Controlling Quantum Information Devices

by

Felix Motzoi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2012

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

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

Abstract

Quantum information and quantum computation are linked by a common mathematical and physical framework of quantum mechanics. The manipulation of the predicted dynamics and its optimization is known as quantum control. Many techniques, originating in the study of nuclear magnetic resonance, have found common usage in methods for processing quantum information and steering physical systems into desired states. This thesis expands on these techniques, with careful attention to the regime where competing effects in the dynamics are present, and no semi-classical picture exists where one effect dominates over the others. That is, the transition between the diabatic and adiabatic error regimes is examined, with the use of such techniques as time-dependent diagonalization, interaction frames, average-Hamiltonian expansion, and numerical optimization with multiple time-dependences. The results are applied specifically to superconducting systems, but are general and improve on existing methods with regard to selectivity and crosstalk problems, filtering of modulation of resonance between qubits, leakage to non-compuational states, multi-photon virtual transitions, and the strong driving limit.

Acknowledgements

This research would not have been possible without the support and hard work of my collaborators and supervisors. I would like to thank my supervisor Frank Wilhelm for his openness to taking a chance on me on as a graduate student. His leadership, methodology, and insight have all had a significant impact on my personal and academic development. My secondary supervisor, Jay Gambetta, has also been instrumental to the main lines of thought in this research and his vision and guidance as well as his perseverence have opened many opportunities (and driven some home run balls!). I would also like to thank Seth Merkel for his guidance and feedback. I have also been fortunate to help assist other graduate students in common research. Peter Groskowski, Botan Khani, Amira Eltony, Yuval Sanders, Likun Hu, Victor Vetch, and Jerry Chow have all collaborated in quantum control projects that have generated many stimulating and valuable conversations. Finally, I would like to thank the fellow graduate students in 2117 for some freshness and brightness in a room that had neither. This research was made possible by NSERC through the discovery grants, QuantumWorks, and a PGS scholarship.

Dedication

This thesis is dedicated to my three muses: my mother Ina who would have written it herself if she needed to, my sister Clairneige who tirelessly listened to rants and ravings about the endless "wonders" of quantum physics growing up, and my girlfriend Kanako who carried me through these arduous Ph.D. years.

Contents

List of Figures

Chapter 1

Introduction

The digital age has brought us the relentless digitization of information. At its very heart, the framework given by the Universal Turing Machine (UTM) and the von Neumann architecture tells us that any piece of information and any algorithm that acts on it can be stored together as a single string of zeros and ones on one memory device [160]. This equivalence principle states that some piece of information "x" that we have cannot be distinguished in and of itself from some algorithm or function "$f(x)$" that acts on it. That the one represents the physical change in the other has no repercussions in terms of its representation or storage.

Beyond this digital paradigm lie a plethora of characterizations and interpretations of what is possible with a digital computer, but fundamentally they are all reducible to operation of a UTM. That is, different statistical measures of 'efficiency' (such as the entropy, outcome probability, information dissipation, or the algorithmic complexity) rely on the computation as a sequence of indivisible steps for which both the information x and the process acting on it $f(x)$ are completely fixed. And yet the statistical properties are not digital but real-valued and moreover actually changing during the atomic steps. Whether such 'statistical properties' of information are physically real was perhaps first questioned by the Greek philosopher Zeno in his arrow paradox [118], which asks whether something can be in one place and moving at the same time. But, because physically observable quantities are effectively continuous (as formalized by calculus), one being the average of the other is not discernible (they are independent variables [99]) and such speculation remained philosphical.

Nevertheless, classical laws of probability and logic are based on notions of discontinuous events and objects (divisible into smaller and smaller *particles*) and these concepts made their way into our understanding of physics. Based on the motion of

individual particles, a statistical physical theory was developed behaving similarly to (but discovered independently from) modern statistical characterizations of digital information, obeying a set of thermodynamic laws with similar concepts of entropy, heat dissipation, and other probabilistic effects. Yet, evidence of the continuousness of physical phenomena remained a thorn in these developments, and in particular a *wave* theory was concurrently developed that characterized analog oscillations and interactions as propagations of contiguous waves.

Quantum mechanics [152] unified the notions of probabilistic properties of individual particles and continuous wave interactions between disconnected particles. In effect, both heat dissipation and wave propagation are dictated by the same Schrödinger equation. Whether something behaves more like a particle (with fixed properties) or more like a (delocalised) wave is primarily a question of *context*, that is, of whether it is bound to a particular location or whether it is free to propagate in all directions. Moreover, just like a physical potential barrier keeps a particle in a specific location, a measurement event causes the measured property to localize to a particular value. A lack of measurement or an escape from a binding potential on the other hand lead to dissipation and propagation of energy and information. The properties of these objects that are measured can also be discontinous (*quantum*) or continuous, depending on whether they describe a bound state or a free one. In the quantum case it seems Zeno's intuition was correct: both a given discrete property and a complementary property describing its change cannot simultaneoulsy be well defined. However, both properties are physically real, and context-specific localisation dictates which of them is well-defined.

A characterization of computers that somehow physically includes both the information (that we read off digitally) and also the change in the information, described by the analog physical process that (infinitesimally) changes one piece of information that we observe into another, seems like it should be qualitatively richer[1], much like quantum mechanics generalizes Newtonian dynamics. No less, such a formulation of computations is necessarily equivalent and isomorphic to quantum mechanics [66]. While the new paradigm does not alter the fundamental power of the UTM and the question of 'tractability' (the resolution of paradoxes such as the Halting Problem or "this statement is false" cannot be found on *any* computer), it turns out that measures of efficiency such as informational entropy, probability, and dissipation are resources that are physically inherently present in the characterization, but perhaps most interestingly, decreased algorithmic complexity is as well.

[1] At the very least, it should contain twice as much information. For example if the information is encoded in terms of a position coordinate, velocity information would also be present. See also 'superdense coding' [84].

For our purposes, we are interested in such complex algorithms. The most complex operations for a computer are computations about computations themselves. In the context of the first paragraph, for some information "x" and some function acting on it "$f(x)$", a complex operation can roughly be understood as a function "$g(f(x))$" that computes some global property of the function $f(x)$ for all possible x. The "brute force" approach to computing this can be described as trying every single possible value of x, computing $f(x)$, and from the total set of results computing $g(f(x))$. When x can take on many different values such an approach is all but impossible (it is exponential in the width of x). An insight in complexity arises from identifying the algorithm $f(x)$ with the *analog physical process* underlying computation. Then, global properties of the process $f(x)$ can be measured directly, using as input a wave spread over all possible values of x rather than feeding in the values individually. For example, one can measure the frequency that the physical process repeats itself with by appending another analog process (the Fourier transform) to the input and simply reading off the answer [158].

As it turns out, the frequency is the key piece of information to breaking encryption functions used in RSA cryptosystems, used all over the planet for security protocols. Similarly, the characteristic energy of the process (its *eigenvalue*) is instrumental to the simulation and characterization of physical systems [81], with such potential areas of applications as biochemistry and materials design. A crowning achievement but still an open problem would be if one could reliably obtain the characteristic input configuration (the *eigenvector*) for the lowest eigenvalue of a process (that one is trying to optimize), then this corresponds to a solution to a given instance of the notoriously hard but ubiquitous NP-complete problems [44,77].

Of course, building such a computer is in many ways far more demanding than a digital computer because the analog process describing operations must not only avoid accidentally flipping a bit, but it must retain the coherence of the phase oscillation (i.e. timing of the peaks and troughs of the wave). By the same token, a classical analog computer is limited, not only by a simple class of bit operations typically available for the chosen encoding, but because heat dissipation limits the precision of such operations to a finite and relatively small number of bits. Using quantum mechanics circumvents this problem with the important caveat that damping and dissipation must be suppressed to remain quantum. The quantum framework not only provides a mathematical model for describing (continuous) computation, but from a hardware point of view it provides numerous prospective physical implementations.

Candidate quantum implementations typically involve some small physical system that can be sufficiently isolated from the environment to allow its wave properties to be manifested and retained. Additionally, the system must be sufficiently cooled to

prevent flipping of bits via energy decay processes. However, this isolation must be contrasted to the ability to introduce energy/phase shifts via control hardware which must couple to the quantum system. The most developed technologies in this regard have benefited from research for many decades such as liquid state Nuclear Magnetic Resonance (NMR) and photonic systems. These systems have and continue to serve well as testbeds for advances in quantum information but are challenged by difficulties in weak coupling between elements and scaling to larger system sizes. Other more complicated systems such as trapped atoms/ions and solid state implementations (e.g. semiconducting, superconducting) are more promising in this regard. Some of these implementation and design choices are discussed in Sec. 2.3.

The mathematical framework of quantum computing consists of defining a quantum bit (qubit) as some (stable) degree of freedom present in the physical system which can be measured in one of two states, but for which the wave properties allow that energy (and information) can be spread over multiple states simultaneously. Timed operations are controlled by the dynamics of the Schrödinger wave equation which allows coupling of qubits to control hardware or other qubits. An overview on quantum mechanics and operations is given in Sec, 2.1. The Schrödinger evolution is typically controlled on a short time scale in order to define reusable high-quality operations (gates)[2], and (remarkably) these operations can be subsequently chained together as building blocks for more complicated operations and algorithms, just like digital gates on a conventional computer. When the timing is performed correctly, the coherence properties of the wave can spread over many qubits allowing for operations (such as the Fourier transform) to now physically act *simultaneously* on arbitrarily many bits[3].

The statement of this thesis is to describe the generation of quantum operations, acting on a small number of qubits, both in terms of the physics involved and the mathematical features that expand on classical computation. The main reason we focus on a small system size is that, by the very definition of the problem, engineering many-bit operations is not efficient using classical computers, which of course is all we have access to without a quantum computer. In fact, most known multi-qubit operations that are required for known algorithms can be decomposed into a series

[2]At this point, arbitrarily high-quality operations can be obtained using measurement of particular "error syndrome" states [62]. This is another difference from classical analog computation in which measurement is only defined in the (final) steady-state of the system. Note that in some cases, such as typical implementations of *adiabatic quantum computation,* long operations are synthesized directly instead and error correction is therefore not used, in particular if the evolution can be shown to withstand the onset of decay and damping processes.

[3]As another example, using a single operation, one qubit can be simultaneously *entangled* with a large number of other qubits. Measuring one qubit measures all the others as well.

of smaller operations. Thus, for the purpose of well-defined operations, the error can be made small by ensuring that the error in the short operations is proportionally smaller. Thus, the main focus of the research is to engineer these short operations to have low errors. Sec. 2.4 introduces some ways of quantifying the error of a quantum operation.

The eventual goal, in most idealized terms, is to build a quantum computer. In this respect, the quantum computer is built up of three main components. The first is the memory of the system, which can be digitally read out, and which is physically encoded in some quantum system. For example, this quantum information can be in the spin (up or down) of a particle, in the energy level of an atom, or in a quantized superconducting current. The second component is the "processor", which is some control hardware programmed to introduce radiation or other fields into the system to change the configuration of the information stored in the physical device. The device must effectively be able to perform the short operations with sufficiently low error, and so is central to this thesis. Thirdly, the software that runs on the computer is some algorithm, typically one that benefits in reduced complexity by running on a quantum computer. Some of the main developments in quantum software are described in Sec. 2.2. All three of these components are very far from being a finished product, let alone fully understood. As much as all three components have gained tremendously from theoretical research and a variety of analytical insights, they are difficult to understand even for small number of bits.

In the context of a small number of qubits, all three components can often be described using a common physical model described by the Hamiltonian of the system. Both by experimentation and by numerical simulation and optimization, it is hoped that new insights will be gained into the Hamiltonian dynamics that will help to scale the components to systems with more bits. In the case of numerics, solutions to reducing error will often have symmetries from which an analytical expression or technique can be deduced. Finding control fields that contain identifiable features such as specific frequency components or separation into a sequence of steps in time can offer clues about how best to characterize problems and their solutions. Such forms are also very useful because they can point out which physical resources are most valuable in the physical implementation and design choices can be based on these in the future. Finding that some imperfection is intrinsic or irreversible or instead that it can be avoided or rendered irrelevant can avoid unnecessary dead-ends in terms of research and allow resources to go to more promising realizations. From a processing point of view, being able to solve simple problems involving few bits analytically means that the solutions can be computed very efficiently and if multiple techniques can be combined then the solutions for larger numbers of bits

can often remain efficient. Specifically, solutions to different problems that involve distinct controls or different basis elements for the same control, or even different bases, can be applied effectively independently, at least to the highest order of any approximations that may be made in the solutions. In some cases, compound errors coming from combing two different kinds of error still belong to the set of correctable errors and can thus be removed to arbitrary accuracy by applying the analytic techniques in sequence. Moreover, the physics is all present in the interaction of very few components and adding in more components does not qualitatively alter the situation, only adds in more combinations. Finally, from a software point of view, simulations and optimizations can offer insight into how an algorithm works and even find faster ways of running the same algorithm, for example by speeding up the individual gates with a proportional decrease in the time of the algorithm or by combining several steps in the algorithm into a single step. Most notably, the sources of error in certain kinds of algorithms such as adiabatic computation impact directly (super-polynomially) on the algorithmic complexity of the solution [44, 77]. The analytical and numerical optimizations also have a direct effect on how and which experiments can be run. In some cases, the experiments even involve more bits than the simulations [21]. The same advantages that come from simulation can often be found in experimentation, as well as new benefits involved such as seeing exactly what can or cannot be realized. Optimizations allow more complex experiments and the effects are compounded, taking us closer to large scale quantum computation.

A more immediate purpose to the study of quantum operations is to map out the dynamics involved on a qualitative level. In a sense, such a study goes beyond a measure of quantitative error but rather aims to find novel (and often less complex) ways in which particular problems (given by a particular Hamiltonian) can be solved. In effect, quantum computation (unitary evolution) is conceptually different from the binary switching processes that take place in digital computation and understanding the similarities and differences is of much theoretical interest. Not only is the amount of information present in the description of the computational system exponentially larger, but as mentioned the change in information is as much a resource as the information itself. All this means that any path from one state to another can take on an infinite number of different trajectories through analog space. The basic study of quantum control, the study of which quantum variables can be used to change or instead retain information in an efficient way, was pioneered in NMR systems since the middle of last century [162]. This was the first experimental application of quantum mechanics exploiting not simply the ensemble phenomena described by the theory but manipulating and controlling individual pieces of quantum information as well. This progress has been steadily forthcoming as more precision and more complexity

has been incorporated into these operations. Initially it was found that irradiating atomic transitions with resonant magnetic fields, one could drive oscillation between a pair of energy states. This turned out to be useful for example in spectroscopy for identifying the chemical structure of samples. Timing was improved so that a particular state could reliably be obtained at a given time. This in turn allowed complex operations in time such that multiple (matrix) transitions could be sequentially traversed and multi-dimensional data could be obtained via spectroscopy such as the strength of the couplings between atoms in molecules. The wave properties of light-matter interaction was also fundamental to the progress. Pulse shaping beyond square monochromatic pulses improved reliability and spectral selectivity by for example using smoother shapes with a smaller more selective bandwidth [55]. The technology developed further with complex sequences of control pulses being used to manipulate information inside molecules and engineer specific chemical configurations. Thus, chemical analysis is where the basic ideas about quantum algorithms were first laid out, with interference playing a prominent role between desired and undesired pathways for information.

The success of the pioneering efforts in quantum control was largely predicated on a clear hierarchy of the importance of different operators in the system. Physically, the atomic spin precession frequencies were much larger than the amplitudes of magnetic fields, which were much larger than the strengths of the inter-atomic couplings in molecules, which were again much larger than damping and decay errors. In this limit, the effects could be separated and independently characterized. Thus, generating slow timescale evolutions could be understood as a pseudo-digital sequence of operations which permit information to move inside or in between spins, decomposing the dynamics into a matrix algebra. This lead, for example, to powerful techniques for using the faster single-spin dynamics to augment the inter-atomic couplings, such as *composite pulse sequences*, which seek to obtain a new operation from a number of smaller operations. Most generally, the effect of the extra controls can often be understood as changing (temporarily) the encoding (and for modeling purposes, the representation) of information from energy to phase or something in between. On the other hand, in the limit of the fast dynamics, the representation is better understood as a continuous phenomenon for which the classical wave properties such as bandwidth describe which information is affected. Yet such linear measures do not tell the whole story, in part because the solution to the Schrödinger equation (describing the spectroscopic response) is a trigonometric function rather than linear, and in part because the matrix mechanics allow information to be moved in a multi-dimensional manner.

In some sense, the combination of the analog and digital mechanics offers the most

interesting and difficult to predict dynamics but also the most promise in terms of directly steering a set of arbitrary initial conditions to desired ones. One effect, known as a Stark shift or a Bloch-Siegert shift, is to change the energies of all the levels in the Hamiltonian matrix by essentially time-dependently rediagonalizing it. However, even this effect is best understood from the separation of time scales between the frequency and amplitude of the pulse. The main difference with regard to this thesis is that we assume systems can now be built (many decades later than pioneering NMR) that can have arbitrarily shaped (analog) and arbitrarily strong couplings between the components, but, with no one term dominating in magnitude over others, very high precision control can in principle still be achieved. Thus the classical information that can be extracted from the dynamics is neither well defined in frequency nor in time and the matrix mechanics must be described in continuous time. In this sense, the regime that will be studied is neither "ultra-strong" coupling nor highly adiabatic (low coupling), but at the cusp of where the adiabatic approximation breaks down, which offers the interest both in terms of finding intuitive descriptions for the complex nature of simultaneous competing effects and practically in terms of pushing the envelope of what can be done in experiment. The techniques that will be used are specializations and refinements of the study of quantum control. Ch. 3 lists all the different control problems that will be analysed in later chapters in terms of which competing effects will be mathematically present and in terms of the physical systems considered. Analytic techniques used to describe and choose different pertinent representations are discussed in Ch, 4. Meanwhile, advances in being able to simulate and optimize quantum operations in an efficient way on classical computers have also blossomed with the field, and some relevant techniques are discussed in Ch. 5.

The problems that are addressed in this thesis will in general be problems associated with this difficult regime, where multiple competing (non-commuting) effects are roughly of the same magnitude, and in systems where this regime exists in part because fast dynamics within and between components can be largely engineered. The rest of the chapters deal with specific kinds of errors that can arise and so are grouped by, in a sense, which operations they affect (thus mathematical or physical commonalities may be found across chapters). Ch. 6 analyses what happens when the variation in some control parameter due to the classical electronics is roughly as slow as the speed of the actual operation. In many physical applications, this is the limiting factor that sets the gate time, but also if changes cannot occur approximately instantaneously then the control terms can cause new errors by not commuting with the rest of the Hamiltonian. Ch. 7 discusses the problem of having crosstalk error, namely controls which unwittingly couple to more than one quantum element.

Sec. 7.1 discusses the case when the frequency difference between elements is roughly on the order of the amplitude of the control while in Sec. 7.2 the amplitude of the control varies from one element to the next (but not necessarily the frequency). Ch. 8 looks at one specific quantum element, but now the element is allowed to contain multiple frequencies, only one of which should be resonant with the control. Once again, the energy difference is roughly of the same order as the amplitude of the control. Ch. 9 continues with the same kind of error, but now compound errors are considered where multiple unwanted transitions are present in the element(s). Ch. 10 also continues along the same tack, but now multiple transitions inside a system are used to define new transitions that do not intrinsically exist already. The cost is that the old transitions are still present but once again at a frequency difference, which once again can be as small or smaller than the amplitude of the control. The last form of strong coupling that is considered is in Ch. 11, where now the amplitude of a single control and its frequency are of the same order and shows that even in one of the simplest quantum systems the dynamics are far from trivial. Finally, in Ch. 12, the findings are summarized and conclusions drawn.

Chapter 2

Universal Quantum Computation

This chapter gives an overview of the main principles and notation used in quantum computation and applied throughout this thesis. For detailed surveys of quantum computation and quantum information see Refs. [129] and [84].

2.1 Quantum gates

2.1.1 Superposition and measurement

Quantum computation differs from conventional computation in that it is able to satisfy two postulates of quantum mechanics simultaneously.

The first is that conventional digital (bit) readout signals 0 and 1 (represented as the $\binom{1}{0}$ and $\binom{0}{1}$ vector basis states, respectively) can be defined simultaneously (as in the *qubit* $\binom{a}{b}$). Specifically, unlike conventional logical operations, the physical processes that map one basis state to another should also be well defined at all intermediate times during the operation. In order for such a decomposition to be possible, the operations must inherently be described using complex numbers, e.g.,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} (1+i)/2 & (1-i)/2 \\ (1-i)/2 & (1+i)/2 \end{pmatrix} \times \begin{pmatrix} (1+i)/2 & (1-i)/2 \\ (1-i)/2 & (1+i)/2 \end{pmatrix}.$$

More abstractly, in the context of group theory where we take operations to be reversible and of unit determinant, instead of being described as permutations (digital computation) or stochastic matrices (probabilistic computation), quantum operations are instead described by the unitary group [66]. Thus, we must have that a and b are complex numbers, which are analogous to amplitudes of a physical wave, and the vector $\binom{a}{b}$ is called the wave vector.

The second constraint is the basis states must be distinguishable. As entrenched in our every day experience, where for example a car is or is not drivable or a woman may or may not be pregnant, two physical states being distinguishable implies that they cannot simultaneously both be true. This distinguishibility in turn ensures that any intermediary states have no interpretation as measurable elements of reality. This is same condition that is true for conventional digital computers (where signals are either on or off) but differs from models of conventional analog computation (though, in practice, limited precision renders measurement digital as well). In the context of the reality of the basis states, this forces us to think of the complex numbers a, b as somehow describing probabilities. The measure chosen by nature (and the only one conserved by unitaries [1]) is the distribution given by 2-norm (with $|a|^2 + |b|^2 = 1$). In effect, an (intended or otherwise) "instantaneous" act of measurement of an intermediary state must impart an *a posteriori* reality to the distinct computational vector states, either $\binom{1}{0}$ with probability $|a|^2$ or $\binom{0}{1}$ with probability $|b|^2$. In between, both realities must somehow be possible.

While these two postulates of quantum mechanics may seem at odds with one another, with objects in effect enacting both waves and particles, they are both well established and fundamental to quantum computation.

2.1.2 Unitary evolution

The advantage of this paradigm does not come from the speed or physical size of the system, nor solely from the parallelism of having probability amplitude distributions over large numbers of possible states (since we can only measure one at the end), but rather from the otherwise unattainable instruction set made possible by (arbitrarily precise) unitary state evolution [1]. Detailing the dynamics of creating these fundamental building blocks is central to this thesis.

For a single bit register (a qubit), the unitary operations can be thought of as rotations, of which the most prevalent are

$$U_\theta^X \equiv \begin{pmatrix} \cos\theta/2 & i\sin\theta/2 \\ i\sin\theta/2 & \cos\theta/2 \end{pmatrix}, U_\theta^Y \equiv \begin{pmatrix} \cos\theta/2 & -i\sin\theta/2 \\ i\sin\theta/2 & \cos\theta/2 \end{pmatrix}, U_\theta^Z \equiv \begin{pmatrix} \exp i\theta/2 & 0 \\ 0 & \exp -i\theta/2 \end{pmatrix} \tag{2.1}$$

These generate the only non-trivial classical single bit operation, the negation (NOT) gate

[1] The measurement operation is also useful for such techniques as error-correction and teleportation which can in practice improve operational precision (see Sec. 2.4) as well as probabilistic algorithms but strictly speaking is not fundamental to the computation.

$$-iU_{\pi}^{X} = X \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

These rotation operations are critical to quantum computation because they create the intermediate non-computational (non-basis) states discussed in the last section. That is, the pure (classical) computational state $|\psi\rangle$ can be transformed to the *superposition* e.g. $U_{\theta}^{X}|\psi\rangle = (\cos(\theta/2)\mathbb{1} + i\sin(\theta/2)X)|\psi\rangle$. Thus, we can see that the instruction set for quantum operations is much richer.

For multiple registers, the outcome probabilities cross multiply which we capture using tensor notation. That is, two qubits with probability amplitudes $\begin{pmatrix} a \\ b \end{pmatrix}$ and $\begin{pmatrix} c \\ d \end{pmatrix}$ respectively give cross amplitudes $\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac \\ bc \\ ad \\ bd \end{pmatrix}$. Again, the two-bit classical conditional-NOT (a.k.a. XOR) operation $(x, y) \rightarrow (x, x \oplus y)$ which reads in matrix notation

$$CNOT \equiv CX \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

generalizes to the class of controlled operations $CU = \mathbb{1}_2 \oplus U_2$, where U_2 is any unitary operation of dimension 2 and $\mathbb{1}_2$ is the identity matrix of dimension 2. In combination with the superposition operators, these unitary operations allow the superpositions to extend over multiple computational elements. Similarly, there is another important class of two bit gates [60, 155, 189], of which the most common is

$$ISWAP \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The states containing cross-probabilities between qubits that can be obtained from these two classes of operations (as well as others), combined with the measurement postulate outlined above imply that exotic states of matter can be obtained. The fact that the wavevector amplitues can spread out over two or more physical elements (e.g. with cross-probabilities $\{(ac)^2, (bc)^2, (ad)^2, (bd)^2\}$) suggests that a shared states exists between them regardless of how far apart they are. In particular,

when only ac and bd are non-zero (or bc and ad), measurement correlations between the two objects (where the random value of the first bit predicts the second) can be seen to be stronger than if one were to assume that the randomness in the shared state was pre-determined (regardless of spatial distance). Such *entangled* states have been experimentally verified.

Higher dimensional operations can also be defined, most notably N-qubit controlled-U operations ($C_N U_2 = \mathbb{1}_{2^N-2} \oplus U_2$), Hadamard operations $\sqrt{\frac{1}{2^N}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{\otimes N}$, and the quantum Fourier transform transform (mapping basis elements $y_k \mapsto \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{2\pi i jk/N} x_j$) which generalizes the Hadamard operation.

2.1.3 Hamiltonians

To obtain the unitary evolution needed above we need to consider how to obtain it from the physical system in use. To this end, we should be able to gauge the different physical components that make up the system. In this mathematical treatment, a unitary matrix can more conveniently be defined as the exponential of an anti-Hermitian operator, or equivalently by $U = \exp(-i\hat{H}t/\hbar)$, where $\hat{H}$ is a Hermitian operator known as the Hamiltonian and t parametrizes the evolution time. The Hamiltonian gives the intrinsic energies of the system, both for the basis states and for the interactions that allow transfer of occupancy from one state to another. In matrix notation, the energy of the states is given along the diagonal (telling how favourable a state is), while off-diagonal elements give the coupling energies between states which tell how fast one state is replaced by another. As mentioned, the physical time scale is irrelevant to any exponential speed-up, and as such we often use dimensionless units as well as take $\hbar = 1$. In general the unitary operation that is performed will change over time and is described as the matrix product of a sequence of infinitesimal unitary operators. This is written by convention as $U = \mathcal{T} \exp\left(-i \int \hat{H}(t) dt\right)$, where the infinite sequence of infinitesimal operators is equivalent to the use of the time-ordering operator $\mathcal{T}$, which can be used to expand the integrand in an analytically tractable way. Alternatively, the Hamiltonian defines the Schrödinger equation $\dot{U} = -i\hat{H}U$ whose integration gives the evolution in time. Many different Hamiltonians will give the same evolution, a key point that will be revisited when complications arise.

In the absence of complications (that will arise), a constant $\hat{H}$, or an $\hat{H}$ that satisfies the commutation relation $[\hat{H}(t), \hat{H}(t')] = 0, \forall t, t'$, will simply give $U = \exp(-i \int \hat{H}(t) dt)$ with which it is easy to predict and even engineer gate operations. Such expressions (approximate or exact) are the reason it is more convenient to work

with the Hamiltonian than the evolution operator, as the calculations involve linear, rather than exponential functions of matrices. In particular, given the Hermiticity of the Hamiltonian, the Hamiltonian of a qubit with unique time dependence must necessarily be of the form

$$\hat{H}(t) = \Omega(t)(a\hat{X} + b\hat{Y} + c\hat{Z}) \tag{2.2}$$

with Ω the Rabi frequency, $\sqrt{a^2 + b^2 + c^2} = 1$, and $\hat{X}, \hat{Y}, \hat{Z}$ the Pauli operators. The unitary evolution at time T can readily be computed via

$$U(0,T) = \hat{\mathbb{1}} \cos\left[\int_0^T \Omega(t)dt\right] - \frac{i}{\Omega}\hat{H} \sin\left[\int_0^T \Omega(t)dt\right] \tag{2.3}$$

This is known as an *area theorem*. It is easy to see that picking a, b, c and the area $\int_0^T \Omega(t)dt$ appropriately it is possible to generate the rotation operators from Eq. 2.1, which fully span the qubit subspace [41]. In practice it is very difficult and even undesirable to keep the time dependence of the components of $\hat{H}$ unique and most of this thesis assumes it is not the case. However, we will see in Ch. 4 several analytical methods to try to reduce complex time-dependence to the tractable, unique time-dependence case.

2.1.4 Non-unitary evolution

In practice, the computational model of a system of finite size is only unitary if all the accessible states are only those found in the system. However in any classical or semi-classical system (i.e. any practical one), information is inevitably removed over time and transfered to states in the surrounding "environment" making the dynamics of the system at best approximately unitary. The lost information can either take the form of escaped energy (*relaxation*) or increased entropy (loss of *coherence*). To capture the latter, we distinguish from macroscopic *classical incoherence* which pertains to subjective or instrumental uncertainty the idea that objective, microscopic uncertainty about a particular subsystem may increase in an irreversible, unmeasurable way (*decoherence*). Now recall that microscopic uncertainty in the form of superposition was actually a wanted resource in the sense that it allowed a steady and continuous change from one basis state to another. Decoherence, on the other hand, is intrinsically unsteady and by virtue of this actually inhibits the resource of superposition from manifesting itself, effectively forcing the system to remain in one basis state or another. For all practical purposes, classical incoherence behaves operationally in the same way, as an unwanted fluctuation, though fundamentally

superpositions are not destroyed and information remains in the system, rather the information we have about the system becomes degraded [137]. On the other hand, information can also be removed from the system by an act of measurement, which like unitary evolution is a wanted form of fluctuation that steadily and continuously takes an arbitrary state to one of the measurement basis states. However the information that is removed is done so in an apparently non-deterministic way, and as such again forces the system into a random measurement basis state. Whether the non-determinism reduces ultimately to classical incoherence or is it solely caused by objective decoherence (up until the moment of observation) is an open interpretational question though is somewhat moot due to the common phenomenological framework.

Now wanted an usable (i.e. deterministic) uncertainty had necessitated the use of vectors to capture the exponential number of combinations of the n pieces of information as possible measurement outcomes. To model the nondeterministic uncertainty requires modeling only pairwise combinations of the deterministic uncertainties. That is, there are $\binom{2^n}{2}$ possible superpositions of two measurable basis states, each with a coherence associated with it telling how much certainty remains. This construct is called the density matrix ρ and is captured for pure states via the outer product of the wave vector ψ with its conjugate transpose

$$\rho = |\psi\rangle\langle\psi|$$

where we now use Dirac notation for the state vector. The diagonal entries of the density matrix give the effective probabilities of the respective basis states and are typically susceptible to relaxation. Uncertainty in the form of decoherence comes from considering a pure state with larger dimension and then removing environmental degrees of freedom (entangled with the computational subsystem) via a partial trace operation to give back a non-pure (a.k.a. mixed) state. To model classical incoherence, on the other hand, we use an ensemble of wave vectors $\{|\psi_i\rangle\}$ with probability distribution $\{p_i\}$, giving

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$$

In many cases, the degradation of the system is itself phenomenologically motivated in that it is derived macroscopically and is found to fit the measurable exponential decays of the diagonal and off-diagonal elements in our density matrix (corresponding to averaging many measurements). Assuming a common Hamiltonian (and an initially unentangled, memoryless environment), it can be simulated with a Markovian master equation

$$\dot{\rho} = -i[H, \rho] + \sum_{j=1}^{M} \left[\frac{1}{T_j} \mathcal{D}[A_j]\rho \right], \tag{2.4}$$

where M decay processes are taking place each with characteristic decay time T_j and associated operator A_j, and $\mathcal{D}$ is the damping super-operator, defined as $\mathcal{D}[A]\rho = A\rho A^\dagger - \frac{1}{2}A^\dagger A\rho - \frac{1}{2}\rho A^\dagger A$. If there are no decay processes (a perfectly isolated system where $M = 0$), then the evolution is equivalent to the Schrödinger unitary evolution.

2.2 Quantum software

2.2.1 Assembly

As we have seen, being able to create (approximately) unitary operations is of utmost importance for running quantum computations. Since there are an uncountably large number of unitaries even for a single qubit, a method is required to generate them with finite resources. This construction is given by the Solovay-Kitaev theorem and is efficient [91]. Naturally, these operations need to scale with system size in a way that is not prohibitive to design. However, generating such operations using high-dimensional Hamiltonians requires classical simulation and optimization, which is contrary to our operating assumption that the quantum computer will scale better than its classical counterpart. Nonetheless, it has been shown that arbitrary single qubit gates and a single entangling two qubit gate form a sufficient basis from which to generate any unitary that could be desired in 2^n dimensions [172], though the construction is not necessarily efficient. In practice, the resource allocation and design of high-dimensional gates follows from the needs of the algorithm and is often efficient for known algorithms.

We wish to have the gate set readily available for compilation of a quantum algorithm. Thus, one can consider the kinds of low dimensional gates that are required for most known algorithms and build an entire instruction set of such possible operations, which can be called upon to generate quantum circuits in the larger Hilbert (state) space [156]. The most generic model for operating quantum software is known as the circuit model and utilizes such a construction. It is built up entirely of a sequence of gates and one final measurement step at the end[2], and thus it has the most stringent requirements in terms of suppression of errors in gate design. For

[2]Incidental techniques such as error-correction and teleportation will use additional measurements during the computation.

the purposes of optimal control as discussed this thesis, it is the most relevant and stands to gain the most from gate design.

On the other hand, high-fidelity gate operations are usually a resource that is difficult to attain (given the chosen hardware) or not directly useful for the type of computations required by the algorithm. In particular, decoherence and relaxation may be prohibitive to achieving quality unitaries. In contrast to the gate model, the adiabatic model is an equivalent construction that relies on engineering quantum states rather than quantum unitary evolution. By changing the effective energy structure of the system (and its respective eigenstates) one can move from the known eigenstate of a Hamiltonian to that of a different Hamiltonian, allowing one to effectively find the previously unknown ground or equilibrated state of the system. This is known as an adiabatic algorithm [44,77]. Many problems can easily be mapped to this methodology though perhaps the most natural are simulation of physical models. As before, changing the eigenstates requires manipulation of the Hamiltonian. In practice, this manipulation will involve the suppression of a variety of unwanted errors and the design can be optimized. In particular, remaining in the ground state of the eigenbasis requires an adiabatic trajectory from initial state to final state, directly related to the spectral selectivity problem (Ch. 7.1).

Another way to avoid difficult to engineer gate operations is by moving to a measurement-based model. Here, the entanglement in the system is generated by some means at the beginning of operation (one option would be through unitary operations) and the remainder of the circuit is executed via measurement, single-qubit gates, and feed-forward of classical information. The measurement process (which typically involves single-qubit gates) can also be optimized though this falls outside the scope of this text. However, as in the other cases, the biggest challenge of this computing model involves designing and optimizing many-body operations for the purposes of entanglement. This model is not studied further in this thesis.

2.2.2 Outlook on algorithms

It is difficult to say what problem-solving applications will be most relevant to quantum information. At the present moment, there is no definitive proof of a substantial technological premium because while there are numerous positive results for efficient quantum algorithms there are no provably negative results with regard to solving most (NP) hard problems using classical techniques. Furthermore, many of the proven quantum algorithms may have classical workarounds or approximation techniques that are often sufficient to the task. What can be said however is that there will be some applications and these will be more significant as the hardware improves.

At the present time, some of the most rewarding techniques include cryptographically encoded communication by using the measurement properties quantum states [9], and magnetic imaging techniques that use resonance protocols to measure quantum signatures in chemical samples [54, 45] (see Sec. 7.1.1). Quantum protocols are also gaining importance in metrology applications and photon detection devices [19, 130].

In the medium term, much can be accomplished with small or imperfect devices. With a moderate number of qubits, by engineering larger but perhaps imperfect quantum structures, it will also be fairly straightforward to run simulations and calculations with respect to physical theories, which can easily perform the unitary evolutions required in the systems [81], as well as computations relating to field theories such as evaluating Jones polynomials [42]. These simulation type algorithms may be well suited in particular to adiabatic computation devices [44, 77]. With an increase to a couple of hundred logical qubits, it is possible to run algorithms that break present state-of-the-art encryption technology, the most famous of which is Shor's factoring algorithm [158]. It is based on a curious counting argument in number theory: it relates prime number factoring to period finding of periodic functions, something that can be found directly (and therefore efficiently) from measuring the peak of the Fourier transform of the function's input. Similar algorithms which use unitaries corresponding to Fourier transforms over different algebraic groups have been formulated, generally referred to as the hidden subgroup problem.

Of course, in the long term, many more inventive, sophisticated, and large scale applications may be found. Whether this is solely for supercomputing tasks or whether more mainstream, universal resources in quantum information can be identified remains to be seen.

2.3 Quantum hardware

There is a vast array of possibilities for implementations which is too extensive to comprehensively analyse in detail here. Instead, we will outline some of the hardware choices that exist and focus on some fundamental design choices that any technology will have to make. In particular, any engineered quantum computing technology will face some tradeoffs early on between between isolation and controllability of the system, between separation of the qubits in frequency or in space and between mediated vs. direct coupling of the qubits.

2.3.1 Candidate implementations

There are many candidate systems containing quantum degrees of freedom that can be controlled. Natural qubit systems include NMR molecules with spin degrees of freedom [8, 61, 73] and photons' polarization or mode [92]. One can also capture these sorts of natural degrees of freedom using trapping potentials, in the form of for example trapped ions [65], optical lattices [169], NV centres, quantum dots [139], and electron spin resonance, which rely on controlling electronic energy states in complex energy landscapes. Perhaps the most flexible from an engineering point of view are human-made systems such as superconducting [29] and semiconducting materials [139], for which the energy landscape can largely be controlled and for which design choices are reflected directly in the control strategies to choose from and their efficacy. Some of these implementations will be discussed in more detail as specific applications to control problems that will be analysed (Ch. 6-11). Since the most general situation is also the most flexible one, where we have control over all the energy and coupling parameters, engineered systems will often be preferentially chosen to illustrate the control/design choices that can be made both within these implementations and between implementations where such parameters are fixed.

2.3.1.1 Superconducting qubits

For these purposes, a concise introduction to superconducting qubits is in order [30, 113, 153, 184]. The quantum degree of freedom involved is the mesoscopic superconducting current of Cooper pairs [79] which can be measured moving either clockwise or counter-clockwise in an electrical circuit. Alternatively, one can measure other properties of the current, such as a relative number of Cooper pairs, a Voltage, or a phase shift imparted on other coupled quantum elements. Since low-frequency superconducting currents are dissipationless, the currents are resilient to (thermal) fluctuation and offer promise of retaining coherence [6,35]. However, while superconductivity is maintained, the measurable degrees of freedom may still suffer population decay from one to another or otherwise lose coherence between each other, an effect which is amplified by these large devices coupling very strongly to their surrounding environment. Nonetheless, the large sizes and electrical nature are a boon to prospective scaling in the number of qubits, thanks to industrial deposition and lithography techniques already being very advanced for the mass production of conventional silicon based devices.

The physical realization of quantum information processing in superconducting circuits has enjoyed remarkable progress over the last decade. While initially decoherence limited single qubits to only a few coherent oscillations [126], high

precision, general quantum control is now possible over single- and few-qubit systems. This is evident by the demonstration of high-fidelity nonclassical states of two-qubit [3, 26, 39, 164] and three-qubit [40, 127] systems, harmonic oscillators [71], and the demonstration of small quantum algorithms [39]. This success is partially due to our current understanding of sources of noise and the development of techniques and systems that are resilient to these noise sources. Examples include the optimum working point [173] and the introduction of low-dispersion qubits like the transmon [95, 154] and the capacitively shunted flux qubit [165]. On the other hand, a promising route to success are qubits that contains only a minimal number of elements, such as the phase qubit [117, 116, 164].

The fundamental electrical component used for superconducting qubits is the Josephson junction, which has a non-linear inductance and linear capacitance. The element ensures that the energy levels of the circuit are not harmonic which is required for distinguishability of the states (see Ch. 8). By varying the relative strength between the inductance and the capacitance one is able to change the dominating degrees of freedom in the Hamiltonian from the number of fluxoids (for large inductance) to the number of Cooper pairs (for large capacitance). Various Hamiltonians, design choices, and control strategies for superconducting systems will be considered in the relevant chapters of the text.

2.3.2 Controllability

In generic terms, the controllability of a physical system refers to the ability to reach any particular state in the system given any input state, much akin to the universality condition for gate sets. We can consider a Hamiltonian

$$\hat{H} = \hat{H}_0 + \sum_{k=1}^{M} \hat{H}_k(t) \tag{2.5}$$

where $\hat{H}_k$ are individual control Hamiltonians that can be tuned between on and off, and determine whether the entire Hilbert space of the computational vectors can be spanned. This can be straightforwardly be checked by seeing whether all the permutations of the H_k generates the entire Lie algebra. Most commonly, the control Hamiltonians are parametrized as $\hat{H}_k(t) = c_k(t)\hat{H}_k$ which is a simplified case known as bilinear control.

In practice, quantum systems usually not completely controllable. Because of the onset of damping and decay processes which are generally at least partially irreversible, only the equilibrated states can be certifiably achieved. Instead, minimizing

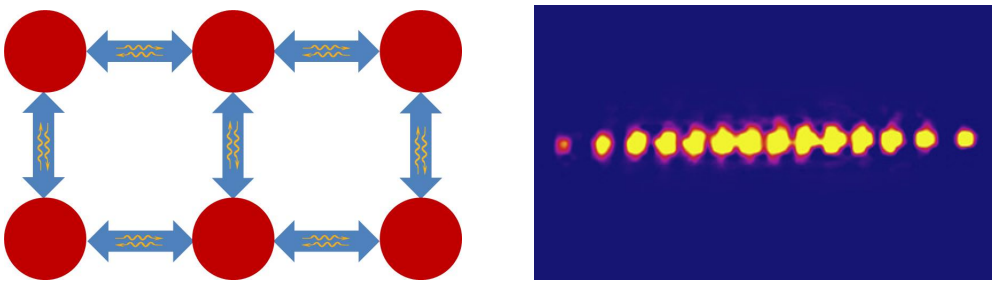

Figure 2.1: Two different coupling topologies: on the left is nearest neighbour coupling while on the right we see (trapped ion) qubits on a common (phonon) bus which can all interchange energy (taken by R. Blatt's group).

the total time for gate operations effectively minimizes the effect of structureless, Markovian decay processes. Here, we meet the first tradeoff between minimizing the time that operations take and maximizing the efficiency of these operations. This is first of all a tradeoff between different physical implementations, as some systems can be very well isolated from the environment (allowing for long operations) but this isolation often only allows for weaker couplings to other elements or external control which increases the amount of time needed for operations. In some cases, the isolation is so good that no direct coupling exists at all between states or between qubits and indirect methods (which are often slower and less efficient) of coupling must be found. On the other hand, quantum elements that couple readily to (various) degrees of freedom in their vicinity (and environment) will often be more controllable in terms of design and run-time parameters. This tradeoff is approximately redressed by speaking instead of the quality factor of quantum operations, namely the number of such operations that can be executed in the characteristic time it takes for the decay process to occur.

Note that although this benchmark is valuable it can sometimes be a simplistic or misleading. Most often, if the source of coupling to the environment can be determined then parameters can be found or strategies laid out that can counteract loss of coherence. These can often involve decreased decay processes at longer gate times and optimal values of multi-dimensional design parameters. Specific examples will be given in Sec. 8.8.2 and Sec. 10.3. In such cases, the decay processes must be included in the error analysis.

2.3.3 Coupling topologies

Choosing an implementation and computational model for the system greatly reduces the freedom in architectural design. Nonetheless, one can try to minimize interactions between elements by having tunable coupling, and moreover, one can decide which elements can even be coupled at all. In general, there is a tradeoff between two particular paradigms, that of trying to couple all pairs of qubits irrespective of physical separation or that of only coupling neighbouring qubits (see Fig. 2.1) Arbitrary couplings typically involve long range interactions (e.g. Rydberg blockade [169] or interactions inside a bus [15, 67]) and are in principle more efficient than nearest neighbour. They typically suffer from imperfect on-off times (e.g. from frequency modulation, see Sec. (6.7)) and/or imperfect decoupling (e.g. from spatial selectivity constraints). Nearest-neighbour type couplings can also be used but can suffer from a polynomial slowdown in computational speed resulting from having to move information around to couple separated logical qubits. However, they may be less susceptible to errors as frequency and spatial selectivity is often not required. Moreover, near-neighbour interaction involve fewer couplings which tends to increase gate efficiency, and also are a natural fit for certain topological qubit schemes (e.g. surface codes [52]). In these, many physical qubits correspond to a computational qubit and benefit from topological protection against information loss due to conservation laws governing certain quantities (such as the number of vortices inside a lattice).

2.3.4 Coupling mechanisms

In order to perform more complicated operations, and in particular to meet the minimum requirements for gate universality as outlined above, the different quantum components in our system have to be able to interact so as to exchange information. There are two architectural choices, either direct coupling of qubits (by physical proximity) or coupling via an intermediary component which can be used to carry information such as a bus (phonon, photon, etc) or auxiliary qubit. In general, two frequency separated qubits will only exchange energy if they are brought on resonance, e.g. for the Hamiltonian

$$\hat{H} = \alpha \hat{X}^1 \hat{X}^2 + \beta \hat{Y}^1 \hat{Y}^2 + \omega_1 \hat{\Pi}^1 + \omega_2 \hat{\Pi}^2 \tag{2.6}$$

with $\omega_1 = \omega_2$, the superscripts indexing the qubit, and the projector $\hat{\Pi}^i = |1\rangle^i\langle 1|^i$. This flip-flop interaction produces an evolution generating a gate in the class of ISWAP operations [155,189]. This is nice because qubits can be brought on resonance this way (allowing beyond-nearest-neighbour interactions) but again suffers from

frequency modulation issues (Sec. 6.7) and the fact that unwanted resonances may be activated when a qubit frequency is changed. On the other hand, one can entangle quantum elements equally well by number-conserving interactions where the presence of one qubit is felt by changing the energy landscape of the second and visa versa, for example with the Hamiltonian

$$H = \gamma \hat{Z}^1 \hat{Z}^2 + \omega_1 \hat{\Pi}^1 + \omega_2 \hat{\Pi}^2$$

known as a ZZ-interaction [189]. Here the interaction occurs irrespective of ω_1 and ω_2, avoiding having to change qubit frequencies (which is hard to do with natural qubits), but an additional mechanism is required to turn the gate on or off. The gate generated is a C-Z which again is equivalent via single qubit operations to ISWAP, CNOT, and other 2-qubit gates.

Natural interactions can arise in nature for two element that are coupled via some force that bonds them together, as occurs between atoms in a molecule or elements in a circuit. They usually give both flip-flop and ZZ interactions but with natural qubits it is difficult to modify the effective qubit frequencies and so ZZ is often more useful. Natural couplings are by definition always on and so a major difficulty is effectively turning them off for single qubit or identity operations. The technique for doing this is known as dynamical decoupling and involves repeatedly "undoing" the interaction (most famously the CPMG pulse sequence [76, 151]). Examples of new results involving decoupling sequences will be given in Sec. 6.7 and Sec. 7.2.

On the other hand, external driving fields can be used to couple devices that normally would not exchange energy. For transitions that exist between the elements but are off-resonance, the field can bring in energy that takes the system from one energy level to another, effectively enabling energy conservation and allowing energy-exchange to take place. Since the transition is effectively higher-order (virtual), this can greatly slow down the gate time but this need not be the case. The energy brought in can also be used to activate a single-qubit transition (to a non-computational state) which may then couple in an interesting way to other qubits, such as for Rydberg states. The main disadvantage here is that higher excitation states are typically prone to more decoherence and relaxation.

Finally, mediated coupling corresponds to using an additional quantum element as a bridge between qubits that (to at least first order) removes the coupling between the elements. When the intermediary device is turned on, activated via a field, or otherwise brought into resonance, it can be used as a bus for information or otherwise to directly (virtually) couple the connected qubits. Examples include using a cavity (Sec. 10.3), another qubit, collective phonon modes, and SQUIDS (Sec. 7.2). Once again, adding in additional elements can have adverse effects in terms of complicating

things by adding in unwanted terms, couplings to the environment, or just generally slowing down operations.

In general, one has an additional operational freedom about the computational basis to work in. Effectively, for the various coupling strategies, this means that one can choose between working in a *bare basis* corresponding to the "physical picture" that one would have of different elements interacting with one another, or in the *dressed basis* corresponding to diagonalizing all interactions and considering the whole system as a single device with time-independent characteristic eigen-energies (only by turning on fields can off-diagonal transitions be activated). The bare frame implies operationally that natural couplings always persist and must be removed via dynamical decoupling, while in the dressed frame off-resonance effects permeate and can either be canceled/ignored or brought into resonance to activate a transition [57]. Since non-computational transitions are typically also diagonalized in the dressed frame, choosing one frame or the other also can have effects in terms of changing the coherence time for components with heterogeneous decay mechanisms.

2.4 Quantum errors

2.4.1 Worst-case error

We have seen that many imperfections can plague quantum operations. An error occurs when a gate generates an undesired output given some input state. Consider an algorithm which consists of M sequential operations. In order for the algorithm to succeed with non-negligible probability we need the error for each operation to be less than $1/M$ where M is a large number. This ensures that the algorithm will succeed some of the time, which we presumably can check by seeing whether the solution output solves the given problem. We can upper bound the worst case by calculating how often the most pessimistically chosen initial state (with perfect preparation) will evolve to the wrong final state (given perfect measurement) given our implemented evolution U_{get}.

$$E_{max} = \sup_{\psi}(1 - \langle\psi|U^{\dagger}U_{get}|\psi\rangle), \tag{2.7}$$

In essence, this figure of merit tells us how faithful our implemented unitary operation is to the one needed in the given quantum algorithm. This error corresponds to a theoretical prediction of what might happen in an experiment. Of course in an actual experiment, there is no way to know that that we have actually hit the bottom

in terms of the lowest error that we get. Theoretically though, given a particular implementation, it is straightforward if not computationally expensive to obtain an estimate for this error. The objective of optimal control is to lower this error as much as possible.

2.4.2 Threshold theorem

Once this error is low enough, the error can be further improved by the techniques of error correction [22, 163]. At the cost of adding in additional auxiliary quantum bits, the ingenuity of this technique comes from the fact that errors can be corrected while avoiding finding out the state, using instead information redundancy. Ultimately, several rounds of error correction can be used to further minimize errors (at the cost of even more quantum resources) but this is only possible if the error is low enough. The quantum threshold theorem states that for a given correction mechanism there exists an error rate below which the error can be made arbitrarily small with the use of more and more quantum bits [93]. Thus it is always possible to get an error much smaller than $1/M$ provided we are willing to make the system large enough.

In practice what this means is that we only need to obtain single gate errors below the error-correction threshold rather than $1/M$. This tacitly assumes that increasing the system size does not further increase the individual gate errors. This may not always be the case but is often justified by most coupling mechanisms decreasing exponentially with distance. Whether no error-correction, some error-correction, or full error-correction is used is mostly dependent on which algorithm we want to run and how many steps it takes, which in turn for now is highly dependent on the error-rates themselves which we can implement and how many quantum bits can fit on our set-up. Thus, any progress that can be made is highly in need of optimizing the lowest error rates possible to benchmark proof of principles in algorithms, implementations and topological design.

2.4.3 Average error

As mentioned, in practice, it is usually very difficult to calculate worst-case error rates, especially for experiments. In simulations, it is possible but scales inefficiently for most systems. In fact, in virtually all cases calculating the average error suffices. This is defined as

$$E_G = 1 - \int d\psi \ \mathrm{Tr}\left[U_{\mathrm{ideal}}|\psi\rangle\langle\psi|U_{\mathrm{ideal}}^{\dagger}\mathcal{E}(|\psi\rangle\langle\psi|)\right], \tag{2.8}$$

where $\mathcal{E}(\rho)$ is the actual process in the full Hilbert space. While certainly a very large worst-case error would be detrimental if every step of an algorithm were to hit this error, this seems unlikely. The reason is that by definition of the average error, the worst-case error is either close to the average error or else very unlikely to happen. One could assume that for a particular input and choice of algorithm, the worst-case error would occur for every single gate. Even in this overly pessimistic scenario, one would still be able to run different inputs successfully, and one would still be able to check whether the program was successful, thereby mitigating the risk of failure. Nonetheless, one can try to minimize norms other than the 2-norm, such as the 3-norm or 4-norm which weigh the worst case errors more strongly [103].

In fact, one can bound the worst-case error from the average error. For low dimension the two are within an order of magnitude of each other, and a very low average error ensures the worst-case is also very low. In fact, virtually every known algorithm uses gates acting on few qubits, as there is no consistent way of generating high-dimensional gates. Most importantly, error correction typically only uses single and double qubit gates and can be used to diminish the errors further (assuming distant qubits do not interact).

For single qubits, average errors can calculated via Ref. [17]

$$E_G = 1 - \frac{1}{6} \sum_{j=\pm x,\pm y,\pm z} \mathrm{Tr}\left[U_{\mathrm{ideal}} \rho_j U_{\mathrm{ideal}}^{\dagger} \mathcal{E}(\rho_j)\right], \tag{2.9}$$

where x, y, z represent the 6 axial states on the Bloch sphere. Clearly the worst-case for qubit error is less than 6 times the average error.

Chapter 3

Elements of Quantum Control

Having looked at quantum information from an operational level, we can look at what is needed to physically control a quantum system. We have already stated that errors per gate below a certain threshold are needed for applications to produce meaningful results. The focus here is therefore to look at quantum information from a control point of view, with an eye towards common settings and their effect on the unitary evolution of a system. Most importantly, we want to distinguish important classes of operators that can all be modeled in a similar way. Through careful analytical and numerical examination, it is anticipated that efficient control strategies can be devised to maximize faithfulness to the desired operations, and that these solutions will scale and combine with each other with larger Hilbert space dimension.

Table 3.1 lists the main physical systems that will be studied in this thesis. The columns identify usual, generic quantum control problems that can be associated with these and other quantum systems. Although, in fact, most of the problems are in fact quite ubiquitous, the models considered in this thesis will capture (only) a few problems at a time, in the interest of both focused and clear solutions.

As shown in table 3.1, it is instructive to separate the sources of imperfection into three broad classes: time independent, deterministically time-dependent, and non-deterministic. As can be seen, these are respectively further subdivided into: on-resonant or off-resonant quantum operators; selective and/or smoothing (semi-classical) control functions; and either static or fast (irreversible) non-unitary evolution. Since the static, time-varying, and non-deterministic aspects of the control are mostly independent these can be combined rather arbitrarily. For example, irreversible errors are the major cause of error in most systems; however, since many systems cannot rectify these errors we choose to only include them when they are relevant to the matter of control. In traditional quantum control (in NMR), where

Ch..	Model \ Error	Dim	Time-indep		Time-dependent		Non-determ.	
			Res.	Off-res.	Select.	Smooth.	Static	Fast
6.5	Waveform shaping	3			X	X		
6.6	Tuning error	4	X	X	X	X		
7.1.3	Frequency selection	2^n		X	X			
7.2.3	Freq.+Space select.	8		X	X	X		
7.2.4	Spatial selection	8	X			X		
8	Leakage	3		X	X		X	X
9.1	Fock state creation	N>3		X	X			
9.2	Multiple leakages	4+		X	X			
9.3	2 qutrit coupling	9		X				
10.1	Raman transition	3			X			
10.2	Sideband transition	4		X	X			
10.3	Bare J-C ladder	9		X	X	X		X
11	Strong coupling	3		X	X	X	X	

Table 3.1: Different physical systems and associated control problems considered. For the different sections which analyse some generic Hamiltonian models, the control problems associated are listed in the columns: time-independent errors (resonant and off-resonant), time-dependent errors (either for selectivity purposes or as a result of smoothing), and non-deterministic errors (either slow or fast).

pulses can be much faster than characteristic evolution times of the Hamiltonian, problems such as unwanted time-independent on-resonant terms and static noise have been redressed emphatically using sequences of short pulses (*composite gates).* Here, we assume that strong coupling between elements can be engineered as desired and therefore that time scales are all of about the same order. Thus, semi-classical time-dependent control fields will play a central role, as will frequency separation of quantum elements.

In the subsections that follow we give basic introductions to the main control errors (those from the table) that will be addressed in the physical models in chapters 6-11. Chapters 4 and 5 will summarize mostly standard analytical and numerical techniques for addressing these problems.

3.1 Time-dependent fields

We have seen with Eq. 2.2 that the fields that control the quantum evolution to generate rotation operators can be time-dependent during the gate, and moreover that the condition $[H(t), H(t')] = 0$, if satisfied, allows a simple method to calculate the unitary evolution. However, one might ask why introduce smooth time-dependence to $\Omega(t)$ in the first place. There are two reasons: either the smoothness is due to experimental imperfection or it is put in place to suppress another kind of error.

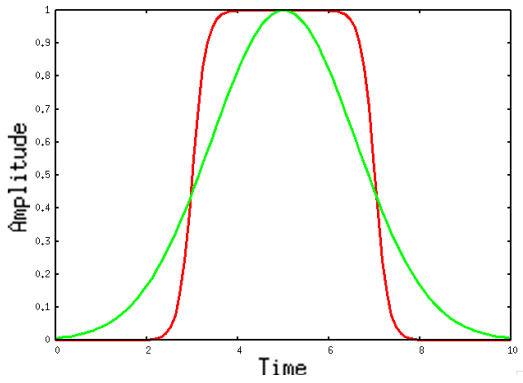

Figure 3.1: Common pulse shapes used in quantum control: a smoothed square (tangential) pulse $\Omega_T(t)$ in red and a Gaussian pulse $\Omega_G(t)$ in green.

In particular, there is a the finite turn-on time to most electronics. That is, it is not strictly possible to specify a discontinuous jump of the field $\Omega(t)$ from one value to another, and instead it typically follows an exponential path, e.g.

$$\Omega_T(t) = \frac{\Omega_{\max}}{2}\left[\tanh\left(\frac{t}{\sigma}\right) - \tanh\left(\frac{T-t}{\sigma}\right)\right], \tag{3.1}$$

where σ is the rise-time and $t \in [0, T]$. Such smoothed square (tangential) pulse (3.2) is plotted in Fig. 3.1 as the red line. More generally, for waveform response functions f(t) that are memoryless (such as a filter), a control signal at input will be replaced by a convolution at the qubit,

$$c(t) = \int_{-\infty}^{\infty} f(t-t')\Omega(t')dt', \tag{3.2}$$

where causality is taken into account by a suitable delay time (with $\Omega(t') = 0, t' > T$). Optimization using response functions will further be discussed in Ch. 6. Even

without the presence of filtering, choosing an $\Omega(t)$ that isn't a constant function can be advantageous when the Hamiltonian is not a simple qubit. Many different examples will be given for control functions, but, due to its small bandwidth, by far the most common is the Gaussian [7, 166]

$$\Omega_G(t) = A \exp\left[-\frac{(t-T/2)^2}{2\sigma^2}\right] \tag{3.3}$$

where σ is the standard deviation and A is chosen such that the correct amount of rotation is implemented (e.g. $A = \pi/\sqrt{2\pi\sigma^2}\mathrm{erf}[T/\sqrt{8}\sigma]$ for a NOT, see Eq. 2.3). A Gaussian pulse is plotted in Fig. 3.1 as the green line.

3.2 Selection error

The most common reason to want to smooth the control field is to improve the frequency selectivity of the pulse. As we have seen fields are often used to activate transitions for single and two-qubit gates. This is usually accomplished by matching the frequency with that of the transition that is driven. Crosstalk typically arises when the spatial (and frequency) extent of the field is larger than the separation between qubits. For a field that is on resonance with a particular transition but with crosstalk to a second transition, a typical Hamiltonian will be of the form

$$H = \Omega(t)X_1 + \lambda\Omega(t)X_2 + \Delta\omega\Pi_2$$

with $\Delta\omega = \omega_1 - \omega_2$ the energy separation between the two qubits/transitions. λ accounts for the difference in the Rabi frequency from one qubit to the next; in particular if the spatial extent of the field is centered around the first qubit then the coupling will be smaller with the second. This problem will be discussed in detail in Ch. 7. Provided we are in the selective regime, namely $\lambda\Omega \ll \Delta\omega$, the selectivity error can be indefinitely reduced provided we pick a large enough gate time (resulting in a smaller Rabi frequency). On the other hand, increasing the time introduces further decoherence and relaxation, and so there is a limit to how many transitions can be put in a given bandwidth. Spectral crowding occurs when the transitions are too close to be individually selected.

The selectivity can be improved by choosing the appropriate waveform for the Rabi frequency, where in general smoother functions will have a smaller bandwidth, and the most common is the Gaussian, Eq. 3.3. The most significant tool for estimating the error is the Fourier transform which is a classical scalar function. This

scalar can be generalized to the quantum case where it given instead by a nonlinear function of the envelope amplitude, as will be shown in Sec. 7.1.

Some common strategies for further reducing the spectral bandwidth beyond the Gaussian envelope are spin pinging [55], Rapid Adiabatic Passage (RAP [10]), half-gaussians [55], and composite pulse Narrow Band schemes (NB1 and NB2 [181]). In this thesis another method is developed: Derivative Removal by Adiabatic Gate (DRAG).

It is noteworthy that selectivity is not limited to addressing qubits but other quantum transitions as well for any given purpose. Examples that will be discussed that include some selectivity criteria are crosstalk for two-qubit gates (in a 3 qubit model, Sec. 7.2), leakage transitions outside the qubit subspace in Ch. 8 and Ch. 9, higher-order virtual transitions (Ch. 10), and counter-rotating terms (Ch. 11).

3.3 Off-resonant operator error

The effect of unwanted transitions is more subtle than unwanted lines in the spectral profile of a pulse. The Hilbert space structure of quantum mechanics implies that adding in additional operators that do not commute with the Hamiltonian will modify the dynamics of the evolution in subtle ways. In particular, the matrices representing the Hamiltonian will have modified eigenvalues (due to the extra terms). This in turn will result in an evolution where an unwanted phase difference will build up between the computational states (as well as other errors due to non-commutativity). These phase (and other) errors do not show up in the semi-classical shaping arguments that make certain pulse sequences more selective than others. However, the extra phase is still correctable, usually by changing the resonance frequency of the needed transitions and/or by using extra gates. In the case of decoupled qubits or virtual transitions (discussed in Ch. 10), these errors will act separately or even disappear altogether and can therefore easily be dealt with in isolation (e.g. with frame transformations, Sec. 4.2). The most significant effect is if and when the operator error cannot be separated from the selection error (they do not commute) whereby techniques such as time-dependent diagonalization transformations (Sec. 4.3) and operator expansions (Sec. 4.4) are required to find solutions.

This is most significant in the case where an unwanted "leakage" transition is connected by a shared energy level to the driven transition. In fact, the concept of a perfectly isolated two-level system is an idealization [182] and so this error is present to some extent in all quantum systems. An extra transition gives is a 3-level Hamiltonian

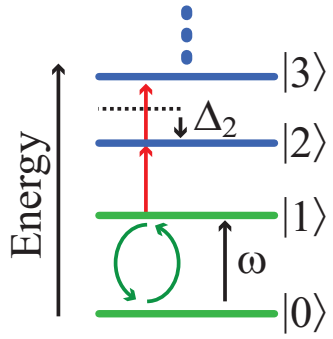

Figure 3.2: Energy level diagram of an anharmonic oscillator. The qubit is formed by the $|0\rangle$ and $|1\rangle$ (green) levels and we aim to have complete control in this subspace when leakage to the $|2\rangle$ and then $|3\rangle$ etc. is possible (red arrows). The dotted black lines indicate the positions the energy levels would be at if the system was a harmonic oscillator of frequency ω.

$$H = a(t)(X^{0,1} + \lambda X^{1,2}) + b(t)(Y^{0,1} + \lambda Y^{1,2}) + c(t)(\Pi^1 + 2\Pi^2) + \Delta_2\Pi^2 \tag{3.4}$$

where $X^{j,,k} = |j\rangle\langle k| + |k\rangle\langle j|, Y^{j,k} = i|j\rangle\langle k| - i|k\rangle\langle j|$ are generalizations of the Pauli matrices in the j-k subspace and Π^l is the projector onto the l-th energy level. A particularly significant case of leakage is found in the anharmonic oscillator (studied in Ch. 8and Ch. 9), where each energy separation is different from the next by an amount Δ, the *anharmonicity,* which is typically quite small compared to other leakage mechanisms.

$$H_{qut} = a(t)\sum \lambda_l X^{l-1,l} + b(t)\sum \lambda Y^{l-1,l} + c(t)\sum_l l\Pi^l + \frac{\Delta}{2}\sum_l (l^2 - l)\Pi^l, \tag{3.5}$$

Fig. 3.2 shows the first few levels of an anharmonic ladder, where the qutrit approximation (Eq. 3.4) is most often used. More complex leakage mechanisms involving multiple leakages will be dealt with in Ch. 9.

Note that there are also examples of situations where there is essentially no selection error but the off-resonant operator error still remains. This can occur for example when the Fourier transform of the rectangular (unshaped) pulse happens to go to zero but leakage transitions (and their effect on the Hilbert space) are still present. An example is given in Sec. 9.3.1.

3.4 Resonant operator error

We have seen that certain kinds of unwanted terms are always present. These terms are usually not driven, and they are generally significant in that they are either diagonal or off-diagonal and on-resonance so that there is no natural suppression mechanism as was the case with selectivity error, where terms average out to zero for longer gate times. This can also appear as another form of crosstalk for unwanted driven couplings (Sec. 7.2). In all these cases, the unwanted terms cannot be compensated for using small parameter expansions as the operators can strongly influence the state evolution. Instead the effect of these terms must somehow be undone, traditionally with temporal sequences of composite pulses that create trajectories in the Hilbert space that undo the unwanted evolution. The preferred mathematical technique for this purpose is average Hamiltonian theory (Sec. 4.4). For identity operations, the canonical strategy involves reversing the unwanted evolution midway through the clock cycle by flipping the qubit , known as a Hahn echo and more generally as *dynamical decoupling* [76, 151]. When a more complicated operation is desired, more complicated decoupling sequences can still be applied to remove a variety of error terms. For multiple couplings that need to be decoupled, multiple spin flips are required to antisymmetrize every pair of qubits. Near resonant transitions that do not satisfy the selectivity criterion $\lambda\Omega \ll \Delta\omega$ can also be dealt with using decoupling type pulses (Sec. (6.7)). Finally, quantum noise can often be dealt with using decoupling sequences which often do not depend on the strength of the unwanted coupling, as will be seen next.

3.5 Static non-deterministic error

Instead of having constant or controllable terms, one can think of having an unknown term in the Hamiltonian. The uncertainty can be constant, as in the case of collective degrees of freedom such as NMR molecules or atoms in an optical lattice. On the other hand, if the uncertainty is the result of noise, than we can distinguish between slow noise, which can be treated again as approximately constant (e.g. fabrication uncertainty), or fast noise which is essentially Markovian and irreversible (see the next section). An interesting case occurs when the noise is roughly as fast as quantum gates, known as telegraph noise. This particular case is addressed using optimal control methods in Ref. [121]. When the noise is very slow or constant, the dynamical decoupling sequences from the previous section often work as they are not dependent on the magnitude of the error term. On the other hand, if the noise is not constant it may be advantageous to perform multiple time reversals to undo the evolution of the

unwanted terms on a fast time scale over which they are approximately constant. The most renowned example of this is the CPMG sequence [76,151]. Once again, for noise that interrupts a rotation, there are other composite sequences to correct this, such as CORPSE [33, 168], for noise in the qubit frequency, and Broad Band composite pulses (BB1 [181]) for correcting noise in the Rabi frequency. In a sense, these can be thought of as the antithetic case of selectivity where narrowband behaviour is sought after. Of course the uncertainty can more generally occur for any parameter in the Hamiltonian, and we will be considering two cases in particular using optimal control: Sec. 8.8.2 deals with structured energy and operator dispersion in an anharmonic oscillator (where coherence and anharmonicity form a tradeoff), and Sec. 11 deals with uncertain phases in the driving field. For these purposes, classical uncertainty is usually dealt with by modeling the system as a set of Hamiltonians $\{H_\phi\}$ obeying a set of evolutions $\{U_\phi = \mathcal{T}\exp(-i\int H_\phi dt)\}$.

3.6 Fast non-deterministic error

Noise that is faster than the dynamics of the system is in general irreversible. One way to think of this noise is as an entanglement between the quantum system and another Hilbert space. If the extra Hilbert space is small [150, 157], then we can include the extra dynamics and try to optimize over the larger Hilbert space (as was done for leakage, or can be done with auxiliary qubits, resonators , etc.). However if the extra space is macroscopic and/or fast then it is generally not possible to model or even undo the dynamics and the best we can do is to avoid them. This can either be done by finding a part of the computational space that is less affected by the noise operators (e.g. Sec. 10.3 or Refs. [102, 131, 187]), or else to try to shorten the operations as much as possible. The latter is generally the approach taken in this thesis, where errors are often plotted against time with the implication that error decreases with gate time while Markovian effects increase and so for any given rate of decay there will be an optimum point that minimizes both the error and the noise (see Sec. 8.8.1). Nonetheless, one can still model the full dynamics of the optimized gate (as in Eq. 2.4) or include the decay in the optimization as in Sec. 8.8.2.

Chapter 4

Frame Transformations

In the limit of the Hamiltonian commuting with itself at different times and low Hilbert space dimension, there is often an analytic form for the evolution given by the matrix exponential (e.g. Eq. 2.3). However, most quantum systems contain multiple qubit couplings and time-dependences and thus are difficult to evolve in time. While numerical techniques can usually work for larger Hilbert space dimensions, even these are limited by an exponential scaling with number of qubits. Thus, finding alternative, tractable representations for Hamiltonian structures with complicated time-dependence or in high dimension has the double advantage of allowing a conceptual understanding of its dynamics and a practical utility of being efficient to calculate (generally polynomial in the number of qubits). Combining analytic error-removal strategies can also be accomplished, provided that the errors approximately commute with each other (which is always true in the limit of small errors). For this purpose, in this chapter time-dependent transformations and expansions are discussed that make such tractable representations possible. These will be used to analyse specific kinds of errors as relating to various physical models. This chapter lays out the detailed theoretical ground work which will be used as required in chapters 6 through 11.

4.1 General theory

We have seen that the physical design of our quantum system will have a large impact on its performance. There is also a mathematical freedom about what we choose as the intrinsic energy basis (that defines our computational states) and what the Hamiltonian operator is that describes their evolution. The two extremes where either the basis states are completely static in time or the wavevectors (combinations

of the basis states) are completely static are known as the Heisenberg and Schrödinger pictures, respectively. A basis that does not contain implicit time-evolution of the basis states is known as a bare basis while allowing such time-dependence gives rise to dressed basis states. However, if all of the evolution of the wavevector is implicit as occurs in the Schrödinger picture, then this will not be useful for (vector) computation. Nonetheless, there is a whole range of possible descriptions in between, and a suitable choice will have important repercussions both in terms of operability of the implementation(see e.g. Sec. 2.3.4) and simplification of modeling of complex quantum problems.

The various descriptions have in common that they must satisfy the Schrödinger equation ($\hbar = 1$),

$$\frac{d}{dt}|\psi\rangle = -iH_\psi|\psi\rangle \tag{4.1}$$

4.1.1 Transformation formula

We want to transform to some new dressed basis $|\phi\rangle = V(t)|\psi\rangle$. Then, using the Schrödinger equation again for the new picture we can derive a compound evolution

$$\begin{aligned}
-iH_\phi|\phi\rangle &= \frac{d}{dt}|\phi\rangle \\
-iH_\phi|\phi\rangle &= \left(\frac{d}{dt}V\right)|\psi\rangle + V\left(\frac{d}{dt}|\psi\rangle\right) \\
H_\phi|\phi\rangle &= (VH_\psi + i\dot{V})|\psi\rangle \\
H_\phi|\phi\rangle &= (VH_\psi V^\dagger + i\dot{V}V^\dagger)|\phi\rangle
\end{aligned}$$

Thus we obtain the general time-dependent transformation formula $H_\phi = VH_\psi V^\dagger + i\dot{V}V^\dagger$, where all terms have dependence on a common time, and which tells us that the total evolution of a system H_ϕ will be given by a combination of the explicit evolution of the wavevectors given by H_ϕ and the implicit evolution of its basis states given by V.

In practice, we can also compute the effect of the transformation on an evolution $U_\phi(0,T)$ from an initial wavevector state $|\phi(0)\rangle$ to a final state $|\phi(T)\rangle$ directly. To see this we apply the time-ordering operator of the exponential map and then transform the Hamiltonian as above:

$$
\begin{aligned}
U_\phi(0,T) &= \lim_{\Delta t\to 0}\prod_{l=\lceil T/\Delta t\rceil}^{0}\exp\left(-iH_\phi(l\Delta t)\Delta t\right)\\
&= \lim_{\Delta t\to 0}\prod_{l=\lceil T/\Delta t\rceil}^{0}\exp\left(\dot{V}(l\Delta t)V^\dagger(l\Delta t)\Delta t\right)\exp\left(-iV(l\Delta t)H_\psi(l\Delta t)V^\dagger(l\Delta t)\Delta t\right)\\
&= \lim_{\Delta t\to 0}\prod_{l=\lceil T/\Delta t\rceil}^{0}V(\Delta t)V(l\Delta t)\exp\left(-iH_\psi(l\Delta t)\Delta t\right)V^\dagger(l\Delta t)\\
&= \lim_{\Delta t\to 0}\prod_{l=\lceil T/\Delta t\rceil}^{0}V((l+1)\Delta t)U_\psi(l\Delta t,(l+1)\Delta t)V^\dagger(l\Delta t)\\
&= V\left(T\right)U_\psi(0,T)V^\dagger\left(0\right)
\end{aligned}
$$

which is the same result we expect given the direct transformation of the vector states,

$$\langle\phi(T)|U_\phi(0,T)|\phi(0)\rangle = \langle\phi(T)|V(T)U_\psi(0,T)V^\dagger(0)|\phi(0)\rangle = 1$$

4.1.2 Equivalent frames

There is an important subclass of transformations which is those that do not modify the computational states. As mentioned, these are important for finding alternative representations of physical models that make analysis more tractable. Our requirement is that the state of the quantum system be identical at the start and at the end of gate operations in both frames, but this leaves complete freedom both in the evolution and its representation of what happens in between. The constraint amounts to

$$V(0) = V(T) = \mathbb{1}.$$

Clearly this also leaves the evolution U unchanged. This class of transformations is fundamental because it provides a way to move terms around in our Hamiltonian to obtain more tractable and intuitive forms without changing the other gates. Most of the explicit transformations that follow in this thesis are specializations of this class of transformations and serve this purpose.

4.1.3 Interaction frames

In between the Schrödinger frame and the Heisenberg frame lies the interaction frame, where some part of but not all of the terms in the Hamiltonian have been transformed away. For a Hamiltonian of the form $H = H_0 + H_1$ where $[H_1(t), H_1(t')] = 0$, then the interaction picture with respect to H_1 is given by the transformation $V = \exp(i\int_0^t H_1 dt)$. Then,

$$\begin{aligned} H_{\text{eff}} &= V(H_0 + H_1)V^\dagger + i\dot{V}V^\dagger \\ H_{\text{eff}} &= VH_0V^\dagger + VH_1V^\dagger - H_1VV^\dagger \\ H_{\text{eff}} &= \exp\left(i\int_0^t H_1 dt\right) H_0 \exp\left(-i\int_0^t H_1 dt\right) \end{aligned}$$

For constant time-dependence, V corresponds to undoing the Schrödinger evolution of the perturbation terms in the Hamiltonian, $V = \exp(iH_1 t)$. If H_1 is cyclic, in the sense that it starts and ends at the same value (integrating to zero), then the interaction frame evolution U_{eff} is equivalent to that of the original frame. Otherwise, the evolution is given by $U_{\text{eff}} = V(T)U$. In practice, one can often think of H_1 as a perturbation to an otherwise known evolution and we will see in the next sections some expansions to gauge its impact on the effective evolution.

4.2 Phase transformations

If we expand on our definition of frame equivalency to indistinguishability by measurement (we loosen the boundary condition to $e^{i\phi_1}V(0) = e^{i\phi_2}V(T) = \mathbb{1}$), we include all transformations which are diagonal. These transformations corresponds to a shifting of the clock reference of the processor, and thus can implement gate operations that shift the phase of a wavevector, essentially 'for free' [43].

4.2.1 Phase tracking

The simplest scenario occurs when we simply want to perform a phase rotation U_θ^Z on a qubit. This can be as part of an algorithm or in order to undo an error that has occurred of the opposite magnitude, $U_{-\theta}^Z$ (we will see such errors occur in later chapters, in particular for unwanted off-resonant operators, Sec. 3.3). The strategy is essentially to commute the phase gate operation with anything to its left so that in the end only an unmeasurable phase is left over when the algorithm is complete. For

example, if we have the want to perform a U_θ^Z followed (to the left) by the operation $U = e^{-iHt}$, then this is equivalent to demanding $UU_\theta^Z = U_\theta^Z U_{\text{eff}}$ from which we see we have to pick $U_{\text{eff}} = U_\theta^Z U U_{-\theta}^Z$ and giving the transformed Hamiltonian

$$\begin{aligned} H_{\text{eff}} &= U_\theta^Z H U_{-\theta}^Z = U_\theta^Z (a(t)X + b(t)Y) U_{-\theta}^Z \\ &= (\cos(\theta)a(t) - \sin(\theta)b(t))X + (\cos(\theta)b(t) + \sin(\theta)a(t))Y \end{aligned}$$

which is nice because no Z operator need be physically implemented on the qubit (as this may be difficult or cumbersome). Of course, the phase accumulates to the left as more operations are chained and therefore must be tracked over the length of the algorithm. In particular, if each operation U produces an error such that it gives in fact UU_θ^Z then M such operations in a row can be corrected via M basis transformations, where the M-th basis transformation for the M-th call to U is given by

$$\begin{aligned} H_{\text{eff}} &= (U_\theta^Z)^M H (U_{-\theta}^Z)^M = U_{M\theta}^Z (a(t)X + b(t)Y) U_{-M\theta}^Z & (4.2) \\ &= (\cos(M\theta)a(t) - \sin(M\theta)b(t))X + (\cos(M\theta)b(t) + \sin(M\theta)a(t))Y & (4.3) \end{aligned}$$

4.2.2 Rotating frames

The analog equivalent to the sequential accumulation of phase $U_{M\theta}^Z$ is a continuous accumulation in the form of $V = \exp(i \sum_h \omega_h |h\rangle\langle h|t)$. Once again, this does not modify the probabilities of measuring a given state $|\psi\rangle$ but only adds in a time-dependent phase $|\phi\rangle = e^{i\omega_\psi t}|\psi\rangle$. For unwanted phase operators that are continually present, this continuous transformation is more suitable and corresponds to an interaction frame with respect to the unwanted diagonal operator.

As the most common example, suppose one wants to perform a X or Y rotation on a qubit where the energy splitting is much larger than the external control field. It is standard to move to an interaction frame with respect to the energy splitting and drive the rotations with sinusoidal fields[1]. We start off with the untransformed lab Hamiltonian

$$H_{\text{lab}} = (2a(t)\cos(\omega_d t + \phi) + 2b(t)\sin(\omega_d t + \phi))X_\pi + (c(t) + \omega_q)\Pi^1 \quad (4.4)$$

[1]This is most easily accomplished in experiment by mixing the shaped waveform signal with a carrier pulse at the frequency of the desired transition

which differs from the conventional form Eq. 2.2. This form is not intuitive and lacks a simple method of exponentiation (due to lack of commutativity). However, using the transformation $V = \exp(i\omega\Pi^1 t + i\phi)$ one can move to the effective Hamiltonian

$$\begin{aligned} H_{\text{eff}} &= e^{i\omega\Pi^1 t}\{(2a(t)\cos(\omega_d t+\phi) + 2b(t)\sin(\omega_d t+\phi))X_\pi + c(t)\Pi^1\}e^{-i\omega\Pi^1 t} \\ H_{\text{eff}} &= (c(t)+\Delta\omega)\Pi^1 \qquad (4.5) \\ &\quad +(a(t)\cos(\delta\omega t) + b(t)\sin(\delta\omega t))X \qquad (4.6) \\ &\quad +(b(t)\cos(\delta\omega t) - a(t)\sin(\delta\omega t))Y \\ &\quad +(a(t)\cos(\omega_\Sigma t+2\phi) + b(t)\sin(\omega_\Sigma t+2\phi))X \qquad (4.7) \\ &\quad +(-b(t)\cos(\omega_\Sigma t+2\phi) + a(t)\sin(\omega_\Sigma t+2\phi))Y \end{aligned}$$

where $\delta\omega = \omega_q - \omega$, $\Delta\omega = \omega_d - \omega$, $\omega_\Sigma = \omega_d + \omega$. The frame in which $\Delta\omega = 0$ is known as the interaction frame (with respect to the energy eigenbasis). It is customary to ignore the fast-oscillating last two lines (the counter-rotating terms) which will self-average to zero under the justification that $\omega_\Sigma \gg a, b, c, \delta\omega$, known as the rotating wave approximation (RWA). We will see this approximation is another example of selectivity/off-resonant operator criteria. On the other hand, picking $\delta\omega = 0$ will be referred to as the rotating frame and doesn't contain any fast oscillations within the frame (beyond any 'slow' time dependence in the envelopes $a(t), b(t), c(t)$) in the RWA limit. At resonance, $\delta\omega = 0$ and $\Delta\omega = 0$, the RWA Hamiltonian is now the same as the simple qubit, Eq. 2.2.

If we have n independently driven qubits then the transformation generalizes to $V = V_1 \otimes V_2 \ldots \otimes V_n$ with $V_j = \exp(i\omega_d^j \Pi^j t + i\phi)$, which is known as a multiply-rotating frame. Once again we can go to an interaction or rotating frame for all the levels. Note that if there are more non-zero matrix elements than there are energy levels or if any transition is driven by more than one drive then it is not possible to find a rotating frame that cancels all fast oscillations in the Hamiltonian. This notation will be formalized further in Ch. 7 and strategies for quantifying and suppressing the counter-rotating terms will be given in Ch.11.

4.2.3 Phase ramping

More generally, one will want to perform arbitrary phase operations $c(t)Z$ on our qubits via transformation methods. Going again to the interaction with respect to the variable, we see we can simulate it with

$$\begin{aligned} H_{\text{eff}} &= \exp\left(i\int_0^t c(t')Zdt'\right)(a(t)X + b(t)Y)\exp\left(-i\int_0^t c(t')Zdt'\right) \\ &= \left(\cos\left(\int_0^t c(t')dt'\right)a(t) - \sin\left(\int_0^t c(t')dt'\right)b(t)\right)X \\ &\quad + \left(\cos\left(\int_0^t c(t')dt'\right)b(t) + \sin\left(\int_0^t c(t')dt'\right)a(t)\right)Y \end{aligned}$$

Alternatively, one can time-dependently change the frequency in the lab Hamiltonian Eq. 4.4 by the same amount $\omega_d^{\text{eff}} = \omega_d + \int c(t)dt$ (under the RWA).

Once again, this similarity transformation will give the same evolution as the untransformed Hamiltonian H when $c(t)$ is cyclic. If $c(t)$ is not cyclic then we must track the phase as in Eq. 4.2. Of course, if this operation is the most common one then it may be easier to consider the transformed frame to be the computational frame so that no phase is tracked but which requires transforming all other operations as well (e.g. the identity operation now introduces a constant phase error).

4.3 Diagonalization

Perhaps the most important transformation that one can perform is the diagonalization of the Hamiltonian. We have seen that diagonalizing the terms in the Hamiltonian can be used to define a new computational dressed basis different from the bare undiagonalized one which may have operational advantages. Also, diagonalization is the primary method for exponentiation of the Hamiltonian given its Hermiticity (Appendix A). We also can use the time-dependent version of the transformation formula to diagonalize the Hamiltonian, $H_{\text{eff}} = DH_\psi D^\dagger + i\dot{D}D^\dagger$, where the effect of off-diagonal terms may be clearer since the evolution is trivially a diagonal operator. If D starts and ends with the identity the transformation is an equivalent frame to the computational basis.

4.3.1 Adiabatic expansion

In order to facilitate the diagonalization, we often want to expand around some small term, which in this case we call the adiabaticity parameter. Usually this is the ratio of two energies in the system, which in foresight (to Sec. 7.1) we denote $\alpha = \frac{\Omega}{\Delta}$. The transformation operator has the role of (block) diagonalizing the Hamiltonian and

takes the form $D = \exp(iS)$. As before, the effective Hamiltonian corresponds to a Taylor expansion around the null transformation,

$$\begin{aligned} H_{\text{diag}} &= \exp(iS)H\exp(-iS) + \dot{D}D^\dagger \\ &= \sum_r (iS)^r H \sum_s (-iS)^s + \dot{S} \\ &= H + \dot{S} + i[S,H] - \frac{1}{2}[S,[S,H]] + O(\alpha^3) \end{aligned}$$

4.3.2 Interaction frame

In addition to being a time-dependent diagonalization, the transformation can also be an interaction frame with respect to some parameter. Dividing the Hamiltonian into $H = H_{\text{big}} + H_{\text{small}}$, we can take the interaction frame with respect to H_{small} which we assume diagonalizes the rest of the Hamiltonian H_{big}. giving after application of $V = \exp(i\int H_{\text{small}}\,dt)$.

$$\begin{aligned} H_{\text{diag}} &= \exp\left(i\int H_{\text{small}}dt\right) H_{\text{big}} \exp\left(-i\int H_{\text{small}}dt\right) \\ &= H_{\text{big}} + i\left[\int H_{\text{small}}dt, H_{\text{big}}\right] - \frac{1}{2}\left[\int H_{\text{small}}dt, \left[\int H_{\text{small}}dt, H_{\text{big}}\right]\right] + \ldots \end{aligned}$$

where in the second line we have assumed the integrand is small and can be expanded around (as in the adiabatic expansion) though this may be superfluous when the diagonalization can be done analytically (e.g. Sec. 7.1.3.4, Sec. 10.1). This expansion is most useful if one can find a small (cyclic) perturbation to add to the imperfect Hamiltonian that diagonalizes unwanted parts. The net effect is then simply a shifting of the eigen-energies, which can readily be corrected for using the techniques of Sec. 4.2. Applications of using such an auxiliary control for the purposes of diagonalization will be seen: to deselect transitions (Sec. 7.1), to offset the effect of leakage (Ch. 8), to suppress the counter-rotating terms in the qubit lab Hamiltonian (Ch. 11), or to enhance virtual transitions (Ch. 10).

4.4 Magnus expansion

In the limit of a small Hamiltonian, fast oscillation or a very small evolution time, the Magnus expansion [64, 110] is one of the most helpful tools in quantum control theory. The expansion of a Hamiltonian is given in orders where the first three are

$$\mathcal{H}^{(1)} = -i\int_0^T H(t)dt$$

$$\mathcal{H}^{(2)} = \frac{1}{2T}\int_0^T dt_1 \int_0^{t_1} dt_2 \Big[-iH(t_1), -iH(t_2)\Big]$$

$$\mathcal{H}^{(3)} = \frac{1}{2T^2}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3 \left([-iH(t_1),[-iH(t_2),-iH(t_3)]] + [-iH(t_3),[-iH(t_2),-iH(t_1)]]\right) \tag{4.8}$$

and the composite evolution is given by

$$U = e^{\left(\mathcal{H}^{(1)} + \mathcal{H}^{(2)} + \mathcal{H}^{(3)} + \mathcal{H}^{(4)} + \cdots\right)}$$

The convergence of the series is difficult to prove but a sufficient condition is that $\left|\mathrm{Tr}\int_0^T H(t)dt\right| < \pi$ [23, 89]. For small enough T this can easily be satisfied. For H that obeys an area theorem or whose evolution can be broken into smaller segments all performing the same operation, the expansion can instead be shown to converge for one such subsegment, which in turn is sufficient to show that the Hamiltonian is applying the correct operation on concatenation.

4.4.1 Interaction frame

On the other hand, one often uses this expansion to gain insight into a Hamiltonian rather than to provide its final evolution. In this case, one tends to divide the Hamiltonian according to a small useful part and an poorly understood large part to with respect to which will be applied the interaction frame $V = \exp\left(i\int H_{\mathrm{big}}dt\right)$.

$$H = H_{\mathrm{small}} + H_{\mathrm{big}}$$

$$H_{eff} = \exp\left(i\int_0^t H_{\mathrm{big}}dt'\right) H_{\mathrm{small}} \exp\left(-i\int_0^t H_{\mathrm{big}}dt'\right)$$

The Magnus expansion is then taken with respect to this interaction Hamiltonian. As a general rule, the larger H_{big} the faster the effective Hamiltonian will rotate and the convergence properties will be better. Note also that unlike for the adiabatic case, the interaction frame is with respect to a large parameter rather than the small one.

Chapter 5

Numerical Simulation and Optimization

The standard quantum control problem is to find a way to implement a desired unitary operation U_{ideal}, given discrete control vectors $\mathbf{u}_k$ each of length N. It is standard to assume that the underlying Hamiltonian is of the form $\hat{H} = \hat{H}_0 + \sum_k c_k(t)\hat{H}_k$ where $\hat{H}_k$ is the control Hamiltonian and $c_k(t)$ is a continuous field which is parameterized by the control $\mathbf{u}_k$. The goal is to choose controls $\mathbf{u}_k$ such that at the final time $t = T$, the dynamics of this system approximates the target unitary operator: U_{ideal}. To find these controls one first needs to simulate the dynamics of the system, given the Hamiltonian. In addition, a cost function C is needed to describe how close a simulated operation is to U_{ideal}. This allows us to evaluate the performance of each $\mathbf{u}_k$ and thus search the space of controls for an optimum. Further refinements to this method will be discussed in Ch. 6.1.

5.1 Time-slicing

In order to numerically integrate the Schrödinger equation it is necessary to discretize the evolution, normally into piecewise constant intervals. The approach traditionally used (e.g. in Ref. [85]) is to also use this discretization to map between the continuous fields and the control vector. That is

$$c_k(t) = \sum_{j=0}^{N-1} u_{k,j} \sqcap_j (t, \Delta t), \tag{5.1}$$

with the rectangle function

$$\sqcap_j (t, \Delta t) \equiv [\Theta(t - j\Delta t) - \Theta(t - (j+1)\Delta t)], \tag{5.2}$$

where Θ is the Heaviside unit step function. Evaluating the Schrödinger evolution using these fields and the Hamiltonian in Eq. (2.5) yields

$$U(0, T) = \prod_{j=N-1}^{0} U_j, \tag{5.3}$$

where

$$U_j = \exp\left[-i\hat{H}(j\Delta t)\Delta t\right], \tag{5.4}$$

the product running backwards is to enforce time-ordering, and $N\Delta t = T$.

5.2 Cost functions

In Eq. (2.7) we saw that the maximum error given the most pessimistic input states was the required figure of merit for using the error threshold theorem. The average error Eq. (2.8) and for qubits Eq. (2.9) were given as appropriate alternatives whenever certain basic conditions such as low Hilbert space dimension or variable input states could be ascertained. Since the time-slicing equation Eq. (5.3) requires the multiplication and exponentiation of N d-dimensional matrices, which scales as $O(Nd^3)$, the criterion of low dimension for control theory can usually be met thereby allowing the use of average error as a cost function.

5.2.1 Overlap fidelity

For the purposes of evaluating quantum gates, it is standard to use another form of the average fidelity, namely the *overlap fidelity*

$$\Phi_4 = \frac{1}{d^2}|\mathrm{Tr}(U^\dagger U_{get}^\dagger(T))|^2, \tag{5.5}$$

where d is the dimension of the Hilbert subspace . It tells how close the two unitary operations are to one other where the squaring effectively removes a global phase and gives the interpretations of a probability (for further discussion of Φ_4 see [85]). This fidelity is equivalent to the average fidelity, Eq. (2.8), via [104]

$$F_g = \frac{1}{d_Q + 1}(1 + d_Q\Phi_4),$$

with d_Q the dimension of the computational subspace. This overlap form is preferred because it does not make reference to the states of the quantum system and can be computed more quickly using fast matrix multiplication.

5.2.2 Subspace fidelity

For larger dimensions where the qubits are embedded in a larger Hilbert space (e.g., containing leakage channels, a cavity, or a non-Markovian environment) we can use instead [148]

$$\Phi_5 = \frac{1}{d_Q^2}|\mathrm{Tr}(U^\dagger \mathbb{P}_Q U_{get}(T)\mathbb{P}_Q)|^2, \tag{5.6}$$

where $\mathbb{P}_Q$ is a projector onto the computational subspace. This formulation reduces the search space because it ignores correctly what happens to states mapped from outside of the computational subspace back onto outside of the computational subspace. This effectively increases the number of possible optimal solutions and reduces the expected search time.

We can also decompose the desired unitary evolution $U_{get}^\dagger(T) = \prod_{j=N-1}^0 U_j = \prod_{j=N-1}^0 U_j \sum_{\psi_i \in B_H} |\psi_i\rangle\langle\psi_i|$ where B_H is a basis for the entire Hilbert space into rather a basis sum where we do not include basis elements outside those of the qubits' subspace B_{H_Q}; the whole Hilbert space dimension is however still used for the simulation.

$$\begin{aligned}
\Phi_6 &= \frac{1}{d_Q^2}|\mathrm{Tr}(U^\dagger \mathbb{P}_Q \prod_{j=N-1}^{0} U_j \sum_{\psi_i \in B_H} |\psi_i\rangle\langle\psi_i|\mathbb{P}_Q)|^2, &(5.7)\\
&= \frac{1}{d_Q^2}|\sum_{\psi_i \in B_{H_Q}} \langle\psi_i|U^\dagger \prod_{j=N-1}^{0} U_j|\psi_i\rangle)|^2 &(5.8)
\end{aligned}$$

Computing the fidelity then requires multiplying into d_Q vectors, where each multiplication is $O(d^2)$. The total fidelity given N time steps then takes $O(Nd_Q d^2)$ which can often be smaller then $O(Nd^{2.8})$ usually required for N matrix multiplications by standard numerical software packages.

5.2.3 Ensemble fidelity

If one wants to include the incoherent dynamics (Sec. 2.1.4) by using numerical rather than analytic techniques (Sec. (3.5)), it is necessary to optimize over an

ensemble of different descriptions of the system. To account for this, we define an average performance index

$$\Phi_7 = \int_p f(p)\Phi_j(p)dp, \tag{5.9}$$

where $\Phi_j(p)$ is one of the given fidelity functions evaluated using a value of the ensemble parameter p , which is unknown for specific instantiations of the quantum computer but occurs according to the probability distribution $f(p)$ (lacking specific knowledge we in this thesis set $f(p) = 1$). In practice, the integral is replaced by summation, which we sample discretely as $p, p_2, \ldots p_M$.

As a consequence of sampling, the optimization may find control sequences that result in discretely different trajectories in the Hilbert space from one value of the ensemble parameter p_i to the next (e.g. a different winding number around the Bloch sphere). Then, values of the parameter in between may have an average trajectory that differs significantly from either of the successful trajectories, thereby making the overall ensemble unsuccessful. While this situation can largely be alleviated by increased sampling, in practice it can also be mitigated by ensuring that the global phase of the operation from one ensemble member to the next remains the same, eliminating at least one common way in which different trajectories can happen. Thus, we move the squaring of the unitary product outside the sum of fidelities, which only disregards the common phase, giving

$$\Phi_8 = \frac{1}{(dM)^2}\left|\sum_{l=1}^{M} \mathrm{Tr}\left(U^\dagger U_{get}^{(p_l)}\right)\right|^2.$$

This fidelity is over-constrained but often gives better results in practice (see Sec. 8.8.2). Of course one can seek further refinements to the constrained solution by optimizing further over Eq. (5.9) once a solution is found. As an additional insurance, one will want to sample non-uniformly over the ensemble parameter p.

5.2.4 Penalties

There are a number of additional constraints that one may want to impose on the controls from a practical point of view. For example one may want to limit the energy of the pulses or avoid population of certain states during the gate. This is often accomplished by adding additional cost elements to the cost function, where the prefactor tells the relative importance of the terms.

$$\Phi_{\mathrm{tot}} = \Phi_j + \sum_k \lambda_k P_k,$$

where Φ_j is the chosen fidelity function and each penalty P_k has a multiplier λ_k associated with it. In the ideal case, λ_k is a Lagrange multiplier, whose value is optimized for the minimum point in the fidelity. However, in practice, choosing a different value may be more fruitful in terms of the trajectory used to get to the optimum point by an optimization algorithm. That is, a large λ_k will enforce the constraint more rigourously in the end, but it will give less freedom in terms of the search path towards an optimum, making it potentially harder to reach.

Unfortunately, any of the constraints that were tried in this thesis proved to be too demanding for the available controls in the sense that the optimization was over-constrained, giving a tradeoff between the cost terms. This is particularly true in the limit of short times where the controls are often already fully constrained. Thus if one wants to add in extra constraints it is usually necessary to increase the times of the pulses which may be more detrimental if irreversible forms of damping are significant in the error. In at least one case where the hard constraints are necessary to include on our pulses, Sec. (6.1), we find that better describing the unitary evolution is a more accurate and elegant way than by manually setting the value of the multiplier to give a realistic pulse.

5.3 Gradient Ascent

The mathematical statement of the optimization problem is to maximize Φ with respect to the vector $\mathbf{u}_k$. It is not immediately apparent how one should perform this optimization. One of the simplest approaches is the steepest ascent (or gradient search) algorithm [20]. If we consider the multi-dimensional surface (control landscape) formed by $\Phi(\mathbf{u}_k)$, steepest ascent is an iterative update procedure that locally examines the landscape at each iteration and provides a new $\mathbf{u}_k$ by moving in the direction that increases Φ the greatest. We take some initial guess for the controls and iteratively update according to the rule $\mathbf{u}_k \rightarrow \mathbf{u_k} + \epsilon \nabla \mathbf{u_k} \Phi(\mathbf{u_k})$, where ϵ is a unitless step matrix. Given an arbitrary initial configuration for the control fields, the algorithm follows a steepest ascent to a local optimum, at least in the case of small ϵ. For simplicity, in this work we take ϵ to be a scalar which is chosen adaptively. It is in no way clear that this procedure will produce anything other than *local* maxima, but remarkably, for the types of Φ's we examine in quantum control, the landscape is sufficiently under-constrained that gradient searches find *global* maxima with high probability (to arbitrary precision), in particular for small dimensions and large gate times. This has implications for the topology of the control landscape [34, 145].

The GRadient Ascent Pulse Engineering (GRAPE) algorithm [85] is an example

of a gradient search technique. A principle insight of GRAPE is that the gradient of Φ can be computed more efficiently than that of an arbitrary function. Specifically, the derivative of Φ_5 with respect to one of our control variables is

$$\begin{aligned}\frac{\partial \Phi_5}{\partial u_{k,j}} =& \frac{2}{d_Q^2}\mathrm{Re}\left[\mathrm{Tr}\left(U_{\mathrm{ideal}}^\dagger \mathbb{P}_Q \frac{\partial U(T)}{\partial u_{k,j}} \mathbb{P}_Q\right)\right. \\ & \left. \times \mathrm{Tr}\left(U_{\mathrm{ideal}} \mathbb{P}_Q U(T)^\dagger \mathbb{P}_Q\right)\right],\end{aligned} \tag{5.10}$$

where

$$\frac{\partial U(T)}{\partial u_{k,j}} = \left(\prod_{n=N-1}^{j+1} U_n\right) \frac{\partial U_j}{\partial u_{k,j}} \left(\prod_{n=j-1}^{0} U_n\right). \tag{5.11}$$

The speedup over gradient searches based on a naïve approach where one queries the fidelity function for each variation in the control values comes from the observation that the forward evolution $U(t,0)$ and the backwards evolution $U(T,t)$ need only be calculated once for a given control configuration. This in turn allows each derivative to be calculated with a constant number of matrix multiplications, as opposed to the N required for the full Schrödinger evolution, leading to an overall scaling of $O(N)$ as opposed to $O(N^2)$.

The exact form of the derivative of the unitary time slice can be found in the original GRAPE paper [85],

$$\begin{aligned}\frac{\partial U_j}{\partial u_{k,j}} &= -i\Delta t \bar{H}_{k,j} U_j, \\ \bar{H}_{k,j} &= \frac{1}{\Delta t}\int_0^{\Delta t} e^{-i\hat{H}(j\Delta t)\tau} \hat{H}_k e^{i\hat{H}(j\Delta t)\tau} \mathrm{d}\tau.\end{aligned} \tag{5.12}$$

For fine-grained control fields, i.e. $\|\hat{H}\Delta t\| \ll 1$, the derivative can be approximated as

$$\frac{\partial U_j}{\partial u_{k,j}} = -i\Delta t \hat{H}_k U_j. \tag{5.13}$$

This approximation can lead to difficulties in finding optimal solutions if the time step Δt is not sufficiently small, but this problem can be circumvented by evaluating the integral in Eq. (5.12) exactly [109]. Calculating this integral using the given form of U_j is straightforward after obtaining the diagonalization of $H(j\Delta t)$ (which is the preferred method of exponentiating the Hamiltonian given its Hermiticity [120]). Thus, the form of the derivative of the time slice U_j with respect to the k-th control [as in Eq. (5.12)] in the diagonal basis of the full Hamiltonian is

$$\langle m_j|\frac{\partial U_j}{\partial u_{k_j}}|n_j\rangle = \langle m_j|\hat{H}_k|n_j\rangle \frac{e^{-i\lambda_{m_j}\Delta t} - e^{-i\lambda_{n_j}\Delta t}}{\Delta t(\lambda_{m_j} - \lambda_{n_j})} \tag{5.14}$$

for $n_j \neq m_j$, or

$$\langle m_j|\frac{\partial U_j}{\partial u_{k,j}}|m_j\rangle = -i\Delta t\langle m_j|\hat{H}_k|m_j\rangle e^{-i\lambda_{m_j}\Delta t} \tag{5.15}$$

if $n_j = m_j$. Here, λ_{m_j} and λ_{n_j} are eigenvalues of the full Hamiltonian at time slice j, associated with eigenvectors $|m_j\rangle$ and $|n_j\rangle$ respectively. The actual derivative is then calculated by reverting the matrix elements back to the original, non-diagonal frame.

The gradient formula, Eq. (5.10), explicitly assumes that the total Hamiltonian, control plus drift, is constant across each slice. For situations where the pulse shaping resolution limited by the pulse generator's sampling rate is slower than the Hamiltonian dynamics, it is important to consolidate the two time-scales. Several approaches have been tried that avoid changing the existing framework. The first is to discretize the dynamics into smaller slices and to penalize the differences between adjacent control values. This would lead to a search which has a dimension that scales with the smallest time-scale. Another approach would be to coarse-grain in a way that completely discards the fast dynamics. For example, one can ignore the counter-rotating terms in the rotating wave approximation (RWA). Sec. 6.1 develops a method that redresses this tradeoff.

5.3.1 Non-unitary control

There are exceptions to the rule that shorter times lead to less decoherence for systems with non-Markovian decoherence [150, 157] or approximately decoherence-free subspaces [102, 131, 187]. In particular, rather than looking at the Hilbert space evolution once can look at the evolution of the Liouvillian superoperator $\mathcal{L}$ obeying the master equation

$$\dot{\mathcal{L}} = -i\mathcal{H}\mathcal{L},$$

which is equivalent to the evolution Eq. (2.4) and where the operator $\mathcal{L}$ gives the evolution of a vector whose entries are the elements of the density matrix ρ, though it is no longer unitary. The solution is given once again by the exponential

$$\mathcal{L} = \mathcal{T}\exp(-i\mathcal{H}\Delta t).$$

This technique is not used in any of the optimizations in this thesis and rather Eq. (2.4) is used for evaluation of the optimized unitary pulses. Note also that penalty functions which avoid the $O(d^4)$ matrix elements in $\mathcal{L}$ for efficiency reasons have been proposed [180] provided the decoherence is weak (which is usually the case for quantum computation).

5.3.2 Robust control

For ensemble control, the fidelity is a linear combination of all the possible Φ_ψ, and therefore the gradient of Φ is found by

$$\frac{d\Phi}{du_{k,j}} = \int_\psi \frac{d\Phi_\psi}{du_{k,j}} d\psi. \tag{5.16}$$

For discrete sampling we must provide the derivative, $\partial\Phi/\partial u_{kj}$, which we calculate by

$$\frac{\partial\Phi}{\partial u_{kj}} = \frac{2}{(dM)^2}\mathrm{Re}\left(\mathrm{Tr}\left[V^\dagger \sum_{l=0}^{M} \frac{\partial U^{(l)}}{\partial u_{kj}}\right] \sum_{l'=0}^{M} \mathrm{Tr}\left[V^\dagger U^{(l')}\right]^*\right). \tag{5.17}$$

Essentially, we must calculate $\partial U^{(l)}/\partial u_k$ for M different evolutions leading to a computational-scaling for evaluating this derivative that is only M times more than standard GRAPE on a d-dimensional system. However, this is significantly faster than optimizing a gate for a dM-dimensional Hilbert space since exponentiating and multiplying $d \times d$ matrices scales as $\mathcal{O}(d^3)$.

5.4 Convergence improvements

Speed was not a central issue in this thesis as most applications were illustrative and did not involve large Hilbert space dimension. Nonetheless, second order techniques such as Newton's method and LBFGS [109] allow for much faster optimization. These approaches are standard in quantum control as they decrease the number of iterations necessary by more accurately predicting the magnitude and direction of the step-size parameter ϵ. The fact that the optimization is multi-dimensional (i.e. over many controls) implies that this gradient/Hessian search is preferable to techniques that optimize over a single control at a time (such as Krotov's method [109]) since the landscape is climbed more efficiently.

Another improvement that will be necessary for further scaling of the numerics with dimension is to properly take into account the sparsity of most Hamiltonians.

This would work well in conjunction with Eq. (5.7). One way to take advantage of sparsity is by the so-called Hamiltonian-splitting method of exponentiation which is discussed in Appendix A.

Finally, given where modern computing is going, most search algorithms can benefit from some form of parallelism. In particular, there are good libraries for matrix multiplications over GPUs or network clusters. This is another much needed improvement for the optimization code developed for this thesis.

Chapter 6

Waveform Shaping

6.1 Introduction

As the control of quantum systems becomes increasingly sophisticated, shorter and shorter pulses which are less prone to decoherence become in principle possible. Given the stringent requirements for quantum computation which require very low errors and very precise timings, it is inevitable to run into the limitations of the (classical) hardware which is used to activate these computations. This is no more true than in engineered solid-state quantum systems, such as superconducting qubits (Sec. 2.3.1.1), where quantum coherence is typically limited to times on the order of microseconds [47], which necessitates very short pulses [27,106,183]. In these systems, the control technology is in the microwave frequency range where current state-of-the-art electronics restrict the variation in controls to a few nanoseconds [28, 107]. However, ideal gates operate exactly in the few nanosecond regime in order to allow thousands of operations in the characteristic coherence time. Thus, the techniques that are otherwise well-developed in other contexts need to be tailored to these constraints. Properly quantifying the shape and controllability of fast pulses using slow modulation is an important component of lowering effective error rates. There has been some progress in this direction with the development of smooth analytic optimal control methods such as DRAG [122] and some preliminary numerics in Refs. [50, 147]. An added benefit is that faster gates linearly speed up computation in these systems.

Most numerical control methods assume that the sampling rate of the pulse shape that can be optimized is identical to the sampling rate of the control field; in other words, either the control fields can be fully shaped in real time, or we shape the envelope of a driving field represented by the Hamiltonian in an appropriate rotating

frame where all remnants of time-dependence on the scale of the driving field disappear. In both cases, the time-evolution across a step of the shaped pulse, a pixel, can be straightforwardly approximated using a time-independent Hamiltonian in either frame.

In this chapter, we are addressing a more general case in which the Hamiltonian across the pixel is time-dependent. This can occur for a wealth of reasons. If the pulse-shaping hardware has an internal filter with a bandwidth that is much smaller than the input control sample rate, then the control fields become smooth rather than a sequence of plateaux (see Sec. 6.2). Secondly, for many applications where one requires a high-fidelity operation, performing a rotating wave approximation (dropping fast rotating terms from a Hamiltonian) becomes invalid when the control pulses are extremely short. That is, the counterrotating terms should be taken into the integration of the Schrödinger equation even though the input controls can only be changed on a time scale which is much slower than these terms. More generally, if multiple Fourier components of a driving field are applied in order to, e.g., induce AC Stark- and Zeeman-shifts, not all of these frequencies can be eliminated by a suitable transformation.

To account for this effect, we introduced an extra level of discretization. We separate the discretization necessary for integrating the quantum evolution from that of the discretization of the control parameters. The method is derived in Sec. 6.3. We show with some examples that only a few extra subdivisions are needed to greatly increase the accuracy of the optimization (in Sec. 6.5 and see also Sec. 11.2). Moreover, since the number of controls remains the same, the search space of possible pulses does not change and the time-cost of the algorithm is only linearly affected with the number of subdivisions. These results are published in Ref. [123].

Another important application of filtering and finite rise-times is in using frequency selection to turn a transition on and off, where they are the main source of error. This is generically discussed in Sec. 6.5 with an analytic solution found using the Magnus expansion of the evolution. The analysis is relevant in particular to frequency coupling of qubits (see Sec. 2.3.4) . This is explored in Sec. 6.7. and the full Hilbert space is also optimized using the numerical procedure.

6.2 Waveform generator example

As a motivating example, we have conducted a “theorist’s experiment" to show the typical response function given by an arbitrary waveform generator (AWG) to a set of digital inputs. In Fig. 6.1, the dashed purple line was inputed into a Tektronix AWG5014 (the AWG) and the response (solid green line) was measured on a LeCroy

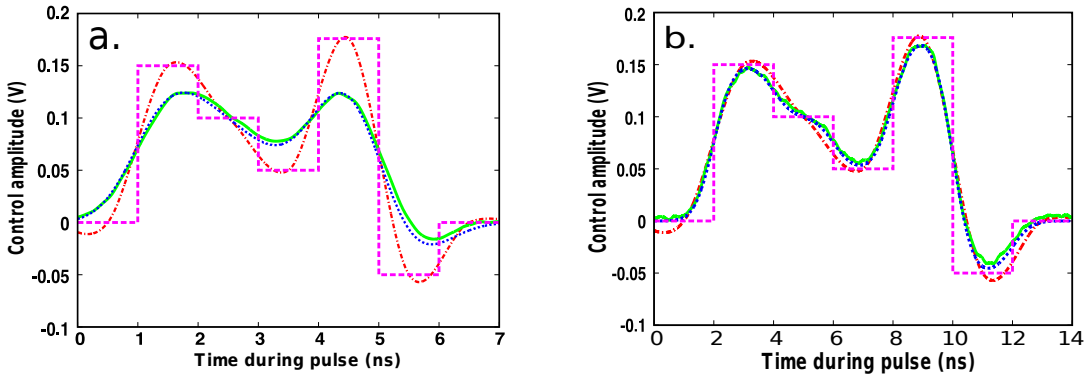

Figure 6.1: The purple dashed line represents the input controls $u_{j,k}$ for a arbitrary chosen 5 pixel pulse with (a) 1ns and (b) 2ns pixels, the green solid line is the output from the Tektronix AWG5014 running at 1 GSample/s (1ns pixels) and 500 MSamples/s (2 ns pixels) respectively with 14 bit vertical resolution, the red dot-dashed line is the cubic spline interpolation, and the blue dotted line is the Gaussian filter approximation (see text for details).

WavePro 725 Zi oscilloscope. This oscilloscope had a bandwidth of 2.5 GHz (-3dB) which was large enough to ensure that all smoothing of the pulse was from the AWG with a 250 MHz bandwidth (-3dB). The pixels were set to 1ns and 2ns in Fig. 6.1 (a) and (b), respectively. Here it is abundantly clear that the piecewise constant input control field (dashed purple line) and the waveform produced by current state-of-the-art AWG (solid green line) are not the same, with significant smoothing taking place. The remaining two lines (red dot-dashed and blue dotted line) represent approximations that will be discussed in Sec. 6.4.

6.3 Numerical optimization with fast fields

Our method is to separate the control parameters from the integration steps. This, for example, occurs when a control field arising from an arbitrary waveform generator (AWG) is mixed with a carrier field of frequency much larger than the allowed sample rate of the AWG, when the control field are a combination of multiple Fourier components [161], or when the control parameters are smoothed by a filter. In general, the field $c_k(t)$ can depend on all input controls $\mathbf{u}_k$, though in many practical situations it only depends on local values of $\mathbf{u}_k$. Furthermore, our drift and control Hamiltonians can be time-dependent.

To numerically integrate the Schrödinger equation, it is a necessity to discretize the continuous fields $c_k(t)$ into $s_{k,l}$, which is independent of our choice of $\mathbf{u}_k$. That is, time is digitized to $t = l\delta t$, where δt is the time scale chosen such that $\|\frac{dH}{dt}\|\frac{\delta t}{\|H\|} \ll 1$ and $M\delta t = T$.

The evolution takes the form of Eq. (5.3) where now

$$U = \prod_{l=M-1}^{0} U_l, \tag{6.1}$$

with the propagator,

$$U_l = \exp\left[-i\left(H_{0,l} + \sum_k s_{k,l} H_{k,l}\right)\delta t\right]. \tag{6.2}$$

Here, $H_{0,l}$ and $H_{k,l}$ are time-sliced versions of $H_0(t)$ and $H_k(t)$. The update rule for $\mathbf{u}_k$ is computed the same as Eq. (5.3), but now the gradient of Φ is found through

$$\frac{\partial \Phi}{\partial u_{k,j}} = \sum_{l=0}^{M-1} \frac{\partial s_{k,l}}{\partial u_{k,j}} \frac{\partial \Phi}{\partial s_{k,l}}, \tag{6.3}$$

with

$$\frac{\partial \Phi}{\partial s_{k,l}} = \frac{2}{d_Q^2} \mathrm{Re}\left[\mathrm{Tr}\left(U_{\mathrm{ideal}}^\dagger \mathbb{P}_Q \frac{\partial U(T)}{\partial s_{k,l}} \mathbb{P}_Q\right) \mathrm{Tr}\left(U_{\mathrm{ideal}} \mathbb{P}_Q U(T) dg \mathbb{P}_Q\right)\right], \tag{6.4}$$

where

$$\frac{\partial U(T)}{\partial s_{k,l}} = \left(\prod_{n=M-1}^{l+1} U_n\right) \frac{\partial U_l}{\partial s_{k,l}} \left(\prod_{n=l-1}^{0} U_n\right). \tag{6.5}$$

Calculating the $\partial U_l / \partial s_{k,l}$ proceeds exactly the same as in Sec. 5.3, either exactly or through a linear approximation. While this method will only ever be a linear approximation to the physical control fields, using the exact derivative as opposed to the linear approximation may speed up the gradient search. The partial derivative $\partial s_{k,l} / \partial u_{k,j}$ is provided by the response function $s_{k,l}(\mathbf{u_k})$.

6.4 Response functions

There are a wealth of response functions for taking the controls $\mathbf{u}_k$ to the continuous fields $c_k(t)$. These response functions do not have to be linear but in many circumstances a linear approximation is valid. For example, the filtering seen in Fig. 6.1 can is well approximated by a linear response function.

A linear response function can be represented by

$$s_{k,l} = \sum_{j=0}^{N-1} T_{k,l,j} u_{k,j}, \tag{6.6}$$

where the entirety of our response function is encapsulated in a transfer matrix, $T_{k,l,j}$. In this case, evaluating Eq. 6.3 is particularly straightforward since

$$\frac{\partial s_{k,l}}{\partial u_{k,j}} = T_{k,l,j}. \tag{6.7}$$

In the case where the input controls $\mathbf{u}_k$ are indexed by time, the response function (such as an interpolating function, a filter, or a carrier function) will simply reparametrize the weights of the individual control pixels. If the response function provides variations on a smaller time scale than the original input pixels then additional sub-pixels $s_{k,l}$ are required for the digitization of the continuous field to be valid. This idea is illustrated in Fig. 6.2, where in particular we can think of the gradient of the (purple-dashed) control pixels as the weighted sum of the gradients (the black arrows) of the sub-pixels (orange bars) inside or near the control pixel but which we cannot directly control. Note that this visualization works only when the control pixels are indexed in time. Otherwise, the input controls and response functions represent different dimensions (e.g. if the controls were Fourier components, the response would be in the time domain) though the same formalism still applies.

6.4.1 Cubic spline interpolation

If there is no detailed description of the transfer, smoothing can be approximated by interpolating between a discrete set of points with a piecewise cubic spline interpolation. Like the piecewise approximation in Eq. 5.1, here we approximate the continuous fields as a piecewise function

$$c_k(t) = \sum_{j=-1}^{N-1} S_{k,j}(t) \sqcap_{j+1/2} (t, \Delta t) \tag{6.8}$$

but now instead of remaining constant over the Δt interval the field is described by a cubic function,

$$S_{k,j}(t) = \tilde{a}_{k,j} t^3 + \tilde{b}_{k,j} t^2 + \tilde{c}_{k,j} t + \tilde{d}_{k,j}. \tag{6.9}$$

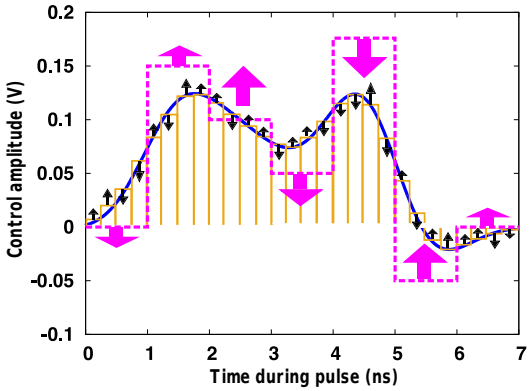

Figure 6.2: The purple dashed line represents the input controls pixels $u_{j,k}$, the solid blue line represents the continuous field $c_k(t)$, and the orange bars represent the sub-pixel approximation ($s_{k,j,l}$) to the continuous field. The black arrows represent the update vector of the individual sub-pixels while the large purple arrows represent the update (weighted average) for the control pixels.

To ensure that the field and its first derivative are continuous across boundaries and to introduce the control parameters, we enforce

$$\begin{aligned} S_{k,j}((j+1/2)\Delta t) =& u_{k,j}, \\ S'_{k,j}((j+1/2)\Delta t) =& \frac{u_{k,j+1} - u_{k,j-1}}{2\Delta t}, \\ S_{k,j}((j+3/2)\Delta t) =& u_{k,j+1}, \\ S'_{k,j}((j+3/2)\Delta t) =& \frac{u_{k,j+2} - u_{k,j}}{2\Delta t}. \end{aligned} \tag{6.10}$$

The box functions are offset by $\Delta t/2$ so that the control parameters indicate the value at the center of the step and we additionally require that the function and its derivatives are zero at the boundaries (which can be enforced by padding the control vector with zeros).

We can now derive the transfer matrix $T_{k,l,j}$ from the above conditions, and it is sparse. For each k, l pair there are only four non-zero transfer matrix elements. There exists a j' such that $(j'+1/2)\Delta t \leq l\delta t < (j'+3/2)\Delta t$. The non-zero matrix

elements of $T_{k,l,j}$ are then

$$\begin{aligned} T_{k,l,j'-1} &= -\frac{\tau}{2\Delta t}(\frac{\tau}{\Delta t}-1)^2 \\ T_{k,l,j'} &= 1+\frac{3\tau^3}{2\Delta t^3}-\frac{5\tau^2}{2\Delta t^2} \\ T_{k,l,j'+1} &= \frac{\tau}{2\Delta t}+\frac{4\tau^2}{2\Delta t^2}-\frac{3\tau^3}{2\Delta t^3} \\ T_{k,l,j'+2} &= \frac{\tau^3}{2\Delta t^3}-\frac{\tau^2}{2\Delta t^2} \end{aligned} \tag{6.11}$$

where $\tau = l\delta t - (j'+1/2)\Delta t$. Spline fits of this form are shown in Fig. 6.1 as the red dot-dashed line. In these figures, the spline is a better approximation to the AWG for the 2ns pulse(Fig. 6.1b) than for the 1ns pulse (Fig. 6.1a); this is primarily because for longer pixels there is less of an issue with the attenuation of the controls (which this model of the spline does not capture).

6.4.2 Filter function

Most electronic systems undergo some amount of filtering. It is important to model this to understand the shape the waveforms will take upon reaching our quantum systems. To represent the effect of the filter the continuous field is

$$\begin{aligned} c_k(t) &= \int_{-\infty}^{\infty} f_k(t-t')u_k(t')dt' \\ c_k(t) &= \mathcal{F}^{-1}[F_k(\omega)\mathcal{F}[u_k(t)]], \end{aligned} \tag{6.12}$$

where $f_k(t)$, and $F_k(\omega)$ are the response function and transfer function respectively of the filter for control k and $\mathcal{F}$ is the Fourier transform. As before, causality (the upper bound of the integral) is implicitly taken into account by a suitable delay time of the response. To ensure that the pulse is zero at the boundaries we add a finite number of control parameters to the start and end of the pulse that are fixed at zero. The number of parameters necessary for such a padding, n_r, depends on the bandwidth of the filter, ω_B, according to $n_r = \lceil 2\pi/(\omega_B\Delta t)\rceil$.

In the frequency domain, the response function is

$$s_{k,l}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u_k(t')\int_{-\infty}^{\infty} F_k(\omega)e^{i\omega(l\delta t-t')}d\omega dt' \tag{6.13}$$

Using this with the piecewise constant control fields in Eq. (5.1) and the assumption that the filter function in the frequency domain is even, the transfer matrix becomes

$$T_{k,l,j} = \int_{-\infty}^{\infty} F_k(\omega) \frac{\cos[\omega(l\delta t - (j + \frac{1}{2}\Delta t)] \sin[\frac{1}{2}\omega\Delta t]}{\pi\omega} d\omega. \tag{6.14}$$

It is easy to pre-compute $T_{k,l,j}$ numerically, and for many filters it can be calculated analytically. Note also if $|l\frac{\delta t}{\Delta t} - j| > n_r$ we can assume that $T_{k,l,j} \approx 0$. We demonstrate a calculation of $T_{k,l,j}$ using a Gaussian filter. In this case, the filter function is given by

$$F_k(\omega) = \exp(-\omega^2/\omega_{0_k}^2) \tag{6.15}$$

where ω_{0_k} is the reference frequency [178] for the k-th control. Performing the integration we find

$$T_{k,l,j} = \frac{1}{2}\left\{ \text{erf}\left[\omega_{0_k}\left(\frac{l\delta t - j\Delta t}{2}\right)\right] - \text{erf}\left[\omega_{0_k}\left(\frac{l\delta t - (j+1)\Delta t}{2}\right)\right]\right\}. \tag{6.16}$$

Gaussian filters are often used to approximate the actual hardware filtering typically found in experiments, which are usually parametrized by their bandwidth ω_B (frequency of 3dB attenuation). For a Gaussian the 3dB attenuation occurs at $\omega_B = 0.5887\omega_0$. Thus for the AWG with a bandwidth of $\omega_B/2\pi = 250$ MHz we find $\omega_0/2\pi = 425.4$ MHz. The Gaussian approximation to the Tektronix AWG5014 is shown in Fig. 6.1 as the blue dotted line. Here, we see that it reasonably well approximates the experiment (much better then the spline) and for the remainder of the chapter, we used this as our benchmark.

6.4.3 Fourier components

In previous sections, we envision the control parameters as the magnitudes of piecewise constant functions, i.e. square pulses, and the approximations as smoothed versions of these fields. In principle, we could instead look at the $u_{k,j}$ as more arbitrary parameters such as Fourier components (see e.g. Ref. [136]), where now we write

$$c_k(t) = \sum_j u_{k,j}(t) \sin(\omega_j t + \phi_j), \tag{6.17}$$

and the transfer matrix is simply

$$T_{k,l,j} = \sin(\omega_j l\delta t + \phi_j). \tag{6.18}$$

Often the the Fourier components $u_{k,j}$ are assumed to be a constant in time. On the other hand, perhaps the most common use of a Fourier component is to match that of the frequency of a qubit or other transition, as is the case of the lab frame Hamiltonian, Eq. 4.4. The repercussions of the fast time-dependence of this system is further analysed using this methodology in Ch. 11. It is also straightforward to combine other response functions such as filters with frequency component mixing. Finally, one can also consider optimizing the value of the phase and frequency of the Fourier components (e.g. Sec. 10.2.2) though this falls outside the domain of linear control.

6.5 Filtered amplitude-modulated coupling

For the first example, we consider filtering at the Rabi frequency of a qubit system. To avoid the trivial area theorem solution to an on-resonant 2-level system (Sec. 3.1) which is not greatly affected by filtering, we consider the incremental complication of a driven qutrit system. This system presents an important and widely applicable control problem in its own right which will be studied in great detail in Ch. 8 but here we are merely concerned with the (adverse) effect of filtering. We begin with the following lab-frame Hamiltonian:

$$H = \left[\mathcal{E}^x(t)\cos(\omega_1 t + \psi) + \mathcal{E}^y(t)\sin(\omega_1 t + \psi)\right](\Gamma + \Gamma^\dagger) + \sum_{j=1,2} \omega_j |j\rangle\langle j|, \tag{6.19}$$

where ω_1 is the $0-1$ transition frequency, ω_2 is the $0-2$ transition frequency, ψ is the unknown phase of the carrier signal at the start of the pulse, Γ is the effective lowering operator $\Gamma = |0\rangle\langle 1| + \sqrt{2}|1\rangle\langle 2|$, and $\mathcal{E}_x(t)$ and $\mathcal{E}_y(t)$ are the controls. With this system we want to perform a π rotation in the subspace formed by the 0 and 1 states without losing population to the third level. It is standard to define this π pulse in the frame rotating at the frequency ω_1 (Sec. 4.2.2). Moving to this frame and making the rotating wave approximation (RWA), we can model this system by the Hamiltonian

$$H = \mathcal{E}^x(t)\frac{(\Gamma + \Gamma^\dagger)}{2} + \mathcal{E}^y(t)\frac{(i\Gamma - i\Gamma^\dagger)}{2} + \Delta|2\rangle\langle 2|, \tag{6.20}$$

where $\Delta = 2\omega_1 - \omega_2$.

Under the RWA with two controls (with infinite bandwidth and sampling rate) it is always possible to find a solution that gives a perfect π pulse [122]. However, when we restrict the control parameters to 1ns pixels (a typical setting for current

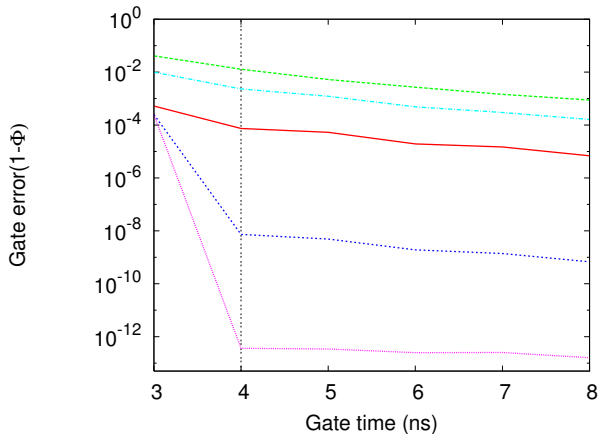

Figure 6.3: Gate error as a function of the gate time for the amplitude-modulation example. The dotted purple line is the optimal grape solution without any filtering. The dashed green line is the predicted error when a the optimal grape solutions are filtered by a Gaussian filter approximating a AWG with a 250 MHz bandwidth. The dot-dashed cyan line is the predicted gate error after filtering with the Gaussian filter for the spline optimization with 2 sub-pixels per control pixel. The solid red line and the dotted blue line are the predicted gate errors after filtering with the Gaussian filter for the Gaussian filter optimization with 2 and 20 sub-pixels per control pixel, respectively. The vertical dashed line indicates the gate time used in Fig. 6.4. Other parameters are given in the text.

microwave AWGs), we find that we need a gate time of least 4ns (4 pixels) to reach errors below 10^{-5}. This is shown in Fig. 6.3 as the dotted purple line where we have taken $\Delta/2\pi = -500$MHz. Naively, one would then predict that these optimized pulses could be perfectly implemented. However, due to the internal filtering imposed by the AWG this is not the case (see Fig. 6.1). To demonstrate this we take the numerically optimized pulses and filter them with a Gaussian filter approximation to a AWG with an internal bandwidth of 250MHz (see Sec. 6.4.2), which for short we will call the 'AWG filter'. The predicted error is then shown in Fig 6.3 as the dotted green line. Here we see that the effect of the filtering is drastic, and hence the optimized pulses will not perform well for quantum operations with the AWG due to shaping error.

Taking the filtering into account during the optimization allows for much better pulses. We can do this either by finding smooth pulses with a cubic spline interpolation or by actually taking into account the filter. To demonstrate this we plot in

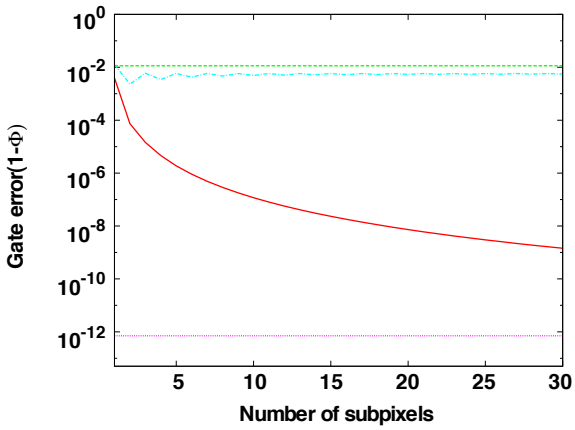

Figure 6.4: Gate error as a function of the number of subpixels for the amplitude-modulation example. The dotted purple line is the optimal grape solution without any filtering (limited by the numerical precision). The dashed green line is the filtered optimal grape solutions when the filter is a Gaussian filter approximating a AWG with a 250 MHz bandwidth. The dot-dashed cyan and solid red lines are the predicted gate errors after filtering with the Gaussian filter for the spline and Gaussian filter optimizations, respectively. Other parameters are given in text.

Fig. 6.3 the error as a function of gate time when optimizing under the cubic spline interpolation (dot-dashed cyan line) and a Gaussian filter (solid red line). Here we included 2 subpixels per control pixel, and after finding the optimized solutions, applied the AWG filter. Increasing the number of subdivisions allows our algorithm to better approximate the AWG filter; this is illustrated by the dotted blue line where we find for 20 sub-pixels a greatly improved performance for all gate times. To get an indication of the performance of our algorithm with the number of subdivisions, we set the gate time to 4ns (vertical dotted line in Fig. 6.3) and plot in Fig. 6.4 the gate error as a function of number of sub-pixels. Here, we observed that for only a few subdivisions the performance of our algorithm reaches very small error rates. We also find that the spline optimization is not as good as the Gaussian filter. This is expected as we have assumed the real situation (AWG filter) to be a Gaussian filter, hence the spline optimization is not necessarily going to find pulses that are consistent with the AWG filter. On the other hand, picking the correct response function improves the situation, even for very few subdivisions.

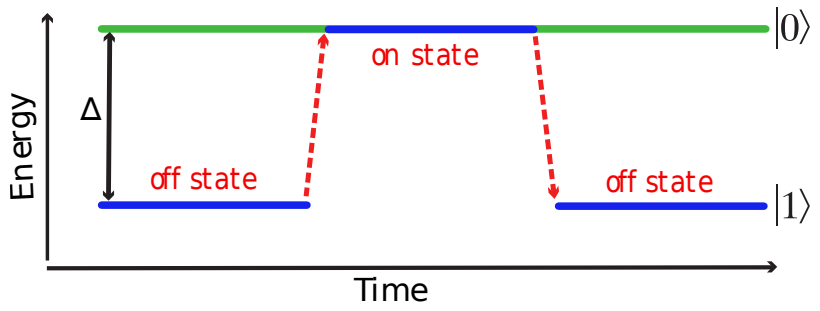

Figure 6.5: The strategy of frequency-modulated coupling is shown. The blue state $|1\rangle$ is brought into resonance with $|0\rangle$, allowed to exchange energy, and brought back to the off-resonant position.

6.6 Filtered frequency-modulated coupling

Now we consider a system where the coupling is constant and the control is achieved by modulation of the transition frequency in and out of resonance. It turns out this is a very simple way to tune multiple transitions on and off, and provides a straightforward method of e.g. couplings many qubits, with the caveat that it suffers from many imperfections (see Sec. 2.3.4). In particular, it is a common method of performing a 2-qubit gate and this will be explored in detail in the next section, 6.7. Here, we consider the generic case of a single such coupling with a biasing frequency control. Fig. 6.5 shows the basic strategy, with the corresponding Hamiltonian

$$H(t) = \frac{\Delta - \delta(t)}{2}\hat{Z} + J\hat{X} \tag{6.21}$$

We want to determine $\delta(t)$ such that at final time $T = \pi/2J$ we have $U(T) = \hat{X}$. The $\delta(t)$ control allows to effectively switch the $\hat{X}$ rotation on and off by putting the transition into resonance. In the ideal case, we simply select a rectangle function for $\delta(t) = \sqcap_0(0, T)\Delta$ which performs the rotation. In the off-state, Δ is chosen to be as large as possible (while avoiding coming to close to other transitions such as qubits or imperfections). Moreover, the off-state can also be optimized by e.g. using dynamical decoupling techniques (see Sec. 3.4).

However, a great difficulty with this strategy is that there will inevitably be filtering on $\delta(t)$ which prohibits an immediate jump from 0 to Δ but rather takes a characteristic rise-time (the yellow line in Fig. 6.6). Instead one must apply a new kind of dynamical decoupling pulse (Sec. 3.4) to cancel the unwanted term with the added complication from standard techniques that it is time-dependent.

Moreover, $\delta(t)$ acts simultaneously as the control we use to dynamically decouple the error and as the error term itself that we want to remove given the filtering. For demonstrative purposes of analytically optimizing the approximate functional form of $\delta(t)$, we impose the simplified condition that $|\dot{\delta}| \leq \mu$ though of course in practice actual functions that one can implement are greatly constrained by the filter.

In this example we divide $\delta(t)$ into three parts, the initial rise, $0 \leq t \leq t^*$, a period of time-independent evolution with $\delta(t) = \Delta$, $t^* \leq t \leq T - t^*$, and finally the ramp back down, $T - t^* \leq t \leq T$. The rise and fall of the field amplitude are the most difficult to assess because the terms in the Hamiltonian do not commute, unlike in the case of filtered amplitude-modulation where a simple area theorem would apply to driving the qubit. Furthermore, a small-parameter expansion is not possible given that $\delta(t)$ varies from very large to very small relative to J.

In order to analytically derive the form of the optimal control field, we use the average Hamiltonian theory and Magnus expansion taken to second order, defined in Sec. 4.4. During each of the three sections of the control pulse we would like the effective Hamiltonian to be $H_{\text{eff}}(t) = J\hat{X}$. This is exactly true for the free evolution when $\delta(t) = \Delta$ but for the ramps we need to look at the conditions given by each order of the Magnus expansion such that this is true.

The first order condition is simply the average Hamiltonian must be proportional $\hat{X}$, which is evaluated by integrating the Hamiltonian and gives rise to the condition $\int_0^{t*}(\Delta - \delta(t))dt = n2\pi$. Thus we have an approximate area theorem which states that we need to set the area above and below $\delta(t) = \Delta$ to be the same. This solution is illustrated in Fig. 6.6 as the red line. The error as a function of the energy separation Δ is plotted in red in Fig. 6.7.

To second order the double integral of the commutator of the Hamiltonian at different times gives the condition

$$\int_0^{t*}\int_0^{t_1}[H(t_1), H(t_2)]dt_2dt_1 = J\int_0^{t*}\int_0^{t_1}(\delta(t_1) - \delta(t_2))dt_2dt_1 = n2\pi,$$

and can be solved by using a cubic polynomial. Similarly, the next order corrections are given by satisfying Eq. 4.8 which give the quintic polynomial solution

$$\delta(\mathrm{t}) = 56\Delta + 20\Delta\frac{t}{\mathrm{t}^*} - 100\Delta\left(\frac{t}{t*}\right)^2 + 200\Delta\left(\frac{t}{t*}\right)^3 - 175\Delta\left(\frac{t}{t*}\right)^4 \tag{6.22}$$

This solution is illustrated in Fig. 6.6 as the blue line. The shape bears resemblance to a Fresnel integral. The error as a function of the energy separation Δ is plotted in blue in Fig. 6.7. We see that both the zeroth order and second order Magnus expansion

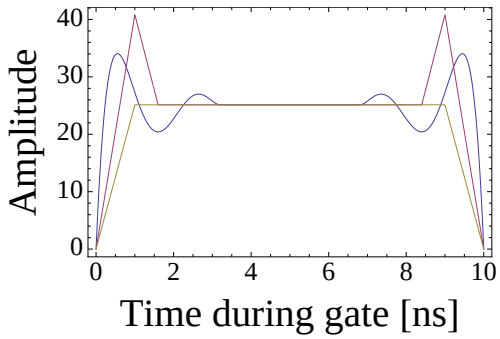

Figure 6.6: Control amplitude as a function of the time during the gate operation for the filtered ISWAP example. The yellow line is the unoptimized sequence with rise-time restriction. The red line represents the solution to the average Hamiltonian assuming the same rise-time of 1ns. The blue line represents the cubic polynomial solution to the second order in the Magnus expansion given by Eq. 6.22 with t*=3ns. The short time of T=10ns is used for visual clarity.

corrections greatly reduce the rise-time error at small splittings but become worse as the initial detuning is increased. This related directly to the observation that at the beginning of the pulse we have the detuning $\delta(t) \gg J$ while near resonance $\delta(t) \ll J$. The latter case is a clear application of average Hamiltonian theory where the expansion converges, while the former is not and requires a different treatment. In fact, it would be preferable to split the evolution into 5 parts, where the first and last are defined as the region where $\delta(t) \gg J$, and their contribution to the error is calculated differently. This type of error is known as off-resonant error and is an application of selectivity theory as will be studied beginning in Sec. 7.1.1 and again in later chapters.

6.7 Application to 2-qubit coupling

The implementation of quantum computing requires a universal set of gates [41]. With single qubit rotations, the $\sqrt{\mathsf{iSWAP}}$ or iSWAP [155] operations are universal two-qubit gates and turn out to be a natural two-qubit gate for certain coupling Hamiltonians [189]. In particular, they occur from an effective XX interaction between two resonant qubits, written as $H_{\text{int}} = -JX$ where $X = \sigma_1^+\sigma_2^- + \sigma_1^-\sigma_2^+$ can be understood as a Pauli matrix on the subspace spanned by $\{|01\rangle, |10\rangle\}$ and J is the coupling strength. Time evolution for a time t under this interaction leads to the unitary $U(\theta) = \exp(+iX\theta)$ with $\theta = t/J$. $\theta = \pi/4$ gives the $\sqrt{\mathsf{iSWAP}}$ and $\theta = \pi/2$

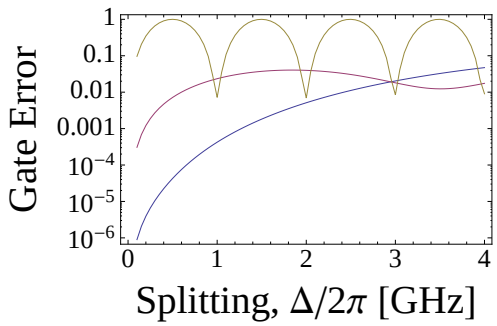

Figure 6.7: Gate error as a function of the splitting between energy levels for the filtered ISWAP example. The yellow line is the unoptimized sequence upon filtering. The red line represents the solution to the average Hamiltonian. The blue line represents the cubic polynomial solution to the second order in the Magnus expansion given by Eq. 6.22. Parameters are taken as J=30MHz and $T = \pi/J$.

gives the iSWAP. The XX interaction does not only occur naturally by the two-dimensional truncation of the exchange interaction. It occurs generically between weakly transversally coupled qubits [60]. This makes it ubiquitous in the interaction between resonantly tuned superconducting qubits [11, 30, 149, 185] as well as qubits and resonators [15, 14, 71, 111, 116, 159]. Precisely timing the pulse to the right angle can be done by either having a tunable coupling strength J [12, 70, 140, 142] or by bringing initially far detuned in frequency qubits into resonance. The latter is the more commonly used method. It is usually assumed that this tunneling is fully non-adiabatic, i.e., the parameters of the Hamiltonian change without any change of state. Realistically, it is slower than that, will exhibit a tendency to follow the adiabatic eigenbasis, and there will be a large error arising from the Landau-Zener tunneling. A similar situation occurs in the exchange-driven gates in the singlet-triplet qubit realized in semiconductor quantum dots [139]. Furthermore if the qubit is made from an anharmonic oscillator which is the case for many superconducting qubits there will be an additional source of error. This error is leakage from the two qubit subspace to the higher levels, controlled by the ratio between the coupling strength between and the anharmonicity; it will be discussed in Ch. 9.

6.7.1 Physical example: capacitively coupled phase qubits

As a physical motivation for how this coupling can arise, consider two superconducting phase qubits connected in parallel by a capacitor (see Sec. 2.3.1.1). Using Kirchoff's laws, it is straightforward [60] to derive circuit equations and from these

to deduce the respective Lagrangian

$$\mathcal{L} = I_{C1}\cos\phi_1 + \frac{C_1+C_2}{2}\frac{\Phi_0}{2\pi}\dot{\phi}_1^2 + I_{b1}\frac{\Phi_0}{2\pi}\phi_1 + I_{C2}\cos\phi_2 + \frac{C_1+C_2}{2}\frac{\Phi_0}{2\pi}\dot{\phi}_2^2 + I_{b1}\frac{\Phi_0}{2\pi}\phi_2 + C_G\frac{\Phi_0}{2\pi}\dot{\phi}_1\dot{\phi}_2$$

where the subscripts b and c denote bias and critical current respectively, the numerical subscript indexes the Josephson junction, and C_G is a coupling capacitor. Using canonical quantization with $E_{Ci}\hat{N}_i = \sqrt{\frac{\Phi_0}{2\pi}}\dot{\phi}_i$ this gives the Hamiltonian description

$$H = \sum_i [E_{C1}\hat{N}_1 - E_{Ji}\cos\hat{\phi}_i - I_{bi}\frac{\Phi_0}{2\pi}\hat{\phi}_i] + 4C_G E_{C1}E_{C2}\hat{N}_1\hat{N}_2$$

with the introduction of the Josephson energy E_{Ji}. This Hamiltonian is well approximated by a polynomial given by the Taylor expansion of the cosine around $\hat{\phi}_i = 0$,

$$H \backsimeq \sum_i [E_{C1}\hat{N}_1 + E_{Ji}\hat{\phi}_i^2 - I_{bi}\frac{\Phi_0}{2\pi}\hat{\phi}_i] + 4C_G E_{C1}E_{C2}\hat{N}_1\hat{N}_2$$

This Hamiltonian in turn is a harmonic oscillator and is exactly solvable. Introducing raising and lowering operators $\hat{\phi}_i = \sqrt{\frac{E_{Ci}}{2E_{Ji}}}(\hat{a}_i + \hat{a}_i^\dagger)$ and $\hat{N}_i = \sqrt{\frac{E_{Ci}}{2E_{Ji}}}i(\hat{a}_i - \hat{a}_i^\dagger)$ we can rewrite

$$H = \omega_1\hat{n}_1 + \omega_2\hat{n}_2 - 2C_G\sqrt{\frac{E_{C1}^3E_{C2}^3}{E_{J1}E_{J2}}}(\hat{a}_1 - \hat{a}_1^\dagger)(\hat{a}_2 - \hat{a}_2^\dagger)$$

with $\omega_i = \sqrt{4E_{Ci}E_{Ji}}$ and n the number of excitations in the qubit eigenbasis. Finally, we note that terms which do not conserve energy can be dropped (fast-rotating terms) and that the situation of variable qubit frequency can be realized by tuning E_C *in situ* by threading a flux through the Josephson junction loop. Truncating the Hilbert space by ignoring all states above $n_i = 1$ gives the two-qubit flip-flop Hamiltonian

$$H_{2Q} = \frac{\omega_1(t)}{2}\hat{Z}_1 + \frac{\omega_2(t)}{2}\hat{Z}_2 - 2C_G\sqrt{\frac{E_{C1}^3E_{C2}^3}{E_{J1}E_{J2}}}(\sigma_1^-\sigma_2^+ + \sigma_2^-\sigma_1^+)$$

6.7.2 Reduction to single qubit analytic treatment

We consider two qubit gates realized by an exchange interaction,

$$H = \sum_i \frac{\omega_i(t)}{2} Z_i + J(\sigma_1^- \sigma_2^+ + \sigma_2^- \sigma_1^+) \tag{6.23}$$

where ω_i are the resonant frequencies of the two qubits which we assume can be varied and J is the coupling rate which we assume is fixed. This interaction is the natural interaction for two superconducting qubits coupled capacitively [116,133,132], inductively [11, 30, 185] or by virtually photons in the circuit QED architecture [15, 111]. The variation in the frequencies can be achieved by changing the operation point of the qubit with external flux lines. The above Hamiltonian can be dramatically simplified by defining a representation of the Pauli algebra given by,

$$\begin{aligned} \hat{I} =& \frac{\sigma_1^z + \sigma_2^z}{2}, \\ \hat{X} =& \sigma_1^- \sigma_2^+ + \sigma_2^- \sigma_1^+, \\ \hat{Y} =& -i\sigma_1^- \sigma_2^+ + i\sigma_2^- \sigma_1^+, \\ \hat{Z} =& \frac{\sigma_1^z - \sigma_2^z}{2}. \end{aligned} \tag{6.24}$$

In this case Eq. 6.25 reduces to

$$H = \Sigma(t)\hat{I} + \Delta(t)\hat{Z} + J\hat{X}, \tag{6.25}$$

where $\Delta(t)$ and $\Sigma(t)$ are respectively the difference and sum of the two qubit frequencies. Effectively we can limit our controls to a single excitation algebra if we choose $\int \Sigma(t)dt = 2n\pi$, and thus the task of designing a two-qubit gate is isomorphic to that of generating a single qubit X-rotation using a Landau-Zener-type sweep [126]. In particular, the analysis of Sec. 6.6 is applicable. As a testament to the simplicity of this strategy, it is widely used for coupling superconducting qubits and the first-order control sequence forms the basis of present-day embodiments of the technology [57].

6.7.3 Numerical optimization

Once again we can apply the numerical optimization gradient ascent technique to optimize the two-qubit gate in the presence of filtering. We optimize a single detuning field $\delta(t)$ in the rotating frame Hamiltonian given by

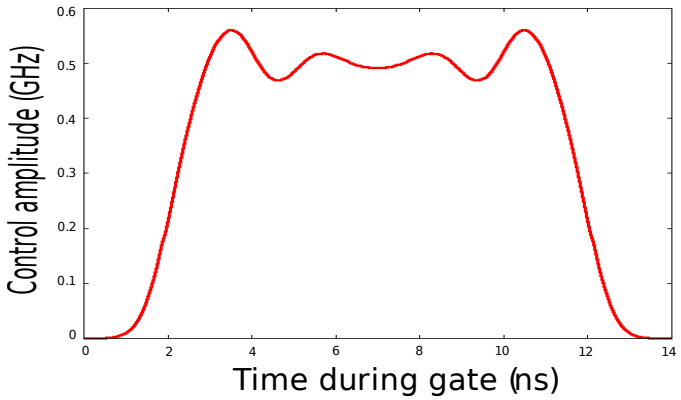

Figure 6.8: Optimized detuning $\delta(t)$ as a function of the gate time for the filtering 2-qubit example. Parameters are given in the text.

$$H = \frac{\Delta - \delta(t)}{2} Z_1 + J(\sigma_1^- \sigma_2^+ + \sigma_2^- \sigma_1^+) \tag{6.26}$$

For typical values of J (on the order of 50 MHz) and typical filtering (about 250MHz bandwidth) it is always possible to fully remove the error. Interestingly, the pulses found are very similar to the polynomial found using the Magnus expansion, Eq. 6.22 and one such pulse is plotted in Fig. 6.8 for $\Delta = 0.5$GHz, and T=12ns, with a 2ns buffer. Using this technique, the error can be reduced arbitrarily given the use of the correct response function and coupling that is not faster than the smoothing.

6.8 Summary

We have derived numerical and analytical techniques for dealing with fast-oscillations in the Hamiltonian underlying slow variation of the control parameters. For gates that are very short to avoid the effect of decoherence, fast oscillations can be particularly detrimental.

One kind of fast dynamics that has been studied in detail in this chapter is filtering. In the case of filtering of amplitude modulated pulses, the effect can be fairly subtle because the Hamiltonian approximately commutes with itself at different times and therefore the filtering manifests itself as rather straightforward shaping error. In the limit of simple time dependence, an area law or other suitable analytic approximation (by e.g. taking the filter as a Gaussian function) allowed relatively little disturbance to the study of single qubit rotations.

On the other hand, if the filtering affects an operator that does not commute approximately with the rest of the Hamiltonian then the effect of filtering is harder to understand. Analytically, a Magnus expansion of the evolution allowed to obtain a pulse shape for the detuning of a coupling transition that permitted to effectively turn the transition on and off as desired. Moreover, the shape can be easily understood as a polynomial satisfying a certain number of integral dynamical decoupling constraints and matched closely to the optimum pulse shape found through optimization using the derived numerical techniques.

Other instances of fast dynamics in the presence of slow controls will be discussed throughout the text. In particular, the presence of unwanted off-resonant transitions can be thought of in this context for fast pulses. Specific applications for the numerical technique will be discussed in Ch. 10 and Ch. 11 to optimize over multiple driving fields at different frequencies, or one such field with multiple harmonics present, respectively.

Chapter 7

Crosstalk

In chapter 6, we saw that imperfections can change the intended shape of the field that is sent to access a quantum coupling, hence causing errors if not accounted for. Now we consider imperfections not in the magnitude of the field but in its spatial extent. Both in real space and in frequency, it is desirable to separate qubits and other coupled energy levels from each other so that they are individually accessible. Crosstalk error occurs when manipulating one quantum element is not done independently of other elements. In order to minimize this effect, engineering qubits that are sufficiently separated in space and in frequency while retaining sufficient coupling and scalability constraints is required (see Sec. 2.3.2). This chapter aims to examine the techniques and tradeoffs of minimizing crosstalk terms while retaining coupling and scalability: section 7.1 deals with crosstalk for addressing individual qubits at different frequencies while in 7.2 spatial crosstalk for two-qubit gates is discussed. In general, crosstalk is ubiquitous across quantum control strategies and physical implementations of quantum information, namely it will reoccur in the context of leakage couplings (Ch. 8 and 9), virtual couplings (Ch. 10), and strong coupling (Ch. 11). In this chapter, we elaborate on the idea that frequency crosstalk consists primarily of semi-classical selectivity error (Sec. 3.2) and shifting of quantum eigen-energies (Sec. 3.3) which are primarily an exercise in pulse-shaping, while spatial crosstalk requires performing trajectories in state space that are able in composite (Sec. 3.4) to act independently of and unwind unwanted operators.

Section 7.1 analyses the selectivity problem (Sec. 7.1.1) first from a semi-classical perspective (Sec. 7.1.2) and with quantum corrections (Sec. 7.1.3). The writing of this section was aided by Frank Wilhelm. Next, Sec. 7.2 introduces the problem of spatially separating (flux) qubits while simultaneously retaining (tunable) coupling between them (the coupling results are derived and analyzed by Peter Groszkowski;

for the publication and the full derivation see Ref. [63]). The coupling landscape is used to demonstrate that sufficiently low error gates can be performed under realistic constraints and minimal spatial separation, both with (Sec. 7.2.3) and without (Sec. 7.2.4) the additional use of frequency separation of the qubits.

7.1 Frequency crosstalk

Spectroscopy is arguably the most commonly used experimental technique in physics [36, 101, 162]. It relies on the concept of resonance - the object of study is exposed to monochromatic radiation and responds if the radiation frequency matches a frequency of that system. In quantum systems, that frequency is the difference of two of the systems' energies. Complex systems usually contain a large wealth of these frequencies. The ability to selectively address these frequencies defines the spectral resolution. The limitation of spectral resolution can be twofold: On the one hand, the frequencies forming spectral lines are intrinsically broadened by decoherence and / or relaxation. On the other hand, the idea of an ideally monochromatic external excitation is only a convenient fiction - in reality, the bandwidth of that external signal is limited by a on a scale of $1/T$ where T is the duration of the experimental pulse.

In continuous wave spectroscopy, T is chosen to be so long that this restriction is not a limiting factor. There are however situations in which this is not possible. In magnetic resonance, e.g., certain spectral lines can only be reached through complex pulse sequences that all need to be executed within the relaxation time of the system. Consequently, a wealth of techniques has been developed that reaches fine spectral selectivity with pulses of limited duration, including the case of 2D-spectroscopy [45, 54, 101] .

Quantum technologies such as quantum computing are often based on spectroscopic ideas [24, 125, 46, 129]. In fact, the already mentioned spin resonance is a primary candidate for the implementation of quantum computing [8, 61, 73].This means that the quantum mechanical transitions corresponding to certain quantum logic operations are typically addressed through their transition frequency. This can occur on the level of single qubits, when the two states representing the qubit are taken out of a complex spectrum with low anharmonicity such as it is the case in superconducting qubits [30, 38, 112, 153, 186] (see Ch. 8). It can occur when multiple qubits are in close spatial proximity, much closer than the spatial resolution of the external field, as it is the case in spin resonance [54, 101, 129, 162]. It can also occur if single elements are multifunctional, e.g., when a single qubit contains transitions pertinent to local rotations as well as to coupling different components [14, 57, 71, 111, 141, 159]

(see Ch. 10). Note that typical scalable quantum computing implementation candidates typically do not rely on spectroscopic resolution alone and at least contain some element of local addressability. Still, clearly, a crowding of the frequency spectrum will be detrimental to both spectroscopy and coherent quantum control. In quantum information, it is a key requirement to perform a large number of highly accurate operations well below the coherence time of the system. Thus, the challenge of reaching good enough spectral addressability in short times is in particular significant.

The above discussion holds essentially for all bit-flip type operations, where transitions between energy eigenstates of a qubit system occur. On a single qubit level, these would e.g. include the NOT gate. Between qubits, these include the controlled NOT and the iSWAP gate [129].

Now a key difference between spectroscopy and quantum control in the pursuit of selective excitation is the following: spectroscopy is an analytic technique to find energy levels through transition frequencies, hence we want to guarantee that beyond a narrow band around the desired transition, excitation profiles are suppressed. In quantum control, the spectrum is well characterized and the positions of undesired transitions are known, hence, it is sufficient to suppress the excitation profile at those frequencies. This chapter aims primarily at the second approach.

7.1.1 Selectivity criteria

The controls that are used to manipulate quantum systems, typical external AC fields, can often neither spatially nor by selection rule distinguish between the quantum transition that is being controlled and other quantum transitions. This can be mitigated if all these transitions have distinct transition frequencies $\omega_{j,k} = E_j - E_k$ where E_j is the energy eigenvalue of state j and here and henceforth we use natural units with $\hbar = 1$. If we now drive the system control indexed by l with a drive frequency ω_l^d that is much closer to a specific transition frequency labelled by $j(l)$, $k(l)$ than to any other, and/or if this control has an appreciable matrix element $\hat{\Gamma}_{jk}^l$ for this transition, only it will be driven, and no other transition. We will quantify this statement below and outline its limitations.

We start by assuming a Hamiltonian $\hat{H}_{\text{tot}} = \hat{H}_0 + \hat{H}_{\text{control}}$ and work in the basis of eigenstate of $\hat{H}_0$. We can formalize the statement about spectral selectivity by assuming that the drive Hamiltonian has some appreciable matrix elements for

multiple quantum elements in the system, that is the control Hamiltonian

$$\hat{H}_{\text{control}}(t) = \sum_{l=0}^{L-1} \Omega_l(t) e^{-i\phi_l} \sum_{m=0}^{M-1} \hat{\Gamma}^l_m + h.c. \tag{7.1}$$

where there are M quantum elements in the system and L drives to control them. This can arise for example if we consider M qubits and a collective drive composed of L frequencies, each of which is meant to address a particular qubit but has unintended crosstalk on the rest of the M qubits, as it e.g. occurs in NMR. In such a case, the lab Hamiltonian will read

$$\hat{H}_{tot} = \sum_l \Omega_l(t) e^{-i\phi_l} \cos\left(\int_0^t \omega_l^d dt\right) \left(\sum_m X_m\right) + \sum_m \omega_{0,1}^m |1\rangle\langle 1|_m \tag{7.2}$$

Moving this to the interaction picture (Sec. 4.2.2), we can better appreciate the selectivity condition by introducing offsets $\Delta_{jkl} = \omega_l^d - \omega_{j,k}$ as

$$\hat{\Gamma}^l_m = \sum_{|j\rangle,|k\rangle \in \mathcal{H}_m} \lambda^l_{j,k} e^{-i\Delta_{jkl}t} |j\rangle\langle k|. \tag{7.3}$$

where the total Hilbert space is $\mathcal{H}_{sys} = \bigotimes_{m=0}^{M-1} \mathcal{H}_m$. In the qubits example, we have $|j\rangle = |0\rangle_m$, $|k\rangle = |1\rangle_m$. More generally, Hilbert spaces where selectivity criteria are used outside the scope of qubit crosstalk will be discussed in in Ch. 8, Ch. 10, and Ch. 11.

Here, it is tacitly assumed that the envelope $\Omega_l = \text{Re}\Omega_l + i\text{Im}\Omega_l$ is complex-valued. For the implementation of simple drive pulses, the phase is typically assumed to be constant which with appropriate choice of reference means $\text{Im}\Omega = 0$, however, later we will explicitly use the ability to control both terms independently. Here, we assume that to each index l we match a transition j, k to which it is almost resonant, identified as $j(l)$ and $k(l)$. In this interaction frame representation, we can quantitatively formulate the selectivity condition: we observe that the terms on the right hand side of Eq. (7.3) are oscillating at frequency Δ_{jkl} and have magnitude $\lambda^l_{j,k}\Omega$. In order for the drive to be effective, it must not oscillate wildly on the time scale of the transition $|\lambda_{j,k}\Omega|^{-1}$ hence giving us the sufficient criterion that we have frequency selectivity for all l if

$$\Delta_{jkl} \gg \left|\lambda^l_{jk}\Omega_l\right| \quad \forall j, k \neq j(l), k(l) \tag{7.4}$$

On the other hand, in order to drive the desired transition efficiently, we demand to be on or close to resonance

$$\Delta_{j(l),k(l),l} \ll \left|\lambda^{l}_{j(l),k(l)}\Omega_l\right| \tag{7.5}$$

As $\omega_{j,k}$ are given by the quantum system under consideration, the choice of driving frequencies l can only maximize the left hand side of Eq. (7.4) to a certain limit while also obeying Eq. (7.5). Thus, obeying these conditions requires to keep the control amplitudes Ω_l low enough. On the other hand, this increases the duration of the control pulse hence makes the transition vulnerable to decoherence and relaxation, thus we practically demand that $\lambda^{l}_{j(l),k(l)}\Omega_l \gg \gamma$ where γ represents typical incoherent rates of the system. This constraint on addressability is a result of spectral crowding and the loss of fidelity due to the need for long pulses degrades spectroscopic techniques as well as the implementation of coherent gates in a quantum computer. Thus, the spectrum sets a speed and fidelity limit on quantum control. Another way to observe this is using the Fourier transform which we will use in the next section.

7.1.2 Semiclassical spectrum analysis

In this section, we develop a classical argument for our technique of excitation suppression and relate it to ideas of signal processing. It is well established [54, 72, 174, 177] that for a system of qubits or spins 1/2 driven by a weak external field, i.e., for small T/Δ, an excellent measure of off-resonant excitation is the Fourier transform

$$S(\Delta) = \int_0^T (\Omega(t)e^{-i\phi})e^{-i\Delta t}dt \tag{7.6}$$

Note that this is a limited-interval Fourier transform that can be consolidated with the regular, infinite-time Fourier transform by assuming the pulse envelope $\Omega(t)$ vanishes outside the integration interval.

Now as another caveat, when applying this measure to the design of gates, we need to be aware that off-resonant levels can also induce phase errors (coming from, e.g., AC Stark shifts) [59]. It is suitable for population transfers, though. Often, phase errors in gates can be compensated for by inserting frame transformations between pulsed operations (Eq. 4.2).

One established way to compensate spectral weight off -resonance while maintaining a pulse of limited length is to use Gaussian profiles for $\Omega(t)$, Eq. 3.3, because they are also well-confined Gaussians in frequency space. In this case, the Gauss function describing $\Omega(t)$ must be suitably chosen to start and end at zero amplitude.

7.1.2.1 Doublet analysis

We will follow a different strategy that, rather than *reducing* off-resonant excitations for a full *band* of energies it *eliminates* excitation for one or more *discrete* frequencies.

Having established the role of the Fourier transform, we can now adapt an idea from signal processing and classical calculus. We start from the excitation profile for detuning Δ, then integrate by parts (IBP), assume that the real part of the envelope vanishes in both the beginning and the end of the pulse and find

$$\begin{aligned} S(\Delta) &= e^{-i\phi}\int_0^T (\mathrm{Re}\Omega(t) + i\mathrm{Im}\Omega(t))e^{-i\Delta t}dt \\ &= ie^{-i\phi}\int_0^T \left(\frac{\mathrm{Re}\dot{\Omega}(t)}{\Delta} + \mathrm{Im}\Omega(t)\right)e^{-i\Delta t}dt \end{aligned} \tag{7.7}$$

Thus, by imposing a derivative

$$\mathrm{Im}\Omega(t) = -\frac{\mathrm{Re}\dot{\Omega}(t)}{\Delta} \tag{7.8}$$

we can exactly cancel the Fourier transform at the relevant frequency, $S(\Delta) = 0$. So within the applicability of the Fourier transform elaborated above, this argument provides an intuitive motivation for the Derivative Removal by Adiabatic Gate (DRAG) strategy for removing unwanted quantum excitations discovered in [122] and first applied experimentally to solid state qubits in [28, 105]. In fact, in the DRAG scheme derived in Refs. [59, 122], equation (7.8)is the lowest order in $(\Delta T)^{-1} = O(\theta/\Delta)$ where θ is the desired rotation angle. This, specifically Ω/Δ, is the same parameter dictated by the applicability of the Fourier transform, Eq. 7.4 [72].

This strategy is illustrated in Fig. (7.1). The zero of the frequency axis is set to the wanted transition, the unwanted transition is placed at -1. The Gaussian definitely has appreciable spectral weight at the unwanted transition. The derivative also has spectral weight there, so the difference with the appropriate weight will be zero. By construction, the derivative of the Gaussian has no spectral weight at the working transition hence does not alter the spectral profile of the working transition.

7.1.2.2 Multiplet analysis

So far our only constraint was to bring $S(\Delta) = 0$ in Eq. 7.7. While the condition (7.8) will certainly do that by canceling the integrand at each point, this is not always possible or necessary. The constraint works for any function that integrates in total

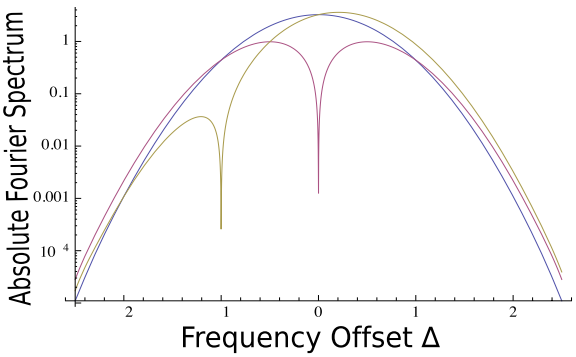

Figure 7.1: Abs Fourier transform of a Gaussian pulse(blue), its derivative(red), and their sum (yellow)

to the same area as the driving field. In particular, the off-resonant transition can be simultaneously driven by additional fields Ω_l at a different frequency difference $\Delta_{jkl} = \omega_l^d - \omega_{j,k}$ which must together satisfy the condition

$$\int_0^T \sum_l^n \Omega_l e^{-i\Delta_{jkl}t} dt = 0 \tag{7.9}$$

for excitation $j - k$ to be suppressed. For multiple transitions $\{jk\} \in S$, the set of unwanted transitions, this sets up a system of equations to solve. For example, for one such unwanted transition, if we choose to put a second tone symmetrically on the other side of the unwanted transition from the driving field ($\Delta_{jk2} = -\Delta_{jk1}$), the relation works out to give

$$\frac{\mathrm{Re}(\dot{\Omega}_1(t))}{\Delta_{jk1}} = \Omega_2(t) e^{-i\Delta_{jk1}T} \tag{7.10}$$

This will not work as well as Eq. 7.8 since it is only an average effect but may be useful in conjunction with other tones to remove multiple transitions. Choosing Ω_2 to be a derivative function rather than a function with area similar to Ω_1 also has the advantage of minimally disturbing the rest of the spectrum since Ω_2 will typically require an energy smaller than Ω_1 by a factor of about $(\Delta_{jk2})^2$.

In general, the spectrum may be quite crowded with unwanted transitions on either side of the working transition. Unfortunately, it is clear from Fig. 7.1 and from the anti-symmetry of Eq. 7.8 around $\Delta = 0$ that the derivative will only decrease the excitation on one side of the working transition while increasing it on the other. On

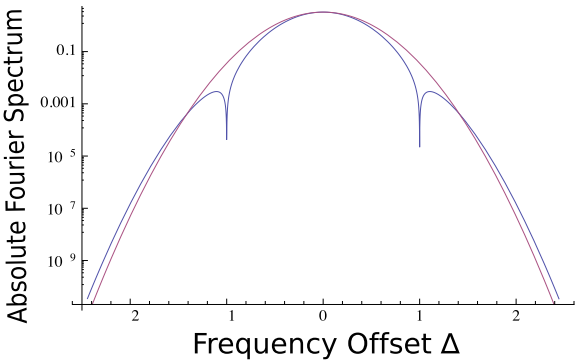

Figure 7.2: Absolute Fourier transform of a Gaussian(red) and the sum of the Gaussian with its 2nd derivative (blue)

the other hand, the second derivative of of the envelope will be symmetric and effect both sides in the same way. Applying IBP to Eq. 7.7 a second time (ignoring now the imaginary part) and setting now the boundary terms in the envelope function and its first derivative to 0 gives the relation

$$\int_0^{\mathrm{T}} \mathrm{Re}\Omega(t)e^{-i\Delta t}dt = \int_0^{\mathrm{T}} -\frac{\mathrm{Re}\ddot{\Omega}(t)}{\Delta^2}e^{-i\Delta t}dt$$

Clearly, supplementing the driving function with another tone in-phase and at the same frequency but proportional to the second derivative of the first will again satisfy Eq. 7.9 for

$$\Omega_2(t) = \frac{\ddot{\Omega}_1(t)}{\Delta^2}. \tag{7.11}$$

The effect is demonstrated in Fig. 7.2. Off resonance, unwanted transitions are cancelled at the chosen $\Delta = \pm 1$. On resonance, setting the area to the desired rotation angle θ via $\int_0^{\mathrm{T}}(\Omega(t) + \ddot{\Omega}(t))dt = \theta$ allows the working transition (at $\Delta = 0$) to remain unchanged. The cost of this technique is that the function and its first derivative both need to begin and end at 0 for IBP to be valid, and the pulse shape may be slightly harder to implement than a simple Gaussian or its first derivative. On the other hand, the strategy requires only one control to shape and synchronize and it obeys an area theorem avoiding higher order effects such as phase shifts and rotation errors on the working transition. An added benefit is that we can see the overall bandwidth (above a given threshold – here 0.001) is decreased compared to a traditional Gaussian where some of the energy has been moved from the selective

region to the tails where it instead falls below threshold. By placing the spectral holes appropriately at various frequency offsets one can engineer the bandwidth that falls above a desired selectivity threshold. This can be useful for narrowband behaviour where there are many transitions to be avoided such as used in resonance imaging and spectroscopy, with a decrease in bandwidth of about 25% compared to the optimal Gaussian with the same threshold.

For multiple unwanted transitions it is straightforward to satisfy Eq. 7.9 to null out a set S of unwanted spectral elements using multiple derivatives all at the same frequency ($\Delta_{jkl} = \Delta_{jk}$). Using the IBP formulae, the system of equations simplifies to a set linear equations

$$1 - \sum_{l=1}^{|S|} (\Delta_{jk})^l a_l = 0, \quad \forall \{jk\} \in S \tag{7.12}$$

$$\Omega_l(t) = (i)^l a_l \frac{d^l}{dt^l} \Omega(t)$$

which can easily be solved and which we will see can be generalized to the quantum case, at least for the first few orders of differentiation. As with simple Gaussian pulses, the IBP strategy ignores phase shifts (in this case both on and off resonance) but these are irrelevant for state transfer or easily compensated by frequency shifting and/or inserting frame transformations (Sec. 4.2). In the next section we examine how this mostly-classical argument can be generalized and derived for the quantum case to account for these and other higher order errors.

7.1.3 Quantum operator diagonalization

7.1.3.1 Rotating frame

We now provide a quantum mechanical treatment starting from the interaction frame Hamiltonian Eq. (7.1). In order to include the Hamiltonian dynamics it is preferable to have time-dependence coming only from the exponentiation of the matrix, and for this purpose remove the explicit time-dependence from the Hamiltonian by moving to an (approximately) static frame. We choose to work in the rotating frame with the frequency of a given drive (Sec. 4.2.2). To switch from one frame to the other and remove any oscillations in the Hamiltonian of frequency $\Delta_{j,k,l} = \omega_l^d - \omega_{j,k}$ (the detuning with respect to drive l) we apply a transformation $O_{j,k,l} = \exp(-i\Delta_{j,k,l}t|k\rangle\langle k|)$ which removes any exponential factors $\exp(-i\Delta_{j,k,l}t)$ from the off-diagonal terms $|j\rangle\langle k|$. For a set of such combinations of drive and non-zero, off-resonant matrix

elements $\mathcal{O} = \{(j,k,l)\}$, the aggregate transformation $O = \prod_{(j,k,l)\in\mathcal{O}} O_{jkl}$ has the effect of removing all the sinusoidal oscillations in the Hamiltonian. Starting with the interaction frame Hamiltonian now written as

$$\hat{H}_{\text{int}}(t) = \sum_{r,s,t} H_{r,s,t}(\Omega_t(t)e^{-i(\phi_t+\Delta_{r,s,t}t)}, 0) \tag{7.13}$$

$$H_{r,s,t}(\Omega,\Delta) = \Omega\lambda^r_{st}|s\rangle\langle t| + \text{h.c.} + \Delta|t\rangle\langle t| \tag{7.14}$$

and applying $\widetilde{H}(t) = O\hat{H}_{\text{int}}(t)O^\dagger + i\dot{O}O^\dagger$ gives

$$\widetilde{H}(t) = \sum_{r,s,t} H_{r,s,t}\left(\Omega_t(t)e^{-i(\phi_t+\Delta^{\text{dif}}_{r,s,t}t)}, \sum_{(j,k,l)\in\mathcal{O}} \delta_{j,r}\delta_{k,s}\delta_{l,t}\Delta_{j,k,l}(t)\right) \tag{7.15}$$

$$\Delta^{\text{dif}}_{r,s,t} = \Delta_{r,s,t} + \sum_{(j,k,l)\in\mathcal{O}} (\delta_{k,s} - \delta_{k,r})\Delta_{jkl} \tag{7.16}$$

where $\delta_{a,b}$ is the Kronecker delta. Unfortunately, if there are more non-zero matrix elements $\lambda_{j,k,l}$ than there are levels and/or more than one drive per level, not all oscillation in the Hamiltonian can be removed in this way; however, we will typically only be concerned with some or few of these elements. In these cases, the rotating wave Hamiltonian simplifies greatly to

$$\widetilde{H}(t) = \sum_{r,s,t} H_{r,s,t}(\Omega_t(t)e^{-i\phi_t}, \Delta_{r,s,t}(t)) \tag{7.17}$$

Note that we have chosen to make the detuning $\Delta_{r,s,t}(t)$ time-dependent in the last line though this was not done throughout the derivation for clarity. To see how do derive the time-dependence in the simple qubit case see Sec. 4.2.2, which constrains several possible implementations.

7.1.3.2 Block diagonal frame

Now we use the intuition from the Fourier transform analysis to apply a procedure that removes these off-resonant coupling terms. We want to see the effect an off-phase control can have and as such move to an interaction frame with respect to it Starting with the matrix transition element

$$H_{jkl}(e^{-i\phi_l}\Omega_l(t), \Delta_{jkl}) = H_{jkl}(e^{-i\phi_l}\mathrm{Re}\Omega_l(t), \Delta_{jkl}) + H_{jkl}(ie^{-i\phi_l}\mathrm{Im}\Omega_l(t), 0)$$

we apply the interaction transformation

$$D_{jkl} = \exp\left(\int_0^t H_{jkl}(e^{-i\phi_l}\mathrm{Im}\Omega_l(t'), 0)dt'\right) \tag{7.18}$$

Assuming we are under the conditions such that Eq. 7.17 applies (matrix elements are not driven simultaneously at two frequencies), the effective Hamiltonian under the transformation reads

$$\begin{aligned} H^{\mathrm{eff}}_{jkl}(t) &= D_{jkl}H_{jkl}(e^{-i\phi_l}\Omega_l(t), \Delta_{jkl})D^\dagger_{jkl} + i\dot{D}_{jkl}D^\dagger_{jkl} \\ &= D_{jkl}H_{jkl}(e^{-i\phi_l}\mathrm{Re}\Omega_l(t), \Delta_{jkl})D^\dagger_{jkl} + H_{jkl}(ie^{-i\phi_l}\mathrm{Im}\Omega_l(t), 0) + i\dot{D}_{jkl}D^\dagger_{jkl} \\ &= D_{jkl}H_{jkl}(e^{-i\phi_l}\mathrm{Re}\Omega_l(t), \Delta_{jkl})D^\dagger_{jkl} \end{aligned} \tag{7.19}$$

The goal is to achieve a perfect on-resonant excitation and block-diagonalize $\tilde{H}^{\mathrm{eff}}(t)$ as described, i.e.,

$$\begin{aligned} H^{\mathrm{eff}}_{j(l),k(l),l} &= \Omega^{\mathrm{eff}}(t)e^{-i\phi}\lambda^l_{j(l),k(l)}|j(l)\rangle\langle k(l)| + \mathrm{h.c.} \\ H^{\mathrm{eff}}_{j\neq j(l),k\neq k(l),l} &= \Delta^{\mathrm{eff}}_{j,k,l}|k\rangle\langle k| \\ \int_0^T \Omega^{\mathrm{eff}}(t)dt &= \theta \end{aligned} \tag{7.20}$$

The driving field $\Omega^{\mathrm{eff}}(t)$ can be thought of as an adiabatic drive, in the sense that the energy input into the system via driving fields does not change the energy populations of unwanted transitions but only temporarily shifts the level structures before reverting to the original energy splitting (the net effect being only an accumulation of dynamical phase). In effect, we can think of these selectivity criteria as restoring 'adiabaticity' to the system.

In the case where the transitions are disconnected (the unwanted transitions are in separate quantum elements), the condition amounts to the removal of crosstalk. In this case the system of equations can be solved exactly in principle, as will be discussed next. For an infinite ladder of connected transitions it is not possible to find an exact solution and instead an iterative approach will be needed (see Sec. 8).

7.1.3.3 Solution basis of higher derivatives

Now a single Ω_l can cancel exactly one excitation. To completely suppress two transitions using the method for a single constraint requires an additional control (for the additional constraint). Here, we again apply the IBP intuition and use a second derivative component. To see the effect of the second derivative control for a single constraint we move again to an interaction frame with respect to it

$$E_{jkm} = \exp\left(-i\int_0^t H_{jkm}(\mathrm{Re}\Omega_m(t'),0)dt'\right)$$

which gives an effective Hamiltonian according to

$$\begin{aligned}
H^{\mathrm{eff}}_{jkl}(\Omega(t),\Delta_{jkl}) =& E_{jkm}\left(H_{jkl}(\mathrm{Re}\Omega_l(t)+\mathrm{Re}\Omega_m(t),\Delta_{jkl})\right)E^\dagger_{jkm} + i\dot{E}_{jkm}E^\dagger_{jkm}\\
=& E_{jkm}H_{jkl}(\mathrm{Re}\Omega_l(t),\Delta_{jkl})E^\dagger_{jkm}\\
=& H_{jkl}(\mathrm{Re}\Omega_l(t),0) + \Delta_{jkl}E_{jkm}|k\rangle\langle k|E^\dagger_{jkm}
\end{aligned}$$

In this case the transformation does not directly diagonalize the unwanted coupling and instead to do this we should apply another (in this case non-interaction frame) transformation

$$\begin{aligned}
H^{\mathrm{eff}}_{jkl}(\Omega(t),\Delta_{jkl}) =& D_{jkl}E_{jkm}H_{jkl}(\mathrm{Re}\Omega_l(t),\Delta_{jkl})E^\dagger_{jkm}D^\dagger_{jkl} + i\dot{D}_{jkl}D^\dagger_{jkl} \qquad (7.21)\\
=& D_{jkl}H_{jkl}(\mathrm{Re}\Omega_l(t),0)D^\dagger_{jkl} + \Delta_{jkl}D_{jkl}E_{jkm}|k\rangle\langle k|E^\dagger_{jkm}D^\dagger_{jkl} + i\dot{D}_{jkl}D^\dagger_{jkl}
\end{aligned}$$

where the first term is diagonalized by cancelation with the real part of the second term, and the imaginary part of the second term cancels out with the third term. To first order D_{jkl} is the same as the transformation defined in Eq. 7.19.

To combine the effect of multiple (derivative) controls, for a set $\{jk\} \in S_l$ of unwanted, off-resonant transitions connected to drive l, we have the compound formula

$$\begin{aligned}
\tilde{H}^{\mathrm{eff}}_{jkl}(\Omega_l(t),\Delta_{jkl}) =& D_{jkl}E_{jkl}H_{jkl}(\Omega_l(t),\Delta_{jkl})E^\dagger_{jkl}D^\dagger_{jkl} \qquad (7.22)\\
&+ iD_{jkl}\dot{E}_{jkl}H_{jkl}(\Omega_l(t),\Delta_{jkl})E^\dagger_{jkl}D^\dagger_{jkl} + i\dot{D}_{jkl}D^\dagger_{jkl}
\end{aligned}$$

where D_{jkl} corresponds to the exponentiation of an imaginary operator and E_{jkl} to the exponential of a real operator, as defined in the equations above. For more

than two unwanted transitions, the control(s) is (are) a function of the higher order derivatives of the original pulse shape (e.g. a Gaussian),

$$\Omega_l(t) = \sum_{r=0}^{|S_l|} \alpha_r \Omega_G^{(r)} + i \sum_{r=1}^{|S_l|} \beta_r \Omega_G^{(r)}$$

where the superscripts denote the order of differentiation and which together with Eq. 7.20 for each value of j, k give conditions for the coefficients α_r, β_r , which will be a function of Ω_G and Δ_{jkl}. In order for the effective frame to be equivalent to the original bare frame we must choose a pulse shape Ω_G whose $|S_l|$ derivatives are zero at the endpoints of the pulse (see Sec. 4.1.2). Note that although the solution is approximately given simply by semi-classical Eq. 7.12, higher order quantum mechanical corrections are needed to fully solve the system. Some of these will be given in the next sections.

7.1.3.4 Doublet analysis

The simplest scenario are two disconnected transitions: two qubits. We put the drive on resonance with the first qubit ($\Delta_A = \omega_0^d - \omega_{0,1}^A \approx 0$) and letting $\Delta_B = \omega_0^d - \omega_{0,1}^B = \Delta_A + \Delta$, Eq. 7.1 gives

$$\hat{H}_{\text{tot}} = \Omega(t) e^{-i\phi} (\lambda_{01}^A (\hat{\sigma}_{0,1}^+)^A \otimes \mathbb{1}^B + \lambda_{01}^B \mathbb{1}^A \otimes (\hat{\sigma}_{0,1}^+)^B e^{-i\Delta t}) + \text{h.c.} \tag{7.23}$$

In the frame rotating at the drive frequency ($\mathcal{O} = \{(00, 01, 1), (10, 11, 1)\}$), this gives

$$\tilde{H}(t) = H_{0,1,0}(e^{-i\phi}\Omega(t), 0) \otimes \mathbb{1}^B + \mathbb{1}^A \otimes H_{0,1,0}(e^{-i\phi}\Omega(t), \Delta_B)$$

Nulling the off-diagonal components for qubit B to give $\tilde{H}_B^{\text{eff}} = \Delta_B^{\text{eff}} |1\rangle\langle 1|$ can easily be satisfied through the diagonalizing transformation, by plugging Eq. 7.19, which for qubits simplifies to

$$\begin{aligned} D = &\left(\cos\Big(\int_0^t e^{-i\phi} \text{Im}\Omega_l(t) \lambda_{jk}^A dt \Big) \hat{\mathbb{1}} + \sin\Big(\int_0^{\text{t}} \text{e}^{-\text{i}\phi} \text{Im}\Omega_{\text{l}}(\text{t}) \lambda_{\text{jk}}^{\text{A}} \text{dt} \Big) \hat{\sigma}_{0,1}^+ + \text{h.c.} \right) \\ &\otimes \left(\cos\Big(\int_0^t e^{-i\phi} \text{Im}\Omega_l(t) \lambda_{jk}^B dt \Big) \hat{\mathbb{1}} + \sin\Big(\int_0^{\text{t}} \text{e}^{-\text{i}\phi} \text{Im}\Omega_{\text{l}}(\text{t}) \lambda_{\text{jk}}^{\text{B}} \text{dt} \Big) \hat{\sigma}_{0,1}^+ + \text{h.c.} \right), \end{aligned} \tag{7.24}$$

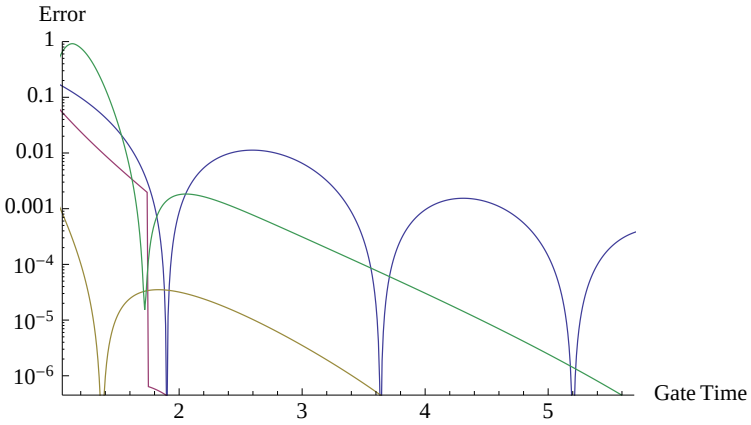

Figure 7.3: Population selection (inversion) error for two uncoupled qubits for energy difference $0.85 \times 2\pi$ (ignoring phase error). The blue line shows the error using standard Gaussian shaping, Eq. 3.3 while the red line uses the pulse shape given by DRAG Eqs. 7.25-7.27 which exactly solves the problem in the adiabatic regime (beyond 1.75). The orange line is using the first order shape given by the second derivative and the green line is first order shape using the fourth derivative.

into Eq. 7.19. Solving Eq. 7.20 for qubit B then gives the solution

$$e^{-i\phi}\lambda^B_{0,1}\int_0^t \mathrm{Im}\Omega(t)dt = \frac{1}{2}\arctan\left(\frac{2e^{-i\phi}\lambda^B_{01}\mathrm{Re}\Omega(\mathrm{t})}{\Delta}\right) \tag{7.25}$$

or equivalently

$$\mathrm{Im}\Omega(t) = \frac{\Delta\mathrm{Re}\dot{\Omega}(\mathrm{t})}{\Delta^2 + 4(\lambda^B_{01}\mathrm{Re}\Omega(\mathrm{t}))^2} \tag{7.26}$$

This is the quantum mechanical equivalent to the IBP formula, Eq. 7.8. In particular the first order Taylor expansion is the same which is often easier to work with.

In the quantum mechanical case, this interaction frame changes the A qubit as well, which must also be compensated for. Luckily, since this is again a qubit this can again be done exactly by now solving Eqs. 7.8 for qubit A:

$$\begin{aligned} DH^A(\Omega(t),\Delta_A)D^\dagger &= \Omega^{\mathrm{eff}}(t)\sigma_x \\ \int_0^T \Omega^{\mathrm{eff}}(t)dt &= \theta. \end{aligned}$$

These equations can readily be satisfied by setting

$$\Delta_A = \frac{1}{2}\left(-\Delta + \sqrt{-16(\Omega(t))^2 + \Delta^2}\right) \tag{7.27}$$
$$\int_0^T \frac{\Omega(t)\Delta}{\Delta_B\sqrt{1+\frac{4\Omega(t)^2}{\Delta_B^2}}} = \theta$$

Fig. 7.3 shows how this error can be fully removed in the adiabatic regime. Δ is chosen to be $0.85 \times 2\pi$ and the selection error is plotted as a function of gate time. The error for a simple Gaussian of correct area is plotted in blue (ignoring again any net phase errors), while in red is the solution given by the technique we call Derivative Removal by Adiabatic Gate (DRAG), which generalizes the procedure we found originally in Ref. [122]. Note that selection error is already suppressed for the Gaussian whenever the Fourier transform at the frequency offset of the second qubit $\int_0^T \Omega(t)e^{-i\Delta_B t}dt$ is nulled (this first happens at about T=2). This is essentially a first order result. For DRAG, the error is always suppressed beyond about $T = 2$ where both qubit A is rotated and B is left intact. The result is exact provided the adiabaticity condition is met, $4\Omega(t) < \Delta_B$ (otherwise the solution is undefined, though we use in Fig. 7.3 the first order in $\frac{4\Omega(t)}{\Delta_B}$ to show an improvement still exists here). In the adiabatic regime, we see that the derivative (imaginary) control suppresses any crosstalk to the second qubit (keeping things adiabatic) while the change of frame essentially shifts the energy levels temporarily during the gate such that the resonance condition needs to be adjusted according to Eq. 7.27. Practically speaking, the DRAG result is clearly more robust than using the single envelope solution as it is immune to changes in frequency or gate time. Note that both solutions contain net phase errors on qubit B that need to be corrected subsequently.

Using the second derivative result, an approximate solution (ignoring a term small in $\frac{\Omega(t)}{\Delta_B}$) can similarly found from Eq. 7.21 with

$$\mathrm{Re}\Omega = \Omega_G + \frac{d}{dt}\frac{\dot{\Omega}_G}{\Delta^2 + 4{\Omega_G}^2}$$

which is plotted in orange in Fig. 7.3 clearly suppressing the Gaussian result and outperforming the first derivative result in the diabatic regime. Note that no change of phase is induced because the main control and auxiliary control are in phase with each other and therefore obey an area theorem (Sec. 2.3).

Using the third derivative result with Eq. 7.22, an approximate solution can be found as

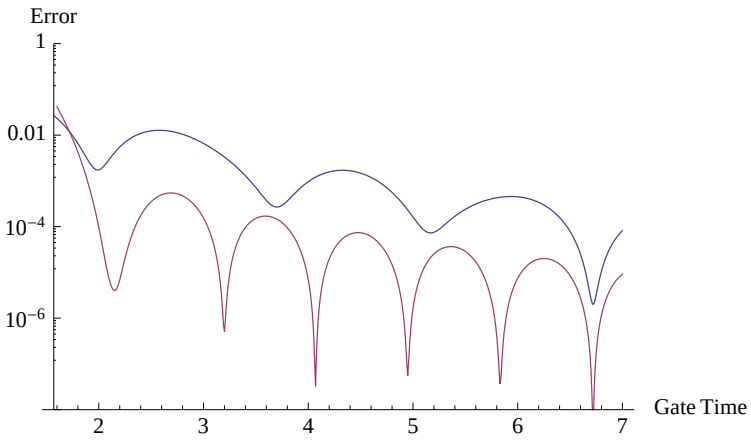

Figure 7.4: Selection error for three uncoupled qubits for energy differences $0.85 \times 2\pi$ and $1.445 \times 2\pi$ from the driven qubit. The blue line shows the error using standard Gaussian shaping, Eq. 3.3 while the red line uses the pulse shape given by DRAG Eqs. 7.25-7.27.

$$\mathrm{Im}\Omega = \frac{\Delta \dddot{\Omega}_G}{(\Delta^2 + 4{\Omega_G}^2)^2}$$

$$\Delta_A = -2\Omega_G \tan\left(2\int_0^t Im\Omega_l(t)\mathrm{dt}\right) \tag{7.28}$$
$$\int_0^T \Omega(t)\sec\left(2\int_0^t Im\Omega_l(t)\mathrm{dt}\right) = \theta$$

which is plotted in green in Fig. 7.3 still outperforming the Gaussian result in the adiabatic (long-time) limit.

7.1.3.5 Multiplet analysis

If more than 2 qubits are in the system, with a single amplitude control it becomes increasingly difficult (at the cost of larger T) to find a gate time where crosstalk is avoided on all other qubits.

For instance, adding in an extra qubit C with $\Delta_C = 1.7\Delta_B$ such that the Hamiltonian (Eq. 7.2) gives in the rotating frame

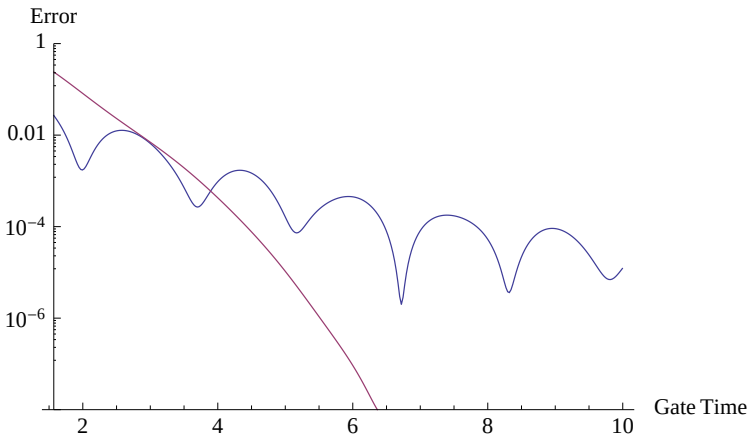

Figure 7.5: Selection error for three uncoupled qubits for energy differences $0.85 \times 2\pi$ and $1.445 \times 2\pi$ from the driven qubit. The blue line shows the error using standard Gaussian shaping, Eq. 3.3 while the red line uses the pulse shape given by DRAG semi-classical Eq. 7.12.

$$\hat{H} = \Omega(t)e^{-i\phi}\left((\hat{\sigma}^+_{0,1})^A + (\hat{\sigma}^+_{0,1})^B + (\hat{\sigma}^+_{0,1})^C + \text{h.c.}\right) \tag{7.29}$$

$$+\Delta_A|1\rangle\langle 1|_A + \Delta_B|1\rangle\langle 1|_B + \Delta_C|1\rangle\langle 1|_C \tag{7.30}$$

and applying the same pulse sequence Eq. 7.25 (which nulls crosstalk at B) is also beneficial to qubit C when it is on the same side of the spectrum from qubit A ($\Delta_C\Delta_B > 0$). In Fig. 7.4, we see that plotting again gate selection error for all three qubits vs. gate time, we get a distinct decrease in errors for all times when applying the DRAG correction. The situation is somewhat reversed from 2 qubits, where now the Gaussian error is somewhat constant but the Gaussian with DRAG has a recurrence of perfect addressability roughly once per unit time, when the Fourier transform of the pulse at C is approximately zero.

Eq. 7.21 is in general difficult to solve exactly but to first order the analysis is the same as using the semi-classical method. In particular, the solution for a single unwanted transition is given still by the second derivative control, Eq. 7.11. For two unwanted transitions $a \leftrightarrow b$ and $c \leftrightarrow d$ it is necessary to solve the pair of conditions given by $H^{\text{eff}}_{abl}(\Omega(t), \Delta_A) = \Delta^{\text{eff}}_A|b\rangle\langle b|$ and $H^{\text{eff}}_{cdl}(\Omega(t), \Delta_C) = \Delta^{\text{eff}}_C|d\rangle\langle d|$. This simplifies to the semiclassical scheme, Eq. 7.12, which for the two unwanted transitions gives

$$\text{Re}\Omega = \Omega_G + \frac{\ddot{\Omega}_G}{\Delta_B\Delta_C} \text{ and } \text{Im}\Omega = \left(\frac{1}{\Delta_B} + \frac{1}{\Delta_C}\right)\dot{\Omega}_G$$

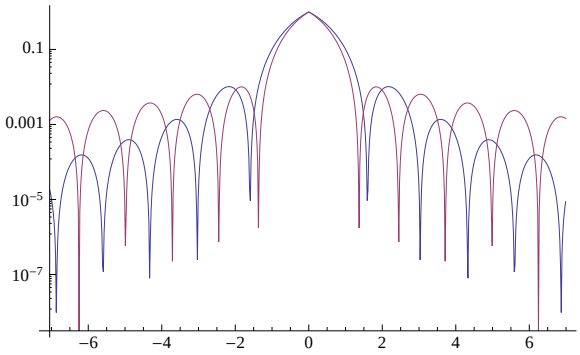

Figure 7.6: Selection error as a function of frequency offset when using the second derivative.

The inversion error for this pulse sequence is plotted in Fig. 7.5 as compared to the simple Gaussian. We see the error is completely suppressed beyond about 5 time steps which is not as early as was the case for using a single derivative. This is because in order to enforce the first derivative starting and ending at zero a slightly less selective basic shape must be used (with wider bandwidth) as a starting point and because we are using the semiclassical result which removes just the first order error which is only dominant for larger times. The result can be improved by considering higher order quantum effects which is the direction of future research. Alternatively, two different derivatives can be also used (e.g. 2nd and 4th) which can remove the commutativity error. Nonetheless, we see the suppression for the semi-classical result is complete beyond the adiabatic limit and the robustness properties are returned. Similar results exist for four or more qubits.

Finally, we confirm the observation that using the second derivative can be used for its even symmetry around the driving transition to decrease the bandwidth of selective pulses. As in the classical case, we see in Fig. 7.6 that appropriate weighting of the second derivative creates holes on either side of the working transition in the excitation profile here (as opposed to the spectrum). Once again, we notice that the energy has been moved from the selective region to the tails where it falls below the chosen selectivity threshold. We see a bandwidth improvement again of about 25%.

7.2 Spatial crosstalk

Alternatively to being suppressed from energy conservation considerations, quantum elements that are spatially separated from the target will be attenuated as a result of

local addressing. An alternative or complementary scheme to avoid crosstalk therefore is to design systems for control lines addressing one section to interact minimally with others. In fact, the spectral crowding limitations developed above imply that some amount of spatial selectivity will be necessary to scale beyond the limitations imposed by the coherence times of the quantum dynamics. Unfortunately, while qubits themselves can be quite small in spatial extent and even engineered to be arbitrarily distanced from each other, the necessity for inter-qubit exchange of information provides a great challenge in terms of maximizing coupling between relevant degrees of freedom while turning off unwanted coupling. In particular, this is exemplified for superconducting qubits, for which both long-range and nearest-neighbour interactions between circuit elements are possible. Moreover, both frequency selective transmission elements and spatially addressable control lines are feasible as building blocks in architectural designs. Thus, there are numerous proposals for designs of superconducting coupling mechanisms [25, 29, 38, 112, 117, 126, 134, 173, 179, 185].

In this section, we look at the motivating example of flux qubits inductively coupled using dc Superconducting QUantum Interference Devices (DC-SQUIDs) used for tunability. The DC-SQUIDs are treated as high excitation energy quantum objects (assumed to always stay in their ground states) where the circulating current mediates inductive coupling between pairs of qubits [5, 75]. Because of the spatial extent of the couplers, there will result crosstalk to other qubits and couplers. This method was first theoretically described in [143], and subsequently demonstrated experimentally in [70]. Similar coupling approaches were presented in [18,69,132,138,170]. This example is a case where the unwanted coupling between qubits can be very strong (in amplitude), and so here we investigate whether it is possible to achieve low error using dynamical decoupling control techniques, with or without the advantage of energy separation between components.

7.2.1 Physical Model

A flux qubit can be constructed from a superconducting loop interrupted by three Josephson junctions [134]. A current flowing in one direction can represent a computational basis quantum state $|0\rangle$ and in the opposite direction, a state $|1\rangle$. The strength of the coupling between qubits is mediated by a circulating current in a DC-SQUID which is inductively coupled to each of the qubits. This circulating current can be controlled by the applied flux as well as the bias current, both of which can be tuned experimentally. Furthermore, the DC-SQUIDs may also be used for qubit readout [108, 135, 171].

The performance of any inductively coupled qubit architecture will strongly de-

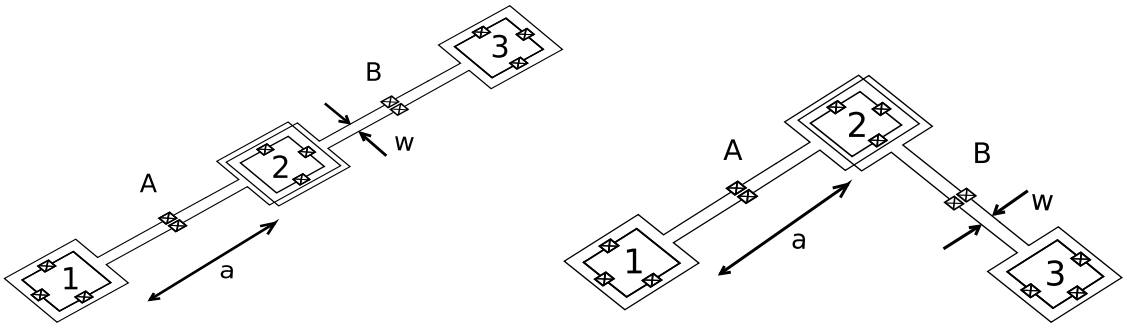

Figure 7.7: Two different arrangements of flux qubits 1, 2, 3 and couplers A, B. "a" represents the arm length of the couplers, and "w" their width. We assume that the wires of couplers A and B overlap at most in two places around qubit 2. a) Straight-line shape. b) Capital "L" shape.

pend on the geometry of the qubit-coupler layout. The couplers themselves are based on the long-distance design proposed in [53]. In this section we investigate two geometries, which can potentially serve as building blocks for topologically motivated approaches to quantum computation [90,146] with high error thresholds [37,52,146,176]. In these approaches the physical qubits are typically laid out in a (for example) 2D square lattice. The geometries we study, which may be used to build such a lattice, are both shown in Fig. 7.7. In a), the coupling DC-SQUIDs form a straight line and in b), they enclose a 90 $^\circ$ angle suggestive of a capital letter "L". In both cases we have two qubits at the end points (labeled "1" and "3") and the third (labeled "2") in the middle. The middle qubit is therefore surrounded by both of the couplers. We consider the qubits at the edge to be nearest-neighbours to the middle qubit. The qubits are placed inside the couplers, hence maximizing the qubit-coupler inductance. Even though the coupling DC-SQUIDs are the same in both geometries, we expect the coupling energies to differ as even for the same length couplers, the distance between qubit 1 and 3 will be smaller in the 'L' shaped scenario.

For concreteness, the following physical parameters are used. The edge length of the square qubits is taken to be $44\mu m$, roughly the same as those used in [70]. The distance between the qubit and the surrounding wire of the coupler is between 2 and 4μm. The arm length a is varied between 25μm and 300μm and the arm width between $w = 48\mu$m and $w = 2\mu$m (although in the latter case a $10 \times 10\mu m^2$ loop in the center of the arm is used to deliver flux through the couplers).

7.2.2 Coupling landscape

The mutual inductance between qubits is composed of the mutual inductance between the currents flowing in the qubits and the product of inductances between each qubit and the coupler. Using Kirchhoff's laws and flux quantization it is possible to link the interaction energy to flux and bias current applied to DC-SQUID, as well the net inductance matrix between elements using finite element simulations using FastHenry [82]. This derivation involves a lengthy calculation which has been performed in Ref. [63].

Fig. 7.8 A shows the coupling strength of the L-shape geometry, where the arm width "w" is 2μm and the arm length "a" is 100μm. The horizontal axis corresponds to the values of coupling energies between the first and second qubit K_{12}, the vertical to the energy between the second and third K_{23} and finally the "out-of-page" direction to the crosstalk term K_{13}. The plot was obtained by varying the fluxes and biasing currents through both couplers. Thousands of data points were calculated, which were then used to obtain a surface that spans different combinations of the three coupling energies. The process was then repeated for varying arm lengths and arm widths of the coupling DC-SQUIDs. Such a map gives a full characterization of a Hamiltonian which we write

$$H(\Phi_A, \Phi_B, I_A, I_B) = \sum_{j=1,2,3} \frac{1}{2}(\epsilon_j \sigma_z + \Delta_j \sigma_x) - \sum_{i \neq j} K_{ij}(\Phi_A, \Phi_B, I_A, I_B)\sigma_z^i \sigma_z^j \quad (7.31)$$

where ϵ_j is the energy bias which depends on the flux threaded through the qubit and Δ_j is the matrix element, which we take to be fixed at fabrication time and usually of the order of a few GHz. The applied flux Φ_n and bias current I_n control completely the coupling between the qubits. From the results, one can note that the crosstalk term K_{13} can be treated as a single-valued function of K_{12} and K_{23}. Hence it is possible to simply work with the possible values of the energies K_{ij} without particular attention to the fluxes and bias currents which will not be single valued.

The first question we wish to address is whether we can keep the coupling between two nearest-neighbour qubits (say 1 and 2) high, the other nearest-neighbour coupling (say 2 and 3) turned off, while eliminating the crosstalk (interaction between qubits 1 and 3). We concentrate on this scenario as we envision doing gates between nearest-neighbors (here qubits 1 and 2), and would like the qubits that are not involved in the gate (qubit 3) to be disturbed as little as possible. In practice, we will find that it is not possible to completely turn off the interactions with the third qubit, and hence all the values of K will need to be varied, but during a gate between qubits 1

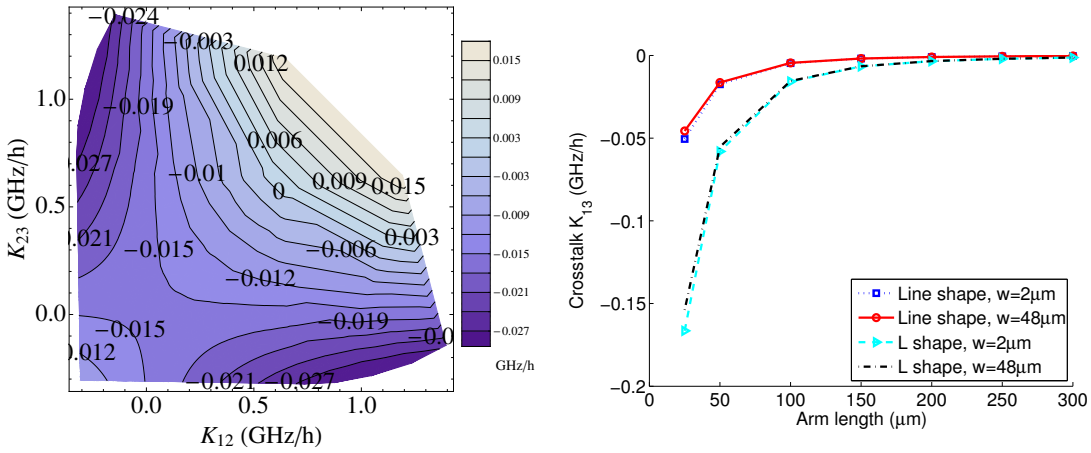

Figure 7.8: (A) Interaction energies K_{12} (horizontal axis), K_{23} (vertical axis) and K_{13} (out-of-page direction) for an L-shape geometry with couplers' arm width of $w = 2\mu$m and arm length $a = 100\mu$m. . (B) A plot consisting of values for the crosstalk K_{13}, while keeping $K_{12} = K_{23} \approx 0$. The horizontal axis represents different arm lengths of the coupling DC-SQUIDs.

and 2 the interaction K_{12} will dominate. We can get the answer to this first question by studying plots like the one presented in Fig. 7.8 A. We first fix K_{23} (vertical axis) at zero, and then traverse in the horizontal direction along the dashed line. In the line-shaped geometry scenario with shortest considered arm length $a = 25\mu$m and K_{23} tuned to approximately 0, we find the crosstalk K_{13} to be 47MHz. The crosstalk only varies slightly as we increase the coupling K_{12}. In the L-shape geometry, the situation is less favorable, as expected, since qubits 1 and 3 are closer together and their direct mutual inductance is larger. Furthermore, they interact with the coupler that is further away from them more than is the case in the line-shaped geometry. The best we can do given short couplers with $a = 25\mu$m is $K_{13} \approx 160$MHz. However, we can reduce the unwanted crosstalk term K_{13} by simply increasing the arm length of the coupling DC-SQUIDs — for example in a case of $a = 300\mu$m, for $K_{12} > 1.0$GHz, the magnitude of K_{13} is smaller than 2MHz in all the geometries. One cannot keep increasing the arm length indefinitely, as the process also decreases the maximum achievable energy that can be mediated between nearest-neighbours qubits resulting in longer gate times and the onset of decoherence. As with most control problems, how much gate error one is willing to tolerate is a function of when the marginal

decrease in error becomes smaller than the increase in incoherent decay of the states. In the next section, we examine how to remove this crosstalk error even when the qubits are closest.

The next question we wish to answer is whether the interaction between all pairs of qubits can be turned off completely. This may be useful when one wants to perform single gate operations on any (or all) of the qubits without affecting the others. To determine this, we once again turn to Fig. 7.8 A (and its analogues for other geometries and values of arm lengths and widths — not explicitly shown). We can now look at the crossing of the two dashed lines, which corresponds to $K_{12} = K_{23} = 0$, and figure out the value of the crosstalk (K_{13}). We plot the results in Fig. 7.8 B for both geometries and two different arm widths. From the data it is clear that there is no complete off-state. The straight-line geometry is much less (over 3 times in a case of shortest arm length we studied) susceptible to the unwanted crosstalk K_{13}. However the crosstalk energy dies off quickly, and in a case of arm length $a > 200\mu$m, its magnitude stays below 5MHz. The error can be very small (negligible when the qubits are off resonant) and is further correctable using standard dynamical decoupling sequences (Sec. 3.4).

7.2.3 Optimization of off-resonant coupling

We look at two gates, the first, the first is the $\sqrt{\text{iSWAP}}$, and the other a CNOT gate. For demonstration purposes, we start with the simplest case, which is a $\sqrt{\text{iSWAP}}$ gate, which is a natural gate for the Hamiltonian. Moreover, we chose the qubit energies to be $\Delta_1/h = 5.0$ GHz, $\Delta_2/h = 5.4$ GHz and $\Delta_3/h = 5.8$ GHz, which provides frequency selectivity.

We start with the system Hamiltonian shown in Eq. 7.31, and we switch basis, to the eigenbasis of σ_x, namely $\{|+\rangle, |-\rangle\}$ and furthermore go into a rotating frame defined by $U = U_1 \otimes U_2 \otimes U_3$ with $U_j = \exp(-i\frac{\Delta_j}{2}\sigma_z^j t)$. We then assume that the qubits can be driven at their respective resonances (meaning $\Delta_j = \omega_j$) and neglect the fast oscillating terms (perform a RWA). All this leads to an effective Hamiltonian in the rotating frame and $\{|+\rangle, |-\rangle\}$ basis

$$\begin{aligned} H_e = &- K_{12}(t)e^{-i(\Delta_1-\Delta_2)t}\left(|100\rangle\langle 010| + |101\rangle\langle 011|\right) + h.c. \\ &- K_{23}(t)e^{-i(\Delta_2-\Delta_3)t}\left(|001\rangle\langle 010| + |110\rangle\langle 101|\right) + h.c. \\ &- K_{13}(t)e^{-i(\Delta_1-\Delta_3)t}\left(|100\rangle\langle 001| + |110\rangle\langle 011|\right) + h.c.. \end{aligned} \tag{7.32}$$

We explicitly make the $K_{ij}(t)$ terms time dependent to stress that they will vary. We apply the GRAPE algorithm (Ch. 5.1) which we constrain by only allowing values

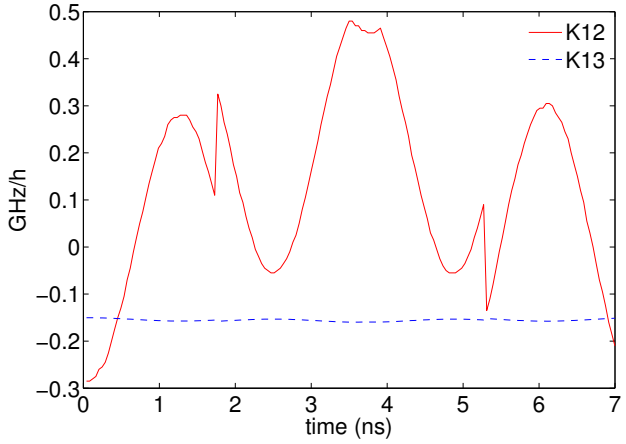

Figure 7.9: A pulse sequence for a $\sqrt{\text{iSWAP}}$ gate between qubits 1 and 2. Since the system coupling is "natural" for this gate, no single-qubit rotations are needed. The pulses lead to the desired gate with 99.9% fidelity. In the calculations we take $\Delta_1/h = 5.0$ GHz, $\Delta_2/h = 5.4$ GHz and $\Delta_3/h = 5.8$ GHz.

of K_{ij} which can be achieved using physical controls (as described in section 7.2.2). We can think of this as staying on a surface spanned by the triplet (K_{12}, K_{23}, K_{13}). From Section 7.2.2, we note that K_{13} can be treated as a single-valued function of K_{12} and K_{23}. This slightly simplifies our procedure of finding the gradient, since we are only dealing with two independent variables that account for all the interactions in the problem. The gradient is not exact since the third variable depends implicitly on the first two but this does not pose a problem in terms of convergence.

The resulting pulse for the $\sqrt{\text{iSWAP}}$ is presented in Fig.. 7.9 which provides a fidelity of 99.9% though this number can be increased with longer optimizations. Only the non-zero controls are explicitly shown. We study the case of the L-shaped geometry with the arm length of 25μm where the unwanted crosstalk coupling K_{13} is the largest (and as can be seen from the plots, never zero). From the plot, it is clear that only coupling 1-2 has to be varied, and coupling 2-3 can be maintained at 0. The shape found by the optimization is reminiscent of the strategy proposed by Ref. [11] which uses a coupling field oscillating at the frequency difference of the qubits to couple the flux qubits and overcome the energy difference between them. However in the analytic case there are many unaccounted errors and the technique is inexact while in the numerical case the pulse contains modulation on top of the carrier which makes it arbitrarily precise. This is an example of a virtual coupling control strategies, more of which will be discussed in Sec. 10.

The unwanted crosstalk coupling K_{13} is significant throughout the gate. However, since we have chosen the qubits to be off-resonance with each other, in the rotating frame the terms K_{23} as well as the crosstalk K_{13} oscillate at frequencies $(\Delta_2 - \Delta_3)$ and $(\Delta_1 - \Delta_3)$ respectively. The combination of small amplitude and high frequency (from the space and frequency separation of the qubits, respectively) effectively eliminates the effect of these terms and it is visible that little correction is needed in the generated controls. Note that operating the qubits at different frequencies does not risk running into spectral crowding limitations because there is only significant coupling between nearest-neighbour and next-to-nearest-neighbour qubits. Thus we see that the combination of space and frequency separation is a powerful way to remove crosstalk and does not pose a problem in terms of scalability.

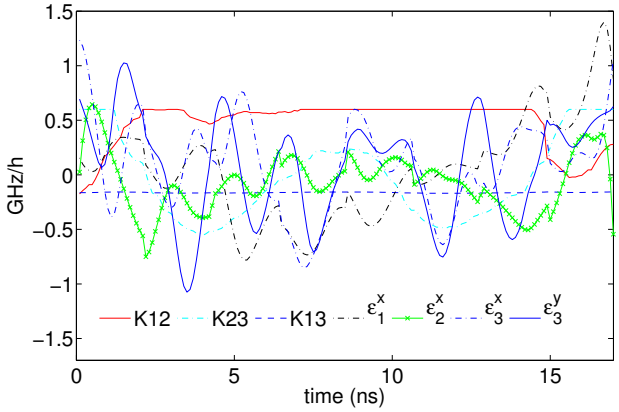

Figure 7.10: A pulse sequence for a CNOT gate between qubits 1 and 2. Single-qubit rotations are applied to all three qubits. For timely convergence of the GRAPE algorithm, we modulate both σ_x and σ_y quadratures on the third qubit. Furthermore, we allow all three interaction energies to vary. The pulses lead to the desired gate with 99.9% fidelity. In the calculations we take $\Delta_1/h = \Delta_2/h = \Delta_3/h = 5.0$ GHz.

.

7.2.4 Optimization of on-resonance coupling

In the second example we introduce two additional complications. In this set-up, we study creating the CNOT gate, which is a "non-natural" gate in that it does not correspond to the evolution of the drift Hamiltonian. However, a CNOT gate is often used to describe the active error correction procedures in the topologically motivated schemes from e.g. Refs. [90, 146, 51] for which this qubit geometry is

a natural contender. Thus, it is useful to optimize this gate as it will be used repeatedly and overall algorithm performance can benefit from making this gate as short as possible. As a second (though perhaps unjustified) complication, we take all qubits to be on resonance — meaning $\Delta_1/h = \Delta_2/h = \Delta_3/h = 5.0$ GHz, removing the convenient energy penalty on the crosstalk terms. The added complexity can only be mitigated by the use of extra controls, in particular single-qubit controls are needed for performing dynamical decoupling. Note that each ϵ_j is time dependent and can be written as

$$\epsilon_j(t) = 4\epsilon_j^x(t)\cos(\omega_j t) + 4\epsilon_j^y(t)\sin(\omega_j t). \tag{7.33}$$

Switching basis to the eigenbasis of σ_x, going into the rotating frame, and performing the RWA, we obtain

$$\begin{aligned} H_e = &\sum_j (\epsilon_j^x(t)\sigma_x^j - \epsilon_j^y(t)\sigma_y^j) - K_{12}(t)\left(|100\rangle\langle 010| + |101\rangle\langle 011| + \text{h.c.}\right) \\ &- K_{23}(t)\left(|001\rangle\langle 010| + |110\rangle\langle 101|\right) - K_{13}(t)\left(|100\rangle\langle 001| + |110\rangle\langle 011|\right) + \text{h.c.}. \end{aligned} \tag{7.34}$$

Fig. 7.10 shows a particular pulse found after optimization. In contrast to the off-resonant case, to achieve high fidelities, qubit 3 must be repeatedly flipped and rotated (via ϵ_3^x and ϵ_3^y) to counteract the crosstalk term, corresponding to a sophisticated dynamical decoupling sequence (Sec. 3.4). Since qubits 1 and 2 are on resonance, the interaction naturally couples them. In turn, the single qubit controls on qubit 1 and 2 (ϵ_1^x and ϵ_2^x) act to convert the flip-flop interaction (naturally drifting into an ISWAP) into a CNOT transition [60, 189]. Of course, rather than applying the single qubit gates before and after the two-qubit operation as is the standard conversion [129], GRAPE is able to find much shorter pulses that are able to perform these rotations simultaneously. Once again, the pulses found perform with a fidelity of 99.9%, a number which can be improved with further optimization. The pulses in general are not unique and can depend on the initial guesses which GRAPE then modifies. Finally, let us note that we have shown how one has to vary the energies in the system to obtain the desired pulses. Alternatively one could concentrate on working with values of physical parameters (currents, fluxes) which might be more useful when trying to devise an experimental demonstration of this system.

Chapter 8

Leakage

8.1 Introduction

Until now we have assumed that quantum elements behave independently and errors come from limitations of the control hardware. In reality the Hilbert space structure of quantum systems is far more complex and the next three chapters analyse the principle ways in which we want to / have to do away with the qubit assumption. The next logical approximation is a three-level system, the qutrit, which we consider in this chapter. We assume that the coupling to the third level is driven by a field, rather than being always-on, primarily because always-on couplings can be diagonalized to give back 2-level systems in a different eigenbasis (of course, in other contexts always-on couplings can still be relevant, see Sec. 2.3.4). In fact, it has been shown that no elementary particle can be a true qubit [182]. In many cases, the coupling to a third level is weak or spectrally distinct from the qubit frequency and contribute error only for very short gates. On the other hand, an important example where the transitions are only slightly spectrally distinct is the weakly anharmonic oscillator.

It is known [49, 74] that the time evolution operator of a linearly driven harmonic oscillator is a combination of a coherent displacement operator tracking the classical trajectory of the driven oscillator and a global phase factor. This evolution encompasses all energy levels and cannot be reduced to a single-qubit rotation (between two of these levels). Spectroscopically, this can be understood as follows: a single qubit rotation is typically implemented by a pulse of radiation resonant with the qubit energy splitting. In a harmonic oscillator, all energy splittings are the same, so driving one transition drives all others at the same time. A system starting initially in an energy eigenstate will quickly be driven into a superposition over many energy

eigenstates. By this token, it is crucial that a qubit is nonlinear [30,86,113,153,173], that is that the transition frequency of the qubit levels is different by an amount Δ from the transition frequencies to the non-qubit levels (the anharmonicity). Spectroscopically, we would expect that whenever the bandwidth of the pulse comes close to Δ, i.e., when its duration becomes short on the scale of $1/\Delta$, we expect significant leakage to higher states. Thus, it is a challenge to implement fast single-qubit gates in weakly anharmonic systems. The implementation of faster gates is important as it allows more gates to be executed in a given coherence time, an important step toward high-fidelity quantum logic.

While superconducting qubits are the most well known example of qubits made from weakly anharmonic oscillators, there are many other examples. Examples of leakage states include higher vibrational states in optical lattices [114], polarized spin states in the singlet-triplet qubits [139], and auxiliary states in ion traps [65] and Rydberg atoms [56,68,169]. Within the framework of superconducting qubits, several relevant current realizations include low-dispersion qubits like the transmon [95,154] and the capacitively shunted flux qubit [165]. On the other hand, a promising route to success are qubits that contains only a minimal number of elements, such as the phase qubit [117, 116, 164].What these systems have in common is a weakly anharmonic energy level structure, i.e., the states that are outside of the qubit subspace spanned by $|0\rangle$ and $|1\rangle$ are only separated from each other and the qubit subspace by energies only slightly different than the qubit frequency.

In this chapter, we outline quantum control strategies to implement single-qubit quantum gates in qubits singled out from the spectrum of an anharmonic oscillator. The method of performing this approximation is summarized in Secs. 8.2 and 8.3. We develop an adiabatic expansion technique that leads to order-by-order constraint equations on a frame transformation and the control fields. This is done in an interaction frame formalism in Sec. 8.5 (published in Ref. [122]) and again using a generalized class of transformations in Sec. 8.6 (published in Ref. [59]). The latter derivations were aided by Seth Merkel and Jay Gambetta. Numerical optimization is performed in Sec. 8.7 to further characterize the controllability of the system. Furthermore, we consider in Sec. 8.8 the influence of relaxation and dispersion in the control of these systems. The model for dispersion permits a deeper analysis of the control problem with further optimization via ensemble techniques which was carried out by Botan Khani and published in Ref. [88,87]. Finally, the techniques developed in this chapter have gained traction in the experimental community, most notably we were able to devise with Jerry Chow et al. at Yale (published in Ref. [28]) ways to benchmark the technique in their laboratory, with results reproduced in Sec. 8.9.

8.2 Mathieu equation

Generically, we can think of anharmonic systems as modeled by a particle in a periodic potential. The Hamiltonian for such a system reads

$$H_0 = \frac{p^2}{2m} + U_0 \sin^2(k_L x), \tag{8.1}$$

where m is the mass of the atom, k_L wave number of the lattice and U_0 is the lattice potential depth. For a given value of p this is exactly solved by the special function, known as the Mathieu function. It is available in mathematical software packages such as Mathematica, but can also be solved using the Bloch theorem and central matrix theorem [4, 87]. For a single well in the potential, the eigenstate corresponds instead to a Wannier state and equals the weighted sum of several of the delocalized Bloch states [96]. For a finite lattice with fixed number of lattice sites and bound momentum states, the eigenvalues are hybrids between the two descriptions.

This Hamiltonian precisely describes the vibrational energy of an alkali atom in a 1D optical lattice, where the periodic potential is given by the standing wave generated by two counter-propagating laser fields, as described for instance in Refs [83, 114, 115].

Another example of a system well described by the lattice potential is the superconducting Cooper pairs on a charge island, given by

$$\hat{H}^{(n_g)} = 4E_C(\hat{n} - n_g)^2 - E_J \cos\hat{\phi} \tag{8.2}$$

where $\hat{n}$ is the number of Cooper pairs on the superconducting islands, $\hat{\phi}$ is the phase difference to the island, E_C and E_J are the charging and Josephson energy, respectively and n_g is the effective gate charge. In addition, if one adds a (constant force) term linear in the space coordinate, giving in dimensionless units

$$H_0 = p^2 + r\cos 2x + gx, \tag{8.3}$$

then one obtains what is known as a tilted washboard potential, which describes both the influence of gravity on a particle with mass, or a current source in a superconducting circuit (a phase qubit). Note also that the physical quantities p, x, and r can contain noise, which as we have seen can either be slow or fast, and originate from inhomogeneity of an ensemble, fluctuations, or uncertainties. We will examine pulses and design choices that render uncertainties in these parameters less significant.

8.3 Qutrit approximation

We consider a qutrit subspace formed by the three lowest levels of the anharmonic oscillator. These levels are separated in energy by $\hbar\omega$, where ω is the transition frequency. The j^{th} higher levels are different to $\hbar j\omega$ by $\hbar\Delta_j$, where Δ_j is known as the anharmonicity. That is, the Hamiltonian for the nonlinear oscillator of dimension d is ($\hbar = 1$)

$$H_{\text{fr}} = \sum_{j=1}^{d-1} (j\omega + \Delta_j)\Pi_j, \tag{8.4}$$

where $\Pi_j = |j\rangle\langle j|$ is the projection operator onto the j^{th} energy level. Without loss of generality we set $\Delta_1 = 0$. This is illustrated in Fig. 3.2 with the qubit levels being $|0\rangle$ and $|1\rangle$ (green). For many nonlinear oscillators the anharmonicity takes the form $\Delta_j = \Delta_2(j-1)j/2$, which we will call the standard nonlinear oscillator (SNO), essentially a Duffing oscillator within the rotating wave approximation [86]. For the lowest few levels, superconducting qubits of the transmon [154], phase qubit [116], and capacitively shunted flux qubit [165] types are well approximated by a SNO. Furthermore, motional states in optical lattices [114], collective modes of ion traps [65] and nano-mechanical oscillators [97] are also described as SNOs.

We will assume that control in this system is due to some dipole-like interaction that only allows single photon transitions. As for harmonic oscillators, this is a good approximation because parity forbids all other transitions. The control Hamiltonian is

$$H_{\text{ct}}(t) = \mathcal{E}(t) \sum_{j=1}^{d-1} \lambda_{j-1}\sigma^x_{j-1,j}, \tag{8.5}$$

where $\mathcal{E}(t)$ is the drive amplitude, $\sigma^x_{j,k} = |j\rangle\langle k| + |k\rangle\langle j|$ is one of the effective Pauli spin operators for levels j and k, and λ_j is a dimensionless parameter that weighs the relative strength of driving the $|j\rangle \to |j+1\rangle$ transition versus the $|0\rangle \to |1\rangle$ transition. In our model we take $\lambda_0 = 1$ and leave the λ_js as input parameters. For a harmonic oscillator controlled via a dipole interaction with an external field $\lambda_j = \sqrt{j}$.

For the functional form of the drive $\mathcal{E}(t)$ we will assume that $|\mathcal{E}(t)| \ll \omega$ (weak driving regime) and introduce envelope shaping of the driving field at carrier frequency ω_d. This leads to

$$\mathcal{E}(t) = \Omega_x(t)\cos(\omega_d t + \phi_0) + \Omega_y(t)\sin(\omega_d t + \phi_0). \tag{8.6}$$

For quantum information processing it is well suited to define operations with respect to the frame rotating at the driving frequency ω_d (Sec. 4.2.2). In this frame

we have three independent controls: $\delta(t) = \omega(t) - \omega_d$ (the qubit detuning), $\Omega_x(t)$ and $\Omega_y(t)$, which, projected to the qubit subspace, control application of the three Pauli spin operators $\sigma^z_{0,1}$, $\sigma^x_{0,1}$, and $\sigma^y_{0,1}$, respectively. For example, the identity operation is achieved by setting $\delta(t) = \Omega_x(t) = \Omega_y(t) = 0$. To move to the rotating frame we define the unitary

$$R(t) = \sum_{j=1}^{d-1} \exp\left[-ij\omega_d t\right] \Pi_j, \tag{8.7}$$

which determines the transformed Hamiltonian $H^R(t) = R^\dagger(t)H(t)R(t) + i\dot{R}^\dagger(t)R(t)$ to be explicitly

$$H^R(t) = \sum_{j=1}^{d-1}(j\delta + \Delta_j)\Pi_j + \left[\frac{\Omega(t)}{2}e^{-i\omega_d t - i\phi_0} + \text{c.c.}\right] \times \sum_{j=1}^{d-1} \lambda_{j-1}\left[|j\rangle\langle j-1|e^{i\omega_d t} + \text{h.c.}\right], \tag{8.8}$$

where c.c. stands for complex conjugate and $\Omega(t) = \Omega_x(t) + i\Omega_y(t)$. Assuming that ω_d is larger then any other rate or frequency in this frame we can perform the rotating wave approximation (i.e. time average the fast rotating terms to zero). For the SNO case this amounts to restricting the dimension d to be less then $\sqrt{2\omega/\Delta}$, specifically $d = 7$. After this approximation we can write the Hamiltonian as

$$H^R(t) = \sum_{j=1}^{d-1}(j\delta(t) + \Delta_j)\Pi_j + \sum_{j=1}^{d-1} \lambda_{j-1}\left[\frac{\Omega_x(t)}{2}\sigma^x_{j-1,j} + \frac{\Omega_y(t)}{2}\sigma^y_{j-1,j}\right], \tag{8.9}$$

where $\sigma^y_{j,k} = -i|j\rangle\langle k| + i|k\rangle\langle j|$ for $k > j$. Here we have included the ϕ_0 into the energy eigenstates ($|j\rangle \to e^{ij\phi_0}|j\rangle$), and we see that within the rotating wave approximation the relative phase between the envelope and the carrier at the start of the operation is irrelevant. We see that if we can restrict the system to the lowest two levels then all rotations in the single qubit space can be achieved by independent controls; however, in general, this is not true. In Sec. 8.4 we show that the higher level transitions lead to a combination of a rotating, phase and leakage error [2, 166, 122]. This has been experimentally measured in Refs. [27] and [106].

8.3.1 Coupling strength from dressing

To obtain explicitly the value of the couplings λ_j beyond the harmonic oscillator approximation we should look at the coupling mechanism. In cases where the system is a SNO and is driven directly, λ_j is well approximated by the harmonic oscillator

matrix elements, namely $\lambda_j \approx \sqrt{j}$. This, for example, occurs for the phase qubit when it is driven by a time-varying bias current [117] and the charge qubit when it is driven by time-varying gate voltage [100]. Essentially, these systems are very nearly harmonic oscillators and the controls are almost proportional to the quadrature operator. On the other hand, when the coupling is done via an intermediary quantum element such as a cavity then one must look at the Hilbert space of the combined system. This is the case for the cavity [68,119] or circuit [15,175] QED architecture.

In the cavity or circuit QED architecture, the anharmonic oscillator is coupled to a resonator and is controlled by driving the resonator far off-resonance. In this case, λ_j can take on essentially any value. To see this, we start by writing the full Hamiltonian for the multi-level anharmonic oscillator and resonator as

$$H_{\rm JC} = \omega_r a^\dagger a + \sum_{j=1}^{d-1}(j\omega + \Delta_j)\Pi_j + \sum_{j=1}^{d-1} g_{j-1,j}\left(|j-1\rangle\langle j|a^\dagger + |j\rangle\langle j-1|a\right), \qquad (8.10)$$

again $\Delta_0 = \Delta_1 = 0$. This is a generalized Jaynes-Cummings Hamiltonian [78] with ω_r being the resonator frequency, and $g_{j,k}$ being the vacuum Rabi coupling for the j to k levels. Again we have assumed that the anharmonic oscillator only allows one-photon transitions.

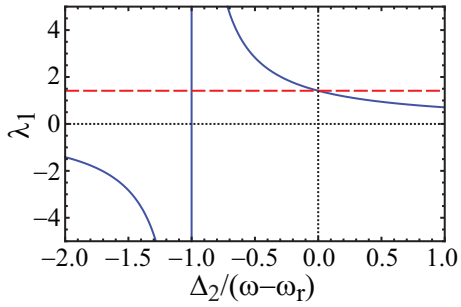

Figure 8.1: Ratio of coupling strength of the $|1\rangle \leftrightarrow |2\rangle$ transition to the $|0\rangle \leftrightarrow |1\rangle$ transition as a function of anharmonicity for control through a resonator (solid blue) and direct drive (dashed red).

If $|\omega'_{j-1,j} - \omega_r| \gg |g_{j-1,j}|$ for all j where $\omega'_{j-1,j} = \omega + \Delta_j - \Delta_{j-1}$ then diagonalization of Eq. (8.10) can be performed to lowest order in $g_{j-1,j}/(\tilde{\omega}_{j-1,j} - \omega_r)$ by the canonical transformation $H_{\rm JC}^{\rm D} = D^\dagger H_{\rm JC} D$ where

$$D = \exp\left[\sum_{j=1}^{d-1} \frac{g_{j-1,j}}{\omega'_{j-1,j} - \omega_r}(a^\dagger|j-1\rangle\langle j| - a|j\rangle\langle j-1|)\right] \qquad (8.11)$$

to give

$$H_{\rm JC}^{D} = (\omega_r - \chi_{01})a^\dagger a + \sum_{j=1}^{d-1} (j\omega + \Delta_j + \chi_{j-1,j})\,\Pi_j + \sum_{j=1}^{d-1} (\chi_{0,1} + \chi_{j-1,j} - \chi_{j,j+1}) a^\dagger a \Pi_j. \tag{8.12}$$

Here $\chi_{j-1,j} = g_{j-1,j}^2/(\omega'_{j-1,j} - \omega_r)$ is the Lamb shift induced on the anharmonic oscillator by the resonator and the last term in Eq. (8.12) is the ac-Stark shift [58]. Assuming that the dressed cavity is in vacuum then Eq. (8.12) is well approximated by Eq. (8.4) with

$$\omega \rightarrow \tilde{\omega} = \omega + \chi_{0,1} \tag{8.13}$$

$$\Delta_j \rightarrow \tilde{\Delta}_j = \Delta_j + \chi_{j-1,j} - j\chi_{0,1} \tag{8.14}$$

and the tilde implies dressed values for the transition frequency and anharmonicity from the resonator.

As stated above, the control is usually through the resonator and is represented by the Hamiltonian

$$H_{\rm dr}(t) = \varepsilon(t)(a + a^\dagger) \tag{8.15}$$

which under the transformation Eq. 8.11 becomes

$$H_{\rm dr}^{D}(t) = \varepsilon(t)\left(a + a^\dagger + \sum_{j=1}^{d-1} \frac{g_{j-1,j}}{\omega'_{j-1,j} - \omega_r} \sigma_{j-1,j}^x\right). \tag{8.16}$$

Assuming that $\varepsilon(t)$ is a sinusoidal with a frequency close to the qubit then Eq. 8.16 is well approximated by Eq. (8.5) with

$$\mathcal{E}(t) = \frac{g_{0,1}}{\omega'_{0,1} - \omega_r}\varepsilon(t) \tag{8.17}$$

$$\lambda_{j-1} = \frac{g_{j-1,j}(\omega'_{0,1} - \omega_r)}{g_{0,1}(\omega'_{j-1,j} - \omega_r)}. \tag{8.18}$$

To demonstrate the functional form of λ_1, we will assume the anharmonic oscillator is in the SNO limit: $\omega'_{j-1,j} = \omega + (j-1)\Delta_2$, and $g_{j-1,j} = \sqrt{j}g_{0,1}$. In this case, Eq. (8.18) becomes

$$\lambda_{j-1} = \frac{\sqrt{j}}{1 + (j-1)\Delta_2/(\omega - \omega_r)}, \tag{8.19}$$

and is plotted in In Fig. 8.1 as a function of the bare anharmonicity in units of $(\omega - \omega_r)$. Here, it is clearly seen that λ_1 is of not equal to $\sqrt{2}$ and actually changes

sign at the point ($\omega - \omega_r = -\Delta_2$). This point invalidates the diagonalization for the second level and can not be treated under this model. Away from this point the values obtained from this model approximate the real situation. This was confirmed in Ref. [28] where the experimental value for λ_1 for the operation point used was found to agree with Eq. 8.18. However, we propose that λ_j should be used as a fitting parameter in any experiment as effects such as higher modes of the resonator and higher order perturbation will result in additional corrections to this value [16].

8.4 Conventional control

Our goal is to implement gates contained within the qubit subspace. That is, we want to shape $\Omega_x(t)$, $\Omega_y(t)$ and $\delta(t)$ in Eq. 8.9 so that according to the selectivity conditions 7.20 we have

$$U_{\text{ideal}} = \mathcal{T} \exp\left[-i \int_0^{t_g} H^R(t) dt\right] = e^{i\phi} U_{\text{qb}} \oplus U_{\text{rest}}, \tag{8.20}$$

where t_g is the gate time, $\mathcal{T}$ is the time ordering operator, U_{qb} is a unitary that acts only in the qubit subspace, U_{rest} acts only outside of the qubit space, and ϕ describes a relative phase. Therefore, U_{rest} as well as the phase ϕ are completely irrelevant for operations in the Hilbert space formed by the qubit.

To demonstrate the typical set of errors we choose $U_{\text{qb}} = \sigma^x_{0,1}$, the NOT gate. For a leakage-free qubit this would be implemented by simply setting $\delta(t) = 0$ and $\Omega_y(t) = 0$, with the only requirement on Ω_x that that $\int_0^{t_g} \Omega_x(t) dt = \pi$. To reduce the leakage to the third level, the standard result prior to Ref. [122] was to use Gaussian modulation of the envelope [7, 166], Eq. 3.3. The motivation for Gaussian shaping is that the small and strictly limited frequency bandwidth ($1/\sigma$) ensures little excitation at the leakage transition frequency. For short pulses however, there is still significant spectral weight at the anharmonicity Δ_2. This is shown in Fig. 8.2 (a, c, e) where the Fourier transforms of $\Omega_G(t)$ are plotted for $\sigma = \{1/3, 2/3, 3/2\} 2\pi/\Delta_2$ and $t_g = 4\sigma$ respectively.

To demonstrate the errors arising from Gaussian shaping we consider a $d = 5$ SNO and numerically calculate the average gate error, Eq. 2.9, for $\sigma = \{1/3, 2/3, 3/2\} 2\pi/\Delta_2$ and $t_g = 4\sigma$. We find gate errors of 0.198, 0.0160, and 0.0030 respectively.

To understand these error values we plot the populations of the first three levels in Fig. 8.2 (b, d, f). The ground state populations are given by the blue solid line; the red dotted line shows the first excited state; the green dashed line is the second excited state. We observe that for the shortest gate, $\sigma = 2\pi/3\Delta_2$, the error after the pulse is mostly residual population of the third and higher level. This is what we refer

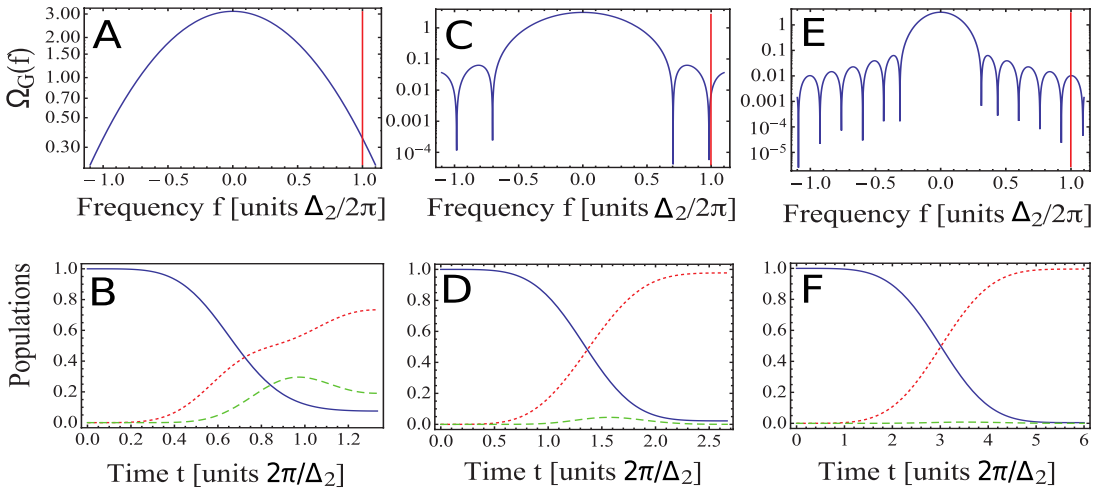

Figure 8.2: Fourier transforms of the control fields (a, c, e) and populations (b, d, f) of the ground (blue solid), first (red dotted), and second excited (green dash) in a simulation of a NOT gate with a Gaussian amplitude pulse for a $d = 5$ SNO. In (a, b) $\sigma = 2\pi/3\Delta_2$, (c, d) $\sigma = 4\pi/3\Delta_2$, and (e, f) $\sigma = 3\pi/\Delta_2$ and the gate time is taken to be 4σ.

to as the leakage error. For longer gates, e.g. $\sigma = 4\pi/3\Delta_2$, the residual population does not account for the calculated error. This error is mostly a combination of phase and rotation error resulting from the finite population of the third level *during* the pulse. Even though the final state is restricted to the computational levels, the admixture of the third level leads to a phase shift on the second level, resulting in a net phase error at the end of the pulse. At the longest time when the population of the third level is nearly negligible there is still a large gate error. From these results we conclude that Gaussian shaping is of limited performance even if the pulse bandwidth is somewhat smaller than Δ_2.

8.5 Adiabatic diagonalization

Having established the importance of spectral considerations in studying the qutrit problem, we are ready to apply the DRAG methodology from Sec. 7.1 to this family of systems. We have seen in Sec. 7.1.3 that the insight from the Fourier transform and the solution given by the integration by parts provides a motivation for using a second control phase shifted from the first, and that the solution will resemble a derivative. In the case of crosstalk, the set of transitions driven by the field formed

a disjoint set, in the sense that no transitions were connected (i.e. qubits). This allowed a diagonalization transformation that did not introduce any new transitions into the operator space. In effect, this means that if there were M transitions whose evolution needed to be fixed, then a basis of M functions would be sufficient. That is only M functions were needed satisfy Eq. 7.20. In contrast, for the set of unwanted transitions S, if these transitions are not simple qubits but the driven transitions are connected to other transitions, then a transformation would remove the unwanted transitions will also introduce new transitions that were previously undriven. In other words, a basis of M fields would not be enough. The possibility for an exact solution is even more difficult for an infinite ladder where every time the transformation were amended to remove the additional transition, yet a new transition would be added in its place, the process continuing ad infinitum. In effect one would need an infinite number of basis fields (e.g. derivatives to the Gaussian) in order to exactly null out all the transitions.

Given the difficulty in defining such an exact transformation, it is wiser to cut one's losses and define an iterative transformation instead. The transformation must be adiabatic in the sense that all values we work with are expanded in a small parameter $\epsilon = \frac{\Omega_l(t)\lambda_{j,k,l}}{\Delta_{j,k,l}}$. For a SNO this is a strong limitation in the sense that for every transition up the ladder, the energy splitting Δ_k decreases and the matrix elements λ_k increase. Nonetheless, an expansion will be seen to be at least somewhat fruitful.

We define the transformation recursively via the iterator h such that

$$\tilde{H}^{D(h)}(t) = \hat{D}_h(t)\tilde{H}^{D(h-1)}(t)\hat{D}_h^\dagger(t) + i\dot{\hat{D}}_h(t)\hat{D}_h^\dagger(t) \tag{8.21}$$

with $\tilde{H}_0^D = \hat{H}^O$ and $\tilde{H}_h{}^D(t) = O(\epsilon^h)$. The condition for selectivity, Eq. 7.20, is weakened to

$$\begin{aligned} \tilde{H}^{D(h)}_{j(l),k(l),l} &= \Omega^{\text{eff(h)}}(t)e^{-i\phi}\hat{\Gamma}_{j(l),k(l),l}(0) + h.c. + O(\epsilon^h) \\ \tilde{H}^{D(h)}_{j\neq j(l),k\neq k(l),l} &= \Delta^{\text{eff(h)}}_{j,k,l}|k\rangle\langle k| + O(\epsilon^h) \\ \int_0^T \Omega^{\text{eff(h)}} dt &= \theta \end{aligned} \tag{8.22}$$

where effective operators are also expanded in the small parameter: $\Omega^{\text{eff(h)}} = \sum_{g=0}^h \Omega^{\text{eff}}_g(t)\epsilon^g$ and $\Delta^{\text{eff(h)}}_{jkl} = \sum_{g=0}^h (\Delta_{jkl})_g(t)\epsilon^g$ (with $\Delta^{\text{eff(h)}}_{j(l),k(l),l} = 0$). Transformations that perform this task are not fully constrained and a general solution is explored in Sec. 8.6 for SNOs. For the case of a generalized Hilbert space the solution

space is even larger. Here, we look only at the transformations in the form of the transformation operators that exactly solved Sec. 7.1.3 but now expanded in the small parameter. That is, we constrain the transformation operator to be in the form

$$D_h = \exp\left(-i\sum_{jk}\int_0^T \tilde{H}_{jkl}(\Omega_h\epsilon^h, 0)\mathrm{dt} + \mathrm{h.c.}\right)$$

Unlike in the case of the qubits, Sec. 7.1.3, we cannot ensure that the\is transformation corresponds to an interaction frame with respect to one of our controls. That is, if the $j-k$ transition corresponds to a control then we can cancel the $+i\dot{\hat{D}}_h(t)\hat{D}_h^\dagger(t)$ term. However if this cancelation is not possible then the term must be left in. Unfortunately, as made clear in Sec. 7.1.1 via the use of integration by parts, the derivative will have the same spectral weight as the original matrix element and not much will be gained by moving to a higher order (only commutator errors will be decreased). Both these situation will come up when studying the SNO.

Returning to the qutrit, we have derived

$$H^R(t) = \sum_{j=1}^{d-1}(j\delta(t)+\Delta_j)\Pi_j + \sum_{j=1}^{d-1}\lambda_{j-1}\left[\frac{\Omega_x(t)}{2}\sigma_{j-1,j}^x + \frac{\Omega_y(t)}{2}\sigma_{j-1,j}^y\right], \tag{8.23}$$

with d set for simplicity to 3 Going to an interaction frame with respect to the out-of-phase control (here $\Omega_y(t) = -\frac{\dot{\Omega}_x(t)}{\Delta}$) with the transformation

$$D_1 = \exp\left(-i\left(\sum_{j=1}^{d-1}\lambda_{j-1}\int_0^t \frac{\Omega_y(t)}{2}dt\right)\sigma_{j-1,j}^y\right)$$

gives the interaction Hamiltonian

$$\begin{aligned} H^{D(1)} = &\left(\Omega_x(t) + \frac{(\lambda^2-2)(\Omega_x(t))^3}{\Delta^2}\right)\sigma_{0,1}^x + \frac{\lambda\Omega_x^2(t)}{2\Delta}\sigma_{0,2}^x + \left(\Delta + 2\delta(t) + \frac{(\lambda^2+2)\Omega_x^2(t)}{\Delta}\right)\Pi_2 \\ &+ \left(\delta(t) - \frac{(\lambda^2-4)\Omega_x^2(t)}{\Delta}\right)\Pi_1 + O(\epsilon^3). \end{aligned} \tag{8.24}$$

In this frame, it is easy to see there are three qubit errors associated with the result. The selection error is corrected with the off-phase derivative control

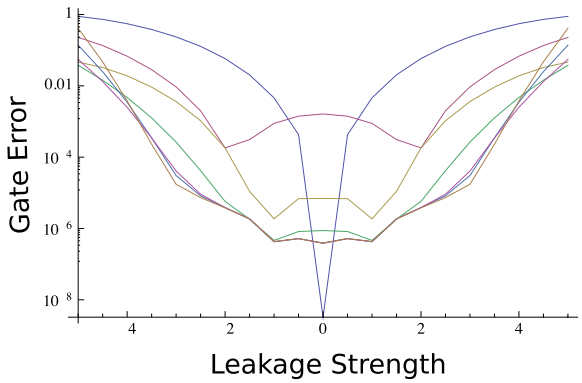

Figure 8.3: For a 3-level system driven by a 7ns pi-pulse, gate error is plotted vs. leakage transition strength λ for solutions to the adiabatic expansion to different orders $H^{D(h)}$ correcting for errors to order h. The top blue line is the simple Gaussian and each line below it (at λ=2) corrects errors to the next order in ϵ. That is the red line under it is correct to order ϵ^2, the yellow under it to $O(\epsilon^4)$, the green under that to $O(\epsilon^5)$, the blue under that to $O(\epsilon^6)$, the red under that to $O(\epsilon^7)$, and the yellow under that to $O(\epsilon^8)$.

$\Omega_y(t) = -\frac{\dot{\Omega}_x(t)}{\Delta}$ which again is exactly the semiclassical result. The phase error can be corrected with $\delta(t) = \frac{(\lambda^2-4)\mathcal{E}_x^2}{\Delta}$ by either shifting the eigen-energies of the system or by a combination of phase ramping and frame compensation (Sec. 4.2). Finally the rotation error is significant and can be corrected by ensuring the area law $\left(\int_0^T(\Omega_x(t) + \frac{(\lambda^2-2)(\Omega_x(t))^3}{\Delta^2}\right) dt = \theta$. To avoid higher order commutator errors, it is even better to satisfy the condition at all times by constraining $\Omega_x(t) = \Omega_G(t) - \frac{(\lambda^2-2)(\Omega_G(t))^3}{\Delta^2}$.

Now a control in the $0-2$ transition is assumed not present in our system. Therefore the next highest error $\frac{\lambda\Omega_x^2(t)}{2\Delta}\sigma_{0,2}^x$ gives an error which according to our arguments is largely uncorrectable with the present formulation. That is we have

$$D_2 = \exp\left(i\frac{\lambda\Omega_x^2(t)}{2\Delta}\sigma_{0,2}^y/\Delta\right)$$

but the transformation leaves in the term $+i\dot{\hat{D}}_2(t)\hat{D}_2^\dagger(t)$. To validate this hypothesis we have calculated all orders of $H^{\tilde{D}(h)}$ using this iterative technique up to$h = 7$. Fig. 8.3 shows the effect of correcting the phase, rotation, and leakage errors in the qubit by order as a function of the strength of the coupling element λ_1. The top blue line is the zeroth order and corresponds to a Gaussian function. The red line

under it corrects the phase and selection error. The yellow line under that corrects the second order rotation error. Each subsequent line corrects errors to the next order assuming a derivative of a function is of the same order as the function itself (rather than one greater). We see that beyond the second or third order, most of the error is pushed into the derivatives which causes the expansion to exhibit asymptotic behaviour. Next we look at some other transformations and expansions.

8.5.1 Exact solution to the qutrit

Of course, if we assume instead that a control in the $0-2$ transition is possible then the expansion no longer behaves asymptotically. To see this, consider again our rotating frame qutrit Hamiltonian with additional control $\Omega_2^x(t)$,

$$H^R(t) = \sum_{j=1}^{2}(j\delta(t) + \Delta_j)\Pi_j + \sum_{j=1}^{2}\lambda_{j-1}\left[\frac{\Omega_x(t)}{2}\sigma_{j-1,j}^x + \frac{\Omega_y(t)}{2}\sigma_{j-1,j}^y\right] + \frac{\Omega_2^x(t)}{2}\sigma_{0,2}^x, \tag{8.25}$$

In practical terms the extra control is readily achievable given that SNOs are not harmonic oscillators. This can be seen for example by taking the Jaynes-Cummings diagonalization in Sec. 8.3.1 to next highest order. Thus, non-adjacent levels will still be coupled but the matrix element will simply be smaller. Note however that we can just absorb the smaller $\lambda_{0,2}$ into $\Omega_2^x(t)$ by simply increasing the amplitude of the field. Note also that $\Omega_2^x(t)$ will be at a different carrier frequency than the adjacent transitions as the energy level difference is approximately doubled.

We reuse the interaction frame transformation with respect to the off-phase control but now we do not need to perform an adiabatic expansion

$$D = \exp\left(-i\sum_{j=1}^{d-1}\lambda_{j-1}\int_0^t \frac{\Omega_y(t)}{2}\,dt\,\sigma_{j-1,j}^y\right)$$

To simplify the resulting expression we assign $A = \sqrt{3}\int_0^t \frac{\Omega_y(t)}{2}dt$ and pick the most common case $\lambda_1 = \sqrt{2}$. The transformed Hamiltonian $H^D = DH^RD^\dagger$ is cumbersome to write but contains 4 conditions to satisfy, the rotation, phase, leakage in $1-2$ and leakage in $0-2$ conditions (Eqs. 7.20), which are explicitly

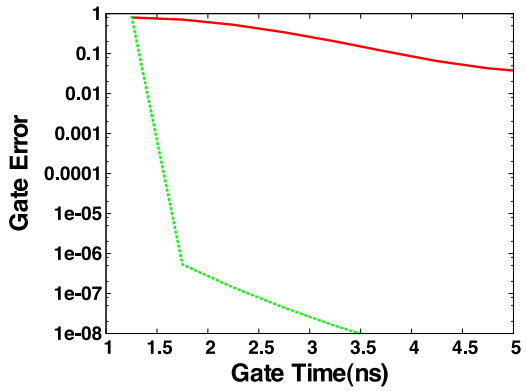

Figure 8.4: Gate error for the implementation of a NOT gate in a $d = 3$ SNO as a function of gate time. The solid red line corresponds to a Gaussian pulse while the dotted green line corresponds to the exact solution given by Eqs. 8.27 and 8.30. Parameters chosen are $\lambda_1 = \sqrt{2}$ and $\Delta/2\pi = 400$MHz.

$$
\begin{aligned}
2\left(\Omega_2^x(t)+\sqrt{2}(2\delta+\Delta)\right)\cos[A]+\left(4\Omega_2^x(t)-\sqrt{2}(\delta(t)+2\Delta)\right)\cos[2A]& \\
+12\Omega_2^x(t)-3\sqrt{2}\delta(t)+2\sqrt{6}\Omega_x(t)\sin[A]-\sqrt{6}\Omega_x(t)\sin[2A]=0,& \\
2\left(\sqrt{2}\Omega_2^x(t)+4\delta(t)+2\Delta\right)\cos[A]+\left(4\sqrt{2}\Omega_2^x(t)-2\delta(t)-4\Delta\right)\cos[2A]& \\
-6\sqrt{2}\Omega_2^x(t)+3\delta(t)+4\sqrt{3}\Omega_x(t)\sin[A]-2\sqrt{3}\Omega_x(t)\sin[2A]=0,& \\
12\Omega_x(t)\cos[A]-3\Omega_x(t)\cos[2A]& \\
-2\sqrt{3}\left(\sqrt{2}\Omega_2^x(t)+4\delta(t)+2\Delta+(2\sqrt{2}\Omega_2^x(t)-\delta(t)-2\Delta)\cos[A]\right)\sin[A]=9\Omega^\theta(t),& \\
3\sqrt{2}\Omega_x(t)\cos[2A]+2\sqrt{3}(-\Omega_2^x(t)-\sqrt{2}(2\delta+\Delta)& \\
+6\sqrt{2}\Omega_x(t)\cos[A]+\left(4\Omega_2^x(t)-\sqrt{2}(\delta(t)+2\Delta)\right)\cos[A])\sin[A]=0,&
\end{aligned}
\tag{8.26}
$$

where $\int_0^T \Omega^\Theta(t)dt = \Theta$. The first three equations can be solved in terms of A, as

in,

$$\begin{aligned}
\delta(t) &= \frac{4}{9}\left(\Delta + \frac{6\left(-3\Delta + 2\Delta\cos[A] + \sqrt{3}\Omega^{\Theta}(t)\left(-2\sin[A] + \sin[2A]\right)\right)}{8 - 3\cos[A] + \cos[3A]}\right) \\
\Omega_2^x(t) &= -\frac{4\sqrt{2}\sin\left[\frac{A}{2}\right]^2\left(2\Delta\cos[A] + \Delta\cos[2A] + 6\left(\Delta + \sqrt{3}\Omega^{\Theta}(t)\sin[A]\right)\right)}{9\left(8 - 3\cos[A] + \cos[3A]\right)} \\
\Omega_x(t) &= \frac{2\left(9\Omega^{\Theta}(t)\cos[A] - 2\sqrt{3}\Delta\sin[A]^3\right)}{3\left(8 - 3\cos[A] + \cos[3A]\right)}
\end{aligned} \tag{8.27}$$

The last equation is a self-consistency equation for A,

$$-3\sqrt{2}\Omega^{\Theta}(t)\left(2 + \cos[3A]\right) - 2\sqrt{6}\Delta\left(\sin[A] - 2\sin[2A]\right) = 0 \tag{8.28}$$

This can be solved by Taylor expansion followed by iterative solution,

$$\begin{aligned}
A^{(0)} &= \frac{\sqrt{3}\Omega^{\Theta}(t)}{2\Delta}, \\
A^{(m)} &= \sum_{n=1}^{m} \frac{-1^n}{(2n)!}(3A^{(m-1)})^{2n}\frac{3^{2n}\Omega^{\theta}(t) + 2\Delta A^{(m-1)}(1 - 2^{2n+1})}{6\Delta}.
\end{aligned} \tag{8.29}$$

The second quadrature control is then easily recovered by taking the derivative of A

$$\Omega_y(t) = \frac{2}{\sqrt{3}}\frac{dA}{dt} \tag{8.30}$$

We see in Fig. 8.4 that the error is exactly removed in the qutrit case when we are in the adiabatic limit (here $T > 1.5$ns) far surpassing what is possible with only the adjacent-level controls. Given the possibility of using a second control in experiment this strategy is clearly preferable to the simple single frequency case. The sum is taken to 7th order before numerical precision becomes an issue.

8.6 Generalized transformation

To find a general scheme that satisfies Eq. 7.20 we write $D(t) = \exp[-iS(t)]$ where $S(t)$ is an arbitrary Hermitian operator that we decompose as

$$S(t) = \sum_{j=1} s_{z,j}(t)\Pi_j + \sum_{j<k} s_{x,j,k}(t)\sigma_{j,k}^x + \sum_{j<k} s_{y,j,k}(t)\sigma_{j,k}^y, \tag{8.31}$$

and by assuming a power series in the small parameter ϵ, we can write each element as

$$s_{\alpha,j,k}(t) = \sum_{n-1}^{h} s^{(n)}_{\alpha,j,k}\epsilon^n,$$

where $\alpha = x, y$, or z. Thus we can find solutions to different orders h.

8.6.1 Zeroth order solutions

To the zeroth order we have Eq. 8.22 which simplifies to

$$\Omega^{\text{eff}(0)}(t) = \Omega_0^{\text{eff}}(t)\epsilon^0, \quad \Delta^{\text{eff}(0)}(t) = 0, \tag{8.32}$$

giving control solutions

$$\Omega_0^{\text{eff}}(t) = \Omega_G(t), \quad \delta(t) = 0. \tag{8.33}$$

These are the controls used in Sec. 8.4 and here we will use them as a benchmark for the higher order solutions. In Fig. 8.5, we plot the error, $1 - F_g$ with F_g given by Eq. (5.5) (blue dashed line), between a NOT gate an a unitary from the control field given by Eq. (3.3) with $A = \pi$ and $t_g = 4\sigma$ for a SNO with $d = 5$. In this figure it is clearly seen that the error associated with these controls is quite large; for fast gate times this error is unacceptable for quantum information processing, and long gate times will have additional error arising from decoherence.

8.6.2 First order solutions

To determine the first order solutions we need to determine the frame transformation conditions for $S^{(1)}(t)$ to remove leakage to first order. Direct substitution leads to

$$s^{(1)}_{x,1,k}(t) = 0, \quad s^{(1)}_{y,1,k}(t) = -\lambda_1 t_g \Omega_G(t)\delta_{k,2}/2 \tag{8.34}$$

where all other parameters $s_{\alpha,j,k}$ are free variables. To satisfy Eq. 8.22, the first order corrections to the control fields are

$$\bar{\Omega}^{(1)}_x(t) = 2\dot{s}^{(1)}_{x,0,1}(t), \tag{8.35}$$

$$\bar{\Omega}^{(1)}_y(t) = 2\dot{s}^{(1)}_{y,0,1}(t) - s^{(1)}_{z,1}(t)t_g\Omega_G(t), \tag{8.36}$$

$$\bar{\delta}^{(1)}(t) = \dot{s}^{(1)}_{z,1}(t) + 2s^{(1)}_{y,0,1}(t)t_g\Omega_G(t) + \frac{\lambda_1^2 t_g^2 \Omega_G^2(t)}{4}. \tag{8.37}$$

$$\tag{8.38}$$

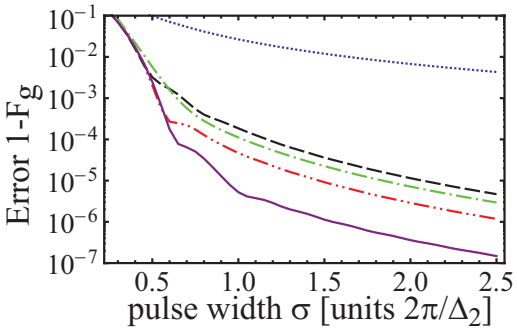

Figure 8.5: Gate error for the implementation of a NOT gate in a $d = 5$ SNO as a function of σ, with $t_g = 4\sigma$, and a Gaussian shaped pulse. The blue dotted line is the zeroth order solution. The black dash line is the first order Z-only correction. The red dash-dot-dot line is the first order Y-only correction. The green dash-dot line is the first order correction from the controls presented in Sec. (8.5). The purple solid line is for the optimal first order correction.

Here we see that there is a continuous family of DRAG pulses as (8.35-8.37) is under-constrained; however, in this section we will consider four particular solutions. In all of these solutions we take $s^{(1)}_{x,0,1}(t) = 0$ as it has no influence on our choice for $\bar{\Omega}^{(1)}_y(t)$ and $\bar{\delta}^{(1)}(t)$.

The first solution we consider is one where the control field $\Omega_y(t) = 0$. This is achieved by setting $s^{(1)}_{z,1}(t) = 2\dot{s}^{(1)}_{y,0,1}(t)/t_g\Omega_G(t)$, resulting in

$$\bar{\delta}^{(1)}(t) = \frac{2\ddot{s}^{(1)}_{y,0,1}(t)}{t_g\Omega_G(t)} - \frac{2\dot{s}^{(1)}_{y,0,1}(t)\dot{\Omega}_G(t)}{t_g\Omega^2_G(t)} + 2s^{(1)}_{y,0,1}(t)t_g\Omega_G(t) + \frac{\lambda_1^2 t_g^2\Omega^2_G(t)}{4}. \tag{8.39}$$

The simplest solution that satisfies $S^{(1)}(0) = S^{(1)}(t_g) = 0$ is $s^{(1)}_{y,0,1}(t) = 0$. In this case the controls become

$$\Omega_x(t) = \Omega_G, \quad \Omega_y(t) = 0, \quad \delta(t) = \frac{\lambda_1^2\Omega^2_G(t)}{4\Delta_2}. \tag{8.40}$$

For the SNO considered in Fig. 8.5, the error for this control set is plotted as the black dashed line. It clearly has a much lower error than the standard Gaussian amplitude modulation control, and we will refer to this as the Z-only correction.

The second control solution we consider is when the control field $\delta(t) = 0$. This is achieved by setting $s^{(1)}_{y,0,1}(t) = -\dot{s}^{(1)}_{z,1}(t)/2t_g\Omega_G(t) - \lambda_1^2 t_g\Omega_G(t)/8$, which results in

$$\bar{\Omega}^{(1)}_y(t) = -\frac{\ddot{s}^{(1)}_{z,1}(t)}{t_g\Omega_G(t)} + \frac{\dot{s}^{(1)}_{z,1}(t)\dot{\Omega}_G(t)}{t_g\Omega^2_G(t)} - \frac{\lambda_1^2 t_g\dot{\Omega}_G(t)}{4} - s^{(1)}_{z,1}(t)t_g\Omega_G(t). \tag{8.41}$$

Again, the simplest solution that satisfies $S^{(1)}(0) = S^{(1)}(t_g) = 0$ is $s^{(1)}_{z,1}(t) = 0$. In this case, the controls become

$$\Omega_x(t) = \Omega_G(t), \quad \Omega_y(t) = -\frac{\lambda_1^2 \dot{\Omega}_G(t)}{4\Delta_2}, \quad \delta(t) = 0, \tag{8.42}$$

and, for the SNO considered in the numerical demonstration, the error for this control set is plotted in Fig. 8.5 as the red dash-dot-dot line. Its error rate is lower than both the standard Gaussian controls and the Z-only correction. This is the control procedure used in Ref. [28] and Ref. [107], where it was referred to as simple DRAG and half derivative respectively. Here we will refer to this as the Y-only correction.

We also present the first order solution found in Sec. 8.5 [122]. This occurs when we choose $s^{(1)}_{x,0,1}(t) = s^{(1)}_{z,1}(t) = 0$ and $s^{(1)}_{y,0,1}(t) = -\Omega_G(t)/2$, resulting in

$$\begin{aligned} \Omega_x(t) &= \Omega_G(t), \quad \Omega_y(t) = -\frac{\dot{\Omega}_G(t)}{\Delta_2}, \\ \delta(t) &= \frac{\Omega_G^2(t)}{4\Delta_2}[\lambda_1^2 - 4]. \end{aligned} \tag{8.43}$$

This solution corresponds to fully removing the selection error by avoiding leaving in any derivative term. In Fig. 8.5, the green dash-dot line shows how the error scales with this control set. We see that for the first order solution, the error is larger than the Y-only correction method. That is, the first order version of Ref. [122] is not optimal. We will see below this can be traced to an increased rotation error.

The final solution we consider is what we refer to as the optimal first order solution. This is achieved by minimizing the elements in the effective Hamiltonian of $O(\epsilon^2)$. The transformation that minimizes these is

$$s^{(2)}_{y,0,2} = -\frac{1}{2} t_g \Omega_G \lambda_1 (t_g \Omega_G + s^{(1)}_{y,0,1}), \tag{8.44}$$

$$s^{(2)}_{x,0,2} = -\frac{1}{2} t_g \Omega_G \lambda_1 s^{(1)}_{x,0,1}, \tag{8.45}$$

$$s^{(2)}_{y,1,2} = -\lambda_1 \dot{s}^{(1)}_{x,0,1}, \tag{8.46}$$

$$s^{(2)}_{x,1,2} = \frac{1}{2}\lambda_1 (2\dot{\Omega}_G + 2\dot{s}^{(1)}_{y,0,1} - 2 t_g \Omega_G s^{(1)}_{z,1} + t_g \Omega_G s^{(1)}_{z,2}).$$

Using these expressions and requiring that the matrix elements of $H^{(2)}_{\text{extra}}(t)$ are zero in the qubit subspace (and elements coupling to the qubit subspace are zero) results in $s^{(1)}_{x,0,1}(t) = s^{(1)}_{z,1}(t) = 0$ and $s^{(1)}_{y,0,1}(t) = -t_g \Omega_G(t) \lambda_1 / 4$. Substituting these into Eq. 8.35

gives the controls fields

$$\begin{aligned}\Omega_x(t) =&\Omega_G(t), \quad \Omega_y(t) = -\frac{\dot{\Omega}_G(t)\lambda_1}{2\Delta_2},\\ \delta(t) =&\frac{\Omega_G^2(t)}{4\Delta_2}[\lambda_1^2 - 2\lambda_1].\end{aligned} \tag{8.47}$$

This optimal first order solution is plotted in Fig. 8.5 as the solid purple line. Its error is substantially lower the the other first order correction methods.

8.6.3 Second order solutions

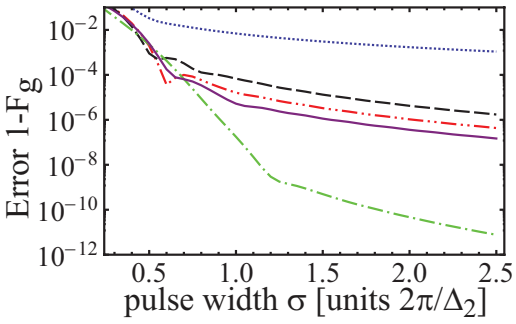

Figure 8.6: Gate error for the implementation of a NOT gate in a $d = 5$ SNO as a function of σ, with $t_g = 4\sigma$, and a Gaussian shaped pulse. The blue dotted line is the zeroth order solution. The black dash line is the second order Z-only correction. The red dash-dot-dot line is the second order Y-only correction. The green dash-dot line is the second order correction from the controls presented in Sec. 8.5. The purple solid line is for the optimal first order correction.

The higher order solutions become impractical to solve in generality because the number of terms grows quickly with increasing order. However, we can easily find the second and higher order corrections to the different first order solutions. We find that the corrections to the above four cases only change the $\Omega_x(t)$ field. In the Z-only case, $\Omega_x(t)$ becomes

$$\Omega_x(t) = \Omega_G(t) + \frac{\lambda_1^2\Omega_G^3}{8\Delta_2^2}, \tag{8.48}$$

in the Y-only case, $\Omega_x(t)$ becomes

$$\Omega_x(t) = \Omega_G(t) - \frac{\lambda_1^2(\lambda_1^2 - 4)\Omega_G^3}{32\Delta_2^2}, \tag{8.49}$$

and the control set presented in Sec. 8.5 gives

$$\Omega_x(t) = \Omega_G(t) + \frac{(\lambda_1^2 - 4)\Omega_G^3}{8\Delta_2^2}. \tag{8.50}$$

To demonstrate these corrections we plot in Fig. 8.6 the second order solutions. The line marking and colours are the same as in Fig. 8.5 with the exception that they now refer to second order solutions (all except the blue dotted line, which remains the zero order solution, and the purple solid line, which is the first order optimal solution). We see that the second order only makes small improvements to the Y-only (red dash-dot-dot) and Z-only (black dashed) first order solutions. Remarkably, the original DRAG scheme from Ref. [122] (green dash-dot) is improved substantially when corrected to second order. It is for this reason that we argue this is the best solution for implementing a DRAG correcting pulse.

We have not proven the DRAG solution to be optimal, and it seems likely that a better second order solution exists. Nonetheless it is clear that this is the only second order solution that completely removes the derivative term thereby minimizing selection error. Moreover, note that although these solutions are second order, the added complexity of implementing them in practice is minimal given that the rotation error does not drastically change the shape of the Rabi envelope.

8.7 Numerical optimization

We have explored the solution landscape to the leakage problem using the intuition from the semiclassical argument and with particular consideration to practicality and duration of the pulses. However, we have found that unlike in the case of frequency crosstalk, Sec. 7.1.1, an exact or optimal solution has been elusive, partly for algebraic complexity reasons (given the infinite Hilbert space) and partly because the transformations we have been able to derive appear to be asymptotic. Here we try to answer these questions using numerical optimization techniques paying particular attention still to practicality and time-duration of the pulses.

8.7.1 Prefactor optimization

Given the ease of implementing the first order solutions, in this section we consider the problem of numerically optimizing a value for the control fields with the following

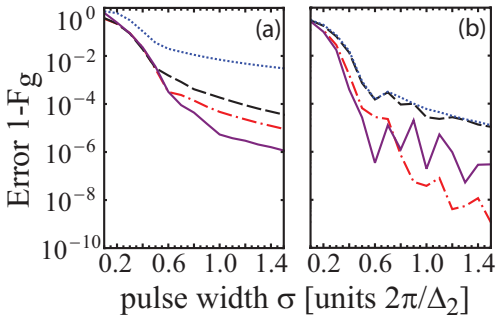

Figure 8.7: Gate error for the implementation of a NOT gate in a $d = 5$ SNO as a function of σ, with $t_g = 4\sigma$, and a Gaussian shaped pulse. Panel (a) is for optimized first order solutions. The blue is for optimized α with $\beta = \gamma = \delta_0 = 0$ (zeroth order solution). The black dashed is for optimized α and γ with $\beta = \delta_0 = 0$ (Z-only solution). Red dash-dot-dot is optimized α and β with $\gamma = \delta_0 = 0$ (Y-only solution). Purple solid line is for optimized α, β, and γ with $\delta_0 = 0$ (optimal first order solution). Panel (b) is the same as (a) but with δ_0 being optimized.

ansatz

$$\begin{aligned} \Omega_x(t) =& \alpha\Omega_G(t), \quad \Omega_y(t) = -\beta\frac{\dot{\Omega}_G(t)}{\Delta_2}, \\ \delta(t) =& \gamma\frac{\Omega_G^2(t)}{\Delta_2} + \delta_0, \end{aligned} \tag{8.51}$$

where α, β, γ and δ_0 are fit parameters. We consider a SNO with $\Delta_2 = -2\pi$, $\lambda_{j-1} = \sqrt{j}$, $d = 5$, and a control field given by Eq. 3.3 with $A = \pi$ and $t_g = 4\sigma$ (same as before). In Fig. 8.7 we plot the gate error as a function of σ for different optimizations. The optimization procedure was done with Mathematica with a working precision of 10. In Fig. 8.7 (a) we consider the case when $\delta_0 = 0$, and we find that optimizing the weighting of the control fields only improves the first order solutions slightly. This is expected as the second order solutions require different functional forms for the controls. However, when we allow δ_0 to be non-zero, we find some interesting results. For the numerical parameters considered, we find that implementing a time varying $\delta(t)$ ($\gamma \neq 0$) does not lead to any improvements. This is seen in Fig. 8.7 (b) where we show that the error arising from an optimized Gaussian with added constant detuning (blue dotted line) is approximately equal to the optimized Z-only correction with an added constant detuning (black dashed line). Furthermore, the optimized Y-only correction with an added constant detuning (red dash-dot-dot) is approximately equal to the optimal first order solution with an

added constant detuning (solid purple). We also find that for the solutions with the derivative for the Y-control (solid purple and red dash-dot-dot) the gate error is much lower then in the other cases (blue dotted and black dashed). This gate error is approximately equal to those found with the second order corrections from Fig. 8.6 (green dash-dot line). We see from these numerics that the optimal DRAG-like solution can be obtained by applying a pulse to the x-axis (and its derivative to the y-axis) with a frequency that is not equal to the transition frequency of qubit, Sec. 4.2. Once again, this corresponds to removing the selection error which effectively moves the control problem back into a qubit problem where we know the right combination of rotation angle and detuning can always find an optimal solution.

8.7.2 Full optimization

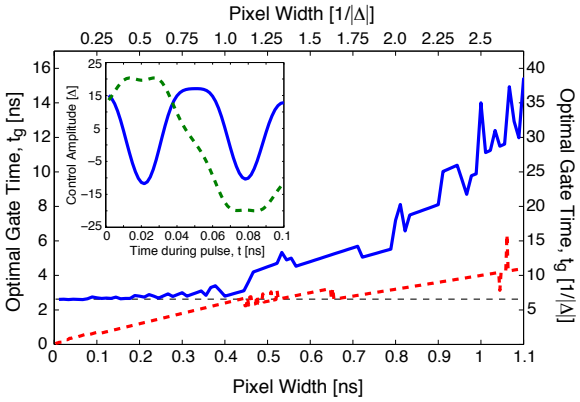

Figure 8.8: Pulse duration vs. pixel width for 1 control (blue solid) and 2 controls (red dashed). Pulse duration is calculated as the minimal time to bring the error down to 10^{-5}. The insert shows the two control GRAPE result for $t_g = 1/4\Delta$ (blue solid is Ω^x, and green dashed is Ω^y).

We apply optimal control theory to our system by employing gradient ascent numerical optimization, Sec. 5.3. Restricting GRAPE to 1ns pixels (which is consistent with current experimental limitations, Sec. 6.1) and setting the initial condition for the Ω^x quadrature to be the Gaussian pulse, Eq. (3.3), the algorithm quickly converges to the optimal solution. After optimization, the shape of the Gaussian in the $\mathcal{E}^x$ quadrature is largely unchanged. However the Ω^y quadrature changes from 0 to a shape that resembles the derivative of the Gaussian. GRAPE essentially takes the idea of DRAG to infinite order and discretizes the pulse. The shapes of the pulses are smooth with no sudden rises or falls in the control amplitudes.

The two quadrature control GRAPE pulse outperforms its one quadrature control version [148]. Fig. 8.8 shows the minimal pulse duration to obtain gate errors smaller than 10^{-5} vs the pixel width for both one (blue solid) and two (red dashed) controls. Here, we see that at small pixel widths (continuous limit) the one control saturates at a time of $2\pi/\Delta$ [148] (black dashed line), whereas the two controls allows for arbitrarily small pulse times, allowing operation to go well into the diabatic regime. This is because, in the one control case, phase cancelation of the leakage is done through phase accumulation by the natural precession of the third level, whereas, for two controls, it can be generated instantaneously. The insert of Fig. 8.8 shows such a continuous two control pulse; this pulse clearly displays a combination of DRAG and composite features. The symmetry of the real control suggests that perhaps a second derivative is being added to the pulse to remove selection error from the $0-2$ transition. At the larger pixels (more experimentally realizable), the two controls offer a substantial improvement over one control. For example, at 1ns pixels, there is a factor of 3.5 improvement. The results are consistent with Sec. 8.6 where we saw a drastic improvement when going from a single control to two controls and we note the third (detuning) control can be generated from the other two controls using phase transformations, Sec. 4.2, offering little if any additional improvement.

8.8 Incoherent effects

Superconducting systems suffer greatly from loss of coherence. To properly benchmark what control of these systems is possible requires taking these effects into account [144]. We begin by considering the effect of decay on our control pulses. Secondly, we consider the effect of parameter uncertainty with regard to the energy landscape on the retention of coherence in the pulses. This has application to many physical systems, in particular optical lattices and other ensemble qubits where the measurement of many particles lends itself vulnerable to inhomogeneity effects. Both these effects are important and can be used to predict what kind of experimental errors to expect.

8.8.1 Relaxation

In realistic physical systems, decoherence cannot be neglected. Here, we include this effect by simulating the evolution by a master equation for each optimal pulse. The

master equation we use is

$$\dot{\rho} = -i[H^R, \rho] + \sum_{j=1,2} \left[\tfrac{1}{T_1^j} \mathcal{D}[\sigma_j^-]\rho + \tfrac{1}{T_\phi^j} \mathcal{D}[\Pi_j]\rho \right] \tag{8.52}$$

where ρ is the density matrix, T_1^j and T_ϕ^j represent relaxation and pure dephasing times respectively, and $\mathcal{D}$ is the damping super-operator, defined as

$$\mathcal{D}[A]\rho = A\rho A^\dagger - \frac{1}{2}A^\dagger A\rho - \frac{1}{2}\rho A^\dagger A. \tag{8.53}$$

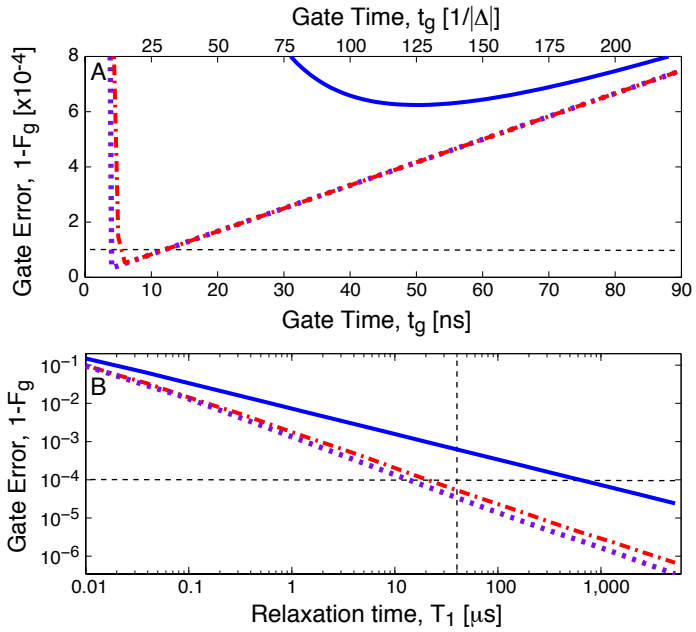

Figure 8.9: A. error vs. pulse duration for $T_1 = 40\mu$s and B. minimum error vs. T_1 for Gaussian (σ=$t_g/2$) (blue solid), Gaussian (σ=$t_g/2$) with DRAG (red dot-dashed), and GRAPE (purple dotted) with 1ns pixels.

For the simulation, we used parameters consistent with the low-dispersion Transmon qubit and as such take pure dephasing to be zero. The results are shown in Fig. 8.9. Fig. 8.9A plots the gate error $(1 - F_g)$ as a function of gate time for $T_1 = 40\mu$s. Here, we see there is an optimal value for the gate time (when the error due to decoherence is about the same as the error due to leakage) and for both the DRAG (green dashed) and the GRAPE (purple dotted) pulse, this is much less then

the standard Gaussian pulse (blue solid) with an optimal error of 5.3×10^{-5} and 3.4×10^{-5} respectively. Fig. 8.9B plots this minimum as a function of T_1 where we observe that at current experimental values $T_1 \approx (1-4)\ \mu$s [47] DRAG and GRAPE still outperform the simple Gaussian by about an order of magnitude.

8.8.2 Dispersion

We wish to address uncertainty with respect to the quasi-momentum in our periodic-potential system (e.g. charge noise for a transmon/charge qubit) coming from dealing with a statistical ensemble of systems either in time or in space (Sec. 3.5). To do this we apply the numerical optimization theory developed in Sec. 5.3.2.

Bloch's theorem states that a periodic Hamiltonian such as Eq. 8.3 has eigenstates of the form, $|\psi_n^{(k)}\rangle$ or

$$\langle x||\psi_n^{(k)}\rangle = \psi_n^{(k)}(x)e^{ikx}, \tag{8.54}$$

where n indicates the energy band and k the quasi-momentum and $\psi_n^{(k)}(m\pi x) = \psi_n^{(k)}(x)$ for all integers m. These Bloch functions satisfy the Schrödinger equation

$$E_n^{(k)}\psi_n^{(k)} = H_0^{(k)}\psi_n^{(k)} \tag{8.55}$$

with the now k-dependent Hamiltonian

$$H_0^{(k)} = (p-k)^2 + \frac{r}{2}(1-\cos 2x), \tag{8.56}$$

The resulting wavefunctions can be calculated using the central matrix method or the Bloch method [4]. Of primary importance is that the quasi-momentum is preserved. The single-particle energy eigenstates of the optical lattice thus, from a control perspective, form an inhomogeneous ensemble of discrete quantum systems with different parameters in the Hamiltonian Eq. 8.56. We shall see later on that the controls at hand also conserve k thus this ensemble decomposition remains consistent even with controls.

To control this system we introduce the parameters $\eta(t)$ and $\phi(t)$ which represent the ratio of intensity with respect to r and the phase, respectively. These correspond to the vertical amplitude and horizontal displacement of the lattice potential, respectively. In particular, they correspond to the amplitude and phase of the control laser in an optical lattice implementation [114], though a similar orthogonal set of controls can be found in Josephson qubits. Rewriting (8.56) in terms of these new

variables we obtain,

$$\begin{aligned} H^{(k)}(t) =&(p-k)^2 + \frac{r}{2}\left[1+\eta(t)\right]\left[1-\cos\left(2x+\phi(t)\right)\right] \\ =&H_0^{(k)} - \frac{r}{2}\left[1-\cos(2x)\right] + \frac{r}{2}\left[1+\eta(t)\right]\left[1-\cos\left(2x+\phi(t)\right)\right] \\ =&H_0^{(k)} + \frac{r}{2}\left[1-(1+\eta(t))\cos\phi(t)\right]\cos(2x) + \frac{r}{2}(1+\eta(t))\sin\phi(t)\sin(2x) + \frac{r}{2}\eta(t). \end{aligned} \tag{8.57}$$

In order to express the Hamiltonian in the standard bilinear form for quantum control, Ref. [85], we reparametrize the control fields in terms of

$$\begin{aligned} \alpha(t) =&\frac{r}{4}\Big[1-\left[1+\eta(t)\right]\cos\phi(t)\Big], \\ \beta(t) =&\frac{r}{4}\Big[\left[1+\eta(t)\right]\sin\phi(t)\Big]. \end{aligned} \tag{8.58}$$

so that the total control Hamiltonian (neglecting the global phase) is

$$H^{(k)}(t) = H_0^{(k)} + 2\alpha(t)\cos(2x) + 2\beta(t)\sin(2x). \tag{8.59}$$

The lattice will be filled with multiple atoms occupying a range of k-values. Measurement of the success of a quantum operation can be performed in the manner suggested in [124], which involves averaging single particle measurements over the entire ensemble. Demanding that we find $\alpha(t)$ and $\beta(t)$ that perform the desired gate irrespective of k ensures that any gates we find will retain their fidelity after this averaging.

States with different quasi-momenta evolve independently under the applied Hamiltonian and so can be viewed as an ensemble of non-interacting particles, labelled by the quasi-momentum. The eigen-energies of this system depend on the quasi-momentum, i.e. $H_0^{(k)}|\psi_n^{(k)}\rangle = E_n^{(k)}|\psi_n^{(k)}\rangle$, leading to dispersion of the resonance frequencies across different quasi-momenta. That is the resonant frequencies of excitations associated with different vertical transitions are not the same, see Fig. 8.10. We characterize the dispersion by the dimensionless quantity

$$\mathcal{D} = 1 - \frac{\Delta E_{01}^{(1)}}{\Delta E_{01}^{(0)}}, \tag{8.60}$$

where $\Delta E_{01}^{(k)} = E_1^{(k)} - E_0^{(k)}$. In addition to energy dispersion there is dispersion in the strengths of the couplings between the bands with respect to the quasi-momentum.

The main consequence of the dispersion is that a single harmonic excitation cannot simultaneously resonantly excite all transitions. This poses a formidable challenge for controlling an ensemble with high dispersion. For the lattice potential, the amount of energy dispersion decreases with an increase in the strength of potential, Fig. 8.10. Unfortunately, the anharmonicity of the lattice, which is essential to resolve different transition frequencies, also decreases with increasing potential depth. In the case of the Transmon, the system is operated in the low dispersion regime (r>20) and dispersion is most often neglected. Here we aim to be more precise by quantifying the trade-off when choosing the depth of the lattice between dispersion and anharmonicity and the degree to which both can be simultaneously corrected.

Note that in superconducting qubits, the CORPSE pulse sequence [33] has been applied to mitigate fluctuations of the energy splittings [32], which has been investigated further numerically [121]. Here we look not only at dispersion in the energy but more comprehensively at changes in the strength of all operators in the Hamiltonian as a function of quasi-momentum.

8.8.2.1 Optimization

Using the techniques outlined in Sec. 5.3.2, we optimized control parameters for the lattice Hamiltonian in Eq. 8.3 for varying potential depths and pulse durations.

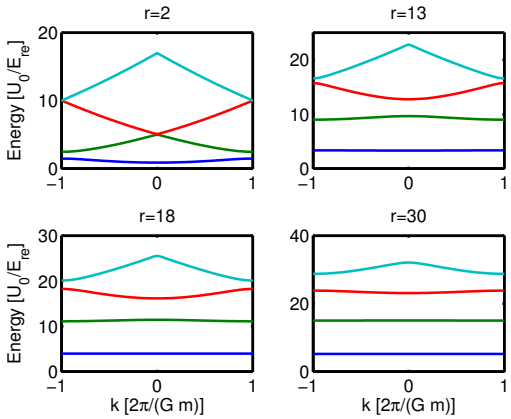

Figure 8.10: The first four energy 1-D band structures are shown in the momentum basis for four different potential depths r. The lower potential depth $r = 2$ (top left) shows the largest amount of energy dispersion with the closest energy crossings. The depth $r = 30$ (bottom right) shows a lesser amount of dispersion and larger energy splittings.

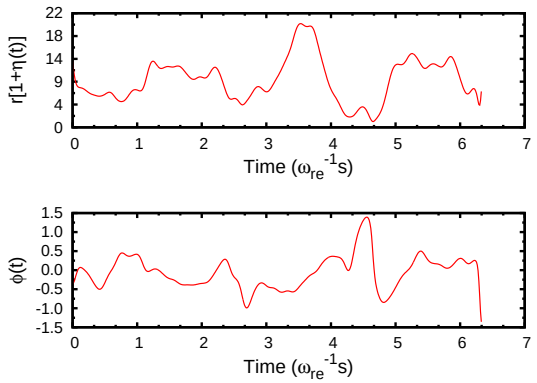

Figure 8.11: Optimized controls for preparing a X_π-gate when the optimization uses only ten points in quasi momentum space. Pulses are well-behaved in the sense that they don't translate the lattice by a full lattice site or possess very large excursions from the initial value for r. (a) An optimized pulse for a potential depth $r = 7$ with 32.7% dispersion and fidelity 98.7% which is calculated by sampling over 100 quasi momentum values. . The duration of the pulse was 5 free oscillations (at k=0).

Our target gate is a X_π-gate on the first two energy bands. We considered both a Rabi pulse assuming no dispersion and bounded random controls fields as the initial guesses for the control fields. The evolution is simulated using only the first six energy bands from our model and sampling over 10 values of the quasi-momentum. While we sampled over only a small number of values for the quasi-momenta in order to have efficiently performed the numerical search, when we calculate the final fidelity we sampled over a finer grid in quasi-momenta space.

Here are two examples of optimized controls for performing X_π-gates on our lattice model. One is a control for a lattice system with potential depth $r = 7$ (which corresponds to a dispersion of 13.2%), as shown in Fig. 8.8.2.1. The duration of this control was 5 free oscillations (at k=0). The gate error across the ensemble was less than 2% with the exception of particles with quasi-momentum near the edges of the Brillouin zone. Even with such high dispersions in the energies and control Hamiltonians we are able to find gates with reasonable fidelities.

We observe the relationship between the maximum fidelity of the solutions and the level of dispersion of the Hamiltonian. Ideally, our gradient search will halt when a local maximum in the fidelity is reached but this may take a considerable amount of computational resources. All our optimizations were halted after either the algorithm converged to a solution or 10^5 updates in the control parameters were

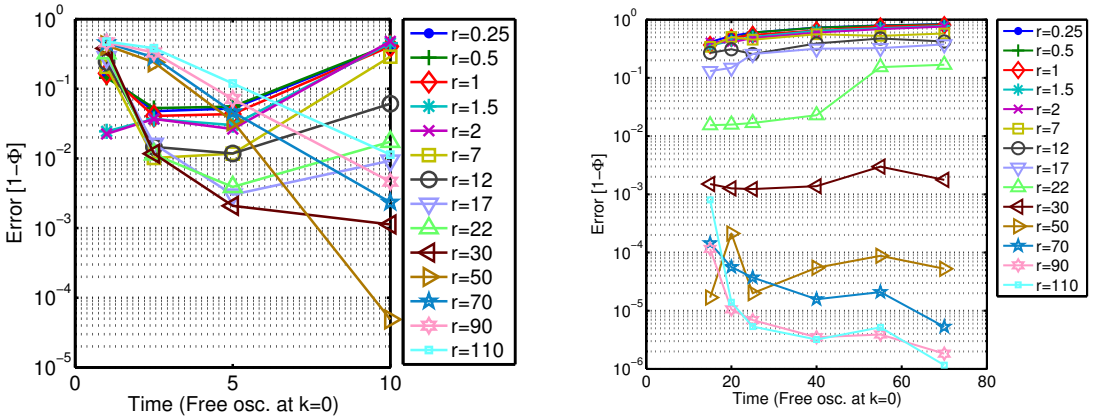

Figure 8.12: The maximum fidelity for an optimized pulse over a range of times, from 1 to 10 free oscillations (at k=0), preparing a X_π-gate as a series of different potential depths. The potential depth ranged from $r = 0.25$ to $r = 110$, giving a dispersion ranging from 93.8% to 0.01%, respectively. Each point is an average of optimized pulses from 1 Rabi and 10 random initial pulse.

performed. Nevertheless, for short times (Fig. 8.12A) we find excellent control fields ($F > 0.90$) for the entire range of potential depths we considered, $0.25 \leq r \leq 110$.

In general, for long gate times, the maximum fidelity solution for X_π-gates became lower as the dispersion was increased, Fig. 8.12B. The optimization becomes less tractable with higher dispersion since we are asking a single set of control fields to solve wildly different problems depending on where the system is in quasi-momentum space. As a result, broadband pulses must be tailored to accommodate a range of possible energies and couplings. We also find that for high dispersion the fidelity becomes worse as the gate time becomes longer. This may seem counterintuitive, but it is due to the fact that as the gate time increases the fidelity can vary more quickly as a function of k. For very short gate times, the fidelity is fairly constant across quasi-momentum space, but for long gates we find that the fidelity is only high for the specific points we optimized, and dips to almost zero in the intermediate regimes.

We can see the connection between gate duration and rate of change of fidelity very easily through the simulation described in Fig. 8.13. Here we run GRAPE for exactly one value of the quasi-momentum and and then look at the performance of this control field across k. W find that as the duration of the gate increases the fidelity function become more and more tightly peaked about the single value we optimized, consistent with the spectral bandwidth of the pulse.

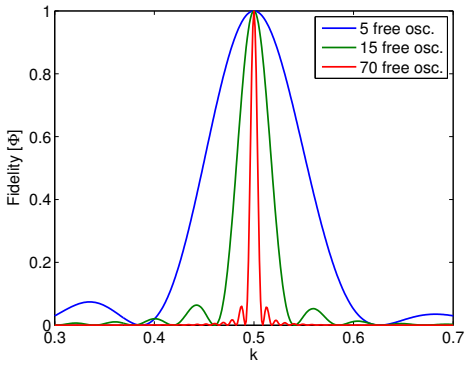

Figure 8.13: Pulses for performing X_π-gate at $r = 2$ for three different times were optimized specifically for $k = 0.5$ and the fidelity response over quasi-momenta shown. Shorter pulses have larger spectral bandwidth and thus affect a larger range of quasi-momenta.

8.9 Implementation

The DRAG result was first implemented in superconducting qubits at Yale with transmons in [28, 39] and shortly after with phase qubits at U.C. Santa Barbara [105]. The transmon qubit is a prime candidate for studying leakage because the anharmonicity Δ is only $\sim 3-5\%$ of $\omega_{0,1}$. For reasons of practicality and simplicity the Y-only solution from Sec. 8.6 is explored in this experiment for demonstration purposes though the full optimal solution would be achievable to implement, in particular increased coherence times.

Although Gaussian control pulses (characterized by a width σ) are often the paradigm due to their localized frequency bandwidths given by $B = 1/2\pi\sigma$, leakage errors can occur as gate times are reduced such that B is comparable to Δ. The correction protocol to the leakage errors as prescribed above is to apply an additional control on the quadrature channel, $\mathcal{E}^y(t) = \beta\dot{\mathcal{E}}^x(t)$ and a dynamical detuning of the drive frequency $\delta(t) = \mathcal{E}^x(t)^2(-4\beta\Delta + \lambda^2)/4\Delta$, where β is a scale parameter. For a qutrit driven without dynamical detuning, the optimal $\beta = \lambda^2/4\Delta$.

The experiments are performed in a circuit QED sample consisting of two transmons coupled to a coplanar waveguide resonator. The sample fabrication and experimental setup are described in Ref. [39]. The two transmons (designated L and R) are detuned from one another with ground to excited state ($0 \leftrightarrow 1$) transition frequencies of $\omega_{0,1}^{\mathrm{L,R}}/2\pi = 8.210, 9.645\,\mathrm{GHz}$ and the ground state cavity frequency is$\omega_\mathrm{C}/2\pi = 6.902\,\mathrm{GHz}$. The anharmonicities of the transmons are found using two-

tone spectroscopy measurements [154] to be $\Delta^{\mathrm{L,R}} = 330,\, 300\,\mathrm{MHz}$ and coherence times are measured to be $T_1^{\mathrm{L,R}} = 1.2,\, 0.9\,\mu\mathrm{s}$ and $T_2^{*\mathrm{L,R}} = 1.5,\, 1.1\,\mu\mathrm{s}$.

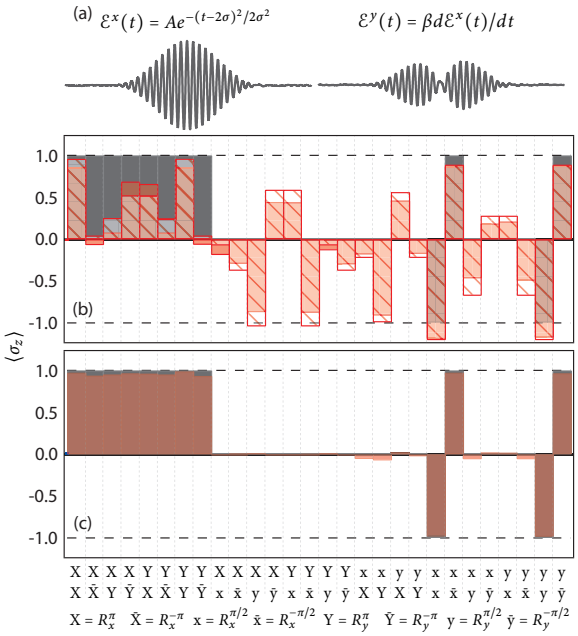

Figure 8.14: (a) Gaussian and derivative pulse shapes applied to the in-phase and quadrature control channels, respectively, for implementing DRAG to first order. (b) Measured $\langle\sigma_z\rangle$ on qubit L (shaded red bars) with Gaussians with $\sigma = 1$ns for a test sequence consisting of pairs of π and $\pi/2$ rotations. The slash filled bars correspond to a master-equation simulation of a three-level system with parameters of the sample tested. The grey filled bars reflect ideal values. (c) Similarly measured $\langle\sigma_z\rangle$ on qubit L but using derivative pulses on the quadrature with a scale factor $\beta^{\mathrm{L}} = 0.4$ (shaded red bars), also overlaid on the ideal values (shaded dark grey bars).

To implement DRAG to first order and perform single-qubit gates on the transmons, we use an arbitrary-waveform generator to shape microwave-frequency pulses with quadrature control, permitting rotations about either the x- or y-axis of each qubit. We fix the drive frequency to $\omega_{0,1}$, and the pulse amplitudes and phases define the rotation angle and axis orientation, respectively. When performing an x rotation, $\mathcal{E}^x(t)$ is a Gaussian pulse shape, Fig. (8.14)(a), while the derivative of the Gaussian is applied simultaneously on the other quadrature, $\mathcal{E}^y(t) = \beta\dot{\mathcal{E}}^x(t)$. All of the pulses are truncated to 2σ from the center with an added buffer time of 5ns to ensure complete separation between concatenated pulses.

A simple test sequence is used to tune up the scale parameter β as well as to demonstrate the effect of using first-order DRAG pulses versus standard single-quadrature Gaussians. The sequence consists of pairs of π and $\pi/2$ pulses around both the x- and$y$-axes. An important feature of this sequence is that the final average z-projection of the single qubit, $\langle\sigma_z\rangle$, will ideally take on values from the set $S = \{+1, 0, -1\}$, making any deviations easily visible.

Using Gaussian pulse shaping with $\sigma = 1\,\text{ns}$ and implementing the test sequence for qubit L, we find significant deviations from S, as shown in the solid red bars of Fig. 8.14 B. The theoretical results for each pair of rotations are shown with solid grey bars in the background. We observe that the largest errors occur when the two rotations are around different axes, which indicates the presence of significant phase error, or a residual z rotation after the first gate.

We repeat the same test sequence, but applying the derivative of the Gaussian to the quadrature channel. By varying β, it is possible to find an optimal value such that the measurements of $\langle\sigma_z\rangle$ agree very well with the theoretical predictions. The shaded red bars of Fig. 8.14 C show measured $\langle\sigma_z\rangle$ for qubit L using $\beta = 0.4$. Here, deviations from the ideal grey bars decrease to $< 2\%$. We have also applied the DRAG protocol for qubit R, finding the optimal value $\beta = 0.25$ (data not shown). From the experimental determination of β and α_1, we can infer the second excited state coupling strengths $\lambda^{\text{L,R}} = 1.82,\ 1.41$. Using λ^{L} and the three-level model of Eq. (8.4), a master equation simulation for the Gaussian shaping gives the red hash-filled bars in Fig.8.14, which demonstrates good agreement with the experiment.

We find excellent agreement for λ with the calculation for the anticipated λ in a cavity-transmon coupled system (Sec. 8.3.1). The cavity modifies the drive strengths $\Omega_{0,1}$ and $\Omega_{1,2}$ due to its filtering effect. Specifically, for a transmon in a cavity, we have

$$\lambda = \frac{g_{j-1,j}(\omega'_{0,1} - \omega_C)}{g_{0,1}(\omega'_{j-1,j} - \omega_C)}. \tag{8.61}$$

where $j = 1, 2$ for the transmon excitation level, $\gamma = \{1, -1\}$ depending on whether the qubit is located at the input or output side of the cavity, and $g_{i,j}$ is the matrix element coupling the $i \leftrightarrow j$ transmon transition to the cavity [94]. Using the relevant parameters of the two transmons in the experiment and including only the fundamental mode of the cavity, we find $\lambda^{\text{L,R)}} = 1.85,\ 1.57$, within 12% of those determined from the test sequence. There are corrections to the drive due to the higher modes of the cavity, but it is difficult to use Eq. 8.61 to estimate as a result of cutoff dependence.

We characterize the degree of improvement to single-qubit gates by using the

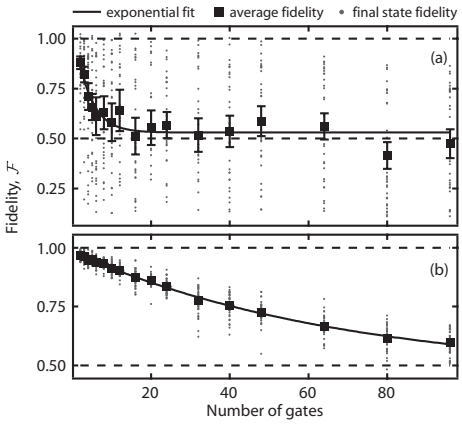

Figure 8.15: Randomized benchmarking for qubit L with $\sigma = 1\,\text{ns}$ using (a) Gaussian pulses, and (b) additional Gaussian derivative pulses on the quadrature channel. The scattered grey points are extracted fidelities for 32 RB sequences, truncated at different numbers of gates. A remarkable reduction in the extracted average error per gate (black squares) of the benchmarking results is observed going from (a) to (b). The error bars indicate the variance of all the RB sequences and are smaller than the squares in (b).

technique of randomized benchmarking (RB) [**?**]. RB allows us to determine the average error per gate through the application of long sequences of alternating Clifford gates ($R^{\pi/2}_{x,y}$) and Pauli gates, chosen from $\{1, R^{\pi}_{x}, R^{\pi}_{y}, R^{\pi}_{z}\}$ [62]. We use the RB pulse sequences originally given in Ref. [?] and adapted to superconducting qubits in Ref. [27] for both the Gaussian and the derivative pulse shaping for transmon L. We truncate the randomized sequences at various lengths and compare the resulting measurement of $\langle\sigma_z\rangle$ to the ideal final state to obtain the fidelity$\mathcal{F}$. There is an exponential decrease in$\mathcal{F}$ with an increasing number of gates in the randomized sequences. This RB protocol is then repeated for various pulse widths, corresponding to different Gaussian standard deviations, $\sigma \in \{1, 2, 3, 4, 6\}\,\text{ns}$.

Using the Gaussian shaping, we find a large reduction in fidelity with the shortest pulses, $\sigma = 1\,\text{ns}$, Fig. 8.15A. The scattered grey points give $\mathcal{F}$ for 32 different randomized sequences applied as a function of the number of gates in the sequences. When averaged together, we observe a simple decay of $\bar{\mathcal{F}}$ as a function of the number of gates (solid black squares). Fitting the data with an exponential decay (solid black line), we extract an average error per gate, EPG $= 1 - \bar{\mathcal{F}}$ of 0.13 ± 0.02. However, when employing the first-order DRAG, we find a dramatic improvement in the gate performance at $\sigma = 1\,\text{nsns}$, Fig. 8.15B. There is a significant reduction in the spread

of the grey points corresponding to all the different randomized sequences, and a fit (solid black line) to the exponential decay of the average fidelity (solid black squares) gives EPG $= 0.007 \pm 0.005$.

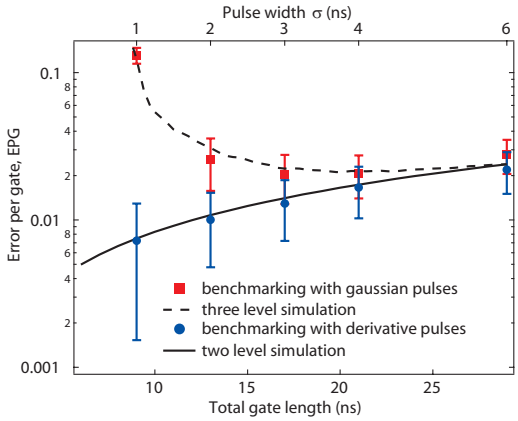

Figure 8.16: Comparison of single-qubit gate errors with and without DRAG. Error per gate for the left qubit extracted from randomized benchmarking for different gate lengths using both Gaussian pulses (red squares) and first-order DRAG pulses (blue squares). Excellent overlap with theory for gate error including qubit decoherence (black curve) suggests that the DRAG pulses successfully eliminate the errors due to the presence of higher levels. Gate errors down to 0.007, which are otherwise unattainable with Gaussian pulses, are reached using DRAG.

Figure 8.16 summarizes the improvement to EPG for different σ by using DRAG. The solid squares are the EPG found using Gaussians, revealing a minimum of0.02 ± 0.007 at $\sigma = 3\,\text{ns}$, before considerably increasing for shorter pulse lengths. Excellent agreement is found with a master equation simulation (dashed line) of the gate error for a qutrit system incorporating only decoherence times and coupling strengths measured in independent experiments. Using first-order DRAG, we find the solid circles in Fig. 8.16, which follow a monotonic decrease in EPG with decreasing σ. Here again we have included a master equation prediction (solid line) of just a single qubit with the same parameters, also giving excellent agreement with the experimentally determined values and demonstrating that DRAG has reduced the response of the system to be like that of a single qubit.

Finally, implementing DRAG on both qubits simultaneously, we can also generate and detect higher quality two-qubit states. Performing state tomography to obtain the two-qubit density matrix ρ via joint readout [26, 48] requires 15 linearly independent measurements, corresponding to the application of all combinations of

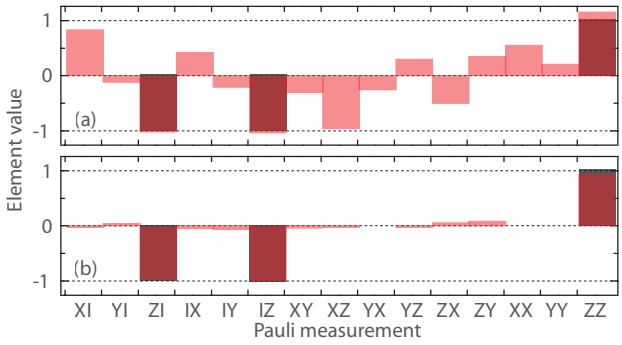

Figure 8.17: Measured two-qubit Pauli sets for preparing the state $|1,1\rangle$ with (a) Gaussian pulses and (b) DRAG pulses ($\beta^{\mathrm{L}} = 0.4$, $\beta^{\mathrm{R}} = 0.25$) applied to the quadrature channels of both transmon L and qubit R. The ideal Pauli set is shown in grey.

I, R_x^{π}, $R_x^{\pi/2}$, and $R_y^{\pi/2}$ on the two qubits prior to measurement. Thus, errors in these analysis rotations in addition to the state preparation pulses can result in incorrect determination of ρ. The two-qubit Pauli set $\vec{P}$ [26] can be used to visualize ρ for the state $|1\rangle_{\mathrm{L}} \otimes |1\rangle_{\mathrm{R}}$ having used Gaussian (Fig. 8.17A) and DRAG (Fig. 8.17B) pulse shaping. $\vec{P}$ consists of ensemble averages of the 15 non-trivial combinations of Pauli operators on both qubits. The ideal $\vec{P}$ of the state is characterized by unit magnitude in$\langle ZI\rangle$, $\langle IZ\rangle$, and $\langle ZZ\rangle$ and zero for all other elements. We can see that with the standard Gaussian pulse shaping, there are substantial ($\sim 50 - 100\%$ of unity) deviations on ideally zero elements, whereas with the DRAG pulses, the Pauli set bars are very close to their ideal values.

By implementing a simple approximation to the optimal control pulses for a multi-transmon coupled-cavity system, we have reduced gate errors below the 10^{-2} level, limited by decoherence. The agreement of the various experiments with and without DRAG pulse shaping with a qutrit model reflects that gate errors due to the coupling to a higher excited state can be minimized while continuing to shorten gate time. Moving forward with optimal control, a tenfold decrease in gate time to approach $\sim 1\,\mathrm{ns}$ through improved electronics or a tenfold increase in coherence times to $\sim 10\,\mu\mathrm{s}$ would place us right at the quoted 10^{-3} fault-tolerant threshold [122]. Furthermore, DRAG is extendable to systems of more than two multi-level atoms for quantum information processing, and has already been employed to enhance single-qubit gates in a circuit QED device with four superconducting qubits [40].

Chapter 9

Multi-Channel Leakage

Complications in the Hilbert space nature of physical systems can be found beyond even the qutrit approximation as a quantum system is scaled to include more elements or when very low error rates are desired. Not only do simultaneous leakages have to be taken into account, but introducing an external driving field may couple different elements together (e.g. crosstalk in Sec. 7.2). In this section we consider three generic multi-channel leakage problems. The first is encountered in Sec. 9.1 when trying to transition between higher energy levels in an anharmonic ladder such as a Duffing oscillator, where now leakages exist both below and above the driven transition (see Fig. 9.1A). This is first done numerically in Sec. 9.1.2 with our results published in Ref. [86] (pulses found by Botan Khani). Next the same problem is studied analytically in Sec. 9.1.3. Secondly, in Sec. 9.2, we find analytic forms for suppressing multiple leakage channels coupled to the same energy level (see Fig. 9.1B). The analytic results of Secs. 9.1.3 and 9.2 are published in Ref. [59] with Seth Merkel and Jay Gambetta contributing to the derivations. Last, in Sec. 9.3, we consider two qutrits and look at the simultaneous removal of both leakages when coupling them (see Fig. 9.1C). This is solved both numerically in Sec. 9.3.2 (solutions found by Amira Eltony) and analytically in Sec. 9.3.3.

9.1 Duffing oscillator

We have seen that anharmonic systems such as those built from Josephson junctions can be well described by the Duffing oscillator (Sec. 6.7.1), with Hamiltonian

$$H = \frac{p^2}{2m} + \frac{m\omega_0^2}{2}x^2 + \delta\frac{m^2\omega_0^2}{3\hbar}x^4 \quad \delta > 0 \tag{9.1}$$

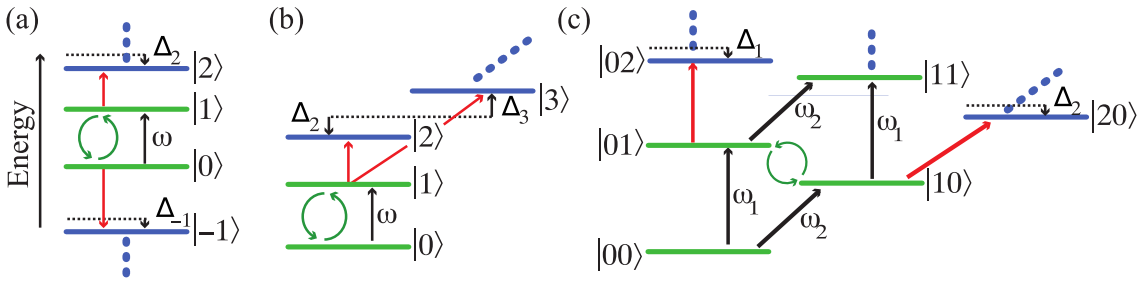

Figure 9.1: In (a) we consider the case when the qubit has leakage from both its $|0\rangle$ and $|1\rangle$ level. In (b) we consider the case when the qubit has leakage from its $|1\rangle$ level to more then one higher level. In (c) two anharmonic oscillators are coupled with $|11\rangle$ leaking to both $|02\rangle$ and $|20\rangle$.

where x and p are coordinate and momentum, respectively, and ω_0 is the resonance frequency, in the limit $\delta \to 0$. Even though we focus on the hard case, $\delta > 0$, we expect similar conclusions to hold for the soft case. This Hamiltonian can be made dimensionless with raising and lowering operators analogous to the harmonic oscillator [31]. It then reads

$$H = \hbar\omega_0 \left(a^\dagger a + \frac{1}{2}\right) + \frac{\hbar\delta}{12}(a + a^\dagger)^4. \tag{9.2}$$

9.1.1 SNO approximation

By expanding the nonlinearity and ordering its terms, we can split the Hamiltonian as $H = H_0(\hat{n}) + \hat{V}$ into a term H_0 that contains only the number operator $\hat{n} = a^\dagger a$ that will dominate perturbation theory for small $\delta/\omega \ll 1$ and corresponds to the rotating wave approximation, and a term $\hat{V}$ that goes beyond that approximation as it contains different powers of a and $a^\dagger$. We find

$$H_0(\hat{n}) = \hbar\omega_0 \left(\hat{n} + \frac{1}{2}\right) + \frac{\hbar\delta}{4}(2\hat{n}^2 + 2\hat{n} + 1). \tag{9.3}$$

$$V = \frac{\delta}{12}(a^{\dagger 4} + 6a^{\dagger 2} + 4a^\dagger(a^2 + a^{\dagger 2})a + 6a^2 + a^4). \tag{9.4}$$

H_0 can be re-parameterized to a simpler form using $\omega = \omega_0 + \delta$, while dropping a constant energy shift

$$H_0 = \hbar\left[\omega\hat{n} + \tfrac{1}{2}\delta\hat{n}(\hat{n} - 1)\right]. \tag{9.5}$$

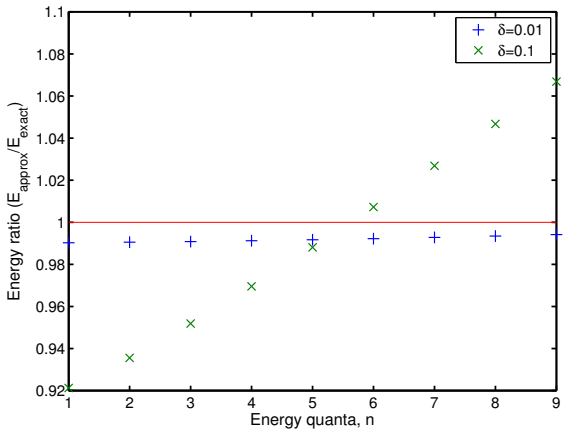

Figure 9.2: Ratio of the eigen-energies of the RWA Hamiltonian, Eq. 9.5 and exact eigenergies of the anharmonic oscillator for the first nine energy levels, assuming nonlinearities $\delta = 0.01$ and $\delta = 0.1$.

where ω is the frequency of the $|0\rangle \leftrightarrow |1\rangle$ transition. H can be diagonalized straightforwardly and its eigenstates $|n\rangle$, $n = 0, 1, 2, \ldots$ are close to the harmonic oscillator eigenstates, see fig. 9.2, at small δ and the eigen-energies are given by the polynomial deriving from Eq. (9.5), $E_n \simeq H_0(n)$. We can note that for small δ, which is the most challenging control problem, the energy splittings between adjacent levels are only slightly different from each other. Denoting the transition frequencies as $\omega_{ij} = (E_i - E_j)/\hbar$, this means that $\omega_{n+1,n} = \omega + \delta n$ only has a weak n dependence by $\delta \ll \omega$.

In the most common physical situation, external driving is given by

$$H_c(t) = F(t)x = \hbar f(t)(a + a^\dagger) \tag{9.6}$$

corresponding to current drive in the Josephson case, a laser field in the ion trap case and an external force in the nano-mechanical case. Here, $f(t)$ is proportional to F [31]. Given that the eigenstates are almost harmonic oscillator eigenstates, this drive predominantly, in the sense of an approximate selection rule, couples adjacent levels, n and $n + 1$.

This setting outlines two difficulties in the task of preparing Fock states starting from the ground state using a straightforward application of H_c. Firstly, there is no direct transition from the ground state to any state higher than $|1\rangle$ unless δ is large enough to make the eigenstate corrections due to $\hat{V}$, Eq. (9.4), significant. Secondly, even for preparing the first excited state, the approximate Liouvillian degeneracy

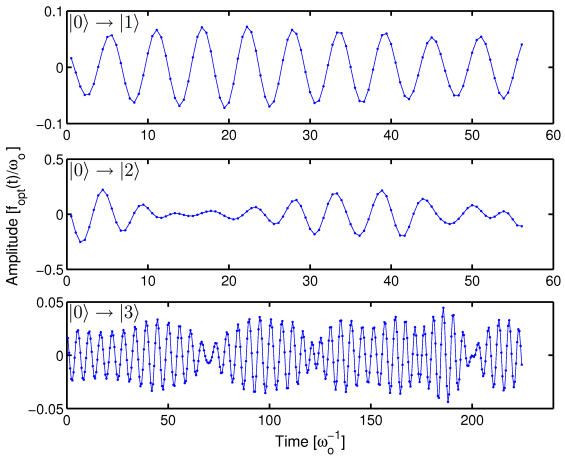

Figure 9.3: The amplitudes of three pulse sequences optimized by GRAPE for preparation of $|1\rangle$ (top), $|2\rangle$ (middle), and $|3\rangle$ (bottom) are shown. Each system had a Hilbert space consisting of 10 eigenstates, with $\delta = 0.12$. Each optimized pulse leads to a fidelity of 0.9999.

(i.e., the weak n-dependence of $\omega_{n+1,n}$, prohibits short pulses: A short resonant π-pulse of length T_r has, by the energy-time uncertainty principle of the Fourier transform, a bandwidth of $\omega_B \simeq \pi/T_r$. If ω_B becomes comparable to the small parameter $|\omega_{n+1,n} - \omega_{n,n-1}| = \delta$, a pulse applied to $\omega_{n,n-1}$ will have a significant Fourier component at $\omega_{n+1,n}$. Hence, driving the transition from $n-1$ to n will inevitably also drive n to $n+1$ inducing leakage to the next higher state $n+1$. This could be overcome by a very long π-pulse that keeps ω_B small, however, in a realistic setting, this will compete with energy relaxation back to the ground state at a rate T_1, effectively limiting $\omega_B \gg 1/T_1$.

9.1.2 Numerical solutions

We have analyzed this setting numerically in a number of cases using the methodology of Ch. 5.1. In all cases, the initial state was chosen to be the ground state, $|\psi_i\rangle = |0\rangle$ and the first, second, and third excited state where chosen as final states, $|\psi_f\rangle = |n = 1, 2, 3\rangle$. We have truncated the energy spectrum at $n = 10$ levels and verified, that changing the number of levels did not lead to a discernible change in the pulse. In Fig. 9.3 we show optimized pulses that display clear beating behavior.

This beating becomes more complex as we go to higher states. More clearly, this structure can be analyzed by its Fourier transform, as shown in fig. 9.4 A.

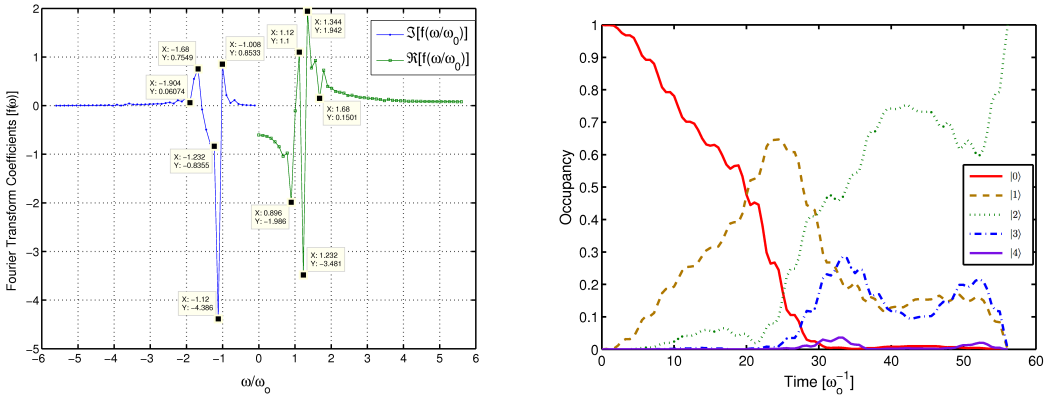

Figure 9.4: A)Fourier spectrum for preparing $|2\rangle$. To the left and right of the zero frequency is the imaginary and real coefficients, respectively, of the decomposition. The main peaks in each pulse spectrum is labeled. The transition frequencies between $|0\rangle$ and $|1\rangle$ ($\omega = 1.12$) are apparent in each pulse. The transition frequency $\omega_{21} = 1.24$ and $\omega_{32} = 1.36$ is apparent in the pulse preparing the $|2\rangle$ and $|3\rangle$ states. B)Populations of the different states during the pulse taking $|0\rangle$ into $|2\rangle$. It is recognized, that $|0\rangle$ is first excited into $|1\rangle$ and that this one then goes into $|2\rangle$ whereas higher states are somewhat occupied during the pulse and then go to zero. Additional fast modulation is due to counter-rotating terms.

We observe a cluster of close discrete lines, confirming the observation that the pulses are beatings. The predominant frequencies in the $|0\rangle \rightarrow |2\rangle$ transition are $\omega_{10} = 1.12$, the slightly higher $\omega_{21} = 1.24$ and $\omega_{32} = 1.36$, with additional sidebands. Analyzing more transitions, we see the same picture, where a transition from $|0\rangle$ to $|n\rangle$ contains frequencies $\omega_{m+1,m}$ for $0 \leq m \leq n$. The state populations during the pulse can be seen in Fig. 9.4B.

This structure can be understood as follows: Due to the oscillator selection rule, the ladder of states needs to be climbed sequentially, it is not possible to go from $|n\rangle$ to $|n+k\rangle$ for any $k > 1$, explaining the frequencies ω_{10} through $\omega_{n,n-1}$ for a transition $0 \rightarrow n$ that correspond to driving π-pulses on those respective transitions. This is clearly visible by the population of $|1\rangle$ in Fig. 9.4B. Due to the bandwidth issue discussed in section 9.1.1, this also drives the transition to level $n+1$, it is also necessary to drive the leakage transition leading out of $|n\rangle$, i.e., at $\omega_{n,n+1}$.

This analysis confirms the picture of sequential occupation of Fock states, climbing an energy ladder, with some intermediate leakage to higher state that corrects itself in the end. Next to these long-scale Rabi dynamics, there is a fast modula-

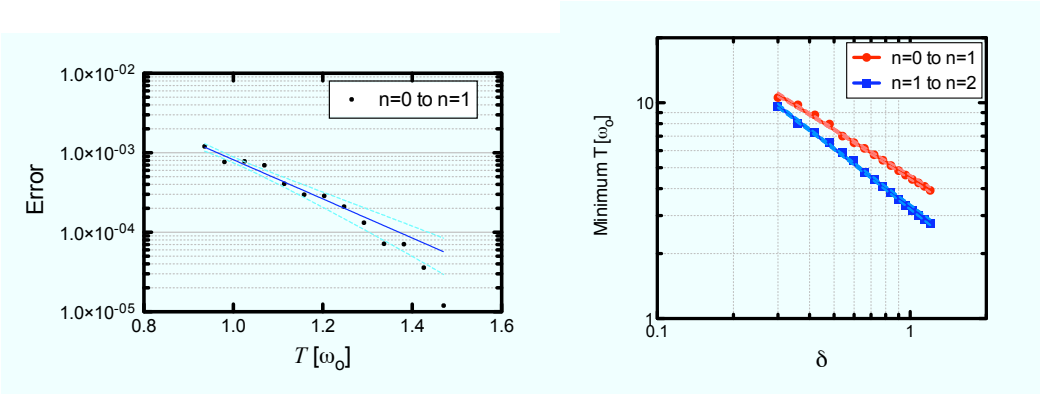

Figure 9.5: (A) Error for a pulse aiming at creating the first excited state for $\delta = 0.12$ along with an exponential fit (solid) and 95% confidence lines (dashed). (B) Scaling of the minimal gate time with the nonlinearity parameter with power-law fits. Exponents are -0.73 for $|0\rangle \rightarrow |1\rangle$ and -0.90 for $|1\rangle \rightarrow |2\rangle$.

tion on the scale of the average driving frequency typical for strong driving where the counter-rotating component of the drive that is not part of standard Rabi physics [31] becomes important. In the end of the pulse, populations of $|1\rangle$ and $|3\rangle$ are brought back to $|2\rangle$.

As mentioned in Sec. 8.1, nonlinearity is a resource for the creation of Fock states. At very short pulse durations t_g state transfer pulses must scale with a constant area, i.e., their amplitudes will scale as $1/t_g$. Thus, if the pulse is too short the driving amplitude will be so high that the δ-term in the Hamiltonian can be neglected and Fock state preparation will not be possible. We have numerically investigated the gate fidelity as a function of pulse duration and found the error to be roughly exponentially growing at short times, see Fig. 9.5A.

Next we investigate the minimal time to reach a fidelity of 0.99999% for pulses $|0\rangle \rightarrow |1\rangle$ and $|1\rangle \rightarrow |2\rangle$, Fig. 9.5B. We see that this minimal time is a power law of the nonlinearity parameter δ; $t_{\min} \propto \delta^{\alpha}$ with $\alpha_{01} = -0.73 \pm 0.029$ and $\alpha_{12} = -0.90 \pm 0.031$

As a reference, we can construct the corresponding power law for non-optimized, single frequency Rabi pulses, starting from the $|0\rangle \rightarrow |1\rangle$ transfer. Transforming to the frame rotating with the driving frequency, the effect of the pulse on the unwanted $|1\rangle \rightarrow |2\rangle$ transition is to switch a transition matrix element on and off. This matrix element has to be compared to the detuning between the transition $|\omega_{21} - \omega_{10}| \simeq \delta$. In order to create a Rabi pulse of fixed area π, the maximal amplitude of the transition matrix elements scales as $1/t_g$, and its rate of change thus scales as $1/t_g^2$. In order to not occupy the $|2\rangle$-state in the end of the pulse, the dynamics of the $|1\rangle \leftrightarrow |2\rangle$-

transition must be fully adiabatic [122], i.e., following a standard result from Landau-Zener theory [98, 167, 188], we demand $1/t_g^2 \ll \delta^2$ leading to a minimal gate time of $t_g \propto 1/\delta$. We would not expect this scaling law to change for the $0 \rightarrow 2$ transition, which is ultimately a sequence of two transitions of length $1/t_g$, only a different prefactor.

Thus, we can conclude that our optimal pulses qualitatively extend the limits of Fock state preparation by changing the minimal time to a softer power law, from $1/\delta$ to $1/\delta^{0.73\pm0.029}$. One can conclude that this is due to quantum interference in the higher levels. The difference is made possible by temporarily occupying higher states and unpopulating them in the course of the pulse. It needs to be remarked, that our pulses qualitatively differ from quantum gate pulses, i.e., rotations of the full basis instead of changing state [122, 148], where a $1/\delta$ scaling of the minimal time is found. This suggests that state preparation, which is less constrained, can use quantum interference more efficiently. A related paper has been posted [80] that looks at preparation of $|1\rangle$ exclusively and finds different envelopes, probably due to additional constraints.

9.1.3 Analytical solutions

We consider a qubit defined in the intermediate states of an anharmonic oscillator [Fig. 9.1 (a)]. This situation is important if the anharmonic oscillator is going to be used for qudit logic, as done in Ref. [128], or for state tomography of the qudit, as done in Ref. [13]. The situation is also more constrained than Sec. 9.1.2 since we consider logical gates rather than state preparation (though the logical gates can be used for state preparation). We rewrite the free Hamiltonian Eq. 9.5 and control Hamiltonian Eq. 9.6 as

$$H_{\rm fr} = \sum_{j=-N}^{N} (j\omega + \Delta_j)\Pi_j, \tag{9.7}$$

$$H_{\rm ct}(t) = \mathcal{E}(t) \sum_{j=-N+1}^{N} \lambda_{j-1}\sigma^x_{js-1,j}, \tag{9.8}$$

where $N = (d-1)/2$ and again we take $\Delta_0 = \Delta_1 = 0$. Moving to a interaction frame similar to Eq. (8.7) (the sum range is changed to be consistent with the above) we

find a dimensionless rotating frame Hamiltonian

$$
\begin{aligned}
H_0 &= \sum_{j=-N,\neq 0,1}^{N} \frac{\Delta_j}{\Delta_2}\Pi_j, && (9.9)\\
H_z &= \sum_{j=-N}^{N} j\Pi_j, && (9.10)\\
H_x &= \sum_{j=-N+1}^{N} \lambda_{j-1}\sigma^x_{j-1,j}, && (9.11)\\
H_y &= \sum_{j=-N+1}^{N} \lambda_{j-1}\sigma^y_{j-1,j}. && (9.12)
\end{aligned}
$$

To find the first order corrections, we follow a similar procedure to Sec. 8.6. The frame constraints for Eq. 8.31 to remove leakage become

$$
\begin{aligned}
s^{(1)}_{x,0,k}(t) =& 0, \quad s^{(1)}_{y,0,k}(t) = -\frac{\lambda_{-1}\Delta_2 t_g \Omega_G(t)\delta_{k,-1}}{2\Delta_{-1}},\\
s^{(1)}_{x,1,k}(t) =& 0, \quad s^{(1)}_{y,1,k}(t) = -\frac{\lambda_1 t_g \Omega_G(t)\delta_{k,2}}{2},
\end{aligned}
\tag{9.13}
$$

and the dimensionless first order control fields are

$$
\begin{aligned}
\bar{\Omega}^{(1)}_x(t) &= 2\dot{s}^{(1)}_{x,0,1}(t), && (9.14)\\
\bar{\Omega}^{(1)}_y(t) &= 2\dot{s}^{(1)}_{y,0,1}(t) - s^{(1)}_{z,1}(t) t_g \Omega_G(t), && (9.15)\\
\bar{\delta}^{(1)}(t) &= \dot{s}^{(1)}_{z,1}(t) + 2s^{(1)}_{y,0,1}(t) t_g \Omega_G(t) && (9.16)\\
&\quad + \frac{t_g^2 \Omega_G^2(t)}{4}\left(\lambda_1^2 - \tfrac{\Delta_2}{\Delta_{-1}}\lambda_{-1}^2\right). && (9.17)
\end{aligned}
$$

From this we find the control fields for Z-only correction are

$$
\begin{aligned}
\Omega_x(t) =& \Omega_G(t), \quad \Omega_y(t) = 0,\\
\delta(t) =& \frac{\Omega_G^2(t)}{4}\left[\tfrac{\lambda_1^2}{\Delta_2} - \tfrac{\lambda_{-1}^2}{\Delta_{-1}}\right],
\end{aligned}
\tag{9.18}
$$

the Y-only are

$$
\begin{aligned}
\Omega_x(t) =& \Omega_G(t), \quad \Omega_y(t) = -\frac{\dot{\Omega}_G(t)}{4}\left[\tfrac{\lambda_1^2}{\Delta_2} - \tfrac{\lambda_{-1}^2}{\Delta_{-1}}\right],\\
\delta(t) =& 0,
\end{aligned}
\tag{9.19}
$$

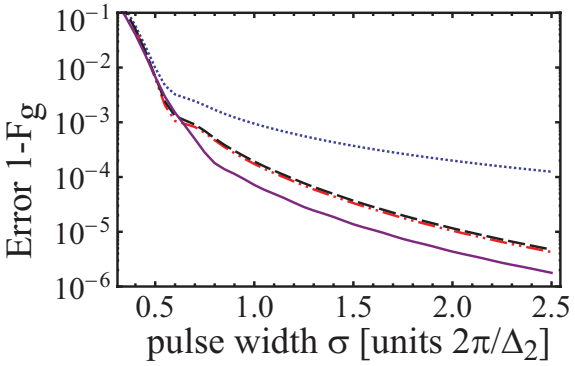

Figure 9.6: Gate error for the implementation of a NOT gate when there is leakage above and below the qubit subspace. The system considered is explained in the text. The blue dotted line is the zeroth order solution. The black dash line is the first order Z-only correction. The red dash-dot-dot line is the first order Y-only correction. The purple solid line is for the optimal first order correction.

and the optimal first order control field corrections (after minimizing second order terms) are

$$\Omega_x(t) = \Omega_G(t), \tag{9.20}$$

$$\Omega_y(t) = -\frac{\dot{\Omega}_G(t)}{2\Delta_2}\sqrt{\lambda_1^2 + \frac{\Delta_2^2}{\Delta_{-1}^2}\lambda_{-1}^2}, \tag{9.21}$$

$$\delta(t) = \frac{\Omega_G^2(t)}{4\Delta_2}\left[\lambda_1^2 - \frac{\Delta_2^2}{\Delta_{-1}^2}\lambda_{-1}\right]. \tag{9.22}$$

To demonstrate an improvement over the zeroth order solution we consider a SNO (of $d = 6$) where we want to control the $2 \rightarrow 3$ transition. In this case, we relabel $j = 2$ to 0 and so on, rescaling the coupling so that the new $0 \rightarrow 1$ transition is unity. This results in setting $\Delta_3 = 3\Delta_2$, $\Delta_{-1} = \Delta_2$, $\Delta_{-2} = 3\Delta_2$, and $\lambda_0 = 1$, $\lambda_1 = \sqrt{4/3}$, $\lambda_2 = \sqrt{5/3}$ $\lambda_{-1} = \sqrt{2/3}$, $\lambda_{-2} = \sqrt{1/3}$. In Fig. 9.6, the gate error for implementing a NOT gate is shown as a function of σ for the same Gaussian shaped pulse as considered in Sec. 8.4. Here we observe that the DRAG technique improves the gate fidelities substantially when compared to the zeroth order solution.

Finally, note that because of the Liouvillian degeneracy, $\Delta_{-1} = \Delta_2$, this system is a prime candidate for using a second derivative pulse (Sec. 7.1.3.5) which is symmetric around the driven transition. This may be addressed in future research.

9.2 Multiple leakages from one level

In the second case shown in Fig. 9.1 B, we consider a qubit made from the lowest 2 levels of a system which is coupled to many other transitions, all transitions having only a small energy cost (approximately Δ_2). This is an interesting example as it shows how this theory can be easily generalized. In this case we write the free and coupling Hamiltonians as

$$H_{\rm fr} = \omega\Pi_1 + \sum_{j=2}^{d-1}(2\omega + \Delta_j)\Pi_j, \tag{9.23}$$

$$H_{\rm ct}(t) = \mathcal{E}(t)\left(\lambda_0\sigma_{0,1}^x + \sum_{j=2}^{d-1}\lambda_{j-1}\sigma_{1,j}^x\right). \tag{9.24}$$

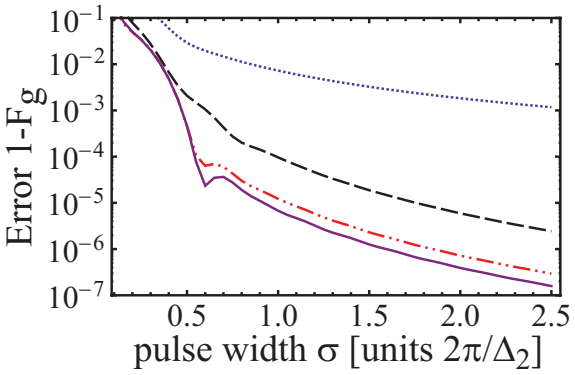

Figure 9.7: Gate error for the implementation of a NOT gate when there are many leakage transitions for the excited state. The system considered is explained in the text. The blue dotted line is the zeroth order solution. The black dash line is the first order Z-only correction. The red dash-dot-dot line is the first order Y-only correction. The purple solid line is for the optimal first order correction.

To eliminate the fast degrees of freedom we move to a rotating frame and make the standard rotating wave approximation. The procedure is similar to Sec. 8.6 with the replacement of Eq. (8.7) by

$$R(t) = \exp\left(-i\omega_d t\right)\Pi_1 + \sum_{j=2}^{d-1}\exp\left(-i2\omega_d t\right)\Pi_j. \tag{9.25}$$

This results in a dimensionless rotating frame Hamiltonian equivalent to Eq. (9.24) with

$$\begin{aligned} H_0 &= \sum_{j=2}^{d-1} \frac{\Delta_j}{\Delta_2} \Pi_j, \quad H_z = \Pi_1 + \sum_{j=2}^{d} 2\Pi_j, \\ H_x &= \lambda_0 \sigma_{0,1}^x + \sum_{j=2}^{d-1} \lambda_{j-1} \sigma_{1,j}^x, \quad H_y = \lambda_0 \sigma_{0,1}^x + \sum_{j=2}^{d-1} \lambda_{j-1} \sigma_{1,j}^y. \end{aligned}$$

Again, we find the zeroth order controls given by Eq. (3.3). To find the first order corrections, we find the first order frame transition to be

$$\begin{aligned} s_{x,0,k}^{(1)}(t) =& 0, \quad s_{y,0,k}^{(1)}(t) = 0, \\ s_{x,1,k}^{(1)}(t) =& 0, \quad s_{y,1,k}^{(1)}(t) = -\frac{\lambda_{k-1} t_g \Omega_G(t) \Delta_2}{2\Delta_k}. \end{aligned} \tag{9.26}$$

This gives the dimensionless first order control fields

$$\begin{aligned} \bar{\Omega}_x^{(1)}(t) =& 2\dot{s}_{x,0,1}^{(1)}(t), \\ \bar{\Omega}_y^{(1)}(t) =& 2\dot{s}_{y,0,1}^{(1)}(t) - s_{z,1}^{(1)}(t) t_g \Omega_G(t), \\ \bar{\delta}^{(1)}(t) =& \dot{s}_{z,1}^{(1)}(t) + 2 s_{y,0,1}^{(1)}(t) t_g \Omega_G(t) \\ &+ \frac{t_g^2 \Omega_G^2(t) \Delta_2}{4} \left(\sum_{k=2}^{d-1} \frac{\lambda_{k-1}^2}{\Delta_k} \right). \end{aligned} \tag{9.27}$$

From this, we find the control fields for the Z-only correction are

$$\begin{aligned} \Omega_x(t) =& \Omega_G(t), \quad \Omega_y(t) = 0, \\ \delta(t) =& \frac{\Omega_G^2(t)}{4} \sum_{k=2}^{d-1} \frac{\lambda_{k-1}^2}{\Delta_k}, \end{aligned} \tag{9.28}$$

for the Y-only correction are

$$\begin{aligned} \Omega_x(t) =& \Omega_G(t), \quad \Omega_y(t) = -\frac{\dot{\Omega}_G(t)}{4} \sum_{k=2}^{d-1} \frac{\lambda_{k-1}^2}{\Delta_k}, \\ \delta(t) =& 0, \end{aligned} \tag{9.29}$$

and the optimal first order control field corrections (after minimizing $H_{\text{extra}}^{(2)}(t)$) are

$$\begin{aligned} \Omega_x(t) =& \Omega_G(t), \quad \Omega_y(t) = -\frac{\dot{\Omega}_G(t)}{2\Delta_2} \sqrt{\sum_{k=2}^{d-1} \frac{\Delta_2^2 \lambda_{k-1}^2}{\Delta_k^2}}, \\ \delta(t) =& \frac{\Omega_G^2(t)}{4\Delta_2} \left[\sum_{k=2}^{d-1} \frac{\Delta_2^2 \lambda_{k-1}^2}{\Delta_k^2} - 2\sqrt{\sum_{k=2}^{d-1} \frac{\Delta_2^2 \lambda_{k-1}^2}{\Delta_k^2}} \right]. \end{aligned} \tag{9.30}$$

We note that these solutions are identical to the previous solutions with a single leakage channel where $\lambda_1 = \tilde{\lambda}$,

$$\tilde{\lambda} \equiv \sqrt{\sum_{k=2}^{d-1} \frac{\Delta_2^2 \lambda_{k-1}^2}{\Delta_k^2}}. \tag{9.31}$$

To numerically demonstrate an improvement over the zeroth order solution, we consider the implementation of a NOT gate for a $d = 6$ system with $\lambda_j = 1$ for all j and $\Delta_3 = 2\Delta_2$, $\Delta_4 = 3\Delta_2$, $\Delta_5 = 4\Delta_2$ (note these are different parameters from the anharmonic oscillator considered in Sec. 9.1.1). The results are plotted in Fig. 9.7, where again, it is clearly seen that the DRAG technique improves the zeroth order solution.

9.3 2-Qubit leakage

9.3.1 General Model

The system that we consider consists of two coupled qubits. Each qubit is made up of the ground and first excited states of a slightly anharmonic oscillator (see Sec. 6.7). We examine the lowest three levels of the spectrum; the third level represents leakage. This results in a single qubit Hamiltonian given by

$$H_{qubit} = \hbar \sum_{i=1,2} (\omega_i^1 \Pi_i^{(1)} + \omega_i^2 \Pi_i^{(2)}) \tag{9.32}$$

where we have projectors $\Pi_j^{(i)} = |j\rangle\langle j|_i$ and transition energies $\hbar\omega_j$. The coupling takes the form of voltage coupling, so that the qubit interaction term can be written as

$$H_{int} = J(t)\left(\sigma_{0,1}^{x,1} + \lambda_1 \sigma_{1,2}^{x,1}\right)\left(\sigma_{0,1}^{x,2} + \lambda_2 \sigma_{1,2}^{x,2}\right) \tag{9.33}$$

where $\sigma_{i,j}^{x,k} = |i\rangle\langle j|_k + |j\rangle\langle i|_k$ and λ_k measures the relative strength of the 1-2 transition to the 0-1 transition for the k^{th} qubit. The parameter $J(t)$ is the coupling strength between the qubits. Control of the system is represented by this variable coupling term. So, in the lab frame our Hamiltonian is

$$H = \hbar \sum_{i=1,2} (\omega_i^1 \Pi_i^{(1)} + \omega_i^2 \Pi_i^{(2)}) + J(t)\left(\sigma_{0,1}^{x,1} + \lambda_1 \sigma_{1,2}^{x,1}\right)\left(\sigma_{0,1}^{x,2} + \lambda_2 \sigma_{1,2}^{x,2}\right) \tag{9.34}$$

By making the rotating wave approximation (RWA), we can write the effective Hamiltonian as

$$H_{\text{eff}} = H_{\text{drift}} + V(t) \tag{9.35}$$

where $V(t)$ is the effective coupling term and H_{drift} is the effective single qubit Hamiltonian. $V(t)$ and H_{drift} are given by

$$V(t) = J(t)\Big[a_1 a_2^\dagger + h.c.\Big] \qquad \text{for } a_i = |0\rangle\langle 1|_i + \lambda|1\rangle\langle 2|_i \tag{9.36}$$

$$H_{\text{drift}} = \Delta\Pi_2^{(1)} + \Delta\Pi_2^{(2)} \tag{9.37}$$

where $\Delta = \omega_2^i - 2\omega_1^i$ is the anharmonicity of the qubits, assumed to be the same for both qubits $i = 1, 2$ which are brought on resonance, and we have set $\hbar = 1$. We have also assumed that $\lambda_1 = \lambda_2 = \lambda$. The subspaces in H are coupled by $V(t)$ in the following way:

$$\begin{aligned} &|01\rangle \quad \text{to} \quad |10\rangle \\ &|12\rangle \quad \text{to} \quad |21\rangle \\ &|11\rangle \quad \text{to} \quad |02\rangle, |20\rangle \quad \text{(coupling across subspaces)} \end{aligned}$$

So, we can write V(t) in terms of a degenerate subspace and another subspace that represents leakage: $V(t) = V_{leak} + V_{deg}$

$$\text{where } V_{leak} = J(t)\Big[\lambda|02\rangle\langle 11| + \lambda|11\rangle\langle 20| + \text{h.c.}\Big] \tag{9.38}$$

$$V_{deg} = J(t)\Big[|01\rangle\langle 10| + \lambda^2|12\rangle\langle 21| + \text{h.c.}\Big] \tag{9.39}$$

In the absence of the third (leakage) level, the coupling terms between the qubits generate an ideal ISWAP gate. Namely, if

$$H = V_{deg} = J\Big[|01\rangle\langle 10| + \lambda^2|12\rangle\langle 21| + h.c.\Big] \qquad \text{with } J = \frac{\pi}{2T} \tag{9.40}$$

a perfect ISWAP unitary propagator is formed $U = \mathcal{T}e^{-i\int_0^T H dt} = e^{-iHT}$ (where T is the gate time). So, if Δ is large, in which case the effect of the third level is essentially averaged out over the gate, very high fidelity can be obtained using a constant coupling strength of $\frac{\pi}{2T}$. For example, for $\Delta = 10$ GRad/s (= 1.59 GHz) and a gate time of 10 ns the fidelity of the ISWAP gate produced with this constant coupling is $\mathcal{F} = 0.99814$. For more practical anharmonicities in the vicinity of 300 MHz, the fidelity drops significantly due to leakage. For example, for $\Delta = 2.2$ GRad/s (= 350 MHz) and a gate time of 10 ns the fidelity is $\mathcal{F} = 0.947449$; doubling the gate time to 20 ns only increases the fidelity to $\mathcal{F} = 0.990626$.

9.3.2 Numerical solution

As a starting point for the GRAPE algorithm, we use a constant coupling $J(t) = \frac{\pi}{2T}$ and choose a pixel width of 0.01 ns. A sample result can be seen in Fig. (9.8), which illustrate the characteristics of solutions for this system found by GRAPE. The solutions exhibit an oscillatory term superimposed onto the constant coupling component $\frac{\pi}{2T}$. In most cases the sinusoidal term has frequency larger than Δ by a small offset and is also slightly amplitude modulated.

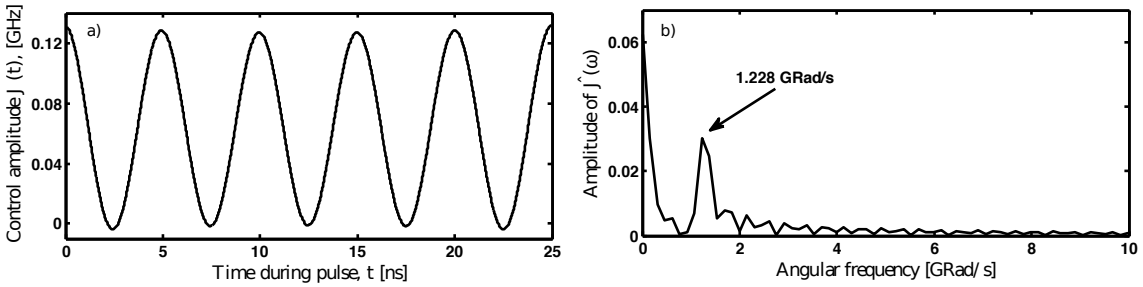

Figure 9.8: (a) Control amplitude (inter-qubit coupling strength) $J(t)$ as a function of time for two coupled qubits both having an anharmonicity of $\Delta = 1$ GRad/s (= 159 MHz). The pulse is obtained numerically using GRAPE with a pixel width of 0.01 ns; the ISWAP gate is produced with fidelity $\mathcal{F} = 0.99997$. This is a significant improvement over the fidelity for a constant coupling strength: $\mathcal{F}_0 = 0.97$. (b) Fourier spectrum (magnitude) of the pulse in (a) showing a peak at frequency $\omega =$ 1.228 GRad/s (= 195 MHz) corresponding to the oscillation frequency of the coupling strength.

Neglecting the small amplitude modulation seen in the GRAPE pulses, we look for solutions of the form:

$$J(t) = \frac{\pi}{2T} + A\cos\Big((\Delta + \epsilon)t\Big) \tag{9.41}$$

By finding optimal solutions, we see that the relative scaling factor does not seem to be present, so $A = \frac{\pi}{2T}$. We also find that increasing the frequency offset ϵ reduces the gate time T necessary for a high fidelity ISWAP gate. However, for each Δ there is a unique $\epsilon = 0.2398\Delta$ yielding a global maximum of the fidelity. These relationship is plotted in Fig. (9.9) which give a clear prescription for removing the leakage error.

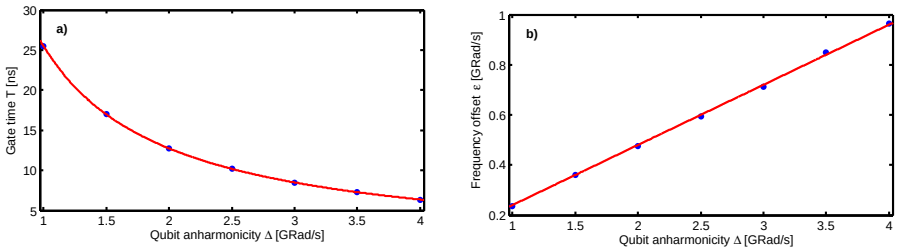

Figure 9.9: For each of the different Δ in these two plots, the frequency offset ϵ and the gate time T are found numerically to produce maximum fidelity (within a reasonable gate time). The fidelities of these points range from $\mathcal{F} = 0.999948$ to 0.999961 which is a large improvement over the range for a constant coupling strength: $\mathcal{F}_0 = 0.9709$ to 0.9713. The resolution is 0.003 GRad/s (= 0.5) MHz for each optimal ϵ and is 0.05 ns for each optimal gate time. (a) Optimal gate time T versus the intrinsic anharmonicity Δ of the qubits (assumed equal) for a number of different Δs (blue point markers). The red line is a fit line of the form: $T(\Delta) = \frac{25.49}{\Delta}$; the root mean squared error (RMSE) of the fit is 0.01719. (b) Optimal frequency offset ϵ versus anharmonicity Δ for a number of different Δ's (blue point markers). The red line is a fit line of the form: $\epsilon(\Delta) = 0.2398\Delta$ (RMSE is 0.006953).

9.3.3 Analytical solution

We wish to motivate the form of solutions found numerically. First of all, one notices that the optimal $T(\Delta) \approx \frac{8\pi}{\Delta}$ so that the leakage transition has minimal spectral weight, $\int_0^T Je^{-i\Delta t}dt \approx 0$. Adding in the second tone at frequency $\Delta + \epsilon$ also does not change the spectral weights of the two transitions as we can see from the two graphs that $\epsilon = 6.11/T \approx 2\pi/T$ and $\Delta + \epsilon \approx 10\pi/T$, giving $\int_0^T J\cos(\epsilon)dt \approx 0$ and $\int_0^T J\cos(\Delta+\epsilon)e^{-i\Delta t}dt \approx 0$. Moreover, since the solutions are constant amplitude solutions, we can neglect the derivative of the envelope function as we did in e.g. Sec. (7.1.1). In fact, if we did have solutions that were not constant in amplitude we would run into a Liouvillian degeneracy as in Sec. (9.1.1) which would deteriorate the quality of the pulses. Thus, with the absence of selectivity or shaping error we expect the main errors and corrections to come from phase accumulation.

To see this we look at the leakage term V_{leak} when $J(t) = J$. Applying a transformation $A = \exp(-iV_{\text{leak}}t)$ to the Hamiltonian gives to first order in the inverse anharmonicity

$$H^A = H_{\text{drift}} + V_{\text{deg}}(t) + \frac{\lambda^2 J^2}{\Delta}(2|11\rangle\langle 11| + |02\rangle\langle 02| + |20\rangle\langle 20|) + O(\frac{1}{\Delta^2}) \qquad (9.42)$$

We want to cancel the phase error (i.e. ac-Stark shift) in the qubit's subspace

$$\frac{2\lambda^2 J^2}{\Delta}|11\rangle\langle 11|$$

Note that the effect is entangling and cannot be canceled with single qubit gates. Instead we expect that adding the second tone $J(t) = J\cos(\Delta + \epsilon)$ has exactly this needed effect. To see this we can move to the frame rotating at frequency $\Delta + \epsilon$ via the rotating frame $R = \exp(-i(\Delta + \epsilon)(\Pi_1^{(1)} + \Pi_1^{(2)} + 2\Pi_2^{(1)} + 2\Pi_2^{(2)})t)$ which gives

$$\begin{aligned} H^R &= J(1 + \cos[(\Delta + \epsilon)t])e^{-i(\Delta+\epsilon)t}(|01\rangle\langle 10| + \lambda|02\rangle\langle 11| + \lambda|11\rangle\langle 20| + \lambda^2|12\rangle\langle 21|) + h.c. \\ &\quad -(\Delta + \epsilon)\Pi_1^{(1)} - (\Delta + \epsilon)\Pi_1^{(2)} - \epsilon\Pi_2^{(1)} + \epsilon\Pi_2^{(2)} \end{aligned}$$

Expanding the cosine into a sum of exponentials and dropping terms with frequency $2(\Delta + \epsilon)$ we can simplify this somewhat to

$$\begin{aligned} H^R &\cong J(e^{-i(\Delta+\epsilon)t} + \frac{1}{2})(|01\rangle\langle 10| + \lambda|02\rangle\langle 11| + \lambda|11\rangle\langle 20| + \lambda^2|12\rangle\langle 21|) + h.c. \\ &\quad -(\Delta + \epsilon)\Pi_1^{(1)} - (\Delta + \epsilon)\Pi_1^{(2)} - \epsilon\Pi_2^{(1)} + \epsilon\Pi_2^{(2)} \end{aligned}$$

From this frame it is possible to deduce the effect of the extra term on the qubit's subspace, which can be seen via the transformation $B = \exp(i(-\frac{iJ}{\Delta+\epsilon}|01\rangle\langle 10| - \frac{i\lambda J}{\epsilon}|02\rangle\langle 11| - \frac{i\lambda J}{\epsilon}|11\rangle\langle 20| + \text{h.c.})t)$ to give

$$\begin{aligned} H^B &= J(e^{-i(\Delta+\epsilon)t})(|01\rangle\langle 10| + \lambda|02\rangle\langle 11| + \lambda|11\rangle\langle 20| + \lambda^2|12\rangle\langle 21|) + h.c \qquad (9.43) \\ &\quad -(\Delta + \epsilon)\Pi_1^{(1)} - (\Delta + \epsilon)\Pi_1^{(2)} - \epsilon\Pi_2^{(1)} + \epsilon\Pi_2^{(2)} \\ &\quad -\frac{\lambda^2 J^2}{4\epsilon}(2|11\rangle\langle 11| + |02\rangle\langle 02| + |20\rangle\langle 20|) \\ &\quad -\frac{J^2}{4(\Delta + \epsilon)}(|01\rangle\langle 01| - |10\rangle\langle 10|) + O(\frac{1}{\Delta^2}) \end{aligned}$$

Comparing this equation with Eq. (9.42)we see that the necessary condition for canceling the phase error on $|11\rangle\langle 11|$ is

$$4\epsilon = \Delta.$$

This corresponds very closely to the relationship found in Fig. (9.9)B, with the slight deviation explainable by higher order corrections. Note that there is also a shift of the $|01\rangle\langle 01|$ and $|10\rangle\langle 10|$ levels but these are suppressed by a factor of $\frac{\Delta}{8\lambda^2(\Delta+\epsilon)}$ which for typical values is about 20 times smaller than the original phase shift, with a similar decrease in the error. In fact, if we include this term in the $|11\rangle\langle 11|$ shift, and choose $4.05\epsilon = \Delta$ this comes even closer to the optimal value. This method can also easily be expanded to the case where the anharmonicities of the qubits are different.

Chapter 10

Virtual Transitions

We have seen that errors in performing gate operations can arise from non-idealities in the driving fields and from complications in the Hilbert space of the qubits that go beyond a 2-level approximation. It turns out that complications in the Hilbert space can also be beneficial. Frame transformations can be used to couple energy levels that do not ordinarily contain a matrix element in the physical lab frame (see e.g. Fig. 10). The advantage of such a situation is that if no matrix element is contained between the qubit levels then relaxation can in principle be greatly suppressed. In other cases, the virtual transition can be used where selectivity conditions would make driving the individual transitions of which it is composed impossible. For virtual transitions, multiple photons are used, with the condition that their energies add up to the energy splitting of the driven transition for energy conservation to hold. It is conventional wisdom to perform an adiabatic approximation for these transitions that any transition that does not conserve energy will be greatly suppressed. This approximation will be quantified in an order expansion as it was in Ch. 8 and Ch. 9. On the other hand, it will be shown in this chapter that two photon (Raman) transitions can be obtained using an exact interaction frame representation where leakage is exactly suppressed and no adiabatic expansion is necessary. When the transitions are driven with equal amplitude, it will be shown that these transitions can be driven in arbitrarily small times. Finally, numerical optimization with response functions (Ch. 5.1) can be used when multiple tones at different frequencies are used to drive the system. In particular, the numerical methodology is useful when not all of the oscillation can be removed from the Hamiltonian description.

The physical examples that will be considered are as follows. In Sec. 10.1, qutrits where the qubit is formed from the first and last levels with driving via an intermediary level will be considered. In Sec. 10.2, quantum elements are coupled using a

virtual sideband transition that couples $|00\rangle$ and $|11\rangle$ for an effective entangling gate. Finally, in Sec. 10.3, the situation where a qubit is isolated from the environment via a cavity and all control is done through the cavity is considered, including virtually driving the qubit.

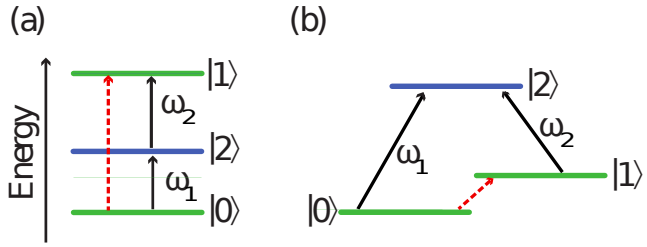

Figure 10.1: A stimulated Raman transition is shown where $|0\rangle$ and $|1\rangle$ qubit states are not directly connected but indirectly via $|2\rangle$. In (a) the transition frequency is given by the sum of the single photon transitions while in (b) it is their difference. Driven transitions between levels exist where there are black arrows while the red arrows indicate the desired transitions.

10.1 Raman transitions

An alternate strategy to assigning two adjacent levels of an energy spectrum to a qubit is to use two levels that are not connected by a matrix element in the Hamiltonian, in order to avoid relaxation error between them. In this case, a temporary level with transitions to both qubit levels can be used to connect them. In practice the auxiliary level will usually have greater decay so it is preferable to minimize intermittent population to the level. To this end virtual multi-photon transitions can be used that to a certain degree of approximation do not populate the intermediary levels. The full dynamics can once again be obtained via a frame transformation.

Most simply, we have an interaction Hamiltonian with two independent drives of the form

$$\hat{H} = (\Omega_1(t)e^{-i\phi_1}\hat{\Gamma^1}_{0,1}(-\delta) + \Omega_2(t)e^{-i\phi_2}\hat{\Gamma^2}_{1,2}(\delta)) + \text{h.c.} \tag{10.1}$$

with $\hat{\Gamma}^l_{a,b}(\delta) = \lambda^l_{a,b}e^{-i\delta t}|a\rangle\langle b|$ as in Eq. 7.3. Without loss of optimality, we pick $\Omega(t) = \text{Re}\Omega_1(t)\lambda^1_{0,1} = \text{Re}\Omega_2(t)\lambda^2_{1,2}$ and moving to frame rotating at the driving frequency we get

$$\tilde{H}(\Omega e^{-i\phi}, \delta) = (\Omega(t)e^{-i\phi}(\hat{\sigma}^{+}_{0,2} + \hat{\sigma}^{+}_{1,2}) + \text{h.c.}) + \delta|1\rangle\langle 1| \tag{10.2}$$

Applying the interaction frame transformation for the off-phase control

$$D = \exp(\int_0^t \tilde{H}(\text{Im}\Omega(t)dt, 0)dt)$$

we obtain a matrix equation for $D\tilde{H}(\text{Re}\Omega(t), 0)D^{\dagger} = \Omega^{\text{eff}}(t)\sigma^{x}_{0,1}$ which can easily be solved for the imaginary control to be

$$\int_0^t \text{Im}\Omega(t)dt = \frac{1}{2\sqrt{2}} \arctan \frac{2\sqrt{2}\text{Re}\Omega(\text{t})}{\delta} \tag{10.3}$$

This transformation leaves the Hamiltonian in the form

$$\begin{aligned} \tilde{H}^{D} \approx & \left(\frac{1}{4} + \frac{3}{4}\sqrt{1 + \frac{8\text{Re}\Omega^2(t)}{\delta^2}}\right) \delta|2\rangle\langle 2| \\ & - \frac{1}{4}\left(-1 + \sqrt{1 + \frac{8\text{Re}\Omega^2(t)}{\delta^2}}\right) \delta\hat{\Gamma}_{0,1}(0) + h.c. \end{aligned} \tag{10.4}$$

Thus we have the simple condition for rotation that

$$\int_0^T \frac{1}{4}\left(-1 + \sqrt{1 + \frac{8\text{Re}\Omega^2(t)}{\delta^2}}\right) \delta dt = \theta \tag{10.5}$$

for an exact rotation, noting that there is no phase error because of the chosen symmetry of the solution between the two drives.

This is an exact solution and completely removes both leakage and rotation error in the pulses. For all gate times and in the limit $T \to 0$, the gate error is identically 0. The solution also tolerates small sampling/shaping error despite the large amplitudes required to drive a rotation at small times. In Fig. 10.2, we have chosen to leave in the small sampling error to illustrate the dramatic decrease in error and gate times from conventional to DRAG pulses. What is most striking about the solution is that it works even for very large energies in the domain that is definitely not adiabatic.

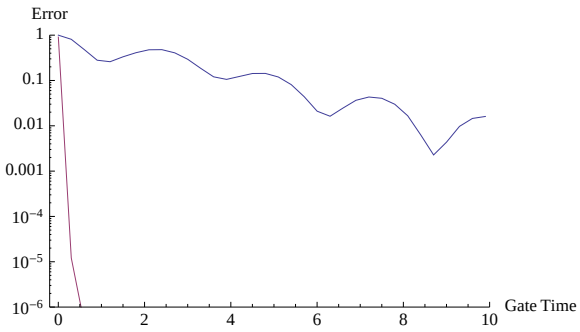

Figure 10.2: Rotation error for Raman transition with two independent drives. The blue line is the error from Gaussian shaping with the red line is the DRAG solution under finite sampling. For infinitesimal sampling this error is identically 0.

This is in contrast to conventional pulses (the blue line) which need to be slow in order to minimize higher order terms in an adiabatic expansion.

For drives not of the same amplitude $\Omega_1(t) \neq \Omega_2(t)$, such as is used in conventional STIRAP pulses [10], it is still exactly solvable but it is unclear whether there exists a closed expression. In this case there may be a similar expansion to Sec. 8.5.1 but given that the more constrained case already solves the problem exactly this has not been explored. Since STIRAP is also adiabatic, a dramatic improvement in unitary transfer is also expected by using instead DRAG.

Note that if one uses a single drive (chosen at the halfway frequency of the two driven transitions) one still gets the same Hamiltonian Eq. 10.2. However one cannot enforce both derivative conditions, Eq. 10.3 with a single first-derivative drive. In effect one runs into the same Liouvillian degeneracy issue as was the case in Sec. 9.1.1and Sec. 9.3.1. As in the other two cases, this would be a good candidate to use the second derivative as prescribed in Sec. 7.1.3.5 to simultaneously deselect both transitions.

In practice, it is better to drive the transitions separately. If there is no separate addressability to the transitions (as in Ch. 7) then we have the more general Hamiltonian

$$\begin{aligned}\hat{H} =&\Omega_1(t)e^{-i\phi_1}(\hat{\Gamma}^1_{0,1}(\Delta+\delta)+\hat{\Gamma}^1_{1,2}(\delta))+\text{h.c.}\\ &+\Omega_2(t)e^{-i\phi_2}(\hat{\Gamma}^2_{0,1}(-\delta)+\hat{\Gamma}^2_{1,2}(\Delta-\delta))+\text{h.c.}\end{aligned} \tag{10.6}$$

Provided Δ is large this does not significantly impact the solution presented, especially when the detunings are on the same side of each of the drive frequencies

(as would occur in Fig. 10A). For Δ comparable to δ it may be possible to improve the result by driving the virtual transition using all four transitions while suppressing all the single photon transitions, and this is a possible direction of future research.

10.2 2-Qubit coupling

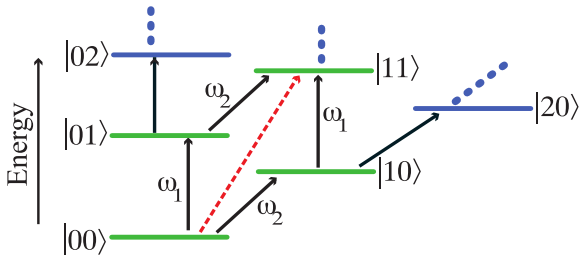

Figure 10.3: We consider the case when the $|00\rangle$ to $|11\rangle$ transition contains no explicit matrix element but by driving the single photon transitions at the correct sum frequency this transition can be activated. Driven transitions between levels exist where there are black arrows while the red arrows indicate the desired transitions.

10.2.1 Ansatz solution

We study the generalized Jaynes-Cummings system [78](also studied in Sec. (8.3.1)) with the purposes of moving an excitation between the qubit and the cavity. As a further complication, we allow the qubit to have leakage to a third level in keeping with superconducting systems as our standard example (see Fig. 10.3). Then,

$$H_{\rm JC} = \omega_r a^\dagger a + \sum_{j=1}^{d-1}(j\omega + \Delta_j)\Pi_j + \sum_{j=1}^{d-1} g_{j-1,j}\left(|j-1\rangle\langle j|a^\dagger + |j\rangle\langle j-1|a\right), \qquad (10.7)$$

where all terms are defined in Sec. (8.3.1). The selected strategy is to drive a sideband transition between the $|00\rangle$ and $|11\rangle$ levels which is a virtual transition [100]. Note that driving the transitions consecutively would not provide a reversible/entangling gate because it involves only single qubit gates. However, it cannot even be used for state preparation (of e.g. a Bell state) because in practice one cannot excite the cavity Fock states directly because it is a (infinite dimensional) harmonic system with no distinct transition energies. That is, photons that are injected into the cavity are driven by a Hamiltonian of the form

$$H_{\rm dr}(t) = \varepsilon(t)(a + a^\dagger) \tag{10.8}$$

Using this drive, three transitions can be combined into one virtual one via $|00\rangle \leftrightarrow |01\rangle \leftrightarrow |10\rangle \leftrightarrow |11\rangle$ where the first and last transition are driven and the middle transition is the natural coupling between the qubit and cavity at strength $g_{0,1}$. The sideband frequency that is used is exactly the halfway energy, $\frac{\omega+\omega_r}{2}$. Rather than calculating the matrix elements which are quite cumbersome and involve calculations very similar to Sec. (10.1) we use the ansatz that the primary errors will involve a derivative selection error on the off-resonant single photon transitions and an ac. Stark shift of the qubit transition. As in Sec. (8.7.1) we simply optimize the prefactors A,B,C,D in front of these terms:

$$\begin{aligned} \epsilon_x(t) &= A\Omega_T(t) \\ \epsilon_y(t) &= B\dot{\Omega}_T(t) \\ \delta\omega(t) &= C\frac{\Omega_T^2(t)}{\omega - \omega_r} + D \end{aligned}$$

with $\Omega_T(t)$ defined in Eq. (3.1).

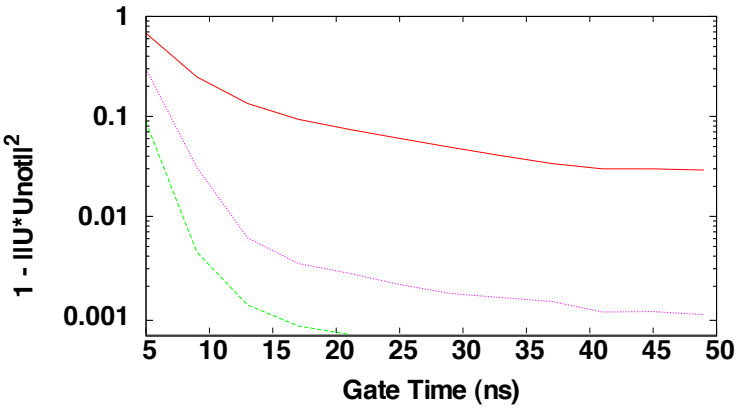

Figure 10.4: Gate error vs. gate time for sideband transition in J-C ladder with ac-Stark shift and derivative corrections.

In Fig. (10.4) we compare gate error vs. gate time for 3 solutions. The red line represents the standard solution of constant driving with 1ns rise time (consistent with electronics) and constant frequency whereby we optimize only A and D. Next we compute the "Z-only" solution (shown in dashed purple) which involves optimizing A,C, and D and where for additional selectivity we use a larger rise-time of 5ns. Finally, the "optimal 1st order" solution is consider in which all prefactors are optimized and a rise-time of 5ns is used once again. We see an improvement of 2 orders of magnitude compared to what is conventionally used.

10.2.2 Optimization with two tones

We now apply full numerical optimization to the problem of coupling two frequency separated qubits using a sideband transition. This example illustrates the need for numerical optimization on a finer grain than the sampling of the controls (as introduced in Sec. 6.1) in the common case that multiple frequency components are used in the control. This is an important problem as including more tones is straightforward way to increase the fidelity of gate operations. We consider two off-resonant qubits coupled by a $XX + YY$ interaction and drive with two different frequencies, one resonant with one of the qubit's transition frequency and the second tuned to exactly the average of the qubits' transition frequencies, in order to generate an $\sqrt{\text{ISWAP}}$ operation. Thus, we see the system has three frequencies, and there is no frame in which all fast dynamics can disappear. Specifically, we start with the Hamiltonian of the form,

$$\begin{aligned} H =& \mathcal{E}_x^{(1)}(t)\sigma_x^{(1)} \cos\left(\omega_d^{(1)} t\right) + \mathcal{E}_y^{(1)}(t)\sigma_x^{(1)} \sin\left(\omega_d^{(1)} t\right) \\ &+ \mathcal{E}_x^{(2)}(t)\sigma_x^{(2)} \cos\left(\omega_d^{(2)} t\right) + \mathcal{E}_y^{(2)}(t)\sigma_x^{(2)} \sin\left(\omega_d^{(2)} t\right) \\ &+ J\left(\sigma_+^{(1)}\sigma_-^{(2)} + \sigma_-^{(1)}\sigma_+^{(2)}\right) + \omega_1\sigma_z^{(1)} + \omega_2\sigma_z^{(1)}, \end{aligned} \tag{10.9}$$

where the superscripts index the qubit, $\sigma_\pm$ are the raising and lowering operators, J is the strength of the coupling, ω_1 and ω_2 are the qubit frequencies, and $\omega_d^{(1)}$ and $\omega_d^{(2)}$ are the drive frequencies.

Choosing the frame rotating at the energies of the two qubits (which is conventional, but not the only available frame) and setting the first drive to $\omega_d^{(1)} = \frac{\omega_1+\omega_2}{2}$

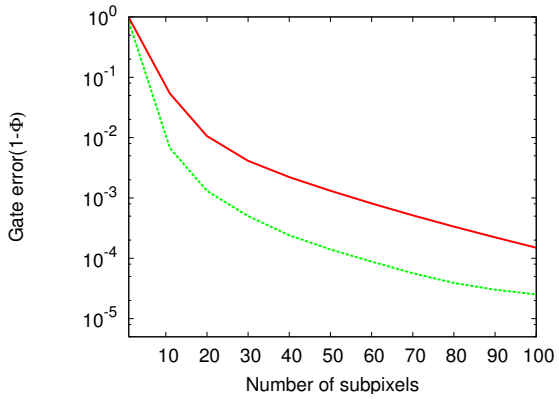

Figure 10.5: Gate error as a function of the number of subpixels for the multi-tone example. The solid red line is for $\Delta/2\pi = 1.0$ GHz and the green dashed line is for $\Delta/2\pi = 0.5$ GHz. Other parameters are given in the text.

and the second drive to $\omega_d^{(2)} = \omega_2$, we find

$$\begin{aligned} H^R =& \frac{\mathcal{E}_x^{(1)}(t)}{2}(\sigma_+^{(1)}e^{-i\Delta t/2} + \sigma_-^{(1)}e^{i\Delta t/2}) + \frac{\mathcal{E}_x^{(2)}(t)}{2}\sigma_x^{(2)} \\ &+ \frac{\mathcal{E}_y^{(1)}(t)}{2}(i\sigma_+^{(1)}e^{-i\Delta t/2} - i\sigma_-^{(1)}e^{i\Delta t/2}) + \frac{\mathcal{E}_y^{(2)}(t)}{2}\sigma_y^{(2)} \\ &+ J(\sigma_+^{(1)}\sigma_-^{(2)}e^{i\Delta t} + \sigma_-^{(1)}\sigma_+^{(2)}e^{-i\Delta t}), \end{aligned} \tag{10.10}$$

where $\Delta = \omega_1 - \omega_2$ is the energy difference between the qubits.

We illustrate the scaling with number of subpixels for this example we choose a gate time of $t_g = 20$ns, $J/2\pi = 94$ MHz, and $\Delta/2\pi = 0.5$ or 1.0, and in Fig. 10.5 we plot the predicted error as a function of the number of sub-pixels for 1ns pixels. This error is calculated by integrating the Schrödinger equation for the optimized controls on a much finer grid. As expected, we see that for the larger detuning much more subpixels are required to achieve the same low error control. Nonetheless, this method provides a valuable tool for quantum control when multiple fields are used, as is the case for but not limited to virtual transitions.

10.3 Driving through a cavity

As the final example of virtual coupling, one can consider a qubit inside a cavity where the two quantum elements are left distinct and the dispersive mode diagonalization is not performed.

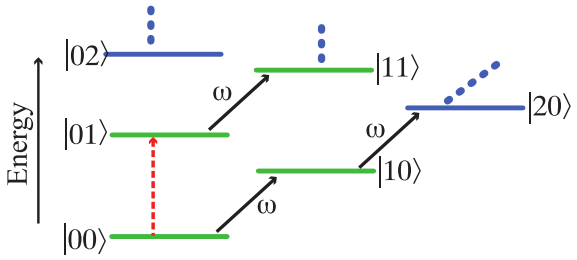

Figure 10.6: We consider the case where a qubit is placed inside a cavity so the $|00\rangle$ to $|01\rangle$ transition forms the qubit. No direct coupling exists between the qubit levels and must be driven by the cavity. Driven transitions between levels exist where there are black arrows while the red arrows indicate the desired transitions.

$$\begin{aligned} H_{\text{JC}} =& \omega_r a^\dagger a + \sum_{j=1}^{d-1}(j\omega + \Delta_j)\Pi_j + \sum_{j=1}^{d-1} g_{j-1,j}\left(|j-1\rangle\langle j|a^\dagger + |j\rangle\langle j-1|a\right) \\ &+ \epsilon_x(t)(a+a^\dagger) + i\epsilon_y(t)(a-a^\dagger), \end{aligned} \tag{10.11}$$

as above. As a design strategy there is no direct drive on the qubit (in order to isolate it from sources of noise) and only the cavity is used to read and write information to the qubit (with the expectation that will decouple most of the noise near the qubit frequency). Here, we operate in the bare basis which means that qubit rotations cannot be driven directly but virtually through the combination of the coupling and the cavity drive.

Optimization is performed using the response function methodology developed in Ch. 6, using pulses with 1ns pixels and a 250MHz filter, consistent limitations in current electronics (Sec. 6.2). An example of a pulse which generates a pi pulse for a 20ns pulse is given in Fig. 10.7 A with arbitrarily small error. Both the X and Y quadratures are needed as was the case for the simple qutrit, with, as before for an X rotation, the ϵ_x being symmetric and ϵ_y being asymmetric. In Fig. 10.7 B, a memory operation is shown. Here the ϵ_y is more clearly approximately the derivative of ϵ_x, with two rotations being indicative of a spin-echo sequence.

This setup has certain advantages compared to the dressed frame, namely that leakage out of the cavity is smaller and can even be further suppressed using optimal control techniques. Thus a major source of incoherent error can be avoided. This is again a selectivity problem where leakage does not come during gate operation but instead during memory operations. Here we expect the main effect to be an a.c. Stark shift since the coupling is constant. Operating in this frame allows for

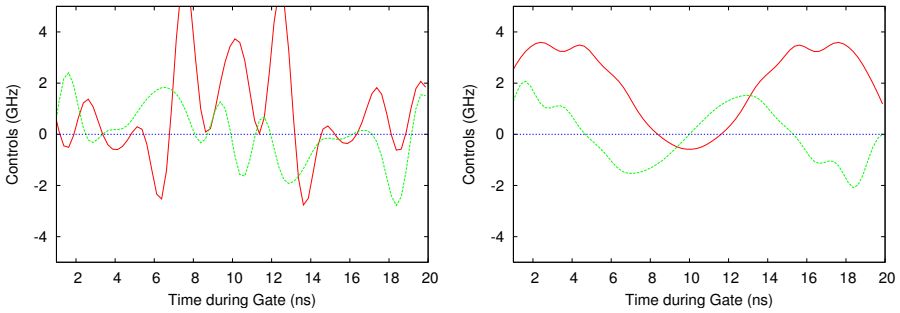

Figure 10.7: Pulse amplitude as a function of the time during gate for a qubit driven through a cavity. The solid red line is the X quadrature control and the green dashed line is the Y. Parameters are given in the text. (A) is a NOT pulse, U_π^X; (B) is a memory operation, $\mathbb{1}$.

drastically better selectivity in this regard, enabling many qubits to be placed inside the cavity. An additional benefit is that the cavity frequency is not shifted by the presence of energy in the qubits, which means radiation coming into the cavity at a certain frequency will not be attenuated given the state of the qubits, resulting in rotation angle errors.

10.4 Summary

In this chapter, we have seen that off-resonant terms and fast dynamics form an important source of error in multi-photon transitions. We have explored three different cases where virtual transitions form an integral part of the computation scheme with a clear benefit from a decrease in these forms of error. Methods such as STIRAP can be used to exchange energy between two levels that do not normally share a transition. This can be extremely useful since relaxation cannot occur where there is no transition and can be greatly suppressed where there are multi-photon transitions. However, the conventional analysis is that these pulses must be adiabatic which can greatly slow down gate times making it prone to decoherence. In contrast, the two-photon DRAG solution derived is exact and works in the diabatic regime as well, allowing for arbitrarily short pulses. Similarly, virtual transitions can be used to entangle a qubit (or anharmonic ladder) to a (harmonic) cavity where single photon transitions in the cavity could not be used for the same effect. Once again, in such a situation there is evidence that using a derivative shaped imaginary control greatly reduces selection error. Finally, working with a non-linear oscillator and a cavity in

the bare frame, it is shown numerically that all gate operations can be achieved using virtual transitions, where once again relaxation can be greatly decreased by avoiding putting population in the cavity where it can spontaneously decay. Such a scheme can have advantages also in terms of how many qubits can fit in the coupling cavity, since placing them at frequencies close to that of the cavity is less of a concern as memory operations can be engineered to undo the effect of the Stark shift from the cavity.

Chapter 11

Strong Coupling

We have seen in Sec. 4.2.2 that transitions are typically driven with monochromatic sinusoidal fields rather than constant or near-constant driving fields. We have seen that in a rotating frame (Sec. 4.2.2) at the frequency of the monochromatic oscillation, a more intuitive representation is given by

$$\begin{aligned} H_{rot} &= e^{i\omega\Pi^1 t}((2a(t)\cos(\omega_d t+\phi)+2b(t)\sin(\omega_d t+\phi))X_\pi + c(t)\Pi^1)e^{-i\omega\Pi^1 t} \\ H_{rot} &= (c(t)+\Delta\omega)\Pi^1 && (11.1) \\ &\quad +\frac{1}{2}(a(t)\cos(\delta\omega t)+b(t)\sin(\delta\omega t))X && (11.2) \\ &\quad +\frac{1}{2}(b(t)\cos(\delta\omega t)-a(t)\sin(\delta\omega t))Y \\ &\quad +\frac{1}{2}(a(t)\cos(\omega_\Sigma t+2\phi)+b(t)\sin(\omega_\Sigma t+2\phi))X && (11.3) \\ &\quad +\frac{1}{2}(-b(t)\cos(\omega_\Sigma t+2\phi)+a(t)\sin(\omega_\Sigma t+2\phi))Y \end{aligned}$$

with the same notation as before and where we stress again that the last two lines (the counter-rotating terms) are usually dropped via energy considerations (the RWA). The statement can be made precise in the formulation of Sec. 7.1.1 with the limiting condition that $\omega_\Sigma \gg a, b$ in order that selectivity hold true. Here, we will retain the terms under the assumption of strong driving that the terms are all of roughly the same order $\omega_\Sigma \gtrapprox a, b$. Following the prescription of Sec. 8.5, the selectivity in the presence of these counter-rotating terms can be improved by the adiabatic expansion technique, as will be done in Sec. 11.1. Next, in Sec. 11.2 the adverse effect of the counter-rotating terms is suppressed using robust numerical optimization techniques with the use of fine-graining.

11.1 Analytic solution

We move to the rotating frame with respect to the off-resonant element with $R_{j,k,l} = \exp(i\sum_{j,l}\omega_\Sigma t|j\rangle\langle j|)$

$$\tilde{H} = \frac{1}{2}\hat{H}_{0,1}(\Omega(t)e^{i\phi}, -\omega_\Sigma) + \frac{1}{2}\widetilde{H}_{0,1,0}(\Omega(t)e^{i\phi}, \omega_\Sigma) \tag{11.4}$$

where we take $\Omega(t)e^{i\phi} = a(t)+ib(t)$ and as before $\widetilde{H}_{0,1,0}(\Omega(t)\phi, \Sigma\omega) = \Omega(t)e^{-i\phi}\sigma^+ + h.c. + \Sigma\omega|1\rangle\langle 1|$ and $\hat{H}_{0,1}(\Omega(t)e^{i\phi}, -\Sigma\omega) = \Omega(t)e^{-i\phi - i\Sigma\omega t}\sigma^+ + h.c. + \Delta\omega|1\rangle\langle 1|$.

Moving to the interaction frame with respect to the imaginary control $\mathrm{Im}\Omega(t) = \Sigma_h\Omega_h(t)\epsilon^h$ and choosing for the counter-rotating terms the selectivity parameter $\epsilon = \frac{\Omega(t)}{2\omega^d}$ defines the transformation

$$D_1 = \exp(-i\sum_{jk}\int_0^t \tilde{H}_{jkl}(i\Omega_1\epsilon, 0)dt + \text{h.c.})$$

which gives upon expansion in the small parameter

$$D_{(1)}\widetilde{H}_{0,1,0}(\mathrm{Re}\Omega(t)\phi, 2\omega^d)D^\dagger_{(1)} = -(\mathrm{Re}\Omega(t))^2/16\omega_d\sigma_z + 2\omega^d|1\rangle\langle 1|_B + O(\epsilon^3)$$

$$\begin{aligned} D_{(1)}\hat{H}_{0,1}(\mathrm{Re}\Omega(t), \phi, -2\omega^d)D^\dagger_{(1)} =& \hat{H}_{0,1}(\mathrm{Re}\Omega(t), \phi, -2\omega^d) - (\mathrm{Re}\Omega(t))^2/16\omega\sigma_z \\ & - (\mathrm{Re}\Omega(t))^3/16\omega^2\sigma^+ + h.c. + O(\epsilon^3) \end{aligned}$$

This gives the second order solution to the selectivity criteria (Eqs. 7.20)

$$\begin{aligned} \Delta_{0,1}(t) &= (\mathrm{Re}\Omega(t))^2/4\omega_d \\ \mathrm{Im}(\Omega(t)) &= \mathrm{Re}(\dot{\Omega}(t))/2\omega_d \\ \int_0^T (\mathrm{Re}\Omega(t))^{\text{eff}} - (\mathrm{Re}\Omega(t))^3/16\omega_d^2 dt &= \theta \end{aligned} \tag{11.5}$$

Fig. 11.1 demonstrates the validity of this strategy. The red line shows gate error vs. gate time when using a Gaussian pulse. The blue line shows the improvement when optimizing using only the derivative control. The yellow line shows the gate error when both the derivative and detuning are applied, Eq. (11.5). In both cases we see the RWA errors are completely suppressed.

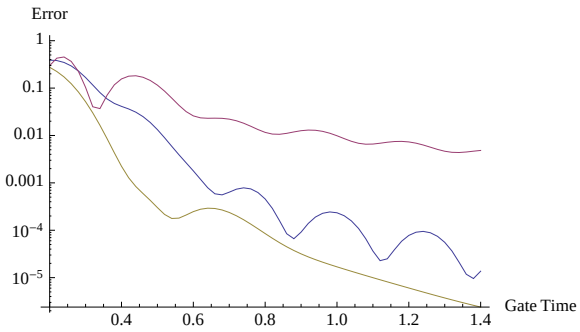

Figure 11.1: Error from counter-rotating terms in qubit frame. Gate error for Gaussian (red), Gaussian with derivative (blue), Gaussian with derivative and detuning (orange) are shown.

11.2 Numerical optimization

Finally, we numerically investigate the errors from making the RWA. We take the same system as in Sec. 6.5, Eq. 6.20, but we optimize without making the RWA. We are concerned with seeing the full extent that the errors arising from the counter-rotating terms can be suppressed. Moreover, the errors arising from the rotating terms demonstrate a different use for carrier modulation optimization (Sec. 6.1). In this case, the Hamiltonian can be written as

$$H = +\Delta|2\rangle\langle 2| + \mathcal{E}^x(t)\frac{[\Gamma(1+e^{-2i\omega_1 t-2i\psi})+\text{h.c}]}{2} + \mathcal{E}^y(t)\frac{[i\Gamma(1+e^{-2i\omega_1 t-2i\psi})+\text{h.c}]}{2}, \tag{11.6}$$

with terms defined as in Eq. 6.20. To demonstrate the error from using the over-simplified Hamiltonian, we first find optimal solutions for the π pulses assuming the RWA (optimization with Eq. 6.20) for $\Delta/2\pi = -500$ MHz, $\omega_1/2\pi = 2.0$ GHz, and control pixels of 0.125 ns. This optimal solution is shown in Fig. 11.2 as the solid purple line as a function of the gate time. For each gate time the controls are then used to evaluate what the fidelity of the operation would be if we did not make the RWA for various phases ψ (evolution under Eq. 11.6). These results are shown in Fig. 11.2 as the dotted blue lines, where each line represents a randomly chosen phase. This figure clearly indicates that for short gate times, neglecting the rotating terms leads to a large error.

To perform the optimization with the rotating terms we simply include them as a carrier function (Sec. 6.4.3). Since we would also like to find pulses which

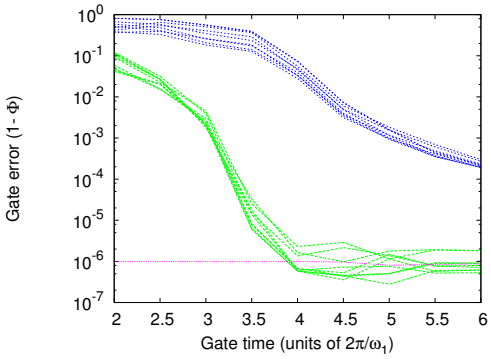

Figure 11.2: Gate error as a function of gate time for the carrier example. The horizontal solid purple line is the (target and achieved) gate error for the optimal solution when we make the RWA. The dotted blue lines are the predicted error when we do not make the RWA but find the optimal solution with the RWA. The lower dashed green lines are the gate errors when we optimize without making the RWA. Each line represents one of the nine different values of the relative phase between the envelope and the carrier: 2.46245, 2.13875, 1.57081, 0.304685, 0.043838, 1.65238, 0.914728, 2.02047, 0.518253.

don't depend on ψ, we also use the robust technique outlined in Sec. 5.3.2. This is essentially an optimization of average fidelity over all possible ψ. In Fig. 11.2 we plot the gate error as a function of gate time (dashed green lines), seeing that it is possible to find pulses that robustly remove the rotation errors. Here we set the number of sub-pixels per control pixel to 100.

Together with the results from the previous subsection, this chapter demonstrates that working in the rotating frame where dynamics are simplified there is a limit to how short pulses can be used. The adiabatic expansion and response function formalism allow to push this limit to very short pulses. In comparison to conventional textbook techniques where many oscillations are assumed, pulses on the order of one or two precessions of the qubit can still achieve the low errors required for computation. Moreover, it was shown numerically that errors from counter-rotating terms and from leakage can both be removed simultaneously. While ideally a perfect system characterization is preferred, this chapter also shows that averaging effects over the order of one or two precessions of the qubit are sufficient to make possible robustness against the effects of the unknown relative phase between the envelope and carrier of shaped pulses.

Chapter 12

Summary and Conclusions

Quantum information promises to deepen our understanding of computation. Fundamentally, existing and future progress in the endeavour relies on the wave-like (near-unitary) nature of the continuous time-evolution of quantum systems. Much of the progress to be made rests in being able to manipulate the complicated interactions between many energy levels in a predictable way. Whether it is possible to perform this task in the presence of non-commuting time-dependent operators and/or large Hilbert space dimension is an important question. Part of the answer lies in being able to effectuate specific transitions where a different transition is also possible and the spectral gap between them is relatively small. Whether a fundamental time limit exists in such (generic) situations or whether quantum control near and into the the diabatic regime is possible may have important consequences in quantum computation. Moreover, methods are desired for properly characterizing and optimizing evolutions of such systems that involve many frequencies and/or simultaneous drives across multiple matrix elements. Most of the research done for this thesis involves control problems of this nature in various physical circumstances, with non-commuting effects competing on approximately the same time scales.

The first question that is addressed in Ch. 6 is whether it is possible to characterize precisely the control hardware of a quantum system. In particular, in the microwave range that is used for solid state systems, technology is not readily available to arbitrarily shape control pulses, with filtering and other transfer functions playing an important limiting role in the electronics. Being able to accurately predict and optimize the control shapes and their quantum effect is crucial across all technologies for the purposes of lowering error rates towards fault-tolerant thresholds. In this chapter, a new method is developed for optimization that separates the time-scales between the digital shaping of the control field and the continuous dynamics that take

place at the quantum level. The method is versatile and usable with many different kinds of controls and linear transfer functions, which includes splines, filters, and Fourier components. The latter is used in later chapters when there are more than one frequency per matrix element and standard numerics cannot be used. Quantifying the effect of filtering, on the other hand, has a drastic effect in terms of fidelity, and pulses that are as short as 3 or 4 digital control pixels become feasible in the presence of non-commuting error (e.g. for qutrits). Rise-times are also studied, in particular in the context of bringing off-resonant qubits into resonance in order to exchange energy. The non-commuting terms in the avoided crossing can be optimized analytically using a Magnus expansion, with a simple pulse with overshoot(s) being found that can straightforwardly be used in experiment, allowing high quality two-qubit gates. Quantifying the full effect of the filtering is also done numerically and confirms the analytic result.

When multiple elements are controlled by a common global drive, being able to control them individually is known as a spectral selectivity problem. The problem is ubiquitous across implementations because while quantum elements can be microscopic, the control fields are often spread out and couple to multiple quantum elements. This is a severe impediment to the ability to scale the technologies as every additional qubit or control line adds crosstalk to the elements around it. An analytic technique called DRAG is derived in Sec. 7.1 which addresses selecting qubits at different operating frequencies. The technique improves greatly on semi-classical approaches based on bandwidth considerations by taking the full time-dependent quantum dynamics into account. For two qubits, the result involves using an auxiliary control $\pi/2$ out of phase with the first and proportional to its derivative, and it is exact in the adiabatic limit. A full basis is produced using higher-order derivatives which allows suppressing multiple qubits, and a general result is derived to first order in the adiabaticity parameter. Even-order derivatives have the property of suppressing transitions both greater and smaller in frequency from the driven transition. This symmetry is useful in several contexts; in particular it can be used to reduce the quantum bandwidth of the pulse, with potential application in magnetic imaging. The more general case where there may not be frequency separation between the qubits is also considered in Sec. 7.2, in the context of crosstalk between inductive couplers for flux qubits in a lattice topology. It is found that separation in the Rabi frequency can be successfully used instead, though in practice the separation in qubit frequency is a valuable resource.

From an engineering point of view, it is desirable to find quantum implementation technologies that can be scaled to very large numbers of qubits. Superconducting technologies benefit from advanced lithographic techniques used in industry but are

impeded by electromagnetic circuit elements functioning as harmonic oscillators. Josephson junctions partially redress this problem but the regime in which they are useful is typically weakly anharmonic. Whether these devices can be accurately controlled has remained an open question due to the absence of strong anharmonicity. However, the DRAG framework has shown conclusively both theoretically and experimentally that fast pulses can be found, many orders shorter than the coherence times of these systems. In Ch. 8, these pulses are derived now using an iterative technique, due to the infinite number of states in the energy ladder. Using a second frequency at twice the frequency of the main driving field, which drives the two-photon transition, an exact analytical solution is found to the qutrit problem which is strictly valid in the adiabatic regime, and still offers a significant improvement in the infinite ladder situation. It is found that, contrary to the conventional belief that an (a.c. Stark) phase shift is the primary effect of off-resonant coupling, at short times the derivative (i.e. selection) error can dominate and removing it is crucial to operation. This is also different from bandwidth arguments that rely on removing selection error on average by the end of the pulse. Since the derivative control commutes with the principle control to generate a phase shift, the derivative control can in fact be used both to remove selection error and phase error. A family of solutions thus exists in the first order of the inverse anharmonicity which is able to partially remove both of these errors; it is useful in practice due to the fact that simpler functional form can be found that make it easier to calibrate. The full optimal solution is also investigated numerically and it is found that the error decreases dramatically with pulse sampling rate, consistent with the result of adding in more higher derivatives. In particular, in the diabatic regime the decrease is exponential and arbitrarily short pulses can be engineered given high enough resolution. The problem of leakage to higher levels is also examined in the context of relaxation and decoherence. While little can be done about direct relaxation other than making the pulses faster, there is a tradeoff between increasing the periodic potential which decreases energy dispersion and decreasing it which increases the anharmonicity. Previous work motivating the transmon qubit has shown that dispersion can be decreased exponentially while still maintaining anharmonicity. The result in this chapter contends that dispersion can also be increased somewhat because the energy dispersion is accompanied by operator dispersion as well. The resultant uncertainty can thereafter be suppressed using robust optimal control techniques while taking advantage of increased anharmonicity such that shorter pulses are found in aggregate. To finish the chapter, the experimental realization is reviewed which verifies the two quadrature analytical result.

Operations involving multiple leakage channels have been even more difficult

to accurately control due to the larger number of constraints involved and as a result two-qubit gates for superconducting qubits have lagged significantly behind single qubit gates. The DRAG framework is generalizable to situations with multiple transitions, as was the case with crosstalk to multiple qubits. In these situations, the derived analytic formulae can once again decrease the errors by orders of magnitude and decrease gate times to the limits set by current pulse-shaping technology. There are essentially two ways in which the extra leakage can manifest itself and these are explored in Ch. 9. The first is when there is a leakage channel connected to either level of the driven transition, typically arising in higher level transitions in an anharmonic ladder. Driving population to higher levels has many potential useful applications including being used for qudit logic (encoding multiple qubits in one non-linear oscillator), using higher level states for coupling qubits, and metrology where there can be a $\sqrt{N}$ factor in the signal strength. Driving these levels was also investigated numerically with low error well into the diabatic regime (for fixed input sampling rate of the control). On the other hand, multiple leakages can be connected to the same level. This is typically the case when multiple ladder elements are combined, such as an anharmonic qubit dressed by a cavity, or when coupling two anharmonic qubits. In the case where the leakage energies are both above and below the computational level (e.g. a Liouvillian degeneracy), an alternate scheme is found that is particularly effective at low coupling strength, namely to use a second tone to create a.c. Stark shifts near the leaking level to counteract the a.c. Stark shifts caused by the leakage transitions. Due to the larger coupling constant at higher levels in (an)harmonic oscillators, and by placing the added tone very close to the phase-shifted level, the leaking level can be corrected with suppressed disturbance to other computational states.

One of the most widely used techniques in quantum control is the use of virtual transitions. In particular, methods such as STIRAP can be used to exchange energy between two levels that do not normally share a transition. This can be extremely useful since relaxation cannot occur where there is no transition and can be greatly suppressed where there are multi-photon transitions. However, the conventional analysis is that these pulses must be adiabatic which can greatly slow down gate times making it prone to decoherence. In contrast, in Ch. 10, the two-photon DRAG solution derived is exact and works in the diabatic regime as well, allowing for arbitrarily short pulses. Other multi-photon schemes are also explored in the context of adding a derivative with decreases in the selectivity error as expected. Finally, working with a non-linear oscillator and a cavity in the bare frame, it is shown numerically that all gate operations can be achieved using virtual transitions, where once again relaxation can be greatly decreased by avoiding putting population

in the cavity where it can spontaneously decay. Such a scheme can have advantages also in terms of how many qubits can fit in the coupling cavity, since placing them at frequencies close to that of the cavity is less of a concern.

The last theme that is explored is the strong driving regime in the rotating frame. The rotating wave approximation is ubiquitous in quantum control because it allows for very high quality control using the weak fields/couplings that are typically available or desired in complex quantum systems. The rotating frame is also very intuitive and avoids the messy non-commutative algebra that is present at shorter times. The DRAG formalism is also applicable to the error from this approximation, as is shown in Ch. 11, with simple and intuitive pulses that are ultra-short, less than two qubit precessions. The error is also examined numerically and can be removed in a way that is robust against the unknown phase between the envelope and carrier of the pulse.

Future research directions will consider control of larger Hilbert spaces and combinations of a larger number of error terms and qubits. Numerically this may be achieved with a speed-up of the simulations and optimizations, with potential improvements including parallelism, Hessian techniques, and sparse matrix algebra. From an analytical point of view, it should be possible to combine analytical solutions when multiple errors are present (such as shaping, off-resonant errors, robustness, etc.). In particular, situations with multiple linked transitions have not yet been investigated using a basis of orders of derivatives of Gaussians as was done for crosstalk to several qubits. A particularly elegant solution may also exist where there is Liouvillian degeneracy, such as climbing up the anharmonic ladder or coupling qutrits, in which case the symmetry of the second derivative solution would make it a prime candidate. Trying different pulse shapes other than Gaussians may also be beneficial, in particular finding shapes where the higher derivatives are zero at the start and end of the pulse. Alternatively, finding different frames where it is not necessary to start and end at zero may be preferable since the bandwidth properties of the base envelope shape would be persevered. Finally, there is strong numerical evidence that higher pulse resolution / more frequency components can greatly improve controllability into the diabatic regime. Investigating the use of extra tones may be fruitful, using techniques derived earlier or new ideas such as using an a.c. Stark shift in the middle of a transition to increase the gap size.

Other potential avenues of future research would include application of optimal control ideas to areas other than optimizing gates. For example, measurement and error correction both involve transfer of information between quantum elements and can be optimized. Moreover, while most of the techniques applied are analog, many circuits or measurement-based approaches involve digital sequences of gates which

could also be optimized together with or independently from the analog optimization of gates. That is, small algorithms that are digital or analog typically are not optimized for specific physical dynamics but doing so would push the envelope of what is presently possible and offer insight into the algorithms.

Appendix A

Matrix Exponentiation

Solving Schrödinger's equation for larger Hilbert spaces can be difficult even when time-dependence is trivial. We examine here some of the known methods of exponentiation.

The first step for any numerical techniques is to subdivide the time into time steps of duration δt. We distinguish this time duration form the pixel sizeΔt (e.g. the inverse of the sampling rate in Ch. 5.1) given in other cases because it is not an intrinsic time scale in the system. Rather, it is an arbitrarily small enough time so that the evolution looks approximately smooth on the coarse-grain. The condition is roughly speaking $|H|\delta t \ll 1$, though this is perhaps a bit strong. In practice, one can see what time step will give convergence for the precision required. In any case, the evolution for a general Hamiltonian with multiple time-dependences of the form can be approximated as

$$H(t) = \sum_{k=0}^{M} c_k(t)\hat{H}_k,$$

where c(t) are continuous time signals such as the convolutions, Eq. 3.2. The evolution can then be calculated via

$$U = \prod_{j=N-1}^{0} U_j = \prod_{j=N-1}^{0} e^{-iH(j\delta t)\delta t}$$

which bears resemblance to Eq. 5.1, with the important distinction that it is an approximation.

A.1 Taylor Expansion

In the case where H is small, the simplest expansion is

$$U = \exp(-iH\Delta t) = 1 - iH\Delta t - \frac{H^2\Delta t^2}{2} + \dots$$

This can be the case for a waveform generator with a large sampling rate ($\Delta t \ll 1$, see Sec. 5.1) or when H is a perturbation such as an auxiliary control (see Sec. 7.1). It is also clearly valid in numerics with small time-slicing (Sec. 6.1, $\Delta t = \delta t$).

If H is not small, it is still possible to use this expansion if desired using the so-called "scaling and squaring technique". Using the special properties of exponential functions, the above exponential can be written as

$$U = (U^{1/N})^N = (\exp(-iH\Delta t/N))^N = (1 - \frac{iH\Delta t}{N} - \frac{H^2\Delta t^2}{2N^2} + \dots)^N.$$

In practice, the Nth power at the end can be accomplished by the repeated squaring technique where the matrix inside the parentheses is recursively multiplied with itself $\lg_2 N$ times.

In general, this technique is most useful when H is sparse and matrix multiplication is cost-effective.

A.2 Diagonalization

Another useful analytic and numeric technique is eigenvector decomposition.

$$H = DH_{\text{diag}}D^\dagger \tag{A.1}$$

which exponentiates trivially to

$$U = \exp(-iH\Delta t) = D\exp(-iH_{\text{diag}}\Delta t)D^\dagger.$$

The diagonal elements of H_{diag} exponentiate like scalars. This *spectral decomposition* is justified by the Hermiticity of the Hamiltonian. The most time-consuming aspect is obtaining the diagonalization D as well as the matrix multiplication. In what concerns the diagonalization, it can often be advantageous to pre-compute the decomposition D when the form of the matrix to be exponentiated is known ahead of time. For $H = a(t)H_0$ and $H_0 = DH_{\text{diag}}D^\dagger$,

$$U = \exp(-iH\Delta t) = D\exp(-ia(t)H_{\text{diag}\}}\Delta t)D^{\dagger}, \tag{A.2}$$

which again only involves scalar exponentiation and matrix multiplication.

A.3 Hamiltonian splitting

In the same vein, we have seen the Hamiltonian will most often be composed of a time-independent drift and several time-dependent controlled operators, as in Eq. 5.1. For short times the different components can be exponentiated separately with

$$U_j = \exp(-iH_j\delta t) \cong \prod_{k=0}^{N} \exp(-iu_{j,k}(t)\widehat{H}_k\delta t).$$

To ensure this approximation is valid the scaling and squaring technique can be used here too. Less scaling can be required by using as many permutations of the the products as possible to ensure as many commutators in a BCH type expansion tend to 0. Calculating the individual exponentials can be done any number of ways though the diagonalization is attractive since the eigen-decomposition need only be found once for each $\widehat{H}_k$ if Eq. A.2 is then used. Once again, if matrix multiplication can be done cheaply due to sparsity then this method may be preferable to diagonalization.

A.4 Other methods

There are numerous other methods [120] which can be used to calculate the evolution including the Padé expansion (used by Matlab for general matrices) and differential equation solvers which may be useful for highly oscillatory functions.

Appendix B

Publications

- B. Khani, S. T. Merkel, **F. Motzoi**, J. M. Gambetta, and F. K. Wilhelm, "High-fidelity quantum gates in the presence of dispersion," *Phys. Rev. A* 85, 022306 (2012). [arXiv:1111.1708]
- **F. Motzoi**, J. M. Gambetta, S. T. Merkel, and F. K. Wilhelm, "Optimal control methods for fast time-varying Hamiltonians," *Phys. Rev. A* 84 (022307) (2011). [arXiv:1102.0584]
- P. Groszkowski, A. Fowler, **F. Motzoi**, and F. K. Wilhelm , "Tunable coupling between three qubits as a building block for a superconducting quantum computer," *Phys. Rev. B* 84, 144516 (2011). [arXiv:1102.0307]
- J. M. Gambetta , **F. Motzoi**, S. T. Merkel, and F. K. Wilhelm, "Analytic control methods for high fidelity unitary operations in a weakly nonlinear oscillator", *Phys. Rev. A* 83 ,012308 (2011). [arXiv:1011.1949]
- J. M. Chow, L. DiCarlo, J. M. Gambetta, **F. Motzoi**, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, "Implementing optimal control pulse shaping for improved single-qubit gates," *Phys. Rev. A* 82, 040305 (2010). [arXiv:1005.1279]
- B. Khani, J. M. Gambetta, **F. Motzoi**, and F. K. Wilhelm, "Optimal generation of Fock states in a weakly nonlinear oscillator," *Physica Scripta*, T137, 014021 (2009). [arXiv:0909.4788]
- **F. Motzoi**, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, "Simple pulses for elimination of leakage in weakly nonlinear qubits", *Phys. Rev. Lett.* 103, 110501 (2009). [arXiv:0901.0534]

Bibliography

[1] Scott Aaronson. Is quantum mechanics an island in theoryspace? *quant-ph/0401062v2*, 2004.

[2] Frederick W. Strauch Anne M. Forney, Steven R. Jackson. Multi-frequency control pulses for multi-level superconducting quantum circuits. 2009. arXiv:0909.1577.

[3] Markus Ansmann, H. Wang, Radoslaw C. Bialczak, Max Hofheinz, Erik Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, A. N. Cleland, and John M. Martinis. Violation of bell's inequality in josephson phase qubits. *Nature*, 461(7263):504–506, September 2009.

[4] N.W. Ashcroft and N.D. Mermin. *Solid state physics*. Holt-Saunders, 1976.

[5] S. Ashhab, AO Niskanen, K. Harrabi, Y. Nakamura, T. Picot, PC De Groot, C. Harmans, JE Mooij, and F. Nori. Interqubit coupling mediated by a high-excitation-energy quantum object. *Phys. Rev. B*, 77(1):14510, 2008.

[6] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Theory of superconductivity. *Phys. Rev.*, 108(5):1175–1204, 1957.

[7] Christopher Bauer, Ray Freeman, Tom Frenkiel, James Keeler, and A. J. Shaka. Gaussian pulses. *J. Magn. Reson.*, 58(3):442–457, 1984.

[8] J. Baugh, J. Chamilliard, C.M. Chandrasekhar, M. Ditty, A. Hubbard, R. Laflamme, M. Laforest, O. Moussa, C. Negrevergne, M. Silva, S. Simmonds, C.A. Ryan, D.G. Cory, J.S. Hodges, and C. Ramanathan. Quantum information processing using nuclear and electron magnetic resonance: Review and prospects. *Phyiscs in Canada*, 63(4):197, 2007.

[9] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on*

Computers, Systems, and Signal Processing, Bangalore, India, pages 175–179, New York, 1984. IEEE.

[10] K. Bergmann, H. Theuer, and B. W. Shore. Coherent population transfer among quantum states of atoms and molecules. *Rev. Mod. Phys.*, 70:1003–1025, 1998.

[11] P. Bertet, C.J.P.M. Harmans, and J.E. Mooij. Parametric coupling for superconducting qubits. *Phys. Rev. B*, 73:064512, 2006.

[12] R.C. Bialczak, M. Ansmann, M. Hofheinz, M. Lenander, E. Lucero, M. Neeley, A.D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, T. Yamamoto, A.N. Cleland, and J.M. Martinis. Fast tunable coupler for superconducting qubits. *Phys. Rev. Lett.*, xx(accepted for publication), 2010.

[13] R. Bianchetti, S.Filipp, M. Baur, J. M. Fink, C. Lang, L. Steffen, M. Boissonneault, A. Blais, and A. Wallraff. Control and tomography of a three level superconducting artificial atom. arXiv:1004.5504, 2010.

[14] A. Blais, J. Gambetta, A. Wallraff, D.I. Schuster, S.M. Girvin, M.H. Devoret, and R.J. Schoelkopf. Quantum-information processing with circuit quantum electrodynamics. *Phys. Rev. A*, 75:032329, 2007.

[15] A. Blais, R-S. Huang, A. Wallraff, S.M. Girvin, and R.J. Schoelkopf. Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computing. *Phys. Rev. A*, 69:062320, 2004.

[16] Maxime Boissonneault, Jay M. Gambetta, and Alexandre Blais. Dispersive regime of circuit qed: Photon-dependent qubit dephasing and relaxation rates. *Phys. Rev. A*, 79(1):013819, 2009.

[17] M. D. Bowdrey, D. K. L. Oi, A. J. Short, K. Banaszek, and Jonathan A. Jones. Fidelity of single qubit maps. *Phys. Lett. A*, 294:258, 2002.

[18] A.M. Brink, AJ Berkley, and M. Yalowsky. Mediated tunable coupling of flux qubits. *New Journal of Physics*, 7:230, 2005.

[19] M. Brune, S. Haroche, V. Lefevre, J. M. Raimond, and N. Zagury. Quantum nondemolition measurement of small photon numbers by rydberg-atom phase-sensitive detection. *Phys. Rev. Lett.*, 65(8):976–979, 1990.

[20] A.E. Bryson and Y-C. Ho. *Applied Optimal Control*. Taylor and Francis, 1975.

[21] C. A. Ryan M. Ditty F. Cyr-Racine W. Power N. Boulant T. Havel D.G. Cory R. Laflamme C. Negrevergne, T.S. Mahesh. Benchmarking quantum control methods on a 12-qubit system. *Physical Review Letters*, 96:170501, 2006.

[22] A. R. Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54:1098–1105, 1996.

[23] Fernando Casas. Sufficient conditions for the convergence of the magnus expansion. *J. Phys. A: Math. Theor.*, 40:15001–15017, 2007.

[24] G. Chen, D.A. Church, B-G. Englert, C. Henkel, B. Rohwedder, M.O. Scully, and M.S. Zubairy. *Quantum Computing Devices: Principles, Designs, and Analysis*. Chapman and Hall/CRC, Boca Raton, 2006.

[25] I. Chiorescu, Y. Nakamura, C.J.P.M. Harmans, and J.E. Mooij. Coherent quantum dynamics of a superconducting flux qubit. *Science*, 299:1869, 2003.

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

[27] J. M. Chow, Jay M. Gambetta, L. Tornberg, Jens Koch, Lev S. Bishop, Andrew A. Houck, B. R. Johnson, L. Frunzio, Steven M. Girvin, and Robert J. Schoelkopf. Randomized benchmarking and process tomography for gate errors in a solid-state qubit. *Phys. Rev. Lett.*, 102(9):090502, 2009.

[28] J.M. Chow, L. DiCarlo, J.M. Gambetta, F. Motzoi, L. Frunzio, S.M. Girvin, and R.J. Schoelkopf. Implementing optimal control pulse shaping for improved single-qubit gates. *Phys. Rev. A*, 82:040305(R), 2010.

[29] J. Clarke and F.K. Wilhelm. Superconducting qubits. *Nature*, 453:1031, 2008.

[30] J. Clarke and F.K. Wilhelm. Superconducting qubits. *Nature*, 453:1031, 2008.

[31] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics*. Wiley Interscience, Weinheim, 1992.

[32] E. Collin, G. Ithier, A. Aassime, P. Joyez, D. Vion, and D. Esteve. Nmr-like control of a quantum bit superconducting circuit. *Phys. Rev. Lett.*, 93:157005, 2004.

[33] H.K. Cummins and J.A. Jones. Use of composite rotations to correct systematic errors in nmr quantum computation. *New J. Phys.*, 2:6, 2000.

[34] P. de Fouquieres and S. G. Schirmer. Quantum control landscapes: A closer look. *arxiv:1004.3492*, 2010.

[35] P.G. de Gennes. *Superconductivity of metals and alloys*. Benjamin, N.Y., 1966.

[36] W. Demtröder. *Laser Spectroscopy*, volume 1. Springer, Berlin, 2008.

[37] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill. Topological quantum memory. *Journal of Mathematical Physics*, 43:4452, 2002.

[38] M.H. Devoret, A. Wallraff, and J.M. Martinis. Superconducting qubits: A short review. cond-mat/0411174.

[39] L. DiCarlo, J. M. Chow, J. M. Gambetta, Lev S Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Demonstration of two-qubit algorithms with a superconducting quantum processor. *Nature*, 460(7252):240–244, Jul 2009.

[40] L. DiCarlo, M. D. Reed, L. Sun, B. R. Johnson, J. M. Chow, J. M. Gambetta, L. Frunzio, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf. Preparation and measurement of three-qubit entanglement in a superconducting circuit. arXiv:1004.4324, 2010.

[41] D.P. DiVincenzo. The physical implementation of quantum computation. *Fortschr. Phys.*, 48:771, 2000.

[42] Zeph Landau Dorit Aharonov, Vaughan Jones. A polynomial quantum algorithm for approximating the jones polynomial. In *Symposium on Theory of Computing, arXiv:quant-ph/0511096v2*, 2006.

[43] R. Martinez C.-H. Tseng E. Knill, R. Laflamme. A cat-state benchmark on a seven bit quantum computer. *arXiv:quant-ph/9908051*, 1999.

[44] Sam Gutmann Joshua Lapan Andrew Lundgren Daniel Preda Edward Farhi, Jeffrey Goldstone. A quantum adiabatic evolution algorithm applied to random instances of an np-complete problem. *Science*, 292:472, 2001.

[45] R.R. Ernst and G. Bodenhausen and. A. Wokaun. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. International series of monographs on chemistry. Oxford University Press, Oxford, 1990.

[46] D. Esteve, J.-M. Raimond, and J. Dalibard, editors. *Superconducting qubits and the physics of Josephson junctions*, volume LXXIX of *Les Houches Session*, Amsterdam, 2004. Universite Joseph Fourier, Elsevier.

[47] A. A. Houck et al. *Phys. Rev. Lett.*, 101(080502), 2008.

[48] S. Filipp et al. *Phys. Rev. Lett.*, 102(200402), 2009.

[49] R.P. Feynman and A.R. Hibbs. *Quantum mechanics and path integrals*. McGraw-Hill, 1965.

[50] R. Fisher, F. Helmer, S. J. Glaser, F. Marquardt, and T. Schulte-Herbrüggen. Optimal control of circuit quantum electrodynamics in one and two dimensions. *Phys. Rev. B*, 81(8):085328, Feb 2010.

[51] A.G. Fowler and K. Goyal. Topological cluster state quantum computing. *Quantum Information and Computation*, 9(9-10):0721–0738, 2009.

[52] Austin G. Fowler, Ashley M. Stephens, and Peter Groszkowski. High-threshold universal quantum computation on the surface code. *Phys. Rev. A*, 80(5):052312, 2009.

[53] Austin G. Fowler, William F. Thompson, Zhizhong Yan, Ashley M. Stephens, B. L. T. Plourde, and Frank K. Wilhelm. Long-range coupling and scalable architecture for superconducting flux qubits. *Phys. Rev. B*, 76,:174507, 2007.

[54] R. Freeman. *Spin Choreography: Basic Steps in High Resolution NMR*. Oxford University Press, New York, 1998.

[55] Ray Freeman. Shaped radiofrequency pulses in high resolution nmr. *Journal of Progress in Nuclear Magnetic Resonance Spectroscopy*, 32:59–106, 1998.

[56] A. Gaëtan, Y. Mirosnychenko, T. Wilk, A. Chotia, M. Viteau, D. Comparat, P. Pillet, A. Browaeys, and P. Grangier. Observation of collective excitation of two individual atoms in the rydberg blockade regime. *Nat. Phys.*, 5:115, 2009.

[57] A. Galiautdinov, A.N. Korotkov, and J.M. Martinis. Resonator/zero-qubit architecture for superconducting qubits. arXiv:1105.3997.

[58] Jay Gambetta, Alexandre Blais, David I. Schuster, Andreas Wallraff, L. Frunzio, J. Majer, Michel H. Devoret, Steven M. Girvin, and Robert J. Schoelkopf. Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting. *Phys. Rev. A*, 74(4):042318, 2006.

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

[60] M.R. Geller, E.J. Pritchett, A. Galiautdinov, and J.M. Martinis. Quantum logic with weakly coupled qubits. *Phys. Rev. A*, 81:012320, 2010.

[61] N.A. Gershenfeld and I.L. Chuang. Bulk spin-resonance quantum computing. *Science*, 275:350, 1997.

[62] Daniel Gottesman. *Stabilizer Codes and Quantum Error Correction.* PhD thesis, Caltech, 1997.

[63] Peter Groszkowski, Austin G. Fowler, Felix Motzoi, and Frank K. Wilhelm. Tunable coupling between three qubits as a building block for a superconducting quantum computer, 2011.

[64] U. Haeberlen and J. S. Waugh. Coherent averaging effects in magnetic resonance. *Phys. Rev.*, 175:453, 1968.

[65] H. Haeffner, C.F. Roos, and R. Blatt. Quantum computing with trapped ions. *Phys. Rep.*, 469:155, 2008.

[66] Lucien Hardy. Quantum theory from five reasonable axioms. *quant-ph/0101012*, 2001.

[67] S. Haroche and D. Kleppner. Cavity quantum electrodynamics. *Physics Today*, page 24, 1989.

[68] Serge Haroche and Jean-Michel Raimond. *Exploring the quantum: atoms, cavities and photons.* Oxford University Press, Oxford, 2006.

[69] R. Harris, A. J. Berkley, M. W. Johnson, P. Bunyk, S. Govorkov, M. C. Thom, S. Uchaikin, A. B. Wilson, J. Chung, E. Holtham, J. D. Biamonte, A. Yu. Smirnov, M. H. S. Amin, and Alec Maassen van den Brink. Sign- and magnitude-tunable coupler for superconducting flux qubits. *Phys. Rev. Lett.*, 98(17):177001, Apr 2007.

[70] T. Hime, PA Reichardt, BLT Plourde, TL Robertson, C.E. Wu, AV Ustinov, and J. Clarke. Solid-state qubits with current-controlled coupling. *Science*, 314(5804):1427–1429, 2006.

[71] Max Hofheinz, H. Wang, Markus Ansmann, Radoslaw C. Bialczak, Erik Lucero, Matthew Neeley, A. D. O'Connell, D. Sank, J. Wenner, John M. Martinis, and A. N. Cleland. Synthesizing arbitrary quantum states in a superconducting resonator. *Nature (London)*, 459(7246):546–9, May 2009.

[72] D.I. Hoult. The solution of the bloch equations in the presence of a varying b1 field — an approach to selective pulse analysis. *J. Magn. Res.*, 35:69, 1979.

[73] R. Hughes, the ARDA Quantum Information Science, and Technology Panel. A quantum information science and technology roadmap. available at http://qist.lanl.gov.

[74] K. Husimi. Miscellanea in elementary quantum mechanics, ii. *Prog. Theor. Phys.*, 9:381, 1953.

[75] C. Hutter, A. Shnirman, Y. Makhlin, and G. Schön. Tunable coupling of qubits: Nonadiabatic corrections. *Europhysics Letters*, 74(6):1088–1094, 2006.

[76] Carr HY and Purcell EM. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Review*, 3:630–638, 1954.

[77] A. P. Young Itay Hen. Exponential complexity of the quantum adiabatic algorithm for certain satisfiability problems. *Phys. Rev. E 84*, 84:061152, 2011.

[78] E. T. Jaynes and Frederick W. Cummings. Comparison of quantum and semiclassical radiation theories with application to the beam maser. *Proc. IEEE*, 51(1):89, 1963.

[79] K.B. Whaley J.I. Korsbakken, F.K. Wilhelm. The size of macroscopic superposition states in flux qubits. *EPL (Europhysics Letters)*, 89(30003), 2010.

[80] H. Jirari, F.W.J. Hekking, and O. Buisson. Optimal control of superconducting n-level quantum systems. arXiv:0903.4028.

[81] K. G. Vollbrecht D. Poulin K. Temme, T. J. Osborne and F. Verstraete. Quantum metropolis sampling. *Nature*, 471:87–90, 2011.

[82] Mattan Kamon, Michael J. Tsuk, and Jacob White. Fasthenry: A multipole-accelerated 3-d inductance extraction program. In *Design Automation Conference*, pages 678–683, 1993.

[83] J F Kanem, S Maneshi, S H Myrskog, and A M Steinberg. Phase space tomography of classical and nonclassical vibrational states of atoms in an optical lattice. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(12):S705, 2005.

[84] P. Kaye, R. Laflamme, and M. Mosca. *An Introduction to Quantum Computing*. Oxford University Press, Oxford, 2007.

[85] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S.J. Glaser. Optimal control of coupled spin dynamics: Design of nmr pulse sequences by gradient ascent algorithms. *J. Magn. Reson.*, 172:296, 2005.

[86] B Khani, J M Gambetta, F Motzoi, and Frank K. Wilhelm. Optimal generation of fock states in a weakly nonlinear oscillator. *Physica Scripta*, T137:014021 (5pp), 2009.

[87] Botan Khani. Robust time-optimal control for the one-dimensional optical lattice for quantum computation. Master's thesis, University of Waterloo, 2011.

[88] Botan Khani, Seth T. Merkel, Felix Motzoi, Jay M. Gambetta, and Frank K. Wilhelm. High-fidelity quantum gates in the presence of dispersion. *Physical Review A*, 85:022306, 2012.

[89] Kaveh Khodjasteh and Daniel A. Lidar. Rigorous bounds on the performance of a hybrid dynamical decoupling-quantum computing scheme. *Phys. Rev. A*, 78:012355, 2008.

[90] A.Y. Kitaev. Quantum measurements and the Abelian stabilizer problem. *Arxiv preprint quant-ph/9511026*, 1995.

[91] A.Y. Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, 1997.

[92] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, 2001.

[93] E. Knill, R. Laflamme, and W. Zurek. Accuracy threshold for quantum computation. *Science,*, 279(5349):342–345, 1998.

[94] 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. Charge-insensitive qubit design derived from the cooper pair box. *Phys. Rev. A*, 76(4):042319, 2007.

[95] Jens Koch, Terri M. Yu, Jay Gambetta, Andrew A. Houck, David I. Schuster, J. Majer, Alexandre Blais, Michel H. Devoret, Steven M. Girvin, and Robert J. Schoelkopf. Charge-insensitive qubit design derived from the cooper pair box. *Phys. Rev. A*, 76(4):042319, 2007.

[96] W. Kohn. Analytic properties of bloch waves and wannier functions. *Physical Review*, 115:809–821, 1959.

[97] I. Kozinsky, H.W.Ch. Postma, O. Kogan, A. Husain, and M.L. Roukes. Basins of attraction of a nonlinear nanomechanical resonator. *Phys. Rev. Lett.*, 99:207201, 2007.

[98] L.D. Landau. A theory of energy transfer on collisions. *Z. Phys. Sowjetunion*, 2:46, 1932.

[99] L.D. Landau and E.M. Lifshitz. *Mechanics*, volume 1 of *Course of Theoretical Physics*. Butterworth-Heinemann, Burlington, MA, 1982.

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

[101] M. Levitt. *Spin Dynamics. Basics of Nuclear Magnetic Resonance*. Wiley, Chichester, 2nd edition, 2008.

[102] D. A. Lidar, I. L. Chuang, and K. B. Whaley. Decoherence-free subspaces for quantum computation. *Phys. Rev. Lett.*, 81(12):2594–2597, Sep 1998.

[103] Frank Wilhelm Likun Hu. Higher norm fidelity functions with extra weight given to worse-case solutions. 2009.

[104] Klaus Mølmer Line Hjortshøj Pedersen, Niels Martin Møller. Fidelity of quantum operations. *Physics Letters A*, 367:47–51, 2007.

[105] E. Lucero, J. Kelly, R.C. Bialczak, M. Lenander, M. Mariantoni, M. Neeley, A.D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, T. Yamamoto, A.N. Cleland, and J.M. Martinis. Reduced phase error through optimized control of a superconducting qubit. *Phys. Rev. A*, 82:042339, 2010.

[106] Erik Lucero, Max Hofheinz, Markus Ansmann, Radoslaw C. Bialczak, N. Katz, Matthew Neeley, A. D. O'Connell, H. Wang, A. N. Cleland, and John M. Martinis. High-fidelity gates in a single josephson qubit. *Phys. Rev. Lett.*, 100(24):247001, 2008.

[107] Erik Lucero, Julian Kelly, Radoslaw C. Bialczak, Mike Lenander, Matteo Mariantoni, Matthew Neeley, A.D. O'Connell, Daniel Sank, H. Wang, Martin Weides, James Wenner, Tsuyoshi Yamamoto, A.N. Cleland, and John Martinis. Reduced phase error through optimized control of a superconducting qubit. arXiv:1007.1690, 2010.

[108] A. Lupascu, E. F. C. Driessen, L. Roschier, C. J. P. M. Harmans, and J. E. Mooij. High-contrast dispersive readout of a superconducting flux qubit using a nonlinear resonator. *Phys. Rev. Lett.*, 96(12):127003, 2006.

[109] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquieres, A. Gruslys, S. Schirmer, and T. Schulte-Herbrüggen. Comparing, optimising and benchmarking quantum control algorithms in a unifying programming framework. arXiv:1011.4874v2, 2010.

[110] W. Magnus. On the exponential solution of differential equations for a linear operator. *Comm. Pure and Appl. Math.*, VII (4):649–673, 1954.

[111] 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. Coupling superconducting qubits via a cavity bus. *Nature*, 449:443, 2007.

[112] Yu. Makhlin, G. Schön, and A. Shnirman. Quantum-state engineering with Josephson-junction devices. *Rev. Mod. Phys.*, 73:357, 2001.

[113] Yuriy Makhlin, Gerd Schön, and Alexander Shnirman. Quantum-state engineering with josephson-junction devices. *Rev. Mod. Phys.*, 73(2):357–400, 2001.

[114] S. Maneshi, J.F. Kanem, C. Zhuang, M. Partlow, and A.M. Steinberg. Efficient vibrational state coupling in an optical tilted-washboard potential via multiple spatial translations and application to pulse echo. *Phys. Rev. A*, 77:022303, 2008.

[115] Samansa Maneshi, Chao Zhuang, Christopher R. Paul, Luciano S. Cruz, and Aephraim M. Steinberg. Coherence freeze in an optical lattice investigated via pump-probe spectroscopy. *Phys. Rev. Lett.*, 105:193001, Nov 2010.

[116] J.M. Martinis. Superconducting phase qubits. *Quantum Information Processing*, 8:81, 2009.

[117] J.M. Martinis, S. Nam, J. Aumentado, and C. Urbina. Rabi oscillations in a large Josephson-junction qubit. *Phys. Rev. Lett.*, 89(11):117901, 2002.

[118] J. Mazur. *Zeno's Paradox: Unraveling the Ancient Mystery Behind the Science of Space and Time*. Plume, 2007.

[119] R. Miller, T. E. Northup, K. M. Birnbaum, A. Boca, A. D. Boozer, and H. J. Kimble. Trapped atoms in cavity qed: coupling quantized light and matter. *J. Phys. B*, 38(9):–551, 2005.

[120] C. Moler and C. van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, 45:3, 2003.

[121] Mikko Möttönen, Rogerio de Sousa, Jun Zhang, and K. Birgitta Whaley. High-fidelity one-qubit operations under random telegraph noise. *Phys. Rev. A*, 73(2):022332, Feb 2006.

[122] F. Motzoi, J. M. Gambetta, P. Rebentrost, and Frank K. Wilhelm. Simple pulses for elimination of leakage in weakly nonlinear qubits. *Phys. Rev. Lett.*, 103(11):110501, 2009.

[123] F. Motzoi, J.M. Gambetta, S.T. Merkel, and F.K. Wilhelm. Optimal control methods for fast time-varying hamiltonians. *Physical Review A*, 84(022307), 2011.

[124] S. H. Myrskog, J. K. Fox, M. W. Mitchell, and A. M. Steinberg. Quantum process tomography on vibrational states of atoms in an optical lattice. *Phys. Rev. A*, 72(1):013615, Jul 2005.

[125] M. Nakahara, S. Kanemitsu, and M.M. Salomaa, editors. *Physical Realizations of Quantum Computing: Are the DiVincenzo Criteria Fulfilled in 2004?*, Singapore, 2006. WorldScientific.

[126] Y. Nakamura, Y.A. Pashkin, and J.S. Tsai. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature*, 398:786, 1999.

[127] M. Neeley, R. C. Bialczak, M. Lenander, E. Lucero, M. Mariantoni, A. D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, Y. Yin, T. Yamamoto, A. N. Cleland, and J. M. Martinis. Generation of three-qubit entangled states using superconducting phase qubits. arXiv:1004.4246, 2010.

[128] Matthew Neeley, Markus Ansmann, Radoslaw C. Bialczak, Max Hofheinz, Erik Lucero, Aaron D. O'Connell, Daniel Sank, Haohua Wang, James Wenner, Andrew N. Cleland, Michael R. Geller, and John M. Martinis. Emulation of a quantum spin with a superconducting phase qudit. *Science*, 325(5941):722–725, 2009.

[129] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, UK, 2000.

[130] J. Niemeyer. Josephson arrays for DC and AC metrology. *Supercond. Sci. and Technology*, 13:546, 2000.

[131] R. Nigmatullin and S.G. Schirmer. Implementation of fault-tolerant quantum logic gates via optimal control. *New. J. Phys*, 11:105032, 2009.

[132] A.O. Niskanen, K. Harrabi, F. Yoshihara, Y. Nakamura, S. Lloyd, and J.S. Tsai. Quantum coherent tunable coupling of superconducting qubits. *Science*, 316:5825, 2007.

[133] A.O. Niskanen, Y. Nakamura, and J.S. Tsai. Tunable coupling scheme for flux qubits at the optimal point. *Phys. Rev. B*, 73:094506, 2006.

[134] T. P. Orlando, J. E. Mooij, L. Tian, C.H. van der Wal, L. S. Levitov, S. Lloyd, and J. J. Mazo. Superconducting persistent-current qubit. *Phys. Rev. B*, 60:15398, 1999.

[135] T.P. Orlando, L. Tian, D.S. Crankshaw, S. Lloyd, C.H. van der Wal, J.E. Mooij, and F.K. Wilhelm. Engineering the quantum-measurement process for the persistent current qubit. *Phyisca C*, 368:294, 2003.

[136] S. Montangero . P. Doria, T. Calarco. arxiv:1003.3750. 2010.

[137] Roger Penrose. *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford University Press, 1989.

[138] B. Peropadre, P. Forn-Díaz, E. Solano, and J. J. García-Ripoll. Switchable ultrastrong coupling in circuit qed. *Phys. Rev. Lett.*, 105(2):023601, Jul 2010.

[139] H.R. Petta, A.C. Johnson, J.M. Taylor, E.A. Laird, A. Yacoby, M.D. Lukin, C.M. Marcus, M.P. Hanson, and A.C. Gossard. Coherent manipulation of coupled electron spins in semiconductor quantum dots. *Science*, 309:2180, 2005.

[140] R.A. Pinto, A.N. Korotkov, M.R. Geller, V.S. Shumeiko, and J.M. Martinis. Analysis of a tunable coupler for superconducting phase qubits. *Phys. Rev. B*, 82, 2010.

[141] J. H. Plantenberg, P. C. de Groot, C. J. P. M. Harmans, and J. E. Mooij. Demonstration of controlled-not quantum gates on a pair of superconducting quantum bits. *Nature*, 447(7146):836–839, 2007.

[142] B.L.T. Plourde, J. Zhang, K.B. Whaley, F.K. Wilhelm, T.L. Robertson, T. Hime, S. Linzen, P.A. Reichardt, C.-E. Wu, and J. Clarke. Entangling flux qubits with a bipolar dynamic inductance. *Phys. Rev. B*, 70:140501(R), 2004.

[143] BLT Plourde, J. Zhang, KB Whaley, FK Wilhelm, TL Robertson, T. Hime, S. Linzen, PA Reichardt, C.E. Wu, and J. Clarke. Entangling flux qubits with a bipolar dynamic inductance. *Phys. Rev. B*, 70(14):140501, 2004.

[144] A. Poudel and M.G. Vavilov. Effect of ohmic environment on optimally controlled flux-biased phase qubit. arXiv:1008.2554.

[145] H. A. Rabitz, M. M. Hsieh, and C. M. Rosenthal. Quantum optimally controlled landscapes. *Science*, 303:1998–2001, March 2004.

[146] R. Raussendorf and J. Harrington. Fault-tolerant quantum computation with high threshold in two dimensions. *Phys. Rev. Lett.*, 98(19):190504, 2007.

[147] P. Rebentrost, I. Serban, T. Schulte-Herbrüggen, and F. K. Wilhelm. Optimal control of a qubit coupled to a non-markovian environment. *Phys. Rev. Lett.*, 102(9):090401, Mar 2009.

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

[149] C. Rigetti, A. Blais, and M. Devoret. Protocol for universal gats in optimally biased superconducting qubits. *Phys. Rev. Lett.*, 94:240502, 2005.

[150] Walter Pötz Robert Roloff, Markus Wenin. Optimal control for open quantum systems: Qubits and quantum gates. *Journal reference: J. Comput. Theor. Nanosci.*, 6:1837, 2009.

[151] Meiboom S. and Gill D. Modified spin-echo method for measuring nuclear relaxation times. *The Review of Scientific Instruments*, 29(8):688–691, 1958.

[152] J.J. Sakurai. *Advanced Quantum Mechanics*. Addison-Wesley, New York, 1967.

[153] R.J. Schoelkopf and S.M. Girvin. Wiring up quantum systems. *Nature*, 451:664, 2008.

[154] J. A. Schreier, Andrew A. Houck, Jens Koch, David I. Schuster, B. R. Johnson, J. M. Chow, Jay M. Gambetta, J. Majer, L. Frunzio, Michel H. Devoret, Steven M. Girvin, and Robert J. Schoelkopf. Suppressing charge noise decoherence in superconducting charge qubits. *Phys. Rev. B*, 77(18):180502, 2008.

[155] Norbert Schuch and Jens Siewert. Natural two-qubit gate for quantum computation using the xy interaction. *Phys. Rev. A*, 67(3):032301, 2003.

[156] T. Schulte-Herbrueggen, A. Spoerl, and S.J. Glaser. Quantum cisc compilation by optimal control and scalable assembly of complex instruction sets beyond two-qubit gates, 2007. arXiv:0712.3227.

[157] T. Schulte-Herbrueggen, A. Spoerl, N. Khaneja, and S. J. Glaser. Optimal control for generating quantum gates in open dissipative systems. *J. Phys. B: At. Mol. Opt. Phys.*, 44:154013, 2011.

[158] P. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. *Proceedings 35th Annual Symposium on Foundations of Computer Science*, page 124, 1994.

[159] M.A. Sillanpää, J.I. Park, and R.W. Simmonds. Coherent quantum state storage and transfer between two phase qubits via a resonant cavity. *Nature*, 449:438, 2007.

[160] Michael Sipser. *Introduction to the Theory of Computation*. Course Technology, 2006.

[161] Thomas E. Skinner and Naum I. Gershenzon. Optimal control design of pulse shapes as analytic functions. *Journal of Magnetic Resonance*, 204(2):248 – 255, 2010.

[162] C.P. Slichter. *Principles of magnetic resonance*. Number 1 in Series in Solid State Sciences. Springer, Berlin, 3. edition, 1996.

[163] Andrew Steane. Simple quantum error correcting codes. *Phys. Rev. A*, 54:4741–4751, 1996.

[164] M. Steffen, M. Ansmann, Radoslaw C. Bialczak, N. Katz, Erik Lucero, R. McDermott, Matthew Neeley, E. M. Weig, A. N. Cleland, and J.M. Martinis. Measurement of the Entanglement of Two Superconducting Qubits via State Tomography. *Science*, 313(5792):1423–1425, 2006.

[165] Matthias Steffen, Shwetank Kumar, David P. DiVincenzo, J.R. Rozen, George A. Keefe, Mary Beth Rothwell, and Mark B. Ketchen. High coherence hybrid superconducting qubit. arXiv:1003.3054v1, 2010.

[166] Matthias Steffen, John M. Martinis, and Isacc L. Chuang. Accurate control of josephson phase qubits. *Phys. Rev. B*, 68(22):224518, 2003.

[167] E.C.G. Stückelberg. Theory of inelastic collisions between atoms. *Helv. Phys. Acta*, 5(369), 1932.

[168] R. Tycko. Broadband population inversion. *Phys. Rev. Lett.*, 51(9):775777, 1983.

[169] E. Urban, T.A. Johnson, T. Henage, L. Isenhower, D.D. Yavuz, T.G. Walker, and M. Saffman. Observation of rydberg blockade between two atoms. *Nat. Phys.*, 5:110, 2009.

[170] S. H. W. van der Ploeg, A. Izmalkov, Alec Maassen van den Brink, U. Huebner, M. Grajcar, E. Il'ichev, H. G. Meyer, and A. M. Zagoskin. Controllable coupling of superconducting flux qubits. *Phys. Rev. Lett.*, 98:057004, 2007.

[171] C.H. van der Wal, A.C.J. ter Haar, F.K. Wilhelm, R.N. Schouten, C.J.P.M.Harmans, T.P. Orlando, S. Lloyd, and J.E.Mooij. Quantum superposition of macroscopic persistent-current states. *Science*, 290:773, 2000.

[172] G. Vidal and C. M. Dawson. Universal quantum circuit for two-qubit transformations with three controlled-not gates. *Phys. Rev. A*, 69(1):010301, 2004.

[173] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and Michel H. Devoret. Manipulating the quantum state of an electrical circuit. *Science*, 296:886, 2002.

[174] R.L. Vold, J.S. Waugh, M.P. Klein, and D.E. Phelps. *J. Chem. Phys.*, 43:3831, 1968.

[175] Andreas Wallraff, David I. Schuster, Alexandre Blais, L. Frunzio, Ren-Shou Huang, J. Majer, S. Kumar, Steven M. Girvin, and Robert J. Schoelkopf. Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature (London)*, 431(7005):162–167, 2004.

[176] David S. Wang, Austin G. Fowler, and Lloyd C. L. Hollenberg. Surface code quantum computing with error rates over 1%. *Phys. Rev. A*, 83(2):020302, Feb 2011.

[177] W.S. Warren. Effect of arbitrary laser or nmr pulse shapes on population inversion and coherence. *J. Chem. Phys.*, 81:5437, 1984.

[178] L. E. Weaver and D. C. Broughto. Gaussian filters for pulse shaping. *Radio and Electronic Engineer*, 41:457, 1971.

[179] G. Wendin and V.S. Shumeiko. *Handbook of Theoretical and Computational Nanotechnology*, chapter Superconducting Quantum Circuits, Qubits and Computing. American Scientific Publishers, 2006.

[180] M. Wenin and W. Pötz. State-independent control theory for weakly dissipative quantum systems. *Physical Review A*, 78(012358), 2008.

[181] S. Wimperis. Broadband, passband, and narrowband pulses for use in advanced nmr experiments. *J. Magn. Reson., Ser. A*, 109:221, 1994.

[182] L.-A. Wu and D.A. Lidar. Qubits as parafermions. *J. Math. Phys.*, 43:4506, 2002.

[183] F. Yoshihara, Y. Nakamura, and J. S. Tsai. Correlated flux noise and decoherence in two inductively coupled flux qubits. *Phys. Rev. B*, 81(13):132502, Apr 2010.

[184] J. Q. You and Franco Nori. Quantum information processing with superconducting qubits in a microwave field. *Phys. Rev. B*, 68(6):064509, 2003.

[185] J.Q. You, Y. Nakamura, and F. Nori. Fast two-bit operations in inductively coupled flux qubits. *Phys. Rev. B*, 71:0242532, 2005.

[186] J.Q. You and F. Nori. Superconducting circuits and quantum information. *Phys. Today*, 58:42, 2005.

[187] P. Zanardi and M. Rasetti. Noiseless quantum codes. *Phys. Rev. Lett.*, 79(17):3306–3309, Oct 1997.

[188] C. Zener. Non-adiabatic crossing of energy levels. *Proc. R. Soc. Lond. A*, 137:696, 1932.

[189] J. Zhang, J. Vala, S. Sastry, and K.B. Whaley. Geometric theory of nonlocal two-qubit operations. *Phys. Rev. A*, 67:042313, 2003.