UMBC

DISSERTATION APPROVAL SHEET

Title of Dissertation: Investigating the geometric gate robustness conjecture and generating robust entangling gates using methods of dynamical error suppression

Name of Candidate: Ralph Kenneth Colmenar

Doctor of Philosophy, 2022

Graduate Program: Physics

Dissertation and Abstract Approved:

Jason Kestner

Jason Kestner

Associate Professor

Department of Physics

4/26/2022 | 7:50:32 AM PDT

NOTE: *The Approval Sheet with the original signature must accompany the thesis or dissertation. No terminal punctuation is to be used.

ABSTRACT

Title of dissertation: INVESTIGATING THE GEOMETRIC GATE ROBUSTNESS CONJECTURE AND GENERATING ROBUST ENTANGLING GATES USING METHODS OF DYNAMICAL ERROR SUPPRESSION

Ralph Kenneth L. Colmenar,
Doctor of Philosophy, 2022

Dissertation directed by: Professor Jason Kestner,
Department of Physics

Noisy intermediate-scale quantum devices suffer primarily from coherent systematic error and stand to benefit from robust quantum control. To this end, methods of dynamical error suppression provide a means to achieve gate error rates that may be sufficient for fault-tolerant quantum computing. We present in this work a collection of theoretical studies that focuses on two methods in particular: composite pulse sequences and filter function formalism. We first employ composite pulses to reduce the effects of static systematic errors in coupled solid-state qubits. We then use the filter function formalism to analyze geometric gates. In particular, we disprove the geometric gate robustness conjecture which states that geometric gates are intrinsically more robust against certain errors than dynamical gates since their accumulated phase is related to some global geometric property of the system's evolution. Finally, we use dynamical invariant theory to develop a framework that is ideal for filter function engineering. We also use this framework to further analyze some properties of geometric gates.

INVESTIGATING THE GEOMETRIC GATE ROBUSTNESS CONJECTURE AND GENERATING ROBUST ENTANGLING GATES USING METHODS OF DYNAMICAL ERROR SUPPRESSION

by

Ralph Kenneth L. Colmenar

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, Baltimore County, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Professor Jason Kestner, Chair/Advisor
Professor Sebastian Deffner
Professor James Franson
Professor Todd Pittman
Dr. Gregory Quiroz

To my family

Acknowledgments

First, I would like to thank my advisor, Dr. Jason Kestner, for all his support and guidance throughout my graduate career. I am especially thankful for his patience and willingness to make time for me whenever I needed help. I am incredibly honored to have worked with and learned from such an amazing mentor.

I would like to thank Dr. Sebastian Deffner, Dr. James Franson, Dr. Todd Pittman, and Dr. Gregory Quiroz for sparing their invaluable time and agreeing to serve on my Ph.D. committee.

Next, I would like to thank my friend and colleague Dr. Utkan Güngördü for serving as my secondary mentor. Working with him was incredibly enriching and I will always treasure all the neat math and programming tricks he taught me. I highly appreciate him for taking the time to respond to my questions in detail even during the weekends. I also want to thank my fellow graduate students. I am grateful for their company during social hours and for all of our discussions, both academic and otherwise. I am especially grateful to Dr. Fernando Calderon-Vargas for helping me during my first year of research. I would also like to thank my longtime friends, Steven and Mar Vyn, for their encouragement and for giving me a semblance of normal social life outside academia.

Finally, I want to express my deepest gratitude to my family for their unconditional love and support. A special thanks goes to our dogs, Princess and Teddy. I am forever grateful for their companionship which helped keep my sanity in check especially during the height of the COVID-19 pandemic.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The commercial success of the semiconductor industry ushered in by the invention of the first transistor-based digital computer led to the ubiquitous presence of computers in our everyday lives. It has become an integral part of our society, so much so that we use it for leisure, transportation, communication, security, education, and many others. One particular application of interest is in research. Modern computers can now perform laborious calculations in a matter of minutes that would otherwise take hours, if not days, to do by hand. This proved a boon across all scientific disciplines. Motivated by the prospect of even higher commercial, scientific, and technological returns, chip fabrication techniques were further refined to a point where, in just a few decades, one can purchase consumer-level processing chips with tens of billion of transistors[1].

Unsurprisingly, solving very complex problems also demanded more computational power. Even before digital computers existed, it was well known that certain systems are impossible to simulate efficiently when enough particles are involved. Indeed, fields of physics such as statistical mechanics and condensed matter physics

[1] Announced in late 2021, Apple's M1 Max chip is equipped with 57 billion transistors!

acknowledge the futility of keeping track of all particle interactions and one must resort to using clever approximations to gain physical insight. This problem is even more pronounced in quantum systems since correlations between quantum states require exponentially more resources to track. In fact, the number of bits required to simulate a system with several hundreds of quantum particles exceeds the estimated number of atoms in the observable universe [1]. It was clear that a new computational paradigm was necessary to tackle such problems. To this end, the idea of using a machine that operates under the rules of quantum mechanics was proposed [2–4]. By taking advantage of quantum phenomena such as superposition and entanglement, it was theorized that quantum computers can outperform traditional computers at certain tasks. This was later proven true in the 1990s after some pioneering works by the likes of Deutsch [5], Simon [6], Shor [7], and Grover [8] which consequently set about a race to develop the very first quantum computer.

While there are different models for doing quantum computation, we will focus on gate-based quantum computing which is viewed as the natural quantum analog of traditional computing. This type of quantum computer has since been proposed in a plethora of physical platforms including photonic systems [9], trapped ions [10], nuclear spin in nuclear magnetic resonance (NMR) [11, 12], electron charge [13] and spin [14] in semiconducting devices, and Josephson junctions in superconducting devices [15–17] among others. Although the underlying physics vary with the particular implementation, all of these platforms satisfy the Divincenzo criteria which outlines the necessary conditions for quantum computation [18]. Put simply, any candidate for a quantum computer must be able to encode its information in a group of well-isolated, interacting 2-state systems whilst having the ability to preserve, measure, and arbitrarily manipulate its state with high accuracy.

Although decades of intense research has brought about the current era of noisy intermediate-scale quantum (NISQ) devices [19], a fully functional quantum computer remains elusive. One of the biggest challenges in making a quantum computer is noise. Noise can manifest in different forms. For example, it can be due to random driving forces from the environment (e.g. Brownian motion), interactions that produce entanglement between the system and the environment, or statistical imprecision in the experimental controls on the system (e.g., timing errors, frequency fluctuations, etc.) [20]. Thus, noise suppression is paramount for any reliable quantum information processor. To understand how this can be done, it is instructive to examine how noise is handled in traditional computing. Suppose we have three bits that are susceptible to bit-flip error where a 0 can randomly flip to 1 and vice versa. Information stored in these bits will inevitably get corrupted if any one of the bit flips its value. On the other hand, suppose that we encode our logical bit states, 0_{L} and 1_{L}, using these three noisy bits [1]. In other words, suppose

$$0_{\mathrm{L}} \longrightarrow 000, \qquad 1_{\mathrm{L}} \longrightarrow 111.$$

Even if any one of the noisy bit flips its value, the encoded information can still be recovered. For sufficiently low bit-flip probability, the information may be recovered through a simple majority ruling. As an example, a 001 corresponds to a 0_{L} that encountered a bit-flip error on the third bit and can be corrected by simply flipping the corresponding state back to the majority bit value which is 0[2].

[2]Measurement of classical bit states is trivial and does not compromise computation. This is no

We would like to employ a similar strategy to correct errors in quantum bits, or qubits. Before we proceed, however, it is worth highlighting some key properties of error correction in the previous example. First, it is necessary to have sufficiently low bit-flip rate for each individual bit. The probability of two bit-flip errors occurring simultaneously has to be much less than a certain error threshold. Otherwise, the majority ruling will fail and the information will get corrupted. Second, error correction requires the encoding of logical information onto multiple physical bits. That is, we trade some computational power in favor of reliability. Fortunately, this no longer a concern for modern processors with billions of transistors to leverage.

Quantum error correction inherits these properties as well. This is somewhat problematic since current NISQ devices offer significantly less flexibility, having only access to as much as hundreds of physical qubits even in state-of-the-art systems[3]. Some argue that the benefits of quantum computing will not be seen until we have quantum computers with at least 100 error-corrected logical qubits. This can translate to over 1000 physical qubits depending on the encoding[4]. Such a feat would require a significant investment in device engineering research and it may take many years to reach a point where fabrication methods can reliably produce quantum processors with thousands of qubits. In addition, just like its classical counterpart, quantum error correction is contingent on the error rate of qubit operations be-

longer the case in qubits since measurements destroy the quantum state under observation, thus making information recovery impossible. Luckily, there are other means to detect errors [1, 20].

[3]On November 2021, IBM Quantum unveiled Eagle, a 127-qubit quantum processor.

[4]For example, Shor's code requires 9 physical qubits and some additional help from ancillary qubits.

ing less than a certain threshold [21]. The least stringent quantum error-correction schemes require qubits to have no more than 1% error per gate [22, 23]. Lowering the average gate error would reduce the overhead cost of implementing error correction in terms of the number of physical qubits required per logical qubit. This can be achieved by adopting better and more robust quantum control protocols which is the primary focus of this dissertation.

We present here a collection of theoretical studies on robust quantum control that primarily involve the use of composite pulse sequences and filter function engineering. In essence, composite pulse sequence [24–28] involves replacing a noisy quantum gate with a sequence of noisy quantum gates such that i) the sequence produces the same intended gate as the original, and ii) the sequence is more robust against noise than the original. The second condition is usually satisfied by carefully designing the sequence components so that the first-order cumulative effect of noise averages to zero. This technique was first pioneered in NMR systems to prolong spin coherence times and has since been routinely used to suppress certain types of coherent systematic errors in many qubit implementations [29–38].

Composite pulses are designed to be robust against noise that fluctuate much slower than the total duration of the sequence, i.e., quasistatic noise. However, noise that fluctuate more rapidly plague many systems as well. Chief among them is $1/f$ noise [39] which is prevalent in solid-state platforms [40–45]. To this end, techniques for characterizing the effects of noise when given the noise power spectral density (PSD) were developed. The filter function formalism [46–48] is one of such error characterization scheme. The general idea is to quantify the net susceptibility of a given control protocol to noise using the overlap in frequency between the noise PSD and the spectral characteristics of the modulation imparted by the control. Thus, one may develop a robust quantum control scheme based on the idea of filter function engineering where the goal is to minimize the frequency overlap.

1.1 Organization of the dissertation

The dissertation is organized as follows. Chapter 2 reviews some of the basic principles of quantum computing and other pertinent concepts. Chapter 3 examines how to dynamically correct a cross-resonance (CR) gate in a system of coupled transmon qubits using composite pulses. Similarly, Chapter 4 focuses on producing a robust entangling operation for a system of capacitively-coupled singlet-triplet qubits with composite pulses in addition to control parameter optimization. Chapter 5 examines a special class of quantum gates called geometric gates. The conjectured intrinsic robustness of geometric gate is invalidated by showing the existence of geometric and dynamical quantum gates with identical filter functions. Chapter 6 develops a framework using filter functions and dynamical invariant theory to efficiently engineer 1-qubit quantum gates that are robust against a specified noise PSD. Finally, Chapter 7 provides the conclusions and summary of the dissertation.

Chapter 2

Quantum Computing Basics

How does one build a quantum computer? This question was addressed by David DiVincenzo in his seminal work where he outlined five requirements that are necessary for the implementation of quantum computation [18]:

- A scalable physical system with well characterized qubits.
- The ability to initialize the state of the qubits to a simple fiducial state such as $|000\ldots\rangle$.
- Long relevant decoherence times, much longer than the gate operation time.
- A "universal" set of quantum gates.
- A qubit-specific measurement capability.

While many leading candidate platforms for a quantum computer have satisfied most of these conditions, it remains very challenging to develop quantum gates with low enough error rates to make quantum error correction feasible. Developing quantum control protocols that are robust against noise can help solve this issue. In this chapter, we briefly review some basic concepts in quantum computing and quantum control.

2.1 Qubits

In classical information theory, the basic unit of information is a bit which can assume one of two distinct values: 0 or 1. The analogous quantity for quantum information theory is a qubit. A qubit is nothing more than a two-level quantum system. The two levels can be physically represented by spin orientation, location of a charged particle in a double-well potential, photon polarization, or flow of superconducting current, among other things. More abstractly, a qubit is an element of the vector space $\mathbb{C}^2$ whose basis vectors are denoted by

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.1}$$

Taking spin orientation as an example, the basis states $|0\rangle$ and $|1\rangle$ can represent the $|\uparrow\rangle$ and $|\downarrow\rangle$ spin states, respectively. One key feature of a qubit that is absent in a classical bit is quantum superposition. An arbitrary qubit state can be expressed as

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle \text{ with } a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1. \tag{2.2}$$

Although a qubit may take infinitely many different states (unlike a bit which can only take two), it contains as much information as a classical bit. Indeed, a measurement to determine whether the qubit is in the state $|0\rangle$ or $|1\rangle$ will result in $|0\rangle$ ($|1\rangle$) with probability $|a|^2$($|b|^2$).

An alternative way of viewing the qubit state is through the Bloch sphere picture. Ignoring the global phase factor, we parameterize $|\psi\rangle$ as

$$|\psi(\theta,\phi)\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle \tag{2.3}$$

Note that we must restrict the values to $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$ to guarantee a unique mapping. This particular form can be thought of as the $+1$ eigenstate of $\vec{n}(\theta,\phi)\cdot\vec{\sigma}$, where the unit vector $\vec{n}$ is

$$\vec{n}(\theta,\phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \tag{2.4}$$

and $\vec{\sigma} = (\sigma_X, \sigma_Y, \sigma_Z)$ is a vector comprising of the Pauli matrices

$$\sigma_X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.5}$$

Thus, the collection of all pure states forms the surface of the Bloch sphere and all the points inside correspond to mixed states. This picture provides a simple way to visualize the action of a quantum gate since they can be interpreted as simple rotations on the Bloch sphere. Unfortunately, this interpretation does not generalize well beyond the 1-qubit example [49].

Let us now consider a group of n qubits. This is often referred to as a quantum register in quantum computing. For a classical register, the state of the system is completely determined by specifying the state of each individual bit. This is no longer the case for a quantum register due to the quantum mechanical phenomenon called quantum entanglement. To demonstrate this, suppose we consider an n-qubit register. If we specify its state in a similar manner as in a classical register, each qubit would be described by a $\mathbb{C}^2$ vector of the form $a_i|0\rangle + b_i|1\rangle$ for a total of $2n$

complex numbers. This yields

$$(a_1 |0\rangle + b_1 |1\rangle) \otimes (a_2 |0\rangle + b_2 |1\rangle) \otimes \cdots \otimes (a_n |0\rangle + b_n |1\rangle), \tag{2.6}$$

where $\otimes$ denotes the Kronecker product operator. More generally, quantum mechanics can allow an n-qubit register to be described by the superposition of such tensor product states

$$|\psi\rangle = \sum_{c_i=0,1} a_{c_1 c_2 \cdots c_n} |c_1\rangle \otimes |c_2\rangle \otimes \cdots \otimes |c_n\rangle. \tag{2.7}$$

It is important to note that there exists superpositions that cannot be decomposed to a form like in Equation (2.6). These states are said to be entangled. Furthermore, the states of an n-qubit register live in a 2^n-dimensional vector space. Even with just $n = 500$, 2^n becomes astronomical and exceeds the estimated number of atoms in the observable universe [1]. This means that it is impossible to store the information of a highly entangled state with sufficiently large n in a traditional computer. Therefore, quantum entanglement and superposition provide an extremely powerful computational resource that is absent in traditional computing.

2.2 Quantum Gates, Gate Decomposition, and Universality

Quantum gates are responsible for manipulating the state of the quantum register during computation. Physically, they correspond to the dynamical evolution of the quantum system representing the qubit. In quantum mechanics, this evolution is governed by the (nonrelativistic) Schrödinger equation[1]

$$i\dot{\mathcal{U}}(t) = H(t)\mathcal{U}(t), \qquad \mathcal{U}(0) = \mathbb{1}, \tag{2.8}$$

[1] We shall adopt natural units for the entire dissertation unless stated otherwise.

where $H(t)$ is the Hamiltonian of the system and $\mathcal{U}(t)$ is the time-evolution operator. A quantum gate is defined to be the action of the operator $\mathcal{U}(T)$ on the logical subspace of the Hilbert space, where T is the duration of the evolution[2]. The Hamiltonian $H(t)$ generally comprises the system's control fields as well as other elements that may be attributed to noise. It can be represented within the logical Hilbert space as a linear combination of n-dimensional Hermitian matrices $H(t) = \sum_i c_i(t)\Lambda_i$, where $c_i(t)$ are functions with units of frequency that are associated with the control and noise parameters, and Λ_i are traceless Hermitian operators. It follow from the Hermitian nature of $H(t)$ that $\mathcal{U}(t)$ is a unitary operator. Since we have a finite dimensional system, $\mathcal{U}(t)$ is an element of the Lie group of unitary matrices $U(n)$[3].

As an example, suppose we encode our qubit on a nuclear spin state which is controlled using external magnetic fields. The Hamiltonian for this system is given by

$$H(t) = \gamma\vec{B}\cdot\vec{S} = \gamma\sum_i B_i(t)S_i, \tag{2.9}$$

where γ is the gyromagnetic ratio, $B_i(t)$ are the magnetic field vector components and $S_i = \sigma_i/2$ are the spin angular momentum operators. Suppose that we let the magnetic field be composed of a large bias field together with a small oscillating transverse field component that can be turned on/off: $\vec{B}(t) = B_0\hat{z} + B_1\cos\omega t\hat{x}$.

[2]We may refer to T as the gate time

[3]Since global phase factors are irrelevant in quantum computing, we can restrict our considerations to the subgroup $SU(n)$ instead

Thus, we obtain the following Hamiltonian

$$H(t) = \frac{\gamma B_0}{2}\sigma_Z + \frac{\gamma B_1}{2}\cos\omega t\sigma_X. \tag{2.10}$$

Solving Equation (2.8) with this Hamiltonian would involve a set of coupled differential equations that is not analytically solvable. However, it is possible to gain more insight by moving to a rotating frame. Consider transforming to rotating frame defined by the transformation $V(t) = \exp\left[i\frac{\omega t}{2}\sigma_Z\right]$. In this frame, the Hamiltonian is given by

$$H_V(t) = VHV^\dagger - iV\dot{V}^\dagger \tag{2.11}$$

$$= \begin{pmatrix} \frac{\omega_0-\omega}{2} & \frac{\omega_1}{2}\left(1+e^{2i\omega t}\right) \\ \frac{\omega_1}{2}\left(1+e^{-2i\omega t}\right) & -\frac{\omega_0-\omega}{2}, \end{pmatrix} \tag{2.12}$$

where we introduce new variables $\omega_0 = \gamma B_0$ and $\omega_1 = \gamma B_1/2$[4]. If the transverse field is at resonance with the qubit ($\omega = \omega_0$), we are left only with the off-diagonal elements. Moreover, since $\omega \gg \omega_1$[5], then Hamiltonian terms that are proportional to $e^{i\omega t}$ oscillate rapidly and average to zero in a very short time. In other words, if we take a coarse-grain time average over a time interval much greater than $1/\omega$, the contribution of the rapidly oscillating terms are negligibly small compared with the slowly varying terms. This is referred to as the rotating wave approximation (RWA). Thus, the Hamiltonian after RWA is

$$H_{\text{RWA}} = \frac{\omega_1}{2}\sigma_X. \tag{2.13}$$

For such a simple Hamiltonian, the time evolution can be obtained simply by exponentiating the integral of H_{RWA} with respect to time. Hence, the total time evolution

[4]The factor of $1/2$ is included for notational convenience.

[5]Since $B_0 \gg B_1$, then $\omega = \omega_0 \gg \omega_1$.

in the original frame is given by

$$\mathcal{U}(t) = e^{-i\frac{\omega t}{2}\sigma_Z} e^{-i\frac{\omega_1 t}{2}\sigma_X}. \tag{2.14}$$

If we set the gate time so that $T = \frac{2n\pi}{\omega}$ for some integer n, the first exponential yields $-\mathbb{1}$. In addition, if we define the variable $\theta = \frac{2n\pi\omega_1}{\omega}$, the corresponding quantum gate is

$$\mathcal{U}(T) = e^{-i\frac{\theta}{2}\sigma_X} \equiv X_\theta \tag{2.15}$$

which corresponds to a rotation by an angle θ about the X-axis of the Bloch sphere.

Other gates can also be produced in this system. The simplest one would be to turn off the transverse field $B_1 = 0$ which yields the following gate

$$\mathcal{U}(T) = e^{-i\frac{\omega_0 t}{2}\sigma_Z} \equiv Z_{\theta'}, \tag{2.16}$$

which is a Z-rotation of angle $\theta' = \omega_0 t$. More broadly, rotation about any axis can be achieved by appropriately choosing ω_0, ω_1, and T. Although such flexibility is generally desirable, this level of control may not be possible in other qubit systems nor is it necessary for quantum computing. Indeed, as indicated by the fourth DiVincenzo criteria, it is only necessary to have a universal set of quantum gates. A set of quantum gates is universal if any operation may be approximated to arbitrary accuracy by a quantum circuit composed of its elements. It was shown by Barenco et al. that the set of all 1-qubit gates, together with the controlled-NOT (CNOT) gate, is universal [50]. It is possible to further reduce this set to a finite number of elements using the Solovay-Kitaev theorem [51] which states that any 1-qubit can be approximated to accuracy ε by the composition of $\mathcal{O}\left(\log^c 1/\varepsilon\right)$ gates obtained from the set $\{\mathcal{H}, \mathcal{S}, \mathcal{T}, \textsc{cnot}\}$[6].

In many implementations, qubit control is designed to produce gate operations on at least two perpendicular axes of rotation. These are the X- and Z-axes in the nuclear spin example. One may decompose the action of an arbitrary 1-qubit operation in terms of these gates. This is called the Euler decomposition of a 1-qubit gate. If we only have access to X- and Z-rotations, we may write the gate $\mathcal{U}$ as a sequence of rotations given by

$$\mathcal{U} = Z_{\theta_3} X_{\theta_2} Z_{\theta_1} \tag{2.17}$$

for some angles θ_i. This decomposition provides a method of producing rotations that are not directly accessible to a particular Hamiltonian (e.g., Y-rotations). In fact, the Euler decomposition is a special case of a broader class known as Cartan decomposition [52, 53]. Let $\mathfrak{g}$ be a semi-simple Lie algebra which may be decomposed into two subspace $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}, \mathfrak{m} = \mathfrak{k}^{\perp}$ satisfying the commutation relations

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, \qquad [\mathfrak{m}, \mathfrak{k}] \subseteq \mathfrak{k}, \qquad [\mathfrak{m}, \mathfrak{m}] \subseteq \mathfrak{k}. \tag{2.18}$$

Such a decomposition is called a Cartan decomposition of $\mathfrak{g}$. Note that $\mathfrak{m}$ is not an algebra since it is not closed under the Lie bracket. Therefore, any subalgebra $\mathfrak{a}$ in $\mathfrak{m}$ is necessarily Abelian[7]. A maximal Abelian subalgebra $\mathfrak{a} \subseteq \mathfrak{m}$ is called a Cartan subalgebra. If G is any connected semisimple Lie group with Lie algebra $\mathfrak{g}$ satisfying

[6] $\mathcal{H} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is the Hadamard gate, $\mathcal{S} = Z_{\frac{\pi}{2}}$, and $\mathcal{T} = Z_{\frac{\pi}{4}}$.

[7] Suppose that a subalgebra $\mathfrak{a}$ of $\mathfrak{g}$ is in a subspace of $\mathfrak{m}$. Since $\mathfrak{a}$ is a subalgebra, then it is closed under the Lie bracket $[\mathfrak{a}, \mathfrak{a}]$. However, if $\mathfrak{a} \subseteq \mathfrak{m}$, then $[\mathfrak{a}, \mathfrak{a}] \subseteq [\mathfrak{m}, \mathfrak{m}] \subseteq \mathfrak{k}$. Since $\mathfrak{m}$ and $\mathfrak{k}$ are orthogonal complements of one another, then $[\mathfrak{a}, \mathfrak{a}] = \{0\}$ which means the subalgebra is Abelian.

Equation (2.18), then each element $\mathcal{U} \in G = e^{\mathfrak{g}}$ admits the following decomposition

$$\mathcal{U} = K_2 A K_1, \tag{2.19}$$

where $K_1, K_2 \in e^{\mathfrak{k}}$, and $A \in e^{\mathfrak{a}}$.

To demonstrate how the Euler decomposition is just a special case of Cartan decomposition, let us consider the Lie algebra $\mathfrak{su}(2) = \text{span}\{-i\sigma_X, -i\sigma_Y, -i\sigma_Z\}$. Note that the choice $\mathfrak{k} = \text{span}\{-i\sigma_Z\}$ and $\mathfrak{m} = \text{span}\{-i\sigma_X, -i\sigma_Y\}$ forms a Cartan decomposition of $\mathfrak{su}(2)$. The Cartan subalgebra of $\mathfrak{m}$ is one-dimensional and can be chosen to be $\mathfrak{a} = \text{span}\{-i\sigma_X\}$. According to Equation (2.19), the elements of the group $SU(2)$ can be expressed as $\mathcal{U} = Z_{\theta_3} X_{\theta_2} Z_{\theta_1}$ which is exactly the Euler decomposition given in Equation (2.17). Naturally, different choices of $\mathfrak{k}, \mathfrak{m}$, and $\mathfrak{a}$ would yield equally valid Euler decompositions (possibly with different rotation angles).

Cartan decomposition is also found to be extremely invaluable in the study of two-qubit operations. In general, the associated Lie algebra $\mathfrak{g} = \mathfrak{su}(4)$ can be Cartan decomposed using the choice[8]

$$\begin{aligned}
\mathfrak{k} &= \text{span}\{i\sigma_{XI}, i\sigma_{YI}, i\sigma_{ZI}, i\sigma_{IX}, i\sigma_{IY}, i\sigma_{IZ}\}, \\
\mathfrak{m} &= \text{span}\{i\sigma_{XX}, i\sigma_{XY}, i\sigma_{XZ}, i\sigma_{YX}, i\sigma_{YY}, i\sigma_{YZ}, i\sigma_{ZX}, i\sigma_{ZY}, i\sigma_{ZZ}\}, \\
\mathfrak{a} &= \text{span}\{i\sigma_{XX}, i\sigma_{YY}, i\sigma_{ZZ}\},
\end{aligned} \tag{2.20}$$

where $\sigma_{ij} \equiv \sigma_i \otimes \sigma_j$ is the Kronecker product of the Pauli matrices $\sigma_X, \sigma_Y, \sigma_Z$,

[8]The choice of $\mathfrak{a}$ here is purely conventional. It is usually chosen to fit the interaction Hamiltonian in consideration. For example, we see a case in Chapter 3 where the choice $\mathfrak{a} = \text{span}\{i\sigma_{ZX}, i\sigma_{XY}, i\sigma_{YZ}\}$ is more appropriate.

and $\mathbb{1}_2$ (the 2×2 identity matrix). Note that $\mathfrak{k} = \mathfrak{su}(2) \oplus \mathfrak{su}(2) \subset \mathfrak{su}(4)$. This is the associated Lie algebra of completely local two-qubit operations, i.e., the group of two-qubit operations that can be decomposed as $\mathcal{U}_1 \otimes \mathcal{U}_2 \in SU(2) \otimes SU(2)$, where the subscript indicate which qubit the operator acts on. In other words, local two-qubit gates are simply one-qubit gates acting on two qubits and do not produce nonlocal effects such as entanglement[9]. Nonlocal effects are produced by the generators in the Cartan subalgebra $\mathfrak{a}$ which produce gates of the form $A = \exp\{i\left(\gamma_X \sigma_{XX} + \gamma_Y \sigma_{YY} + \gamma_Z \sigma_{ZZ}\right)\}$.

Recall that the CNOT gate along with the appropriate one-qubit gate set are universal. Strictly speaking, the CNOT gate itself is not necessary. It is only noteworthy because of its entangling property. Specifically, it belongs to a class of special gates called maximally entangling gates which maximizes the entangling power $EP(\mathcal{U})$ given by [54]

$$\begin{aligned} EP(\mathcal{U}) = &-\frac{1}{18}\cos(4\gamma_X)\cos(4\gamma_Y) - \frac{1}{18}\cos(4\gamma_Z)\cos(4\gamma_Y) \\ &- \frac{1}{18}\cos(4\gamma_X)\cos(4\gamma_Z) + \frac{1}{6}. \end{aligned} \tag{2.21}$$

For reference, local two-qubit gates have no entangling power[10] while maximally entangling gates such as the CNOT gate have $EP = 2/9$. Note that the entangling power is strictly determined by the parameters $\{\gamma_X, \gamma_Y, \gamma_Z\}$ of the gate A. Moreover, if $\mathcal{U}$ is maximally entangling then it is possible to produce a CNOT from just two

[9]Nonlocality of a gate does not necessarily imply that it is entangling. The SWAP gate is an example of a two-qubit gate that is highly nonlocal yet nonentangling.

[10]The SWAP gate also has no entangling power even though the A in its Cartan decomposition is nontrivial.

calls to $\mathcal{U}$ via the following decomposition [54]:

$$\textsc{cnot} = K_3\, \mathcal{U}\, K_2\, \mathcal{U}\, K_1, \tag{2.22}$$

where K_i are local two-qubit gates. In fact, any maximally entangling gates can be decomposed in this manner. Thus, a gate set remains universal if we replace CNOT with any maximally entangling gate. This is of practical importance since it is not always possible to achieve a CNOT gate directly from whatever interaction Hamiltonian is attainable in a qubit implementation.

2.3 Summary

In this chapter, we reviewed some important concepts in quantum computing and quantum control. We looked at some basic properties of qubits and emphasized its quantum properties which classical bits lack. We then examined how quantum gates are produced. We considered the simple case of a qubit encoded onto a nuclear spin which is controlled by an external magnetic field. Finally, we discussed the idea of universal gate sets and how techniques such as Cartan decomposition can be used to realize them.

We have ignored the effects of noise to the gates so far. In many instances, coherent systematic noise are the leading cause of errors in one- and two-qubit gates. In the next two chapters, we will discuss how composite pulse sequences can be used to suppress their effects. We measure the quality of our corrected gates by computing their fidelity. The gate fidelity is a measure of “closeness” between the ideal gate and the noisy gate[11]. A simple way of doing so is by taking the Hilbert-Schmidt inner product, $(U_1, U_2) \equiv \operatorname{tr}\left(U_2^\dagger U_1\right) / \operatorname{tr}\left(U_1^\dagger U_1\right)$, which is also referred to as the trace fidelity. Notice that this requires the evolution operator to compute. Although this is readily available in theoretical calculations, we generally do not have access to this information in practice. To give an idea of how fidelity estimation is done in practice, we use in Chapter 3 a technique called Clifford randomized benchmarking (RB) [55]. The main idea behind Clifford RB is to apply an increasingly long series of noisy Clifford gates onto a fiducial state such that the cumulative product of the Clifford gates results in an identity. The fidelity is estimated by computing the probability decay of measuring the same fiducial state. For the remainder of the dissertation, it is sufficient for our purpose to simply use the trace fidelity[12].

[11]There are many ways to define the fidelity $\mathcal{F}$. However, all valid definitions share some properties such as symmetry ($\mathcal{F}(U_1, U_2) = \mathcal{F}(U_2, U_1)$) and boundedness ($0 \leq \mathcal{F} \leq 1$), to name a few.

[12]We may also use the term “infidelity” which is simply $1 - \mathcal{F}$.

Chapter 3

Simulated Randomized Benchmarking of a Dynamically Corrected Cross-Resonance Gate

We theoretically consider a cross-resonance (CR) gate implemented by pulse sequences proposed by Calderon-Vargas & Kestner, *Phys. Rev. Lett. 118, 150502* (2017). These sequences mitigate systematic error to first order, but their effectiveness is limited by one-qubit gate imperfections. Using additional microwave control pulses, it is possible to tune the effective CR Hamiltonian into a regime where these sequences operate optimally. This improves the overall feasibility of these sequences by reducing the one-qubit operations required for error correction. We illustrate this by simulating randomized benchmarking for a system of weakly coupled transmons and show that while this novel pulse sequence does not offer an advantage with the current state of the art in transmons, it does improve the scaling of CR gate infidelity with one-qubit gate infidelity. This chapter is based on the paper Phys. Rev. A **102**, 032626 [56].

3.1 Introduction

The ability to implement high-fidelity gates is a necessary requirement for creating a fully functional quantum information processor. To this end, fixed-frequency superconducting transmons [57] show great promise [58, 59], as they have been used to theoretically and experimentally demonstrate one-qubit gates [60–64] with fidelities as high as 99.97% [64]. However, generating two-qubit entangling operations with similarly high fidelities remains a challenge. A standard approach to entangling fixed-frequency transmons is through the cross-resonance (CR) effect [65–68]. The CR effect can be observed in a system of two off-resonant fixed-frequency transmons with a small static coupling (e.g., through a quantum bus [69]). By irradiating one transmon at the transition frequency of the other, the coupling is modified by a factor whose magnitude is roughly proportional to the ratio of the microwave drive amplitude and the interqubit detuning.

Theoretical considerations have shown that the CR gate is significantly affected by systematic errors attributed to high-energy excitations of the weakly anharmonic transmon and to crosstalk induced by the CR microwave drive [70, 71]. These processes give rise to unwanted terms in the CR effective Hamiltonian. This necessitates the use of control techniques such as composite pulse sequences [72, 73] in order to isolate the desired entangling dynamics. In the case of a CR gate, such gate errors can be eliminated by a secondary control pulse on the target qubit which, in conjunction with pulse sequences, can result in CR gate fidelities exceeding 99% [74]. However, the pulse sequence used in Ref. [74] is not capable of addressing all coherent systematic errors to leading order.

In this chapter, we analyze how well a different, recently discovered generic composite pulse sequence [75] would perform in the specific application of fixed-frequency transmons coupled via the CR effect as opposed to the conventional approach. This new sequence inserts local π rotations between repeated application of an entangling gate to dynamically correct all coherent systematic errors in that entangling gate, but in practice there is a tradeoff between that reduction of error and the introduction of errors coming from the insertion of imperfect local π pulses. The purpose of this paper is to examine this tradeoff for the case of CR-gated transmons and determine the conditions for which there is a net benefit.

We theoretically simulate standard Clifford randomized benchmarking (RB) to assess the CR gate performance and show that, while there is no benefit to using the sequence of Ref. [75] with current transmon noise levels and single-qubit fidelities, as single-qubit fidelities improve the new pulse sequence could provide better two-qubit RB fidelities than the currently used dynamical correction scheme.

3.2 Dynamical Error Correction via Pulse Sequences

We begin by summarizing the formalism developed in Ref. [75]. We are interested in developing a protocol that allows us to dynamically correct coherent systematic error affecting an arbitrary two-qubit entangling gate. To this end, Ref. [75] presented a family of composite pulse sequences that are composed using repetitions of the nonlocal gate $(\theta)_{ab} = \exp\left[-i\left(\theta/2\right)\sigma_{ab}\right]$, where $a, b \in \{X, Y, Z\}$, which can be generated from any arbitrary two-qubit coupling along with appropriate one-qubit rotations [76, 77]. In practice, the building block $(\theta)_{ab}$ may contain errors, which we only consider up to the leading order. Thus, we have

$$(\theta)_{ab} = \exp\left[-i\frac{\theta}{2}\sigma_{ab}\right]\left(I + i\sum_{i,j\in\{\mathrm{I,X,Y,Z}\}}\epsilon_{ij}\sigma_{ij}\right), \tag{3.1}$$

where ϵ_{ij} is constant in time and is hereafter referred to as the error in the ij error channel. The pulse sequences have the general form

$$\begin{aligned}
&\sigma_{\text{echo}}^{(n)}(\theta)_{ab}\,\sigma_{\text{echo}}^{(n)}\sigma_{\text{echo}}^{(n-1)}(\theta)_{ab}\,\sigma_{\text{echo}}^{(n-1)}\ldots\sigma_{\text{echo}}^{(1)}(\theta)_{ab}\,\sigma_{\text{echo}}^{(1)}\\
&= \exp\left[-i\frac{\theta}{2}\sum_{l=1}^{n}\xi_l\sigma_{ab}\right]\\
&\times\left\{I + i\sum_{i,j}\epsilon_{ij}\sigma_{ij}\sum_{m=1}^{n}\zeta_m^{ij}\exp\left[i\frac{\theta}{2}(\chi_{ij}-1)\sum_{l=1}^{m-1}\xi_l\sigma_{ab}\right]\right\},
\end{aligned} \tag{3.2}$$

where $\sigma_{\text{echo}}^{(l)}$ denotes a local π rotation of the form $\sigma_{cd} \equiv \sigma_c \otimes \sigma_d$ with $c, d \in \{I, X, Y, Z\}$ hereafter referred to as an echo pulse, and

$$\xi_l \equiv \begin{cases} +1, & \text{if}\left[\sigma_{\text{echo}}^{(l)}, \sigma_{ab}\right] = 0,\\ -1, & \text{if}\left\{\sigma_{\text{echo}}^{(l)}, \sigma_{ab}\right\} = 0,\end{cases} \tag{3.3}$$

$$\zeta_m^{ij} \equiv \begin{cases} +1, & \text{if}\left[\sigma_{\text{echo}}^{(l)}, \sigma_{ij}\right] = 0,\\ -1, & \text{if}\left\{\sigma_{\text{echo}}^{(l)}, \sigma_{ij}\right\} = 0,\end{cases} \tag{3.4}$$

$$\chi_{ij} \equiv \begin{cases} +1, & \text{if}\,[\sigma_{ij}, \sigma_{ab}] = 0,\\ -1, & \text{if}\,\{\sigma_{ij}, \sigma_{ab}\} = 0.\end{cases} \tag{3.5}$$

We refer to a sequence containing n applications of the noisy entangling operation as a "length-n" sequence. To eliminate the effects of the ij error channel to leading order, we require

$$\sum_{m=1}^{n}\zeta_m^{ij}\exp\left[i\frac{\theta}{2}(\chi_{ij}-1)\sum_{l=1}^{m-1}\xi_l\sigma_{ab}\right] = 0. \tag{3.6}$$

To simplify this robustness condition, Ref. [75] considered two cases: one where only commuting errors are present ($\chi_{ij} = 1$) and one where only anticommuting errors are present ($\chi_{ij} = -1$).

Let us first consider the case where we only have errors that commute with the entangling operation $(\theta)_{ab}$. In this case, Equation (3.6) reduces to

$$\sum_{m=1}^{n} \zeta_m^{ij} = 0. \tag{3.7}$$

This immediately suggests that the robustness constraint is satisfied only for even values of n. The robustness condition in Equation (3.7) for a length-2 sequence requires $\zeta_1^{ij} = -\zeta_2^{ij}$. Setting $\zeta_1^{ij} = 1$ implies that the first echo pulse commutes with all the errors. Without loss of generality, we can choose the first pulse to be the identity operator for simplicity. Note that, in order to have a non-identity operation, the second pulse must commute with σ_{ab}, i.e., $\xi_2 = 1$. The second pulse must also anticommute with all the errors in order to satisfy the robustness condition. If every potential commuting errors are present, this is not possible since there is no choice of $\sigma_{\text{echo}}^{(2)}$ that will simultaneously anticommute with all commuting errors, $[\sigma_{\text{echo}}^{(2)}, \sigma_{ij}] = 0 \,\forall\, ij \ni [\sigma_{ij}, \sigma_{ab}] = 0$. A length-2 sequence can cancel four of the seven commuting error terms while producing an entangling operation, which may be all that is necessary in certain situations, but no more. (This can be quickly verified for any specific choice of σ_{ab} by simply listing all possibilities, but see Appendix A for the general proof.)

Nonetheless, with the exception of error in the *ab* channel itself, all errors that commute with σ_{ab} can be eliminated to first order by using two nested applications of a length-2 sequence, i.e., a length-4 sequence. For instance, the length-4 sequence

$$\begin{aligned}\mathcal{U}^{(4)}\left[(\theta)_{ab}\right] &\equiv (\theta)_{ab}\,\sigma_{aI}\,(\theta)_{ab}\,\sigma_{aI}\sigma_{cc}\,(\theta)_{ab}\,\sigma_{aI}\,(\theta)_{ab}\,\sigma_{aI}\sigma_{cc} \\ &= (\theta)_{ab}\,\sigma_{aI}\,(\theta)_{ab}\,\sigma_{bc}\,(\theta)_{ab}\,\sigma_{aI}\,(\theta)_{ab}\,\sigma_{bc} \\ &= \exp\left[-i\frac{4\theta}{2}\sigma_{ab}\right]\left(I+\mathcal{O}\left(\epsilon^2\right)\right), \end{aligned} \tag{3.8}$$

where $\{\sigma_{cc}, \sigma_{ab}\} = 0$, eliminates all commuting error channels to first order except for the ab channel itself.

We now consider the complementary case where all the errors instead anticommute with the entangling operation $(\theta)_{ab}$. The robustness constraint in Equation (3.6) becomes

$$\sum_{m=1}^{n} \zeta_m^{ij} \exp\left[-i\theta \sum_{l=1}^{m-1} \xi_l \sigma_{ab}\right] = 0. \tag{3.9}$$

Ref. [75] showed that a nontrivial solution can be found when $n = 5$, $\xi_l = 1$, $\zeta_{(1,2,4,5)} = \pm 1$, $\zeta_3 = \mp 1$, and $\theta = \theta_0 \equiv \arccos\left[(\sqrt{13}-1)/4\right] \approx 0.27\pi$. A set of echo pulses that correspond to these values are $\sigma_{\text{echo}}^{(1,2,4,5)} = I$ and $\sigma_{\text{echo}}^{(3)} = \sigma_{ab}$. Thus, a length-5 sequence that corrects all anticommuting errors to leading order is given by

$$\begin{aligned}\mathcal{U}^{(5)}\left[(\theta_0)_{ab}\right] &\equiv (\theta_0)_{ab}\,(\theta_0)_{ab}\,\sigma_{ab}\,(\theta_0)_{ab}\,\sigma_{ab}\,(\theta_0)_{ab}\,(\theta_0)_{ab} \\ &= \exp\left[-i\frac{5\theta_0}{2}\sigma_{ab}\right]\left[I+\mathcal{O}\left(\epsilon_{anticomm}^2\right)\right]. \end{aligned} \tag{3.10}$$

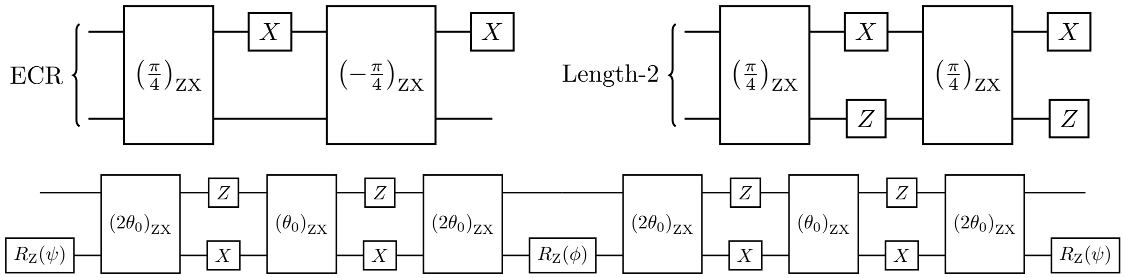
The resulting gate in Equation (3.10) is nearly maximally entangling, but it is not locally equivalent to a CNOT. We can, however, construct a gate locally equivalent to a CNOT that can serve as a two-qubit Clifford group generator by using two applications of the dynamically corrected gate:

$$\mathcal{U}_{\mathrm{Clif}_2} = \exp\left[-i\frac{\psi}{2}\sigma^{'}\right]\mathcal{U}^{(5)}\exp\left[-i\frac{\phi}{2}\sigma^{'}\right]\mathcal{U}^{(5)}\exp\left[-i\frac{\psi}{2}\sigma^{'}\right], \tag{3.11}$$

where $\psi = 2\arctan[(\sqrt{-57+16\sqrt{13}})/(4-\sqrt{13}+2\sqrt{-7+2\sqrt{13}})] \approx 0.36\pi$, $\phi = -2\arccos[-1/(2\sqrt{-14+4\sqrt{13}})] \approx -1.56\pi$, and $\sigma^{'} \in \{\sigma_{IX}, \sigma_{IY}, \sigma_{IZ}, \sigma_{XI}, \sigma_{YI}, \sigma_{ZI}\}$ such that $\{\sigma^{'}, \sigma_{ab}\} = 0$.

It is possible to combine a length-2 (or length-4) sequence with a length-5 sequence in order to generate a length-10 (or length-20) sequence that can address both commuting and anticommuting error channels simultaneously. Furthermore, all of these pulses can also be combined with a BB1-like pulse sequence in order to correct the ab channel errors. First-order error in this channel can manifest from gate mistiming or fluctuations in the effective interqubit coupling, both of which result in over/under-rotation of the entangling operation. We refer the reader to Ref. [75] for a more detailed discussion.

Finally, we wish to emphasize that although the rest of this manuscript focuses on the application of the length-5 pulse sequence to fixed frequency transmon qubits, similar considerations apply in any other scenario having the key feature that the errors in the entangling gate anticommute with the entangling operator. For example, in a silicon-based system of two double quantum dots (DQDs), each containing a single spin, coupled through a resonator [78]. The resonator is coupled to only one of the quantum dots, which makes the effective coupling dependent on the magnetic gradient within the DQD. Imperfections on the magnetic field gradient, which can be caused by either an anisotropy in the electron g-tensor or misalignment



Length-5

Figure 3.1: Circuit diagrams for the two-qubit Clifford generator implemented using the ECR, the length-2 sequence, and the length-5 sequence (see Equation (3.11)). The top (bottom) line corresponds to the control (target) qubit. The X and Z gates are the usual Pauli gates and $R_{\mathrm{Z}}(\theta)$ is a rotation about the Z-axis by an angle θ. The generated Clifford gate is locally equivalent to a CNOT gate in all cases.

of the local micromagnet, causes systematic commuting and anticommuting errors to emerge. These can be addressed by a length-5 sequence or a combination of a length-2 and a length-5 sequence depending on the severity of the error. However, from this point on we use numbers appropriate for the CR gated transmon case.

3.3 Dynamically Corrected CR Gate

We now apply the formalism we summarized in Section 3.2 to a CR gate. We consider a system of two fixed off-resonant transmons that are weakly coupled to a bus resonator. We then apply a constant-amplitude microwave driving field on one qubit, the control qubit, at the transition frequency of the other qubit, the target qubit. In the weak driving limit, a block-diagonal effective Hamiltonian for a CR gate can be perturbatively constructed using the Schrieffer-Wolff transformation [70]:

$$H_{\text{eff}}^{\text{CR}} = \frac{1}{2}h_{ZI}\sigma_{ZI} + \sum_{j\in\{X,Y,Z\}} \left(\frac{1}{2}h_{Ij}\sigma_{Ij} + \frac{1}{2}h_{Zj}\sigma_{Zj}\right), \tag{3.12}$$

where the expressions for h_{ij} in terms of the physical parameters are given in the appendix of Ref. [70]. This approach differs from previous derivation of the CR Hamiltonian [65] in that it yields coherent error terms pertaining to higher-energy level leakage. We note that the Hamiltonian belongs in the embedding $\mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{u}(1) \subset \mathfrak{su}(4)$. In particular, $\mathfrak{u}(1)$ is generated by σ_{ZI} which yields a factor in the time-evolution operator that can be removed by applying a local Z-rotation on the first qubit. For this reason, we will ignore the effects of the σ_{ZI} term. The entangling term here is the σ_{ZX} term which, if factored out, yields

$$U(t) = \exp\left[-\frac{i}{2}th_{ZX}\sigma_{ZX}\right]\left[I + i\sum_{j\in\{X,Y,Z\}}(\epsilon_{Ij}\sigma_{Ij} + \epsilon_{Zj}\sigma_{Zj})\right], \tag{3.13}$$

where ϵ_{ij} can be calculated analytically up to a desired order using the Baker-Campbell-Hausdorff (BCH) formula. The pulse sequence building block, $(\theta)_{\text{ZX}}$, can then be obtained by setting $t = \theta/h_{\text{ZX}}$.

In experiments, the microwave drive acting on the control qubit often leaks into the target qubit which results in on-resonant crosstalk. This introduces large IX and IY terms in the effective Hamiltonian. Thus, in practice, the commuting errors are in the ZX and IX channels, while the anticommuting ones are in the IY, IZ, ZY, and ZZ channels [74]. Neglecting the ZX error channel for the moment, which is reasonable provided that the errors are static and the evolution time of the entangling gate is properly compensated through calibration, the result in Section 3.2 suggests that we apply a length-2 sequence with a σ_{XZ} echo pulse to eliminate the IX channel. In addition, this choice of echo pulse can also eliminate any higher-order ZI channel errors that may accumulate due to the presence of large anticommuting error terms. Although alternatives such as σ_{XY} can serve the same purpose, we choose σ_{XZ} in order to take advantage of novel control methods that allow implementation of near-perfect virtual Z-gates via abrupt phase modulation of the microwave control drive [64].

We note that Refs. [74] and [73] use an operationally distinct pulse sequence called an echoed CR (ECR) gate which has the same effect as the above length-2 sequence with an XZ echo pulse. The key operational difference between the ECR scheme and our length-2 sequence is the sign reversal in the entangling operation,

$$ECR \equiv \left(\frac{\pi}{4}\right)_{ZX} \sigma_{XI} \left(-\frac{\pi}{4}\right)_{ZX} \sigma_{XI}, \tag{3.14}$$

which can be implemented experimentally by reversing the signal of the microwave drive ($\Omega \rightarrow -\Omega$). Unlike in our length-2 scheme, a σ_{XI} echo pulse, which anticommutes with σ_{ZX}, is applied to avoid implementing a purely local gate. We show in Appendix B that this yields a mathematically equivalent pulse as the length-2 sequence in the case of a CR Hamiltonian. So in the remainder of our discussions, the length-2 sequence and the ECR gate are equivalent.

Another approach involves applying a secondary microwave pulse onto the target qubit so as to eliminate particular terms in the Hamiltonian [74]. This cancellation pulse is calibrated such that it eliminates the h_{IX}, h_{IY}, and h_{ZY} terms. Using the experimental parameters provided in Ref. [74], it can be verified numerically that the remaining terms have different scales, $h_{IZ} \ll h_{ZZ} \ll h_{ZX}$. Thus, the dominant source of remaining error comes from the h_{ZZ} term of the effective Hamiltonian, which translates to errors in the ZZ and IY channels. Although a length-2 sequence with an XZ echo pulse (i.e., ECR) can partially suppress these anticommuting errors, one can instead get complete first-order correction using the length-5 sequence of Equation (3.10), $\mathcal{U}^{(5)}[U(\theta_0/h_{ZX})]$, with $ab = ZX$. Then, to obtain a two-qubit Clifford generator, we use Equation (3.11) with $\sigma' = \sigma_{IZ}$, where we again make use of virtual Z gates. This yields an entangling Clifford gate compensated for all relevant coherent systematic errors to leading order.

It is important to keep in mind that the theory we summarized in Section 3.2 assumes that the echo pulses can be implemented perfectly. This is not the case in practice and echo pulse errors can be detrimental to the sequence's efficacy. Even though a longer and theoretically better sequence can be obtained by combining the length-2 and length-5 sequences, the resulting length-10 sequence requires more potentially noisy one-qubit gates to implement. So, depending on the level of one-qubit error, a length-2 or length-5 sequence can be more effective than a length-10 sequence. The supplemental material of Ref. [75] indicates that one-qubit gate errors on the order of, at most, 10^{-5} are required in order to build a CNOT out of a length-20 sequence with gate error below 10^{-3}. This may be very difficult to realize in the near future, which is why we focus our discussion to the length-2 and the length-5 sequence. In principle, a $\left(\frac{\pi}{2}\right)_{\text{ZX}}$ gate generated using the length-2 sequence requires 4 one-qubit gates, while $\mathcal{U}_{\text{Cliff}_2}$ requires 14 one-qubit gates. However, by taking advantage of virtual Z-gates, we can reduce this to 2 and 4 physical one-qubit gates, respectively.

Finally, we note that the two-qubit Clifford gate generated by the length-2 sequence, $\left(\frac{\pi}{2}\right)_{\mathrm{ZX}}$, and by the length-5 sequence, U_{Clif_2}, are different up to local Clifford rotations. In both cases, though, the local invariants of the resulting composite gate are equal to that of a CNOT gate. We present in Fig. 3.1 a circuit diagram for each of the cases we discussed.

3.4 Simulated Randomized Benchmarking

To assess the performance of our dynamically corrected gate, we simulate standard Clifford randomized benchmarking (RB) [55] using

$$\left\{I, X_{\pm\frac{\pi}{2}}, X_{\pm\pi}, Z_{\pm\frac{\pi}{2}}, Z_{\pm\pi}, \mathcal{U}_{\mathrm{Clif}_2}\right\}$$

as our generating set where, as an example, X_π denotes a π rotation about the X-axis. We include quasistatic error in all local X-rotations using the following noise model:

$$X_\theta \to \exp\left[-i\frac{\varepsilon}{2}\frac{r_x\sigma_X + r_y\sigma_Y + r_z\sigma_Z}{\sqrt{r_x^2 + r_y^2 + r_z^2}}\right]X_\theta, \tag{3.15}$$

where $\{r_x, r_y, r_z\}$ are sampled uniformly from $[-1, 1]$, and ε is sampled from a normal distribution centered at 0 with standard deviation $\delta\theta$. We present in Appendix C an analytical formula that relates the one-qubit RB infidelity to $\delta\theta$. On the other hand, Z-rotations are performed with no error, corresponding to the virtual gate method described in Ref. [64]. We also calculate the trace infidelity, which is not efficiently accessible in experiment, but is a less computationally demanding measure for theory, especially in the limit of very weak noise.

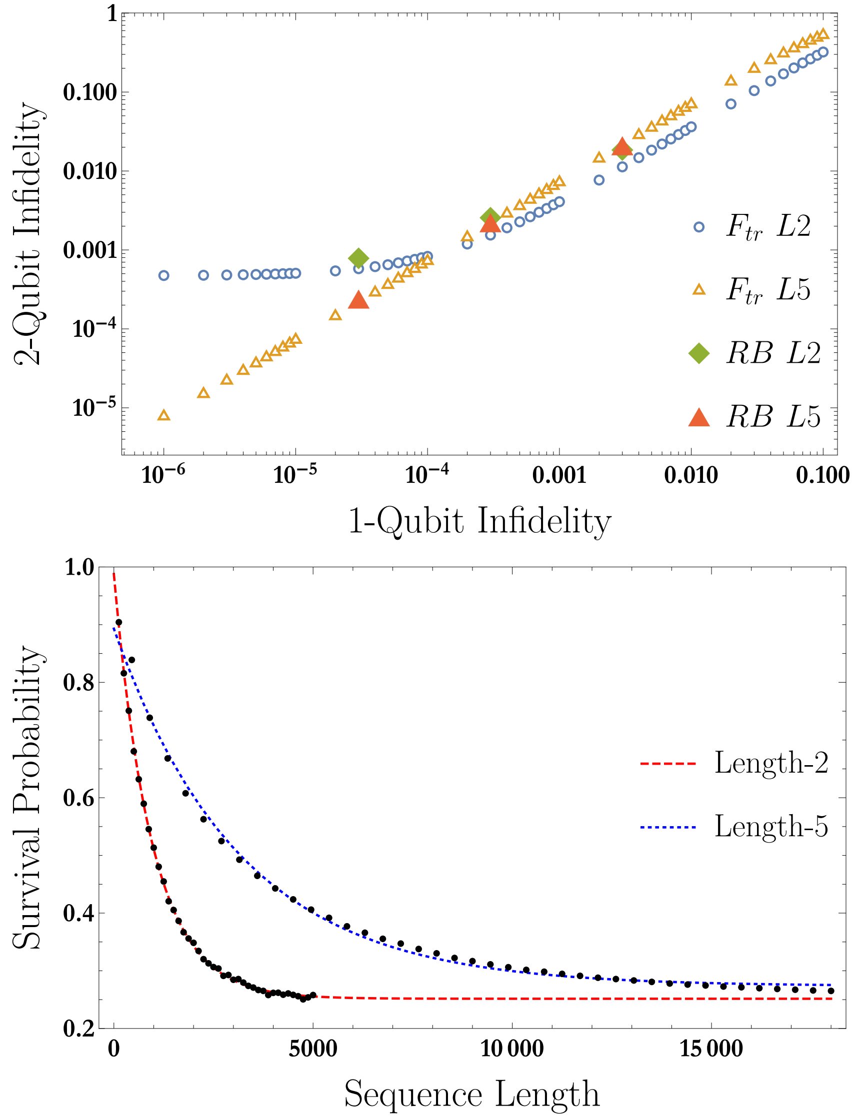


Figure 3.2: (TOP) Solid symbols show randomized benchmarking (RB) two-qubit Clifford gate infidelity as a function of one-qubit RB infidelity for the length-2 and length-5 sequences. A current experimentally attainable value of one-qubit infidelity is 3×10^{-4} [64], corresponding to the RB points in the center of the plot. Open symbols show the trace fidelity. (BOTTOM) A standard RB decay plot comparing the length-2 and length-5 sequences for the case where one-qubit infidelity is set to 3×10^{-5}.

We use the experimental parameters reported in Ref. [74]: $\omega_1/2\pi = 5.114$ GHz, $\omega_2/2\pi = 4.914$ GHz, $\delta_1/2\pi = \delta_2/2\pi = -0.330$ GHz, $\Omega/2\pi = 60$ MHz, and $J/2\pi = 3.8$ MHz. The evolution time for the building block of the length-2 sequence, $\left(\frac{\pi}{4}\right)_{\mathrm{ZX}}$, is $t = \pi/(4h_{ZX}) = 49.2$ns, while that of the length-5 sequence, $(\theta_0)_{\mathrm{ZX}}$, is $t = \theta_0/h_{ZX} = 54$ns. In order to simulate the effect of the cancellation pulse, we only include h_{IZ}, h_{ZX}, and h_{ZZ} in our effective Hamiltonian. Furthermore, we ignore relaxation errors in our simulations and focus solely on coherent systematic error. Each point in the decay curve of our RB simulations are averaged over 1000 different sequences and noise realizations. The sequence length is set just enough to find a good fit for the survival probability function $ap^k + b$, where a, b, and p are fitting parameters and k is the sequence length. The results of our simulations are presented in Fig. 3.2.

Note that we do not include the initial portion of the decay from 100% down to around 90% (not plotted) in the fit, since there is non-exponential behavior there, particularly in the case of the length-5 decay. Non-exponential decay is commonly attributed to gate-dependent errors or low-frequency time-dependent noise [79], both of which are present in our simulations. Gate-dependent errors are present because we simulated RB with perfect Z-gates but noisy X-gates, as in experiments. Low-frequency noise effects appear because we perform each individual RB run with a fixed set of randomly generated noisy one-qubit Clifford group, again corresponding to the likely experimental case. We keep generating new sets of noisy Clifford gates until we exhaust all of our RB sequences. This builds statistics consistent with the distribution from which the noisy one-qubit gates are generated. A non-exponential decay is obtained by averaging over this ensemble of RB data.

We find that the length-2 sequence performs similarly to the length-5 sequence when the one-qubit RB error is set to 3×10^{-4} to match Ref. [64]. The length-2 sequence yields a fidelity of 99.7% and the length-5 sequence, which takes about five times as long (540ns $+4t_{1\mathrm{Q}}$ vs 98ns $+2t_{1\mathrm{Q}}$, where $t_{1\mathrm{Q}}$ denotes the echo pulse gate time), yields 99.8% [1]. However, if the one-qubit errors were reduced, we see that the length-5 sequence increasingly outperforms the length-2.

We can gain further insight by comparing the performance of the two pulse sequences in the limit where there are no one-qubit errors. For this task we use the trace fidelity since simulated randomized benchmarking requires simulating increasingly long sequences to obtain enough fidelity decay to fit as the one-qubit gate error is reduced. We rearrange Equation (3.1) and isolate the error terms:

$$\begin{aligned}\delta U &= \exp\left[\frac{i}{2} t h_{ZX} \sigma_{ZX}\right] U(t) - I \\ &= i \sum_{j \in \{I,X,Y,Z\}} \epsilon_{Ij}\sigma_{Ij} + \epsilon_{Zj}\sigma_{Zj}.\end{aligned}$$

We numerically calculate this for both schemes, assuming perfect one-qubit gates,

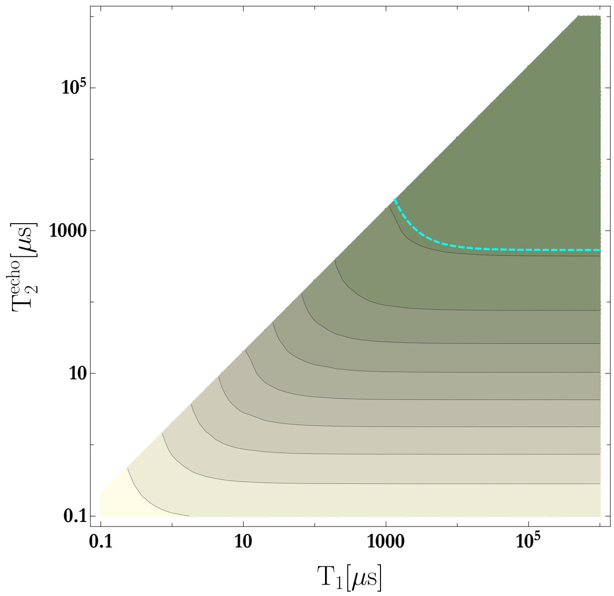
[1]The reason our simulated length-2 sequence fidelity is slightly higher than the experimental one reported in Ref. [74] is simply because we used the more recent lower one-qubit noise value. If we use the conditions of Ref. [74], our trace fidelity calculation yields an error of roughly 6×10^{-4} which is consistent with the experimentally observed values.

and get

$$\delta U_{L2} = -2.4 \times 10^{-4} I + .015i(\sigma_{IY} - \sigma_{ZZ})$$
$$+ 7.5i \times 10^{-4}(\sigma_{IZ} + \sigma_{ZY}) + 3.5i \times 10^{-4}\sigma_{ZX}$$
$$\delta U_{L5} = -2i \times 10^{-5}\sigma_{IX} - 4.8i \times 10^{-4}\sigma_{ZX},$$

where we have omitted any error terms with magnitudes below 10^{-5}. Since the gate infidelity is proportional to ϵ_{ij}^2, we see that the length-5 scheme can reach error rates on the order of 10^{-7} at best, while the length-2 scheme can only reach 10^{-4}. The much lower ideal infidelity of the length-5 sequence is because it cancels all the leading order errors in $U(t)$, whereas the length-2 sequence is not capable of eliminating the anticommuting error channels IY and ZZ. Of course, the actual performance of both sequences is highly dependent on the severity of the one-qubit gate imperfections, as is evident in Fig. 3.2, but one can see there that the trace infidelity of the length-5 sequence keeps decreasing with decreasing single-qubit error while the length-2 sequence plateaus in the 10^{-4} region. Moreover, the crossing point where the length-5 is predicted to outperform the length-2 sequence occurs when the one-qubit infidelity is roughly 1×10^{-4}. This indicates that the length-5 sequence may be experimentally viable in the near future if one-qubit gate fidelities can be brought above 99.99%.

One caveat to this conclusion is that, as previously noted, the two-qubit Clifford generated from length-5 sequences is about five times slower than its length-2 counterpart. Thus, the length-5 sequence will suffer more from T_1 relaxation error, and its contribution to gate infidelity goes roughly as T_{gate}/T_1 [80]. One could con-



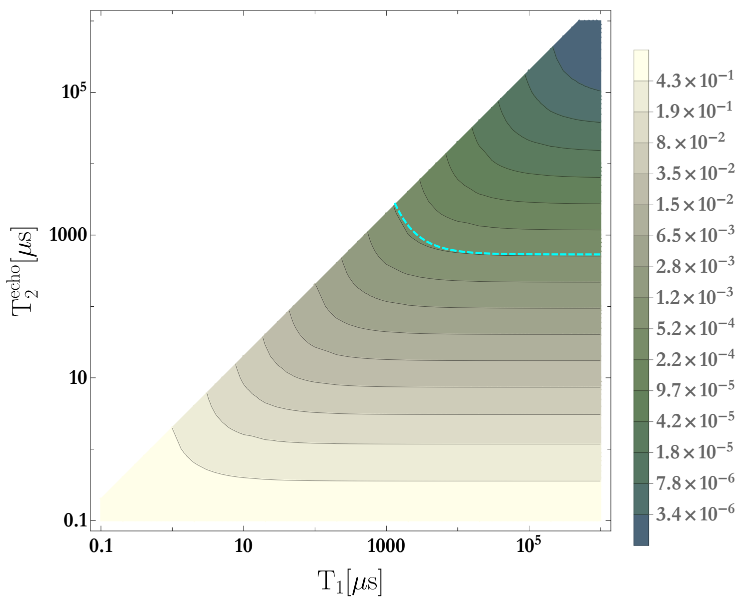


Figure 3.3: A contour plot of two-qubit infidelity as a function of relaxation time, T_1, and echoed dephasing time, T_2^{CPMG}, for the length-2 sequence (TOP) and the length-5 sequence (BOTTOM), assuming one-qubit gates with no coherent or leakage errors and an average gate time of 30ns. Unphysical regions where $T_2^{\mathrm{CPMG}} > 2T_1$ are excluded. The dashed cyan contour indicates the crossing point where the two sequences have equal infidelities. The same colorscale is used for both panels.

sider increasing the CR drive amplitude Ω to speed up the gate. Numerical analysis of the CR gate in the strong driving regime indicates that the Hamiltonian terms that we considered as systematic error cannot be treated perturbatively when using a naïve cosine ramp model for the drive [71]. These terms can potentially be minimized while reducing the CR gate time by using pulse shapes derived from optimal control schemes which can result in CR gates under 100ns [81]. Alternatively, one could also consider increasing the coupling between the qubits to speed up the gate, since the corresponding increase in σ_{ZZ} crosstalk due to unwanted excitations to higher energy transmon states would anyways be canceled by the length-5 sequence, but the issue is that the diminished qubit addressability would likely lower *one-qubit* echo pulse fidelities. However, at least the task of engineering a high-fidelity two-qubit gate is then effectively reduced to the problem of engineering high-fidelity local gates.

Without those sort of changes to speed up transmon operations, one has to consider in more detail the trade-off between reduction of coherent error by the length-5 sequence and increased incoherent error due to the longer gate time of the sequence. We aim to quantify this now by analyzing the effects of decoherence. For simplicity, we only consider dephasing and relaxation from the first transmon excited state to the ground state. Using the same parameters as above and setting the ground state energy to zero, we simulate the evolution by a Lindblad master equation

$$\dot{\rho} = -i\left[H_{\text{eff}}^{\text{CR}}, \rho\right] + \frac{1}{T_1}\sum_{j=1,2}\mathcal{D}\left[\sigma_j^-\right]\rho + \frac{1}{T_2^{\text{CPMG}}}\mathcal{D}\left[\Pi_j^1\right]\rho, \tag{3.16}$$

where ρ is the density matrix, T_1 is the relaxation time of the two qubits, T_2^{CPMG} is the dephasing time measured via Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence (used here as a lower bound on T_2), σ_j^- (Π_j^1) is the j^{th} qubit's lowering operator (projection operator to the $|1\rangle$ state), and $\mathcal{D}$ is the damping superoperator

$$\mathcal{D}\left[A\right]\rho = A\rho A^\dagger - \frac{1}{2}A^\dagger A\rho - \frac{1}{2}\rho A^\dagger A. \tag{3.17}$$

In order to focus on the role of decoherence, we assume that each one-qubit gate in the sequence is implemented without coherent or leakage errors and with an average gate time of 30 ns [74] during which the transmons can relax. The average two-qubit infidelity can then be calculated [82]:

$$\langle F\rangle = \frac{1}{16}\left[4 + \frac{1}{5}\sum_{i,j=\mathrm{I,X,Y,Z}} \mathrm{tr}\left[U\sigma_{ij}U^\dagger \mathcal{M}\left(\sigma_{ij}\right)\right]\right], \tag{3.18}$$

where U is the ideal unitary time-evolution operator, $\mathcal{M}$ is a trace-preserving linear map, and $\sigma_{ij} = \sigma_i \otimes \sigma_j$ are the 15 non-identity Kronecker products of Pauli matrices. We plot the results in Fig. 3.3.

The current state-of-the-art transmons can achieve average coherence times of $T_1 = 0.23$ms and $T_2^{\mathrm{CPMG}} = 0.38$ms [83]. For those values, as opposed to the case of purely coherent error considered in Fig. 3.2, Fig. 3.3 indicates that even in the absence of one-qubit coherent gate error, the length-5 sequence does not outperform the length-2 sequence due to the effects of incoherent error over the longer gate time. However, at increased coherence times of $T_1, T_2^{\mathrm{CPMG}} \approx 1.6$ms, the fidelity of the length-5 sequence begins to surpass that of the length-2 sequence. For $T_1, T_2^{\mathrm{CPMG}} \gg 1$ms the performance of the length-2 sequence plateaus at 3.8×10^{-4}, while the length-5 sequence continues to show improvement until it also eventually plateaus at roughly 3×10^{-7}, consistent with what we observed in Fig. 3.2(a).

Thus, while the length-5 sequence is not currently practical, given the rate of improvement in coherence times in recent years (roughly an order of magnitude every three years) [84] and the amount of attention being devoted to this task [85], it is reasonable to expect the length-5 sequence to become a viable option in the near future.

3.5 Summary

We have shown how to dynamically correct a CR gate using a recently developed composite pulse sequence and we theoretically simulated a randomized benchmarking protocol for an experimentally accessible comparison of its performance with the standard ECR scheme, which is equivalent to a length-2 pulse sequence. The application of a cancellation pulse onto the target qubit eliminates a significant amount of coherent systematic error from the effective CR Hamiltonian. The length-2 sequence cannot address all of the remaining dominant errors, all of which anticommute with the entangling operation, but the newly developed length-5 sequence can, at the cost of additional local rotations and a slower entangling gate. We find that both sequences perform similarly against coherent error when using one-qubit gates with currently achievable fidelities. However, we also show that the length-5 sequence performance could scale much better than the length-2 sequence when one-qubit gates are improved.

The pulse sequences we presented can be easily extended to systems with more than two fixed transmon qubits. Ideally, any given pair of control and target qubit must be decoupled from the remaining idle qubits when generating a two-qubit operation. In cases where more than two qubits share the same bus, the static always-on coupling can lead to spurious Z interactions with one or more of the idle qubits. One work-around to this problem is by performing the CR operation on the control and target qubit while decoupling the idle qubits through Hahn-echo-like pulses [86, 87]. We can apply the same idea to the length-2 and length-5 sequence in order to simultaneously address entangling gate errors within the control-target subspace and spurious errors with the idle qubits. However, the additional echo pulses required to implement this makes the sequence even longer than it already is.

The long gate time of the length-5 sequence already makes it impractical for current coherence times, as the improvement the sequence is designed to produce against coherent errors is outweighed by the increased susceptibility to incoherent errors. However, once coherence times are increased beyond 1ms, the sequence we have presented in this paper will become useful for increasing overall two-qubit gate fidelity.

Chapter 4

Stroboscopically Robust Gates For Capacitively Coupled Singlet-Triplet Qubits

Recent work on Ising-coupled double-quantum-dot spin qubits in GaAs with voltage-controlled exchange interaction has shown improved two-qubit gate fidelities from the application of oscillating exchange along with a strong magnetic field gradient between adjacent dots [88]. By examining how noise propagates in the time-evolution operator of the system, we find an optimal set of parameters that provide passive stroboscopic circumvention of errors in two-qubit gates to first order. We predict over 99% two-qubit gate fidelities in the presence of quasistatic and $1/f$ noise, which is an order of magnitude improvement over the typical unoptimized implementation. This work was based on the paper Phys. Rev. A **99**, 012347 [89].

4.1 Introduction

Quantum dot spin qubits provide a promising platform for quantum computing due to their potential scalability and relatively long coherence times. For single-spin qubits [90], one-qubit operations with gate fidelities exceeding the fault-tolerant threshold have been realized in single-spin qubits [91], but two-qubit gates have much lower fidelities [92, 93]. Likewise, for singlet-triplet spin qubits [94, 95], which we focus on below, a recent two-qubit experiment reported only up to 90% entangling gate fidelity [88]. This can be improved by circumventing the effects of the two main noise sources, namely fluctuations in the electric confining potential and fluctuations in the Zeeman energy difference between the quantum dots.

The fluctuation in the confining potential is often attributed to thermal fluctuations in the occupation of nearby charge traps, i.e., charge noise, thus leading to fluctuations in the local electric field [96]. Relative to the time-scale of spin qubit rotation times, these fluctuations can be treated quasistatically as a first approximation, but the actual power spectral density of charge noise in these qubit systems has been measured to behave like $1/f^{0.7}$ in GaAs [40] and $1/f$ in Si [41, 97] out to tens or even hundreds of kHz. The quasistatic part of the noise can be addressed by applying composite pulse sequences, where noisy gate operations are applied sequentially such that the gate errors conspire to cancel one another. These sequences, however, typically only suppress noise that is slow on the timescale of the sequence, and amplify noise that is faster [47].

The Zeeman fluctuations manifest in two ways depending on how the gradient is generated. When the gradient comes from the Overhauser effect due to the hyperfine coupling of the dot electron with the nuclear spin of the host semiconductor, such as in GaAs-based architectures using dynamical nuclear spin polarization [98–100], electron-mediated nuclear spin flip-flops produce $1/f^2$ noise [101, 102] that is essentially quasistatic. When the gradient comes from a micromagnet structure

[103], as used in some GaAs devices [104, 105] and which is necessary for silicon-based architectures with far fewer spinful nuclei [106], it is possible for charge noise to also couple in via small shifts in the dot position, again resulting in higher frequency noise [107].

Two-qubit gate fidelity in singlet-triplet systems is mostly limited by charge noise when the qubit dynamics is dominated by the exchange interaction [94, 108]. Recent work on capacitively-coupled, double-quantum dot spin qubits with gate-controlled exchange coupling between the spins has demonstrated suppression of charge noise by applying a strong magnetic gradient between the two dots in each qubit that is much stronger than the exchange interaction [88]. An analytical expression for the full time-evolution operator of this particular system can be obtained by using the rotating-wave approximation (RWA) [109].

In this chapter, we analyze how perturbations in the control parameters of a capacitively-coupled singlet-triplet system affect the time-evolution and present a strategy to minimize those effects. In Section 4.2, we derive the time-evolution operator using the RWA. We consider in Section 4.3 two different parameter regimes for two qubits with similar energy splitting: when the magnetic field gradient dominates the splitting, and when the exchange interaction dominates instead. We calculate the leading order errors and show that certain parameter choices result in a synchronization of oscillating error terms such that a passive reduction of gate errors occurs at specific times. In Section 4.4 we examine the effects of the optimization in the presence of both quasistatic noise and $1/f$ noise. We find that our optimization isolates the effects of noise into particular $\mathfrak{su}(4)$ basis elements, allowing us to prescribe composite pulse sequences to mitigate the remaining errors. In principle, this work allows the improvement of experimental two-qubit gate fidelities to above 99%. While most of our work is presented in the limit of zero pulse rise time, we show in Appendix D that typical finite rise times do not pose a challenge to the stroboscopic error suppression.

4.2 The Time-Evolution Operator

We consider a system of capacitively-coupled singlet-triplet qubits, which corresponds directly to the experimental setup in Ref. [88], but our results are also applicable to any system similarly described by a static Ising coupling and local driving fields. The effective two-qubit Hamiltonian is given by

$$\mathcal{H} = \sum_{i=1}^{2} \left(\frac{J_i + j_i \cos[\omega_i t]}{2} \sigma_Z^{(i)} + \frac{h_i}{2} \sigma_X^{(i)} \right) + \alpha\, \sigma_{ZZ}, \tag{4.1}$$

where $\sigma_{ij} \equiv \sigma_i^{(1)} \otimes \sigma_j^{(2)}$ with $\{i, j\} \in \{I, X, Y, Z\}$ collectively form a 15-dimensional $\mathfrak{su}(4)$ basis. The exchange interaction between two spins in the i^{th} qubit is a function of the difference in electrochemical potential between the dots, ε_i, which can vary in time. By oscillating ε_i, the exchange is caused to oscillate at a driving frequency ω_i, which makes the effective exchange interaction oscillate about an average value J_i with an amplitude j_i. The static, longitudinal magnetic field gradient is denoted by h_i; this can be generated by using either a micromagnet or, in GaAs, through the hyperfine interaction between the dot electrons and the nuclear spins in the semiconductor. Thus, the static part of a qubit's total energy splitting is $\Omega_i \equiv \sqrt{h_i^2 + J_i^2}$. Finally, α is the electrostatic coupling strength between the adjacent qubits, which is proportional to the product of the two qubits' electric dipole moments.

Ref. [109] reported an approximate time-evolution operator for the aforementioned Hamiltonian using the RWA. There it was implicitly assumed that $\frac{j_i J_i}{2\Omega_i} \ll \Omega_i$. We lift this assumption and apply the same formalism to find a more general description of the time evolution. We begin by first performing a local rotation to align the x-axis along the vector sum of the combined local static fields

$$U = \exp\left[\frac{\imath}{2}\sum_{i=1}^{2}\phi_i\sigma_Y^{(i)}\right]U_1\exp\left[-\frac{\imath}{2}\sum_{i=1}^{2}\phi_i\sigma_Y^{(i)}\right], \tag{4.2}$$

where $\phi_i \equiv \tan^{-1}(J_i/h_i)$ and U is the lab-frame evolution operator. We then transform to the rotating frame

$$U_1 = \exp\left[-\imath\sum_{i=1}^{2}\left(\frac{\omega_i t + \xi_i(t)}{2}\right)\sigma_X^{(i)}\right]U_2, \tag{4.3}$$

where the inclusion of $\xi_i(t) = \frac{j_i J_i \sin(\omega_i t)}{\omega_i \Omega_i}$ generalizes Ref. [109]. We perform the RWA by doing a coarse-grain time-average over a time scale $1/\alpha \gg \tau \gg \max\{1/\omega_i\}$. The addition of $\frac{\xi_i(t)}{2}$ in the local rotation causes some of the terms in the rotating-frame Hamiltonian to have nontrivial averages. The time-averaged evolution operator is given by

$$\begin{aligned} U_2 = \exp\Bigg[-\imath t\Bigg(&\sum_{i=1}^{2}\left(\chi_i\sigma_Z^{(i)} + \frac{\Omega_i - \omega_i}{2}\sigma_X^{(i)}\right) \\ &- \frac{h_1 J_2\alpha}{\Omega_1\Omega_2}\mathcal{J}_1\left[\frac{j_1 J_1}{\omega_1\Omega_1}\right]\sigma_{ZX} - \frac{h_2 J_1\alpha}{\Omega_1\Omega_2}\mathcal{J}_1\left[\frac{j_2 J_2}{\omega_2\Omega_2}\right]\sigma_{XZ} \\ &+ \frac{J_1 J_2\alpha}{\Omega_1\Omega_2}\sigma_{XX} + \frac{h_1 h_2\alpha}{2\Omega_1\Omega_2}\left(\mathcal{I}_{YY}\sigma_{YY} + \mathcal{I}_{ZZ}\sigma_{ZZ}\right)\Bigg)\Bigg], \end{aligned} \tag{4.4}$$

where $\mathcal{J}_i[z]$ is the i^{th} order Bessel function of the first kind, $\chi_i \equiv \frac{h_i\omega_i\mathcal{J}_1\left[\frac{j_i J_i}{\omega_i\Omega_i}\right]}{2J_i}$ is the

Rabi frequency, and

$$\mathcal{I}_{YY} = \frac{1}{\tau}\int_0^\tau 2\sin\left(\omega_1 t + \xi_1\right)\sin\left(\omega_2 t + \xi_2\right)\mathrm{d}t \tag{4.5}$$

$$\mathcal{I}_{ZZ} = \frac{1}{\tau}\int_0^\tau 2\cos\left(\omega_1 t + \xi_1\right)\cos\left(\omega_2 t + \xi_2\right)\mathrm{d}t. \tag{4.6}$$

We require $\omega_i \gg \left\{\left|\frac{h_i j_i}{2\Omega_i}\right|, \alpha\right\}$ to ensure the validity of the RWA.

To gain a better understanding of the entangling dynamics, we take another transformation to eliminate the remaining local operators in the Hamiltonian:

$$U_2 = \exp\left[-\imath t\sum_{i=1}^{2}\left(\frac{\Omega_i - \omega_i}{2}\sigma_X^{(i)} + \chi_i\sigma_Z^{(i)}\right)\right]U_3. \tag{4.7}$$

We set the control field at resonance with the energy splitting, $\omega_i = \Omega_i$, thus eliminating the $\sigma_X^{(i)}$ terms. Note that by completely dropping this off-resonant term below, we have limited the validity of our analysis to cases where perturbations in Ω_i are much less than χ_i. Lifting this assumption would not permit us to obtain a time-independent Hamiltonian. Nonetheless, this is not an unrealistic assumption. At this point, we can proceed the same way as in Ref. [109]. We apply another round of the RWA which requires $|\chi_i| \gg \alpha$. If $||\chi_1| - |\chi_2|| \ll \alpha$, the average time-evolution operator is given by

$$U_3 = \exp\left[\frac{-\imath\alpha t}{2}\left(\frac{h_1h_2\mathcal{I}_{YY} + 2J_1J_2}{2\Omega_1\Omega_2}(\sigma_{XX} + \sigma_{YY}) + \frac{h_1h_2}{\Omega_1\Omega_2}\mathcal{I}_{ZZ}\sigma_{ZZ}\right)\right], \tag{4.8}$$

but if $||\chi_1| - |\chi_2|| \gg \alpha$, we instead have

$$U_3 = \exp\left[-\imath t\frac{\alpha h_1h_2}{2\Omega_1\Omega_2}\mathcal{I}_{ZZ}\sigma_{ZZ}\right], \tag{4.9}$$

This reduces to the result of Ref. [109] in the regime $h_i \gg J_i$, which is experimentally relevant [88], but it becomes quite different when the exchange is dominant, as in earlier experiments [94, 108].

The entangling dynamics depend on whether the qubit energy splittings, Ω_i, are nearly equal or not. If the difference between the two energy splittings is much larger than α, $|\Omega_1 - \Omega_2| \gg \alpha$, $\mathcal{I}_{ZZ}$ and $\mathcal{I}_{YY}$ become small. Looking at Equation (4.8) and (4.9), one can avoid a suppressed coupling rate by setting the Rabi frequencies equal to one another, $\chi_1 = \chi_2$, and operating in the large exchange regime, $J_i \gg h_i$. On the other hand, if the two qubits have similar energy splittings, the effective coupling rate is $\sim \alpha$ regardless of which parameter dominates.

4.3 First-Order Error Channels

As previously mentioned, the magnetic field gradient, h_i, in singlet-triplet systems is produced by either micromagnets, as demonstrated in a silicon-based experiment [106], or the hyperfine interaction between the quantum dot electron and the nuclear spins, as has often been used in the case of GaAs [98–100]. Whereas the latter case allows some fine-tuning of h_i through dynamic nuclear polarization, the same is not true for micromagnets. Thus, we consider two main cases of experimental relevance – when h_i is tunable and when it is not. Furthermore, the sensitivity of the qubits to fluctuations depends on the parameter regime at work. If J_i and h_i are completely uncorrelated, the fluctuation on the qubit energy splitting is given by

$$\delta\Omega_i^2 = \frac{J_i^2 \delta J_i^2 + h_i^2 \delta h_i^2}{\Omega_i^2}. \tag{4.10}$$

Note that when either J_i or h_i completely dominates the energy splitting, the noise due to the weaker one is suppressed by a factor of their ratio. We know from experiments that δh_i is mostly quasistatic on the timescale of the gates [101, 102] and δJ_i contains both a quasistatic and a $1/f$ component [40]. Thus, it is best to suppress the $1/f$ δJ_i errors by choosing $h_i \gg J_i$ and then correct the residual quasistatic errors with spin echo protocols. This is consistent with the improvement reported in Ref. [88] when the magnetic field gradient was increased.

As discussed in the previous section, rapid entanglement in the $h_i \gg J_i$ regime only occurs when the two qubit energy splittings are tuned close to one another ($h_1 \approx h_2$). If one is forced to work with fixed but very different gradients ($|h_1 - h_2| \gtrsim \min\{h_i\}$), which is a possible scenario when micromagnets are used, then one must work in the $J_i \gg h_i$ regime. Therefore, we will limit our discussion to these two cases: when h_i is dominant and when J_i is dominant. We assume similar qubit energy splittings in both cases for convenience, particularly when simplifying Equations (4.5) and (4.6).

4.3.1 Similar qubits with $h_i \gg J_i$

We consider a system of similar qubits ($\Omega_1 = \Omega_2$) where the magnetic field gradient dominates the energy splitting ($h_i \gg J_i, \Omega_i \simeq h_i$) and the driving frequencies are equal and at resonance with the energy splitting ($\omega_1 = \omega_2 \equiv \omega = \Omega_i$) in the absence of noise. For simplicity, we take the case where the Rabi frequencies of the two qubits are dissimilar (Equation (4.9)), although our analysis can be extended to the similar Rabi case easily. In this parameter regime, we can expand $\mathcal{J}_1[z]$ to first-order and obtain $\chi_i \approx \frac{h_i j_i}{4\Omega_i}$, and $\xi_i(t) \approx 0$ which allows us to evaluate

Table 4.1: First-order errors obtained by projecting Δ onto an $\mathfrak{su}(4)$ basis formed by Kronecker products of Pauli operators.

σ_{IX}	$\left(\frac{\imath(h_2\delta J_2-J_2\delta h_2)\cos(\omega t)}{2\Omega_2^2}-\frac{\imath(h_2\delta h_2+J_2\delta J_2)}{h_2 j_2}\right)\cos\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)\sin(2\chi_2 t)$
σ_{IY}	$\frac{\imath(h_2\delta J_2-J_2\delta h_2)\left(\cos\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)\cos(\omega t)\cos(2\chi_2 t)-1\right)}{2\Omega_2^2}+\frac{2\imath(h_2\delta h_2+J_2\delta J_2)\cos\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)\sin^2(\chi_2 t)}{h_2 j_2}$
σ_{IZ}	$\frac{\imath(h_2\delta J_2-J_2\delta h_2)\left(j_2J_2t-2\Omega_2\sin(\omega t)\right)}{4\Omega_2^3}-\frac{\imath h_2 t\delta j_2}{4\Omega_2}$
σ_{XI}	$\left(\frac{\imath(h_1\delta J_1-J_1\delta h_1)\cos(\omega t)}{2\Omega_1^2}-\frac{\imath(h_1\delta h_1+J_1\delta J_1)}{h_1 j_1}\right)\cos\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)\sin(2\chi_1 t)$
σ_{XX}	0
σ_{XY}	0
σ_{XZ}	$\left(\frac{\imath(h_1\delta J_1-J_1\delta h_1)\cos(\omega t)\cos(2\chi_1 t)}{2\Omega_1^2}+\frac{2\imath(h_1\delta h_1+J_1\delta J_1)\sin^2(\chi_1 t)}{h_1 j_1}\right)\sin\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{YI}	$\frac{\imath(h_1\delta J_1-J_1\delta h_1)\left(\cos\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)\cos(\omega t)\cos(2\chi_1 t)-1\right)}{2\Omega_1^2}+\frac{2\imath(h_1\delta h_1+J_1\delta J_1)\cos\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)\sin^2(\chi_1 t)}{h_1 j_1}$
σ_{YX}	0
σ_{YY}	0
σ_{YZ}	$\left(\frac{\imath(J_1\delta h_1-h_1\delta J_1)\cos(\omega t)}{2\Omega_1^2}+\frac{\imath(h_1\delta h_1+J_1\delta J_1)}{h_1 j_1}\right)\sin\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)\sin(2\chi_1 t)$
σ_{ZI}	$\frac{\imath(h_1\delta J_1-J_1\delta h_1)\left(j_1J_1t-2\Omega_1\sin(\omega t)\right)}{4\Omega_1^3}-\frac{\imath h_1 t\delta j_1}{4\Omega_1}$
σ_{ZX}	$\left(\frac{\imath(h_2\delta J_2-J_2\delta h_2)\cos(\omega t)\cos(2\chi_2 t)}{2\Omega_2^2}+\frac{2\imath(h_2\delta h_2+J_2\delta J_2)\sin^2(\chi_2 t)}{h_2 j_2}\right)\sin\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{ZY}	$\left(\frac{\imath(J_2\delta h_2-h_2\delta J_2)\cos(\omega t)\sin(2\chi_2 t)}{2\Omega_2^2}+\frac{\imath(h_2\delta h_2+J_2\delta J_2)\sin(2\chi_2 t)}{h_2 j_2}\right)\sin\left(\frac{h_1h_2\alpha t}{\Omega_1\Omega_2}\right)\sin(2\chi_2 t)$
σ_{ZZ}	$\frac{\imath h_2J_1\alpha t(h_1\delta J_1-J_1\delta h_1)}{2\Omega_1^3\Omega_2}+\frac{\imath h_1J_2\alpha t(h_2\delta J_2-J_2\delta h_2)}{2\Omega_1\Omega_2^3}-\frac{\imath h_1h_2t\delta\alpha}{2\Omega_1\Omega_2}$

$\mathcal{I}_{YY} = \mathcal{I}_{ZZ} \approx 1$. Thus, combining Equations (4.2), (4.3), (4.7), and (4.9), the total time-evolution can be written as

$$U(t) = R_1(t) \exp\left[- \imath t \frac{\alpha h_1 h_2}{2\Omega_1 \Omega_2} \sigma_{ZZ} \right] R_2(t), \tag{4.11}$$

where the purely local operators $R_1(t)$ and $R_2(t)$ are given by

$$\begin{aligned} R_1(t) &= \exp\left[\frac{\imath}{2} \sum_{i=1}^{2} \phi_i \sigma_Y^{(i)} \right] \exp\left[-\imath \sum_{i=1}^{2} \frac{\omega_i t}{2} \sigma_X^{(i)} \right] \exp\left[-\imath t \sum_{i=1}^{2} \chi_i \sigma_Z^{(i)} \right], \\ R_2(t) &= \exp\left[\frac{-\imath}{2} \sum_{i=1}^{2} \phi_i \sigma_Y^{(i)} \right]. \end{aligned} \tag{4.12}$$

Since Equation (4.11) is already canonically decomposed into local and nonlocal components [110], it is clear to see how to "undo" the local part of the evolution that accompanies the entangling gate. By applying additional local operations, $R_1^\dagger$ and $R_2^\dagger$, in the absence of coupling, we obtain a purely nonlocal σ_{ZZ} gate,

$$R_1^\dagger(t) U(t) R_2^\dagger(t) = \exp\left[- \imath t \frac{\alpha h_1 h_2}{2\Omega_1 \Omega_2} \sigma_{ZZ} \right]. \tag{4.13}$$

So far we have been careful to distinguish between Ω_1 and Ω_2 so as to allow for the perturbative effect of noise, but other than that we have not discussed the effect of such a perturbation. Noise during the original entangling operation produces errors in both the nonlocal phase of Equation (4.11) and in its accompanying local operations given in Equation (4.12). The pre- and post-applied locals, $R_i^\dagger$, only undo the ideal local rotations accompanying the entangling gate, but any random perturbations are left uncanceled. By expanding each term in Equations (4.11) and (4.12) to first order in perturbations δJ_i, δj_i, δh_i, and $\delta\alpha$, and commuting all of the perturbations to the right, we may write the effect of the noise in the form

$$\begin{aligned} U_{nl}(t) &= R_1^\dagger(t)U(t)\left(\mathbb{1}+\Delta_0\left(t\right)\right)R_2^\dagger(t) \\ &= \exp\left[-\imath t\frac{\alpha h_1 h_2}{2\Omega_1\Omega_2}\sigma_{ZZ}\right]\left(\mathbb{1}+\Delta\left(t\right)\right) \\ &\simeq \exp\left[-\imath t\frac{\alpha}{2}\sigma_{ZZ}\right]\left(\mathbb{1}+\Delta\left(t\right)\right). \end{aligned} \tag{4.14}$$

where $\mathbb{1}$ is the identity operator, Δ_0 contains the first-order perturbation of the physical entangling operation U, and $\Delta \equiv R_2\Delta_0 R_2^\dagger$ is the resulting perturbation in the purely nonlocal operation. The approximate equality makes use of the fact that powers of J_i/h_i are negligibly small compared to the dominant errors we wish to correct. The error Δ due to the perturbations is reported in Table 4.1 in terms of its projections onto the 15 $\mathfrak{su}(4)$ basis elements, henceforth referred to as error channels,

$$\Delta = \frac{1}{4}\sum_{ij}\mathbf{tr}(\sigma_{ij}\Delta)\sigma_{ij}. \tag{4.15}$$

One prominent feature of these error channels is their oscillatory behavior. Notice that one can, for example, choose parameters such that $\sin(\chi_i t) = 0$ at the end of the entangling gate. By doing so, one effectively eliminates several error terms in Table 4.1. If we also choose parameters such that $\cos(\omega t) = 0$ at the time that the gate is complete, all but five of the error channels in Table 4.1 (σ_{ZI}, σ_{IZ}, σ_{YI},σ_{IY}, and σ_{ZZ}) will be synchronized to vanish at the gate time. We are thus left with a gate that is partially corrected, for both quasistatic and $1/f$ noise. This stroboscopic circumvention of error requires no knowledge of the errors involved, only that they are small enough for the higher-order terms in the error expansion to remain insignificant.

Table 4.2: The same errors reported in Table 4.1 after substituting the optimized parameters.

σ_{IY}	$\imath\frac{J_2\delta h_2 - h_2\delta J_2}{2\Omega_2^2}$
σ_{IZ}	$\imath\frac{\left((-1)^m h_2 - 2n_2 J_2\pi\right)(J_2\delta h_2 - h_2\delta J_2)}{2h_2\Omega_2^2} - \frac{\imath h_2 t\delta j_2}{4\Omega_2}$
σ_{YI}	$\imath\frac{J_1\delta h_1 - h_1\delta J_1}{2\Omega_1^2}$
σ_{ZI}	$\imath\frac{\left((-1)^m h_1 - 2n_1 J_1\pi\right)(J_1\delta h_1 - h_1\delta J_1)}{2h_1\Omega_1^2} - \frac{\imath h_1 t\delta j_1}{4\Omega_1}$
σ_{ZZ}	$\frac{\imath h_2 J_1\alpha t(h_1\delta J_1 - J_1\delta h_1)}{2\Omega_1^3\Omega_2} + \frac{\imath h_1 J_2\alpha t(h_2\delta J_2 - J_2\delta h_2)}{2\Omega_1\Omega_2^3} - \frac{\imath h_1 h_2 t\delta\alpha}{2\Omega_1\Omega_2}$

Specifically, stroboscopic error elimination can be achieved by choosing

$$t = (m + 1/2)\pi/\omega, \tag{4.16}$$

$$j_i = \frac{4n_i\Omega_i\omega}{h_i(m + 1/2)} \simeq \frac{4n_i\omega}{(m + 1/2)}, \tag{4.17}$$

where m and n_i are integers. We also want to produce a given nonlocal phase, $\exp\left[\imath\frac{\theta}{2}\sigma_{ZZ}\right]$, at the end of the operation. So, we have another constraint from Equation (4.13), which we can satisfy to good approximation by choosing m such that

$$\left|\frac{(m + 1/2)\pi}{\omega}\alpha - \theta\right|. \tag{4.18}$$

is minimized. Due to the typically weak coupling, $\alpha/\omega \ll 1$, the minimum value is likewise small and occurs at a large value of integer m (corresponding to a gate time containing many cycles of the driving field).

As mentioned earlier, we must take care to stay within a parameter regime where the RWA is valid. We use some of the remaining free parameters to ensure that the RWA remains valid for the choices above that lead to error cancellation.

We enforce the RWA condition of resonant driving ($\omega = \Omega_1 = \Omega_2$) by setting

$$h_2 = \sqrt{h_1^2 + J_1^2 - J_2^2} \simeq h_1 \tag{4.19}$$

with the values of J_i still free as of yet other than being small compared to h_i. We enforce the RWA conditions on the driving amplitude of $|\chi_i| \gg \alpha$ and $||\chi_1| - |\chi_2|| \gg \alpha$ by taking the integers of Equation (4.17) such that $n_2 = 2n_1$ in order to maximize the difference in Rabi frequencies while keeping both large (which can be ensured via the choice of n_1). In the case of detuning-controlled singlet-triplet qubits, due to the empirically exponential dependence of the exchange interaction on the detuning [40, 108], $\delta J_i \propto J_i$ and it is advantageous to choose small values of J_i, but while still maintaining $J_i > j_i$ in order to avoid calling for negative exchange. So, we will choose values of J_i slightly larger than j_i. Without loss of generality, and for the sake of concreteness, we take $j_1 = 2j_2, J_1 = 2J_2$. Finally, another physical consideration specific to the capacitively-coupled singlet-triplet system is the treatment of perturbations in the coupling, $\delta\alpha$. Since α is proportional to the product of the derivatives of the exchange interactions in each qubit and the proportionality constant is such that $\delta\alpha$ is about two order of magnitude smaller than δJ_i [108], its effects are negligible and can safely be ignored.

We summarize and combine all of the constraints above in the following set of robustness conditions:

$$
\begin{gathered}
h_1, J_1, n_1, \alpha, \theta \text{ are free and subject to} \\
\alpha \ll j_i < J_i \ll h_i \text{ with } n_1 \in \mathbb{Z}, \\
\omega = \sqrt{h_1^2 + J_1^2} \simeq h_1, \\
J_1 = 2J_2, j_1 = 2j_2, \\
h_2 = \sqrt{h_1^2 - 3J_1^2} \simeq h_1 \\
m = \operatorname{nint}\left(\frac{\theta}{\pi}\frac{\omega}{\alpha} - \frac{1}{2}\right), \\
t = (m + 1/2)\pi/\omega, \\
j_1 = \frac{4n_1\Omega_1\omega}{h_1(m+1/2)} \simeq \frac{4n_1\omega}{(m+1/2)},
\end{gathered}
\tag{4.20}
$$

where it suffices to meet the approximate equalities due to the condition $J_i \ll h_i$, and $\operatorname{nint}(x)$ is the nearest integer function. The effect of these constraints on the first-order error channels is shown in Table 4.2. With the parameter choices of Equation (4.20), the surviving five error channels are left with terms that are approximately proportional to $\frac{\delta j_i}{\alpha}$, $\frac{\delta J_i}{h_i}$, $\frac{J_i}{h_i}\frac{\delta h_i}{h_i}$, $\frac{J_i}{h_i}\frac{\delta J_i}{h_i}$, $\left(\frac{J_i}{h_i}\right)^2\frac{\delta J_i}{h_i}$, and $\left(\frac{J_i}{h_i}\right)^2\frac{\delta h_i}{h_i}$. The last four terms in the list are clearly negligible. By invoking the exponential behavior of the exchange interaction, we have $\delta J_i = \frac{\mathrm{d}J_i}{\mathrm{d}\varepsilon_i}\delta\varepsilon_i \propto J_i\delta\varepsilon_i$, which indicates that the second term in the list is also suppressed for $J_i \ll h_i$. However, the first term in the list is not necessarily small. Errors from δj_i accumulate linearly with the gate time and are, consequently, effectively proportional to $1/\alpha$. Again noting that the empirically exponential nature of the exchange implies $\delta j_i \propto j_i$, it is possible to avoid unnecessarily large δj_i by choosing the free integer n_1 that appears in j_i to be as small as possible while still maintaining the RWA condition of $|\chi_i| \gg \alpha$.

The low-frequency content of the remaining δj_i error can be removed by inserting a refocusing π-pulse about the x-axis of each qubit in between two entangling gates. This is a well-known strategy [31, 75, 111], making use of the fact that the local σ_{XX} insertion commutes with the nonlocal σ_{ZZ} phase but anticommutes with the σ_{IZ} and σ_{ZI} error terms.

Since we are left with only five error channels, extracting the first-order error of the refocused entangling gate like in Equation (4.14) is analytically straightforward. The refocusing process shuffles these errors among the $\mathfrak{su}(4)$ basis elements, some of which appear in the σ_{XX}, σ_{YY}, σ_{XY}, and σ_{YX} channels. These errors commute with the nonlocal σ_{ZZ} phase, which suggests that concatenating with a local π-pulse about the z-axis of either qubit, e.g. σ_{ZI}, can be used to further correct the residual errors in the refocused gate.

4.3.2 Similar Qubits with $J_i \gg h_i$

We follow the same process as before but now we assume that the magnetic field gradients are fixed. Since we are taking $J_i \gg h_i$, the terms in the evolution operator that are proportional to $\frac{h_1 h_2}{\Omega_1 \Omega_2}$ are negligibly small. Thus, to generate an entangling gate, it is preferable for us to take the case where $||\chi_1| - |\chi_2|| \ll \alpha$ (Equation (4.8)). Ignoring the negligible terms, the time-evolution is

$$U(t) = R_1(t) \exp\left[-\imath t \frac{\alpha J_1 J_2}{2\Omega_1 \Omega_2} \left(\sigma_{XX} + \sigma_{YY}\right)\right] R_2(t) \tag{4.21}$$

$$\simeq R_1(t) \exp\left[-\imath t \frac{\alpha}{2} \left(\sigma_{XX} + \sigma_{YY}\right)\right] R_2(t), \tag{4.22}$$

where the purely local operators $R_1(t)$ and $R_2(t)$ are given by

$$R_1(t) = \exp\left[\frac{-\imath}{2}\sum_{i=1}^{2}\phi_i\sigma_Y^{(i)}\right]\exp\left[-\imath\sum_{i=1}^{2}\frac{\omega_i t+\xi_i(t)}{2}\sigma_X^{(i)}\right]\exp\left[-\imath t\sum_{i=1}^{2}\chi_i\sigma_Z^{(i)}\right],$$
$$R_2(t) = \exp\left[\frac{\imath}{2}\sum_{i=1}^{2}\phi_i\sigma_Y^{(i)}\right]. \tag{4.23}$$

The error channels for this evolution can be calculated in a similar fashion as in the previous case; the results are reported in Appendix E.

We proceed to our goal of synchronizing the error terms so that they vanish at the gate time. We can eliminate a number of error terms by choosing our parameters so that $\sin(\chi_i t)$ and $\cos(\omega_i t + \xi_i(t))$ simultaneously vanish at the gate time. However, as before, a significant amount of error remains in the σ_{IZ} and σ_{ZI} channels. In this case, though, we cannot simply apply a refocusing π-pulse since these error channels do not commute with the entanglement generator $\sigma_{XX} + \sigma_{YY}$. Fortunately, Ref. [75] offers a sequence of 10 local π-pulses interspersed between short entangling operations that can deal with these anticommuting errors to first-order while reducing the entanglement generator to σ_{XX}. Therefore, it is again possible in principle to generate high-fidelity entangling gates from a combination of stroboscopic decoupling and composite pulses in this parameter regime.

However, we must note that the assumption following Equation (4.7) of $\delta\Omega_i \ll \chi_i$ is likely unrealistic in this $J_i \gg h_i$ case for the charge noise levels currently reported in singlet-triplet qubits. Quasistatic fluctuations in the detuning, $\delta\varepsilon$, typically have a standard deviation of several μV [40] and around $J \sim$GHz this can cause $\delta\Omega_i \sim 10$MHz, whereas in this regime $\chi_i \simeq h_i/4 \sim 10$MHz as well. We estimate that roughly an order of magnitude decrease in the charge noise strength, down to

Table 4.3: Average CPHASE fidelity in the presence of 20neV magnetic noise and 8μV quasistatic charge noise with a $1/f^{0.7}$ component of $0.9\text{nV}/\sqrt{\text{Hz}}$ at 1MHz.

Sequence	$\langle F\rangle_{unoptimized}$	$\langle F\rangle_{optimized}$
No refocusing	.768	.811
Singly refocused	.950	.974
Doubly refocused	.944	.996

under a microvolt, would be required in order to safely neglect off-resonance errors. Note that the previous case of $h_i \gg J_i$ did not have this problem because there $\delta\Omega_i$ is dominated by magnetic noise, which is typically $\sim$ 10neV, whereas in that regime $\chi_i \lesssim j_i/4 \sim$ 100neV. Therefore, the case of similar qubits with $h_i \gg J_i$ is a more feasible operating regime for our proposed high-fidelity two-qubit gates in a double quantum dot singlet-triplet system. In the context of silicon singlet-triplet qubits with micromagnet gradients, this along with our discussion at the beginning of Section 4.3 means that the silicon devices must be engineered to either allow enough tunability of the magnetic differences across each qubit (via dot positioning, etc.) for them to be equalized in situ, or to physically reduce charge noise in the device. The former seems an easier target.

4.4 Simulations

We now examine the effects of our optimization in the presence of quasistatic magnetic noise and $1/f^{0.7}$ charge noise [40]. We will simulate the fidelity of CPHASE gates generated by a single-shot pulse, a single spin echo composite pulse, and a double spin echo composite pulse for both unoptimized and stroboscopically optimized parameters.

We report in Table 4.3 a summary of the calculated fidelities. The magnetic noise was generated from a normal distribution with a standard deviation of 20neV [112, 113]. To generate the charge noise, we superimposed 20 random telegraph noises with the appropriate weighting [114] and relaxation times ranging from 1MHz to 1GHz [88] evenly spaced on a logarithmic scale with an amplitude of $0.9\text{nV}/\sqrt{\text{Hz}}$ at 1MHz. An additional quasistatic noise component is added to ensure that the integrated power spectral density from 0 to 1MHz is consistent with the experimentally reported noise amplitude of 8μV [40]. Finally, we translated the noise in detuning ϵ into noise in exchange J by using an exponential fit on the data reported in Ref. [40].

We numerically solve for the time-evolution operator using the unapproximated, time-dependent Hamiltonian in Equation (4.1) with the optimal parameters predicted by the RWA analysis above, and then convert it to a CPHASE gate by using the perfect local operations prescribed by the RWA, as in the left-hand side of Equation (4.13). Note that for these numerical calculations we do not assume that the RWA is accurate; e.g, we do not assume now that the right-hand side of Equation (4.13) holds. We calculate the average two-qubit gate fidelity [82]

$$\langle F\rangle = \frac{1}{16}\left[4 + \frac{1}{5}\sum_{\sigma_{ij}} \mathrm{Tr}\left[U_1\sigma_{ij}U_1^\dagger U_2\sigma_{ij}U_2^\dagger\right]\right], \tag{4.24}$$

where U_1 is the ideal CPHASE and U_2 is the actual noisy evolution, which we obtain purely numerically for a given set of parameter values and averaging over 1000 different noise realizations. Any error due to the RWA is also included in that

Table 4.4: Local parameters used in the simulations.

Parameters ($\frac{1}{2\pi}$MHz)	Unoptimized, all cases	Optimized, no refocusing / singly refocused	Optimized, doubly refocused
J_1	266	80	150
J_2	320	40	75
j_1	69	74	147
j_2	36	37	73
h_1	922	1000	1500
h_2	905	1002	1506

fidelity.

A summary of all the parameter values used in the simulations are provided in Table 4.4. We have taken $\alpha = 2\pi \times 2.3$MHz in all cases for consistency. For all pulse sequences the same unoptimized parameters are used, obtained from Ref. [109] consistent with the range reported in experiment [88]. On the other hand, the optimized parameters are chosen following the rules in Equation (4.20). We choose the free parameters $h_1 = 1$GHz, $J_1 = 80$MHz, and $n_1 = 4$ for the no refocusing and singly refocused case, ensuring that $h_1 \gg J_1 > j_1$. On the other hand, we take $h_1 = 1.5$GHz, $J_1 = 150$MHz, and $n_1 = 2$ for the doubly refocused case in order to compensate for the shorter gate time needed. These immediately determine the values of ω, J_2, and h_2 shown in Table 4.4. The value of θ can either be $\pi/2$, $\pi/4$, or $\pi/8$, depending on which composite pulse sequence is being performed, as we discuss below.

As previously mentioned, all the simulations target a CPHASE gate. When applying the singly refocusing pulse, we replace the simple CPHASE gate $U_{nl}\left(t_{\pi/2}\right)$

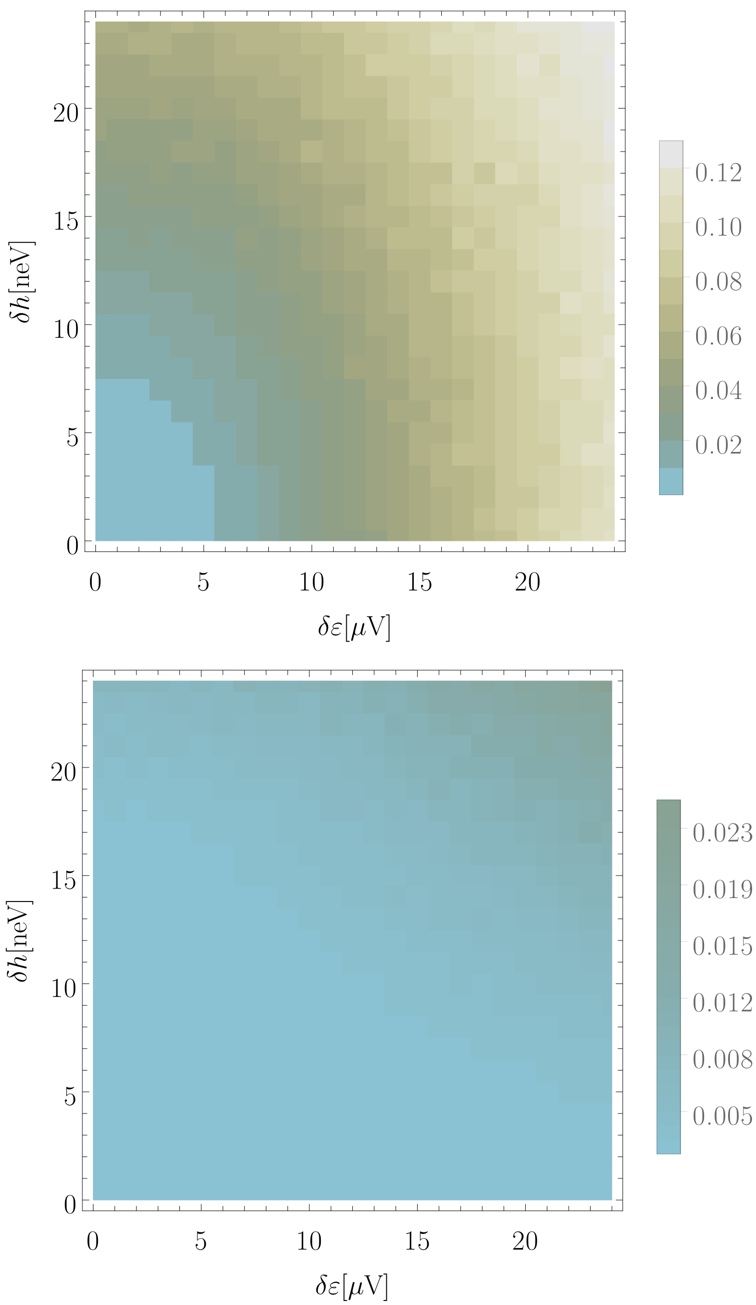


Figure 4.1: Average infidelity as a function of noise strength for the unoptimized (TOP) and optimized (BOTTOM) case after applying a doubly refocusing π-pulse. The values in the axes indicate the strength of quasistatic noise. $1/f$ noise is added to the exchange with an amplitude $0.9\text{nV}/\sqrt{\text{Hz}}$ at 1MHz.

with the composite CPHASE gate

$$U_{nl}(t_{\pi/4})\sigma_{XX}U_{nl}(t_{\pi/4})\sigma_{XX}, \tag{4.25}$$

where $U_{nl}(t_\theta)$ is the noisy entangling gate targeting a nonlocal phase θ and σ_{ab} is a local π rotation about the a-axis of the first qubit and the b-axis of the second qubit. The doubly refocused composite pulse requires twice as many component gates, but note that the entangling time is not any longer since each entangling component is shorter,

$$\begin{aligned} &\left[U_{nl}(t_{\pi/8})\sigma_{XX}U_{nl}(t_{\pi/8})\sigma_{XX}\right]\sigma_{ZI}\left[U_{nl}(t_{\pi/8})\sigma_{XX}U_{nl}(t_{\pi/8})\sigma_{XX}\right]\sigma_{ZI} \\ &\quad = U_{nl}(t_{\pi/8})\sigma_{XX}U_{nl}(t_{\pi/8})\sigma_{YX}U_{nl}(t_{\pi/8})\sigma_{XX}U_{nl}(t_{\pi/8})\sigma_{YX}. \end{aligned} \tag{4.26}$$

We further examine how our optimization behaves under a range of noise amplitudes. We keep the amplitude of the $1/f^{0.7}$ charge noise component the same as before for consistency, but we generate quasistatic noise with amplitudes ranging from 0 to 24 neV (μV) for magnetic (charge) noise. A contour plot of the average infidelity as a function of quasistatic noise strength for the case of doubly refocused gates is provided in Fig. 4.1. We find that combining our optimization scheme with the doubly refocusing pulse yields an order of magnitude improvement in fidelity compared to the unoptimized case. We emphasize that this improvement can be attributed to the isolation of error onto specific channels presented in Table 4.2. In fact, if one can further reduce the average fluctuations in the magnetic field gradient (e.g. down to 8neV [112]), it is possible to generate a CPHASE gate with average fidelities over 99% using only the singly refocusing pulse.

4.5 Conclusion

We theoretically analyze the first-order effects of errors in two capacitively-coupled singlet-triplet qubits by perturbing parameters in the time-evolution operator derived using the RWA. We examined two extreme regions of the parameter space and showed that it is better to operate in the parameter regime where the magnetic field gradient dominates the exchange than the opposite case.

We find that certain choices of parameter lead to passive, stroboscopic circumvention of errors. This enables the isolation of the errors onto specific basis elements of $\mathfrak{su}(4)$, consequently allowing the application of composite pulse sequence to mitigate the residual errors. Our numerical simulations show that our analytic prescription produces CPHASE gates with fidelities above 99% using only 4 applications of local π pulses on each qubit, which is an order of magnitude improvement over an unoptimized implementation.

Finally, we comment on our strategy of using composite pulse sequences to produce robust entangling operations. First, this method is greatly dependent on the availability of nearly perfect one-qubit operations. We saw in Chapter 3 that the efficacy of our proposed pulse sequences greatly diminishes as the one-qubit gate fidelity worsens. This is especially problematic in cases where the total composite gate time is comparable to the qubit's decoherence time. In this chapter, we were able to circumvent some of these issues by optimizing the control parameters. This eliminated many error channels that anticommute with the entangling operation which consequently made simple echo pulses sufficient to address the remaining error channels. We emphasize that this was only possible because of our analytical insight on the evolution operator. Therefore, it is important that we investigate alternative methods for suppressing systematic noise. This will be the subject of the next two chapters.

Chapter 5

Investigating the Robustness Conjecture for Geometric Quantum Gates

Geometric quantum gates are conjectured to be more resilient than dynamical gates against certain types of error, which makes them ideal for robust quantum computing. However, there are conflicting claims within the literature about the validity of that robustness conjecture. Here we use dynamical invariant theory in conjunction with filter functions in order to analytically characterize the noise sensitivity of an arbitrary quantum gate. For any control Hamiltonian that produces a geometric gate, we find that under certain conditions one can construct another control Hamiltonian that produces an equivalent dynamical gate with identical noise sensitivity (as characterized by the filter function). Our result holds for a Hilbert space of arbitrary dimensions, but we illustrate our result by examining experimentally relevant single-qubit scenarios and providing explicit examples of equivalent geometric and dynamical gates. This work was based on the paper arXiv:2105.02882 [115].

5.1 Introduction

One of the biggest roadblocks in quantum computing is developing techniques that enable control of quantum information under a certain error threshold [116]. Among the plethora of potential candidates for robust quantum control, geometric quantum computation (GQC) [117] stands out owing to its elegant formulation in terms of concepts from differential geometry and topology. Put simply, a geometric quantum gate is a type of quantum gate for which it is possible to attribute a geometric interpretation to the accumulated phase. The usual paradigm is to generate a desired quantum gate in a basis of cyclic states. After an adiabatic [118] or non-adiabatic [119] cyclic evolution, these states accumulate a phase that depends on the qubit's spectrum. If the computational basis is encoded in an energetically non-degenerate (degenerate) subspace of the total Hilbert space, the computational basis accumulates an Abelian (non-Abelian) phase [120, 121]. This phase can be decomposed into a dynamical and a geometric component. A geometric gate is naturally produced when the dynamical component of the total phase is trivial, though that condition is not necessary [122]. Further extension to noncyclic evolution has also been made [123]. Experimentally, geometric gates have been realized in nuclear magnetic resonance [124–126], trapped-ion [127, 128], solid-state [129–134], and superconducting qubits [135–139].

The primary motivation for using GQC is the robustness conjecture which claims that geometric gates are intrinsically more robust than dynamical gates [140]. This is typically supported by the reasoning that since geometric phase is a global feature of quantum evolution, then it must be intrinsically resilient to noise that only generates local perturbations in the system's evolution path [141]. Thus, a majority of the effort on GQC focuses on finding experimentally feasible ways of eliminating dynamical phase contributions in a gate. Numerous studies on geometric gates, both theoretical [117, 141–150] and experimental [138, 151, 152], have shown evidence to support the robustness conjecture. However, there are also studies that report control situations in which geometric gates are not intrinsically more robust than dynamical gates [153–157] and, in certain scenarios, their sensitivity to noise deteriorates [140, 158–161].

Here we analyze the robustness of geometric and dynamical quantum gates to coherent noise. We use dynamical invariant theory [162] in conjunction with filter functions [47, 163] in order to analytically characterize how noise sensitivity changes with the type of accumulated phase. We show that for any geometric gate it is possible to find, under certain conditions, an equivalent gate with the same filter function but with a phase whose nature can be continuously varied from purely geometric to purely dynamical. In other words, we show within our framework that noise robustness and phase type are unrelated concepts. This consequently invalidates the most general form of the robustness conjecture for geometric gates, and our analysis applies equally to adiabatic and non-adiabatic geometric gates. We explicitly illustrate our result in experimentally relevant single-qubit cases, including both Abelian and non-Abelian geometric gates. We also discuss how the presence of control constraints can give rise to preferential phase robustness. Our result may reconcile decades of seemingly contradictory claims on geometric gate robustness within the literature. Furthermore, our result calls into question the primary motivation for using GQC.

5.2 Theory

5.2.1 Dynamical Invariants

We begin by briefly describing how a geometric gate is generated. We temporarily restrict our attention to the Abelian case. A natural framework for considering geometric phase is through the theory of dynamical invariants [162, 164]. Although dynamical invariants have previously been used in the context of quantum control [165–171], we only use it here as a convenient way to describe the dynamical and geometric phases[1]. Consider a qubit system whose evolution is governed by some Hamiltonian $H(t)$. A dynamical invariant $I(t)$ is a solution to the Liouville-von Neumann equation

$$i\frac{\partial I(t)}{\partial t} - [H(t), I(t)] = 0, \tag{5.1}$$

where we use units such that $\hbar = 1$. The eigenvectors $|\phi_n(t)\rangle$ of $I(t)$ are related to the solutions of the Schrödinger equation by a global phase factor: $|\psi_n(t)\rangle = e^{i\alpha_n(t)}|\phi_n(t)\rangle$, where $\alpha_n(t)$ are the Lewis-Riesenfeld phases given by [172]

$$\alpha_n(t) = \alpha_{n,g}(t) + \alpha_{n,d}(t), \tag{5.2}$$

$$\alpha_{n,g}(t) = \int_0^t \langle\phi_n(t')|i\partial_{t'}|\phi_n(t')\rangle\, \mathrm{d}t', \tag{5.3}$$

$$\alpha_{n,d}(t) = -\int_0^t \langle\phi_n(t')|H(t')|\phi_n(t')\rangle\, \mathrm{d}t', \tag{5.4}$$

[1] We provide a more detailed discussion of dynamical invariants in Chapter 6.

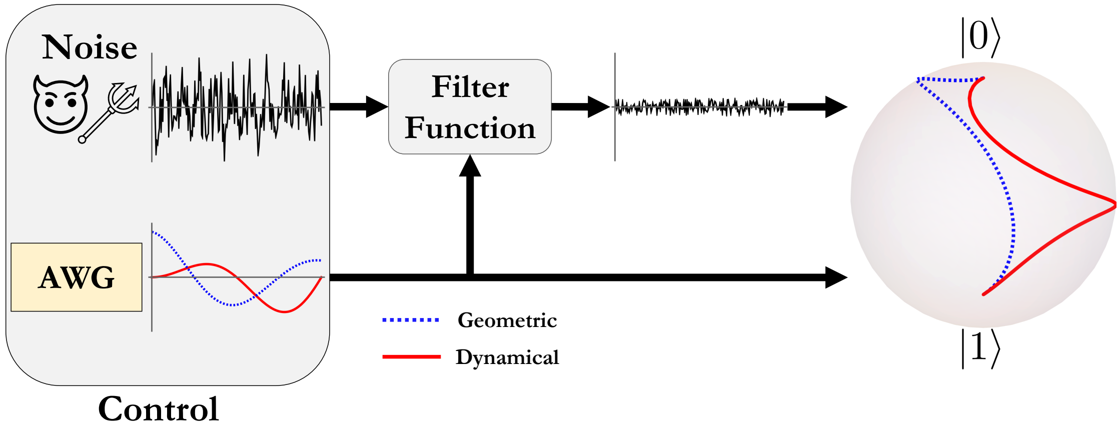


Figure 5.1: A schematic diagram summarizing our main result. The qubit's state is manipulated by applying control fields, e.g. using an arbitrary waveform generator (AWG), that are subject to some noise process (schematically represented here by a demon). The control fields determine an object called a filter function which characterizes the control's sensitivity to noise. In this diagram, the control is robust since it suppresses the effects of noise. In its most general form, the geometric gate robustness conjecture is that gates based on a geometric phase are naturally more robust than gates based on a dynamical phase. Our main result rejects this by showing how to construct different control fields, producing different evolution paths with accrued phases ranging from purely geometric to purely dynamical, that all result in the same gate and noise sensitivity.

and the subscripts g and d denote the geometric and dynamical phase, respectively. We fix the $U(1)$ gauge freedom on our choice of $|\phi_n(t)\rangle$ by setting $|\phi_n(0)\rangle = |\phi_n(T)\rangle$, where T is the gate time. This particular choice is consistent with Berry's adiabatic geometric phase [118] and generalizations thereof [119, 120, 173]. One can show that, unlike the dynamical phase which generally depends on T, the geometric phase is independent of T and is completely determined by the underlying geometric/topological property of the evolution path in Hilbert space. Within this framework, the evolution operator $U(t)$ can be expressed as

$$U(t) = \sum_n \mathrm{e}^{i\alpha_n(t)} |\phi_n(t)\rangle \langle\phi_n(0)| \,. \tag{5.5}$$

A geometric gate is produced if the final accumulated dynamical phase is trivial, which can be ensured by, for example, carefully choosing the Hamiltonian so that the integral in Equation (5.4) vanishes or by using composite pulses [174].

5.2.2 Filter Functions

The validity of the robustness conjecture can be tested using filter functions [47, 163], which provide a convenient method of quantifying the gate fidelity's susceptibility to noise of a given spectral composition. A noisy $\mathfrak{su}(N)$ Hamiltonian can be decomposed as

$$H(t) = H_c(t) + H_e(t), \tag{5.6}$$

where $H_c(t)$ is the ideal deterministic control Hamiltonian and $H_e(t)$ is the stochastic error Hamiltonian,

$$H_c(t) = \boldsymbol{h}_c(t) \cdot \boldsymbol{\sigma}, \quad H_e(t) = \sum_q \delta_q(t) \boldsymbol{\chi}_q \left[\boldsymbol{h}_c(t)\right] \cdot \boldsymbol{\sigma}, \tag{5.7}$$

where q indexes a set of uncorrelated stochastic variables $\delta_q(t)$, $\boldsymbol{\chi}_q$ is the vector describing the first-order sensitivity of the control Hamiltonian to $\delta_q(t)$, and $\boldsymbol{\sigma}$ is a vector comprising the $N^2 - 1$ traceless Hermitian generators of $\mathfrak{su}(N)$. Most commonly the sensitivity vector $\boldsymbol{\chi}_q$ is of the general linear form

$$\boldsymbol{\chi}_q \left[\boldsymbol{h}_c(t)\right] = \boldsymbol{a}_q + M_q(t) \boldsymbol{h}_c(t), \tag{5.8}$$

where $\boldsymbol{a}_q$ is independent of the control (i.e., additive noise) and M_q is likewise a real matrix accounting for sensitivity linearly proportional to some subset of the control (e.g., multiplicative noise).

For sufficiently weak noise, we can compactly express the ensemble averaged gate infidelity as

$$\langle \mathcal{I} \rangle \approx \frac{1}{2\pi} \sum_q \int_{-\infty}^{\infty} \mathrm{d}\omega \, S_q(\omega) F_q(\omega), \tag{5.9}$$

where $S_q(\omega)$ denotes the power spectral density for the stochastic variable $\delta_q(t)$ and $F_q(\omega)$ is the corresponding filter function. This is true only when the Magnus expansion of the evolution operator converges and higher-order noise contributions to the average gate infidelity are negligible [47]. Fortunately, these conditions can be easily satisfied in a well-prepared system such as in many state-of-the-art quantum devices which routinely achieve gate fidelities above 99%. Thus, it is safe to focus only on the first order term.

Denote the $N \times N$ unitary evolution operator generated by the control Hamiltonian, H_c, in the absence of noise as the time-ordered exponential

$$U_c(t) = \mathcal{T} e^{-i \int_0^t \mathrm{d}t' H_c(t')} \equiv e^{-i\boldsymbol{\theta}(t)\cdot\boldsymbol{\sigma}/2}. \tag{5.10}$$

U_c can also be represented via its adjoint representation, R, defined through

$$U_c \left(\boldsymbol{x}\cdot\boldsymbol{\sigma}\right) U_c^\dagger \equiv \left(R\boldsymbol{x}\right)\cdot\boldsymbol{\sigma} \implies R_{ij} = \mathrm{tr}\big(\sigma_i U_c \sigma_j U_c^\dagger\big)/N. \tag{5.11}$$

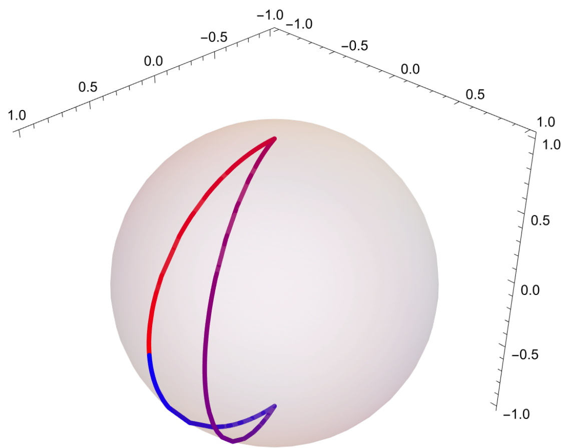
For example, in the case of $N = 2$, $R(t) = \exp\left(\boldsymbol{\theta}(t)\cdot\boldsymbol{L}\right)$, where $\boldsymbol{L}$ is the vector of generators of $\mathfrak{so}(3)$ isomorphic to $\boldsymbol{\sigma}$. In general, the filter function can be interpreted geometrically as the magnitude of a complex vector

$$F_q(\omega) = \boldsymbol{R}(\omega)\cdot\boldsymbol{R}(\omega)^*, \tag{5.12}$$

$$\boldsymbol{R}(\omega) = \int_0^T R^\intercal(t)\boldsymbol{\chi}_q\left[\boldsymbol{h}_c(t)\right]\mathrm{e}^{-i\omega t}\mathrm{d}t. \tag{5.13}$$

5.2.3 Invalidating the Robustness Conjecture

We can invalidate the robustness conjecture if we show that for any control scheme $\boldsymbol{h}_c(t)$ that generates a gate geometrically there exists a different control $\tilde{\boldsymbol{h}}_{\boldsymbol{c}}(t)$ that generates the same gate dynamically *with an identical filter function.* The existence of such a control consequently proves that a gate's phase type and noise robustness are unrelated. To this end, we calculate the filter function of two different control Hamiltonians, $H_c(t)$ and $\tilde{H}_c(t)$. Since for any arbitrary pair of time-dependent hermitian operators H_c and $\tilde{H}_c$ one can define a unitary $V = V_1 V_2$ where $\dot{V}_1 = -i\tilde{H}_c V_1$ and $\dot{V}_2 = +iH_c V_2$, the two can be related without loss of generality



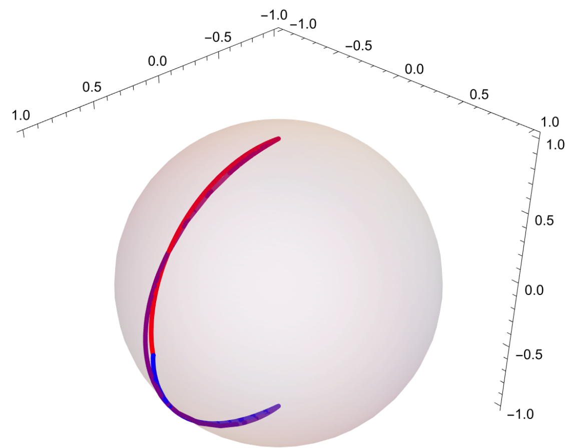


Figure 5.2: An illustration of a dynamical invariant eigenvector's evolution along the Bloch sphere for the geometric $X_{\frac{\pi}{2}}$ gate (LEFT) and the dynamical $X_{\frac{\pi}{2}}$ gate (RIGHT). The path is traversed twice and its orientation is determined by the color gradient which begins with red and ends with blue. We see that $|\phi_+(t)\rangle$ traces out a loop with nonzero area in the geometric case. In contrast, the loop in the dynamical case encloses zero area.

via a quantum canonical transformation

$$\tilde{H}_c = VH_cV^\dagger - iV\dot{V}^\dagger. \tag{5.14}$$

Note that we are not invoking a frame transformation here – indeed, the whole Hamiltonian is not subject to this transformation, only the control part – we are only using a convenient mathematical way to encapsulate in V the difference between any two control Hamiltonians within the same frame. The geometric and dynamical phases produced by the two different control fields differ by a shift that is easily expressed in terms of V [164, 175]:

$$\tilde{\alpha}_{n,g}(t) = \alpha_{n,g}(t) + \int_0^t \left\langle \phi_n(t') \middle| iV^\dagger\dot{V} \middle| \phi_n(t') \right\rangle \mathrm{d}t', \tag{5.15}$$

$$\tilde{\alpha}_{n,d}(t) = \alpha_{n,d}(t) - \int_0^t \left\langle \phi_n(t') \middle| iV^\dagger\dot{V} \middle| \phi_n(t') \right\rangle \mathrm{d}t'. \tag{5.16}$$

We emphasize again that, although it is well-known [175–177] that a frame transformation leaves the total Lewis-Riesenfeld phase invariant while shifting the dynamical and geometric phases, we are merely noting that different control Hamiltonians within a fixed frame generally have different dynamical and geometric phases and the difference is elegantly quantified in terms of the relationship V between the Hamiltonians. As we will return to below, the noise model of Equation (5.8) remains fixed, as it must in a fixed frame.

The control evolution operator induced by $\tilde{H}_c$ can be written in terms of that induced by H_c as $\tilde{U}_c(t) = V(t)U_c(t)V^\dagger(0)$. Likewise, the adjoint representations of the evolutions can be related, denoting the adjoint representation of V as Q, as

$$\tilde{R}(t) = Q(t)R(t)Q^\intercal(0). \tag{5.17}$$

Our goal now can be stated as finding two different control Hamiltonians in the same frame that satisfy three conditions:

1. The same final gate is produced for both control Hamiltonians, $\tilde{U}_c(T) = U_c(T)$.
2. A geometric phase is traded for a dynamical one via Equations (5.15)-(5.16).
3. The relevant filter function(s) produced are the same for both control Hamiltonians.

We can satisfy the first condition by requiring $V(0) = V(T) = \mathbb{1}$ and likewise for Q. The second condition can be satisfied by finding a Q such that a geometric evolution $R(t)$ is related to a dynamical evolution $\tilde{R}(t)$ by Equation (5.17). The third condition can be satisfied if, for a given noise source δ_q, the integrands in Equation (5.13) are equal; i.e., combining with Equation (5.17), it suffices that

$$Q^\intercal(t)\boldsymbol{\chi}_q\left[\tilde{\boldsymbol{h}}_c(t)\right] = \boldsymbol{\chi}_q\left[\boldsymbol{h}_c(t)\right]. \tag{5.18}$$

Note that the sensitivity vector $\boldsymbol{\chi}_q$ has a fixed functional dependence on its control input. (If we were making a frame transformation, this term would be transformed in the same way as the control Hamiltonian.) This reflects the fact that the underlying noise mechanism is fixed by the physics of the device, and is not under the control of the user. As an example, if we have some control field $h_1(t)$ with an error model $h_1(t) \to h_1(t)(1 + \delta_1(t))$, then any other choice of that control field must have the same noise dependence: $\tilde{h}_1(t) \to \tilde{h}_1(t)(1 + \delta_1(t))$. More generally, the sensitivity vector is simply evaluated as a function of the new control input

$$\tilde{\boldsymbol{h}}_{\boldsymbol{c}}(t) = Q(t)\boldsymbol{h}_{\boldsymbol{c}}(t) + \boldsymbol{h}_{\boldsymbol{Q}}(t),\ h_{Q,i}(t) \equiv \mathrm{tr}\Big(-iQ\dot{Q}^T\Lambda_i\Big)/N, \tag{5.19}$$

with $\boldsymbol{\Lambda}$ as the (N^2-1)-dimensional vector of generators of $\mathfrak{su}(N)$ isomorphic to $\boldsymbol{\sigma}$. In conjunction with Equation (5.18), this yields the condition

$$\begin{aligned} &Q^{\intercal}(t)\boldsymbol{a}_q + Q^{\intercal}(t)M_q(t)(Q(t)\boldsymbol{h_c}(t) + \boldsymbol{h_Q}(t)) \\ &\qquad = \boldsymbol{a_q} + M_q(t)\boldsymbol{h_c}(t). \end{aligned} \tag{5.20}$$

If we can find a $Q(t)$ that satisfies Equation (5.20), we will have two Hamiltonians that result in identical gates and filter functions but with different phase types. This simple fact is the crux of this paper.

While the existence of a solution to the nonlinear Equation (5.20) is not obvious, we simplify by taking the more restrictive condition that the first (second) term on the lhs must separately equal the first (second) term on the rhs. Thus, one should choose $Q(t)$ such that i) $\boldsymbol{a_q}$ is an eigenvector of $Q(t)$, ii) $[Q(t), M_q(t)] = 0$, and iii) $\boldsymbol{h_Q}(t)$ is in the null space of $M_q(t)$. Parameterizing as $Q(t) = \mathcal{T}\exp\{\int \boldsymbol{\omega}(t)\cdot\boldsymbol{\Lambda}\mathrm{d}t\}$, these conditions become i) $\omega_i(t) = 0$ if $\boldsymbol{a_q} \in \operatorname{Col}\Lambda_i$, ii) $[\boldsymbol{\omega}(t)\cdot\boldsymbol{\Lambda}, M_q(t)] = 0$, and iii) $\boldsymbol{\omega}(t) \in \operatorname{null} M_q(t)$. In practice it is typically easy to satisfy these conditions.

5.3 Examples

5.3.1 Abelian Case

To illustrate, consider a single qubit under a generic $\mathfrak{su}(2)$ control Hamiltonian

$$\boldsymbol{h_c}(t) = \frac{1}{2}\begin{pmatrix} \Omega(t)\cos(\varphi(t)) \\ \Omega(t)\sin(\varphi(t)) \\ \Delta(t) \end{pmatrix}. \tag{5.21}$$

This form is pertinent to a variety of qubit implementations such as superconducting qubits [57], quantum dot spin qubits [178], and NMR qubits [179] to name a few, corresponding to the rotating wave approximation for a two-level system driven by an oscillating field with amplitude Ω at a carrier frequency detuned from resonance by Δ, and with phase φ. Suppose that this qubit is subject to independent additive fluctuations in the resonance frequency, $\Delta \to \Delta + \delta_\Delta$, and in the phase, $\varphi \to \varphi + \delta_\varphi$, as well as multiplicative amplitude noise, $\Omega \to \Omega(1+\delta_\Omega)$, i.e., in terms of Equation (5.8),

$$\boldsymbol{a_\Delta} = \frac{1}{2}\hat{\boldsymbol{z}}, \quad M_\Delta = 0, \tag{5.22}$$

$$\boldsymbol{a_\varphi} = \boldsymbol{0}, \quad M_\varphi = \frac{1}{2}\left(\hat{\boldsymbol{y}}\hat{\boldsymbol{x}}^T - \hat{\boldsymbol{x}}\hat{\boldsymbol{y}}^T\right), \tag{5.23}$$

$$\boldsymbol{a_\Omega} = \boldsymbol{0}, \quad M_\Omega = \frac{1}{2}\left(\hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^T + \hat{\boldsymbol{y}}\hat{\boldsymbol{y}}^T\right). \tag{5.24}$$

Note that this noise model encompasses most, if not all, scenarios treated in the literature of non-adiabatic geometric gates.

The three flavors of Equation (5.20) corresponding to $q = \Delta, \varphi, \Omega$ are all satisfied by choosing $Q(t) = \mathrm{e}^{\nu(t)\Lambda_z}$ (and hence $\boldsymbol{h_Q} = \dot{\nu}(t)\hat{\boldsymbol{z}}/2$) where Λ_z is the z-rotation generator in $\mathfrak{so}(3)$ and $\nu(t)$ is any function satisfying $\nu(0) = \nu(T) = 0$. Thus, for any geometric gate produced by a particular choice of $\Omega(t)$, $\varphi(t)$, and $\Delta(t)$ in Equation (5.21), one can implement the same gate with identical robustness by using the modified control

$$\tilde{\boldsymbol{h}}_{\boldsymbol{c}}(t) = \frac{1}{2}\begin{pmatrix} \Omega(t)\cos(\varphi(t) + \nu(t)) \\ \Omega(t)\sin(\varphi(t) + \nu(t)) \\ \Delta(t) + \dot{\nu}(t) \end{pmatrix} \tag{5.25}$$

and the free parameter $\nu(t)$ allows a way to tune the nature of the Lewis-Riesenfeld phase as indicated in Equations (5.15) and (5.16).

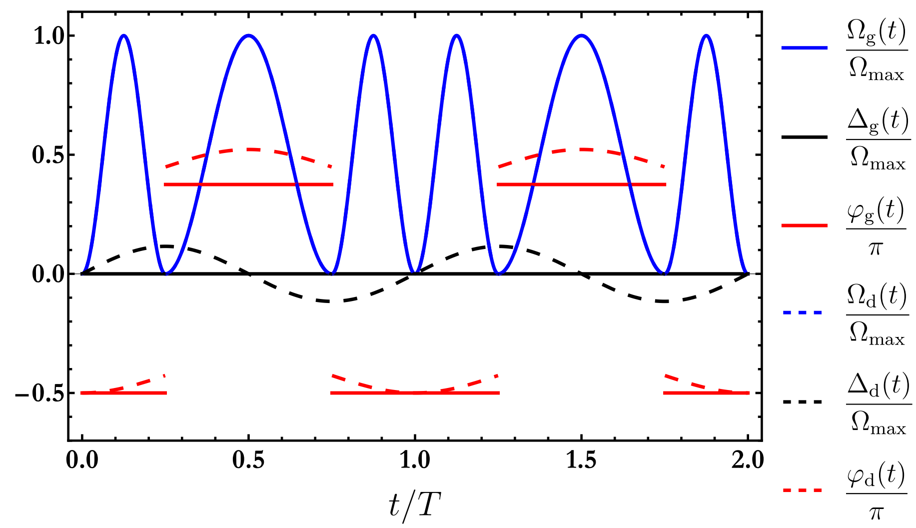


Figure 5.3: A plot of the control parameters that generate an $X_{\frac{\pi}{2}}$ gate. The subscript "g" ("d") denote the control parameters that generate a geometric (dynamical) gate. The values are normalized by Ω_{max} which denote the maximum value of $\Omega_g(t)$. Note that $\Omega_g(t) = \Omega_d(t)$ and that $\Delta_d(t)$ is non-trivial.

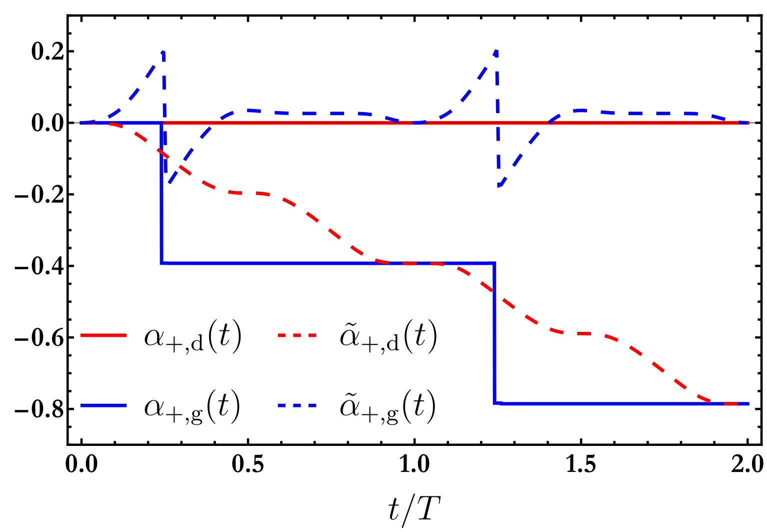


Figure 5.4: A comparison of the geometric and dynamical phases generated by the state $|\phi_+(t)\rangle$. The variables with (without) tilde correspond to the dynamical (geometric) $X_{\frac{\pi}{2}}$ gate.

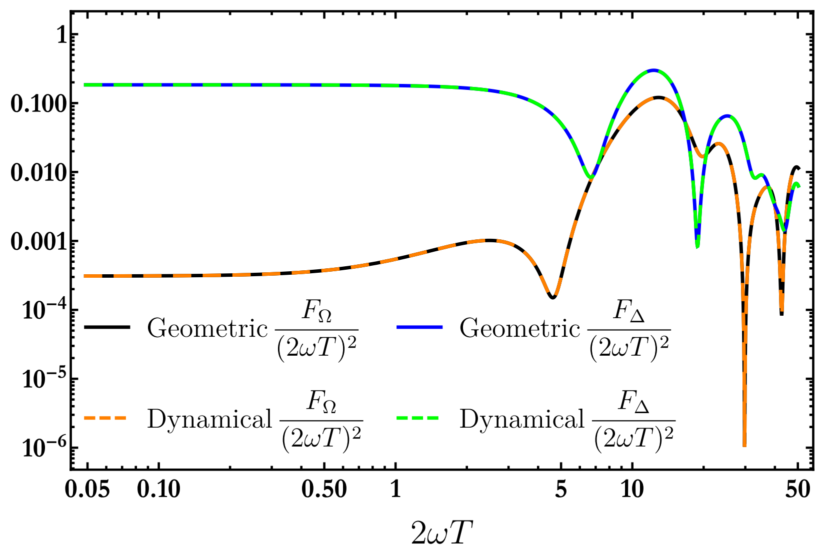


Figure 5.5: A comparison of the geometric and dynamical $X_{\frac{\pi}{2}}$ gate filter functions for additive dephasing and multiplicative amplitude noise when $\Omega_{max} = 1$. We verify that the two control Hamiltonians produce the same filter functions.

5.3.1.1 Orange-slice scheme

We further reinforce our claim by providing explicit examples of matching the noise sensitivity of a geometric gate with an equivalent dynamical gate. We begin by considering the Abelian geometric gate proposed in Ref. [180]. The dynamical invariant eigenstate traces out an orange-slice path along the Bloch sphere (see Fig. 5.2), and the geometric phase is equivalent to half the enclosed area [181]. We present in Equations (5.26)-(5.28) the corresponding control constraints in terms of the Hamiltonian of Equation (5.21) with $\Delta(t) = 0$:

$$t \in [0, T_1] \qquad \int_0^{T_1} \Omega \mathrm{d}t = \theta \qquad \varphi = \eta - \frac{\pi}{2}, \tag{5.26}$$

$$t \in [T_1, T_2] \qquad \int_{T_1}^{T_2} \Omega \mathrm{d}t = \pi \qquad \varphi = \eta + \gamma + \frac{\pi}{2}, \tag{5.27}$$

$$t \in [T_2, T] \qquad \int_{T_2}^{T} \Omega \mathrm{d}t = \pi - \theta \qquad \varphi = \eta - \frac{\pi}{2}. \tag{5.28}$$

We denote the generated evolution operator by $U_0(t) = e^{i\gamma \boldsymbol{n}\cdot\boldsymbol{\sigma}}$, where

$$\boldsymbol{n} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta).$$

We can produce a two-part composite gate that suppresses additive dephasing ($\Delta \to \Delta + \delta_\Delta$) and multiplicative amplitude noise ($\Omega \to \Omega + \delta_\Omega \Omega$) by applying the same evolution twice [148]: $U(2T) = U_0^2(T)$. We target a geometric $X_{\frac{\pi}{2}}$ gate which can be achieved by setting $\gamma = -\frac{\pi}{8}, \theta = \frac{\pi}{2}$, and $\eta = 0$. We assume the pulse shape $\Omega(t) = \sin^2(\pi t/\tau)$ where τ is the length of the relevant time interval. To generate its purely dynamical equivalent, we use the modified Hamiltonian of Equation (5.25) with an arbitrary choice of $\nu(t) = c\sin^2(\pi t/T)$, numerically tuning c until the geometric phase is zero, which occurs at $c \approx 0.461875$. We present in Fig. 5.3 a plot of the control parameters for both geometric and dynamical $X_{\frac{\pi}{2}}$ gate. We use Equations (5.3) and (5.4) to verify that the modified Hamiltonian produces a purely dynamical gate. The dynamical invariant eigenvectors $|\phi_\pm(t)\rangle$ are determined using the inverse engineering scheme in Ref. [165]:

$$|\phi_+(t)\rangle = \begin{pmatrix} \cos\left(\frac{\gamma(t)}{2}\right)\exp\left(-i\beta(t)\right) \\ \sin\left(\frac{\gamma(t)}{2}\right) \end{pmatrix}, \tag{5.29}$$

$$|\phi_-(t)\rangle = \begin{pmatrix} \sin\left(\frac{\gamma(t)}{2}\right) \\ -\cos\left(\frac{\gamma(t)}{2}\right)\exp\left(i\beta(t)\right) \end{pmatrix}, \tag{5.30}$$

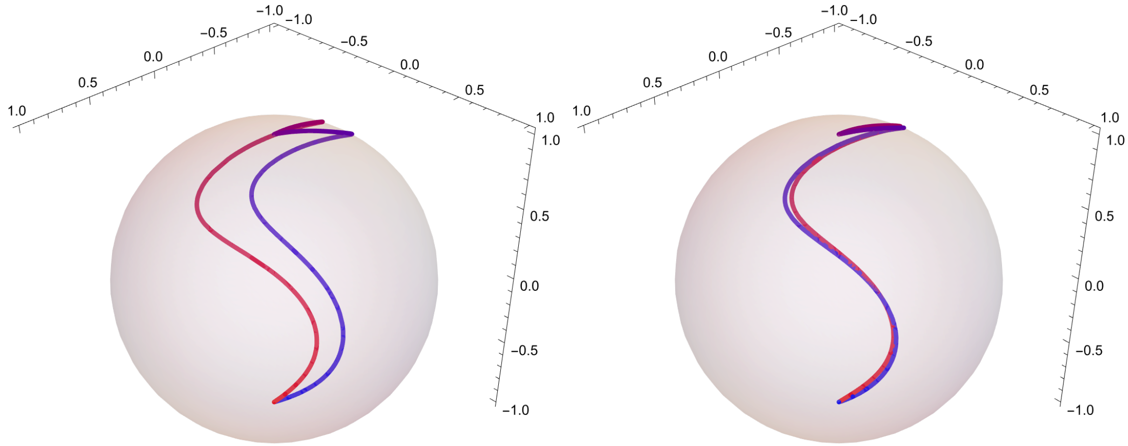


Figure 5.6: An illustration of a dynamical invariant eigenvector's evolution along the Bloch sphere for the geometric T-gate (LEFT) and the dynamical T-gate (RIGHT). The path orientation is determined by the color gradient which begins with red and ends with blue. Just like in the previous example, we find that $|\phi_+(t)\rangle$ traces out a loop with nonzero area in the geometric case and zero area in the dynamical case.

where the parameters γ and β obey the following coupled differential equations:

$$\dot{\gamma} = -\Omega \sin(\beta - \varphi), \tag{5.31}$$

$$\dot{\beta} = \Delta - \Omega \cot\gamma \cos(\beta - \varphi). \tag{5.32}$$

We set the boundary conditions so that $\gamma(0) = \frac{\pi}{2}$ and $\beta(0) = 0$ which corresponds to an eigenvector of $X_{\frac{\pi}{2}}$. The effect of this evolution on $|\phi_+(t)\rangle$ is shown in Fig. 5.2. Finally, a comparison of the geometric and dynamical phases for both gates is shown in Fig. 5.4 and their corresponding filter functions for dephasing and amplitude noise in Fig. 5.5.

5.3.1.2 Inverse engineering with optimal control

Next, we consider the case of a non-adiabatic Abelian geometric gate that is produced using Hamiltonian inverse engineering and optimal control theory [182]. Suppose that we target a T-gate ($Z_{\frac{\pi}{4}}$) as in Ref. [182]. The optimized inverse-engineered control parameters are given by

$$\Omega = -\frac{\dot{\gamma}}{\sin(\beta - \varphi)}, \tag{5.33}$$

$$\varphi = \beta - \arctan\left(\frac{\dot{\gamma}}{\dot{\beta}} \cot\gamma\right), \tag{5.34}$$

$$\Delta = 0, \tag{5.35}$$

which depend on the piecewise-defined functions $\gamma(t)$ and $\beta(t)$ that satisfy

$$t \in [0, T/2]: \quad \gamma(t) = \pi \sin^2(\pi t/T), \tag{5.36}$$

$$\beta(t) = -\frac{4}{3}\cos\left(\frac{\pi}{2}\cos\left(\frac{2\pi t}{T}\right)\right), \tag{5.37}$$

$$t \in [T/2, T]: \quad \gamma(t) = \pi \sin^2(\pi t/T), \tag{5.38}$$

$$\beta(t) = -\frac{4}{3}\cos\left(\frac{\pi}{2}\cos\left(\frac{2\pi t}{T}\right)\right) + \frac{\pi}{8}. \tag{5.39}$$

The gate time is $T = \frac{\sqrt{17}\pi^2}{\Omega_{\max}}$, where $\Omega_{\max}$ is the maximum value of $\Omega(t)$. To generate its purely dynamical equivalent, we again use the modified Hamiltonian of Equation (5.25) with an arbitrary choice of $\nu(t) = c\sin\left(\frac{2\pi t}{T}\right)$, numerically tuning c until the geometric phase is zero, which occurs at $c \approx 0.220530$. We set the boundary conditions so that $\gamma(0) = \beta(0) = \pi$ which corresponds to an eigenvector of $Z_{\frac{\pi}{4}}$. The effect of this evolution on $|\phi_+(t)\rangle$ is shown in Fig. 5.6. We present in Fig. 5.7 a plot of the geometric and the dynamical T-gate's control parameters. Finally, a compar-

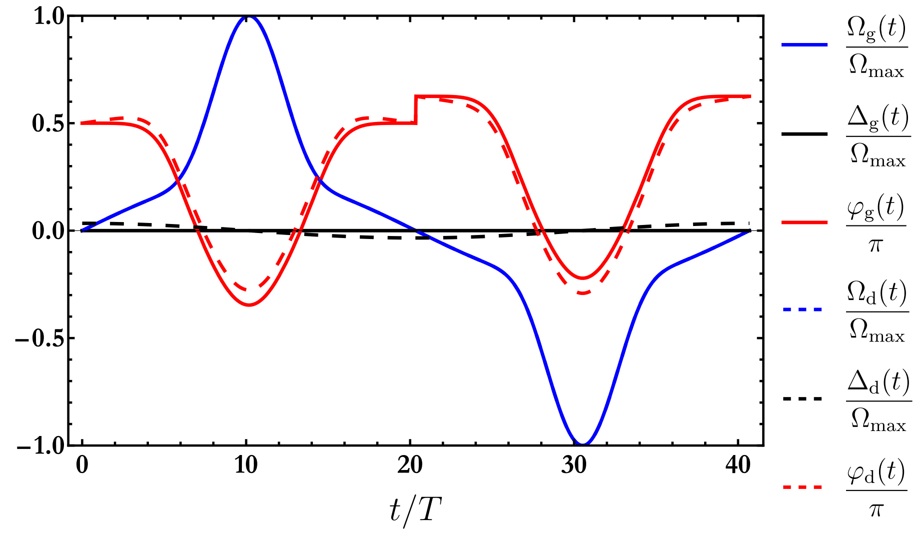


Figure 5.7: A plot of the control parameters that generate a $Z_{\frac{\pi}{4}}$ or T-gate. The subscript "g" ("d") denote the control parameters that generate a geometric (dynamical) gate. The values are normalized by Ω_{max} which denote the maximum value of $\Omega(t)$. We note again that $\Omega_g(t) = \Omega_d(t)$ and that $\Delta_d(t)$ is non-trivial.

ison of the geometric and dynamical phases for both gates is shown in Fig. 5.8 and their corresponding filter functions for dephasing and amplitude noise in Fig. 5.9.

5.3.2 Non-Abelian case

We now extend this treatment to the non-Abelian case. Unlike the Abelian case, where ensuring the dynamical phase is zero at the final time is a constraint on the geometric pulse design, non-Abelian geometric quantum computing typically encodes the computational basis in an energetically degenerate subspace of the full Hilbert space such that any dynamical phase is either automatically zero at all times or can be treated as a global phase factor. We generalize our previous framework

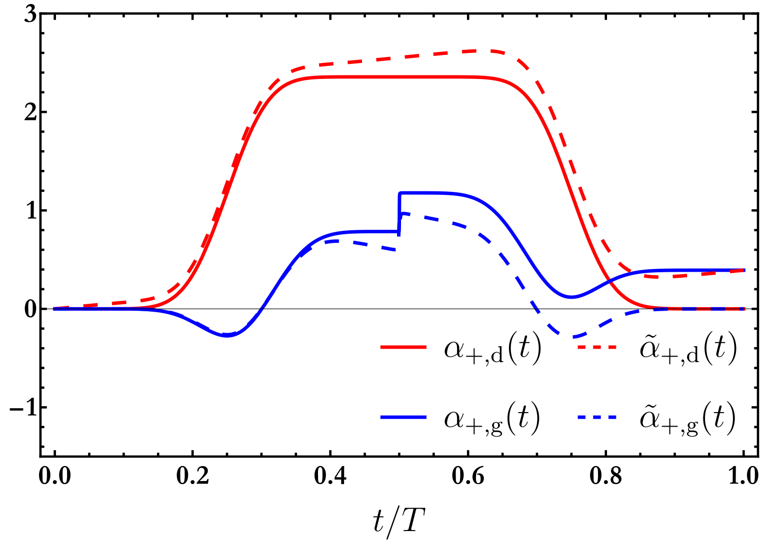


Figure 5.8: A comparison of the geometric and dynamical phases generated by the state $|\phi_+(t)\rangle$. The variables with (without) tilde correspond to the dynamical (geometric) T-gate.

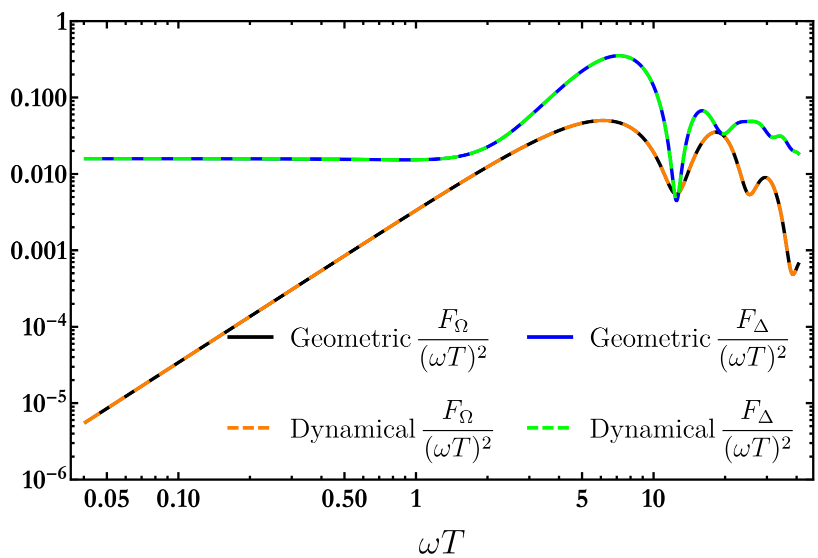


Figure 5.9: A comparison of the geometric and dynamical T-gate filter functions for additive dephasing and multiplicative amplitude noise when $\Omega_{max} = 1$. We verify that the two control Hamiltonians produce the same filter functions.

and denote the eigenvectors of $I(t)$ by $|\phi_{n;a}(t)\rangle$ where $a \in \{1, 2, \dots d_n\}$ labels the orthonormal basis vectors of a d_n-fold degenerate subspace corresponding to the n^{th} eigenvalue. The propagator of Equation (5.5) generalizes to [164]

$$U(t) = \sum_n \sum_{a,b=1}^{d_n} u_{n;ab}(t) \, |\phi_{n;a}(t)\rangle \, \langle \phi_{n;b}(0)| \,, \tag{5.40}$$

where the eigenstates accumulate a non-Abelian phase $u_n(t)$ given by

$$u_n(t) = \mathcal{T} e^{i \int_0^t \mathcal{A}_n(t') + \mathcal{E}_n(t') \mathrm{d}t'}, \tag{5.41}$$

$$\mathcal{A}_{n;ab}(t) = \langle \phi_{n;a}(t) | i\partial_t | \phi_{n;b}(t) \rangle \,, \tag{5.42}$$

$$\mathcal{E}_{n;ab}(t) = - \langle \phi_{n;a}(t) | H(t) | \phi_{n;b}(t) \rangle \,. \tag{5.43}$$

Thus, if we again consider the effect of changing the Hamiltonian as expressed by a quantum canonical transformation (again, without loss of generality) via time-dependent unitary V, we get an expression for the change in the geometric and dynamical components of the Lewis-Riesenfeld phase similar to Equations (5.15) and (5.16):

$$\tilde{\mathcal{A}}_{n;ab}(t) = \mathcal{A}_{n;ab}(t) + \left\langle \phi_{n;a}(t) \middle| i V^\dagger \dot{V} \middle| \phi_{n;b}(t) \right\rangle , \tag{5.44}$$

$$\tilde{\mathcal{E}}_{n;ab}(t) = \mathcal{E}_{n;ab}(t) - \left\langle \phi_{n;a}(t) \middle| i V^\dagger \dot{V} \middle| \phi_{n;b}(t) \right\rangle . \tag{5.45}$$

We note that the dynamical and geometric contributions to the phase are easily separable in the Abelian case, as in Equations (5.15) and (5.16). In the non-Abelian case, the gate accumulates matrix-valued dynamical and geometric phase components at each time step, as seen in Equation (5.41), which generally do not commute. Thus, the inseparability of the phase's time-ordered integral can lead to nontrivial dynamical contributions even if the integral of $\mathcal{E}(t)$ in Equation (5.43) is zero. This is why purely geometric non-Abelian gates are typically defined to have $\mathcal{E}(t) = 0$ within the computational basis [181, 183]. Therefore, to illustrate that a non-Abelian gate is no longer purely geometric, it is sufficient to show that the transformation V produces a nontrivial $\tilde{\mathcal{E}}$ in Equation (5.45). Specifically, it is sufficient to show that $\tilde{\mathcal{E}}_{n;ab}(t) \neq 0$ where a and b indexes the computational basis states.

On the other hand, it is possible to simultaneously diagonalize both matrices in special cases where $[\mathcal{A}(t), \mathcal{E}(t)] = 0$ which allows the decoupling of the geometric and dynamical phase contribution (for example, in adiabatic non-Abelian geometric gates [121]). In such cases, it is again straightforward to separate and tune the two types of phase.

5.3.2.1 Non-adiabatic case

As an example, consider a three-level system in a Λ configuration where the states $|0\rangle$ and $|1\rangle$ are coupled to an excited state $|e\rangle$ [183]. The $k \leftrightarrow e$ transition ($k = 0, 1$) is separately driven by a laser pulse with fixed polarization and frequency. Following our notation in Equation (5.7), the system-laser interaction is described by the following rotating-frame Hamiltonian belonging to an $\mathfrak{su}(3)$ algebra:

$$\boldsymbol{h_c}(t) = \begin{pmatrix} 0 \\ 0 \\ \frac{\Delta_0 - \Delta_1}{2} \\ \Omega\left(t\right) \cos\left(\frac{\varphi}{2}\right) \sin\left(\frac{\theta}{2}\right) \\ \Omega\left(t\right) \sin\left(\frac{\varphi}{2}\right) \sin\left(\frac{\theta}{2}\right) \\ -\Omega\left(t\right) \cos\left(\frac{\varphi}{2}\right) \cos\left(\frac{\theta}{2}\right) \\ \Omega\left(t\right) \sin\left(\frac{\varphi}{2}\right) \cos\left(\frac{\theta}{2}\right) \\ \frac{\Delta_0 + \Delta_1}{2\sqrt{3}} \end{pmatrix}, \tag{5.46}$$

where θ and φ are fixed angles that describe the relative strength and relative phase of the $k \leftrightarrow e$ transitions, Δ_k are detunings which can be independently varied, $\Omega(t)$ describes the pulse amplitude envelope, and $\boldsymbol{\sigma}$ is chosen to comprise the Gell-Mann matrices. If we impose the constraint that $\int_0^T \Omega(t)\mathrm{d}t = \pi$ and drive the qubit at resonance ($\Delta_k = 0$), the evolution produces a purely geometric gate which, when projected in the computational space spanned by $\{|0\rangle, |1\rangle\}$, yields [183]

$$\mathrm{proj}_{\{|0\rangle,|1\rangle\}}\left[U(T)\right] = \begin{pmatrix} \cos\theta & e^{-i\varphi}\sin\theta \\ e^{i\varphi}\sin\theta & -\cos\theta \end{pmatrix}. \tag{5.47}$$

It is possible to generate any single-qubit operation by applying Equation (5.47) with different values of θ and φ.

Suppose that this qubit is subject to independent additive fluctuations in the laser detunings, $\Delta_k \to \Delta_k + \delta_{\Delta_k}$, in their relative strength, $\theta \to \theta + \delta_\theta$, and in their relative phase, $\varphi \to \varphi + \delta_\varphi$, as well as multiplicative amplitude noise, $\Omega \to \Omega(1+\delta_\Omega)$.

Then, in terms of Equation (5.8), we have

$$(\boldsymbol{a_{\Delta_k}})_i = (-1)^k \pi \delta_{i,3} + \frac{\pi}{\sqrt{3}} \delta_{i,8}\,, \quad M_{\Delta_k} = 0, \tag{5.48}$$

$$\boldsymbol{a_\Omega} = \mathbf{0}, \quad M_\Omega = E_{4,4} + E_{5,5} + E_{6,6} + E_{7,7}, \tag{5.49}$$

$$\boldsymbol{a_\theta} = \mathbf{0}, \quad M_\theta = \frac{1}{2}\left(E_{5,7} - E_{7,5} + E_{6,4} - E_{4,6}\right), \tag{5.50}$$

$$\boldsymbol{a_\varphi} = \mathbf{0}, \quad M_\varphi = \frac{1}{2}\left(E_{5,4} - E_{4,5} + E_{6,7} - E_{7,6}\right), \tag{5.51}$$

where $E_{i,j}$ is a square matrix which show the value 1 at the position (i, j) and zeros elsewhere [184]. It is straightforward to verify that the transformation $Q(t) = \mathrm{e}^{\nu(t)\Lambda_3}$, where Λ_i are the adjoint representations of the Gell-Mann matrices [2], uniquely satisfies all the previously specified criteria. (If any one of these noise sources is irrelevant, there is more freedom in the transformation.) Thus, for any non-Abelian gate produced by a particular choice of θ, φ, and $\Omega(t)$ in Equation (5.47), one can implement the same gate with identical robustness using the modified control

$$\tilde{\boldsymbol{h}}_{\boldsymbol{c}}(t) = \begin{pmatrix} 0 \\ 0 \\ \frac{\Delta_0 - \Delta_1 + \nu'(t)}{2} \\ \Omega\left(t\right) \cos\left(\frac{\varphi + \nu(t)}{2}\right) \sin\left(\frac{\theta}{2}\right) \\ \Omega\left(t\right) \sin\left(\frac{\varphi + \nu(t)}{2}\right) \sin\left(\frac{\theta}{2}\right) \\ -\Omega\left(t\right) \cos\left(\frac{\varphi + \nu(t)}{2}\right) \cos\left(\frac{\theta}{2}\right) \\ \Omega\left(t\right) \sin\left(\frac{\varphi + \nu(t)}{2}\right) \cos\left(\frac{\theta}{2}\right) \\ \frac{\Delta_0 + \Delta_1}{2\sqrt{3}} \end{pmatrix} \tag{5.52}$$

where the free parameter $\nu(t)$ breaks the degeneracy of an equal-detuning setting, and similar to the Abelian case, provides a way to tune the nature of the Lewis-

[2] We distinguish the adjoint representation of a group which is defined in Equation (5.11) from the adjoint representation of a Lie algebra which can be calculated using the structure constants of the algebra f_{ijk} obeying $[\sigma_i, \sigma_j] = \sum_k i f_{ijk} \sigma_k$ as $[\mathrm{ad}(\sigma_i)]_{jk} = -i f_{ijk}$.

Riesenfeld phase as indicated in Equations (5.44) and (5.45).

As previously mentioned, we need only show that our transformation yields $\tilde{\mathcal{E}}_{n;ab}(t) \neq 0$ within the computational subspace to guarantee that the gate is no longer purely geometric. Since $H(t)$ commutes with itself at all times, we can calculate the resulting evolution operator $U(t)$ analytically

$$U(t) = \exp\left[-i\overline{\Omega}(t)\begin{pmatrix} 0 \\ 0 \\ 0 \\ \cos\left(\frac{\varphi}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\varphi}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\varphi}{2}\right)\cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\varphi}{2}\right)\cos\left(\frac{\theta}{2}\right) \\ 0 \end{pmatrix}\cdot\boldsymbol{\sigma}\right], \tag{5.53}$$

where we denote $\overline{\Omega}(t) \equiv \int_0^t \Omega(s)\mathrm{d}s$. We note that a dynamical invariant can be constructed by using the cyclic states of $U(T)$ as its eigenbasis [164]. To proceed, we first compute the eigenvectors and eigenvalues of $U(t)$:

$$\lambda_1(t) = 1 \qquad |\lambda_1(t)\rangle = \begin{pmatrix} e^{-i\frac{\varphi}{2}}\cos\left(\frac{\theta}{2}\right) \\ e^{i\frac{\varphi}{2}}\sin\left(\frac{\theta}{2}\right) \\ 0 \end{pmatrix}, \tag{5.54}$$

$$\lambda_2(t) = e^{-i\overline{\Omega}(t)} \qquad |\lambda_2(t)\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-i\frac{\varphi}{2}}\sin\left(\frac{\theta}{2}\right) \\ -e^{i\frac{\varphi}{2}}\cos\left(\frac{\theta}{2}\right) \\ 1 \end{pmatrix}, \tag{5.55}$$

$$\lambda_3(t) = e^{i\overline{\Omega}(t)} \qquad |\lambda_3(t)\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -e^{-i\frac{\varphi}{2}}\sin\left(\frac{\theta}{2}\right) \\ e^{i\frac{\varphi}{2}}\cos\left(\frac{\theta}{2}\right) \\ 1 \end{pmatrix}. \tag{5.56}$$

The cyclic states $|\phi_i(t)\rangle$ of $U(T)$ are linear combinations of $|\lambda_i\rangle$ up to global time-dependent phase which we choose so that $|\phi_i(0)\rangle = |\phi_i(T)\rangle$:

$$|\phi_1(t)\rangle = |\lambda_1(t)\rangle, \tag{5.57}$$

$$|\phi_2(t)\rangle = e^{i\overline{\Omega}(t)}\frac{\lambda_2(t)|\lambda_2(t)\rangle - \lambda_3(t)|\lambda_3(t)\rangle}{\sqrt{2}}, \tag{5.58}$$

$$|\phi_3(t)\rangle = e^{i\overline{\Omega}(t)}\frac{\lambda_2(t)|\lambda_2(t)\rangle + \lambda_3(t)|\lambda_3(t)\rangle}{\sqrt{2}}. \tag{5.59}$$

Consequently, we can define a dynamical invariant as $I(t) = \sum_i c_i |\phi_i(t)\rangle \langle\phi_i(t)|$ where c_i are arbitrary constants. Using the eigenvectors of $I(t)$, we can calculate the change in the dynamical phase contribution under the transformation $V = e^{-i\frac{\nu(t)}{2}\lambda_3}$ (or, equivalently, by $Q = e^{\nu(t)\Lambda_3}$ in the adjoint representation) using Equation (5.45):

$$
\tilde{\mathcal{E}}(t) = \frac{1}{4}\begin{pmatrix} -2\cos(\theta)\nu'(t) & -\sin(\theta)\,\nu'(t)\left(1+e^{i2\overline{\Omega}(t)}\right) & -\sin(\theta)\,\nu'(t)\left(1-e^{i2\overline{\Omega}(t)}\right) \\ -\sin(\theta)\,\nu'(t)\left(1+e^{-i2\overline{\Omega}(t)}\right) & 2\cos(\theta)\cos^2(\overline{\Omega}(t))\nu'(t) & -4\Omega(t) - i\cos(\theta)\sin\left(2\overline{\Omega}(t)\right)\nu'(t) \\ -\sin(\theta)\,\nu'(t)\left(1-e^{-i2\overline{\Omega}(t)}\right) & -4\Omega(t) + i\cos(\theta)\sin\left(2\overline{\Omega}(t)\right)\nu'(t) & 2\cos(\theta)\sin^2\left(\overline{\Omega}(t)\right)\nu'(t) \end{pmatrix} \tag{5.60}
$$

Since $\tilde{\mathcal{E}}(t)$ is nontrivial in the computational subspace, then the gate $\tilde{U}(T)$ must not be purely geometric by definition, though its filter function is the same as $U(T)$.

5.3.2.2 Adiabatic case

We next consider the case of an adiabatic non-Abelian geometric gate. Specifically, we consider a four-level system with three ground or metastable states coupled to a single excited state as in Ref. [185, 186]. The system is controlled using three distinctly polarized and resonantly driven lasers which, in the rotating frame, yields the following control Hamiltonian:

$$
\tilde{\boldsymbol{h}}_{\boldsymbol{c}}(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \Omega(t)\cos\varphi(t)\sin\theta(t) \\ \Omega(t)\sin\varphi(t)\sin\theta(t) \\ \Omega(t)\cos\theta(t) \end{pmatrix}, \tag{5.61}
$$

where θ controls the relative strength between the lasers, φ controls their relative phases, and $\boldsymbol{\sigma}$ is chosen to comprise of the $\mathfrak{so}(4)$ generators:

$$
\begin{aligned}
e_1 &= i(E_{2,1} - E_{1,2}), & e_2 &= i(E_{3,1} - E_{1,3}), \\
e_3 &= i(E_{3,2} - E_{2,3}), & e_4 &= E_{4,1} + E_{1,4}, \\
e_5 &= E_{4,2} + E_{2,4}, & e_6 &= E_{4,3} + E_{3,4}.
\end{aligned} \tag{5.62}
$$

We assumed for simplicity that all laser magnitudes are constant through the evolution and are sufficiently large to ensure adiabaticity. The control parameters θ and φ are then tuned cyclically so that at the gate time $t = T$ we have $H(0) = H(T)$. The dynamics of the system can be described using the eigenvectors of $H(t)$:

$$
\lambda = 0 \qquad |\lambda_{1,0}\rangle = \begin{pmatrix} \cos\theta(t)\cos\varphi(t) \\ \cos\theta(t)\sin\varphi(t) \\ -\sin\theta(t) \\ 0 \end{pmatrix}, \tag{5.63}
$$

$$
\lambda = 0 \qquad |\lambda_{1,1}\rangle = \begin{pmatrix} -\sin\varphi(t) \\ \cos\varphi(t) \\ 0 \\ 0 \end{pmatrix}, \tag{5.64}
$$

$$
\lambda = -\Omega(t) \qquad |\lambda_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos\varphi(t)\sin\theta(t) \\ \sin\varphi(t)\sin\theta(t) \\ \cos\theta(t) \\ -1 \end{pmatrix}, \tag{5.65}
$$

$$
\lambda = \Omega(t) \qquad |\lambda_3\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos\varphi(t)\sin\theta(t) \\ \sin\varphi(t)\sin\theta(t) \\ \cos\theta(t) \\ 1 \end{pmatrix}, \tag{5.66}
$$

with $|\lambda_{1,0}\rangle$ and $|\lambda_{1,1}\rangle$ spanning the energetically degenerate computational basis. In this basis, $\mathcal{E}_{n;ab}(t) = \varepsilon_n(t)\delta_{ab}$ which consequently decouples $\mathcal{E}$ and $\mathcal{A}$ in Equation (5.41). Thus, the adiabatic evolution operator is given by

$$
U(t) = \sum_{n=1}^{3} \sum_{a,b=0}^{1} u_{n;ab} \left|\lambda_{n;a}(t)\right\rangle \left\langle\lambda_{n;b}(t)\right|, \tag{5.67}
$$

where

$$u_n(t) = e^{-i\int_0^t \varepsilon_n(t)} \mathcal{T} e^{\int_0^t \mathcal{A}_n(t')\mathrm{d}t'}, \tag{5.68}$$

$$\mathcal{A}_{n;ab}(t) = \langle \lambda_{n;a}(t) | i\partial_t | \lambda_{n;b}(t) \rangle . \tag{5.69}$$

As a result, any accumulated dynamical phase in the computational basis can be treated as a global phase factor. Therefore, the nontrivial effects of the evolution is completely due to the non-Abelian geometric phase.

Suppose that this system is subject to independent additive fluctuations in the lasers' relative strength, $\theta \to \theta + \delta_\theta$, in their relative phase, $\varphi \to \varphi + \delta_\varphi$, as well as multiplicative amplitude noise, $\Omega \to \Omega(1 + \delta_\Omega)$. Then, in terms of Equation (5.8), we have

$$\boldsymbol{a_\Omega} = \mathbf{0}, \qquad M_\Omega = E_{4,4} + E_{5,5} + E_{6,6}, \tag{5.70}$$

$$\boldsymbol{a_\theta} = \mathbf{0}, \qquad M_\theta = -\tan\theta(t) E_{6,6}, \tag{5.71}$$

$$\boldsymbol{a_\varphi} = \mathbf{0}, \qquad M_\varphi = \frac{1}{2}\left(E_{5,4} - E_{4,5}\right). \tag{5.72}$$

It can be easily verified that the transformation $V = e^{-i\nu(t)e_1}$ (or, equivalently, by $Q = e^{2\nu(t)\mathrm{ad}(e_1)}$) satisfies the conditions we outlined in the main text, where $\mathrm{ad}(e_1)$ denotes the adjoint representation of e_1. We can calculate the corresponding change in the dynamical phase contribution using Equation (5.45):

$$\tilde{\mathcal{E}}(t) = \begin{pmatrix} 0 & i\cos(\theta(t))\,\nu'(t) & 0 & 0 \\ -i\cos(\theta(t))\,\nu'(t) & 0 & -\frac{i\sin(\theta(t))\nu'(t)}{\sqrt{2}} & -\frac{i\sin(\theta(t))\nu'(t)}{\sqrt{2}} \\ 0 & \frac{i\sin(\theta(t))\nu'(t)}{\sqrt{2}} & -\Omega(t) & 0 \\ 0 & \frac{i\sin(\theta(t))\nu'(t)}{\sqrt{2}} & 0 & \Omega(t) \end{pmatrix}, \tag{5.73}$$

where the computational subspace is located in the upper 2×2 block. We again see that $\tilde{\mathcal{E}}(t)$ is nontrivial in the computational subspace which indicates that the gate $\tilde{U}(T)$ is not purely geometric even though its filter function is the same as $U(T)$. We further note that this transformation yields a control Hamiltonian with non-degenerate energy levels. Thus, the geometric and dynamical components of the phase integral in Equation (5.41) are no longer decoupled which is in contrast with the non-Abelian version of the gate.

5.4 Discussion

We have demonstrated in Section 5.3 that in many experimentally relevant scenarios there exist families of solutions to Equation (5.20) which implies that the notion of noise robustness is independent of phase type. This invalidates the most general form of the robustness conjecture since it is always possible in principle to find two Hamiltonians that produce the same gate and noise sensitivity but with polar opposite phase types. In other words, there is nothing particularly special about geometric gates when it comes to robustness. At first glance, this may seem to contradict the significant evidence in the literature that supports the robustness conjecture for geometric gates. It is crucial to note, however, that the solution set of Equation (5.20) is only non-trivial when the control Hamiltonian is not severely constrained. Here constraints refer to the physical limitations of a specific qubit implementation such as control parameter bounds, only two-axis control, or bandwidth limitations. Depending on the error model in consideration and the severity of the constraints, there can be scenarios where the only solution to Equation (5.20) is a trivial one ($Q(t) = \mathbb{1}$). This clearly happens, for example, with an unusual error

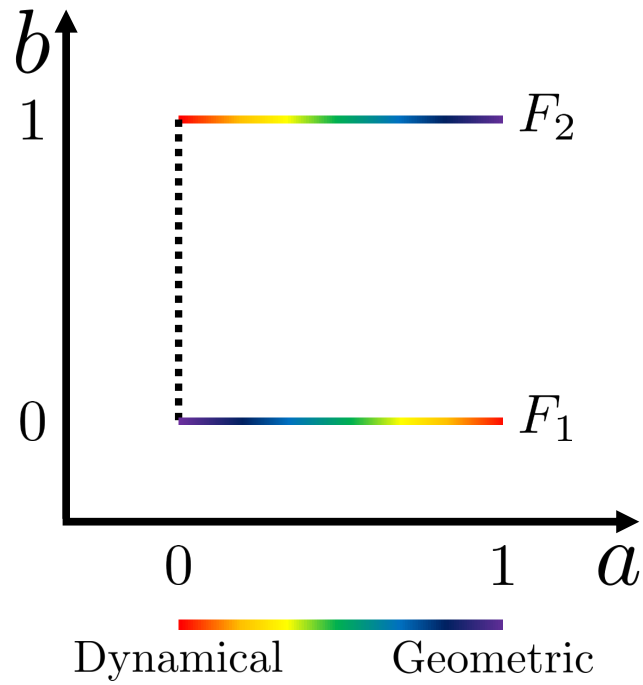


Figure 5.10: A schematic illustration demonstrating how control constraints can give rise to preferential phase robustness. Here we consider the set of control Hamiltonians described by Equation (5.75). The colored horizontal lines represent families of control Hamiltonians that preserve the filter function, F_1 and F_2 respectively. The color gradient indicates the gate's phase type, which can range from purely dynamical to purely geometric. The dotted line represents a subset of the control space that is physically accessible in a given experiment as a result of strict constraints (generally this dotted line will broaden into an extended region of the plane). In this case, the constraint prohibits one from traversing the horizontal lines, so one would obtain a different noise sensitivity (i.e., filter function) for a geometric gate versus a dynamical gate.

model of $\boldsymbol{a_q} = 0$ and $M_q = \mathbb{1}$ in Equation (5.8). Likewise, geometric gates naturally emerge as superior in the particular case of a strictly two-axis control Hamiltonian with static multiplicative amplitude error [174]. In such special cases, the correspondence between filter function and phase type is unique, i.e., phase preference emerges.

We can illustrate how phase preference emerges using Fig. 5.10. Generally one has many tunable parameters in the control Hamiltonian (e.g., even for only a single time-dependent control field, one has the value of the field over each infinitesimal time step). However, for the sake of being able to sketch an illustration, consider a control Hamiltonian with only two free parameters,

$$\boldsymbol{h_c}(t; a, b) = Q(t)\left[b\boldsymbol{h_g}(t) + (1-b)\boldsymbol{h_d}(t)\right] + \boldsymbol{h_Q}(t), \tag{5.74}$$

$$Q(t) = \mathcal{T}\exp\left\{a\int\left[b\boldsymbol{\omega_g}(t) + (1-b)\boldsymbol{\omega_d}(t)\right]\cdot\boldsymbol{\Lambda}\mathrm{d}t\right\}, \tag{5.75}$$

where $\boldsymbol{h_c}$ ($\boldsymbol{h_d}$) denotes a specific physically accessible control Hamiltonian that produces a particular target gate geometrically (dynamically), $\boldsymbol{\omega_g}$ ($\boldsymbol{\omega_d}$) denotes a rotation axis vector that determines an operator Q which solves Equation (5.20) such that it produces the same target gate with an identical filter function but a different phase type (assuming such a solution exists for the relevant error model, as in the examples of Section 5.3), and a, b parameterize a continuous deformation between these four specific points in the control space.

The horizontal lines of Fig. 5.10 are sets of control fields that all yield the same filter function (labeled F_1 and F_2 in the figure). The color gradient indicates the phase type across these lines which can range from purely dynamical to purely geometric. However, physical constraints may only allow access to some subarea of the $a - b$ plane. For example, in a severely constrained case, one may only have access to Hamiltonians with $a = 0$ in Equation (5.75), indicated by the dotted line in Fig. 5.10. So, although in this paper we have shown that a geometric gate generally has a dynamical equivalent with equal noise sensitivity (and vice versa), a strict control constraint could prohibit one from accessing these equivalent controls in practice. As depicted in Fig. 5.10, one would then observe different noise sensitivities for different phase types.

A notable example where this behavior is observed is in Ref. [140]. Here the authors considered the control Hamiltonian in Equation (5.21) with the constraint that $\varphi = \varphi_0 t$ and Ω, Δ, and φ_0 are constants. In addition, they assumed that Ω and Δ are subject to multiplicative noise. By further imposing the restriction that $\varphi = \left[\Delta \pm \sqrt{\Delta^2 - \eta(\Omega^2 + \Delta^2)}\right]/\eta$ and $\Delta = \Omega\sqrt{\eta/(1-\eta)} + \Delta_0$ where η and Δ_0 are constants, it is possible for them to switch between a geometric and a dynamical gate. However, it can be easily verified that the constraints they take do not permit the control to be changed as prescribed in Equation (5.25). Thus, when they change the gate's phase type (move up/down the dotted line), they also changed the gate's filter function. In their case, they found that geometric gates performed better than dynamical gates. In the situation of Ref. [154], where a strictly two-axis control scheme with additive noise on both axes was considered, the constraints again preclude moving along the horizontal lines of Fig. 5.10 but in this case it is dynamical gates that were found to perform better than geometric gates.

Thus, our result can be used to reconcile seemingly contradictory claims in the literature regarding the robustness conjecture of geometric gates, in that studies that support the conjecture considered a constrained model in which the solution set of Equation (5.20) favor geometric gates, whereas studies that do not support the conjecture considered a constrained model that favor dynamical gates instead. We emphasize, however, that the apparent superiority of either phase type is merely a consequence of restricting the solution set of Equation (5.20) since we have shown in Section 5.3 that they are generically equivalent in the absence of constraints. Determining in which scenario dynamical gates or geometric gates are superior can only be done on a case-by-case basis as it is determined by the noise model as well as the particular constraints.

Realizing equivalent Hamiltonians may require degrees of freedom in the control to be present in one that are not present in the other. For example, a tunable detuning, $\Delta(t)$, is necessary to produce the dynamical gates seen in Fig. 5.3 and 5.7. While controlling the detuning is not entirely common, this level of control has already been achieved in superconducting qubits [187] and in quantum dot charge qubits [188].

In addition, although our analysis is only strictly valid for coherent noise models, it can also be applied to dissipative processes. We have verified through Lindblad master equation simulations that the equivalent geometric and dynamical gates are identically affected by dephasing and relaxation. This behavior is expected since, by construction, the two Hamiltonians produce gates with the same duration and filter function.

5.5 Conclusion

In summary, we examine the broadband noise-resilience of geometric and dynamical gates using filter functions and show that there exists no intrinsic advantage for one or the other – for any control Hamiltonian producing a geometric gate one can find a different control Hamiltonian that produces a completely equivalent dynamical gate in the same frame. We illustrate this explicitly in a one-qubit scenario for both the Abelian and non-Abelian case. Our argument applies to both adiabatic and non-adiabatic gates and does not impose any speed restriction on the control. We argue how the presence of control constraints can give rise to preferential phase robustness. Our result reconciles the apparent contradictory claims in the current literature regarding the robustness of geometric gates. Since geometric gates are not inherently more robust than dynamical gates, then the use of geometric quantum computing becomes a question of experimental convenience.

Chapter 6

Efficient reverse engineering of one-qubit filter functions with dynamical invariants

We derive an integral expression for the filter-transfer function of an arbitrary one-qubit gate through the use of dynamical invariant theory and Hamiltonian reverse engineering. We use this result to define a cost functional which can be efficiently optimized to produce one-qubit control pulses that are robust against specified frequency bands of the noise power spectral density. We demonstrate the utility of our result by generating optimal control pulses that are designed to suppress broadband detuning and pulse amplitude noise. We report an order of magnitude improvement in gate fidelity in comparison with known composite pulse sequences. More broadly, we also use the same theoretical framework to prove the robustness of nonadiabatic geometric quantum gates under specific error models and control constraints. This work was based on the paper arXiv:2204.08457 [189].

6.1 Introduction

Accurate manipulation of noisy quantum systems is an important problem in optimal control theory with potential applications in the field of chemical reaction control [190–192], quantum sensing [193, 194], and quantum information processing (QIP) [1] to name a few. In QIP, a typical strategy for suppressing errors due to noise is to use dynamical decoupling [195–200] and composite pulse sequences [28, 36, 37, 72, 75, 201]. These techniques are designed to perturbatively suppress noise with correlation time scales that are much longer than the target evolution time (quasistatic noise). In many instances, however, quantum devices also suffer from non-static noise that fluctuates on the order of the evolution time or faster [40, 41, 200, 202]. Composite pulses have limited efficacy in such cases [203] and can even be detrimental to the quality of the generated quantum gate [204].

An alternative solution to these control problems is to use pulse shaping techniques [62, 165, 205–210]. The main idea of this approach is to find, either analytically or numerically, an appropriate set of time-dependent control Hamiltonian parameters that produces a desired evolution. Since the time-dependent Schrödinger equation (TDSE) is generally not analytically tractable, analytical solutions are typically limited to simple pulse shapes [211] or in restricted settings (e.g., for static error [208, 210] or state transfer protocols [206]). Numerical solutions offer much more flexibility in the control landscape. When combined with the formalism of filter functions [47], which characterizes the sensitivity of a control protocol to the power spectral density of the noise, it is possible to generate quantum gates that are robust against a specified spectral region of noise. Specifically, robust quantum gates are obtained by minimizing the overlap between the control's filter function and the noise power spectral density (PSD) in frequency space. This may be used, along with any control field constraints, to define a cost functional to be minimized using, for example, gradient-based methods. Optimization algorithms that are designed for deep learning and are implemented in platforms such as TensorFlow [212] or Julia's Flux package [213] are especially well-suited for these tasks owing to their built-in automatic differentiation capability. The power and flexibility offered by deep neural networks for solving quantum control problems has been demonstrated in a variety of recent works [214–219]. However, filter function engineering typically involves solving the TDSE for the time evolution operator. It is possible to circumvent this, for example, using Hamiltonian reverse engineering based on the theory of dynamical invariants [165]. Thus, it is possible to further reduce the computational workload of the optimization framework by reparameterizing the cost functional in terms of dynamical invariant parameters.

In this work, we use dynamical invariant theory and Hamiltonian reverse engineering to derive an integral expression for the filter function of an arbitrary one-qubit gate and explore its theoretical and practical applications. Our work is structured as follows. We begin Section 6.2 by reviewing the theory of dynamical invariants. We follow this up with a derivation of the one-qubit filter function for an arbitrary noise model in terms of the dynamical invariant parameters. We explore the practical applications of our results in Section 6.3 by numerically searching for optimal control solutions using deep neural networks. Specifically, we consider noise models with a $1/f$ noise spectrum [39] which is prevalent in solid-state qubits [40–45]. In addition, we discuss in Section 6.4 some theoretical implications of our result by proving the robustness of geometric quantum gates against certain noise models under a strict only two-axis driving constraint. We then conclude and summarize our findings in Section 6.5.

6.2 Dynamical invariants

We consider as our starting point a general one-qubit control Hamiltonian with three-axis driving,

$$H_c(t) = \frac{1}{2}\begin{bmatrix} \Delta(t) & \Omega(t)\mathrm{e}^{-i\varphi(t)} \\ \Omega(t)\mathrm{e}^{i\varphi(t)} & -\Delta(t) \end{bmatrix}. \tag{6.1}$$

This particular form is relevant in systems such as superconducting qubits [57], quantum dot spin qubits [178], and NMR qubits [179] to name a few, corresponding to the rotating wave approximation for a two-level systm that is driven by an oscillating field with amplitude Ω at a carrier frequency detuned from resonance by Δ, and with phase φ. The solution to the time-dependent Schrödinger equation with this Hamiltonian is not analytically tractable in general. It is possible, however, to use the theory of dynamical invariants to reformulate this problem so as to specify a resulting unitary evolution and then analytically calculate a time-dependent Hamiltonian that would produce it [165]. A dynamical invariant $I(t)$ is a solution to the Liouville-von Neumann equation [162]

$$i\frac{\partial I(t)}{\partial t} - [H_c(t), I(t)] = 0. \tag{6.2}$$

The eigenvectors $|\phi_n(t)\rangle$ of $I(t)$ are related to the solutions of the Schrödinger equation by a global phase factor: $|\psi_n(t)\rangle = \mathrm{e}^{i\alpha_n(t)} |\phi_n(t)\rangle$, where $\alpha_n(t)$ are the Lewis-Riesenfeld phases given by [172]

$$\alpha_n(t) = \int_0^t \left\langle \phi_n(s) \middle| i\frac{\partial}{\partial s} - H_c(s) \middle| \phi_n(s) \right\rangle \mathrm{d}s. \tag{6.3}$$

Within this framework, the time evolution operator $U_c(t)$ can be expressed as

$$U_c(t) = \sum_{n=\pm} \mathrm{e}^{i\alpha_n(t)} |\phi_n(t)\rangle \langle \phi_n(0)| \,. \tag{6.4}$$

Thus, the theory of dynamical invariants effectively transforms the problem of solving the time-dependent Schrödinger equation to finding an appropriate $I(t)$ that satisfies Equation (6.2). As a consequence, we are free to choose a parameterization for $U_c(t)$ by choosing the $|\phi_n(t)\rangle$ appropriately. Suppose that we choose

$$|\phi_+(t)\rangle = \cos\left(\frac{\gamma(t)}{2}\right) \mathrm{e}^{-i\beta(t)} |0\rangle + \sin\left(\frac{\gamma(t)}{2}\right) |1\rangle \,, \tag{6.5}$$

$$|\phi_-(t)\rangle = \sin\left(\frac{\gamma(t)}{2}\right) |0\rangle - \cos\left(\frac{\gamma(t)}{2}\right) \mathrm{e}^{i\beta(t)} |1\rangle \,, \tag{6.6}$$

where $I(t)\,|\phi_n(t)\rangle = \pm\Omega_0/2\,|\phi_n(t)\rangle$ and Ω_0 is an arbitrary constant with units of frequency. This allows us to express $I(t)$ in a form similar to Equation (6.1)

$$I(t) = \frac{\Omega_0}{2} \begin{pmatrix} \cos(\gamma) & \sin(\gamma)\mathrm{e}^{-i\beta} \\ \sin(\gamma)\mathrm{e}^{i\beta} & -\cos(\gamma) \end{pmatrix}. \tag{6.7}$$

If we require Equations (6.1) and (6.7) to satisfy Equation (6.2), we are left with two coupled auxiliary equations [165]

$$\dot{\gamma} = -\Omega \sin(\beta - \varphi) \tag{6.8}$$

$$\Delta - \dot{\beta} = \Omega \cot(\gamma) \cos(\beta - \varphi), \tag{6.9}$$

which, along with the appropriate boundary conditions, can be used to determine the control parameters $\Omega(t)$, $\Delta(t)$, and $\varphi(t)$ that targets a desired $U_c(t)$. This choice of parametrization allows us to write $U_c(t)$ strictly in terms of the dynamical invariant parameters and the Lewis-Riesenfeld phase:

$$U_c(t) = \mathrm{e}^{-i\frac{\beta(t)}{2}\sigma_Z}\mathrm{e}^{-i\frac{\gamma(t)}{2}\sigma_Y}\mathrm{e}^{i\frac{\zeta(t)-\zeta(0)}{2}\sigma_Z}\mathrm{e}^{i\frac{\gamma(0)}{2}\sigma_Y}\mathrm{e}^{i\frac{\beta(0)}{2}\sigma_Z}, \tag{6.10}$$

where $\alpha = \alpha_+ = -\alpha_-$ and we introduce a new parameter

$$\begin{aligned} \zeta(t) &= 2\alpha(t) - \beta(t) \\ &= -\beta(0) + \int_0^t \frac{\dot{\gamma}\cot(\beta-\varphi)}{\sin\gamma}\,\mathrm{d}t' \end{aligned} \tag{6.11}$$

The auxiliary equations provide a family of control solutions that allow us to reverse engineer a desired quantum gate. Since the gate only depends on the boundary values of the dynamical invariant parameters, there are infinitely many ways to generate the gate. It is desirable to use this freedom in the control Hamiltonian such that the resulting evolution is also robust against noise. To this end, filter functions provide a convenient method of quantifying the gate fidelity's susceptibility to noise with respect to its spectral properties [47]. The total one-qubit Hamiltonian in the presence of noise can be written as

$$H(t) = H_c(t) + H_e(t), \tag{6.12}$$

where $H_c(t)$ is the ideal deterministic control Hamiltonian and $H_e(t)$ is the stochastic error Hamiltonian. More explicitly, $H_e(t)$ can generally be expressed as

$$H_e(t) = \sum_q \sum_{i=1}^{3} \delta_q(t)\chi_{q,i}(t)\sigma_i, \tag{6.13}$$

where q indexes a set of uncorrelated stochastic variables $\delta_q(t)$, $\chi_{q,i}(t)$ contains the sensitivity of the control parameters (which generally can be a function of the parameters themselves) to $\delta_q(t)$, and σ_i are Pauli operators. For sufficiently weak noise, the average gate infidelity $\langle\mathcal{I}\rangle$ of the noisy evolution $U(t)$, which satisfies $i\dot{U}(t) = H(t)U(t)$ where $U(0) = \mathbb{1}$, can be compactly expressed as (see Appendix F)

$$\langle\mathcal{I}\rangle \approx \frac{1}{2\pi}\sum_q \int_{-\infty}^{\infty} S_q(\omega)F_q(\omega)\,\mathrm{d}\omega, \tag{6.14}$$

where $S_q(\omega)$ denotes the noise PSD for the stochastic variable $\delta_q(t)$ and $F_q(\omega)$ is the corresponding filter function which can be calculated using the following equations:

$$F_q(\omega) = \sum_k \left|R_{q,k}(\omega)\right|^2, \tag{6.15}$$

$$R_{q,k}(\omega) = \sum_i \int_0^T \chi_{q,i}(t)R_{ik}(t)\mathrm{e}^{i\omega t}\,\mathrm{d}t, \tag{6.16}$$

$$R_{ik}(t) = \frac{1}{2}\,\mathrm{tr}\left(U_c^\dagger(t)\sigma_i U_c(t)\sigma_k\right), \tag{6.17}$$

where T is the gate time.

Combining Equations (6.4) and (6.17) allows us to express Equation (6.16) as

$$R_{q,k}(\omega) = \frac{1}{2}\sum_{i,n,n'} \langle\phi_n(0)|\sigma_k|\phi_{n'}(0)\rangle \int_0^T \mathrm{e}^{i(\alpha_n(t)-\alpha_{n'}(t)+\omega t)}\chi_{q,i}(t)\,\langle\phi_{n'}(t)|\sigma_i|\phi_n(t)\rangle\,\mathrm{d}t. \tag{6.18}$$

Thus, the filter function corresponding to $\delta_q(t)$ is given by

$$\begin{aligned}
F_q(\omega) &= \sum_k R_{q,k}(\omega)R^*_{q,k}(\omega)\\
&= \frac{1}{4}\int_0^T\int_0^T \sum_{\substack{i,j,k,\\ n,m,n',m'}} \langle\phi_n(0)|\sigma_k|\phi_{n'}(0)\rangle\,\langle\phi_m(0)|\sigma_k|\phi_{m'}(0)\rangle\\
&\times \mathrm{e}^{i(\alpha_n(t_1)-\alpha_{n'}(t_1)+\omega t_1)}\chi_{q,i}(t_1)\,\langle\phi_{n'}(t_1)|\sigma_i|\phi_n(t_1)\rangle\,\mathrm{d}t_1\\
&\times \mathrm{e}^{i(\alpha_m(t_2)-\alpha_{m'}(t_2)-\omega t_2)}\chi_{q,j}(t_2)\,\langle\phi_{m'}(t_2)|\sigma_j|\phi_m(t_2)\rangle\,\mathrm{d}t_2.
\end{aligned} \tag{6.19}$$

For a given n, n', m, and m', the k-dependent factors of this sum yields

$$\sum_k \langle\phi_n(0)|\sigma_k|\phi_{n'}(0)\rangle \langle\phi_m(0)|\sigma_k|\phi_{m'}(0)\rangle$$
$$= \begin{cases} 1 & \text{if } \{n,n',m,m'\} = \{\pm,\pm,\pm,\pm\} \\ -1 & \text{if } \{n,n',m,m'\} = \{\pm,\pm,\mp,\mp\} \\ 2 & \text{if } \{n,n',m,m'\} = \{\pm,\mp,\mp,\pm\} \\ 0 & \text{otherwise} \end{cases}. \tag{6.20}$$

We can use Equations (6.5), (6.6), (6.20) as well as the fact that $\langle\phi_\pm(t)|\sigma_k|\phi_\pm(t)\rangle = -\langle\phi_\mp(t)|\sigma_k|\phi_\mp(t)\rangle$ to simplify Equation (6.19) into

$$\begin{aligned} F_q(\omega) &= \sum_{i,j} \left(\int_0^T \langle\phi_+(t)|\sigma_i|\phi_+(t)\rangle\, \chi_{q,i}(t) e^{\imath\omega t}\,\mathrm{d}t\right) \left(\int_0^T \langle\phi_+(t)|\sigma_j|\phi_+(t)\rangle\, \chi_{q,j}(t) e^{-\imath\omega t}\,\mathrm{d}t\right) \\ &+\frac{1}{2}\left(\int_0^T \langle\phi_-(t)|\sigma_i|\phi_+(t)\rangle\, \chi_{q,i}(t) e^{\imath 2\alpha(t)+\imath\omega t}\,\mathrm{d}t\right)\left(\int_0^T \langle\phi_+(t)|\sigma_j|\phi_-(t)\rangle\, \chi_{q,j}(t) e^{-\imath 2\alpha(t)-\imath\omega t}\mathrm{d}t\right) \\ &+\frac{1}{2}\left(\int_0^T \langle\phi_+(t)|\sigma_i|\phi_-(t)\rangle\, \chi_{q,i}(t) e^{-\imath 2\alpha(t)+\imath\omega t}\,\mathrm{d}t\right)\left(\int_0^T \langle\phi_-(t)|\sigma_j|\phi_+(t)\rangle\, \chi_{q,j}(t) e^{\imath 2\alpha(t)-\imath\omega t}\,\mathrm{d}t\right). \end{aligned} \tag{6.21}$$

Finally, substituting in Equations (6.5) and (6.6) allows us to compactly write Equation (6.2) in the following vectorized expression:

$$F_q(\omega) = \left\| \int_0^T \Lambda(t) \begin{bmatrix} \chi_{q,X}(t) \\ \chi_{q,Y}(t) \\ \chi_{q,Z}(t) \end{bmatrix} e^{i\omega t}\,\mathrm{d}t \right\|^2, \tag{6.22}$$

where the entries of the matrix Λ are given by

$$\Lambda = \begin{pmatrix} \cos\beta\sin\gamma & \sin\beta\sin\gamma & \cos\gamma \\ -\cos\beta\cos\gamma\cos\zeta - \sin\beta\sin\zeta & -\sin\beta\cos\gamma\cos\zeta + \cos\beta\sin\zeta & \sin\gamma\cos\zeta \\ -\cos\beta\cos\gamma\sin\zeta + \sin\beta\cos\zeta & -\sin\beta\cos\gamma\sin\zeta - \cos\beta\cos\zeta & \sin\gamma\sin\zeta \end{pmatrix}. \tag{6.23}$$

This is our main result and we show in the following sections some examples of its utility. Before we proceed, we comment on the form of Equation (6.22). First, although the similarity between Equations (6.15)-(6.16) and (6.22) might seem to suggest that $R(t)$ and $\Lambda(t)$ are identical and we have not really simplified anything, in fact what we have done is to note that the dependence of $R(t)$ on the value of the dynamical invariant parameters evaluated at $t=0$ does not affect the filter function value, and $\Lambda(t)$ does not carry that extraneous dependence. Second, certain error models admit an alternative interpretation for Equation (6.22). For example, suppose we consider the dephasing and over-rotation noise models. The former can be induced by an additive shift to the qubit detuning, $\Delta(t) \to \Delta(t) + \delta_\Delta(t)$, and the latter can be induced by a multiplicative shift in the pulse amplitude, $\Omega(t) \to \Omega(t)\,(1+\delta_\Omega(t))$. The corresponding error sensitivities are $\boldsymbol{\chi}_\Delta(t) = \frac{1}{2}\,[0,0,1]^\intercal$ and $\boldsymbol{\chi}_\Omega(t) = \frac{1}{2}\,[\Omega\cos\varphi, \Omega\sin\varphi, 0]^\intercal$. Substituting these expressions onto Equation (6.22) yields the following filter functions

$$F_\Delta(\omega) = \left\| \int_0^T \frac{1}{2} \begin{bmatrix} \cos\gamma \\ \sin\gamma\cos\zeta \\ \sin\gamma\sin\zeta \end{bmatrix} e^{i\omega t}\,\mathrm{d}t \right\|^2, \tag{6.24}$$

$$F_\Omega(\omega) = \left\| \int_0^T \frac{1}{2} \begin{bmatrix} \dot\zeta\sin^2\gamma \\ \dot\zeta\sin\gamma\cos\gamma\cos\zeta + \dot\gamma\sin\zeta \\ \dot\zeta\sin\gamma\cos\gamma\sin\zeta - \dot\gamma\cos\zeta \end{bmatrix} e^{i\omega t}\,\mathrm{d}t \right\|^2. \tag{6.25}$$

Up to a scalar factor, the detuning filter function in Equation (6.24) can be reinterpreted as a position vector with constant speed[1]

$$\dot{\vec{r}} = [\cos\gamma, -\sin\gamma\cos\zeta, -\sin\gamma\sin\zeta]\,. \tag{6.26}$$

If robustness at a certain noise frequency is defined by a vanishing filter function

[1] The sign difference in comparison with Equation (6.24) is a consequence of our choice of parameterization for the dynamical invariant eigenvectors and is irrelevant since only the magnitude of $\dot{\vec{r}}$ matters.

value, robustness against static detuning noise (i.e., at $\omega = 0$) is equivalent to having the position vector trace a closed three-dimensional curve whose curvature is given by $\Omega(t)$. Such a geometric interpretation has been noted previously in the literature [157, 220–225].

A similar observation can be made for the pulse amplitude filter function. Note that the vector in the integrand of Equation (6.25) is equivalent to $\dot{\vec{r}} \times \ddot{\vec{r}}$. This can be rewritten as $\Omega\,\vec{b}$ [226], where $\vec{b}$ is the binormal vector corresponding to $\vec{r}$ and we have used the fact that the curvature $\kappa = \Omega$. Therefore, constructing a quantum gate that is simultaneously robust against static detuning and pulse amplitude noise is mathematically equivalent to finding a closed three-dimensional curve such that $\int_0^T \Omega(t)\vec{b}(t)\,\mathrm{d}t = \vec{0}$. As far as we know, this has not been noted before.

6.3 Broadband Noise Optimization

We demonstrated in Section 6.2 that it is possible through Hamiltonian reverse engineering to analytically calculate the filter function of an arbitrary one-qubit gate in terms of the dynamical invariant parameters $\beta(t)$, $\gamma(t)$, and $\zeta(t)$ as well as the sensitivity $\chi_{q,i}(t)$. One immediate implication of this result is the possibility of filter function engineering which can be used for error suppression [214–217] or quantum sensing [227]. In the context of error suppression, we can use Equation (6.22) to define a cost functional which can be minimized in spectral regions where the noise PSD is dominant. This approach allows us to target any robust one-qubit gate provided that we can find an appropriate $\gamma(t)$ and $\beta(t)$. Furthermore, this is different from previous filter function engineering results since calculating the evolution operator is no longer necessary, which helps to reduce the computational workload of the optimization framework.

We consider again as an example the case where our system is subject to detuning and pulse amplitude noise. Note that both Equations (6.24) and (6.25) depend only on γ and ζ. This means that β is a free parameter up to the boundary conditions imposed by the reverse engineering process. This extra degree of freedom can be used to impose control restrictions such as strict two-axis control. Combining Equations (6.8), (6.9), and (6.11) provides us with the reverse engineered Hamiltonian parameters in terms of the dynamical invariant parameters:

$$\Omega = \sqrt{\dot{\gamma}^2 + \dot{\zeta}^2 \sin^2 \gamma} \tag{6.27}$$

$$\varphi = \beta - \arctan \frac{\dot{\gamma}}{\dot{\zeta} \sin \gamma} \tag{6.28}$$

$$\Delta = \dot{\beta} - \dot{\zeta} \cos \gamma. \tag{6.29}$$

We can set $\Delta = 0$ by solving the differential equation $\dot{\beta} = \dot{\zeta} \cos \gamma$ for β with the boundary condition $\beta(0) = -\zeta(0)$. Thus, all properties of the output gate is determined by γ and ζ.

Restricting β in this manner does not necessarily diminish our ability to target arbitrary one-qubit gates. In practice, a finite set of quantum gates are used to target arbitrary operations. Although we can engineer γ and ζ to target gates directly, it is worth pointing out that many qubit qubit implementations have access to virtual Z (VZ) gates [64, 228–230]. These zero-duration gates are essentially perfect and implemented through abrupt changes to the reference phase. It can be shown that any one-qubit gate can be decomposed into the product of Z gates and two $X_{\frac{\pi}{2}}$ [64]: $Z_{\theta_1} X_{\frac{\pi}{2}} Z_{\theta_2} X_{\frac{\pi}{2}} Z_{\theta_3}$. We can rewrite the engineered evolution operator in Equation (6.10) as

$$\begin{aligned} U_c(t) &= Z_{\beta(t)} Y_{\gamma(t)} Z_{\zeta(0)-\zeta(t)} Y_{-\gamma(0)} Z_{-\beta(0)} \\ &= Z_{\psi_1} X_{\theta} Z_{\psi_2}, \end{aligned} \tag{6.30}$$

where

$$\cos(\theta) = \cos(\zeta(0) - \zeta(t)) \sin(\gamma(t)) \sin(\gamma(0)) + \cos(\gamma(t)) \cos(\gamma(0)), \tag{6.31}$$

and ψ_1 and ψ_2 are angles that depend on the target gate. By setting $\theta = \frac{\pi}{2}$, we can replace all $X_{\frac{\pi}{2}}$ in the gate decomposition with $U_c(T)$ and combining all neighboring Z gates. Since Z gates may be executed virtually, we only need one physical gate, $U_c(T)$ with $\theta = \frac{\pi}{2}$, to produce any one-qubit operation.

Our goal is to minimize the following cost functional:

$$\begin{aligned} \text{cost} = &\int_{-\infty}^{\infty} F_\Delta(\omega) S_\Delta(\omega)\, \mathrm{d}\omega + \int_{-\infty}^{\infty} F_\Omega(\omega) S_\Omega(\omega)\, \mathrm{d}\omega \\ &+ |\cos(\zeta(0) - \zeta(T)) \sin\gamma(T) \sin\gamma(0) + \cos\gamma(T) \cos\gamma(0)| \\ &\left|\frac{\Omega(0)}{\Omega_{\max}}\right| + \left|\frac{\Omega(T)}{\Omega_{\max}}\right| + \sum_i \max\left(0, \frac{\Omega(t_i)}{\Omega_{\max}} - 1\right). \end{aligned} \tag{6.32}$$

The first two terms correspond to the infidelity integrals for detuning and amplitude noise with noise PSD S_Δ and S_Ω, respectively. The third term is the constraint that targets $\theta = \frac{\pi}{2}$. The fourth and fifth term sets the boundary value of the pulse amplitude to zero[2]. Finally, the sixth term imposes a maximum value $\Omega_{\max}$ on Ω

[2]These constraints are not necessary but they help with the overall experimental feasibility of the pulses we produce.

by discretizing the interval $[0, T]$ and evaluating Ω at each time value. The cost penalizes any point where $\Omega(t_i) > \Omega_{\max}$ through the function

$$\max(0, x) = \begin{cases} 0 & x \leq 0 \\ x & x > 0 \end{cases}. \tag{6.33}$$

We demonstrate the flexibility of our approach by considering two examples. We first consider a case where the goal is to produce a gate that acts as a stopband filter against $1/f$ detuning and pulse amplitude noise. We then consider a case where the goal is to produce a gate that is optimal in the presence of $1/f$ pulse amplitude noise and a static detuning noise. To this end, we employ deep neural networks [231, 232] as our optimization framework. The power of neural networks originate from their ability to represent complex ideas as a hierarchy of simpler concepts. This allows them to efficiently identify key abstract properties of a problem, which is highly coveted in tasks such as pattern recognition [233]. It has also been proven that neural networks with sufficient neurons and layers can act as a universal function approximator [234, 235]. This is ideal for our purpose since it eliminates the nontrivial task of finding suitably paramemeterized ansatz function to optimize over that will yield convergent solutions. Furthermore, machine learning frameworks tend to have built-in automatic differentiation capabilities which can be utilized for gradient-based optimization.

In particular, we use a feedforward neural network (sometimes referred to as multilayer perceptron) which is constructed using layers of interconnected computational units called neurons such that information travels only in one direction;

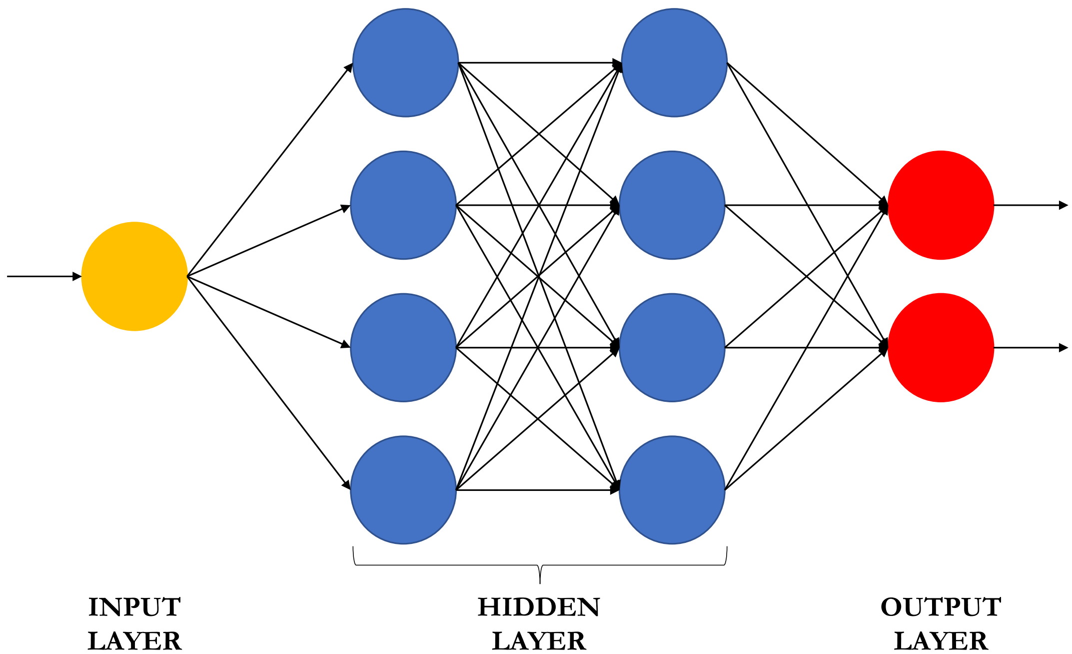


Figure 6.1: A schematic diagram of a feedforward deep neural network with one input neuron, two output neurons, and two hidden layers with four neurons each. A network is deep if it has at least two hidden layers. As information flows from the input layer, each subsequent layer nonlinearly transforms incoming information and returns a value. The goal is to train the neural network so that the final output optimizes the cost. In our case, we would like to train a neural network to take time as input and return the optimized dynamical invariant parameters γ and ζ.

starting with an input layer, then a series of hidden layers, and finally onto an output layer. A schematic diagram of a feedforward neural network is shown in Fig. 6.1. A neural network is deep if it has at least two hidden layers. Each adjacent layers act as a function that takes a vector input and produces a vector output using the following model

$$\boldsymbol{x}_{i+1} = \sigma\left(W_i \boldsymbol{x}_i + \boldsymbol{b}_i\right), \tag{6.34}$$

where $\boldsymbol{x}_i$ is the input in the i^{th} layer, W_i is a matrix that describes the neural connections between the i^{th} and $(i+1)^{\text{th}}$ layer, $\boldsymbol{b}_i$ is a bias vector, and $\sigma(\cdot)$ is a nonlinear activation function such as $\max(0, \cdot)$ or $\tanh(\cdot)$. Our goal is to train the neural network using machine learning algorithms to return the optimized dynamical invariant parameters γ and ζ on the output layer by feeding in time on the input layer. For our optimization we use a feedforward deep neural network with one input neuron, two hidden layers with 16 neurons each and a tanh activation function, and two output neurons for a total of 338 parameters[3].

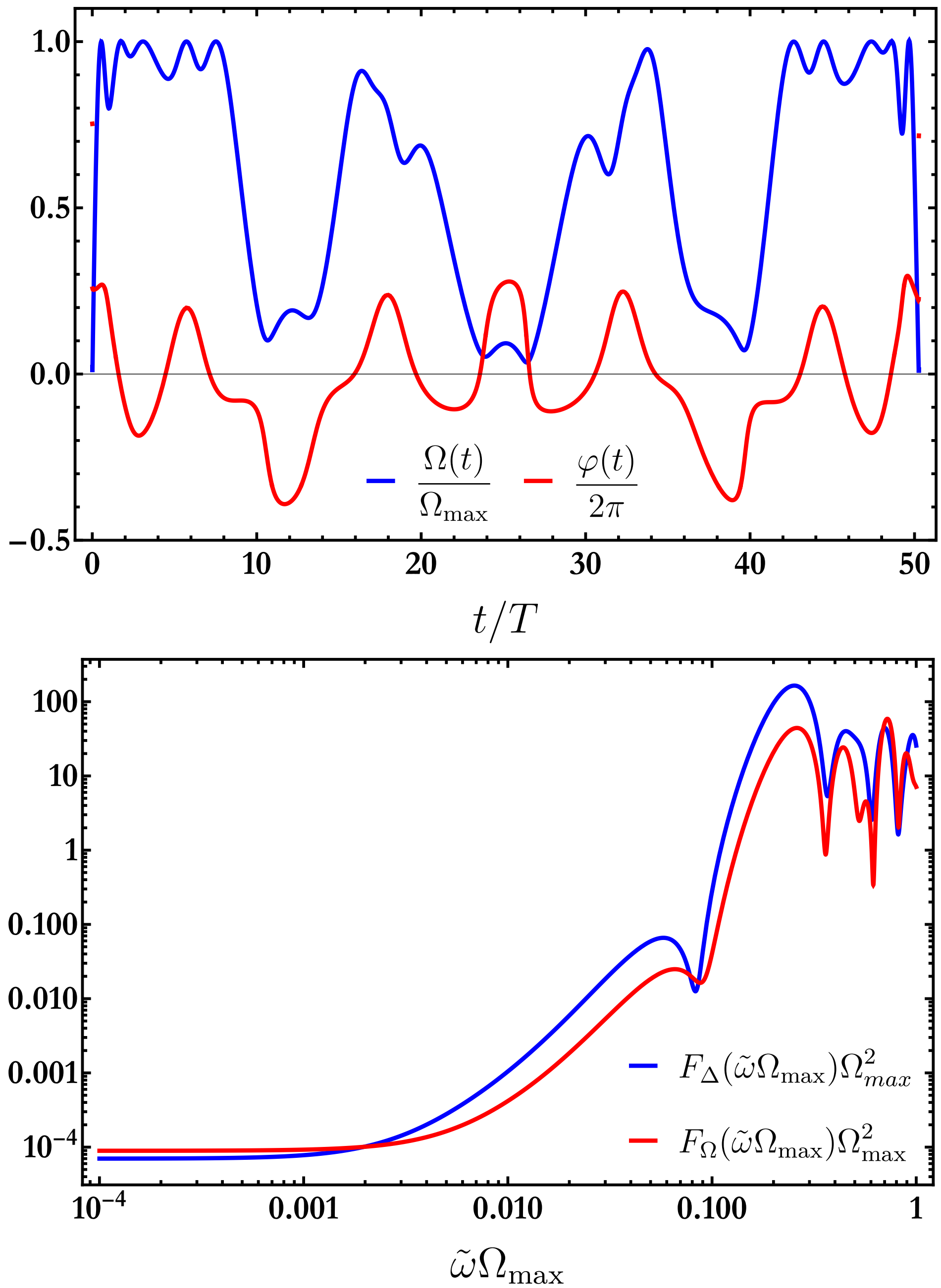


Figure 6.2: A plot of the optimized Hamiltonian (TOP) and filter function (BOTTOM) for the case of simultaneous $1/f$ detuning and pulse amplitude noise over a finite frequency range.

6.3.1 $1/f$ stopband filter for both detuning and pulse amplitude noise

For our first example, we consider identical noise PSD for detuning and amplitude noise:

$$S_\Delta(\omega) = S_\Omega(\omega) = \begin{cases} \frac{A}{\omega} & \omega_0 \leq |\omega| \leq \omega_c \\ \\ 0 & \text{otherwise} \end{cases}, \tag{6.35}$$

where $[\omega_0, \omega_c]$ defines the frequency stopband in which we wish to suppress noise. We set $\omega_0 = 10^{-9}\Omega_{\max}$, $\omega_c = 10^{-1}\Omega_{\max}$, and $T = 16\pi/\Omega_{\max}$. We present in Fig. 6.2 a plot of the optimized control fields and filter functions. The details of our numerical optimization scheme is provided in Appendix G.

We see from Fig. 6.2 that the control pulse we produced satisfies the imposed constraints. We compare the total infidelity of our optimized pulse with that of known pulse sequences in the literature that address either detuning noise, pulse amplitude noise, or both. We present in Table 6.1 a summary of these comparisons. We find that our broadband optimized pulse yields an infidelity that is at least an order of magnitude lower than than any other pulse sequences. Specifically, the minimum improvement is roughly a factor of 14 which is a comparison with the concatenated CORPSE [29, 236, 237] and BB1 [238] pulse sequence (CinBB) [239]. CinBB is able to address static detuning and pulse amplitude noise simultaneously.

[3] In feedforward neural networks, each neural connection adds one parameter. Furthermore, with the exception of the input neurons, each neuron contains an additional bias parameter. Thus, if we have a 1-3-2 network (one input neuron, one hidden layer with three neurons, and two output neurons), we have $(1*3+3)+(3*2+2) = 14$ free parameters to optimize. In our work, we used a 1-16-16-2 network which has $(1*16+16)+(16*16+16)+(16*2+2) = 338$ free parameters.

We attribute the improvement in our result to the fact that pulse sequences are generally designed to suppress static noise. Although they can be used to address noise in the quasitatic regime, their ability to suppress noise that fluctuate on the order of $\Omega_{\max}$ severely limited. However, it is worth noting that our pulse's performance comes at the cost of increased noise sensitivity in frequency regions beyond the indicated stopband.

6.3.2 Static detuning and $1/f$ pulse amplitude noise

For our second example, we consider the case where we have a static detuning noise as well as a $1/f$ pulse amplitude noise:

$$S_\Delta(\omega) = 10A\delta(\omega), \tag{6.36}$$

$$S_\Omega(\omega) = \begin{cases} 0 & 0 \leq |\omega| \leq \omega_0 \\ \frac{A}{\omega} & \omega_0 \leq |\omega| \leq \omega_c \\ \frac{A\omega_c}{\omega^2} & \omega_c \leq |\omega| \end{cases}, \tag{6.37}$$

where we have assumed an order of magnitude difference in the detuning and pulse amplitude noise strength. Here we set $\omega_0 = 10^{-9}\Omega_{\max}$, $\omega_c = 10^{-1}\Omega_{\max}$, and $T = 5\pi/\Omega_{\max}$. We present in Fig. 6.3 a plot of the optimized control fields and filter functions. We again compare our optimized pulse with known pulse sequences and the results are summarized in Table 6.1.

Unlike the previous case, we only see a minimum improvement in infidelity by a factor of 7. This is primarily attributed to penalizing the value of the filter function in regions where $\omega_c \leq |\omega|$. In the previous case, the improvement is due

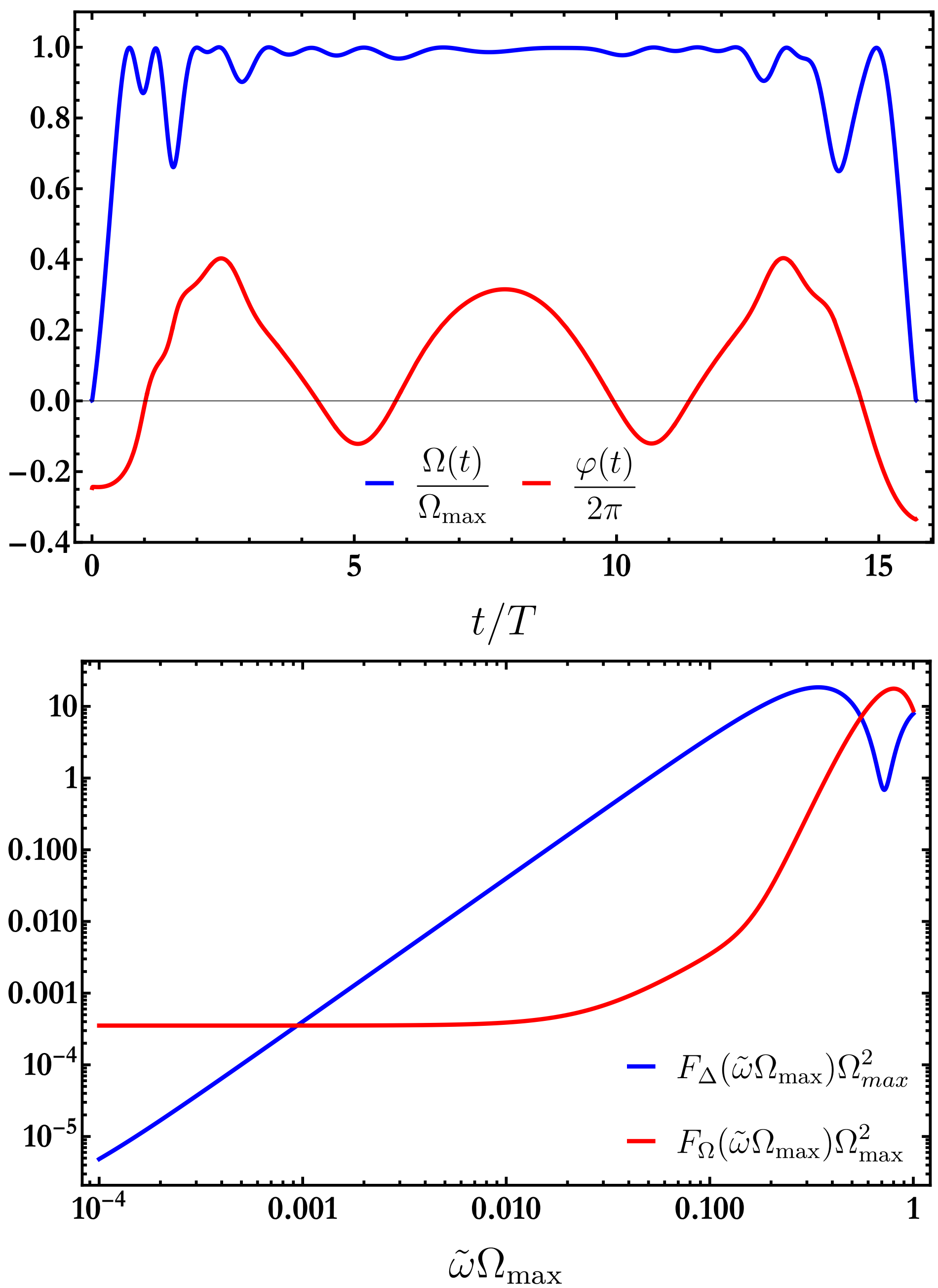


Figure 6.3: A plot of the optimized Hamiltonian (TOP) and filter function (BOTTOM) for the case of static detuning noise and $1/f$ pulse amplitude noise. Unlike the previous example, the $1/f$ spectrum here has a $1/f^2$ tail which penalizes large filter function values in the $\omega_c \leq |\omega|$ region.

Table 6.1: A comparison of infidelities between our deep neural network output and known composite pulse sequences. The ratio $\mathcal{I}_i/\mathcal{I}_{\text{DNN}}$ compares the infidelity of naive and composite pulses targeting a $X_{\frac{\pi}{2}}$ gate against our optimized gate. The subscript A/B indicates whether the case of Section 6.3.1 or that of Section 6.3.2 is in consideration. We also indicate which pulses are robust against static detuning and/or pulse amplitude noise. We report a substantial decrease in infidelity in all cases we considered.

Pulse	$\mathcal{I}_{\text{A}}/\mathcal{I}_{\text{DNN}}$	$\mathcal{I}_{\text{B}}/\mathcal{I}_{\text{DNN}}$	Robust to δ_Δ?	Robust to δ_Ω?
Naive (Square)	223	14	No	No
Short CORPSE	1110	68	Yes	No
BB1	110	9	No	Yes
CinBB	14	7	Yes	Yes
CinSK	57	15	Yes	Yes

to the fact that filter function values outside the stopgap do not contribute to the infidelity. This is no longer true in this case due to the presence of a $1/f^2$ tail which penalizes large filter function values for frequency values greater than ω_c.

6.4 Robustness of geometric phases

We can also apply our result in Section 6.2 to explore the robustness properties of geometric quantum gates. Geometric gates are quantum gates with a trivial dynamical phase which means they rely on the geometric phase to produce unitary dynamics. Geometric gates are of practical interest since they are conjectured to be ideal for robust quantum computation owing to the global nature of the accumulated phase. The validity of this conjecture is the subject of many studies with many showing support for the conjecture. However, there are also studies that report situations in which geometric gate are not intrinsically more robust than dynamical gates [153–157] and, in certain scenarios, their sensitivity to noise deteriorates [140, 158–161]. We showed in Chapter 5 that the noise sensitivity of geometric and dynamical gates are generically equal. However, when control constraints are present (e.g. strict 2-axis or piecewise constant control), it is possible for a particular phase type to become preferable and naturally robust to some noise.

Here we demonstrate preferential phase robustness in nonadiabatic Abelian geometric gates as a consequence of control constraints. To clarify, we say that a quantum gate is robust against the noise process q at a particular frequency ω if $F_q(\omega) = 0$. Our theoretical framework is ideal for this type of analysis because the notion of geometric and dynamical phases are naturally developed in the theory of dynamical invariants. Using the Lewis-Riesenfeld phase in Equation (5.2), the eigenvectors in Equations (6.5) and (6.6), as well as the auxiliary equations in Equations (6.8) and (6.9), we can define the geometric phase, $\alpha_{n,g}$, and dynamical phase, $\alpha_{n,d}$, accumulated by the eigenvector $|\phi_n\rangle$ during the evolution as

$$\begin{aligned}\alpha_{n,g}(T) &= \int_0^T \left\langle \phi_\pm(t) \middle| i\frac{\partial}{\partial t} \middle| \phi_\pm(t) \right\rangle \mathrm{d}t \\ &= \pm\alpha(T) \mp \int_0^T \frac{\dot{\zeta} - \dot{\beta}\cos\gamma}{2}\,\mathrm{d}t, \end{aligned} \tag{6.38}$$

$$\begin{aligned}\alpha_{n,d}(T) &= -\int_0^T \langle \phi_\pm(t)|H(t)|\phi_\pm(t)\rangle \,\mathrm{d}t \\ &= \pm\int_0^T \frac{\dot{\zeta} - \dot{\beta}\cos\gamma}{2}\,\mathrm{d}t. \end{aligned} \tag{6.39}$$

Suppose we consider the special case of a constant detuning Δ, which is a fairly common constraint in works considering geometric gates [140, 180, 182, 240].

We prove the following theorem for that special case by analyzing the filter function expressions that we derived:

Theorem. *Consider the control Hamiltonian in Equation (6.1) <u>under the constraint that Δ is constant</u>. Any gate that is robust to static multiplicative amplitude noise (δ_Ω) as well as static additive or multiplicative detuning noise (δ_Δ) is necessarily geometric.*

Proof. Using Equation (6.29), we can rewrite the dynamical phase integral in Equation (6.39) as

$$\alpha_{n,d}(T) = \pm \int_0^T \frac{\dot{\zeta} - \dot{\beta}\cos\gamma}{2}\,\mathrm{d}t \tag{6.40}$$

$$= \pm \int_0^T \frac{-\Delta\cos\gamma + \dot{\zeta}\sin^2\gamma}{2}\,\mathrm{d}t$$

$$= \mp\frac{\Delta}{2}\int_0^T \cos\gamma\,\mathrm{d}t \pm \frac{1}{2}\int_0^T \dot{\zeta}\sin^2\gamma\,\mathrm{d}t. \tag{6.41}$$

We begin by considering the case where there is additive detuning and multiplicative pulse amplitude noise. Imposing simultaneous robustness against these noise sources would require $F_\Delta(0) = F_\Omega(0) = 0$. However, we see in Equations (6.24) and (6.25) that the filter function is strictly nonnegative and the only way to achieve robustness against static noise is if every integral vanishes. Specifically, robustness against static additive detuning noise requires $\int_0^T \cos\gamma\,\mathrm{d}t = 0$, while robustness against static amplitude noise requires $\int_0^T \dot{\zeta}\sin^2\gamma\,\mathrm{d}t = 0$. Notice, however, that these are precisely the integral expressions in Equation (6.41). Thus, simultaneous robustness against static detuning and pulse amplitude error necessarily requires the dynamical phase to vanish, i.e., the gate must be geometric.

Next, we consider the case where there is multiplicative detuning and pulse amplitude noise. The multiplicative detuning filter function can be found using Equation (6.29) and is given by

$$\begin{aligned} F_{\Delta,\times}(\omega) &= \left\| \int_0^T \frac{\Delta}{2} \begin{bmatrix} \cos\gamma \\ \sin\gamma\cos\zeta \\ \sin\gamma\sin\zeta \end{bmatrix} e^{i\omega t}\,\mathrm{d}t \right\|^2 \\ &= \left\| \int_0^T \frac{\dot{\beta} - \dot{\zeta}\cos\gamma}{2} \begin{bmatrix} \cos\gamma \\ \sin\gamma\cos\zeta \\ \sin\gamma\sin\zeta \end{bmatrix} e^{i\omega t}\,\mathrm{d}t \right\|^2 . \end{aligned} \tag{6.42}$$

Robustness to static noise would require $F_{\Delta,\times}(0) = 0$. We focus in particular on the first integral which can be rewritten as

$$\frac{1}{2}\int_0^T \dot{\beta}\cos\gamma - \dot{\zeta} + \dot{\zeta}\sin^2\gamma\,\mathrm{d}t. \tag{6.43}$$

Just like before, we note that imposing robustness against static pulse amplitude noise requires $\int_0^T \dot{\zeta}\sin^2\gamma\,\mathrm{d}t = 0$ which eliminates the last term in expression above. Setting the remaining terms to zero is equivalent to setting Equation (6.40) to zero. Therefore, imposing simultaneous robustness against static multiplicative detuning and pulse amplitude noise necessitates a geometric gate. □

We make the following observations. First, this theorem is consistent with other results in the literature. It was previously noted in Refs. [174, 239] that composite pulse sequences with detuning fixed to zero that are designed to be robust against multiplicative pulse amplitude noise (and are trivially robust against multiplicative detuning noise since $\Delta = 0$) are indeed geometric quantum gates. Second, we note that in that special case of $\Delta = 0$, the first term in Equation (6.41) vanishes regardless of the value of the integral. In other words, if we don't require robustness to pulse amplitude noise, it *is* possible to obtain dynamical gates that are robust to static detuning noise. A well-known example is the CORPSE family of composite pulses which are designed to be robust against additive detuning noise [29, 236, 237]. Third, gates that are robust against static multiplicative pulse amplitude noise are necessarily geometric but the converse isn't true. One specific example of this is the orange-slice geometric gate presented in Ref. [180]. It was shown in Section 5.3 that the pulse amplitude filter function in this particular case does not vanish at $\omega = 0$ despite being a geometric gate. Fourth, we note that the parallel transport condition ($\langle \phi_\pm(t)|H(t)|\phi_\pm(t)\rangle = 0$) is not necessary to achieve a robust geometric gate; the dynamical phase integral simply has to vanish at the gate time. Finally, this theorem is consistent with the results of Chapter 5. It is argued there that in the absence of control constraints, geometric and dynamical gates are generically equivalent when it comes to noise sensitivity, and preferential phase robustness can only emerge in the presence of control constraints. In this case, the constraint is considering a strictly constant Δ. Removing the constraint on Δ turns β into a free parameter. According to Equations (6.38) and (6.39), the geometric and dynamical component of the total phase is directly dependent on our choice of β. Thus, in the absence of constraints, we can freely tune the phase type from dynamical to geometric. Moreover, the filter functions in Equations (6.24) and (6.25) are independent of β. This indicates that noise sensitivity, as quantified by the filter function, is independent of the phase type in the absence of control constraints as was also shown more generally in Chapter 5.

6.5 Conclusions

We make use of dynamical invariant theory in order to analytically reverse engineer a qubit's control Hamiltonian and calculate its corresponding filter function. This allows us to define a cost function strictly in terms of the dynamical invariant parameters which can be optimized to create filter functions with desirable properties. The primary limitation of our theory is its currently limited applicability to two-level systems, with no provision for operations on more than one qubit or correction of population leakage to higher energy levels. (The effects of *virtual* transitions to higher energy levels do not pose a problem, since they can be incorporated in an effective one-qubit Hamiltonian [241].) In those cases a generalized approach such as Ref. [214] is preferable. However, for the specific task of constructing local rotations with robustness against high frequency noise bands, our method is a useful and efficient tool.

We demonstrate the utility of our theory by generating control pulses that are optimized to operate in the presence of broadband noise. One example we considered is creating a stopband filter for both detuning and pulse amplitude noise. We report at least an order of magnitude improvement in infidelity when our optimized pulse is compared with known composite pulse sequences that are designed to address one or both noise types. Although filter funciton engineering itself is not a novel concept [214], our approach is efficient since the reverse engineering process circumvents the need to compute the evolution operator during the optimization process. The optimizer only requires that we calculate a simple integral expression with the engineered parameters as its input. Furthermore, the engineered parameters offer adequate flexibility to simultaneously target abitrary qubit gates while considering control parameter constraints. In principle, more complicated constraints, such as using different basis functions (Chebyshev, Walsh, Slepian, etc.), time-symmetric or antisymmetric control [208, 210, 242], or spectral-phase-only optimization [207] to name a few, can also be incorporated into our theory. Our results can also be applied to quantum sensing where instead the goal is to maximize the filter function in a limited noise spectral bandwidth [227, 243].

More broadly, we used our theoretical framework to analyze the robustness of geometric gates to detuning and pulse amplitude errors. We proved a theorem for the special case of a control constraint under which geometric gates are necessarily superior to dynamical gates. We emphasize that the robustness we report is not a generic property of geometric gates but rather a consequence of imposing control constraints.

Chapter 7

Summary

NISQ devices suffer primarily from coherent systematic error and stand to benefit from robust quantum control derived using the theory of dynamical error suppression. By lowering the gate error rates, we improve prospects of scalability in current and future quantum devices which is necessary for the successful implementation of fault-tolerant quantum computing. This dissertation explored two methods of dynamical suppression in particular: composite pulse sequences and filter function formalism.

Composite pulse sequences provide a simple method to compensate for the effects of static control perturbations. In this method, robustness is achieved by replacing a noisy quantum gate with a series of other noisy gates that are cleverly designed to ensure that the gate error works against itself. We devoted Chapters 3 and 4 to the application of composite pulses in solid-state systems. In particular, we examined how to dynamically correct entangling gates using simple Hahn-echo like pulses. Although our method was successful in providing significant improvements to the gate fidelity, there are several issues that must be addressed. First, although concatenating different pulse sequences allows us to address more complex situations, it comes at the cost of significantly increasing gate time. This is particularly troublesome in two-qubit gates since qubit coupling is usually weak in some systems (e.g. transmons and singlet-triplet qubits). We saw in Chapter 3 that concatenated pulse sequences in weakly-coupled transmon qubits require coherence times that are not yet achievable with current devices. Fortunately, we were able to avoid this issue in Chapter 4 by using a previously derived analytical model for the entangling operation produced by capacitively-coupled singlet-triplet qubits. The additional analytical insight allowed us to recognize special timings in the entangling operation that can be used to design gates that are stroboscopically robust to certain errors. This approach removed the need for long concatenated pulses since the remaining errors can be efficiently handled by simple Hahn-echo like pulses.

Second, our composite pulses heavily relied on the availability of near-perfect one-qubit operations. We saw in Chapter 3 how severely limiting one-qubit gate errors are to the efficacy of composite pulses. This strongly motivates the need for better one-qubit control protocols. It is conjectured that geometric quantum computing is an excellent candidate for this task. Geometric quantum computing is motivated by the robustness conjecture which states that geometric gates are intrinsically robust to noise since the phase they accumulate is related to some global property of the system's evolution. To test the validity of this conjecture, we used the notion of a filter function. We demonstrated in Chapter 5 that there is no intrinsic robustness to neither geometric nor dynamical gates. Preferential phase robustness exist due to the presence of control constraints. In the absence of constraints, it is possible to find families of Hamiltonians that produce identical quantum gates and filter functions but with varying phase types ranging from purely geometric to purely dynamical. The existence of such families concsequently invalidates the robustness conjecture.

Finally, we were able to use dynamical invariant theory in conjunction with filter function formalism in Chapter 6 to derive an integral expression for a general one-qubit filter function. We use this result to prove a theorem which outlines control scenarios in which geometric gates are naturally superior to dynamical gates. More practically, we used our result to reverse engineer a control Hamiltonian that is robust against a specified noise PSD. We trained a neural network to produce a control pulse that is optimal against broadband $1/f$ detuning and pulse amplitude noise. Our method successfully generated gates that are consistently better than known composite pulse sequences. Furthermore, our method is very general and flexible enough to incorporate many relevant control parameter constraints.

Appendix A: Analysis of the Length-2 Sequence

In Section 3.2 we noted that if all potential commuting errors are present then there exists no $\sigma_{\text{echo}} = \sigma_{cd}$ with $c, d \in \{I, X, Y, Z\}$ that can satisfy

$$\{\sigma_{\text{echo}}, \sigma_{ij}\} = 0 \,\forall\, ij \ni [\sigma_{ij}, \sigma_{ab}] = 0. \tag{A.1}$$

To prove this, we begin by making the observation that there exists a maximal embedding $\mathfrak{a} \oplus \mathfrak{b} \subset \mathfrak{su}(4)$ [244], where '$\oplus$' implies commutation of elements between the respective subalgebras. Let us take $\mathfrak{a} = \mathfrak{u}(1)$ with σ_{ab} as its generator. This means all the generators of $\mathfrak{b}$, which we denote as $\boldsymbol{b} = \{b_1, b_2, \ldots b_n\}$, commute with σ_{ab}. Thus, not only do error channels that commute with σ_{ab} belong in this embedding, but the echo pulse also must belong to the corresponding group embedding since it also needs to commute with σ_{ab} in order for the sequence to produce a non-identity operation. Without losing any generality, we can partition $\boldsymbol{b}$ into two subsets as

$$\boldsymbol{b} = \{\overbrace{b_1, b_2, \ldots}^{\{\sigma_{ij}\}} | \overbrace{\ldots b_n}^{\boldsymbol{b}\backslash\{\sigma_{ij}\}}\},$$

where the left partition contains all the commuting errors that are relevant in the system while the right partition is the coset containing the remaining elements of the generating set of the subalgebra. A σ_{echo} which anticommutes with everything on the left partition and commutes with σ_{ab} can only exist in the coset. If all the possible commuting errors are present, then all the elements of $\boldsymbol{b}$ must belong to the left partition, which leaves no possibility for σ_{echo}. In other words, if all commuting errors are present, then there exists no σ_{echo} that can satisfy Equation (A.1) which proves our claim.

When $\mathfrak{b}$ is a semi-simple Lie-algebra, not all the elements of the coset $\boldsymbol{b} \setminus \{\sigma_{ij}\}$ will necessarily anticommute with all elements of $\{\sigma_{ij}\}$. Nonetheless, if the coset happens to contain a σ_{echo} which anticommutes with all the relevant error channels σ_{ij}, it can be used for error correction. We now show that this is actually the case for the length-2 sequence. But before proceeding further, we first remark that the construction in Section 3.2 relies on the commutation and anticommutation relations of two-qubit Pauli operators. For this reason, we restrict ourselves to embeddings which contain the subalgebra $\mathfrak{b}$ with spinor representation: $\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ [245]. Given our choice $\mathfrak{a} = \text{span}(\sigma_{ab})$ as the $\mathfrak{u}(1)$ subalgebra, there are two choices for $\mathfrak{b}$.

In the first case with $a = b$, all the elements of the generating set

$$\{\sigma_{mI}, \sigma_{Im}, \sigma_{nn}, \sigma_{pp}, \sigma_{mp}, \sigma_{pm}\},$$

where $m = a$, n, and p are mutually distinct and arranged cyclically (e.g. $m = Z, n = X, p = Y$), commute with σ_{aa}. We can define the generators of the commuting $\mathfrak{su}(2)$ subalgebras as $\sigma^{\pm}_{\tilde{X}} \equiv (\sigma_{mI} \pm \sigma_{Im})/2$, $\sigma^{\pm}_{\tilde{Y}} \equiv (\sigma_{np} \pm \sigma_{pn})/2$, and $\sigma^{\pm}_{\tilde{Z}} \equiv (\sigma_{pp} \mp \sigma_{nn})/2$.

In the second case with $a \neq b$, all the elements of the generating set

$$\{\sigma_{mI}, \sigma_{In}, \sigma_{nm}, \sigma_{np}, \sigma_{pm}, \sigma_{pp}\}$$

commute with σ_{ab}, where now we have $m = a$ and $n = b$. The generators can be defined as $\sigma^{\pm}_{\widetilde{X}} \equiv (\sigma_{mI} \pm \sigma_{In})/2$, $\sigma^{\pm}_{\widetilde{Y}} \equiv (\sigma_{np} \mp \sigma_{pm})/2$, and $\sigma^{\pm}_{\widetilde{Z}} \equiv (\sigma_{pp} \pm \sigma_{nm})/2$.

In either case, any error channel or echo pulse lies in the subspace spanned by the "+" and "-" generators (e.g., $\sigma_{aI} = \sigma^{+}_{\widetilde{X}} + \sigma^{-}_{\widetilde{X}}$). Therefore, any given echo pulse can only anticommute with errors belonging to a different subspace. As an example, since the σ_{aI} echo pulse belong to the subspace spanned by the $\sigma_{\widetilde{X}}$ generators, then only errors that belong in the $\sigma_{\widetilde{Y}}$ and $\sigma_{\widetilde{Z}}$ subspaces can be eliminated by a length-2 sequence. Therefore, if all the present commuting errors belong to at most two subspaces only, then the length-2 sequence is sufficient for fixing the errors to first order. The transmon qubit in the main text falls in this category.

A more precise statement of our initial claim is that the length-2 sequence is not capable of correcting errors from all three subspaces. However, placing the initial length-2 sequence inside another length-2 sequence which uses an echo pulse that anticommutes with the initial one allows us to eliminate errors from all three subspaces simultaneously. If we use σ_{aI} for our first sequence's echo pulse, the second echo pulse must be in the $\sigma_{\widetilde{Y}}$ or $\sigma_{\widetilde{Z}}$ subspace in order to satisfy the robustness condition. Clearly, though, the $\mathfrak{u}(1)$ term can not be corrected by a length-2 sequence since it commutes with every allowable echo.

Appendix B: Equivalence of the Length-2 sequence and the ECR scheme

In this section we will show that the ECR scheme is mathematically equivalent to a length-2 sequence with a σ_{XZ} echo pulse. We begin by noting that the h_{IZ} and h_{ZZ} terms in the effective Hamiltonian of a CR gate are proportional to Ω^2, whereas the h_{ZX} and h_{IX} are only proportional to Ω. Thus, in the absence of noise, the evolution can be generally expressed as

$$U\left(\Omega, t\right) = \exp\left[-it\left(\Omega^2\left(a\sigma_{IZ} + b\sigma_{ZZ}\right) + \Omega\left(c\sigma_{ZX} + d\sigma_{IX}\right)\right)\right], \tag{B.1}$$

where a, b, c, and d are given in App. C of Ref. [70]. Using the pulse sequence in Eq. (3.14), we have

$$U(\Omega, t)\sigma_{XI}U(-\Omega, t)\sigma_{XI}.$$

The change from $\Omega \rightarrow -\Omega$ flips the sign of terms that are linearly proportional to Ω. Furthermore, the two σ_{XI} surrounding $U(-\Omega, t)$ flip the sign of terms in the exponential which anticommutes with σ_{XI}. Thus, the cumulative effect of this sequence is

$$
\begin{aligned}
&U(\Omega,t)\sigma_{XI}U(-\Omega,t)\sigma_{XI} \\
&= \exp\Big[-it\big(\Omega^2\big(a\sigma_{IZ}+b\sigma_{ZZ}\big)+\Omega\big(c\sigma_{ZX}+d\sigma_{IX}\big)\big)\Big] \\
&\times \exp\Big[-it\big(\Omega^2\big(a\sigma_{IZ}-b\sigma_{ZZ}\big)-\Omega\big(-c\sigma_{ZX}+d\sigma_{IX}\big)\big)\Big].
\end{aligned} \tag{B.2}
$$

On the other hand, a length-2 sequence with a σ_{XZ} echo pulse yields

$$
\begin{aligned}
&U(\Omega,t)\sigma_{XZ}U(\Omega,t)\sigma_{XZ} \\
&= \exp\Big[-it\big(\Omega^2\big(a\sigma_{IZ}+b\sigma_{ZZ}\big)+\Omega\big(c\sigma_{ZX}+d\sigma_{IX}\big)\big)\Big] \\
&\times \exp\Big[-it\big(\Omega^2\big(a\sigma_{IZ}-b\sigma_{ZZ}\big)+\Omega\big(c\sigma_{ZX}-d\sigma_{IX}\big)\big)\Big],
\end{aligned} \tag{B.3}
$$

where now we flip the sign of terms in the second exponential which anticommute with σ_{XZ}. We see that in either case the final products are exactly equivalent.

Appendix C: Analytical Expression for One-Qubit Clifford RB Fidelity

We now present an analytical expression for the Clifford RB fidelity of a one-qubit gate under the error model given in Equation (3.15). In summary, the goal of Clifford RB is to provide a simple, robust and scalable method for benchmarking the full set of Clifford gates through randomization. The randomization process, also known as twirling, produces a depolarizing channel whose average fidelity can be modeled and experimentally measured. Since the average fidelity of a quantum operation is invariant under the twirling process [246, 247], the measured fidelity is representative of the original untwirled operation. For a more detailed discussion of Clifford RB, we refer the reader to Ref. [55].

The key to creating a depolarizing error channel lies in the fact the uniform probability distribution over the Clifford group, $\mathcal{C}$, comprises a unitary two-design. By definition, this gives the twirling condition

$$\frac{1}{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{C}|}\left(\mathcal{C}_i^{\dagger}\Lambda\mathcal{C}_i\right)(\rho)=\int_{U(d)}\left(U^{\dagger}\Lambda U\right)(\rho)\mathrm{d}U, \tag{C.1}$$

where $\mathcal{C}_i$ are elements of the Clifford group, Λ is an arbitrary quantum channel acting on the system, and the integral is taken with respect to the Haar measure on $U(2^n)$ with n being the number of qubits. The integral in Equation (C.1) produces a unique depolarizing channel Λ_d with the same average fidelity as Λ [246, 247]. The depolarizing channel is modeled by

$$\Lambda_{\mathrm{d}}(\rho) = p\rho + (1-p)\frac{I}{2^n},$$

whose fidelity (as well as Λ's) is given by

$$\mathcal{F} = p + \frac{1-p}{2^n}.$$

To estimate the average fidelity of one-qubit under the error model given in Equation (3.15), we simply replace Λ accordingly and evaluate the sum:

$$\begin{aligned}
&\frac{1}{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{C}|}\left(\mathcal{C}_i^\dagger \exp\left[-i\frac{\varepsilon}{2}\hat{\mathbf{r}}\cdot\vec{\boldsymbol{\sigma}}\right]\mathcal{C}_i\right)(\rho) \\
&= \frac{1}{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{C}|}\mathcal{C}_i^\dagger \exp\left[-i\frac{\varepsilon}{2}\hat{\mathbf{r}}\cdot\vec{\boldsymbol{\sigma}}\right]\mathcal{C}_i\rho\mathcal{C}_i^\dagger \exp\left[i\frac{\varepsilon}{2}\hat{\mathbf{r}}\cdot\vec{\boldsymbol{\sigma}}\right]\mathcal{C}_i \\
&= \frac{I}{2} + \frac{1+2\cos(\theta)}{3}\frac{\hat{\boldsymbol{\rho}}\cdot\vec{\boldsymbol{\sigma}}}{2}.
\end{aligned} \tag{C.2}$$

Since $\rho = \frac{I+\hat{\boldsymbol{\rho}}\cdot\vec{\boldsymbol{\sigma}}}{2}$, then we must have

$$p = \frac{1+2\cos(\theta)}{3}.$$

Thus, the average Clifford RB fidelity is

$$\mathcal{F} = \frac{2+\cos(\theta)}{3}. \tag{C.3}$$

In the main text we noted that we used virtual gates in our simulations. This means that any Z-gates in the Clifford group are treated as noiseless gates. Thus, we can approximate the fidelity when using virtual Z-gates by appropriately weighting the fidelity of the 24 one-qubit Clifford gates:

$$\mathcal{F}_{\mathrm{VZ}} = \frac{20\overline{\mathcal{F}}+4}{24} = \frac{13+5\cos(\theta)}{18}, \tag{C.4}$$

where we assumed that we had 4 noiseless gates $(I, Z, Z_{\pm\frac{\pi}{2}})$. Assuming a Gaussian noise model with a standard deviation $\delta\theta$, we can average over noise realizations and get

$$\begin{aligned}\overline{\mathcal{F}_{\mathrm{VZ}}} &= \frac{1}{\sqrt{2\pi\delta\theta^2}}\int_{-\infty}^{\infty}\exp\left[\frac{-\theta^2}{2\delta\theta^2}\right]\frac{13+5\cos(\theta)}{18}\mathrm{d}\theta\\ &= \frac{13+5\exp\left[-\frac{\delta\theta^2}{2}\right]}{18}.\end{aligned} \tag{C.5}$$

Appendix D: Effects of Exchange Ramping Evolution

In the main text we only considered the case when the exchange $J_i(t)$ is controlled using rectangular pulses both in the beginning and the end of the evolution. Realistically, there is a finite rise time, τ, to go from $J_i(-\tau) \approx 0$ up to $J_i(0) = J_i + j_i$, and, since Equation (4.20) tells us that the exchange should have gone through a odd number of half cycles at the end of the gate, back down from $J_i(t_{gate}) = J_i$ to $J_i(t_{gate} + \tau) \approx 0$. We now consider the effects of the evolution during the finite ramp on our optimization scheme. We will show that the effects are negligible, assuming typical values for the coupling, noise, and rise time.

We choose the well-studied Rosen-Zener pulse shape [248–250] for our ramp:

$$J_i(t) = \begin{cases} J_{i,u} \operatorname{sech}\left(\frac{2\pi t}{\tau}\right), & -\tau < t < 0 \\ J_{i,d} \operatorname{sech}\left(\frac{2\pi(t - t_{gate})}{\tau}\right), & t_{gate} < t < t_{gate} + \tau, \end{cases} \tag{D.1}$$

where $J_{i,u} = J_i + j_i$ is the upward ramp amplitude and $J_{i,d} = J_i$ is the downward ramp amplitude. In addition, since there is a rough proportionality between the average capacitive coupling and the average exchanges, $\alpha \propto J_1 J_2$ [108], the coupling also has a finite ramping time. However, we take $\tau = 1\,\text{ns}$ which is consistent with experimental ramp times in spin qubits [251], and so a typical coupling that ranges up to $1 - 2$ MHz [88, 108] has a negligible effect on such a short time scale. Thus the evolution during the ramp is dominated by the local terms, and the ramping Hamiltonian takes the form

$$\mathcal{H} = \sum_{i=1}^{2}\left(\frac{J_i(t)}{2}\sigma_Z^{(i)} + \frac{h_i}{2}\sigma_X^{(i)}\right). \tag{D.2}$$

We first consider the case where the exchange is ramped up. We begin by noting that since the spin operators for each qubit commute, then we can separate the evolution operator into $U(t) = U_1(t)U_2(t)$. Each of these evolution operators are solutions to

$$\imath\frac{d}{dt}U_i(t) = \left(\frac{J_{i,u}\,\mathrm{sech}\left(\frac{2\pi t}{\tau}\right)}{2}\sigma_Z^{(i)} + \frac{h_i}{2}\sigma_X^{(i)}\right)U_i(t). \tag{D.3}$$

In order for us to use known analytical results, we first rotate to a frame so that

$$U_i(t) = \exp\left[\imath\frac{\pi}{4}\sigma_Y^{(i)}\right]\exp\left[\imath t\frac{h_i}{2}\sigma_Z^{(i)}\right]U_i'(t). \tag{D.4}$$

This allows us to write two coupled differential equations

$$\begin{cases} \imath\,\dot{s}(t) = \frac{J_{i,u}\,\mathrm{sech}\left(2\pi t/\tau\right)}{2}\mathrm{e}^{-\imath h_i t}p(t), \\ \imath\,\dot{p}(t) = \frac{J_{i,u}\,\mathrm{sech}\left(2\pi t/\tau\right)}{2}\mathrm{e}^{\imath h_i t}s(t), \end{cases} \tag{D.5}$$

where $U_i'(t)\psi'(t_o) = (s(t), p(t))^{\mathrm{T}}$ and $\psi'(t_o)$ is the initial wavefunction. Using the results from Refs. [248, 250], we can write the time-evolution in the rotating frame for $t \leq 0$ as

$$U_i'(t) = U_{\mathrm{I}}\mathbb{1} + U_{\mathrm{X}}\sigma_X^{(i)} + U_{\mathrm{Y}}\sigma_y^{(i)} + U_{\mathrm{Z}}\sigma_Z^{(i)}, \tag{D.6}$$

where

$$U_{\mathrm{I}} = \frac{1}{2}\left\{ {}_2F_1\left[-\frac{J_{i,u}\tau}{2}, \frac{J_{i,u}\tau}{2}; \frac{1-\imath h_i\tau}{2}; z\right] + {}_2F_1\left[-\frac{J_{i,u}\tau}{2}, \frac{J_{i,u}\tau}{2}; \frac{1+\imath h_i\tau}{2}; z\right]\right\}$$

$$\begin{aligned} U_{\mathrm{X}} &= \frac{1}{4}J_i\tau\,\mathrm{sech}\left[\frac{t}{\tau}\right]\left\{\frac{\mathrm{e}^{-\imath h_i t}{}_2F_1\left[1-\frac{J_{i,u}\tau}{2}, 1+\frac{J_{i,u}\tau}{2}; \frac{3-\imath h_i\tau}{2}; z\right]}{h_i\tau+\imath}\right. \\ &\qquad \left. - \frac{\mathrm{e}^{\imath h_i t}{}_2F_1\left[1-\frac{J_{i,u}\tau}{2}, 1+\frac{J_{i,u}\tau}{2}; \frac{3+\imath h_i\tau}{2}; z\right]}{h_i\tau-\imath}\right\} \\ U_{\mathrm{Y}} &= \imath\frac{1}{4}J_i\tau\,\mathrm{sech}\left[\frac{t}{\tau}\right]\left\{\frac{\mathrm{e}^{-\imath h_i t}{}_2F_1\left[1-\frac{J_{i,u}\tau}{2}, 1+\frac{J_{i,u}\tau}{2}; \frac{3-\imath h_i\tau}{2}; z\right]}{h_i\tau+\imath}\right. \\ &\qquad \left. + \frac{\mathrm{e}^{\imath h_i t}{}_2F_1\left[1-\frac{J_{i,u}\tau}{2}, 1+\frac{J_{i,u}\tau}{2}; \frac{3+\imath h_i\tau}{2}; z\right]}{h_i\tau-\imath}\right\} \end{aligned} \tag{D.7}$$

$$U_{\mathrm{Z}} = \frac{1}{2}\left\{ {}_2F_1\left[-\frac{J_{i,u}\tau}{2}, \frac{J_{i,u}\tau}{2}; \frac{1+\imath h_i\tau}{2}; z\right] - {}_2F_1\left[-\frac{J_{i,u}\tau}{2}, \frac{J_{i,u}\tau}{2}; \frac{1-\imath h_i\tau}{2}; z\right]\right\},$$

and ${}_2F_1[a, b; c; d]$ is Gauss's hypergeometric function and $z = \frac{1}{2}\left(1+\tanh\left[\frac{t}{\tau}\right]\right)$. We note that $U_i'(t)$ satisfies the initial condition $U_i'(-\infty) = \mathbb{1}$. In order to get the actual solution to equation (D.3) with $U_i(-\tau) = \mathbb{1}$, we use the composition property of time-evolution operators:

$$U_i(t; -\tau) = U_i(t; -\infty)U_i^\dagger(-\tau; -\infty), \tag{D.8}$$

where $U(t; t_o)$ indicates the evolution from t_o to t. More explicitly, the upward ramp evolution operator that corresponds to the Hamiltonian in equation (D.2) is approximately given by

$$\begin{aligned} U_u(t; -\tau) = \exp&\left[\imath\frac{\pi}{4}\sum_{i=1}^{2}\sigma_Y^{(i)}\right]\exp\left[\imath t\sum_{i=1}^{2}h_i\sigma_Z^{(i)}\right]U_1'(t) \\ &\times U_1'^\dagger(-\tau)U_2'(t)U_2'^\dagger(-\tau) \\ &\times \exp\left[\imath\tau\sum_{i=1}^{2}h_i\sigma_Z^{(i)}\right]\exp\left[-\imath\frac{\pi}{4}\sum_{i=1}^{2}\sigma_Y^{(i)}\right]. \end{aligned} \tag{D.9}$$

To solve for the downward ramp evolution, we first note that there is a relationship between the upward and downward ramp Hamiltonian when their amplitudes are similar: $\mathcal{H}_d(t) = \mathcal{H}_u(t_{gate} - t)$. Using this, then we can write the time-evolution of the downward ramp as

$$U_d(t) = \mathcal{T} \exp\left[-\imath \int_{t_{gate}}^{t} H_d(t')\mathrm{d}t'\right]. \tag{D.10}$$

where $\mathcal{T}$ denotes the time-ordering operator. Using a simple change of variable and using the composition property of time-evolution operators, we can express the evolution of the downward ramp in terms of the upward ramp:

$$\begin{aligned}
U_d(t) &= \mathcal{T} \exp\left[-\imath \int_{t_{gate}}^{t} H_u(t_{gate} - t')\mathrm{d}t'\right] \\
&= \mathcal{T} \exp\left[\imath \int_{0}^{t_{gate}-t} H_u(t'')\mathrm{d}t''\right] \\
&= \mathcal{T} \exp\left[\imath \int_{-\tau}^{t_{gate}-t} H_u(t'')\mathrm{d}t''\right] \\
&\quad \times \mathcal{T} \exp\left[\imath \int_{0}^{-\tau} H_u(t'')\mathrm{d}t''\right] \\
&= \mathcal{T} \exp\left[\imath \int_{-\tau}^{t_{gate}-t} H_u(t'')\mathrm{d}t''\right] \\
&\quad \times \left(\mathcal{T} \exp\left[\imath \int_{-\tau}^{0} H_u(t'')\mathrm{d}t''\right]\right)^{\dagger} \\
&= \bar{U}_u(t_{gate} - t; -\tau)\bar{U}_u^{\dagger}(0, -\tau).
\end{aligned}$$

where the bar indicates change from $J_{i,u} \to -J_{i,d}$, and $h_i \to -h_i$. Therefore, the downward ramp evolution operator is given by

$$\begin{aligned} U_d(t;t_{gate}) = &\exp\left[\imath\frac{\pi}{4}\sum_{i=1}^{2}\sigma_Y^{(i)}\right]\exp\left[-\imath(\tau-t)\sum_{i=1}^{2}h_i\sigma_Z^{(i)}\right] \\ &\times \bar{U}_1'(\tau-t)\bar{U}_1'(0)\bar{U}_2'(\tau-t)\bar{U}_2'(0) \\ &\times \exp\left[-\imath\frac{\pi}{4}\sum_{i=1}^{2}\sigma_Y^{(i)}\right]. \end{aligned} \tag{D.11}$$

Now that we have an analytical expression for the ramp evolution operators, we can finally address how they affect the error channels and our optimization. In the presence of noise, it can be verified numerically with the parameters provided in Section 4.4 that perturbations in J_i results in infidelities that are one to two orders of magnitude smaller than the infidelities we report in the main text. This can be mainly attributed to the fact that $h_i \gg J_i$ and τ is relatively short. Thus, the dominant source of error in the ramp evolution is due to perturbations in the magnetic gradient δh_i. However, if we assume 1 ns ramp times and a standard deviation $\delta h_i = 8\text{neV}$ [112], the resulting infidelities are also found to be an order of magnitude smaller than those discussed in the main text. Thus, provided that $\delta h_i \tau$ is much less than the remaining errors in Table 4.2, then the errors associated with the ramp can be neglected.

Finally, we address how the unperturbed ramp evolution affect the error channels. The total evolution of the qubits is given by

$$U(t) = U_u(t)R_1(t)\exp\left[-\imath t\frac{\alpha h_1 h_2}{\Omega_1\Omega_2}\sigma_{ZZ}\right]R_2(t)U_d(t). \tag{D.12}$$

We can rewrite this in terms of our optimized gate given in equation (4.14):

$$U(t) = U_u(t)R_1(t)U_{nl}(t)R_2(t)U_d. \tag{D.13}$$

Since U_u and U_d are purely local operations and provided that the ramp errors are negligible, then applying an initial local rotation $R_1^\dagger(t)U_u^\dagger(t)$ and a final local rotation $U_d^\dagger(t)R_2^\dagger(t)$ ensures that our optimized gate $U_{nl}(t)$ and its errors are unperturbed by the ramps.

Appendix E: Error Channels

We present here a table of error channels for the dissimilar qubit case in Section 4.3.

Table E.1: First-order errors for the similar qubit case with $J_i \gg h_i$. Due to the complexity of the error channels, we had only shown the errors due to fluctuations in the first qubit. To find the effects of perturbations in the second qubit, one need only generate a second table where the labels are swapped ($1 \leftrightarrow 2$ and $\sigma_{ij} \leftrightarrow \sigma_{ji}$).

σ_{IX}	0
σ_{IY}	0
σ_{IZ}	$\left(\frac{\imath\left((J_1\delta h_1-h_1\delta J_1)\sin(\omega_1 t+\xi_1)-2\Omega_1^2 t\left(\frac{\partial\chi_1}{\partial h_1}\delta h_1+\frac{\partial\chi_1}{\partial J_1}\delta J_1\right)\right)}{2\Omega_1^2}-\imath t\frac{\partial\chi_1}{\partial j_1}\delta j_1\right)\sin^2\left(\frac{J_1J_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{XI}	$\left[\left(\frac{\imath(h_1\delta J_1-J_1\delta h_1)\cos(\omega_1 t+\xi_1)}{2\Omega_1^2}-\frac{\imath(h_1\delta h_1+J_1\delta J_1)}{4\chi_1\Omega_1}\right)\sin(2\chi_1 t) -\frac{\imath}{2}\cos(2\chi_1 t)\left(\frac{\partial\xi_1}{\partial h_1}\delta h_1+\frac{\partial\xi_1}{\partial J_1}\delta J_1+\frac{\partial\xi_1}{\partial j_1}\delta j_1\right)\right]\cos\left(\frac{J_1J_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{XX}	$\frac{\imath h_1J_2\alpha t(J_1\delta h_1-h_1\delta J_1)}{2\Omega_1^3\Omega_2}-\imath\frac{J_1J_2t\delta\alpha}{2\Omega_1\Omega_2}$
σ_{XY}	$\left(\frac{\imath\left((h_1\delta J_1-J_1\delta h_1)\sin(\omega_1 t+\xi_1)+2\Omega_1^2 t\left(\frac{\partial\chi_1}{\partial h_1}\delta h_1+\frac{\partial\chi_1}{\partial J_1}\delta J_1\right)\right)}{4\Omega_1^2}+\frac{\imath}{2}t\frac{\partial\chi_1}{\partial j_1}\delta j_1\right)\sin\left(\frac{2J_1J_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{XZ}	0
σ_{YI}	$\left[\frac{\imath(h_1\delta J_1-J_1\delta h_1)(\cos(2\chi_1 t)\cos(\omega_1 t+\xi_1))}{2\Omega_1^2}+\frac{\imath(h_1\delta h_1+J_1\delta J_1)\sin^2(\chi_1 t)}{2\chi_1\Omega_1} +\frac{\imath}{2}\sin(2\chi_1 t)\left(\frac{\partial\xi_1}{\partial h_1}\delta h_1+\frac{\partial\xi_1}{\partial J_1}\delta J_1+\frac{\partial\xi_1}{\partial j_1}\delta j_1\right)\right]\cos\left(\frac{J_1J_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{YX}	$\left(\frac{\imath\left((J_1\delta h_1-h_1\delta J_1)\sin(\omega_1 t+\xi_1)-2\Omega_1^2 t\left(\frac{\partial\chi_1}{\partial h_1}\delta h_1+\frac{\partial\chi_1}{\partial J_1}\delta J_1\right)\right)}{4\Omega_1^2}-\frac{\imath}{2}t\frac{\partial\chi_1}{\partial j_1}\delta j_1\right)\sin\left(\frac{2J_1J_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{YY}	$\frac{\imath h_1J_2\alpha t(J_1\delta h_1-h_1\delta J_1)}{2\Omega_1^3\Omega_2}-\imath\frac{J_1J_2t\delta\alpha}{2\Omega_1\Omega_2}$
σ_{YZ}	0
σ_{ZI}	$\left(\frac{\imath\left((J_1\delta h_1-h_1\delta J_1)\sin(\omega_1 t+\xi_1)-2\Omega_1^2 t\left(\frac{\partial\chi_1}{\partial h_1}\delta h_1+\frac{\partial\chi_1}{\partial J_1}\delta J_1\right)\right)}{2\Omega_1^2}-\imath t\frac{\partial\chi_1}{\partial j_1}\delta j_1\right)\cos^2\left(\frac{J_1J_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{ZX}	$\left[\frac{\imath(J_1\delta h_1-h_1\delta J_1)\cos(2\chi_1 t)\cos(\omega_1 t+\xi_1)}{2\Omega_1^2}-\frac{\imath(h_1\delta h_1+J_1\delta J_1)\sin^2(\chi_1 t)}{2\chi_1\Omega_1} -\frac{\imath}{2}\sin(2\chi_1 t)\left(\frac{\partial\xi_1}{\partial h_1}\delta h_1+\frac{\partial\xi_1}{\partial J_1}\delta J_1+\frac{\partial\xi_1}{\partial j_1}\delta j_1\right)\right]\sin\left(\frac{J_1J_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{ZY}	$\left[\left(\frac{\imath(h_1\delta J_1-J_1\delta h_1)\cos(\omega_1 t+\xi_1)}{2\Omega_1^2}-\frac{\imath(h_1\delta h_1+J_1\delta J_1)}{4\chi_1\Omega_1}\right)\sin(2\chi_1 t) -\frac{\imath}{2}\cos(2\chi_1 t)\left(\frac{\partial\xi_1}{\partial h_1}\delta h_1+\frac{\partial\xi_1}{\partial J_1}\delta J_1+\frac{\partial\xi_1}{\partial j_1}\delta j_1\right)\right]\sin\left(\frac{J_1J_2\alpha t}{\Omega_1\Omega_2}\right)$
σ_{ZZ}	0

Appendix F: Estimating gate fidelity using filter functions

We now provide a more detailed derivation of the average gate infidelity provided in Equation (6.14) which was reported in Ref. [47]. We begin by writing the noisy Hamiltonian as

$$H(t) = H_c(t) + H_e(t), \tag{F.1}$$

where $H_c(t)$ is the deterministic control Hamiltonian and $H_e(t)$ is the stochastic error Hamiltonian which can generally expressed as in Equation (6.13). By moving to the interaction frame, we can write the noisy time evolution $U(t) = U_c(t)U_e(t)$ where each factors are solutions to the following Schrödinger equations:

$$i\dot{U}_c(t) = H_c(t)U_c(t) \tag{F.2}$$

$$i\dot{U}_e(t) = \left(U_c^{\dagger}(t)H_eU_c(t)\right)U_e(t). \tag{F.3}$$

For sufficiently weak noise, we can perturbatively expand $U_e(t)$ using the Magnus expansion and write

$$U_e(T) \approx \exp\left[-i\int_0^T U_c^{\dagger}(t)H_e(t)U_c(t)\mathrm{d}t\right]. \tag{F.4}$$

The average gate infidelity is given by

$$
\begin{aligned}
\langle \mathcal{I} \rangle &= \langle 1 - F_{\mathrm{tr}} \rangle = \left\langle 1 - \left| \mathrm{tr}\left(U_c^\dagger U\right) / \mathrm{tr}\left(U_c^\dagger U_c\right) \right|^2 \right\rangle \\
&= \left\langle 1 - \left| \mathrm{tr}\, U_e / 2 \right|^2 \right\rangle \\
&\approx \left\langle \mathrm{tr} \int_0^T \int_0^T U_c^\dagger(t_1) H_e(t_1) U_c(t_1) U_c^\dagger(t_2) H_e(t_2) U_c(t_2) \mathrm{d}t_1 \mathrm{d}t_2 \right\rangle. \qquad \text{(F.5)}
\end{aligned}
$$

We can use the adjoint representation of $U_c(t)$ defined through

$$
R_{ij}(t) = \frac{1}{2} \mathrm{tr}\left(U_c^\dagger(t) \sigma_i U_c(t) \sigma_j\right) \qquad \text{(F.6)}
$$

and Equation (6.13) to rewrite Equation (F.5) into

$$
\begin{aligned}
\langle \mathcal{I} \rangle \approx \sum_{q,i,j,k} \int_0^T \int_0^T \left\langle \delta_q(t_1) \delta_q(t_2) \right\rangle \chi_{q,i}(t_1) \chi_{q,j}(t_2) \\
\times R_{ik}(t_1) R_{jk}(t_2) \mathrm{d}t_1 \mathrm{d}t_2. \qquad \text{(F.7)}
\end{aligned}
$$

We may invoke the Wiener-Khinchin theorem for a wide-sense stationary noise process to express the autocorrelation function of $\delta_q(t)$ as the Fourier transform of its PSD: $\langle \delta_q(t_1), \delta_q(t_2) \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_q(\omega) \mathrm{e}^{i\omega(t_2 - t_1)} \mathrm{d}\omega$. If we further define

$$
R_{q,k}(\omega) := \sum_i \int_0^T \chi_{q,i}(t) R_{ik}(t) \mathrm{e}^{i\omega t} \mathrm{d}t, \qquad \text{(F.8)}
$$

we can finally compactly write the gate infidelity as

$$
\langle \mathcal{I} \rangle \approx \frac{1}{2\pi} \sum_q \int_{-\infty}^{\infty} S_q(\omega) F_q(\omega) \mathrm{d}\omega, \qquad \text{(F.9)}
$$

where $F_q(\omega) \equiv \sum_k |R_{q,k}(\omega)|^2$.

Appendix G: Numerical optimization method

We describe here the details of our numerical optimization in Section 6.3. We used Julia's DiffEqFlux package to create a feedforward deep neural network with one input neuron, two output neurons, and two hidden layers with 16 neurons each. Our goal is to minimize the cost given in Equation (6.32). The infidelity integral of a noise process q in the first two terms of Equation (6.32) can be expressed as

$$\langle \mathcal{I}_q \rangle \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_0^T \int_0^T \left(\Lambda(t_1)\vec{\chi}_q(t_1)\right)^\intercal \Lambda(t_2)\vec{\chi}_q(t_2) S_q(\omega) e^{i\omega(t_1-t_2)} \, \mathrm{d}t_1 \, \mathrm{d}t_2 \, \mathrm{d}\omega, \quad \text{(G.1)}$$

where $\vec{\chi} = [\chi_{q,X}, \chi_{q,Y}, \chi_{q,Z}]^\intercal$ is the error sensitivity vector. In the main text, the noise PSD assumes one of two nontrivial forms: $\frac{A}{\omega}$ and $\frac{A\omega_c}{\omega^2}$. We can evaluate the frequency integrals analytically which are given by

$$\int_{\omega_0}^{\omega_c} \frac{A}{\omega} e^{i\omega t} \, \mathrm{d}t = 2 \left(\mathrm{Ci}\left(\omega_c t\right) - \mathrm{Ci}\left(\omega_o t\right) \right), \quad \text{(G.2)}$$

$$\int_{\omega_c}^{\infty} \frac{A\omega_c}{\omega^2} e^{i\omega t} \, \mathrm{d}t = -\pi\omega_c t + 2\cos\left(\omega_c t\right) + 2\omega_c t \, \mathrm{Si}\left(\omega_c t\right), \quad \text{(G.3)}$$

where $\mathrm{Ci}(t)$ and $\mathrm{Si}(t)$ are the cosine and sine integral function, respectively. Let us define $g_q(t_1 - t_2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_q(\omega) e^{i\omega(t_1-t_2)} \, \mathrm{d}\omega$. This allows us to express Equation (G.1) as

$$\int_0^T \int_0^T g_q(t_1 - t_2) \left(\Lambda(t_1)\vec{\chi}_q(t_1)\right)^\intercal \Lambda(t_2)\vec{\chi}_q(t_2) \, \mathrm{d}t_1 \, \mathrm{d}t_2. \quad \text{(G.4)}$$

We can approximate the integrals by converting them into a series of matrix multiplications. In particular, we can treat each time integral as an integral operator which has g_q as its kernel and takes in $\boldsymbol{v}_q = \Lambda\vec{\chi}_q$ as input. Therefore, the average infidelity may be rewritten in the following bilinear form

$$\langle \mathcal{I}_q \rangle \approx \boldsymbol{v}_q^\intercal \mathbb{L} \boldsymbol{v}_q, \tag{G.5}$$

where $\mathbb{L}$ is a matrix that approximates the double time integral.

In our work, the cost is completely vectorized by evaluating the cost terms in evenly spaced intervals of time. The infidelity integrals are evaluated using Equation (G.5) while derivatives, which are used in evaluating quantities such as Ω in Equation (6.27), are implemented using finite differences. Thus, the speed and accuracy of optimization may be controlled by choosing an appropriate level of time discretization.

Bibliography

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[2] R. P. Feynman, International Journal of Theoretical Physics **21**, 467 (1982).

[3] P. Benioff, Journal of Statistical Physics **29**, 515 (1982).

[4] D. Deutsch and R. Penrose, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **400**, 97 (1985).

[5] D. Deutsch and R. Jozsa, Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences **439**, 553 (1992).

[6] D. Simon, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (1994) pp. 116–123.

[7] P. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (1994) pp. 124–134.

[8] L. K. Grover, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96 (Association for Computing Machinery, New York, NY, USA, 1996) p. 212–219.

[9] I. L. Chuang and Y. Yamamoto, Phys. Rev. A **52**, 3489 (1995).

[10] J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).

[11] S. Lloyd, Science **261**, 1569 (1993).

[12] D. P. DiVincenzo, Phys. Rev. A **51**, 1015 (1995).

[13] T. Hayashi, T. Fujisawa, H. D. Cheong, Y. H. Jeong, and Y. Hirayama, Phys. Rev. Lett. **91**, 226804 (2003).

[14] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).

[15] A. Shnirman, G. Schön, and Z. Hermon, Phys. Rev. Lett. **79**, 2371 (1997).

[16] D. Averin, Solid State Communications **105**, 659 (1998).

[17] Y. Makhlin, G. Scöhn, and A. Shnirman, Nature **398**, 305 (1999).

[18] D. P. DiVincenzo, Fortschritte der Physik **48**, 771 (2000).

[19] J. Preskill, Quantum **2**, 79 (2018).

[20] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, 2013).

[21] P. Shor, in *Proceedings of 37th Conference on Foundations of Computer Science* (1996) pp. 56–65.

[22] D. S. Wang, A. G. Fowler, and L. C. L. Hollenberg, Phys. Rev. A **83**, 020302 (2011).

[23] A. G. Fowler, A. C. Whiteside, and L. C. L. Hollenberg, Phys. Rev. Lett. **108**, 180501 (2012).

[24] R. Freeman, S. P. Kempsell, and M. H. Levitt, Journal of Magnetic Resonance (1969) **38**, 453 (1980).

[25] M. H. Levitt and R. Freeman, Journal of Magnetic Resonance (1969) **33**, 473 (1979).

[26] M. H. Levitt and R. Freeman, Journal of Magnetic Resonance (1969) **43**, 65 (1981).

[27] M. H. Levitt and R. Freeman, Journal of Magnetic Resonance (1969) **43**, 502 (1981).

[28] M. H. Levitt, Progress in Nuclear Magnetic Resonance Spectroscopy **18**, 61 (1986).

[29] H. K. Cummins, G. Llewellyn, and J. A. Jones, Phys. Rev. A **67**, 042308 (2003).

[30] I. Roos and K. Mølmer, Phys. Rev. A **69**, 022321 (2004).

[31] C. D. Hill, Phys. Rev. Lett. **98**, 180501 (2007).

[32] N. Timoney, V. Elman, S. Glaser, C. Weiss, M. Johanning, W. Neuhauser, and C. Wunderlich, Phys. Rev. A **77**, 052334 (2008).

[33] P. Rebentrost and F. K. Wilhelm, Phys. Rev. B **79**, 060507 (2009).

[34] J. A. Jones, Progress in Nuclear Magnetic Resonance Spectroscopy **59**, 91 (2011).

[35] S. S. Ivanov and N. V. Vitanov, Opt. Lett. **36**, 1275 (2011).

[36] X. Wang, L. S. Bishop, J. P. Kestner, E. Barnes, K. Sun, and S. Das Sarma, Nature Communications **3**, 997 (2012).

[37] J. P. Kestner, X. Wang, L. S. Bishop, E. Barnes, and S. Das Sarma, Phys. Rev. Lett. **110**, 140502 (2013).

[38] X. Wang, L. S. Bishop, E. Barnes, J. P. Kestner, and S. D. Sarma, Phys. Rev. A **89**, 022310 (2014).

[39] P. Dutta and P. M. Horn, Rev. Mod. Phys. **53**, 497 (1981).

[40] O. E. Dial, M. D. Shulman, S. P. Harvey, H. Bluhm, V. Umansky, and A. Yacoby, Phys. Rev. Lett. **110**, 146804 (2013).

[41] J. Yoneda, K. Takeda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, T. Honda, T. Kodera, S. Oda, Y. Hoshi, N. Usami, K. M. Itoh, and S. Tarucha, Nature Nanotechnology **13**, 102 (2018).

[42] K. W. Chan, W. Huang, C. H. Yang, J. C. C. Hwang, B. Hensen, T. Tanttu, F. E. Hudson, K. M. Itoh, A. Laucht, A. Morello, and A. S. Dzurak, Phys. Rev. Applied **10**, 044017 (2018).

[43] D. J. Van Harlingen, T. L. Robertson, B. L. T. Plourde, P. A. Reichardt, T. A. Crane, and J. Clarke, Phys. Rev. B **70**, 064517 (2004).

[44] F. Yoshihara, K. Harrabi, A. O. Niskanen, Y. Nakamura, and J. S. Tsai, Phys. Rev. Lett. **97**, 167001 (2006).

[45] M. Kim, H. J. Mamin, M. H. Sherwood, K. Ohno, D. D. Awschalom, and D. Rugar, Phys. Rev. Lett. **115**, 087602 (2015).

[46] M. J. Biercuk, A. C. Doherty, and H. Uys, Journal of Physics B: Atomic, Molecular and Optical Physics **44**, 154002 (2011).

[47] T. J. Green, J. Sastrawan, H. Uys, and M. J. Biercuk, New Journal of Physics **15**, 095004 (2013).

[48] A. Soare, H. Ball, D. Hayes, J. Sastrawan, M. C. Jarratt, J. J. McLoughlin, X. Zhen, T. J. Green, and M. J. Biercuk, Nature Physics **10**, 825 (2014).

[49] O. Gamel, Phys. Rev. A **93**, 062320 (2016).

[50] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Phys. Rev. A **52**, 3457 (1995).

[51] A. Y. Kitaev, Russian Mathematical Surveys **52**, 1191 (1997).

[52] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces* (American Mathematical Society, 2001).

[53] S. Kais, K. B. Whaley, A. R. Dinner, and S. A. Rice, *Quantum Information and Computation for Chemistry*, Vol. 154 (John Wiley and Sons, 2014).

[54] C. P. Williams, S. H. Clearwater, *et al.*, *Explorations in Quantum Computing* (Springer London, London, 2011).

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

[56] R. K. L. Colmenar, U. Güngördü, and J. P. Kestner, Phys. Rev. A **102**, 032626 (2020).

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

[58] J. A. Schreier, A. A. Houck, J. Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. B **77**, 180502(R) (2008).

[59] A. A. Houck, J. A. Schreier, B. R. Johnson, J. M. Chow, J. Koch, J. M. Gambetta, D. I. Schuster, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. Lett. **101**, 080502 (2008).

[60] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, Phys. Rev. Lett. **111**, 080502 (2013).

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

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

[63] J. M. Chow, L. DiCarlo, J. M. Gambetta, F. Motzoi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **82**, 040305(R) (2010).

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

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

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

[67] G. S. Paraoanu, Phys. Rev. B **74**, 140504(R) (2006).

[68] J. Li, K. Chalapat, and G. S. Paraoanu, Phys. Rev. B **78**, 064503 (2008).

[69] J. Majer, J. M. Chow, J. M. Gambetta, J. Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Nature **449**, 443 (2007).

[70] E. Magesan and J. M. Gambetta, Phys. Rev. A **101**, 052308 (2020).

[71] V. Tripathi, M. Khezri, and A. N. Korotkov, Phys. Rev. A **100**, 012301 (2019).

[72] L. M. K. Vandersypen and I. L. Chuang, Rev. Mod. Phys. **76**, 1037 (2005).

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

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

[75] F. A. Calderon-Vargas and J. P. Kestner, Phys. Rev. Lett. **118**, 150502 (2017).

[76] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Phys. Rev. A **67**, 042313 (2003).

[77] M. R. Geller, E. J. Pritchett, A. Galiautdinov, and J. M. Martinis, Phys. Rev. A **81**, 012320 (2010).

[78] A. Warren, E. Barnes, and S. E. Economou, Phys. Rev. B **100**, 161303(R) (2019).

[79] M. A. Fogarty, M. Veldhorst, R. Harper, C. H. Yang, S. D. Bartlett, S. T. Flammia, and A. S. Dzurak, Phys. Rev. A **92**, 022326 (2015).

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

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

[82] R. Cabrera and W. E. Baylis, Physics Letters A **368**, 25 (2007).

[83] A. P. M. Place, L. V. H. Rodgers, P. Mundada, B. M. Smitham, M. Fitzpatrick, Z. Leng, A. Premkumar, J. Bryon, A. Vrajitoarea, S. Sussman, G. Cheng, T. Madhavan, H. K. Babla, X. H. Le, Y. Gang, B. Jäck, A. Gyenis, N. Yao, R. J. Cava, N. P. de Leon, and A. A. Houck, Nature Communications **12**, 1779 (2021).

[84] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Annual Review of Condensed Matter Physics **11**, 369 (2020).

[85] P. Jurcevic, A. Javadi-Abhari, L. S. Bishop, I. Lauer, D. F. Bogorin, M. Brink, L. Capelluto, O. Günlük, T. Itoko, N. Kanazawa, A. Kandala, G. A. Keefe, K. Krsulich, W. Landers, E. P. Lewandowski, D. T. McClure, G. Nannicini, A. Narasgond, H. M. Nayfeh, E. Pritchett, M. B. Rothwell, S. Srinivasan, N. Sundaresan, C. Wang, K. X. Wei, C. J. Wood, J.-B. Yau, E. J. Zhang, O. E. Dial, J. M. Chow, and J. M. Gambetta, Quantum Science and Technology **6**, 025020 (2021).

[86] M. Takita, A. D. Córcoles, E. Magesan, B. Abdo, M. Brink, A. Cross, J. M. Chow, and J. M. Gambetta, Phys. Rev. Lett. **117**, 210505 (2016).

[87] M. Takita, A. W. Cross, A. D. Córcoles, J. M. Chow, and J. M. Gambetta, Phys. Rev. Lett. **119**, 180501 (2017).

[88] J. M. Nichol, L. A. Orona, S. P. Harvey, S. Fallahi, G. C. Gardner, M. J. Manfra, and A. Yacoby, npj Quantum Information **3**, 3 (2017).

[89] R. K. L. Colmenar and J. P. Kestner, Phys. Rev. A **99**, 012347 (2019).

[90] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).

[91] M. Veldhorst, J. C. C. Hwang, C. H. Yang, A. W. Leenstra, B. de Ronde, J. P. Dehollain, J. T. Muhonen, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, Nature Nanotechnology **9**, 981 (2014).

[92] T. F. Watson, S. G. J. Philips, E. Kawakami, D. R. Ward, P. Scarlino, M. Veldhorst, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, Nature **555**, 633 EP (2018).

[93] D. M. Zajac, A. J. Sigillito, M. Russ, F. Borjans, J. M. Taylor, G. Burkard, and J. R. Petta, Science **359**, 439 (2018).

[94] J. 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, Science **309**, 2180 (2005).

[95] B. M. Maune, M. G. Borselli, B. Huang, T. D. Ladd, P. W. Deelman, K. S. Holabird, A. A. Kiselev, I. Alvarado-Rodriguez, R. S. Ross, A. E. Schmitz, M. Sokolich, C. A. Watson, M. F. Gyure, and A. T. Hunter, Nature **481**, 344 (2012).

[96] A. V. Kuhlmann, J. Houel, A. Ludwig, L. Greuter, D. Reuter, A. D. Wieck, M. Poggio, and R. J. Warburton, Nature Physics **9**, 570 (2013).

[97] K. Eng, T. D. Ladd, A. Smith, M. G. Borselli, A. A. Kiselev, B. H. Fong, K. S. Holabird, T. M. Hazard, B. Huang, P. W. Deelman, I. Milosavljevic, A. E. Schmitz, R. S. Ross, M. F. Gyure, and A. T. Hunter, Science Advances **1** (2015).

[98] D. J. Reilly, J. M. Taylor, E. A. Laird, J. R. Petta, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Phys. Rev. Lett. **101**, 236803 (2008).

[99] L. Cywiński, W. M. Witzel, and S. Das Sarma, Phys. Rev. Lett. **102**, 057601 (2009).

[100] E. Barnes, L. Cywiński, and S. Das Sarma, Phys. Rev. Lett. **109**, 140403 (2012).

[101] J. Medford, L. Cywiński, C. Barthel, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Phys. Rev. Lett. **108**, 086802 (2012).

[102] F. K. Malinowski, F. Martins, L. Cywiński, M. S. Rudner, P. D. Nissen, S. Fallahi, G. C. Gardner, M. J. Manfra, C. M. Marcus, and F. Kuemmeth, Phys. Rev. Lett. **118**, 177702 (2017).

[103] J. Yoneda, T. Otsuka, T. Takakura, M. Pioro-Ladrière, R. Brunner, H. Lu, T. Nakajima, T. Obata, A. Noiri, C. J. Palmstrøm, A. C. Gossard, and S. Tarucha, Applied Physics Express **8**, 084401 (2015).

[104] M. Pioro-Ladrière, T. Obata, Y. Tokura, Y.-S. Shin, T. Kubo, K. Yoshida, T. Taniyama, and S. Tarucha, Nature Physics **4**, 776 (2008).

[105] R. Brunner, Y.-S. Shin, T. Obata, M. Pioro-Ladrière, T. Kubo, K. Yoshida, T. Taniyama, Y. Tokura, and S. Tarucha, Phys. Rev. Lett. **107**, 146801 (2011).

[106] X. Wu, D. R. Ward, J. R. Prance, D. Kim, J. K. Gamble, R. T. Mohr, Z. Shi, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, Proceedings of the National Academy of Sciences **111**, 11938 (2014).

[107] E. Kawakami, T. Jullien, P. Scarlino, D. R. Ward, D. E. Savage, M. G. Lagally, V. V. Dobrovitski, M. Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, Proceedings of the National Academy of Science **113**, 11738 (2016).

[108] M. D. Shulman, O. E. Dial, S. P. Harvey, H. Bluhm, V. Umansky, and A. Yacoby, Science **336**, 202 (2012).

[109] F. A. Calderon-Vargas and J. P. Kestner, Phys. Rev. B **97**, 125311 (2018).

[110] B. Kraus and J. I. Cirac, Phys. Rev. A **63**, 062309 (2001).

[111] M. J. Bremner, D. Bacon, and M. A. Nielsen, Phys. Rev. A **71**, 052312 (2005).

[112] H. Bluhm, S. Foletti, I. Neder, M. Rudner, D. Mahalu, V. Umansky, and A. Yacoby, Nature Physics **7**, 109 (2010).

[113] F. K. Malinowski, F. Martins, P. D. Nissen, E. Barnes, L. Cywinski, M. S. Rudner, S. Fallahi, G. C. Gardner, M. J. Manfra, C. M. Marcus, and F. Kuemmeth, Nature Nanotechnology **12**, 16 EP (2016).

[114] S. Kogan, *Electronic Noise and Fluctuations in Solids* (Cambridge University Press, 1996).

[115] R. K. Colmenar, U. Güngördü, and J. Kestner, Invalidating the robustness conjecture for geometric quantum gates (2021), arXiv:2105.02882 .

[116] E. Knill, Nature **434**, 39 (2005).

[117] A. Ekert, M. Ericsson, P. Hayden, H. Inamori, J. A. Jones, D. K. L. Oi, and V. Vedral, Journal of Modern Optics **47**, 2501 (2000).

[118] M. V. Berry, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **392**, 45 (1984).

[119] Y. Aharonov and J. Anandan, Phys. Rev. Lett. **58**, 1593 (1987).

[120] J. Anandan, Physics Letters A **133**, 171 (1988).

[121] P. Zanardi and M. Rasetti, Physics Letters A **264**, 94 (1999).

[122] S.-L. Zhu and Z. D. Wang, Phys. Rev. Lett. **91**, 187902 (2003).

[123] D. Kult, J. Åberg, and E. Sjöqvist, Phys. Rev. A **74**, 022106 (2006).

[124] J. A. Jones, V. Vedral, A. Ekert, and G. Castagnoli, Nature **403**, 869 (2000).

[125] J. Du, P. Zou, and Z. D. Wang, Phys. Rev. A **74**, 020302(R) (2006).

[126] Y. Ota, Y. Goto, Y. Kondo, and M. Nakahara, Phys. Rev. A **80**, 052311 (2009).

[127] D. Leibfried, B. DeMarco, V. Meyer, D. Lucas, M. Barrett, J. Britton, W. M. Itano, B. Jelenković, C. Langer, T. Rosenband, and D. J. Wineland, Nature **422**, 412 (2003).

[128] C.-Y. Guo, L.-L. Yan, S. Zhang, S.-L. Su, and W. Li, Phys. Rev. A **102**, 042607 (2020).

[129] C. Zu, W.-B. Wang, L. He, W.-G. Zhang, C.-Y. Dai, F. Wang, and L.-M. Duan, Nature **514**, 72 (2014).

[130] L. Wang, T. Tu, B. Gong, C. Zhou, and G.-C. Guo, Scientific Reports **6**, 19048 (2016).

[131] Y. Sekiguchi, N. Niikura, R. Kuroiwa, H. Kano, and H. Kosaka, Nature Photonics **11**, 309 (2017).

[132] N. Ishida, T. Nakamura, T. Tanaka, S. Mishima, H. Kano, R. Kuroiwa, Y. Sekiguchi, and H. Kosaka, Opt. Lett. **43**, 2380 (2018).

[133] K. Nagata, K. Kuramitani, Y. Sekiguchi, and H. Kosaka, Nature Communications **9**, 3227 (2018).

[134] Y.-Y. Huang, Y.-K. Wu, F. Wang, P.-Y. Hou, W.-B. Wang, W.-G. Zhang, W.-Q. Lian, Y.-Q. Liu, H.-Y. Wang, H.-Y. Zhang, L. He, X.-Y. Chang, Y. Xu, and L.-M. Duan, Phys. Rev. Lett. **122**, 010503 (2019).

[135] X. Tan, D.-W. Zhang, Z. Zhang, Y. Yu, S. Han, and S.-L. Zhu, Phys. Rev. Lett. **112**, 027001 (2014).

[136] C. Song, S.-B. Zheng, P. Zhang, K. Xu, L. Zhang, Q. Guo, W. Liu, D. Xu, H. Deng, K. Huang, D. Zheng, X. Zhu, and H. Wang, Nature Communications **8**, 1061 (2017).

[137] Y. Xu, W. Cai, Y. Ma, X. Mu, L. Hu, T. Chen, H. Wang, Y. P. Song, Z.-Y. Xue, Z.-q. Yin, and L. Sun, Phys. Rev. Lett. **121**, 110501 (2018).

[138] Y. Xu, Z. Hua, T. Chen, X. Pan, X. Li, J. Han, W. Cai, Y. Ma, H. Wang, Y. P. Song, Z.-Y. Xue, and L. Sun, Phys. Rev. Lett. **124**, 230503 (2020).

[139] A. A. Abdumalikov Jr, J. M. Fink, K. Juliusson, M. Pechal, S. Berger, A. Wallraff, and S. Filipp, Nature **496**, 482 (2013).

[140] S.-L. Zhu and P. Zanardi, Phys. Rev. A **72**, 020301(R) (2005).

[141] J. Pachos and P. Zanardi, International Journal of Modern Physics B **15**, 1257 (2001).

[142] A. Carollo, I. Fuentes-Guridi, M. F. Santos, and V. Vedral, Phys. Rev. Lett. **90**, 160402 (2003).

[143] G. De Chiara and G. M. Palma, Phys. Rev. Lett. **91**, 090404 (2003).

[144] G. De Chiara, A. Łoziński, and G. M. Palma, The European Physical Journal D **41**, 179 (2007).

[145] Z. S. Wang, C. Wu, X.-L. Feng, L. C. Kwek, C. H. Lai, C. H. Oh, and V. Vedral, Phys. Rev. A **76**, 044303 (2007).

[146] J. T. Thomas, M. Lababidi, and M. Tian, Phys. Rev. A **84**, 042335 (2011).

[147] Z.-T. Liang, X. Yue, Q. Lv, Y.-X. Du, W. Huang, H. Yan, and S.-L. Zhu, Phys. Rev. A **93**, 040305(R) (2016).

[148] T. Chen and Z.-Y. Xue, Phys. Rev. Applied **10**, 054051 (2018).

[149] B.-J. Liu, X.-K. Song, Z.-Y. Xue, X. Wang, and M.-H. Yung, Phys. Rev. Lett. **123**, 100501 (2019).

[150] T. Chen and Z.-Y. Xue, Phys. Rev. Applied **14**, 064009 (2020).

[151] S. Berger, M. Pechal, A. A. Abdumalikov, C. Eichler, L. Steffen, A. Fedorov, A. Wallraff, and S. Filipp, Phys. Rev. A **87**, 060303(R) (2013).

[152] F. Kleißler, A. Lazariev, and S. Arroyo-Camejo, npj Quantum Information **4**, 49 (2018).

[153] A. Nazir, T. P. Spiller, and W. J. Munro, Phys. Rev. A **65**, 042303 (2002).

[154] A. Blais and A. M. S. Tremblay, Phys. Rev. A **67**, 012308 (2003).

[155] Y. Ota and Y. Kondo, Phys. Rev. A **80**, 024302 (2009).

[156] S.-B. Zheng, C.-P. Yang, and F. Nori, Phys. Rev. A **93**, 032313 (2016).

[157] W. Dong, F. Zhuang, S. E. Economou, and E. Barnes, PRX Quantum **2**, 030333 (2021).

[158] P. Solinas, P. Zanardi, and N. Zanghì, Phys. Rev. A **70**, 042316 (2004).

[159] A. Carollo, I. Fuentes-Guridi, M. F. Santos, and V. Vedral, Phys. Rev. Lett. **92**, 020402 (2004).

[160] J. Dajka, M. Mierzejewski, and J. Łuczka, Journal of Physics A: Mathematical and Theoretical **41**, 012001 (2007).

[161] M. Johansson, E. Sjöqvist, L. M. Andersson, M. Ericsson, B. Hessmo, K. Singh, and D. M. Tong, Phys. Rev. A **86**, 062322 (2012).

[162] H. R. Lewis, Phys. Rev. Lett. **18**, 510 (1967).

[163] H. Ball and M. J. Biercuk, EPJ Quantum Technology **2**, 11 (2015).

[164] A. Mostafazadeh, *Dynamical Invariants, Adiabatic Approximation and the Geometric Phase* (Nova Science Publishers, Inc., New York, 2001).

[165] X. Chen, E. Torrontegui, and J. G. Muga, Phys. Rev. A **83**, 062116 (2011).

[166] E. Torrontegui, S. Ibáñez, X. Chen, A. Ruschhaupt, D. Guéry-Odelin, and J. G. Muga, Phys. Rev. A **83**, 013415 (2011).

[167] X. Chen and J. G. Muga, Phys. Rev. A **86**, 033405 (2012).

[168] U. Güngördü, Y. Wan, M. A. Fasihi, and M. Nakahara, Phys. Rev. A **86**, 062312 (2012).

[169] E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, Phys. Rev. A **89**, 043408 (2014).

[170] A. Levy, A. Kiely, J. G. Muga, R. Kosloff, and E. Torrontegui, New Journal of Physics **20**, 025006 (2018).

[171] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, Rev. Mod. Phys. **91**, 045001 (2019).

[172] H. R. Lewis and W. B. Riesenfeld, Journal of Mathematical Physics **10**, 1458 (1969).

[173] F. Wilczek and A. Zee, Phys. Rev. Lett. **52**, 2111 (1984).

[174] T. Ichikawa, M. Bando, Y. Kondo, and M. Nakahara, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **370**, 4671 (2012).

[175] L.-F. Wei, J.-Q. Liang, and B.-Z. Li, Il Nuovo Cimento B (1971-1996) **110**, 1357 (1995).

[176] T. Bitter and D. Dubbers, Phys. Rev. Lett. **59**, 251 (1987).

[177] G. Giavarini, E. Gozzi, D. Rohrlich, and W. Thacker, Physics Letters A **138**, 235 (1989).

[178] A. Laucht, R. Kalra, S. Simmons, J. P. Dehollain, J. T. Muhonen, F. A. Mohiyaddin, S. Freer, F. E. Hudson, K. M. Itoh, D. N. Jamieson, J. C. McCallum, A. S. Dzurak, and A. Morello, Nature Nanotechnology **12**, 61 (2017).

[179] N. A. Gershenfeld and I. L. Chuang, Science **275**, 350 (1997).

[180] P. Z. Zhao, X.-D. Cui, G. F. Xu, E. Sjöqvist, and D. M. Tong, Phys. Rev. A **96**, 052316 (2017).

[181] E. Sjöqvist, V. Azimi Mousolou, and C. M. Canali, Quantum Information Processing **15**, 3995 (2016).

[182] J. Xu, S. Li, T. Chen, and Z.-Y. Xue, Frontiers of Physics **15**, 41503 (2020).

[183] E. Sjöqvist, D. M. Tong, L. M. Andersson, B. Hessmo, M. Johansson, and K. Singh, New Journal of Physics **14**, 103035 (2012).

[184] W. Pfeifer, *The Lie Algebras* $su(N)$ (Birkhäuser, Basel, 2003).

[185] L.-M. Duan, J. I. Cirac, and P. Zoller, Science **292**, 1695 (2001).

[186] J. Zhang, T. H. Kyaw, D. M. Tong, E. Sjöqvist, and L.-C. Kwek, Scientific Reports **5**, 18414 (2015).

[187] 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, Phys. Rev. A **82**, 042339 (2010).

[188] D. Kim, D. R. Ward, C. B. Simmons, J. K. Gamble, R. Blume-Kohout, E. Nielsen, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, Nature Nanotechnology **10**, 243 (2015).

[189] R. K. L. Colmenar and J. P. Kestner, Efficient reverse engineering of one-qubit filter functions with dynamical invariants (2022), arXiv:2204.08457 .

[190] R. J. Gordon and S. A. Rice, Annual Review of Physical Chemistry **48**, 601 (1997).

[191] A. Assion, T. Baumert, M. Bergt, T. Brixner, B. Kiefer, V. Seyfried, M. Strehle, and G. Gerber, Science **282**, 919 (1998).

[192] S. A. Rice and S. P. Shah, Phys. Chem. Chem. Phys. **4**, 1683 (2002).

[193] F. Poggiali, P. Cappellaro, and N. Fabbri, Phys. Rev. X **8**, 021059 (2018).

[194] M. M. Müller, S. Gherardini, and F. Caruso, Scientific Reports **8**, 14278 (2018).

[195] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **82**, 2417 (1999).

[196] K. Khodjasteh and D. A. Lidar, Phys. Rev. Lett. **95**, 180501 (2005).

[197] M. J. Biercuk, H. Uys, A. P. VanDevender, N. Shiga, W. M. Itano, and J. J. Bollinger, Nature **458**, 996 (2009).

[198] G. de Lange, Z. H. Wang, D. Ristè, V. V. Dobrovitski, and R. Hanson, Science **330**, 60 (2010).

[199] B. Naydenov, F. Dolde, L. T. Hall, C. Shin, H. Fedder, L. C. L. Hollenberg, F. Jelezko, and J. Wrachtrup, Phys. Rev. B **83**, 081201 (2011).

[200] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. G. Cory, Y. Nakamura, J.-S. Tsai, and W. D. Oliver, Nature Physics **7**, 565 (2011).

[201] K. R. Brown, A. W. Harrow, and I. L. Chuang, Phys. Rev. A **70**, 052318 (2004).

[202] R. W. Simmonds, K. M. Lang, D. A. Hite, S. Nam, D. P. Pappas, and J. M. Martinis, Phys. Rev. Lett. **93**, 077003 (2004).

[203] C. Kabytayev, T. J. Green, K. Khodjasteh, M. J. Biercuk, L. Viola, and K. R. Brown, Phys. Rev. A **90**, 012316 (2014).

[204] U. Güngördü and J. P. Kestner, Phys. Rev. B **98**, 165301 (2018).

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

[206] D. Daems, A. Ruschhaupt, D. Sugny, and S. Guérin, Phys. Rev. Lett. **111**, 050404 (2013).

[207] Y. Guo, D. Dong, and C.-C. Shu, Phys. Chem. Chem. Phys. **20**, 9498 (2018).

[208] U. Güngördü and J. P. Kestner, Phys. Rev. A **100**, 062310 (2019).

[209] T. Nöbauer, A. Angerer, B. Bartels, M. Trupke, S. Rotter, J. Schmiedmayer, F. Mintert, and J. Majer, Phys. Rev. Lett. **115**, 190801 (2015).

[210] E. Barnes, X. Wang, and S. Das Sarma, Scientific Reports **5**, 12685 (2015).

[211] B. T. Torosov and N. V. Vitanov, Phys. Rev. A **83**, 053420 (2011).

[212] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems (2015), software available from tensorflow.org.

[213] M. Innes, E. Saba, K. Fischer, D. Gandhi, M. C. Rudilosso, N. M. Joy, T. Karmali, A. Pal, and V. Shah, Fashionable modelling with flux (2018), arXiv:1811.01457 .

[214] H. Ball, M. J. Biercuk, A. R. R. Carvalho, J. Chen, M. Hush, L. A. D. Castro, L. Li, P. J. Liebermann, H. J. Slatyer, C. Edmunds, V. Frey, C. Hempel, and A. Milne, Quantum Science and Technology **6**, 044011 (2021).

[215] C. D. B. Bentley, H. Ball, M. J. Biercuk, A. R. R. Carvalho, M. R. Hush, and H. J. Slatyer, Advanced Quantum Technologies **3**, 2000044 (2020).

[216] Y. Baum, M. Amico, S. Howell, M. Hush, M. Liuzzi, P. Mundada, T. Merkh, A. R. Carvalho, and M. J. Biercuk, PRX Quantum **2**, 040324 (2021).

[217] A. R. R. Carvalho, H. Ball, M. J. Biercuk, M. R. Hush, and F. Thomsen, Phys. Rev. Applied **15**, 064054 (2021).

[218] U. Güngördü and J. P. Kestner, Robust quantum gates using smooth pulses and physics-informed neural networks (2020), arXiv:2011.02512 .

[219] D. W. Kanaar, S. Wolin, U. Güngördü, and J. P. Kestner, Phys. Rev. B **103**, 235314 (2021).

[220] J. Zeng, X.-H. Deng, A. Russo, and E. Barnes, New Journal of Physics **20**, 033011 (2018).

[221] J. Zeng and E. Barnes, Phys. Rev. A **98**, 012301 (2018).

[222] J. Zeng, C. H. Yang, A. S. Dzurak, and E. Barnes, Phys. Rev. A **99**, 052321 (2019).

[223] D. Buterakos, S. Das Sarma, and E. Barnes, PRX Quantum **2**, 010341 (2021).

[224] E. Barnes, F. A. Calderon-Vargas, W. Dong, B. Li, J. Zeng, and F. Zhuang, Dynamically corrected gates from geometric space curves (2021), arXiv:2103.16015 .

[225] F. Zhuang, J. Zeng, S. E. Economou, and E. Barnes, Quantum **6**, 639 (2022).

[226] E. Kreyszig, *Differential Geometry*, Dover Books on Mathematics (Dover Publications, 2013).

[227] L. M. Norris, D. Lucarelli, V. M. Frey, S. Mavadia, M. J. Biercuk, and L. Viola, Phys. Rev. A **98**, 032315 (2018).

[228] E. Knill, R. Laflamme, R. Martinez, and C.-H. Tseng, Nature **404**, 368 (2000).

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

[230] B. R. Johnson, M. P. da Silva, C. A. Ryan, S. Kimmel, J. M. Chow, and T. A. Ohki, New Journal of Physics **17**, 113019 (2015).

[231] Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).

[232] J. Schmidhuber, Neural Networks **61**, 85 (2015).

[233] A. Krizhevsky, I. Sutskever, and G. E. Hinton, Commun. ACM **60**, 84–90 (2017).

[234] G. Cybenko, Mathematics of Control, Signals and Systems **2**, 303 (1989).

[235] K. Hornik, Neural Networks **4**, 251 (1991).

[236] H. K. Cummins and J. A. Jones, New Journal of Physics **2**, 006 (2000).

[237] H. Cummins and J. Jones, Journal of Magnetic Resonance **148**, 338 (2001).

[238] S. Wimperis, Journal of Magnetic Resonance, Series A **109**, 221 (1994).

[239] M. Bando, T. Ichikawa, Y. Kondo, and M. Nakahara, Journal of the Physical Society of Japan **82**, 014004 (2013).

[240] S.-L. Zhu and Z. D. Wang, Phys. Rev. Lett. **89**, 097902 (2002).

[241] J. R. Schrieffer and P. A. Wolff, Phys. Rev. **149**, 491 (1966).

[242] N. E. Bonesteel, D. Stepanenko, and D. P. DiVincenzo, Phys. Rev. Lett. **87**, 207901 (2001).

[243] V. M. Frey, S. Mavadia, L. M. Norris, W. de Ferranti, D. Lucarelli, L. Viola, and M. J. Biercuk, Nature Communications **8**, 2189 (2017).

[244] R. Slansky, Physics Reports **79**, 1 (1981).

[245] P. Ramond, *Group Theory: A Physicist's Survey* (Cambridge University Press, 2010).

[246] M. A. Nielsen, Physics Letters A **303**, 249 (2002).

[247] J. Emerson, R. Alicki, and K. Życzkowski, Journal of Optics B: Quantum and Semiclassical Optics **7**, S347 (2005).

[248] N. Rosen and C. Zener, Phys. Rev. **40**, 502 (1932).

[249] R. T. Robiscoe, Phys. Rev. A **17**, 247 (1978).

[250] B. T. Torosov and N. V. Vitanov, Phys. Rev. A **76**, 053404 (2007).

[251] I. van Weperen, B. D. Armstrong, E. A. Laird, J. Medford, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Phys. Rev. Lett. **107**, 030506 (2011).