)\rangle_{01} = (\tau_{01} - w_\gamma \tau_{00}\tau_{10}^*\tau_{11}) |01\rangle , \tag{H.10b}$$

$$|\chi(T)\rangle_{11} = (\tau_{11} - w_\gamma \tau_{00}^*\tau_{01}\tau_{10}) |11\rangle , \tag{H.10c}$$

$$|\chi(T)\rangle_{10} = (\tau_{10} - w_\gamma \tau_{00}\tau_{01}^*\tau_{11}) |10\rangle . \tag{H.10d}$$

H.2 Liouville Space Functional

Based on the work presented in chapter 7, an optimization functional for a holonomic gate in Liouville space can be formulated. The optimization success is tracked by the three states $\hat{\rho}_1$, $\hat{\rho}_2$, $\hat{\rho}_3$ defined in Eq. (7.12). The optimization functional contains one term for each of these matrices,

$$J_T = w_1 J_1 + w_2 J_2 + w_3 J_3 \,. \tag{H.11}$$

with weights w_1, w_2, w_3.

When the optimization target is reached,

$$\hat{\mathsf{U}} \hat{\rho}_1 \hat{\mathsf{U}}^\dagger = \hat{\rho}_1 \,, \qquad \hat{\mathsf{U}} \hat{\rho}_3 \hat{\mathsf{U}}^\dagger = \hat{\rho}_3 \,. \tag{H.12}$$

The state $\hat{\rho}_3$ ensures that no population is lost from the logical subspace, $\hat{\rho}_1$ optimizes for a diagonal gate. Therefore, we choose

$$\begin{aligned} J_{1,3} &= 1 - \frac{1}{\operatorname{tr}\left[\hat{\rho}_{1,3}^2\right]} \mathfrak{Re}\left[\left\langle\!\!\left\langle \hat{\rho}_{1,3} \,\middle|\, \mathcal{E}\hat{\rho}_{1,3} \right\rangle\!\!\right\rangle\right] \\ &= 1 - \frac{1}{\operatorname{tr}\left[\hat{\rho}_{1,3}^2\right]} \frac{1}{2} \left[\left\langle\!\!\left\langle \hat{\rho}_{1,3} \,\middle|\, \mathcal{E}\hat{\rho}_{1,3} \right\rangle\!\!\right\rangle + \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_{1,3} \,\middle|\, \hat{\rho}_{1,3} \right\rangle\!\!\right\rangle\right] \,, \end{aligned} \tag{H.13}$$

using $\mathcal{E}$ to indicate the dynamical map, and a braket-notation for the Hilbert-Schmidt product,

$$\left\langle\!\!\left\langle \hat{\mathsf{A}} \,\middle|\, \hat{\mathsf{B}} \right\rangle\!\!\right\rangle \equiv \operatorname{tr}\left[\hat{\mathsf{A}}^\dagger \hat{\mathsf{B}}\right] \,. \tag{H.14}$$

The remaining state $\hat{\rho}_2$ is used to track the relative phases and to optimize the entanglement. If $\hat{\mathsf{U}}$ is diagonal as in Eq. (2.29), then

$$\hat{\mathsf{U}} \hat{\rho}_2 \hat{\mathsf{U}}^\dagger = \sum_{i,j} e^{i(\phi_i - \phi_j)} \left|i\right\rangle\!\left\langle j\right| ; \quad i,j = \{00, 01, 10, 11\} \,. \tag{H.15}$$

With $\gamma = (\phi_{00} - \phi_{01}) + (\phi_{00} - \phi_{10}) + (\phi_{11} - \phi_{00})$ and $\hat{\mathsf{P}}_{i,j} \equiv \left|i\right\rangle\!\left\langle j\right|$, we find

$$e^{i\gamma} = \left\langle\!\!\left\langle \hat{\mathsf{P}}_{00,01} \,\middle|\, \hat{\mathsf{U}}\hat{\rho}_2\hat{\mathsf{U}}^\dagger \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \hat{\mathsf{P}}_{00,10} \,\middle|\, \hat{\mathsf{U}}\hat{\rho}_2\hat{\mathsf{U}}^\dagger \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \hat{\mathsf{P}}_{11,00} \,\middle|\, \hat{\mathsf{U}}\hat{\rho}_2\hat{\mathsf{U}}^\dagger \right\rangle\!\!\right\rangle \tag{H.16}$$

$$e^{-i\gamma} = \left\langle\!\!\left\langle \hat{\mathsf{U}}\hat{\rho}_2\hat{\mathsf{U}}^\dagger \,\middle|\, \hat{\mathsf{P}}_{00,01} \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \hat{\mathsf{U}}\hat{\rho}_2\hat{\mathsf{U}}^\dagger \,\middle|\, \hat{\mathsf{P}}_{00,10} \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \hat{\mathsf{U}}\hat{\rho}_2\hat{\mathsf{U}}^\dagger \,\middle|\, \hat{\mathsf{P}}_{11,00} \right\rangle\!\!\right\rangle \tag{H.17}$$

With Eq. (H.4) and the generalization $\hat{\mathsf{U}}\hat{\rho}_2\hat{\mathsf{U}}^\dagger \rightarrow \mathcal{E}\hat{\rho}_2$, we obtain the functional for $\hat{\rho}_2$,

$$\begin{aligned} J_2 = 2 &+ \left\langle\!\!\left\langle \hat{\mathsf{P}}_{00,01} \,\middle|\, \mathcal{E}\hat{\rho}_2 \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \hat{\mathsf{P}}_{00,10} \,\middle|\, \mathcal{E}\hat{\rho}_2 \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \hat{\mathsf{P}}_{11,00} \,\middle|\, \mathcal{E}\hat{\rho}_2 \right\rangle\!\!\right\rangle \\ &+ \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_2 \,\middle|\, \hat{\mathsf{P}}_{00,01} \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_2 \,\middle|\, \hat{\mathsf{P}}_{00,10} \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_2 \,\middle|\, \hat{\mathsf{P}}_{11,00} \right\rangle\!\!\right\rangle \,. \end{aligned} \tag{H.18}$$

J_1 and J_3 take values in $[0,1]$, whereas J_2 takes values in $[0,4]$.

Equivalently to Eq. (H.8), the boundary condition for the backward-propagated states in Krotov's method is

$$\hat{\sigma}_i(T) = -\frac{\partial J_T}{\partial \langle\!\langle \mathcal{E}\hat{\rho}_i \vert}; \quad i = \{1,2,3\}\,, \tag{H.19}$$

which yields

$$\hat{\sigma}_1(T) = \frac{w_1}{2}\hat{\rho}_1\,, \tag{H.20}$$

$$\begin{aligned}
\hat{\sigma}_2(T) = &- \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_2 \,\middle\vert\, \hat{\mathsf{P}}_{00,10} \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_2 \,\middle\vert\, \hat{\mathsf{P}}_{11,00} \right\rangle\!\!\right\rangle w_2\, \hat{\mathsf{P}}_{00,01} \\
&- \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_2 \,\middle\vert\, \hat{\mathsf{P}}_{00,01} \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_2 \,\middle\vert\, \hat{\mathsf{P}}_{11,00} \right\rangle\!\!\right\rangle w_2\, \hat{\mathsf{P}}_{00,10} \\
&- \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_2 \,\middle\vert\, \hat{\mathsf{P}}_{00,10} \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \mathcal{E}\hat{\rho}_2 \,\middle\vert\, \hat{\mathsf{P}}_{00,01} \right\rangle\!\!\right\rangle w_2\, \hat{\mathsf{P}}_{11,00}\,,
\end{aligned} \tag{H.21}$$

$$\hat{\sigma}_3(T) = \frac{w_3}{2}\hat{\rho}_3\,. \tag{H.22}$$

I

List of Publications

1. **Michael H. Goerz**, Tommaso Calarco, and Christiane P. Koch.
   The quantum speed limit of optimal controlled phasegates for trapped neutral atoms
   J. Phys. B: At. Mol. Opt. Phys. **44**, 154011 (2011).
   Special issue on quantum control theory for coherence and information dynamics.

2. Michał Tomza, **Michael H. Goerz**, Monica Musiał, Robert Moszyński, and Christiane P. Koch.
   Optimal production of ultracold ground-state molecules: Stabilization employing potentials with ion-pair character and strong spin-orbit coupling.
   Phys. Rev. A **86**, 043424 (2012).

3. **Michael H. Goerz**, Daniel M. Reich, and Christiane P. Koch.
   Optimal control theory for a unitary operation under dissipative evolution.
   New J. Phys. **16**, 055012 (2014).
   Special issue on coherent control of complex quantum systems.

4. Henning A. Fürst, **Michael H. Goerz**, Ulrich G. Poschinger, Michael Murphy, Simone Montangero, Tommaso Calarco, Ferdinand Schmidt-Kaler, Kilian Singer, and Christiane P. Koch.
   Controlling the transport of an ion: Classical and quantum mechanical solutions.
   New J. Phys. **16**, 075007 (2014).
   Special issue on coherent control of complex quantum systems.
5. Georg Jäger, Daniel M. Reich, **Michael H. Goerz**, Christiane P. Koch, and Ulrich Hohenester.
   Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Comparison of GRAPE and Krotov optimization schemes.
   Phys. Rev. A **90**, 033628 (2014).
6. **Michael H. Goerz**, Eli J. Halperin, Jon M. Aytac, Christiane P. Koch, and K. Birgitta Whaley.
   Robustness of high-fidelity Rydberg gates with single-site addressability.
   Phys. Rev. A **90**, 032329 (2014).
   Editor's suggestion.
7. Paul Watts, Jiří Vala, Matthias M. Müller, Tommaso Calarco, K. Birgitta Whaley, Daniel M. Reich, **Michael H. Goerz**, and Christiane P. Koch.
   Optimizing for an arbitrary perfect entangler. I. Functionals.
   arXiv:1412.7347. Accepted in Phys. Rev. A.
8. **Michael H. Goerz**, Giulia Gualdi, Daniel M. Reich, Christiane P. Koch, Felix Motzoi, K. Brigitta Whaley, Jiří Vala, Matthias M. Müller, Simone Montangero, and Tommaso Calarco.
   Optimizing for an arbitrary perfect entangler. II. Application.
   arXiv:1412.7350. Accepted in Phys. Rev. A.
9. **Michael H. Goerz**, K. Birgitta Whaley, and Christiane P. Koch
   Hybrid Optimization Schemes for Quantum Control.
   arXiv:1505.05331. Submitted to EPJ Quantum Technol.

Chapters 4, 6, 7 of this thesis are based on the publications (3), (7) and (8), and (6), respectively. In parts of the chapters, text has been been taken verbatim from the corresponding publications, but only if it was the author's original work.

References

[1] J. P. Dowling and G. J. Milburn, Phil. Trans. R. Soc. A **361**, 1655 (2003)

[2] P. Brumer and M. Shapiro, *Principles and Applications of the Quantum Control of Molecular Processes*, Wiley Interscience, 2003

[3] E. Collini and G. D. Scholes, Science **323**, 369 (2009)

[4] M. Sarovar, A. Ishizaki, G. R. Fleming, and K. B. Whaley, Nat Phys **6**, 462 (2010)

[5] M. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000

[6] M. Tomza, M. H. Goerz, M. Musiał, R. Moszynski, and C. P. Koch, Phys. Rev. A **86**, 043424 (2012)

[7] H. A. Fürst, M. H. Goerz, U. G. Poschinger, M. Murphy, S. Montangero, T. Calarco, F. Schmidt-Kaler, K. Singer, and C. P. Koch, New J. Phys. **16**, 075007 (2014)

[8] G. Jäger, D. M. Reich, M. H. Goerz, C. P. Koch, and U. Hohenester, Phys. Rev. A **90**, 033628 (2014)

[9] D. J. Tannor, *Introduction to Quantum Mechanics: A Time-Dependent Perspective*, University Science Books, Sausalito, California, 2007

[10] R. P. Feynman, F. L. Vernon, and R. W. Hellwarth, J. Appl. Phys. **28**, 49 (1957)

[11] L. S. Pontryagin, V. G. Boltyanskii, G. R. V., and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Interscience, New York, NY, 1962

[12] N. Khaneja, R. Brockett, and S. J. Glaser, Phys. Rev. A **63**, 032308 (2001)

[13] M. Lapert, Y. Zhang, M. Braun, S. J. Glaser, and D. Sugny, Phys. Rev. Lett. **104**, 083001 (2010)

[14] E. Assémat, M. Lapert, Y. Zhang, M. Braun, S. J. Glaser, and D. Sugny, Phys. Rev. A **82**, 013415 (2010)

[15] N. V. Vitanov, T. Halfmann, B. W. Shore, and K. Bergmann, Annu. Rev. Phys. Chem. **52**, 763 (2001)

[16] R. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, NJ, 1957

[17] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, 1999

[18] S. Boyd and L. Bandenberghe, *Convex Optimization*, Cambridge University Press, 2004

[19] V. F. Krotov, *Global Methods in Optimal Control*, Dekker, New York, NY, USA, 1996

[20] S. Rice and M. Zhao, *Optical Control of Molecular Dynamics*, John Wiley & Sons, 2000

[21] T. E. Skinner, T. O. Reiss, B. Luy, N. Khaneja, and S. J. Glaser, J. Magnet. Res. **163**, 8 (2003)

[22] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magnet. Res. **172**, 296 (2005)

[23] Z. Tošner, T. Vosegaard, C. Kehlet, N. Khaneja, S. J. Glaser, and N. C. Nielsen, J. Magnet. Res. **197**, 120 (2009)

[24] C. Brif, R. Chakrabarti, and H. Rabitz, New J. Phys. **12**, 075008 (2010)

[25] G. E. Moore, Electronics **38**, 114 (1965)

[26] R. Feynman, Int. J. Theor. Phys. **21**, 467 (1982)

[27] T. Schaetz, C. R. Monroe, and T. Esslinger, New J. Phys. **15**, 085009 (2013)

[28] T. Johnson, S. Clark, and D. Jaksch, EPJ Quantum Technol. **1**, 10 (2014)

[29] I. M. Georgescu, S. Ashhab, and F. Nori, Rev. Mod. Phys. **86**, 153 (2014)

[30] A. M. Turing, Proc. Lond. Math. Soc. **s2-42**, 230 (1937)

[31] D. Deutsch, Proc. R. Soc. A **400**, 97 (1985)

[32] P. Kaye, R. Laflamme, and M. Mosca, *An Introduction to Quantum Computing*, Oxford University Press, New York, NY, USA, 2007

[33] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982)

[34] P. Shor, SIAM J. Comput. **26**, 1484 (1997)

[35] C. Pomerance, Notices Amer. Math. Soc. **43**, 1473 (1996)

[36] G. L. Miller, J. Comput. System Sci. **13**, 300 (1976)

[37] R. L. Rivest, A. Shamir, and L. Adleman, Commun. ACM **21**, 120 (1978)

[38] L. Grover, Phys. Rev. Lett. **79**, 325 (1997)

[39] S. Arora and B. Barak, *Computational Complexity: A Modern Approach*, Cambridge University Press, New York, NY, USA, 1st edition, 2009

[40] E. L. Lawler, J. K. Lenstra, A. H. G. R. Kan, and D. B. Shmoys, *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*, Wiley-Interscience, New York, 1985

[41] T. R. Jensen and B. Toft, *Graph coloring problems*, John Wiley & Sons, New York, 1995

[42] L. Grover, Phys. Rev. Lett. **80**, 4329 (1998)

[43] G. Brassard and P. Hoyer, An exact quantum polynomial-time algorithm for simon's problem, in *Proceedings of the Fifth Israel Symposium on the Theory of Computing Systems (ISTCS '97)*, page 12, IEEE Computer Society, 1997

[44] D. G. Cory, M. D. Price, and T. F. Havel, Physica D **120**, 82 (1998)

[45] A. Imamoğlu, D. Awschalom, G. Burkard, D. DiVincenzo, D. Loss, M. Sherwin, and A. Small, Phys. Rev. Lett. **83**, 4204 (1999)

[46] J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995)

[47] A. Nizovtsev, S. Kilin, F. Jelezko, T. Gaebal, I. Popa, A. Gruber, and J. Wrachtrup, Opt. Spectr. **99**, 233 (2005)

[48] N. Gisin and R. Thew, Nature Photon. **1**, 165 (2007)

[49] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002)

[50] C. Shannon, Bell System Tech. J. **28**, 656 (1949)

[51] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, Nature Commun. **2**, 349 (2011)

[52] C. Elliott, New J. Phys. **4**, 46 (2002)

[53] D. Stucki, M. Legré, F. Buntschu, B. Clausen, N. Felber, N. Gisin, L. Henzen, P. Junod, G. Litzistorf, P. Monbaron, L. Monat, J.-B. Page, D. Perroud, G. Ribordy, A. Rochas, S. Robyr, J. Tavares, R. Thew, P. Trinkler, S. Ventura, R. Voirol, N. Walenta, and H. Zbinden, New J. Phys. **13**, 123001 (2011)

[54] L. Oesterling, D. Hayford, and G. Friend, Comparison of commercial and next generation quantum key distribution: Technologies for secure communication of information, in *IEEE Conference on Technologies for Homeland Security*, page 156, 2012

[55] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science **292**, 472 (2001)

[56] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, Adiabatic quantum computation is equivalent to standard quantum computation, in *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, page 42, 2004

[57] M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk, E. M. Chapple, C. Enderud, J. P. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, C. J. S. Truncik, S. Uchaikin, J. Wang, B. Wilson, and G. Rose, Nature **473**, 194 (2011)

[58] W. van Dam, Nature Phys. **3**, 220 (2007)

[59] T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, Science **345**, 420 (2014)

[60] S. W. Shin, G. Smith, J. A. Smolin, and U. Vazirani, arXiv:1401.7087 (2014)

[61] H. Neven, V. S. Denchev, G. Rose, and W. G. Macready, arXiv:0912.0779 (2009)

[62] H. Neven, G. Rose, and W. G. Macready, arXiv:0804.4457 (2008)

[63] J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, Addison-Wesley, 2 edition, 2011

[64] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems*, Oxford University Press, 2007

[65] Y. Ohtsuki, W. Zhu, and H. Rabitz, J. Chem. Phys. **110**, 9825 (1999)

[66] S. Kallush and R. Kosloff, Phys. Rev. A **73**, 032324 (2006)

[67] J. C. Tremblay and P. Saalfrank, Phys. Rev. A **78**, 063408 (2008)

[68] D. J. Gorman, K. C. Young, and K. B. Whaley, Phys. Rev. A **86**, 012317 (2012)

[69] A. del Campo, I. L. Egusquiza, M. B. Plenio, and S. F. Huelga, Phys. Rev. Lett. **110**, 050403 (2013)

[70] L. B. Levitin and T. Toffoli, Phys. Rev. Lett. **103**, 160502 (2009)

[71] N. Margolus and L. B. Levitin, Physica D **120**, 188 (1998)

[72] K. Bhattacharyya, J. Phys. A **16**, 2993 (1983)

[73] M. H. Goerz, T. Calarco, and C. P. Koch, J. Phys. B **44**, 154011 (2011)

[74] D. P. DiVincenzo, Fortschr. Phys. **48**, 771 (2000)

[75] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998)

[76] D. P. DiVincenzo, Phys. Rev. A **51**, 1015 (1995)

[77] D. M. Reich, G. Gualdi, and C. P. Koch, Phys. Rev. A **88**, 042309 (2013)

[78] D. M. Reich, G. Gualdi, and C. P. Koch, Phys. Rev. Lett. **111**, 200401 (2013)

[79] D. J. Egger and F. K. Wilhelm, Phys. Rev. A **90**, 052331 (2014)

[80] M. Goerz, Optimization of a controlled phasegate for ultracold calcium atoms in an optical lattice, diploma thesis, Freie Universität Berlin, 2010

[81] B. Kraus and J. I. Cirac, Phys. Rev. A **63**, 062309 (2001)

[82] R. Feynman, Found. Phys. **16**, 507 (1986)

[83] D. D'Alessandro, *Introduction to quantum control and dynamics*, Taylor & Francis Ltd, Hoboken, NJ, 2007

[84] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Phys. Rev. A **52**, 3457 (1995)

[85] C. M. Dawson and M. A. Nielsen, Quantum Info. Comput. **6**, 81 (2006)

[86] D. Deutsch, A. Barenco, and A. Ekert, Proc. R. Soc. A **449**, 669 (1995)

[87] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Phys. Rev. Lett. **91**, 027903 (2003)

[88] G. Vidal and C. M. Dawson, Phys. Rev. A **69**, 010301 (2004)

[89] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Phys. Rev. Lett. **93**, 020502 (2004)

[90] Y. Makhlin, Quantum Inf. Process. **1**, 243 (2002)

[91] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Phys. Rev. A **67**, 042313 (2003)

[92] P. Watts, M. O'Connor, and J. Vala, Entropy **15**, 1963 (2013)

[93] A. Childs, H. Haselgrove, and M. Nielsen, Phys. Rev. A **68**, 052311 (2003)

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

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

[96] F. Bloch, Phys. Rev. **70**, 460 (1946)

[97] K.-A. Suominen, Open quantum systems and decoherence, in *Quantum Information and Coherence*, edited by E. Andersson and P. Öhberg, Scottish Graduate Series, pages 247–282, Springer, 2014

[98] D. Walls and G. Milburn, *Quantum Optics*, Springer-Verlag, Berlin, 1994

[99] H.-P. Breuer, E.-M. Laine, and J. Piilo, Phys. Rev. Lett. **103**, 210401 (2009)

[100] H.-P. Breuer, J. Phys. B **45**, 154001 (2012)

[101] S. Huelga and M. Plenio, Contemp. Phys. **54**, 181 (2013)

[102] A. Bartana, R. Kosloff, and D. J. Tannor, J. Chem. Phys. **99**, 196 (1993)

[103] D. A. Lidar, I. L. Chuang, and K. B. Whaley, Phys. Rev. Lett. **81**, 2594 (1998)

[104] L. Viola, E. M. Fortunato, M. A. Pravia, E. Knill, R. Laflamme, and D. G. Cory, Science **293**, 2059 (2001)

[105] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **82**, 2417 (1999)

[106] S. Deffner and E. Lutz, Phys. Rev. Lett. **111**, 010402 (2013)

[107] T. Caneva, M. Murphy, T. Calarco, R. Fazio, S. Montangero, V. Giovannetti, and G. E. Santoro, Phys. Rev. Lett. **103**, 240501 (2009)

[108] S. J. Devitt, W. J. Munro, and K. Nemoto, Rep. Prog. Phys. **76**, 076001 (2013)

[109] M. A. Nielsen, Phys. Lett. A **303**, 249 (2002)

[110] L. H. Pedersen, N. M. Møller, and K. Mølmer, Phys. Lett. A **367**, 47 (2007)

[111] A. Usman, M. Luján, L. Freeman, and J. Gurd, Performance evaluation of storage formats for sparse matrices in fortran, in *High Performance Computing and Communications*, edited by M. Gerndt and D. Kranzlmüller, volume 4208 of *Lecture Notes in Computer Science*, page 160, Springer, 2006

[112] I. S. Duff, A. M. Erisman, and J. K. Reid, *Direct Methods for Sparse Matrices*, Numerical Mathematics and Scientific Computation, Oxford University Press, 1989

[113] E. Anderson, Z. Bai, C. Bischof, L. S. Blackford, J. Demmel, J. J. Dongarra, J. Du Croz, S. Hammarling, A. Greenbaum, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999

[114] R. Kosloff, J. Chem. Phys. **92**, 2087 (1988)

[115] R. Kosloff, Quantum molecular dynamics on grids, in *Dynamics of Molecules and Chemical Reactions*, edited by R. E. Wyatt and J. Z. Zhang, Marcel Dekker, New York, 1996

[116] E. Fattal, R. Baer, and R. Kosloff, Phys. Rev. E **53**, 1217 (1996)

[117] V. Kokoouline, O. Dulieu, R. Kosloff, and F. Masnou-Seeuws, J. Chem. Phys. **110**, 9865 (1999)

[118] M. Nest and H.-D. Meyer, Chem. Phys. Lett. **352**, 486 (2002)

[119] K. Willner, O. Dulieu, and F. Masnou-Seeuws, J. Chem. Phys. **120**, 548 (2004)

[120] S. Kallush and R. Kosloff, Chem. Phys. Lett. **433**, 221 (2006)

[121] K. Singer, U. Poschinger, M. Murphy, P. Ivanov, F. Ziesel, T. Calarco, and F. Schmidt-Kaler, Rev. Mod. Phys. **82**, 2609 (2010)

[122] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, Dover, Mineola, NY, 2001

[123] J. D. Lambert, *Numerical Methods for Ordinary Differential Systems: The Initial Value Problem*, John Wiley & Sons, New York, NY, 1991

[124] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN; The Art of Scientific Computing*, Cambridge University Press, New York, NY, 1993

[125] A. Leggett, Rev. Mod. Phys. **73**, 307 (2001)

[126] J. Demmel, O. Marques, B. Parlett, and C. Vömel, SIAM J. Sci. Comput. **30**, 1508 (2008)

[127] H. Tal-Ezer and R. Kosloff, J. Chem. Phys. **81**, 3967 (1984)

[128] A. Gil, J. Segura, and N. Temme, *Numerical Methods for Special Functions*, Society for Industrial and Applied Mathematics, 2007

[129] M. Ndong, H. Tal-Ezer, R. Kosloff, and C. P. Koch, J. Chem. Phys. **130**, 124108 (2009)

[130] W. E. Arnoldi, Q. Appl. Math. **9**, 17 (1951)

[131] M. Ndong, H. Tal-Ezer, R. Kosloff, and C. P. Koch, J. Chem. Phys. **132**, 064105 (2010)

[132] G. Ashkenazi, R. Kosloff, S. Ruhman, and H. Tal-Ezer, J. Chem. Phys. **103**, 10005 (1995)

[133] R. Kosloff, Annu. Rev. Phys. Chem. **45**, 145 (1994)

[134] M. Berman, R. Kosloff, and H. Tal-Ezer, J. Phys. A **25**, 1283 (1992)

[135] W. Huisinga, L. Pesce, R. Kosloff, and P. Saalfrank, J. Chem. Phys. **110**, 5538 (1999)

[136] H. Tal-Ezer, SIAM J. Sci. Comput. **29**, 2426 (2007)

[137] L. Reichel, BIT **30**, 332 (1990)

[138] J. P. Palao and R. Kosloff, Phys. Rev. A **68**, 062308 (2003)

[139] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1989

[140] J. Kennedy and R. Eberhart, Particle swarm optimization, in *Proceedings of the IEEE International Conference on Neural Networks*, volume 4, page 1942, 1995

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

[142] J. A. Nelder and R. Mead, Comput. J. **7**, 308 (1965)

[143] A. M. Weiner, Rev. Sci. Instrum. **71**, 1929 (2000)

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

[145] T. Caneva, T. Calarco, and S. Montangero, Phys. Rev. A **84**, 022326 (2011)

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

[147] C. G. Broyden, IMA J. Appl. Math. **6**, 76 (1970)

[148] R. Fletcher, Comput. J. **13**, 317 (1970)

[149] D. Goldfarb, Math. Comput. **24**, 23 (1970)

[150] D. F. Shanno, Math. Comput. **24**, 647 (1970)

[151] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, A limited memory algorithm for bound constrained optimization, Technical report nam-08, Northwestern University, 1994

[152] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, ACM Trans. Math. Softw. **23**, 550 (1997)

[153] P. de Fouquières, S. Schirmer, S. Glaser, and I. Kuprov, J. Magnet. Res. **212**, 412 (2011)

[154] A. Konnov and V. Krotov, Autom. Rem. Contr. **60**, 1427 (1999)

[155] D. M. Reich, M. Ndong, and C. P. Koch, J. Chem. Phys. **136**, 104103 (2012)

[156] J. P. Palao, R. Kosloff, and C. P. Koch, Phys. Rev. A **77**, 063412 (2008)

[157] S. E. Sklarz and D. J. Tannor, Phys. Rev. A **66**, 053619 (2002)

[158] J. P. Palao, D. M. Reich, and C. P. Koch, Phys. Rev. A **88**, 053409 (2013)

[159] D. M. Reich, J. P. Palao, and C. P. Koch, J. Mod. Opt. **61**, 822 (2014)

[160] M. M. Müller, H. Haakh, T. Calarco, C. P. Koch, and C. Henkel, Quantum Inf. Process. **10**, 771 (2011)

[161] R. Eitan, M. Mundt, and D. J. Tannor, Phys. Rev. A **83** (2011)

[162] T. E. Skinner and N. I. Gershenzon, J. Magnet. Res. **204**, 248 (2010)

[163] M. H. Goerz, E. J. Halperin, J. M. Aytac, C. P. Koch, and K. B. Whaley, Phys. Rev. A **90**, 032329 (2014)

[164] K. Kobzar, T. E. Skinner, N. Khaneja, S. J. Glaser, and B. Luy, J. Magn. Reson. **170**, 236 (2004)

[165] K. Kobzar, T. E. Skinner, N. Khaneja, S. J. Glaser, and B. Luy, J. Magn. Reson. **194**, 258 (2008)

[166] M. Saffman, T. G. Walker, and K. Mølmer, Rev. Mod. Phys. **82**, 2313 (2010)

[167] T. Wilk, A. Gaëtan, C. Evellin, J. Wolters, Y. Miroshnychenko, P. Grangier, and A. Browaeys, Phys. Rev. Lett. **104**, 010502 (2010)

[168] L. Isenhower, E. Urban, X. L. Zhang, A. T. Gill, T. Henage, T. A. Johnson, T. G. Walker, and M. Saffman, Phys. Rev. Lett. **104**, 010503 (2010)

[169] M. DePue, C. McCormick, S. Winoto, S. Oliver, and D. Weiss, Phys. Rev. Lett. **82**, 2262 (1999)

[170] S. Bergamini, B. Darquié, M. Jones, L. Jacubowiez, A. Browaeys, and P. Grangier, J. Opt. Soc. Am. B **21**, 1889 (2004)

[171] A. Gaëtan, Y. Miroshnychenko, T. Wilk, A. Chotia, M. Viteau, D. Comparat, P. Pillet, A. Browaeys, and P. Grangier, Nature Phys. **5**, 115 (2009)

[172] S. Whitlock, R. Gerritsma, T. Fernholz, and R. Spreeuw, New J. Phys. **11**, 023021 (2009)

[173] K. D. Nelson, X. Li, and D. S. Weiss, Nature Phys. **3**, 556 (2007)

[174] W. S. Bakr, J. I. Gillen, A. Peng, S. Fölling, and M. Greiner, Nature **462**, 74 (2009)

[175] J. Kruse, C. Gierl, M. Schlosser, and G. Birkl, Phys. Rev. A **81**, 060308 (2010)

[176] C. Weitenberg, M. Endres, J. F. Sherson, M. Cheneau, P. Schauß, T. Fukuhara, I. Bloch, and S. Kuhr, Nature **471**, 319 (2011)

[177] E. Urban, T. A. Johnson, T. Henage, L. Isenhower, D. D. Yavuz, T. G. Walker, and M. Saffman, Nature Phys. **5**, 110 (2009)

[178] D. Jaksch, J. I. Cirac, P. Zoller, S. L. Rolston, R. Côté, and M. D. Lukin, Phys. Rev. Lett. **85**, 2208 (2000)

[179] M. M. Müller, M. Murphy, S. Montangero, T. Calarco, P. Grangier, and A. Browaeys, Phys. Rev. A **89**, 032334 (2014)

[180] X. Zhang, A. Gill, L. Isenhower, T. Walker, and M. Saffman, Phys. Rev. A **85**, 042310 (2012)

[181] E. Brion, L. H. Pedersen, and K. Mølmer, J. Phys. A **40**, 1033 (2007)

[182] D. D. B. Rao and K. Mølmer, Phys. Rev. A **89**, 030301 (2014)

[183] U. Gaubatz, P. Rudecki, S. Schiemann, and K. Bergmann, J. Chem. Phys. **92**, 5363 (1990)

[184] D. Møller, L. B. Madsen, and K. Mølmer, Phys. Rev. Lett. **100**, 170504 (2008)

[185] I. I. Beterov, M. Saffman, E. A. Yakshina, V. P. Zhukov, D. B. Tretyakov, V. M. Entin, I. I. Ryabtsev, C. W. Mansell, C. MacCormick, S. Bergamini, and M. P. Fedoruk, Phys. Rev. A **88**, 010303 (2013)

[186] P. Aliferis, D. Gottesman, and J. Preskill, Quant. Inf. Comput. **8**, 181 (2008)

[187] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. A **67**, 052109 (2003)

[188] T. Corcovilos and D. S. Weiss, Rydberg calculations, private communication

[189] B. W. Shore, *Manipulating Quantum Structures Using Laser Pulses*, Cambridge University Press, 2011

[190] K. Bergmann, H. Theuer, and B. W. Shore, Rev. Mod. Phys. **70**, 1003 (1998)

[191] G. Grigoryan and Y. Pashayan, Opt. Comm. **198**, 107 (2001)

[192] M. Saffman, Rydberg shifts, private communication

[193] M. M. Müller, D. M. Reich, M. Murphy, H. Yuan, J. Vala, K. B. Whaley, T. Calarco, and C. P. Koch, Phys. Rev. A **84**, 042315 (2011)

[194] E. Campani, G. Degan, and G. Gobini, Lett. Nuovo Cim. **23**, 187 (1978)

[195] M. Ortiz and J. Campos, J. Quant. Spectr. Rad. Trans. **26**, 107 (1981)

[196] M. H. Goerz, D. M. Reich, and C. P. Koch, New J. Phys. **16**, 055012 (2014)

[197] M. J. Wright, P. L. Gould, and S. D. Gensemer, Rev. Sci. Instrum. **75**, 4718 (2004)

[198] D. Gottesman, Quantum Info. Comput. **14**, 1338 (2014)

[199] B. W. Reichardt, Algorithmica **55**, 517 (2009)

[200] M. Tinkham, *Introduction to Superconductivity*, McGraw-Hill, 1996

[201] M. H. Devoret and J. M. Martinis, Quantum Inf. Process. **3**, 163 (2004)

[202] L. Frunzio, A. Wallraff, D. Schuster, J. Majer, and R. Schoelkopf, IEEE Trans. on Appl. Superc. **15**, 860 (2005)

[203] P. J. Hirst and R. G. Humphreys, High temperature superconductor films: processing techniques, in *Handbook of Superconducting Materials*, edited by D. A. Cardwell and D. S. Ginley, volume I, Institute of Physics Publishing, 2003

[204] B. D. Josephson, Adv. Phys. **14**, 419 (1965)

[205] M. H. Devoret, Quantum fluctuations in electrical circuits, in *Quantum Fluctuations*, edited by S. Reynaud, E. Giacobino, and J. Zinn-Justin, Les Houches, Session LXIII, 1995, pages 353–386, Elsevier, 1997

[206] Y. Nakamura, Y. A. Pashkin, T. Yamamoto, and J. S. Tsai, Phys. Rev. Lett. **88**, 047901 (2002)

[207] M. H. Devoret, A. Wallraff, and J. M. Martinis, arXiv:0411174 (2004)

[208] V. Bouchiat, D. Vion, P. Joyez, D. Esteve, and M. H. Devoret, Phys. Scripta **1998**, 165 (1998)

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

[210] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. Devoret, Superconducting quantum bit based on the cooper pair box, in *New Directions in Mesoscopic Physics (Towards Nanoscience)*, edited by R. Fazio, V. Gantmakher, and Y. Imry, volume 125 of *NATO Science Series*, pages 173–195, Springer Netherlands, 2003

[211] D. I. Schuster, *Circuit Quantum Electrodynamics*, PhD thesis, Yale University, 2007

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

[213] A. Cottet, *Implementation of a quantum bit in superconducting circuit*, PhD thesis, Université Paris VI, 2002

[214] J. M. Chow, *Quantum Information Processing with Superconducting Qubits*, PhD thesis, Yale University, 2010

[215] S. M. Girvin, Superconducting qubits and circuits: Artificial atoms coupled to microwave photons, Lectures delivered at Ecole d'Eté Les Houches, 2011

[216] S. Richer, Perturbative analysis of two-qubit gates on transmon qubits, Master's thesis, RWTH Aachen, 2013

[217] L. S. Bishop, *Circuit Quantum Electrodynamics*, PhD thesis, Yale University, 2010

[218] F. Beaudoin, J. M. Gambetta, and A. Blais, Phys. Rev. A **84**, 043832 (2011)

[219] B. R. Johnson, *Controlling Photons in Superconducting Electrical Circuits*, PhD thesis, Yale University, 2011

[220] A. A. Houck, H. E. Tureci, and J. Koch, Nature Phys. **8**, 292 (2012)

[221] J. Majer, J. M. Chow, J. M. Gambetta, J. Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Nature **449**, 443 (2007)

[222] S. Poletto, J. M. Gambetta, S. T. Merkel, J. A. Smolin, J. M. Chow, A. D. Córcoles, G. A. Keefe, M. B. Rothwell, J. R. Rozen, D. W. Abraham, C. Rigetti, and M. Steffen, Phys. Rev. Lett. **109**, 240505 (2012)

[223] D. Basilewitsch, Theoretische Untersuchung von Quantengattern für Transmon-Qubits, Bachelor thesis, Universität Kassel, 2013

[224] G. Zhu, D. G. Ferguson, V. E. Manucharyan, and J. Koch, Phys. Rev. B **87**, 024510 (2013)

[225] A. Dewes, F. R. Ong, V. Schmitt, R. Lauro, N. Boulant, P. Bertet, D. Vion, and D. Esteve, Phys. Rev. Lett. **108**, 057002 (2012)

[226] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Nature **460**, 240 (2009)

[227] J. M. Chow, J. M. Gambetta, A. W. Cross, S. T. Merkel, C. Rigetti, and M. Steffen, New J. Phys. **15**, 115012 (2013)

[228] P. J. Leek, S. Filipp, P. Maurer, M. Baur, R. Bianchetti, J. M. Fink, M. Göppl, L. Steffen, and A. Wallraff, Phys. Rev. B **79**, 180511 (2009)

[229] A. Blais, J. Gambetta, A. Wallraff, D. I. Schuster, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, Phys. Rev. A **75**, 032329 (2007)

[230] H. Häffner, C. Roos, and R. Blatt, Phys. Rep. **469**, 155 (2008)

[231] F. Schmidt-Kaler, H. Häffner, M. Riebe, S. Gulde, G. P. T. Lancaster, T. Deuschle, C. Becher, C. F. Roos, J. Eschner, and R. Blatt, Nature **422**, 408 (2003)

[232] J. M. Gambetta, in *Bulletin of the American Physical Society, APS March Meeting 2012*, volume 57, 2012

[233] C. Zener, Proc. R. Soc. A **137**, 696 (1932)

[234] M. V. Berry, Proc. R. Soc. A **392**, 45 (1984)

[235] P. Zanardi and M. Rasetti, Phys. Rev. A **264**, 94 (1999)

[236] S. M. Girvin and A. Shabani, private communication

[237] D. J. Egger and F. K. Wilhelm, Supercond. Sci. Technol. **27**, 014001 (2014)

[238] D. Reich, Optimising the nonlocal content of a two-qubit gate, diploma thesis, Freie Universität Berlin, 2010

[239] P. Watts, J. Vala, M. M. Müller, T. Calarco, K. B. Whaley, D. M. Reich, M. H. Goerz, and C. P. Koch, arXiv:1412.7347 (2014)

[240] M. H. Goerz, G. Gualdi, D. M. Reich, C. P. Koch, F. Motzoi, K. B. Whaley, J. Vala, M. M. Müller, S. Montangero, and T. Calarco, arXiv:1412.7350 (2014)

[241] M. Musz, M. Kuś, and K. Życzkowski, Phys. Rev. A **87**, 022111 (2013)

[242] C. Rigetti, J. M. Gambetta, S. Poletto, B. L. T. Plourde, J. M. Chow, A. D. Córcoles, J. A. Smolin, S. T. Merkel, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Phys. Rev. B **86**, 100506 (2012)

[243] D. M. Reich, N. Katz, and C. P. Koch, arXiv:1409.7497 (2014)

[244] D. M. Reich, *Efficient Characterisation and Optimal Control of Open Quantum Systems. Mathematical Foundations and Physical Applications*, PhD thesis, Universität Kassel, 2015

[245] Y. Ohtsuki, New J. Phys. **12**, 045002 (2010)

[246] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser, J. Phys. B **44**, 154013 (2011)

[247] H. F. Hofmann, Phys. Rev. Lett. **94**, 160504 (2005)

[248] J. Fiurášek and M. Sedlák, Phys. Rev. A **89**, 012323 (2014)

[249] J. Lawrence, Č. Brukner, and A. Zeilinger, Phys. Rev. A **65**, 032320 (2002)

[250] S. Bandyopadhyay, P. O. Boykin, V. P. Roychowdhury, and F. Vatan, Algorithmica **34**, 512 (2002)

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

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

[253] M. J. Peterer, S. J. Bader, X. Jin, F. Yan, A. Kamal, T. J. Gudmundsen, P. J. Leek, T. P. Orlando, W. D. Oliver, and S. Gustavsson, Phys. Rev. Lett. **114**, 010501 (2015)

[254] J. Gambetta, A. Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **74**, 042318 (2006)

[255] J. Gambetta, A. Blais, M. Boissonneault, A. A. Houck, D. I. Schuster, and S. M. Girvin, Phys. Rev. A **77**, 012112 (2008)

[256] M. Boissonneault, A. C. Doherty, F. R. Ong, P. Bertet, D. Vion, D. Esteve, and A. Blais, Phys. Rev. A **85**, 022305 (2012)

[257] F. Motzoi, private communication

[258] J. M. Martinis, K. B. Cooper, R. McDermott, M. Steffen, M. Ansmann, K. D. Osborn, K. Cicak, S. Oh, D. P. Pappas, R. W. Simmonds, and C. C. Yu, Phys. Rev. Lett. **95**, 210503 (2005)

[259] Y. Shalibo, Y. Rofe, D. Shwa, F. Zeides, M. Neeley, J. M. Martinis, and N. Katz, Phys. Rev. Lett. **105**, 177001 (2010)

[260] H. Krovi, O. Oreshkov, M. Ryazanov, and D. A. Lidar, Phys. Rev. A **76**, 052117 (2007)

[261] E. Ferraro, H.-P. Breuer, A. Napoli, M. A. Jivulescu, and A. Messina, Phys. Rev. B **78**, 064309 (2008)

[262] S. Lorenzo, F. Plastina, and M. Paternostro, Phys. Rev. A **87**, 022317 (2013)

[263] P. Rebentrost, I. Serban, T. Schulte-Herbrüggen, and F. K. Wilhelm, Phys. Rev. Lett. **102**, 090401 (2009)

[264] E. Asplund and T. Klüner, Phys. Rev. Lett. **106**, 140404 (2011)

[265] R. Schmidt, A. Negretti, J. Ankerhold, T. Calarco, and J. T. Stockburger, Phys. Rev. Lett. **107**, 130404 (2011)

[266] F. F. Floether, P. de Fouquières, and S. G. Schirmer, New J. Phys. **14**, 073023 (2012)

[267] G. Gualdi and C. Koch, Phys. Rev. A **88**, 022122 (2013)

[268] E. O. Brigham, *The Fast Fourier Transform and Its Applications*, Prentice-Hall, Upper Saddle River, NJ, 1988

[269] P. Duhamel and M. Vetterli, Sig. Process. **19**, 259 (1990)

[270] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Academic Press Professional, San Diego, CA, 1990

Index

Hiermit versichere ich, dass ich die vorliegende Dissertation selbstständig, ohne unerlaubte Hilfe Dritter angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Dritte waren an der inhaltlich-materiellen Erstellung der Dissertation nicht beteiligt; insbesondere habe ich hierfür nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden.

Mai 2015, Kassel

Michael Goerz