Title
Evaluation of Composite Pulse Techniques to Generate Robust Single-Qubit Gates

Author
Williams, Kajsa M

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Evaluation of Composite Pulse Techniques to Generate Robust Single-Qubit Gates

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Science in Chemistry

by

Kajsa M. Williams

2022

ABSTRACT OF THE THESIS

Evaluation of Composite Pulse Techniques to Generate
Robust Single-Qubit Gates

by

Kajsa M. Williams
Master of Science in Chemistry
University of California, Los Angeles, 2022
Professor Louis-Serge Bouchard, Chair

Superconducting transmon qubits are a promising architecture for large-scale fault-tolerant computation: transmons possess favorable properties for scalability, and mature microwave control electronics are readily available. Recent advances in coherent control of these devices have produced optimally calibrated microwave control pulses with fidelities approaching fault-tolerance thresholds. However, miscalibrations and drift in the amplitude or frequency of the microwave drive may result in suboptimal gate performance. Compensation schemes employing open-loop optimal control to increase robustness toward these types of frequency detuning and field amplitude errors have been under development in the NMR community since the 1970s. Composite and adiabatic pulses are two such robust pulse design techniques that may benefit coherent control of superconducting architectures.

In this thesis, we demonstrate applications of composite and adiabatic pulses for the implementation of robust inversion gates on a transmon qubit. We characterized the robustness, fidelities, seepage and leakage rates of the inversion pulses using simulated and experimental transition probabilities and randomized benchmarking methods. Both adiabatic and composite pulses were able to compensate for a broader range of systematic drive amplitude and off-resonance errors compared to standard gates. Several composite pulse schemes also improved on-resonance fidelity.

The thesis of Kajsa M. Williams is approved.

Benjamin Joel Schwartz

Daniel Neuhauser

Louis-Serge Bouchard, Committee Chair

University of California, Los Angeles

2022

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Louis-Serge Bouchard for his guidance and invaluable feedback during my graduate studies at UCLA. Chapter 4 of this thesis emerged from a recently submitted paper [1], which could not have been completed without his excellent direction and editorial assistance as the principal investigator. I am also indebted to my committee members, Daniel Neuhauser and Benjamin Joel Schwartz, for their knowledgeable review of my thesis and for inspiring my interest in physical chemistry, both inside the classroom and out.

CHAPTER 1

Introduction

Quantum information processing is a promising technology to solve certain classically intractable problems. Quantum algorithms such as Shor's algorithm for prime-number factorization offer nearly exponential speedups over their classical counterparts [2], and simulations of entangled many-particle systems that would be impossible classically are theoretically feasible with quantum computers [3, 4]. Current quantum processors operate in the noisy intermediate scale quantum (NISQ) regime and are not of sufficient size or accuracy to realize the quantum supremacy promised by such canonical applications. Even so, a variety of algorithms demonstrating quantum advantage have already been implemented on NISQ architectures. Entire economic sectors such as quantum finance have emerged due to quadratic speedups with quantum implementations of Monte Carlo algorithms [5]. Transcending the NISQ regime ultimately requires improvements both in the scalability of quantum information processors and in the fidelity of quantum control operations in order to satisfy the requirements of error-correction schemes.

NISQ superconducting architectures possess several attributes that make these systems a promising platform on which to develop next-generation quantum information processors. Superconducting qubits offer favorable scalability, well-characterized tunable coupling schemes, long coherence times, and mature electronics for qubit-specific measurement and for crafting custom microwave control fields [6]. NISQ-era improvements in the scalability and intrinsic coherence times of superconducting qubits are largely attributable to the development of transmission line shunted plasma oscillation (transmon) qubits. Transmons are quantized superconducting LC-circuits with an anharmonic potential shaped by the relative strength of the nonlinear inductance of the superconducting Josephson junction (L_J) and the

capacitance of a shunt capacitor (C_s). NISQ transmon qubits operate in the regime where the Josephson junction energy greatly exceeds that of the shunt capacitor. This reduces the sensitivity of transmons to charge noise and allows for improved coherence times over pre-NISQ superconducting qubits, which had higher shunt capacitor energies. Microwave control has also progressed in the NISQ era. Sophisticated control pulse design techniques have been developed that attain gate fidelities approaching fault-tolerance thresholds [7, 8]; however, these nominal fidelities may be reduced in the presence of amplitude or frequency errors in the microwave control field.

Compensation schemes are required to reduce the sensitivity of control pulses to frequency detuning and field amplitude errors, which can arise from a multitude of sources, including drift in the instrument calibration over time, random jitter, or programming errors. Well-established techniques to compensate for miscalibration of the control field have been used for decades in the NMR community, where control field inhomogeneities have historically been a dominant source of error [9]. Robust control of other quantum information processing architectures has previously benefited from the implementation of pulsed NMR techniques [10, 11, 12, 13]. However, these techniques have never been applied to NISQ transmon qubit architectures. Although the transmon qubit is encoded in a subspace of the anharmonic potential such that leakage to higher energy states must be considered to fully evaluate control pulse performance, the idealized single-qubit gate problem from transmons is still well modeled by a two-level system completely analogous to the NMR control problem. Robust NMR techniques are therefore readily applicable to coherent control of transmon qubits and may offer benefits similar to those previously demonstrated on other architectures.

In this thesis, we present an application of two microwave control field design techniques — composite pulses (CPs) and adiabatic pulses (APs) — to generate robust single-qubit gates on a transmon qubit. We implemented industry-standard methods to characterize the robustness, fidelity, and leakage of the control pulses based on simulated and experimental inversion profiles, interleaved randomized benchmarking, and leakage randomized bench-

marking. We demonstrate that composite and adiabatic pulses generally improve robustness towards frequency and amplitude errors and can additionally increase on-resonance fidelity. We also consider examples of distinct classes of composite and adiabatic pulses designed to either perform transformations between well-defined initial and final states ('point-to-point' transformations) or general (universal) rotations. The latter class of general rotation pulses is a particularly promising foundation for constructing universal gatesets endowed with greater tolerance towards miscalibrations of the control field.

The main experimental results of this work are presented in chapter 4, which contains the findings of the robustness, fidelity, and leakage measures for composite and adiabatic pulses. Details of the experimental implementation of microwave control field shaping techniques and performance characterization methods are also discussed in that section. Chapter 3 presents a more detailed theoretical background on randomized benchmarking, and chapter 2 summarizes the necessary quantum mechanical formalism. Chapters 2 and 3 are derivative, and no originality is claimed for the content of these sections.

CHAPTER 2

Mathematical Preliminaries

This chapter is intended to give an overview of the quantum mechanical formalism used in the remainder of this thesis. A more comprehensive discussion of the principles of quantum computing introduced herein can be found in [14], and additional resources on the applications of unitary t-designs in quantum channel characterization are available [15, 16, 17].

2.1 Quantum Systems and Hilbert Spaces

Quantum mechanical systems are represented with respect to Hilbert spaces, which are most generally defined as a vector space that is complete and endowed with a norm. A d-dimensional isolated quantum system is completely described by a $(d \times 1)$ column vector $|\psi\rangle$ in an associated state space $\mathcal{H}_d$, which is a complex Hilbert space of dimension d. The inner product structure that defines the norm on this space is the vector inner product $\langle\phi|\psi\rangle$ where $\langle\phi|$ is a member of the adjoint space $\mathcal{H}_d^*$. There is a one-to-one correspondence between the vectors $|\psi\rangle$ in the state space $\mathcal{H}_d$ and the unique functionals $\mathfrak{f}_\phi[\,] = \langle\phi|$ in the adjoint space $\mathcal{H}_d^*$, whose action $\mathfrak{f}_\phi[\psi] = \langle\phi|\psi\rangle$ is to take an input vector and return a complex (c) number, with the aid of the inner product.

In this thesis, we will deal with finite-dimensional Hilbert spaces that can be equipped with an orthonormal basis $\{\phi_i\}$ such that any state vector can be written as a linear combination of orthonormal basis elements

$$|\psi\rangle = \sum_i c_i \, |\phi_i\rangle \, .$$

In this representation, elements of $\mathcal{H}_d^*$ are formed as the complex conjugate transpose of the elements of $\mathcal{H}_d$, in accordance with the typical definition of the Hermitian conjugate $\psi^\dagger = (\psi^*)^T$, i.e.

$$\langle\psi| \equiv |\psi\rangle^\dagger .$$

Realistic quantum systems often include system-environment interactions that cannot be easily represented in the state-vector formalism. Density operators were proposed by von Neumann in order to describe pure and mixed states. Such operators are positive and have unity trace ($\rho \in \mathcal{O}_d$),

$$\rho = \sum_j p_j \, |\psi_j\rangle \, \langle\psi_j| .$$

Pure states are characterized by $\mathrm{Tr}[\rho^2]$=1 whereas mixed states have $0 \leq \mathrm{Tr}[\rho^2] < 1$. Here, $|\psi_j\rangle \langle\psi_j|$ is the outer product, and $\mathcal{O}_d$ is the space of linear operators (represented by $d \times d$ matrices) acting on $\mathcal{H}_d$. $\mathcal{O}_d$ is itself also a Hilbert space with an inner product structure given by the Hilbert-Schmidt or trace inner product,

$$\langle\sigma|\tau\rangle = \mathrm{Tr}\left(\sigma^\dagger \tau\right)$$

for $\sigma, \tau \in \mathcal{O}_d$. The strength of the density operator formalism is that it can be used to represent systems in an unknown state, which is described by an ensemble of probabilities and state vectors $\{p_i, |\psi_i\rangle\}$. Pure states are idempotent $\rho^2 = \rho$ and can be written as a rank-1 projector $|\psi\rangle \langle\psi|$.

2.1.1 Composite Systems

The state space of a composite system $\mathcal{H}_{AB}$ formed from separable Hilbert spaces $\mathcal{H}_A$, with basis $\{|\psi_n\rangle\}$, and $\mathcal{H}_B$, with basis $\{|\phi_m\rangle\}$, is normally constructed from the tensor product of $\mathcal{H}_A$ and $\mathcal{H}_B$:

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B.$$

The composite state space $\mathcal{H}_{AB}$ is of dimension $\dim(\mathcal{H}_A) \times \dim(\mathcal{H}_B)$ and is spanned by the tensor product basis $\{|\psi_1\rangle \otimes |\phi_1\rangle, \ldots |\psi_1\rangle \otimes |\phi_m\rangle, \ldots, |\psi_n\rangle \otimes |\phi_1\rangle, \ldots |\psi_n\rangle \otimes |\phi_m\rangle\}$. The multipartite composition of separable Hilbert spaces is formed by simple extension such that the composite state space is the tensor product of the state spaces describing each party.

Given operators $A \in \mathcal{O}_{d_A}(\mathcal{H}_A)$ and $B \in \mathcal{O}_{d_B}(\mathcal{H}_B)$ the operator on the composite Hilbert space is defined as the tensor product of the operators and acts on states of the composite system $|\psi\rangle \otimes |\phi\rangle$ as shown.

$$(A \otimes B)(|\psi\rangle \otimes |\phi\rangle) = A|\psi\rangle \otimes B|\phi\rangle$$

The shorthand notation $|\psi\rangle|\phi\rangle$ or $|\psi\phi\rangle$ may also be used to represent tensor product states $|\psi\rangle \otimes |\phi\rangle$.

2.1.2 Purification of Density Matrices

Composite systems are of central importance to the field of quantum computing, not only for their use in describing the composition of physical systems of interest in multi-qubit computations but also for their utility in modeling system-environment interactions. Purification of density matrices is a powerful tool that treats such system-environment interactions from a purely mathematical perspective in order to simplify calculations involving mixed states [14].

Definition 2.1.1. *Purification*: Given a mixed state ρ^A of a system with a state space $\mathcal{H}_A$, we may introduce a fictitious reference system R on the same state space such that the composite system is in a pure state $|AR\rangle\langle AR|$.

The state of the principle system ρ^A is recovered from the composite system by performing a partial trace over the degrees of freedom of the system R.

Definition 2.1.2. *Partial Trace*: The partial trace Tr_B operation for a composite system $\mathcal{H}_A \otimes \mathcal{H}_B$ is defined for an operator on the composite space $A \otimes B$, where $A \in \mathcal{H}_A$ and $B \in \mathcal{H}_B$,

as

$$\mathrm{Tr}_B(A \otimes B) = \mathrm{Tr}(B)A$$

The purification procedure can be validated by considering decompositions of the states ρ^A and $|AR\rangle$. The density matrix ρ^A is a positive operator, and as such has an orthonormal decomposition $\rho = \sum_i p_i |i_A\rangle \langle i_A|$ by the spectral theorem for normal operators. Here $\{|i_A\rangle\}$ is an orthonormal basis of $\mathcal{H}_A$. For the pure state $|AR\rangle$, a particular form of decomposition known as the Schmidt decomposition [14] may be applied.

Defintion 2.1.3. *Schmidt Decomposition*: For any pure state $|\psi\rangle$ of a composite system AR there exists an orthonormal set of states $\{|i_A\rangle\}$ for the system A and a set of orthonormal states for the system R $\{|i_R\rangle\}$ such that $|\psi\rangle$ can be decomposed over the tensor product states $|i_A\rangle |i_R\rangle$ with non-negative, real expansion coefficients λ_i. The λ_i are called *Schmidt coefficients* and satisfy $\sum_i \lambda_i^2 = 1$

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_R\rangle .$$

Defining the Schmidt decomposition for the pure state of the composite system as

$$|AR\rangle \equiv \sum_i \sqrt{p_i} |i_A\rangle |i_R\rangle$$

where the p_i are the eigenvalues of ρ^A, the reduced density matrix ρ^A is then obtained from the partial trace Tr_R in the given basis.

$$\mathrm{Tr}_R(|AR\rangle \langle AR|) = \sum_k \langle k_R| \sum_{i,j} \sqrt{p_i p_j} |i_A\rangle |i_R\rangle \langle j_A| \langle j_R| |k_R\rangle =$$

$$\sum_{i,j,k} \sqrt{p_i p_j} |i_A\rangle \langle j_A| \delta_{ik}\delta_{jk} = \sum_i p_i |i_A\rangle \langle i_A| = \rho^A.$$

2.2 Evolution

Several interconnected formalisms for describing the evolution of quantum systems have been developed. These methods negotiate a balance between generality in the type of quantum evolution they can represent and the specificity of the knowledge they provide regarding the time evolution of the system under consideration. The unitary propagator formalism is well suited to descriptions of ideal, isolated quantum dynamics and, as such, provides a useful tool with which to optimize ideal qubit dynamics under the influence of a control field (see appendix B). However, realistic quantum systems are typically not isolated (open), and it is important to understand the influence of interactions with the environment on the evolution of the principal system. The quantum operations formalism and master equation formalism both allow such non-unitary dynamics to be characterized. Quantum operations are the most general representation of quantum evolution as they are time-independent descriptions of discrete state changes. In this thesis, we represent open quantum dynamics using the quantum operations formalism and use the unitary propagator formalism to describe ideal qubit control.

2.2.1 Unitary Evolution of State Vectors

The continuous-time dynamics of state vector evolution in closed systems are governed by the Schrödinger equation,

$$\frac{d\,|\psi\rangle}{dt} = -iH\,|\psi\rangle$$

to which the solution in the case of a time-dependent Hamiltonian may be written using the time-ordering operator $\mathcal{T}$. Terms in the expansion of the exponential are organized by the action $\mathcal{T}$ such that operators corresponding to earlier times appear to the right of those for later times. Note, we use Hartree atomic units ($\hbar = 1$) throughout this thesis.

$$|\psi(t)\rangle = \mathcal{T}\exp\left[-i\int_0^t H(t')dt'\right]|\psi_0\rangle$$

The general form of the solution to this equation suggests an alternative description of the system evolution in terms of a unitary propagator $U(t)$ that depends only on the initial and final times and transforms the state $|\psi_0\rangle$ into $|\psi(t)\rangle$.

$$U(t) = \mathcal{T} \exp\left[-i \int_0^t H(t')dt'\right]$$

Isolated quantum systems evolve according to unitary dynamics governed by this propagator, which satisfies the definition of a unitary operator on $\mathcal{H}_d$ ($UU^\dagger = U^\dagger U = \mathbb{I}_d$).

2.2.2 Density Matrix Evolution and Master Equations

Continuous-time evolution of density matrices in closed systems can be described by a reformulation of the Schrödinger equation known as the Liouville master equation, in which $[A, B] = AB - BA$ is the commutator and $\mathcal{L}$ is the Liouville superoperator.

$$\frac{d}{dt}\rho(t) = -i\,[H(t), \rho(t)] = \mathcal{L}\hat{\hat{\rho}}(t)$$

Definition 2.2.1. *Superoperator*: A superoperator $\mathcal{S}$ belonging to the Hilbert space of superoperators $\mathcal{S}_d$ is a linear operator on the elements $\rho \in \mathcal{O}_d(\mathcal{H}_d)$ of the Hilbert space of linear operators on the state-space of the system. Elements of $\mathcal{S}_d$ can be represented as $(d^2 \times d^2)$ matrices that act on the vectorized density matrix $\hat{\hat{\rho}}$, which may be represented as a $(d^2 \times 1)$ column vector.

Definition 2.2.2. *Vectorization*: Vectorization is a transformation that converts matrices $\rho \in \mathcal{O}_d(\mathcal{H}_d)$ into $(d^2 \times 1)$ column vectors $\hat{\hat{\rho}}$. The vector space of $(d^2 \times 1)$ column vectors is referred to as Liouville space. Here, we follow the row-stacking convention

$$\hat{\hat{\rho}} = \mathrm{vec}(\rho) = \begin{pmatrix} \rho_{00} \\ \rho_{01} \\ \rho_{10} \\ \rho_{11} \end{pmatrix} \qquad \text{where } \rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}$$

The solution to the Liouville equation is obtained by direct integration of the vectorized equation

$$\frac{d}{dt}\hat{\hat{\rho}}(t) = -\left(H(t)\otimes\mathbb{I}_d^T - \mathbb{I}_d\otimes H(t)^T\right)\hat{\hat{\rho}}(t) = \mathcal{L}(t)\hat{\hat{\rho}}(t)$$

as

$$\hat{\hat{\rho}}(t) = \mathcal{T}\exp\left[-i\int_0^t \mathcal{L}(t')dt'\right]\hat{\hat{\rho}}(0) = \mathcal{T}\exp\left[-i\int_0^t \left(H(t')\otimes\mathbb{I}_d^T - \mathbb{I}_d\otimes H(t')^T\right)dt'\right]\hat{\hat{\rho}}(0)$$

Here, we have used the vector triple product identity $\mathrm{vec}(ABC) = A\otimes C^T\mathrm{vec}(B)$ to write the matrix representation of the superoperator $\mathcal{L}$. Since, $H(t')\otimes\mathbb{I}_d^T$ and $\mathbb{I}_d\otimes H(t')^T$ commute for all times t' and $f(A\otimes\mathbb{I}_d) = f(A)\otimes\mathbb{I}_d$ for any analytic function f, we obtain a simplified expression for the propagation superoperator $\mathcal{U}(t)$

$$\begin{aligned}\hat{\hat{\rho}}(t) &= \left(\mathcal{T}e^{-i\int_0^t H(t')dt'}\otimes\mathbb{I}_d\right)\left(\mathbb{I}_d\otimes\mathcal{T}e^{i\int_0^t H(t')^T dt'}\right)\hat{\hat{\rho}}(0)\\ &= \underbrace{\left(\mathcal{T}e^{-i\int_0^t H(t')dt'}\otimes\mathcal{T}e^{i\int_0^t H(t')^T dt'}\right)}_{\mathcal{U}(t)}\hat{\hat{\rho}}(0)\end{aligned}$$

Reversing the vectorization yields an expression for the evolution of the density matrix in terms of the the same general form of the unitary propagator $U(t)$ obtained by solving the Schrödinger equation. The discrete-time unitary evolution of the initial state $\rho(0)$ is thus

$$\rho(t) = \mathcal{U}(t)\hat{\hat{\rho}}(0) = U(t)\rho(0)U^\dagger(t)$$

The density matrix formalism is, however, a much more general tool for describing the evolution of quantum systems than the state-vector formalism derived from the Schrödinger equation as it can also be used to represent non-unitary dynamics. Differential equations governing the evolution of an open subsystem ($\mathcal{H}_S$) can be derived from the Liouville equation by defining a closed composite system consisting of the system and its environment ($\mathcal{H}_T = \mathcal{H}_S\otimes\mathcal{H}_E$) that is subject to unitary dynamics. By decomposing the Hamiltonian operator on the composite state-space $H_T = H\otimes\mathbb{I}_E + \mathbb{I}_S\otimes H_E + \alpha H_I$ into Hamiltonians for the

system $H \in \mathcal{O}(\mathcal{H}_S)$, environment $H_E \in \mathcal{O}(\mathcal{H}_E)$, and interaction $H_I \in \mathcal{O}(\mathcal{H}_T)$, it is possible to formulate the evolution of the total density matrix in terms of the interaction Hamiltonian and interaction strength parameter α. Integration and iterative substitution of the integral solution to the resulting differential equation produce a tractable form of the equation of motion that can be truncated to a given order in the small interaction strength parameter α. Performing a partial trace over the environment variables finally yields a master equation for the evolution of the subsystem density matrix alone, which is commonly written in the Lindblad form

$$\frac{d}{dt}\rho(t) = -\frac{i}{\hbar}\left[H(t), \rho(t)\right] - \sum_i \left(L_i \rho(t) L_i^\dagger - \frac{1}{2}\left\{ L_i^\dagger L_i, \rho(t) \right\} \right) = \left(\mathcal{L}(t) + \mathcal{D}(t)\right) \hat{\rho}(t),$$

where $\{A, B\} = AB + BA$ is the anti-commutator. A complete derivation can be found in [18]. The Lindblad (or jump) operators L_j describe the coupling of the principal system to its environment and give rise to the non-unitary $\mathcal{D}$ portion of the propagation superoperator $\mathcal{S}(t)$, which takes the form

$$\mathcal{S}(t) = \mathcal{T} \exp\left[\int_0^t (\mathcal{L}(t') + \mathcal{D}(t')) dt' \right].$$

The Lindblad propagation superoperator provides a good description of the dynamics of open quantum systems provided that the system and environment are initially uncorrelated, and correlations with the environment decay rapidly relative to the system dynamics (Markovian approximation) [19].

Propagators are an effective tool for describing discrete changes between quantum states at different times; however, the propagators themselves are still time-dependent. Often in quantum computing, it is desirable to present state changes in a time-agnostic manner using the quantum operations formalism, in which state changes are described by linear maps on $\mathcal{O}_d$. Quantum operations therefore provide a more general (not necessarily Markovian) and often simplified language in which to represent the same discrete state changes described by the propagator solutions to master equations.

2.2.3 Quantum Operations Formalism

The quantum operations formalism is a powerful tool for describing discrete state changes in quantum systems, including evolution and measurement. In this formalism, quantum operations, or quantum channels, are represented by linear maps $\mathcal{E} : \mathcal{O}_d \to \mathcal{O}_d$. Quantum channels (operations) represented by completely positive trace-preserving (CPTP) maps provide the most general language with which to describe evolution; they can be used to model the same non-unitary dynamics of open quantum systems as the master equation formalism and do not depend on time.

Definition 2.2.3. *Positivity*: A positive map $\mathcal{E} : \mathcal{O}_d \to \mathcal{O}_d$ is positive if for any positive operator $\rho \in \mathcal{O}_d$, $\mathcal{E}(\rho)$ is positive.

Definition 2.2.4. *Complete Positivity*: Consider the extension $\mathcal{E}'$ of a map $\mathcal{E} : \mathcal{O}_{d_A}(\mathcal{H}_A) \to \mathcal{O}_{d_A}(\mathcal{H}_A)$ to the operator space of a composite system $\mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_B$ is the state space of an ancillary system of arbitrary dimension d_B. The map $\mathcal{E}$ is completely positive if the extension $\mathcal{E}'(\rho') = (\mathcal{E} \otimes \mathbb{I}_{d_B})(\rho')$ given by the tensor product of $\mathcal{E}$ with the identity operator $\mathbb{I}_{d_B}$ on $\mathcal{H}_B$ is positive for all positive $\rho' \in \mathcal{O}_{d_A d_B}(\mathcal{H}_A \otimes \mathcal{H}_B)$.

Definition 2.2.5. *Trace-Preserving*: A map is a trace-preserving map if $\mathrm{Tr}(\mathcal{E}(\rho)) = \mathrm{Tr}(\rho)$ for all ρ.

2.2.3.1 Kraus Operator Representation of Quantum Operations

The non-unitary evolution of open quantum systems can be ascribed to interactions between the principal system and the environment. If the system of interest is redefined as the closed composite system-environment space that contains all these interactions, it is thus possible to represent the evolution of the composite quantum state by unitary dynamics ($U \in \mathcal{O}(\mathcal{H}_S \otimes \mathcal{H}_E)$). Assuming that the initial state of the system and environment are uncorrelated (ie. the composite system is in an initial state $\rho \otimes \rho_E$), the quantum operation

$\mathcal{E}(\rho)$ on the system alone is defined using the partial trace over the environment s

$$\mathcal{E}(\rho) = \mathrm{Tr}_E \left[U(\rho \otimes \rho_E) U^\dagger \right]$$

The CPTP map $\mathcal{E}(\rho)$ also has a convenient representation in terms of operators $A_k \in \mathcal{O}(\mathcal{H}_S)$ that act on the state space of the principal system alone.

$$\mathcal{E}(\rho) = \sum_k \langle e_k | \, U(\rho \otimes |e_0\rangle \langle e_0|) U^\dagger \, |e_k\rangle$$

$$= \sum_k A_k \rho A_k^\dagger$$

Here, $|e_k\rangle$ is a complete basis for the environment, and the initial state of the environment $\rho_E = |e_0\rangle \langle e_0|$ is chosen to be a pure state without loss of generality since any mixed state of the environment can be purified without effect on the dynamics of the principal system. The operators $A_k = \langle e_k | U | e_0 \rangle$ are Kraus operators, and the representation of $\mathcal{E}(\rho)$ is known as the Kraus operator or operator-sum representation. The Kraus operators satisfy the completeness constraint

$$\sum_k A_k^\dagger A_k = \mathbb{I}$$

which ensures that the map $\mathcal{E}$ is trace-preserving. The Choi-Kraus representation theorem proves that any completely positive trace-preserving map has a Kraus operator representation, and conversely, any map that can be decomposed into an operator sum representation is completely positive and trace-preserving [14].

Clearly, the unitary propagation superoperator $\mathcal{U}(\rho)$ previously presented satisfies these requirements

$$\mathcal{U}(\rho) = U(t)\rho(0)U(t)^\dagger \qquad \text{and } U(t)^\dagger U(t) = \mathbb{I}$$

Such a map is, in fact, also unital (i.e. it preserves the identity element $\mathbb{I}_d \in \mathcal{O}_d$ of the vector space $\mathcal{U}(\mathbb{I}_d) = \mathbb{I}_d$) because it satisfies the additional requirement that $\sum_k A_k A_k^\dagger = \mathbb{I}$.

The Lindblad propagation superoperator $\mathcal{S}(\rho)$ can also be given an operator sum repre-

sentation [18, 19], for which the time dynamics of the Kraus operators are known.

$$(\mathcal{L} + \mathcal{D})(\rho) = \sum_k A_k(t)\rho(0)A_k(t)^\dagger$$

The more complete knowledge of the time dynamics provided by this picture comes at the cost of an additional assumption of Markovianity for system-environment interactions.

2.3 Measurement

Quantum measurements are described by a set of measurement operators $\{M_m\}$ acting on the state space of the system being measured, where m indexes the possible outcomes of the measurement. The probability of obtaining a given outcome for a prepared state ρ is

$$p(m) = \mathrm{Tr}[M_m^\dagger M_m \rho]$$

where the operators satisfy a completeness relation

$$\sum_m M_m^\dagger M_m = \mathbb{I} \text{ such that } \sum_m p(m) = 1.$$

The post-measurement state of the system is

$$\rho_m = \frac{M_m \rho M_m^\dagger}{p(m)}$$

General measurements in quantum mechanics are described by positive operator-valued measures (POVMs) that consist of a set of positive operators $\{E_m\}$, whose elements are defined in terms of the measurement operators M_m as

$$E_m \equiv M_m^\dagger M_m$$

The POVM $\{E_m\}$ is sufficient to describe the probability of all possible measurement outcomes $p(m) = \mathrm{Tr}[E_m\rho]$ and also satisfies the completeness relation necessary for $\{p(m)\}$ to be a properly normalized probability distribution.

The POVM formalism simplifies in the case that a measurement is described by a complete set of orthonormal projectors $P_m = |\psi_m\rangle \langle\psi_m|$ onto the state space of the system. Such sets $\{M_m\} = \{P_m\}$ describe observables (M), which are Hermitian operators on the state space of the system and thus admit a spectral decomposition.

$$M = \sum_m m P_m.$$

The eigenvalues of the operator correspond to the observable outcomes of the measurement described by $\{P_m\}$. By the orthonormality of the projectors $P_i P_j = \delta_{ij} P_i$, the set of POVM elements $\{E_m\} = P_m^\dagger P_m = P_m$ are the same the set measurement operators themselves. This type of measurement is known as projective measurement.

2.4 Distance Measures for Quantum Information

Characterizing the accuracy with which quantum states are prepared or information is transmitted through quantum channels requires the introduction of distance measures for the distinguishability of quantum states. Distance measures fall into two broad classes: static measures, which quantify the similarity between two quantum states, and dynamic measures, which quantify the preservation of information under some quantum operation.

Describing distinguishability measures as distances implies that these functions are metrics or closely related to metrics on $\mathcal{O}_d$.

Definition 2.4.1. *Metric*: a metric on a set X is a function on that set $d : X \times X \to \mathbb{R}$ for which the following axioms hold

(i) *non-negativity:* $d(x, y) \geq 0, \quad \forall\ x, y \in X,$

(ii) *identity of indiscernibles:* $d(x, y) = 0 \iff x = y$

(iii) *symmetry:* $d(x,y) = d(y,x), \quad \forall\ x, y \in X$

(iv) *triangle inequality:* $d(x,y) \leq d(x,z) + d(z,y), \quad \forall\ x, y, z \in X$

2.4.1 Static Distance Measures

Two of the most commonly used static distance measures for quantum information are the trace distance and the fidelity. The trace distance is itself a metric on $\mathcal{O}_d$, and although the fidelity is not, it readily gives rise to a valid metric and possesses other useful properties that motivate its use [14]. Both of these measures have classical analogues that are distance measures on classical probability distributions, and indeed, as the quantum measures can be related to classical measures on the eigenvalue spectrum of the operators, it is instructive to introduce both.

Definition 2.4.2. *Classical Trace Distance*: The classical trace distance $\mathcal{D}_\mathcal{C}$ between two probability distributions $\{p_x\}$ and $\{q_x\}$ over the same set of indices $\{x\}$ is defined as half the l_1 distance between the distributions.

$$\mathcal{D}_\mathcal{C}(\{p_x\}, \{q_x\}) = \frac{1}{2} \sum_x |p_x - q_x|$$

From this expression, the relationship $\mathcal{D}_\mathcal{C}(\{p_x\}, \{q_x\}) \geq \frac{1}{2} \left(\sum_x p_x - \sum_x q_x\right)$, suggests an equivalent definition for the classical trace fidelity.

$$\mathcal{D}_\mathcal{C}(\{p_x\}, \{q_x\}) = \max_{S \subseteq \{x\}} \left(\sum_{x \in S} p_x - \sum_{x \in S} q_x\right)$$

This offers the intuitive interpretation of the trace distance as the maximized difference between the probability that an event S occurs according to the probability distribution $\{p_x\}$ as opposed to the same event occurring according to the probability distribution $\{q_x\}$.

Definition 2.4.3. *Quantum Trace Distance*: The trace distance $\mathcal{D}$ for quantum mechanical states ρ and σ is defined as half the trace norm of the difference of the density

matrices

$$\mathcal{D}(\rho,\sigma) = \frac{1}{2}\mathrm{Tr}\,|\rho-\sigma|$$

where the absolute value for $A \in \mathcal{O}_d$ is defined per usual as $|A| = \sqrt{A^\dagger A}$. $\mathcal{D}$ satisfies all the properties of a metric and additionally is invariant under unitary transformations

$$\mathcal{D}(\rho,\sigma) = \mathcal{D}(U\rho U^\dagger, U\sigma U^\dagger).$$

In analogy with the classical trace distance, the quantum mechanical trace distance can be defined as the maximization over all positive operators or projectors $P \leq \mathbb{I}$

$$\mathcal{D}(\rho,\sigma) = \max_P \ \mathrm{Tr}(P(\rho-\sigma)).$$

The quantum trace distance is, therefore, the maximum difference in probability that a measurement outcome corresponding to a POVM element P occurs for state σ as opposed to state ρ. From this, it is apparent that the probability distribution of observing the eigenvalues of P for the states ρ and σ should also be maximally separated, which allows the quantum mechanical and classical trace distances to be related as follows.

$$\mathcal{D}(\rho,\sigma) = \max_{\{E_m\}} \ \mathcal{D}_\mathcal{C}(p_m, q_m)$$

Here, the maximization is performed over the set of POVM elements $\{E_m\}$ corresponding to probability distributions of measurement outcomes $p_m = \mathrm{Tr}(\rho E_m)$ and $q_m = \mathrm{Tr}(\sigma E_m)$.

Definition 2.4.4. *Classical Fidelity*: The classical fidelity $\mathcal{F}_\mathcal{C}$ of two probability distributions $\{p_x\}$ and $\{q_x\}$ over the same set of indices $\{x\}$ is

$$\mathcal{F}_\mathcal{C}(\{p_x\},\{q_x\}) = \sum_x \sqrt{p_x q_x}$$

The fidelity is nonzero, and is in fact maximized, in the case $\{p_x\} = \{q_x\}$. It is therefore not a valid metric by condition (ii) of the definition of a metric, though it satisfies the remaining

requirements.

Definition 2.4.5 *Quantum Mechanical Fidelity*: The fidelity of two quantum states ρ and σ is

$$\mathcal{F}(\rho,\sigma) = \mathrm{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}$$

The quantum fidelity also satisfies all properties of a metric except condition (ii) and is invariant under unitary transformations.

$$\mathcal{F}(U\rho U^\dagger, U\sigma U^\dagger) = \mathcal{F}(\rho,\sigma) = \mathrm{Tr}\sqrt{\sqrt{U\rho U^\dagger}(U\sigma U^\dagger)\sqrt{U\rho U^\dagger}} =$$

$$\mathrm{Tr}\sqrt{U\sqrt{\rho}U^\dagger U\sigma U^\dagger U\sqrt{\rho}U^\dagger} = \mathrm{Tr}\left(U^\dagger U\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right) = \mathrm{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}.$$

Here we have made repeated use of the positive operator identity $\sqrt{UAU^\dagger} = U\sqrt{A}U^\dagger$ and the cyclic property of the trace.

If one of the quantum states is a pure state $\rho = |\psi\rangle\langle\psi|$, the fidelity formula may be further simplified to the square root of the overlap of σ and $|\psi\rangle$.

$$\mathcal{F}(|\psi\rangle\langle\psi|,\sigma) = \mathrm{Tr}\sqrt{\langle\psi|\sigma|\psi\rangle|\psi\rangle\langle\psi|} = \sqrt{\langle\psi|\sigma|\psi\rangle}.$$

As in the case of the trace-distance, it is possible to relate the classical and quantum fidelity by considering the probability distributions induced by a measurement:

$$\mathcal{F}(\rho,\sigma) \leq \min_{\{E_m\}} \mathcal{F}_{\mathcal{C}}(p_m, q_m),$$

where $\{E_m\}$ is again a complete set of POVM elements and $p_m = \mathrm{Tr}(\rho E_m)$ and $q_m = \mathrm{Tr}(\sigma E_m)$ are the probability distributions corresponding to measurement with $\{E_m\}$.

2.4.2 Dynamic Distance Measures

Dynamic distance measures describe how well a quantum operation preserves information. Commonly, these measures are framed in terms of the fidelity of quantum channels rather

than the trace distance. One way of characterizing the channel fidelity is to use the static fidelity to determine the distance between the state prepared by the quantum channel $\mathcal{E}_{\mathcal{U}}(\rho) = \mathcal{E} \circ \mathcal{U}$, where $\mathcal{E}$ is the noise operator associated with the channel and the state that would be prepared by the ideal implementation of the channel $\mathcal{U}(\rho)$. However, the channel fidelity should ideally describe the performance of an operation independent of any particular initial state. This can be accomplished by defining an ensemble-average fidelity over a set of initial states $\{\rho_j\}$, prepared with probability p_j.

Definition 2.4.6 *State-Averaged Channel Fidelity*: The channel fidelity is a dynamic fidelity measure that describes the ensemble-average performance of a quantum operation $\mathcal{E}_{\mathcal{U}}$ relative to the ideal unitary superoperator $\mathcal{U}$. Note that the ensemble $\{p_j, \rho_j\}$ should be selected to appropriately sample the state space of possible initial states

$$\overline{\mathcal{F}_{\mathcal{E}_{\mathcal{U}},\mathcal{U}}} = \sum_j p_j \left[\mathcal{F}(\mathcal{U}(\rho_j), \mathcal{E}(\rho_j))\right]^2 = \sum_j p_j \left(\mathrm{Tr}\sqrt{\sqrt{\mathcal{E}_{\mathcal{U}}(\rho)}\mathcal{U}(\rho)\sqrt{\mathcal{E}_{\mathcal{U}}(\rho)}}\right)^2.$$

By the invariance of the fidelity under unitary transformations, this is equivalent to the channel fidelity of the noise operator $\mathcal{E}$ to the identity $\mathcal{I}$

$$\overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}} = \sum_j p_j \left(\mathrm{Tr}\sqrt{\sqrt{\mathcal{E}(\rho)}\mathcal{I}(\rho)\sqrt{\mathcal{E}(\rho)}}\right)^2.$$

The role of the squared fidelity in the definition of the channel fidelity is not immediately obvious but can be understood by considering the second commonly used measurement of dynamic fidelity – entanglement fidelity. The entanglement fidelity is premised on the idea that a quantum operation that preserves information well also preserves entanglement. If instead of defining the fidelity by comparing the action of a quantum channel's noise operator $\mathcal{E}$ on an initial state ρ to the action of the identity $\mathcal{I}$ on that same state, we instead consider that $\mathcal{E}$ and $\mathcal{I}$ act replicas of the initial state ρ and $\tilde{\rho}$, it becomes evident that $\mathcal{E}$ is close to $\mathcal{I}$ if it preserves the initial entanglement of ρ and $\tilde{\rho}$. More generally, $\mathcal{E}$ has high entanglement fidelity if it preserves entanglement with any ancillary system.

The expression for the entanglement fidelity is therefore determined by first purifying the initial state ρ of system A by introducing a fictitious system R such that the composite system AR is in a pure state.

Definition 2.4.7 *Entanglement Fidelity* The entanglement fidelity is a dynamic fidelity measure that depends on the extent to which a system of interest A in an initial state ρ and a reference system R remain entangled after A is acted upon by the noise channel $\mathcal{E}$ and R by the identity $\mathcal{I}$.

$$\mathcal{F}_E(\mathcal{E},\mathcal{I}) = \langle AR| \left[(\mathcal{E}\otimes\mathcal{I}) |AR\rangle \langle AR| \right] |AR\rangle$$

The right hand side is equal to the square of the static fidelity

$$\mathcal{F}(|AR\rangle\langle AR|, (\mathcal{E}\otimes\mathcal{I})(|AR\rangle\langle AR|) = \sqrt{\langle AR|(\mathcal{E}\otimes\mathcal{I})(|AR\rangle\langle AR|)|AR\rangle}$$

between the initial state of the composite system $|AR\rangle\langle AR|$ and the final state $(\mathcal{E}\otimes\mathcal{I})(|AR\rangle\langle AR|)$.

The entanglement fidelity therefore provides a useful interpretation of the "information source" that is preserved under high fidelity quantum operations as the entanglement between a system and its environment. In the remainder of this thesis, however, we exclusively employ the channel fidelity since that is the more commonly used dynamic fidelity measure in randomized benchmarking protocols.

2.4.3 Uniform Sampling and the Haar Measure

In order to obtain the ensemble average channel fidelity introduced in the previous section, it is necessary to define a measure μ_{FS} to uniformly sample the complex projective space $\mathbb{CP}^{d-1}$ of pure states $|\psi\rangle\langle\psi|$. This measure is the Fubini-Study measure μ_{FS}, which is the unique unitarily invariant measure on pure states [20].

The state-averaged channel fidelity between an ideal gate described by the unitary superoperator $\mathcal{U}$ and the noisy implementation described by a quantum channel formed by the

composition $\mathcal{E}_{\mathcal{U}} = \mathcal{E} \circ \mathcal{U}$ can then be written as an integral over μ_{FS} [15].

$$\overline{\mathcal{F}_{\mathcal{E}_{\mathcal{U}},\mathcal{U}}} = \int_{\mathbb{CP}^{d-1}} \mathrm{Tr}\sqrt{\sqrt{\mathcal{E}_{\mathcal{U}}(|\psi\rangle\langle\psi|)}\,\mathcal{U}(|\psi\rangle\langle\psi|)\,\sqrt{\mathcal{E}_{\mathcal{U}}(|\psi\rangle\langle\psi|)}}\,d\mu_{FS}.$$

From the unitary invariance of the fidelity, this is equal to the average fidelity of the noise channel $\mathcal{E}$ to the identity $\mathcal{I}$

$$\overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}} = \int_{\mathbb{CP}^{d-1}} \mathrm{Tr}\left(|\psi\rangle\langle\psi|\,\mathcal{E}(|\psi\rangle\langle\psi|)\right)\,d\mu_{FS}.$$

Uniform sampling of the state space can also be achieved by applying random unitaries U selected from the Haar measure μ_H on a fixed pure state $|\phi\rangle\langle\phi|$ such that $|\psi\rangle\langle\psi| = U(|\phi\rangle\langle\phi|)U^\dagger$. This leads to an equivalent form for the average channel fidelity integrated over the Haar measure on the unitary group of $(d \times d)$ matrices $\mathfrak{U}_d$.

$$\overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}} = \int_{\mathfrak{U}(d)} \mathrm{Tr}\left(U|\phi\rangle\langle\phi|U^\dagger\mathcal{E}(U|\phi\rangle\langle\phi|U^\dagger)\right)d\mu_H$$

2.4.4 Unitary t-Designs

In practice, generating a Haar random unitary is inefficient because the number of gates required grows exponentially with the number of qubits [15]. However, it is possible to exploit the properties of groups of operations known as unitary t-designs to circumvent the generation of Haar random unitaries, resulting in more efficient quantum channel characterization protocols [15, 16].

Definition 2.4.8 *Group*: A group $\mathcal{G}$ is a set of group operations $\{g_j\}$ together with a combinatorial operation $\cdot$ (group multiplication) such that four axioms are satisfied [21]:

(i) *closure*: $\forall g_i, g_j \in \mathcal{G},\ g_i \cdot g_j \in \mathcal{G}$

(ii) *associativity*: $\forall g_i, g_j, g_k \in \mathcal{G}, (g_i \cdot g_j) \cdot g_k = g_i \cdot (g_j \cdot g_k)$

(iii) *identity*: there is a group operation $\mathcal{I} \in \mathcal{G}$ with the property that $g_i \cdot I = I \cdot g_i = g_i$

(iv) *inverse*: $\forall g_i \in \mathcal{G}$ there exists an inverse operation $g_i^{-1} \in \mathcal{G}$ such that $g_i \cdot g_i^{-1} = g_i^{-1} \cdot g_i = I$

A group of operation elements is a unitary t-design if it satisfies the following definitiion.

Definition 2.4.9 *Unitary t-Design*: A unitary t-design is a finite set of unitary operators $\{U_k\} \subset \mathfrak{U}(d)$ on the d-dimensional field of complex numbers $\mathbb{C}^d$ such that for every polynomial $P_{(t,t)}(U)$ of degree at most t in the matrix elements of U and their complex conjugates

$$\frac{1}{K}\sum_{k=1}^{K} P_{(t,t)}(U_k) = \int_{\mathfrak{U}(D)} d\mu_H P_{(t,t)}(U)$$

where $d\mu_H$ denotes the Haar measure on the unitary group $\mathfrak{U}(d)$.

2.4.4.1 Unitary 2-Design Twirl

Definition 2.4.10 *Twirl*: The twirl of a quantum channel $\mathcal{E}$ is the average superoperator $\mathcal{E}$ under conjugation by randomly chosen unitary superoperators $\mathcal{U}$ sampled according to the probability measure μ (typically chosen to be μ_H)

$$\overline{\mathcal{E}}(\rho) = \int_{\mathfrak{U}(d)} \mathcal{U}^\dagger \circ \mathcal{E} \circ \mathcal{U}(\rho)\ d\mu$$

Here, $\mathcal{U}^\dagger \circ \mathcal{E} \circ \mathcal{U}$ denotes the composition $\circ$ of operations such that U is applied first, followed by $\mathcal{E}$, and then by $U^\dagger$. Using this ordering and the Kraus operator representation of the unitary superoperator $\mathcal{U}(\rho) = U\rho U^\dagger$, the definition of the twirl takes a form suggestive of its relationship to the average channel fidelity [15, 16]

$$\overline{\mathcal{E}}(\rho) = \int_{\mathfrak{U}(d)} U^\dagger \mathcal{E}\left(U\rho U^\dagger\right) U d\mu.$$

The probability measure over which twirling is performed is typically chosen to be the Haar measure since it can be proven that the channel has a simple representation as a depolarizing

channel in this case [22]

$$\overline{\mathcal{E}}(\rho) = \int_{\mathfrak{U}(D)} \mathcal{U}^\dagger \circ \mathcal{E} \circ \mathcal{U}(\rho)\ d\mu_H = p\rho + (1-p)\frac{\mathbb{I}_d}{d}.$$

Here p is the depolarizing parameter that governs the probability with which the quantum channel replaces the input state ρ with the maximally mixed state $\mathbb{I}_d$. The depolarizing channel fidelity is also a simple function of the depolarizing parameter and dimension of the system

$$\overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}} = p + \frac{1-p}{d}.$$

The equivalence between the twirl of superoperators over the Haar measure and depolarizing channels is the foundation of all randomized benchmarking (RB) methods. The relationship between the average gate fidelity obtained by RB and the depolarizing channel fidelity is discussed in greater detail in chapter 3.

One more twirling identity is, however, required. The above definition of the twirl still relies on unitaries sampled from the Haar measure, which is exponentially inefficient. For unitary 2-designs, the discrete twirl over a uniformly sampled, finite set of unitaries $\{U_k\}$ is equal to the twirl over the full Haar measure on the unitary group $\mathfrak{U}(d)$ [15].

$$\frac{1}{K}\sum_k^K U_k^\dagger \circ \mathcal{E} \circ U_k(\rho) = \int_{\mathfrak{U}(d)} U^\dagger \circ \mathcal{E} \circ \mathcal{U}(\rho)\ d\mu_H.$$

Random circuits of elements selected from a unitary 2-design thus provide an efficient twirling protocol with which to implement RB.

2.4.5 The Clifford Group

The Clifford hierarchy consists of groups of operations important to the theory of quantum error-correcting codes and fault-tolerant computation [23]. In this thesis, the first two levels of this hierarchy, the Pauli and Clifford groups, are of significance.

Definition 2.4.11 *Pauli Group*: The n-qubit Pauli group $\mathfrak{P}_n$ is a unitary 1-design with

a set of group operators $\mathfrak{P}_n = \{X, Y, Z, I\}^{\otimes n} \times \{\pm 1, \pm i\}$ consisting of the n-fold tensor product of elements of the single-qubit Pauli group $\mathfrak{P}_1$ along with multiplicative phases. The elements of $\mathfrak{P}_1 = \{X, Y, Z, I\} \times \{\pm 1, \pm i\}$ have representations as 2×2 matrices where

$$X = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad Y = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad Z = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \mathbb{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Definition 2.4.12 *Clifford Group*: The n-qubit Clifford group $\mathfrak{C}_n$ is a unitary 2-design consisting of the normalizers of the Pauli group: an operator C of dimension $d = 2^n$ is an element of the $\mathfrak{C}_n$ if and only if $CPC^\dagger \in \mathfrak{P}_n$ for all $P \in \mathfrak{P}_n$.

A group can be described in terms of its generating set, which is a subset of group operations such that any operation in the group can be constructed by group multiplication of elements in the subset. For the single-qubit Clifford group $\mathfrak{C}_1$, the minimal generating set consists of the Hadamard H and phase gate S. For higher dimensions $n \geq 2$, an additional CNOT gate is required, and elements of $\mathfrak{C}_n$ are formed as n-fold tensor products of the generating gates $\{H, S, CNOT\}$.

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad S = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix} \qquad CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The Clifford group is possessed of several properties that make it a suitable set of operations with which to implement randomized benchmarking methods. First and foremost, it is a unitary 2-design, so the discrete twirl over elements of $\mathfrak{C}_n$ is equal to the full twirl over the Haar measure. Sampling from the discrete group eliminates the inefficiency associated with Haar-random unitary generation. The evolution of quantum states under the action of the Clifford group can also be efficiently simulated on a classical computer, as proven in the Gottesman-Knill theorem [23]. Simulated results of Clifford-based randomized benchmarking under particular noise models provide a valuable reference for analyzing the types

of error processes present in quantum information processors [8]. Moreover, though the Clifford group is not itself a basis for universal quantum computation, arbitrary unitaries can be generated by the addition of a single gate (ex. $\pi/8$ or Toffoli gates) outside the gateset. Typical fault-tolerant architectures based on stabilizer codes are dominated by Clifford gates and are most sensitive to errors in the elementary elements of $\mathfrak{C}_n$. Characterizing the average error associated with Clifford gates, therefore, provides a good estimate of the performance of the universal gateset as a whole.

2.4.6 DiVincenzo Criteria

The centrality of the Clifford group to universal quantum computation and error correction schemes is a defining feature of this group. More generally, any quantum information processing architecture and control scheme may be evaluated according to several criteria that must be satisfied to implement universal quantum computation.

DiVincenzo's criteria explicitly enumerate the requirements of universal quantum computation and provide a concrete rubric with which to identify developmental targets for NISQ architectures [24]. The five Divincenzo criteria for universal computation are as follows. First, a QIP must be a scalable physical system with qubits (quantum two-level systems) whose physical parameters are well-characterized. Physical parameters are considered to be well characterized if the single-qubit Hamiltonian and coupling Hamiltonians for interactions with other qubits or external fields are accurately known. Second, it must be possible to initialize the system in a well-defined fiducial state such that the qubits are in a known state at the start of the computation. Third, decoherence, which arises due to system-environment interactions in the open quantum system, should occur on a much longer timescale than the gate operation time. Fourth, it must be possible to implement a set of universal quantum gates, which are control operations capable of implementing sequences of arbitrary unitary transformations that each act on a small number of qubits. Error correction schemes require these control operations to be implementable in parallel such that gate operations involving a finite fraction of the qubits may be performed simultaneously. Additionally, the

combined systematic and random error rate per gate operation must be constrained below a threshold value in the vicinity of 10^{-4} - 10^{-5}, depending on the noise model. Finally, qubit-specific measurement must be possible so that the state of individual qubits can be accurately determined.

In this thesis, we focus primarily on the fourth criterion and attempt to reduce error rates associated with quantum operations. However, it should be noted that superconducting architectures already rank highly with respect to many of the DiVincenzo criteria due to their scalability, well-characterized tunable coupling schemes, long coherence times, and mature microwave electronics for qubit-specific measurement and control. Incremental improvements in the fidelity of quantum operation therefore represent substantive progress towards universal quantum computation.

CHAPTER 3

Randomized Benchmarking

Accurate characterization of noise in quantum operations is an essential step towards fault-tolerant computation. Threshold theorems for a variety of noise models [25, 26] promise that arbitrarily large quantum computations may be performed efficiently provided that the error associated with individual gates is below a given rate. Quantum process tomography (QPT) is one traditional method of noise characterization that strives to completely reconstruct the cumulative noise operator associated with a process [27, 28]. This is achieved by tomographic measurement of the output of that process acting on a set of (d^2) input states that span the operator space $\mathcal{O}_d$ on $\mathcal{H}_d$, and the number of experiments therefore scales exponentially (d^{2n}) with the number of qubits (n), preventing the application of QPT to the characterization of larger systems [22]. Additionally, the process matrix in QPT absorbs state preparation and measurement (SPAM) errors such that the infidelity associated with the target transformation can not be separately determined. Randomized benchmarking (RB) methods present an alternative approach to partial noise characterization that is scalable and insensitive to SPAM errors [29].

RB methods exploit properties of unitary t-designs to efficiently obtain information about the average superoperator of a noise channel. Provided the noise is Markovian and the variance between the noise associated with individual elements in the unitary t-design is within certain bounds, which are explored in detail in [30, 29], these methods well characterize the performance of a gateset. Extensions of randomized benchmarking methods have also been developed: leakage randomized benchmarking (LRB) allows leakage out of the computational subspace to be estimated, and interleaved randomized benchmarking (IRB) allows the fidelity of individual gates within the gateset to be determined.

3.0.1 Randomized Benchmarking Protocol

Standard randomized benchmarking provides information about the average fidelity of a quantum channel under the action of a set of unitaries. The Haar average fidelity may be obtained efficiently by taking advantage of the twirling identity introduced in section 2.4.4.1 for a set of gates whose members are elements of a unitary 2-design $\mathcal{G}$. Frequently, $\mathcal{G}$ is designated to be the Clifford group $\mathfrak{C}_n$ since this group is a unitary 2-design, which can be made into a basis for universal quantum computation through the addition of a single gate outside the group [29]. The noisy implementation of the Clifford elements can be written as a composition $\mathcal{S} = \mathcal{E} \circ \mathcal{C}$ of the ideal Clifford element $\mathcal{C}$ with a noise channel $\mathcal{E}$. The RB protocol introduced in [29] and outlined here then allows the characterization of the average noise channel as follows.

1. Choose an array of sequence lengths $m \in \mathbb{N}$ over which to repeat the procedure.

2. Construct a sequence $\mathbf{i_m} = \{\mathcal{C}_{i_1}, \mathcal{C}_{i_2} ... \mathcal{C}_{i_m}, \mathcal{C}_{i_{m+1}}\}$ with m elements randomly selected from $\mathcal{G}$ and an $(m+1)$-th element that is the inverse $\mathcal{C}_{m+1} = \mathcal{C}_1^\dagger \circ \ldots \circ \mathcal{C}_m^\dagger$ of the sequence $\{\mathcal{C}_{i_1}, \mathcal{C}_{i_2} ... \mathcal{C}_{i_m}\}$. In the absence of noise, $\mathbf{i_m}$ is thus the identity operation. The noisy sequence for a noise channel $\mathcal{E}_{i_j,j}$, which may be time (j) and gate $(\mathcal{C}_{i_j})$ dependent, is then

$$\mathcal{S}_{\mathbf{i_m}} = \bigcirc_{j=1}^{m+1} \left(\mathcal{E}_{i_j,j} \circ \mathcal{C}_{i_j} \right)$$

3. Compute the survival probability $\mathrm{Tr}\left[E_\psi \mathcal{S}_{\mathbf{i_m}}(\rho_\psi)\right]$ for $\mathbf{i_m}$, where E_ψ is the POVM element corresponding to the projector onto the initial state in the presence of measurement errors, and ρ_ψ is the density matrix for the initial pure state (typically $|0\rangle\langle 0|$) in the presence of state-preparation errors.

4. Repeat steps *2* & *3* for k independent random sequences $\{\mathbf{i_m}\}$ to determine the average survival probability over the set of sequences

$$\overline{F}_{\mathrm{seq}}(m, \psi) = \mathrm{Tr}\left[E_\psi \mathcal{S}_m(\rho_\psi)\right]$$

where $\mathcal{S}_m$ is the sequence operation $\mathcal{S}_{\mathbf{i_m}}$ averaged over k realizations of the sequence.

$$\mathcal{S}_m = \frac{1}{|\{\mathbf{i_m}\}|} \sum_{\mathbf{i_m}}^{|\{\mathbf{i_m}\}|} \mathcal{S}_{\mathbf{i_m}}$$

5. Fit the average sequence fidelity to either the zeroth or first order model in the gate-dependence of the noise channel.

$$\overline{F}^{(0)}_{\text{seq}}(m, \psi) = A_0 p^m + B_0$$

$$\overline{F}^{(1)}_{\text{seq}}(m, \psi) = A_1 p^m + C_1 (m-1)(q-p^2) p^{m-2} + B_1$$

The coefficients $A_0, B_0, A_1, B_1,$ and C_1 absorb state preparation and measurement errors, and the depolarizing parameter p is related to the strength of the noise. The average error rate r per gate is one minus the depolarizing channel fidelity

$$r = 1 - p - \frac{(1-p)}{d}.$$

3.0.2 Derivation of the Randomized Benchmarking Model

The full derivation of the standard RB protocol is presented in [29], and restrictions on the validity of the zeroth and first order models are elaborated in [30]. Here, we present a brief overview of the connection between the average sequence fidelity and the fidelity of a depolarizing channel.

The relationship between the random sequence $S_{\mathbf{i_m}}$ and the twirl over $\mathcal{G}$ can be derived by repeated applications of the identity operation $\mathcal{C}_{i_j} \circ \mathcal{C}_{i_j}^{\dagger}$. In the simplified case of gate-independent noise, the noise at each step $\mathcal{E}_{i_j,j}$ is equal to the average noise channel $\mathcal{E}$.

$$\mathcal{E} = \frac{1}{|\{i_j, j\}|} \sum_{i_j,j}^{|\{i_j,j\}|} \mathcal{E}_{i_j,j}$$

The sequence $\mathcal{S}_{\mathbf{i_m}}$ is then $\mathcal{S}_{\mathbf{i_m}} = \bigcirc_{j=1}^{m+1} \mathcal{E} \circ \mathcal{C}_{i_j} = \mathcal{E} \circ \mathcal{C}_{i_{m+1}} \circ \mathcal{E} \circ \mathcal{C}_{i_m} \circ \ldots \mathcal{E} \circ \mathcal{C}_{i_2} \circ \mathcal{E} \circ \mathcal{C}_{i_1}$

$$= \mathcal{E} \circ \mathcal{C}_{i_{m+1}} \circ \mathcal{C}_{i_m} \circ \ldots \mathcal{C}_{i_1} \circ \mathcal{C}_{i_1}^\dagger \ldots \mathcal{C}_{i_m} \circ \mathcal{C}_{i_m}^\dagger \circ \mathcal{E} \circ \mathcal{C}_{i_m} \circ \mathcal{C}_{i_{m-1}} \circ \ldots \mathcal{C}_{i_1} \circ \mathcal{C}_{i_1}^\dagger \ldots \circ \mathcal{C}_{i_{m-1}}^\dagger \circ \mathcal{E}$$

$$\ldots\ldots\ldots \circ \mathcal{E} \circ \mathcal{C}_{i_2} \circ \mathcal{C}_{i_1} \circ \mathcal{C}_{i_1}^\dagger \circ \mathcal{E} \circ \mathcal{C}_{i_1}$$

$$= \mathcal{E} \circ \bigcirc_{j=1}^{m} \mathcal{D}_{i_j}^\dagger \circ \mathcal{E} \circ \mathcal{D}_{i_j}$$

where the conjugation operations $\mathcal{D}_{i_j}^\dagger \circ \mathcal{E} \circ \mathcal{D}_{i_j}$ are defined in terms of new gates $\mathcal{D}_{i_j} = \mathcal{C}_{i_j} \circ \mathcal{C}_{i_{j-1}} \circ \ldots \mathcal{C}_{i_1}$.

The average sequence operation $\mathcal{S}_m$ then takes the form of a twirling operation and can therefore be written as an m-fold composition of depolarizing channels.

$$\mathcal{S}_m = \frac{1}{K} \sum_{i_j} \mathcal{E} \circ \bigcirc_{j=1}^{m} \mathcal{D}_{i_j}^\dagger \circ \mathcal{E} \circ \mathcal{D}_{i_j} = \mathcal{E} \circ \sum_{i_j} \frac{\widetilde{\mathcal{E}}_{i_j}}{K} = \mathcal{E} \circ \mathcal{E}_{twirl}^{\circ m}$$

Here, the abbreviated notation $\widetilde{\mathcal{E}}_{i_j} = \mathcal{D}_{i_j}^\dagger \circ \mathcal{E} \circ \mathcal{D}_{i_j}$ has been introduced for the conjugation operations, and K has been used to denote the number of elements in the finite set $\{\mathcal{D}_{i_j}\}$. Replacing $\mathcal{E}_{twirl}$ with its definition as a depolarizing channel $\mathcal{E}_{dep} = p\rho + (1-p)\frac{\mathbb{I}}{d}$ and calculating the average sequence fidelity yields the zeroth order model.

$$\overline{F}_{\text{seq}}^{(0)}(m, \psi) = \text{Tr}\left[E_\psi \circ \mathcal{E} \circ \mathcal{E}_{twirl}^{\circ m}\right] = \text{Tr}\left[E_\psi \circ \mathcal{E}\left(p^m \rho_\psi + (1-p^m)\frac{\mathbb{I}}{d}\right)\right]$$

$$= \text{Tr}\left[E_\psi \circ \mathcal{E}\left(\rho_\psi - \frac{\mathbb{I}}{d}\right)\right] p^m + \text{Tr}\left[E_\psi \circ \mathcal{E}\left(\frac{\mathbb{I}}{d}\right)\right]$$

The coefficients $A_0 = \text{Tr}\left[E_\psi \circ \mathcal{E}\left(\rho_\psi - \mathbb{I}/d\right)\right]$ and $B_0 = \text{Tr}\left[E_\psi \circ \mathcal{E}\left(\mathbb{I}/d\right)\right]$ absorb the errors associated with measurement and state preparation such that p provides an estimate of the fidelity of the depolarizing channel unaffected by these elements.

In the case that the noise operator is gate-dependent, higher order corrections can be determined by performing a perturbative expansion about the average gate error using $\delta\mathcal{E}_{i_j,j} = \mathcal{E}_{i_j,j} - \mathcal{E}$. The derivation follows the same general procedure with the exception

that different forms of the average sequence operation arise depending on whether the perturbation is applied on the first gate, at a point within the sequence, or on the last gate [29].

3.0.3 Interleaved Randomized Benchmarking

The standard randomized benchmarking protocol only provides information about the average noise channel associated with transformations in the group and cannot be used to characterize individual gates. In order to characterize the fidelity associated with a single gate within the gateset, as we strive to do for our pulse gates, the RB gate sequence can be modified by alternating random Clifford elements $\mathcal{C}_{i,j} \in \mathcal{G}$ with a gate of interest $\mathcal{C} \in \mathcal{G}$ to construct interleaved randomized benchmarking circuits. Note that the recovery operation in IRB is also modified to invert the interleaved sequence $\mathcal{C}_{i_{m+1}} = \mathcal{C}_1^\dagger \circ \mathcal{C}^\dagger \ldots \mathcal{C}_m^\dagger \circ \mathcal{C}^\dagger$ such that the ideal operation performed by the interleaved circuit is still the identity.

Interleaved randomized benchmarking also relies on the properties of the twirl over a unitary 2-design; however, in this case, the twirl is performed over the composition $(\mathcal{E}_\mathcal{C} \circ \mathcal{E}_{i,j})$ of the noise channel for the gate of interest $\mathcal{E}_\mathcal{C}$ with that of a random member of the group $\mathcal{E}_{i,j}$. The quantum operation for each IRB sequence $\mathbf{i_m}$ is the following,

$$\mathcal{V}_{\mathbf{i_m}} = \mathcal{E}_{i_{m+1}} \circ \mathcal{C}_{i_{m+1}} \circ \left(\bigcirc_{j=1}^{m} \left[\mathcal{C} \circ \mathcal{E}_C \circ \mathcal{E}_{i_j} \circ \mathcal{C}_{i_j} \right] \right)$$

which can be rewritten as a composition of conjugation operations and averaged over sequences in the same manner as for standard RB to depolarize $\left(\mathcal{E}_\mathcal{C} \circ \mathcal{E}_{i_j}\right)$. Fitting the average survival probability $\overline{F}_{\text{seq}}$ of the interleaved sequence to the zeroth or first order models yields the depolarizing parameter $p_\mathcal{C}$ for the composite channel. This depolarizing parameter $p_\mathcal{C}$ is then compared to the depolarizing parameter p obtained from a standard randomized benchmarking experiment without the interleaved element to generate an estimate for the

error associated with the interleaved gate.

$$r_{\mathcal{C}} = \frac{(d-1)(1-p_{\mathcal{C}}/p)}{d}$$

For sufficiently small variance in the noise operators of the Clifford elements $\mathcal{E}_{i_j} \approx \mathcal{E}$, the estimated error of the interleaved gate can be bounded to lie in the range $[r_{\mathcal{C}} - E, r_{\mathcal{C}} + E]$ where

$$E = \min \begin{cases} \dfrac{(d-1)\left[|p - p_{\mathcal{C}}/p| + (1-p)\right]}{d} \\ \dfrac{2(d^2-1)(1-p)}{pd^2} + \dfrac{4\sqrt{1-p}\sqrt{d^2-1}}{p} \end{cases}$$

The full derivation of the interleaved gate error and bounds can be found in [31].

3.0.4 Leakage Randomized Benchmarking

Multiple extensions of randomized benchmarking protocols designed to characterize average leakage rates for unitary t-designs have been presented [7, 32, 33, 8]. Early experimental implementations discussed in references [7, 32] require direct measurement of the leakage subspace and rely on phenomenological decay models for their analysis of leakage rates. Subsequent rigorous mathematical treatments of leakage rates [33, 8] validate the general form of the empirical models and yield improved leakage characterization protocols. Of these, the protocol we have chosen to implement is the LRB model of Wood et al. [8], which provides a method of separately estimating leakage rates into and out of the computational subspace, even in the case where the leakage subspace cannot be directly measured. This method also allows simultaneous estimation of the average gateset fidelity.

The aim of the LRB model is to characterize the average error rate for a quantum channel $\mathcal{E}$ that couples a computational subspace χ_1 of dimension d_1 and a leakage subspace χ_2 of dimension d_2. The cumulative state leakage $L(\rho)$ of a density matrix $\rho \in \chi$ on the full direct-sum state space of the system $\chi = \chi_1 \oplus \chi_2$ can be calculated using projectors $\mathbb{I}_1$ and

$\mathbb{I}_2$ onto χ_1 and χ_2 respectively

$$L(\rho) = \mathrm{Tr}\left[\mathbb{I}_2\rho\right] = 1 - \mathrm{Tr}\left[\mathbb{I}_1\rho\right].$$

The noise associated with the quantum channel $\mathcal{E}$ can be characterized by considering the effect of $\mathcal{E}$ on the state leakage of elements in the computational and leakage subspaces $\{|\psi_1\rangle\langle\psi_1|\} \in \mathcal{O}_{d_1}(\chi_1)$ and $\{|\psi_2\rangle\langle\psi_2|\} \in \mathcal{O}_{d_2}(\chi_2)$. The average change in state leakage over elements of χ_1 yields an estimate of the leakage out of the subspace (L_1),

$$L_1(\mathcal{E}) = \int L\left(\mathcal{E}\left(|\psi_1\rangle\langle\psi_1|\right)\right) d\psi_1 = L\left(\mathcal{E}\left(\frac{\mathbb{I}_1}{d_1}\right)\right)$$

and the average over elements of χ_2 yields an estimate of seepage (L_2) back into the computational subspace

$$L_2(\mathcal{E}) = 1 - \int L\left(\mathcal{E}\left(|\psi_2\rangle\langle\psi_2|\right)\right) d\psi_2 = 1 - L\left(\mathcal{E}\left(\frac{\mathbb{I}_2}{d_2}\right)\right).$$

In the above, $d\psi_1$ and $d\psi_2$ indicate uniform selection of states from the Haar measure on χ_1 and χ_2 respectively. The average rates $L_1(\mathcal{E})$ and $L_2(\mathcal{E})$ often provide a good idea of the channel performance in practice, and it is furthermore possible to bound the worst-case leakage rates over all input states in terms of the average rates [33]. A full derivation of the model for the zeroth order approximation $\mathcal{E}_{i_j} = \mathcal{E}$ can be found in [8].

The LRB protocol may be implemented as follows.

1. Choose an array of sequence lengths $m \in \mathbb{N}$ and construct sequences $\mathbf{i_m}$ of m random elements of $\mathcal{G}$ whose $(m+1)$-th elements are defined as the recovery operation $\mathcal{C}_{m+1} = \mathcal{C}_1^\dagger \circ \ldots \circ \mathcal{C}_m^\dagger$.

2. Apply the sequences to an initial pure state ρ_ψ (typically $|0\rangle\langle 0|$) in the computational subspace $\mathcal{O}_{d_1}(\chi_1)$, and perform measurements over all projectors E_j onto the compu-

tational subspace ($j = 0, 1 \ldots d_1 - 1$) to estimate the survival probabilities.

$$\mathcal{P}_j(\mathbf{i_m}) = \mathrm{Tr}\left[E_j \mathcal{S}_{\mathbf{i_m}}(\rho_\psi)\right]$$

Here $\mathcal{S}_{i_m}$ is the quantum operation for the sequence, as defined in standard RB.

- ✱ In the case where high-fidelity measurements are limited to a single two-outcome POVM $\{E_0, \mathbb{I} - E_0\}$, LRB may still be implemented by modifying the recovery operation $\mathcal{C}_{m+1}$ such that the population of state $|j\rangle$ is rotated onto the projective space of E_0. Defining a unitary matrix U_j such that $E_j = U_j E_0 U_j^\dagger = \mathcal{U}_j(E_0)$, we modify the recovery option $\mathcal{C}_{m+1}^{(j)} = \mathcal{U}_j^\dagger \circ \mathcal{C}_{m+1}$ to obtain a new sequence $\mathbf{i_m}^{(j)} = \mathcal{U}_j^\dagger \circ \mathbf{i_m}$. This yields

$$\mathcal{P}_j(\mathbf{i_m}) = \mathrm{Tr}\left[E_j \mathcal{S}_{\mathbf{i_m}}(\rho_\psi)\right] = \mathrm{Tr}\left[E_0 \mathcal{S}_{\mathbf{i_m}}^{(j)}(\rho_\psi)\right]$$

3. Sum the survival probabilities $\mathcal{P}_j$ to estimate the population of χ_1

$$\mathcal{P}_{\mathbb{I}_1}(\mathbf{i_m}) = \sum_j \mathcal{P}_j(\mathbf{i_m}) = \mathrm{Tr}\left[\mathbb{I}_1 \mathcal{S}_{\mathbf{i_m}}(\rho_\psi)\right]$$

4. Repeat steps *1,2,* & *3* for k independent random sequences $\mathbf{i_m}$ to obtain an estimate of the average survival probability $\mathcal{P}_{\mathbb{I}_1}$ of the computational subspace.

$$\mathcal{P}_{\mathbb{I}_1}(m, \psi) = \frac{1}{|\{\mathbf{i_m}\}|} \sum_{\mathbf{i_m}}^{|\{\mathbf{i_m}\}|} \mathcal{P}_{\mathbb{I}_1}(\mathbf{i_m})$$

5. Fit the average survival probability to the decay model

$$\mathcal{P}_{\mathbb{I}_1}(m, \psi) = Cp^m + B$$

and compute the average leakage and seepage rates for the gateset as

$$L_1(\mathcal{E}) = (1 - B)(1 - p)$$

$$L_2(\mathcal{E}) = B(1 - p)$$

6. LRB may additionally be used to estimate the average gateset fidelity by fitting the average survival probability for the initial pure state, here taken to be $|0\rangle\langle 0|$, to the model $\mathcal{P}_0(m, \psi) = B_0 + C_0 p_1^m + A_0 p_2^m$. Here, the value of p_1 is fixed by the value obtained in step *3*, and the coefficients obey the relationships $0 \leq B_0 \leq B$, $0 \leq A_0 \leq 1$, and $0 \leq A_0 + B_0 + C_0 \leq 1$. The average gate error is then given by

$$r = 1 - \frac{1}{d_1}\left[(d_1 - 1)p_2 + 1 - L_1\right]$$

For weak leakage, this simplifies to the standard RB decay model $\mathcal{P}_0(m, \psi) = B_0 + A_0 p_2^m$.

CHAPTER 4

Quantification of Robustness, Fidelity, and Leakage for Composite and Adiabatic Gates

4.1 Introduction

In the last two decades, superconducting qubits have emerged as a promising architecture to achieve large-scale, fault-tolerant quantum computation. Improvements to quantum circuit design have dramatically increased intrinsic coherence times from nanoseconds to well over 100 μs on modern devices [6]. As coherence times improve, the depth of quantum circuits can be increased; however, longer sequences of operations lead to the rapid accumulation of pulse errors. Large cumulative errors may result in erroneous circuit operation. It has therefore become increasingly important to minimize control errors such that error rates are kept below fault-tolerance thresholds. High fidelity operations for optimally calibrated microwave pulses have already been implemented on superconducting systems, with average single-qubit gate fidelities as high as ~ 0.999 [8, 7]. However, variability (drift) and inaccuracies in the amplitude or frequency of the microwave field may significantly decrease fidelities, leading to control pulses that must be recalibrated frequently and are only accurate over narrow bandwidth and amplitude ranges. That said, a plethora of well-developed compensation schemes exist to mitigate the effects of miscalibration. Since the 1970s, NMR pulse design techniques have been optimized to tolerate control field imperfections, which have historically been a dominant source of error [9]. Robust control of quantum information processors inspired by pulsed NMR methods has been implemented on a variety of architectures including trapped ions [10], doped solids [11], NMR ensembles [12, 9] as well as pre-NISQ era superconducting

qubits [13]. These techniques have so far benefited each type of quantum processor they were tested on.

Composite pulses (CPs) were originally developed in NMR to achieve uniform inversion in the presence of rf field inhomogeneity [34]. CPs consist of short sequences of phase-modulated pulses designed to correct for errors in the rotation angle or axis of rotation. Trains of only a handful of pulses can compensate for relatively large miscalibrations of the control field by careful selection of pulse phases and flip angles [34]. General methods of CP design using propagator expansions can generate longer sequences resulting in gates of arbitrary precision [9] and robustness with respect to variations in amplitude and frequency [12]. CPs have been used to generate numerous broadband, narrowband, and passband schemes in population inversion experiments [35]. A class of composite pulses known as general rotors (or universal rotations) that produce desired rotations independent of the initial state have also been developed for quantum computing applications where insensitivity to the initial state is critical [12]. Such universal rotation pulses are typically designed to perform very high fidelity rotations in the presence of small errors but may be less robust than "point-to-point" transformation pulses frequently used in NMR, where specific initial conditions are often anticipated [12].

Adiabatic pulses (APs) are a class of control pulses possessing inherent robustness; their potential in quantum computing has already been demonstrated: adiabatic gates have been realized with high fidelity ≈ 0.997 in neutral atom qubits [36] and have been also been used to stimulate coherent population transfer between superconducting qubits [37]. APs are designed with continuous amplitude and frequency modulation functions, defined so that the frequency modulation as a function of time $\Delta_q(t)$ gradually alters the rotation axis while maintaining the qubit in an eigenstate colinear with this axis. For inversion pulses, the axis is swept from north to south orientations. Several amplitude and frequency modulation functions have been proposed to perform adiabatic rotations more rapidly, which is essential because circuit depth is limited by the gate duration relative to the intrinsic coherence time of the qubit. These adiabatic rapid passage pulses are primarily used to perform excitation,

adiabatic half-passage (AHP), and population inversion, adiabatic full-passage (AFP) on a well defined initial state.

Conventional APs are not general rotors; they only produce predictable rotations for the component of the initial state that is colinear with the effective microwave drive and induce inhomogeneous phase shifts in the transverse component. However, composite AHPs known and B1-insensitive rotation (BIR) pulses have been developed in which these inhomogeneous phase shifts cancel such that BIR sequences can effect rotations on arbitrary initial states by arbitrary flip angles [38, 11]. Understanding the performance of single adiabatic pulses on superconducting architectures is integral to the future implementation of composite general rotation gates. Of the AFP pulses, the hyperbolic secant (HS1) pulse is the most popular one due to its insensitivity to amplitudes above the threshold value required to satisfy the adiabaticity condition and its near uniform inversion inside the borders governed by the frequency sweep [38]. We characterized HS1 pulses with various parametrizations of the amplitude and frequency modulation functions using the same inversion profile contours and randomized benchmarking techniques as employed for composite pulses.

In this work, we demonstrate that both composite and adiabatic pulses implemented on a transmon qubit can be used to improve the robustness of single-qubit NISQ gates and that certain composite pulse sequences also produce higher on-resonance fidelity. We tested a selection of general rotors consisting of CORPSE (Compensation for Off-Resonance with a Pulse SEquence) and BB1 (Broad Band 1) pulses as well as conventional broadband inversion pulse sequences including $90_x180_y90_x$, $180_{120}180_{240}180_{120}$, and Knill pulses. We used industry-standard methods to characterize the robustness, fidelity, and leakage of CPs based on simulated and experimental inversion profiles, interleaved randomized benchmarking, and leakage randomized benchmarking. We also optimized HS1 pulses that satisfy the adiabaticity condition on shorter than typical timescales [11, 37]; this provides a foundation for composite AP universal rotors. Both composite and AP pulse design techniques can be extended to allow general unitary rotations, which could be used to define entire gatesets endowed with the increased robustness conferred by these schemes.

4.2 Transmon Hamiltonian

The idealized single-qubit gate problem for transmons is a 2-level system ($|0\rangle$, $|1\rangle$) that maps to the NMR problem for spin-1/2 nuclei (the spin Hamiltonians are similar). NMR techniques are therefore readily applicable to coherent control of transmon qubits. Although transmon circuits are anharmonic oscillators and have accessible states outside the computational subspace that must be considered to evaluate leakage rates, the qubit-drive interaction in the rotating frame of the qubit resonant frequency is well approximated by the spin Hamiltonian

$$\hat{H}_d = \frac{\sigma_z}{2}\Delta_q(t) + \Re\left[\Omega(t)\right]\frac{\sigma_x}{2} + \Im\left[\Omega(t)\right]\frac{\sigma_y}{2} \tag{4.1}$$

where $\Delta_q(t) = \omega(t) - \omega_0$ is the detuning from the resonant frequency, $\Re\left[\Omega(t)\right]$ and $\Im\left[\Omega(t)\right]$ are the real and imaginary parts of the microwave drive, and $(\sigma_x, \sigma_y, \sigma_z)$ are the Pauli spin matrices [6].

The propagator for this Hamiltonian can be represented in the convenient form of angle-axis rotations on the Bloch sphere by defining an effective field Ω_{eff} and rotation axis $\mathbf{n}(t) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$.

$$\Omega_{\text{eff}} = \sqrt{|\Omega(t)|^2 + \Delta_q(t)^2} \qquad \theta(t) = \arctan\left(\frac{|\Omega(t)|}{\Delta_q(t)}\right).$$

The propagator for each time step is then

$$U_d(t, t+\Delta t) = \exp\left(-i\frac{\Omega_{\text{eff}}}{2}(\mathbf{n}\cdot\boldsymbol{\sigma})\Delta t\right) \tag{4.2}$$

where the time dependence of $\Omega_{\text{eff}}(t)$ and $\mathbf{n}(t)$ have been suppressed, and $\boldsymbol{\sigma}$ is the vector of Pauli matrices $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. The relationship between the effective rotation axis and the ideal rotation axis in the transverse plane $\mathbf{n}_0 = (\cos\phi, \sin\phi, 0)$ is then determined by error in the phase ϕ, which rotates the axis in the transverse plane, and the detuning Δ_q, which introduces a z-component to the axis. The flip angle about the effective axis for a given

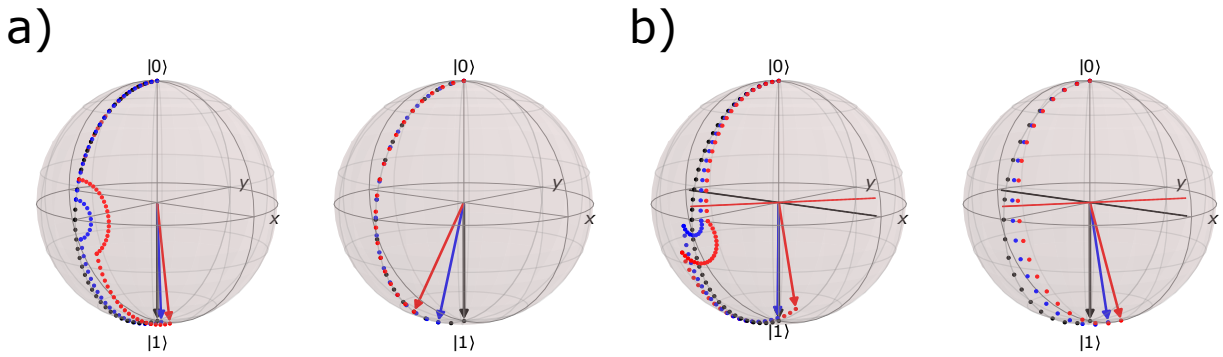

Figure 4.1: a) Bloch sphere rotations for composite $90_x180_y90_x$ square-pulse and single 180x square-pulse with amplitude offsets. Trajectories correspond to on-resonance $\Delta_q \simeq 0$ pulses with amplitude offsets Ω/Ω_0= 0.9 (blue) and 0.8 (red). b) Composite pulse $90_x180_y90_x$ and single 180_x pulse with detuning from resonant frequency. Trajectories are plotted for frequency offsets $\Delta_q/\Omega_0 = 0.09$ (blue) and 0.15(red), and the influence of detuning on the x rotation axis is shown

timestep $d\beta_{\text{eff}}$ can also be related to the nominal flip angle $d\beta_\phi$.

$$d\beta_{\text{eff}} = \sqrt{\frac{|\Omega(t)|^2 + \Delta_q(t)^2}{|\Omega(t)|^2}}\, d\beta_\phi.$$

This representation of the effect of systematic errors in the microwave field allows compensation by composite pulses to be understood from a geometric perspective, as shown for the $90_x180_y90_x$ pulse in Figure 4.1. Here, we employ β_ϕ notation to indicate the nominal flip angle β and phase ϕ (or corresponding axis in the transverse plane $0° = x$, $180° = y$) for each single pulse in sequence.

4.3 Experimental Methods

Experimental data used to characterize the performance of the composite and adiabatic pulses studied in this paper were collected using the IBM Quantum Experience platform and the open-source software development kit Qiskit. Qubit 2 of the five transmon qubit (falcon r4T) IBM quantum backend `ibmq_lima` was used for all data collection. This qubit has an average resonant frequency ω_0 of 5.247 GHz, an estimated anharmonicity of 333.6 MHz,

and T_1 and T_2 coherence times of 114.35 and 104.89 μs respectively. The average readout assignment error is 2.02%, and the error associated with ground state preparation is 0.94% [39].

The quantum processor `ibmq_lima` offers pulse-level control of operations via Qiskit's `quantum_circuit.add_calibration` method, which allows custom pulse schedules used in transpiling the circuit to be specified for any gate. Pulse schedules can be constructed with library functions, including DRAG and Gaussian profiles, or with custom waveforms. Complex valued waveforms can be constructed as piecewise-constant pulses with a dwell-time between samples of 0.22 ns, allowing control pulses with nearly-continuous amplitude and phase modulation to be implemented. Phase evolution provides a method of modulating the carrier frequency of the waveform, which we have set to be the resonant frequency of the qubit (ω_0).

4.3.1 Calibration Experiments

In order to calibrate the carrier frequency, we performed a frequency sweep in steps of 6 MHz over 30 MHz range centered on the most recent IBM calibration of the qubit frequency while probing the readout resonator to observe state-dependent shifts in the kerneled measurement signal [32].

Additional calibration experiments were performed to determine the optimal parameters for the shaped composite pulses. Single pulses were implemented both as Gaussian and DRAG (Derivative Reduction by Adiabatic Gate) pulses to facilitate comparison of the composite pulse leakage rates in the presence and absence of leakage reduction efforts. Gaussian pulses were constructed using the `qiskit.pulse.library` Gaussian class as truncated Gaussian profiles

$$\Omega(t) = Ae^{-(t-T/2)^2/2\sigma^2}, \quad t \in [0, T]$$

with duration $T = 35.2$ ns and standard deviation $\sigma = 8.8$ ns, which are the same as T and σ for the default X gate on `ibmq_lima`. The nominal π-amplitude of the Gaussian was

determined by sweeping A over a range of 200 amplitudes between -0.5 and 0.5 a.u. and fitting the resulting Rabi oscillations to a sinusoid to find the period, half of which gave the π-amplitude. The conversion factor from a.u. to rad/sec was also determined from the Rabi frequency using the relationship

$$\Omega_{conv.}\left[\frac{\text{rad/sec}}{\text{a.u}}\right] = \frac{\beta}{\int_0^T \Omega_{a.u}(t)dt}$$

where β is the nominal flip angle. The determined $\Omega_{conv.}$ was 964,474,196 rad/sec per a.u. The nominal π-pulse amplitude Ω_0 is the product of the calibrated π-amplitude A in a.u with Ω_{conv}.

For DRAG pulses, an additional parameter α controlling the derivative component was also calibrated. The DRAG pulse profile is defined as

$$\Omega(t) = Ae^{-(t-T/2)^2/2\sigma^2} + i\alpha\frac{d}{dt}\left(Ae^{-(t-T/2)^2/2\sigma^2}\right), \quad t \in [0, T]$$

where the duration and standard deviation have again been defined as $T = 35.2$ ns and $\sigma = 8.8$ ns. The parameter α was calibrated by repetitions of the identity rotation block $R_p(\beta)R_m(\beta)$, where $R_p(\beta)$ is a rotation by a positive angle β and R_m is a rotation by the negative angle of the same magnitude. A range of α values was scanned to determine the value that minimized the undesired z rotations for the sequence, as described in Ref. [7]. Standard DRAG compensation schemes are calibrated to minimize these phase errors, which are caused by coupling of computational and noncomputational sublevels under the influence of the microwave drive, since phase errors rather than leakage have been identified as the dominant source of gate infidelity [40]. Leakage reduction still occurs for phase-optimal α [7]

4.3.2 Pulse Simulation

Pulse simulations were performed in Python using the spin Hamiltonian in the ω_0 frame given by Eq. (4.1), which can be rewritten in terms of the amplitude and phase of the microwave

drive as

$$\hat{H}_d = \frac{\sigma_z}{2}\Delta_q(t) + \frac{|\Omega(t)|}{2}\left(\sigma_x \cos(\phi(t)) + \sigma_y \sin(\phi(t))\right).$$

Transition probabilities were simulated for a range of frequency offsets Δ_q and drive amplitudes $A\Omega_{conv}$, where the dimensionless scaling factor A (a.u.) was specified to correspond with the range of experimental drive amplitudes in a.u. and Δ_q to the range of experimental detunings. The complex-valued microwave drive waveform was discretized into n timepoints $\{t_i\}$ separated by time intervals Δt equal to the dwell-time of the IBMQ backend (0.22 ns). The amplitude $|\Omega(t)|$ and phase $\angle\Omega(t)$ were calculated at each timepoint, and the propagator for each time t_i was determined $U(t_i, t_{i-1}) = \exp\left(-i\hat{H}_d(t_i)\Delta t\right)$. The discretized approximation to the overall action of the pulse $U(t_n, t_0)\,|0\rangle = U_{t_n} \ldots U_{t_0}\,|0\rangle$ on the initial state $|0\rangle = (1,0)^\dagger$ was then projected onto the desired final state $|1\rangle = (0,1)^\dagger$ to determine the transition probability.

4.3.3 Adiabatic Pulse Design

The HS1 pulse is defined in the ω_0 rotating frame by the complex-valued function

$$\Omega(t) = \Omega_{\max}\left[\mathrm{sech}(\beta(\frac{2t}{T} - 1)\right]^{1+i\mu},$$

where μ and β are real valued parameters, $\Omega^{\max}$ is the maximum amplitude, and T is the pulse duration. This can be rewritten in terms of the amplitude and phase modulation functions $\Omega(t) = |\Omega(t)|\, e^{i\phi(t)}$

$$|\Omega(t)| = \Omega_{\max} \cdot \mathrm{sech}(\beta(\frac{2t}{T} - 1))$$

$$\phi(t) = \mu \ln\left[\mathrm{sech}(\beta(\frac{2t}{T} - 1))\right],$$

from which the frequency sweep is

$$\Delta_q(t) = \frac{d\phi}{dt} = -\frac{2\mu\beta}{T}\tanh(\beta(\frac{2t}{T} - 1)).$$

The rate of the frequency sweep and effective drive strength Ω_{eff} determine whether the pulse satisfies the adiabaticity condition

$$\Omega_{\text{eff}} \gg \left| \frac{d\theta(t)}{dt} \right|$$

throughout its duration. The minimum difference $(\Omega_{\text{eff}} - |d\theta(t)/dt|)$ is typically attained when the drive is resonant with the qubit, at which point Ω_{eff} reaches its minimum value $|\Omega(t)|$. For this reason, it is necessary to specify the parameters T, β, and μ such that adiabaticity is preserved at this critical time. These parameters additionally determine the amplitude of the frequency sweep $2\mu\beta/T$, which defines the pulse bandwidth since HS1 pulses are designed to perform nearly uniform inversion for all isochromats (off-resonance qubits) within the frequency range of the sweep [38].

In our implementation, we prioritized simultaneously reducing HS1 duration, maximizing the simulated transition probability, maximizing the minimum distance specified by the adiabiticity condition, and increasing the range of addressable isochromats over explicit specification of either the duration of bandwidth. For our parameter search, we used the Python Genetic Algorithm library PyGad to perform a stochastic optimization of HS1 pulses with a gene space defined by durations T between 95 and 169 ns and β and μ parameters ranging from 1 to 30. The cost function was constructed as a sum of positive terms proportional to the simulated inversion probability and amplitude of the frequency sweep and terms negative terms proportional to the pulse duration and the reciprocal minimum difference $\Omega_{\text{eff}} - |d\theta(t)/dt|$ defining the adiabaticity condition. The best solutions produced by the genetic algorithm were ordered by their duration and by their simulated inversion probability, and the best performing pulses with respect to both metrics were selected.

4.3.4 Randomized Benchmarking

Randomized benchmarking protocols have become an industry standard for noise characterization in quantum operations due to their scalability and insensitivity to state preparation

and measurement (SPAM) errors. In addition to standard randomized benchmarking, which provides an estimate for the average fidelity of a gateset, extensions have been developed to characterize leakage and to estimate the fidelity of individual gates. These techniques are leakage randomized benchmarking (LRB) [8] and interleaved randomized benchmarking (IRB) [31], which we employed in this study to characterize the performance of our selected composite and AFP pulses.

4.3.4.1 Interleaved Randomized Benchmarking

Interleaved randomized benchmarking was performed according to the protocol proposed in [31] with a Clifford generating set $\{\mathbb{I}, X, Y, Z, H, S, S\dagger\}$ and random circuits with sequence lengths of $m = 1, 21, 41 \ldots 341$ random Clifford elements $\mathcal{C}_{i_j} \in \mathfrak{C}_n$ for composite pulses and $m = 1, 11, 21 \ldots 341$ for adiabatic full passage pulses, for which finer graining was necessary to capture the more rapid survival probability decay. A recovery operation $\mathcal{C}_{i_{m+1}}$ computed as the inverse of the random sequence was applied for each sequence length and $\langle \sigma_z \rangle$ was measured to obtain the survival probability of the initial state $|0\rangle$. Average survival probabilities for each m were determined by averaging over 50 independent random sequences $\{i_m\}$. Reference circuits composed solely of the random Clifford elements and interleaved circuits, where the gate of interest $\mathcal{C}$ is inserted between the $\mathcal{C}_{i_j}$, were constructed using a fixed random seed such that the interleaved and standard circuits were identical with the exception of the interleaved element $\mathcal{C}$.

In our implementation of the IRB protocol, there were two additional considerations to account for. First, interleaved randomized benchmarking is designed to estimate fidelities for individual elements of $\mathfrak{C}_n$; however, the majority of the composite pulses and the AFP pulses to which we applied this protocol are not elements of this group. Second, RB protocols are typically implemented for gatesets with Clifford elements of the same average duration since the depolarization of unitary gate errors assumed by these protocols depends on the time at which the depolarizing projection is performed [8]. The composite and AFP pulses we benchmarked are longer than the average Clifford gate, as defined by the default transpiler

instructions associated with these gates on `ibmq_lima`.

To address the first consideration, we have defined our interleaved element as a pseudo-identity schedule (named 'customI') consisting of two composite or AFP pulses in sequence (Fig. 4.2 c). An ideal inversion pulse is its own inverse, so this schedule can be represented as $\mathcal{I} \in \mathfrak{C}_1$ for the purpose of computing the recovery operation. However, unlike applying the composite pulse and its time-reversed complex conjugate, over or under rotations are not reversed but rather magnified by applying two pulses in sequence. We therefore report the gate error obtained by fitting the survival probability decay curves for $\mathcal{C}$ defined by two sequential inversion pulses as well as the expected fidelity of a single inversion pulse, which we estimate to be half that of the two back-to-back pulses.

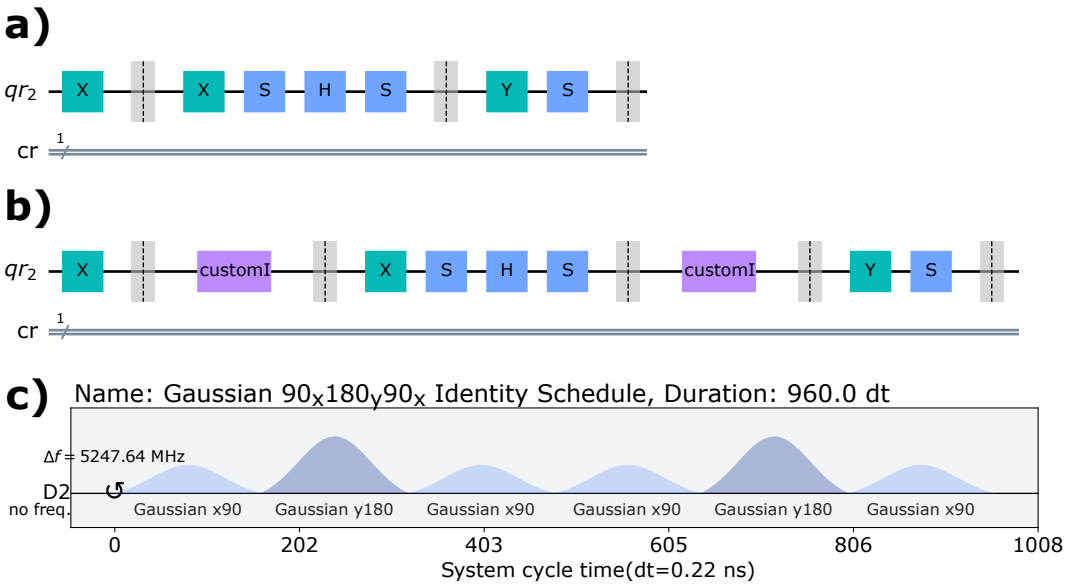

Figure 4.2: Interleaved randomized benchmarking schematic a) Excerpt of 3 random Clifford elements in standard RB circuit decomposed in the Clifford generating set $\{\mathbb{I}, X, Y, Z, H, S, S\dagger\}$ b) Excerpt of IRB circuit with gate of interest 'customI' interleaved between random clifford elements c) Pulse schedule used to define the 'customI' operation for interleaved randomized benchmarking of Gaussian $90_x180_y90_x$ composite pulse

The second consideration was resolved by redefining the transpiler instructions associated with the basis gateset such that the average Clifford duration was the same as the duration of the 'customI' schedule. To execute the standard RB and IRB circuits on `ibmq_lima`, the Clifford generating gates $\{\mathbb{I}, X, Y, Z, H, S, S\dagger\}$ were first decomposed into the basis gateset

$\{\mathbb{I}, \mathrm{CNOT}, R_Z, X, \sqrt{X}\}$, of which X and $\sqrt{X}$ are the only single-qubit gates with non-zero duration. By counting the number of basis gates of each type into which the 50 random sequences of length m=341 are decomposed, the average number of basis gates per Clifford was determined and used to compute the average $\mathcal{C}_{i_j}$ duration. The duration of the pulses executing the X and $\sqrt{X}$ rotations were then increased based on the average composition of $\mathcal{C}_{i_j}$ in terms of these schedules so that the average Clifford duration was equal to that of the 'customI' schedule for the particular AFP or custom pulse under consideration. These longer pulse schedules were defined, like the default schedules, as DRAG pulses with a 4:1 ratio between the duration and the standard deviation. The amplitude and α parameter for the X and $\sqrt{X}$ schedules was calibrated using the methods previously outlined to ensure that the longer schedules performed the same rotations as the replaced defaults.

The remainder of the protocol was implemented as described in Ref. [31]. Potential gate-dependence in the per-Clifford error was investigated by fitting the average survival probability $\overline{F}_{\mathrm{seq}}$ to the zeroth and first order models [29]

$$\overline{F}^{(0)}_{\mathrm{seq}} = A_0 p^m + B_0$$

$$\overline{F}^{(1)}_{\mathrm{seq}} = A_1 p^m + B_1 + C_1(m-1)(q-p^2)p^{m-2}$$

and comparing the reduced chi-square χ^2_ν criteria. All survival probability decay curves were fit well by the zeroth order model with a lower χ^2_ν value than the first-order model. The depolarizing parameter of the zeroth order fit for the standard and interleaved sequences was therefore used to obtain the average error rate of the interleaved element

$$r_{\mathcal{C}} = \frac{(d-1)(1-p_{\mathcal{C}}/p)}{d}$$

and the bounds on the error rate $[r_{\mathcal{C}} - E, r_{\mathcal{C}} + E]$ where

$$E = \min \begin{cases} \dfrac{(d-1)\left[|p - p_{\mathcal{C}}/p| + (1-p)\right]}{d} \\ \dfrac{2(d^2-1)(1-p)}{pd^2} + \dfrac{4\sqrt{1-p}\sqrt{d^2-1}}{p} \end{cases}.$$

The average gate fidelity of the inversion pulse was then calculated from the error rate of the two inversion pulse interleaved element as $\mathcal{F} = (1 - r_{\mathcal{C}})/2$.

4.3.4.2 Leakage Randomized Benchmarking

Transmon qubits are multi-level systems whereby the lowest two levels ($|0\rangle$, $|1\rangle$) are selectively addressed by tuning the center frequency. However, some amount of leakage to the upper levels ($|2\rangle$, $|3\rangle$, etc.) may occur depending on the pulse characteristics. This can be assessed by measuring the leakage rate, another important metric for pulse performance. Leakage errors pose a distinct obstacle to fault-tolerant computation and require unique, resource-intensive error correction schemes [26]. Interactions with leaked states can also lead to logical errors more deleterious to the gate fidelity than leakage itself, such as the phase errors resulting from the AC Stark shift that the DRAG scheme is designed to mitigate [41]. We therefore seek to characterize the leakage rates for our selected composite and AFP pulses using the LRB protocol developed in [8]. Although other RB protocols designed to estimate leakage have been proposed [33], the LRB protocol was selected because it can be implemented without direct measurement of the leakage subspace via a modification of the RB circuits and because it estimates both the leakage (L_1) rate out of the computational subspace and rate of return (L_2) of the leaked population to the computational subspace [8]. These rates are average rates over a gateset and do not provide direct estimates of the leakage rates associated with the individual gates we aim to benchmark. Nevertheless, variations in the average are expected to reflect the performance of the composite or AFP pulse included in the Clifford generating set.

The LRB protocol was implemented using 50 independent random sequences with lengths

of $m = 1, 11, \ldots 551$ Clifford elements. LRB requires the measurement of the probabilities $\mathcal{P}_j$ that the final state prepared by the RB circuit is a state $|j\rangle$ belonging to either the computational $\chi_1 = \{|0\rangle, |1\rangle\}$ or leakage $\chi_2 = |2\rangle$ subspaces. The survival probability of the χ_1 subspace is then determined by summing the survival probabilities $\mathcal{P}_0 + \mathcal{P}_1 = \mathcal{P}_{\chi_1}$ and fitting the resulting decay in the average survival probability of the computational subspace to the model $\overline{\mathcal{P}_{\chi_1}}(m) = Cp^m + B$. The leakage and seepage rates are determined from this model using the equations

$$L_1(\mathcal{E}) = (1 - B)(1 - p), \quad L_2(\mathcal{E}) = B(1 - p).$$

In the event that it is not possible to fully discriminate the $|0\rangle, |1\rangle, |2\rangle$ states, as was the case for our measurements on `ibmq_lima`, the protocol may still be implemented using the method described for the special case where only the POVM $\{E_0, \mathbb{I} - E_0\}$ (where $E_0 = |0\rangle\langle 0|$) can be measured with high fidelity. To implement the modified protocol, the unitary that rotates the projector E_j for the state $|j\rangle$ to the projector onto the ground state is first determined $E_0 = U_j E_j U_j^\dagger$. The recovery operation $\mathcal{C}_{i_{m+1}}$ is then modified as $\mathcal{C}^{(j)}_{i_{m+1}} = \mathcal{U}_j^\dagger \circ \mathcal{C}_{i_{m+1}}$ so that the population of $|j\rangle$ is rotated to $|0\rangle$ prior to measurement. In our case, $|j\rangle = |1\rangle$, and the recovery operation should be modified by $\mathcal{U}_j = \sigma_x$ to allow $\mathcal{P}_1$ to be measured by the projector onto the ground state. We therefore constructed two sets of circuits: one with the unmodified recovery operation $\mathcal{C}_{i_{m+1}}$ to measure $\mathcal{P}_0$ and one with an additional X gate modifying the recovery operator $\mathcal{C}^{(1)}_{i_{m+1}}$ to measure $\mathcal{P}_1$.

As for our implementation of IRB, we define a custom identity pulse schedule from each composite and AFP pulse. Instead of systematically interleaving this element, we included it as an identity operation in the Clifford generating set $\{\mathcal{I}, X, Y, Z, S, S^\dagger\}$. To account for the fact that the randomly interspersed custom identity schedule is not the same length as the other Clifford gates, we again assigned calibrated pulse schedules to the X and $\sqrt{X}$ gates in the transpiled circuits such that the average Clifford duration was the same as the identity pulse schedule. After attaching these transpiler instructions to the circuits, $\mathcal{P}_0$ and $\mathcal{P}_1$ were measured by performing 1024 measurements of the POVM $\{E_0, \mathbb{I} - E_0\}$ and recording $\langle E_0 \rangle$.

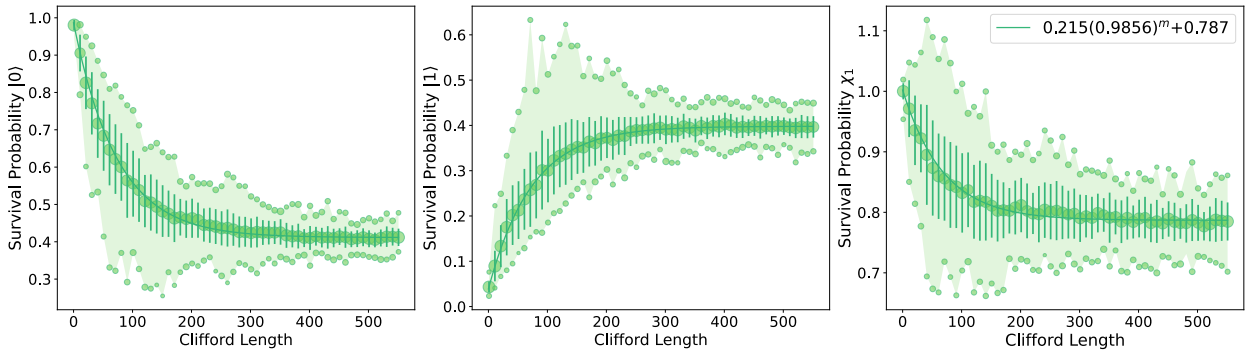

Figure 4.3: Leakage randomized benchmarking measurements of $|0\rangle$ and $|1\rangle$ and fitted computational subspace population χ_1 for HS1 pulse T:148 ns β:5.72 μ:1.85. Survival probabilities of $|0\rangle$ and $|1\rangle$ were determined from 1024 measurements of the POVM $\{E_0, \mathbb{I} - E_0\}$ for RB circuits with the standard $\mathcal{C}_{i_{m+1}}$ and modified $\mathcal{C}^{(1)}_{i_{m+1}}$ recovery operations. Since $\mathcal{P}_0$ and $\mathcal{P}_1$ were determined from separate experiments $\mathcal{P}_{\chi_1} = \mathcal{P}_0 + \mathcal{P}_1$ may exceed 1. Any resulting baseline in the mean values of $\mathcal{P}_{\chi_1}$, determined by the difference $\overline{\mathcal{P}_{\chi_1}(m=1)} - 1$, has been subtracted prior to fitting in order to ensure valid average survival probabilites and leakage rates. Reciprocal standard deviations of the data at each length were supplied as weights for least-squares fitting of the model $\overline{\mathcal{P}_{\chi_1}} = Cp^m + B$. Extracted parameters, $p = 0.9856 \pm 0.0006$, $B = 0.787 \pm 0.001$, yielded leakage and seepage rates of $L_1 = 3.1 \times 10^{-3} \pm 2 \times 10^{-4}$ and $L_2 = 1.13 \times 10^{-2} \pm 5 \times 10^{-4}$.

The survival probability of the computational subspace $\mathcal{P}_{\chi_1} = \mathcal{P}_0 + \mathcal{P}_1$ was then calculated; however, since $\mathcal{P}_0$ and $\mathcal{P}_1$ were recorded in separate experiments, values of $\mathcal{P}_{\chi_1}$ exceeding 1 were possible. Any such baseline in the mean values $\overline{\mathcal{P}_{\chi_1}(m)}$, determined by the difference $\overline{\mathcal{P}_{\chi_1}(m=1)} - 1$, was subtracted prior to fitting to ensure valid average survival probabilities and leakage rates.

4.4 Results

4.4.1 Composite Pulses

The majority of the composite pulses tested are more robust than the Gaussian 180_x and DRAG 180_x single pulses with respect to at least one of the dimensionless figures of merit, namely the amplitude compensation bandwidth $\Theta(\Omega/\Omega_0)$ or the frequency compensation bandwidth $\Xi(\Delta_q/\Omega_0)$. Conceptually, the robustness of control pulses can be defined in terms of the derivatives of the transition probability with respect to the drive parameters,

evaluated at the optimal drive conditions ($\Delta_q/\Omega_0 = 0,\ \Omega/\Omega_0 = 1$). Robust pulses are characterized by small values of these derivatives so that the inversion probability remains high over a broader range of amplitudes and detunings. Analysis of these ranges is facilitated by employing dimensionless quantities reduced by Ω_0. Ω_0 is the nominal drive amplitude, defined as the product of the calibrated amplitude of the shaped π-pulse (0.145 a.u for the experiments in Tables 4.1 and 4.3) and the conversion factor Ω_{conv}. The compensation bandwidths ($\Theta,\ \Xi$) reported in Table 4.1 represent the range of amplitudes Ω/Ω_0 at zero detuning or range of frequencies Δ_q/Ω_0 at nominal amplitude for which the experimental inversion is greater than or equal to 0.9. Ranges reported in bold type are more robust than the single pulse of the same shape while those in italics are more restricted. Both the Gaussian and DRAG $180_{120}180_{240}180_{120}$ pulses and the BB1 pulse are more robust towards amplitude errors than single pulses, while the Gaussian and DRAG $90_x180_y90_x$ sequences and Knill pulse are more robust towards both amplitude and frequency errors.

4.4.1.1 Gaussian and DRAG $90_x180_y90_x$ Pulses

The $90_x180_y90_x$ inversion pulses constructed using single pulses with Gaussian and DRAG profiles are more robust to frequency and amplitude errors than the single 180_x pulses of the same respective shapes. The Ω/Ω_0 range increased by 60% for the Gaussian and 83% for the DRAG $90_x180_y90_x$ pulse. The frequency compensation bandwidth more than doubled in both cases. Composite $90_x180_y90_x$ pulses are also more tolerant to combinations of frequency and amplitude errors, as represented graphically in Fig. 4.4. Both the simulations and experimental results show an increase in the area of the 0.9 contour level of the first Rabi peak over that of the single pulse. The difference in the robustness for DRAG and Gaussian constituent pulses is minimal, with simulated and experimental inversion contours for the different single pulse shapings being nearly identical. The fidelities obtained by randomized benchmarking of the on-resonance pulses at nominal amplitude are higher for the $90_x180_y90_x$ CPs than those of the single 180_x pulses despite the increase in gate duration. DRAG and Gaussian $90_x180_y90_x$ sequences again performed comparably with respect to fidelity.

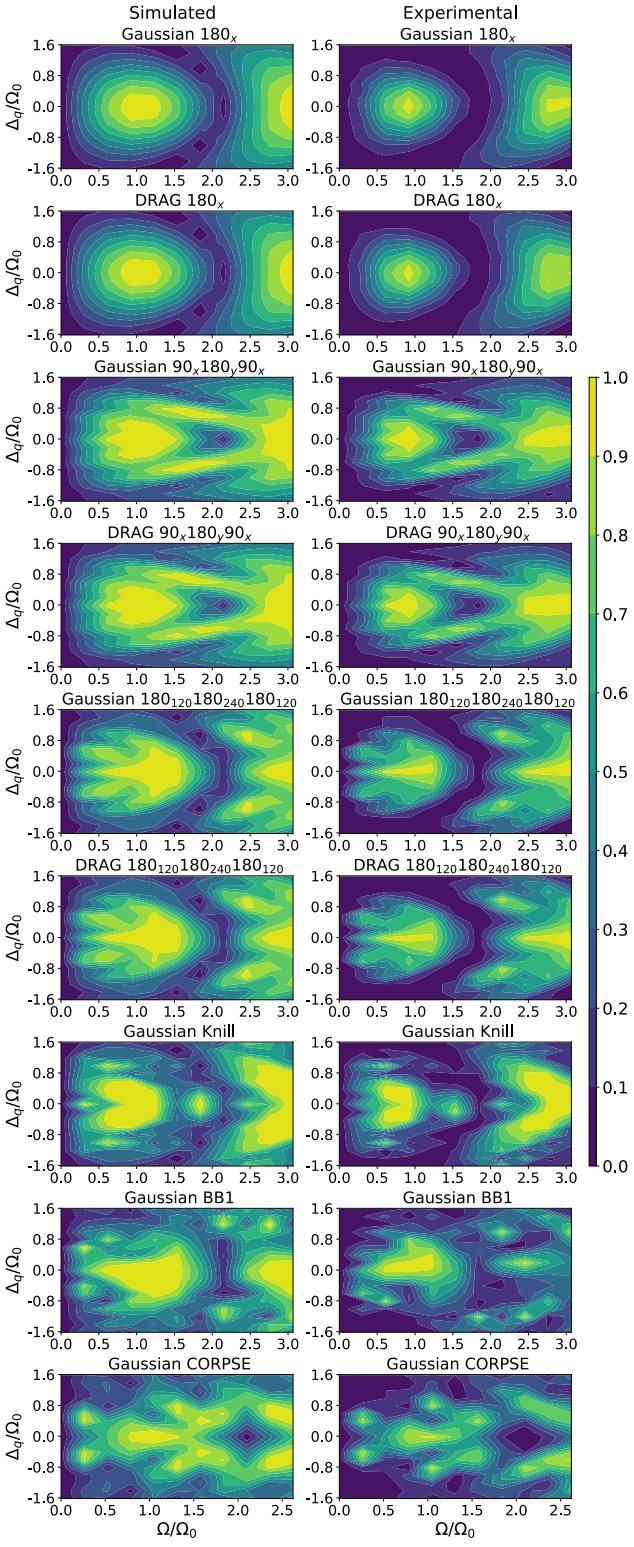

Figure 4.4: Simulated (left) and experimental (right) contours for composite pulses. Experimental frequencies ranged from -40 to 40 MHz in steps of 5 MHz, for a resolution in dimensionless units of ($\Delta\Delta_q/\Omega_0 = 0.2$). Amplitudes ranged from 0 to 0.5 a.u. (or to max π-pulse power supported by `ibmq_lima` for BB1 and CORPSE) in steps of 0.05 a.u. ($\Delta\Omega/\Omega_0 = 0.3$). 1024 trials were performed for each amplitude, frequency pair to obtain the average inversion probability.

Table 4.1: Experimental compensation bandwidths of composite pulses[1]

Composite Pulse	$\Theta(\Omega/\Omega_0)$[2]	$\Xi(\Delta_q/\Omega_0)$[3]
Gaussian Composite Pulses		
180_x	{0.825,1.200}	{-0.179,0.220}
$90_x180_y90_x$	**{0.650,1.250}**	**{-0.404,0.404}**
$180_{120}180_{240}180_{120}$	**{0.500,1.300}**	*{-0.090,0.135}*
Knill	**{0.750,1.175}**	**{-0.449,0.449}**
BB1	**{0.525,1.375}**	*{0.000,0.310}*
CORPSE	*{0.875,1.200}*	*{-0.090,0.090}*
DRAG Composite Pulses		
180_x	{0.800,1.100}	{-0.134,0.179}
$90_x180_y90_x$	**{0.675, 1.225}**	**{-0.404,0.404}**
$180_{120}180_{240}180_{120}$	**{0.525, 1.325}**	*{-0.135,0.090}*

[1] compensation bandwidths in bold are broader than the single pulse of the same profile shape; those in italics are narrower

[2] amplitude sweep on-resonance performed in steps of $\Delta\Omega/\Omega_0 = 0.025$

[3] frequency sweep at nominal amplitude performed in steps $\Delta\Delta_q$ of 1 MHz ($\Delta\Delta_q/\Omega_0$=0.045)

4.4.1.2 Knill Pulse

The Knill inversion pulse consists of sequence of five pulses that perform an overall 180° rotation with an additional 60° phase shift:

$$180_0^{\text{ideal}}(60)_z = (180)_{30}(180)_0(180)_{90}(180)_0(180)_{30}.$$

This pulse has previously been demonstrated to improve robustness to an expanded range of frequency and amplitude errors in spin qubit control [11]. Our results also indicate an increase in the amplitude and frequency compensation bandwiths for this pulse. The Ω/Ω_0

Table 4.2: Randomized benchmarking estimates for gate fidelity and leakage rate of composite pulses

Composite Pulse	Custom Identity Gate Error[4]	Bounds on Gate Error	Inversion Gate Fidelity	Leakage Rate (L_1)[5]	Seepage Rate (L_2)
Gaussian 180_x	0.016 ± 0.002	[0, 0.032]	0.9919	5×10^{-5}	9×10^{-3}
DRAG 180_x	0.019 ± 0.002	[0, 0.037]	0.9907	$2{\times}10^{-5}$	2×10^{-3}
Gaussian $90_x180_y90_x$	0.0069 ± 0.0004	[0, 0.0138]	0.9966	$5{\times}10^{-5}$	4×10^{-3}
DRAG $90_x180_y90_x$	0.0061 ± 0.0003	[0, 0.0122]	0.9970	1.6×10^{-4}	1.3×10^{-3}
Gaussian $180_{120}180_{240}180_{120}$[6]	0.0010 ± 0.0001	[0, 0.0055]	0.9995	2×10^{-5}	1.12×10^{-3}
DRAG $180_{120}180_{240}180_{120}$	0.0010 ± 0.0001	[0, 0.0055]	0.9995	3.1×10^{-4}	1.7×10^{-3}
Gaussian Knill	0.0024 ± 0.0003	[0, 0.0106]	0.9990	3×10^{-4}	2×10^{-2}
Gaussian BB1	0.023 ± 0.004	[0, 0.046]	0.9886	1.4×10^{-4}	7×10^{-3}
Gaussian CORPSE	0.0037 ± 0.0003	[0, 0.0074]	0.9982	1.34×10^{-4}	9×10^{-3}

[4] Error estimated by nonlinear error propagation of fitted parameter variances

[5] uncertainty in L_1 and L_2 in last reported significant figure

[6] variance-covariance matrix could not be used to reliably estimate the error for L_1, L_2, significant figures reported are based on the precision of DRAG $180_{120}180_{240}180_{120}$

range only increased by a modest 13% relative to the Gaussian 180_x pulse, but the range of frequency offsets tolerated at nominal amplitude is the largest any of the composite pulses tested. The fidelity was comparable to the single 180_x pulse.

4.4.1.3 Gaussian and DRAG $180_{120}180_{240}180_{120}$ Pulses

Composite $180_{120}180_{240}180_{120}$ pulses are more robust towards amplitude errors, increasing $\Theta(\Omega/\Omega_0)$ relative to the single 180_x pulse by 113% and 167% for the Gaussian and DRAG profiles but only over a restricted range of frequencies at nominal drive amplitude. Interestingly, the $180_{120}180_{240}180_{120}$ CPs were able to compensate for a broader range of frequencies for amplitudes slightly above the nominal value, as shown by the broadening of the simulated and experimental contours near $\Omega/\Omega_0 = 1.2$. The fidelity of the $180_{120}180_{240}180_{120}$ pulses was the highest of any of the composite or adiabatic pulses tested at $\mathcal{F} = 0.9995$.

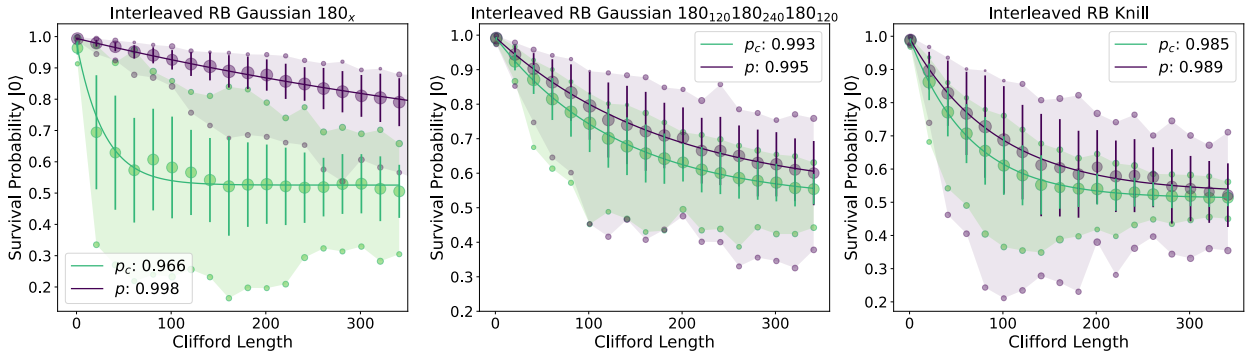

Figure 4.5: Interleaved randomized benchmarking survival probability decay curves for Gaussian 180_x (left), Gaussian $180_{120}180_{240}180_{120}$ (center), and Gaussian Knill (right) composite pulses. Average ground state survival probabilities over 50 random sequences for reference circuits without the interleaved gate (purple) and interleaved circuits (green) are both plotted against the random Clifford length vector $m = 1, 21, \ldots 341$. Error bars represent the standard deviation of the data at each length, and the shaded region denotes the range. Clifford gates $\mathcal{C}_{i_j}$ in the reference and interleaved circuits have average durations equal to that of the interleaved element: 70.4 ns (left), 211 ns (center), 352 ns (right). All data were well described by the zeroth order model $\overline{F}^{(0)}_{\text{seq}}$. Depolarizing parameters for the interleaved ($p_{\mathcal{C}}$) and reference (p) circuits were Gaussian 180_x $p_{\mathcal{C}}, = 0.966 \pm 0.004$, $p = 0.99845 \pm 4 \times 10^{-5}$; Gaussian $180_{120}180_{240}180_{120}$ $p_{\mathcal{C}} = 0.9925 \pm 0.0001$, $p = 0.9945 \pm 0.002$; Gaussian Knill $p_{\mathcal{C}} = 0.9847 \pm 0.0003$, $p = 0.9893 \pm 0.0005$

The robustness and fidelity were again comparable or identical for the Gaussian and DRAG $180_{120}180_{240}180_{120}$ CPs.

4.4.1.4 BB1 Pulse

The BB1 pulse is a broadband pulse designed to tolerate a wide range of amplitudes that theoretically reduces the sensitivity to rotation angle errors to sixth order in the error [35]. The BB1 sequence is a general rotor and can also perform arbitrary rotations of angle β_0^{ideal} by defining the phase of the composite pulses in the sequence as a function to the desired rotation angle:

$$\beta_0^{\text{ideal}} = (180)_\phi(360)_{3\phi}(180)_\phi(\beta)_0, \quad \phi = \cos^{-1}(-\beta/4\pi).$$

In our implementation, β_0^{ideal} is 180°. Experimental results for the BB1 inversion pulse bear out the expected increase in robustness to amplitude errors with a 127% increase in $\Theta(\Omega/\Omega_0)$ over the single Gaussian 180_x pulse, albeit over a narrower range of frequencies than the single pulse. Although the proportional increase in $\Theta(\Omega/\Omega_0)$ for the BB1 pulse is less than the DRAG $180_{120}180_{240}180_{120}$ single pulse over the DRAG 180_x, the absolute amplitude compensation bandwidth of the BB1 pulse is the largest recorded, and the tolerated frequency range is also less diminished. The on-resonance BB1 fidelity ($\mathcal{F} = 0.9886$) is slightly lower than the single Gaussian 180_x pulse ($\mathcal{F} = 0.9919$).

4.4.1.5 CORPSE Pulse

The compensation for off resonance pulse sequence (CORPSE) is another general rotor capable of producing arbitrary flip angles β_0^{ideal}

$$\beta_0^{\text{ideal}} = \left(\frac{\beta}{2} - \psi\right)_0 (2\pi - 2\psi)_\pi \left(2\pi + \frac{\beta}{2} - \psi\right)_0$$

$$\psi = \sin^{-1}\left(\sin(\beta/2)/2\right)$$

in the presence of detuning from the resonant frequency. Again, β_0^{ideal} is 180°. While the simulated contour for the CORPSE sequence displays an increased area of the 0.8 contour level, high fidelity inversion occurs only for a narrower range of amplitudes and frequencies relative to the single pulse. Experimental CORPSE compensation bandwidths are lower than those of a single pulse as well. The on-resonance error rate is, however, also lower than the single 180_x pulse .

4.4.2 Adiabatic Pulses

HS1 pulses proved to be incredibly robust towards frequency and amplitude errors. Several of the parameterizations implemented attain 90% inversion at thresholds below the nomi-

Table 4.3: Experimental robustness parameters of HS1 pulses

HS1 Pulse Parameters	Threshold Amp. Ω/Ω_0[7]	Freq. Sweep $-2\mu\beta/T$ (MHz)	Bandwidth Δ_q (MHz)[8]	Compensation Range $\Xi(\Delta_q/\Omega_0)$
HS1 T:98.6 ns β:10.92 μ:4.81	3.92	1067	330	{-7.4, 1.79}
HS1 T:120 ns β:5.72 μ:1.85	0.84	177	28	{-0.628,0.628}
HS1 T:130 ns β:10.92 μ:1.20	1.19	201	26	{-0.538,0.628}
HS1 T:148 ns β:5.72 μ:1.85	0.70	143	22	{-0.448,0.538}
HS1 T:165 ns β:12.90 μ:2.49	1.82	389	88	{-1.79,1.97}

[7] amplitude sweep on-resonance in steps of $\Delta\Omega/\Omega_0$=0.07, threshold at 0.9 inversion probability

[8] frequency sweep at nominal amplitude performed in steps $\Delta\Delta_q$ of 2 MHz ($\Delta\Delta_q/\Omega_0$=0.09) for rows 2 through 5 and 5 MHz ($\Delta\Delta_q/\Omega_0$=0.22) for row 1

nal π-pulse amplitude and maintain complete amplitude insensitivity up to the maximum microwave power deliverable. The breadth of frequencies addressable by HS1 pulses is also much larger than that of the composite pulse sequences, ranging from 1.1 to 10.2 times even the largest composite pulse compensation bandwidth (Knill pulse). Moreover, the duration of all of the HS1 pulses tested are comparable to the composite pulse durations. Although AFP pulses have often been discounted as they traditionally have durations an order of magnitude or more longer than single rectangular pulses [11], we found that pulses only 2.8 to 5 times the duration of the default `ibmq_lima` X gate are capable of driving robust inversion. The inversion fidelities obtained via randomized benchmarking are lower than those obtained for composite pulses, with the highest fidelity AFP pulse ($\mathcal{F} = 0.9897$) just approaching the fidelity of a single DRAG pulse ($\mathcal{F} = 0.9907$). Leakage rates are also higher for the AFP pulses relative to the composite pulses. However, considering that the bandwidth Δ_q of these pulses is larger than that of the composite pulses, higher leakage rates are expected. The frequency sweep for the two HS1 pulses with the highest leakage rate also crosses the anharmonic transition at 333.6 MHz, even though the recorded experimental inversion bandwidths are narrower than the full range of the frequency sweep and do not reach this detuning.

Table 4.4: Randomized benchmarking estimates for gate fidelity and leakage rate of HS1 pulses

HS1 Pulse Parameters	Custom Identity Gate Error[9]	Bounds on Gate Error	Inversion Gate Fidelity	Leakage Rate (L_1)[10]	Seepage Rate (L_2)
HS1 T:98.6 ns, β:10.9, μ:4.8	0.09±0.01	[0,0.19]	0.953	1.06×10^{-2}	1.6×10^{-2}
HS1 T:120 ns, β:5.7, μ:1.9	0.062±0.003	[0,0.123]	0.9835	7.7×10^{-3}	2.00×10^{-2}
HS1 T:130 ns, β:10.9, μ:1.2	0.044±0.001	[0,0.089]	0.978	7.6×10^{-3}	2.0×10^{-2}
HS1 T:148 ns, β:5.7, μ:1.9	0.02069±0.0006	[0,0.0414]	0.9897	3.1×10^{-3}	1.13×10^{-2}
HS1 T:165 ns, β:12.90, μ:2.5	0.0357±0.002	[0,0.071]	0.982	9.9×10^{-3}	1.9×10^{-2}

[9]error estimated by nonlinear error propagation of fitted parameter variances

[10]uncertainty in L_1 and L_2 in last reported significant figure

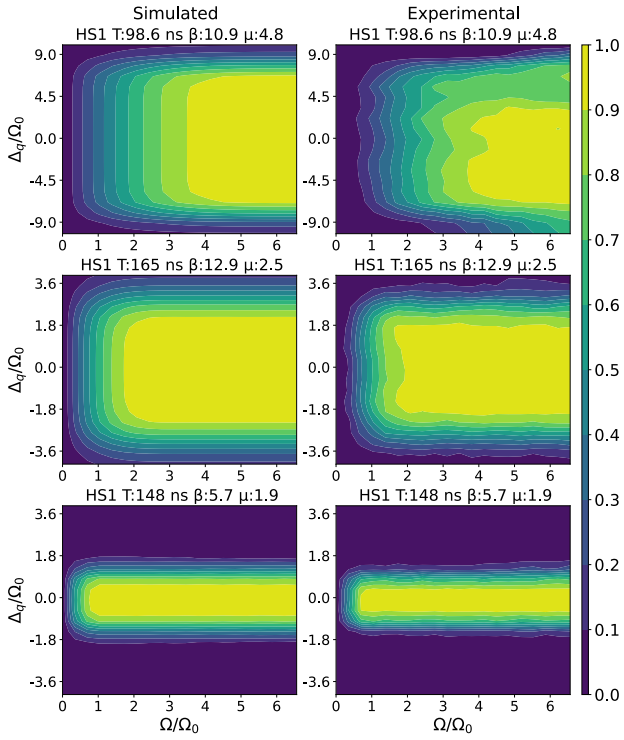

Figure 4.6: Simulated (left) and experimental (right) contours for selected HS1 pulses. Contours used an array of amplitudes ranging from 0 to 1 in steps of 0.05 ($\Delta\Omega/\Omega_0 = 0.3$) for all pulses. Frequencies for the 165 and 148 ns pulses ranged from -90 to 90 MHz with steps of 10 MHz ($\Delta\Delta_q/\Omega_0 = 0.45$), and those for the 98.6 ns HS1 pulse ranged from -225 to 225 in steps of 25 MHz ($\Delta\Delta_q/\Omega_0 = 1.12$).1024 trials were performed for each amplitude, frequency pair to obtain the average inversion probability.

4.5 Discussion

Overall, composite pulses and AFP pulses both proved effective at increasing the range of amplitude and frequency errors tolerated by single qubit inversion gates. Several of the

composite pulses ($90_x180_y90_x$, $180_{120}180_{240}180_{120}$, and CORPSE) also have improved on-resonance fidelities relative to the single 180_x pulses. Given the flexibility of many of the pulse sequences developed for quantum computing applications not only to act on general states but to perform arbitrary rotations, it is possible to implement general gates with improved robustness and potentially fidelity using composite pulses. Our results do not indicate that composite pulses formed of DRAG single pulses improve performance with respect to robustness, fidelity, or leakage. Within the precision afforded by our implementation of the LRB protocol, DRAG pulses had comparable to slightly elevated leakage rates, which is not unexpected for the longer pulses studied in this paper. Although the basis gateset of `ibmq_lima` uses DRAG pulses with 35.2 ns durations, experimental analyses have demonstrated that DRAG has no distinguishable effect in pulses longer than 20 ns since leakage in this regime is the result of incoherent processes [7]. In the incoherent leakage regime, leakage rates increase as a function of pulse duration. This trend is evident in the composite pulse data, where composite sequences were generally found to have higher leakage rates than single pulses. If the duration of the single pulses used in composite sequences were reduced below the default gate duration of 35.2 ns, it is possible that DRAG shaping or other leakage-optimized profiles with even shorter durations (4.16 ns) [42] could be employed to construct robust composite pulses with reduced leakage rates. AFP pulses proved to be the most robust pulses tested. Several achieved amplitude insensitive inversion at threshold microwave power below the nominal amplitude required to drive inversion in composite pulses, and all possessed frequency compensation bandwidths broader than the most robust of the composite pulses. Additionally, the successful implementation of AFP pulses only 2.8 to 5 times longer than single pulses makes composite AFP schemes possible, when otherwise such pulses would consume an untenable proportion of the intrinsic coherence time. The specific parameterizations of the HS1 pulses tested in this study were affected by higher leakage rates and lower fidelities than the composite pulses. This is not surprising given their exceedingly large bandwidth. It may be possible to negotiate a balance between bandwidth robustness and leakage such that these errors are reduced.

CHAPTER 5

Conclusion and Future Prospects

5.1 Conclusion

Composite and adiabatic pulses are both robust control methods that can be readily extended to implement general unitary rotations. The improvements in robustness and, in certain schemes, fidelity observed in this study illustrate the utility of composite and AFP pulses in coherent control of transmon qubits. Our optimization of reduced duration adiabatic pulses and previously proposed methods of shortening nonadiabatic gate time [42] make implementing composite pulses with higher concatenation levels feasible for adiabatic or nonadiabatic single pulses. The improvements in fidelity and robustness observed for many of the inversion pulses tested suggest that more robust gatesets could be constructed from composite or AFP universal rotors. In order to simultaneously compensate for leakage, further investigation of single pulse shaping within composite sequences is needed. Gateset-average leakage rates determined by implementation of the theoretical extension to the LRB protocol proposed in [8] proved sufficiently sensitive to reveal explicable trends in the leakage as a function of pulse duration and bandwidth even without the ability to directly measure the leakage subspace. However, a protocol to benchmark leakage of individual gates would undoubtedly benefit future robust gate design efforts because the building blocks for universal rotations are often point-to-point transformations. Pulsed NMR techniques present a well-developed template for robust gate design that can be easily applied to NISQ superconducting qubit architectures, and further development of these techniques has the potential to benefit control on these processors.

Future prospects for robust adiabatic and composite control pulse design naturally cen-

ter on the realization of universal gatesets composed of these types of pulses. Given the improvements in robustness and fidelity observed for many of the composite and adiabatic inversion pulses studied in this work, it is evident that gatesets constructed entirely of composite or adiabatic general rotors have the potential to achieve fault-tolerant fidelities over a wider range of control field calibrations. Previous implementation of universal, fault-tolerant gatesets on other architectures using composite pulses also bears out this potential: single and two-qubit composite pulse gates on diamond spin qubits have been optimized to achieve fidelities in excess of $\mathcal{F} = 0.99$ in the presence of control errors [43]. Unlike in the case of superconducting qubits, fidelities for NV centers in the absence of control errors are typically below fault-tolerance thresholds, so it may be possible to construct even higher fidelity composite pulse gatesets for superconducting transmon qubits. Characterization methods based on randomized benchmarking are also better suited to the estimation of gateset-average properties. While we were able to implement interleaved randomized benchmarking and leakage randomized benchmarking to characterize the performance of individual inversion pulses, LRB only yielded gateset-average rates and IRB results may also have been influenced by gate-dependence in the error rates of different Clifford gates. A homogeneous gateset consisting entirely of composite or adiabatic pulses would be better modeled by average fidelities and leakage rates. Improved characterization accuracy opens the door to further optimization and would be particularly useful for closed-loop optimization, which has been successfully used in conjunction with LRB to design leakage-optimal short single pulses [42]. The development of an array of similar single-pulse design techniques has produced a sophisticated system for coherent control of superconducting transmon qubits that routinely yields gate fidelities approaching fault tolerance thresholds. Composite and adiabatic pulses are inherently robust and compatible with many of these schemes. Incorporating pulsed NMR techniques in superconducting qubit control combines decades of experience in both fields and may ultimately produce gatesets with robustness and fidelity greater than that realized in either field alone.

CHAPTER 6

Appendices

6.1 Appendix A: Transmon Quantum Electromagnetic Circuit Hamiltonian

Transmission line shunted plasma oscillation (transmon) qubits are superconducting LC-circuit-based qubits whose properties are determined by the inductive and capacitive energies of circuit elements. Superconducting quantum electromagnetic circuit qubits belong to a fascinating class of mesoscopic systems: mesoscopic systems are macroscopic in the sense that they contain a large number of atoms, but they also possess collective degrees of freedom that exhibit quantum mechanical behavior, as would be expected for microscopic particles.

In this appendix, we discuss the process of deriving the quantum mechanical Hamiltonian for macroscopic superconducting circuit-based qubits and connect the quantized transmon circuit Hamiltonian to the spin Hamiltonian representation used throughout this thesis. Derivations are adapted from [6].

6.1.1 Quantization of the Circuit Hamiltonian

The relationship between macroscopic inductive and capacitive elements and the quantum mechanical behavior of these circuits can be derived using the branch-flux method, which is premised on circuit laws and classical Lagrangian mechanics [44]. In this case, the generalized coordinate is the flux $\Phi(t)$, and the canonically conjugate momentum coordinate is the charge $Q(t)$

$$\Phi(t) = \int_{-\infty}^{t} V(t')dt' \quad \text{and} \quad Q(t) = \int_{-\infty}^{t} I(t')dt'$$

where $V(t')$ and $I(t')$ are the voltage and current flowing across the transmon at time t'. The total instantaneous energy of the transmon circuit

$$E(t) = \int_{-\infty}^{t} V(t')I(t')dt'$$

can be divided into expressions for the potential $U_L(t)$ and kinetic energy $\tau_C(t)$ in terms of the generalized coordinate using the relationship between the voltage across the capacitor $V = LdI/dt$ and current through the inductor $I = CdV/dt$.

$$U_L(t) = \int_{-\infty}^{t} L\frac{dI(t')}{dt'}I(t')dt' = \frac{1}{2}LI(t)^2 = \frac{1}{2L}\Phi(t)^2$$
$$\tau_C(t) = \int_{-\infty}^{t} C\frac{dV(t')}{dt'}V(t')dt' = \frac{1}{2}CV(t)^2 = \frac{1}{2}C\dot{\Phi} \quad \text{where} \quad \dot{\Phi} = \frac{d\Phi(t')}{dt'}.$$

The Lagrangian $\mathcal{L}_g$ is constructed as the difference between the kinetic and potential energy terms

$$\mathcal{L}_g = \tau_C - U_L = \frac{1}{2}C\dot{\Phi}^2 - \frac{1}{2L}\Phi^2.$$

The Hamiltonian of quantized LC-circuit is obtained using a Legendre transformation to find the momentum conjugate of the flux, which is the charge on the capacitor in this case

$$\frac{\partial\mathcal{L}_g}{\partial\dot{\Phi}} = C\dot{\Phi} = Q(t).$$

The classical Hamiltonian is defined in terms of the Lagrangian as

$$H = Q\dot{\Phi}(t) - \mathcal{L}_g = \frac{Q^2}{2C} + \frac{\Phi^2}{2L},$$

from which the quantum Hamiltonian is constructed by promoting the canonically conjugate flux and charge coordinates to operators ($Q(t) \to \hat{Q},\ \Phi(t) \to \hat{\Phi}$) and making an association between the classical Poisson bracket $\{\Phi, Q\}$ and the commutator of quantum mechanics $[\hat{\Phi}, \hat{Q}]$.

$$\{\Phi, Q\} = \frac{\delta\Phi}{\delta\Phi}\frac{\delta Q}{\delta Q} - \frac{\delta Q}{\delta\Phi}\frac{\delta\Phi}{\delta Q} \Longleftrightarrow \frac{1}{i}[\hat{\Phi}, \hat{Q}] = \frac{1}{i}\left(\hat{\Phi}\hat{Q} - \hat{Q}\hat{\Phi}\right).$$

Evaluating the Poisson bracket in terms of the functional derivatives $\frac{\delta}{\delta\Phi}$ and $\frac{\delta}{\delta Q}$ yields the commutation relationship for the operators $[\hat{\Phi}, \hat{Q}] = i$. The quantized circuit Hamiltonian of a general LC-circuit is then

$$\hat{H} = \frac{\hat{Q}^2}{2C} + \frac{\hat{\Phi}^2}{2L}.$$

For a transmon LC-circuit, the form of the inductance is nonlinear, resulting in a modification of the potential energy term. The superconducting Josephson junction in a transmon qubit is a nonlinear, dissipationless inductor with a small self-capacitance C_J that satisfies special relations between the current, voltage, and flux across the junction.

$$I(t) = I_c \sin\left(\frac{2\pi\Phi(t)}{\Phi_0}\right) \quad \text{and} \quad V(t) = \frac{\hbar}{2e}\frac{d}{dt}\left(\frac{2\pi\Phi(t)}{\Phi_0}\right).$$

Here I_c is the critical current through the superconductor and $\Phi_0 = h/2e$ is the flux quantum for Cooper pairs of superconducting electrons (Note for this section of the appendix only, we explicitly write $\hbar, e$, which are otherwise equal to 1 in Hartree atomic units, to support intuition regarding the quantized flux of Cooper pairs). Evaluating the expression for the potential energy using the Josephson relation for the current yields the quantized Hamiltonian of a transmon qubit

$$\hat{H}_{tr} = \frac{\hat{Q}^2}{2C} - \frac{I_c\Phi_0}{2\pi}\cos\left(\frac{2\pi\hat{\Phi}}{\Phi_0}\right)$$

where the total capacitance $C = C_s + C_J$ now includes the Josephson junction self-capacitance in addition to the shunt capacitor.

This Hamiltonian can also be represented in terms of the reduced charge $\hat{n} = \hat{Q}/2e$ and reduced flux $\hat{\phi} = 2\pi\hat{\Phi}/\Phi_0$

$$\hat{H}_{tr} = 4E_c\hat{n}^2 - E_J\cos(\hat{\phi}).$$

The reduced operators $\hat{n}$ and $\hat{\phi}$ correspond to the number of Cooper pairs and the gauge-invariant phase across the Josephson junction respectively. The terms E_C and E_J represent the stored capacitive and inductive energies and are defined as $E_C = e^2/2C$ and $E_J =$

$I_c\Phi_0/2\pi$.

6.1.2 Second Quantized Transmon Hamiltonian

As an intermediate step in the derivation of the spin Hamiltonian for transmon qubits, it is necessary to rewrite the quantized circuit Hamiltonian in terms of creation and annihilation operators (second quantization formalism). The creation and annihilation operators $\hat{c}^\dagger$ and $\hat{c}$ for the transmon are defined for the energy levels $|j\rangle$ of the anharmonic potential as

$$\hat{c}^\dagger = \sum_j \sqrt{j+1}\,|j+1\rangle\langle j| \quad \text{and} \quad \hat{c} = \sum_j \sqrt{j}\,|j-1\rangle\langle j|\,.$$

For a transmon, the eigenstates $|j\rangle$ are defined by the difference in the number of Cooper pairs (excess charge quanta) between the superconducting islands of the Josephson junction. The raising operator increases this charge difference, and the lowering operator reduces the charge difference [45]. In terms of these raising and lowering operators, the reduced charge and phase are

$$\hat{n} = in_{zpf}(\hat{c}^\dagger - \hat{c}) \quad \text{and} \quad \hat{\phi} = \phi_{zpf}(\hat{c}^\dagger + \hat{c}), \quad \text{where}$$

$$n_{\text{zpf}} = \left(\frac{E_J}{32E_c}\right)^{1/4} \quad \text{and} \quad \phi_{\text{zpf}} = \left(\frac{2E_c}{E_J}\right)^{1/4}.$$

Further simplification of $\hat{H}_{tr}$ is possible by considering that $\hat{\phi} \ll 1$ since, by design, transmon qubits operate in the regime $E_J/E_C \gg 1$. This allows us to Taylor expand $\cos\hat{\phi}$ to get

$$\hat{H}_{tr} = -4E_c n_{zpf}^2(\hat{c}^\dagger - \hat{c})^2 - E_J\left(1 - \frac{1}{2}\phi_{zpf}^2(\hat{c}^\dagger + \hat{c})^2 + \frac{1}{24}\phi_{zpf}^4(\hat{c}^\dagger + \hat{c})^4 + \dots\right)$$

$$\approx \sqrt{8E_cE_J}\left(\hat{c}^\dagger\hat{c} + \frac{1}{2}\right) - E_J - \frac{E_c}{12}(\hat{c}^\dagger + \hat{c})^4.$$

Here, the substitutions $8E_cn_{zpf}^2 = E_J\phi_{zpf}^2 = \sqrt{2E_cE_J}$ have been performed and the rotating wave approximation has been applied to drop rapidly rotating terms (uneven number of c and $c^\dagger$), resulting in the approximate Hamiltonian in the second line. Finally, defining the qubit resonant frequency $\omega_0 = \sqrt{8E_cE_J}$ and the anhamonicity $\delta = -E_c$ in terms of the

capacitive and inductive energies, the second quantized transmon Hamiltonian is obtained

$$\hat{H}_{tr} = \omega_0 \hat{c}^\dagger \hat{c} + \frac{\delta}{2}((\hat{c}^\dagger \hat{c})^2 + \hat{c}^\dagger \hat{c}) = \left(\omega_0 + \frac{\delta}{2}\right) \hat{c}^\dagger \hat{c} + \frac{\delta}{2}(\hat{c}^\dagger \hat{c})^2.$$

The frequency of the transmon is given by $\omega \equiv \omega_0 + \delta$ such that the spacing between each pair of energy levels differs according to $\omega_{j+1} - \omega_j = \omega + \delta j$. A last simplification of the transmon Hamiltonian can be made by using ω and replacing the number operator $\hat{c}^\dagger \hat{c}$ by its definition $\hat{c}^\dagger \hat{c} = \sum_j j \left|j\right\rangle \left\langle j\right|$

$$\hat{H}_{tr} = \omega \hat{c}^\dagger \hat{c} + \frac{\delta}{2} \hat{c}^\dagger \hat{c}(\hat{c}^\dagger \hat{c} - 1) = \sum_j \left(\left(\omega - \frac{\delta}{2}\right) j + \frac{\delta}{2} j^2 \right) \left|j\right\rangle \left\langle j\right| = \sum_j \omega_j \left|j\right\rangle \left\langle j\right|.$$

From this Hamiltonian, the transmon energy levels as a function of the anharmonicity are found to be $\omega_j = \left(\omega - \frac{\delta}{2}\right) j + \frac{\delta}{2} j^2$.

6.1.3 Fermion to Spin Hamiltonian Mapping

Although the transmon is an anharmonic oscillator with accessible higher energy states outside of two-level the computational subspace $\{\left|0\right\rangle, \left|1\right\rangle\}$, it is reasonable to truncate the Hamiltonian provided the anharmonicity $\delta = -E_c$ is sufficiently large so as to approximately isolate the first two states. The anharmonicity decreases the spacing between higher energy levels so that the frequency necessary to drive a transition outside the computational subspace (ie. $\left|1\right\rangle$ to $\left|2\right\rangle$) is reduced relative to the transition from the ground to first excited state. As a result, the frequency of the control field can be tuned to address a unique transition, and the second quantized transmon Hamiltonian can be restricted to the first two energy levels

$$\hat{H}_{tr} = \sum_{\omega_j=0}^{1} \omega_j \left|j\right\rangle \left\langle j\right| = 0 \left|0\right\rangle \left\langle 0\right| + [(\omega - \frac{\delta}{2}) + \frac{\delta}{2}] \left|1\right\rangle \left\langle 1\right| = 0 \left|0\right\rangle \left\langle 0\right| + \omega_0 \left|1\right\rangle \left\langle 1\right|.$$

The Hamiltonian can be symmetrized around zero without influencing the dynamics by subtracting half the qubit frequency ω_0.

$$\hat{H}_{tr} = -\frac{1}{2}\omega_0 \left|0\right\rangle \left\langle 0\right| + \frac{1}{2}\omega_0 \left|1\right\rangle \left\langle 1\right|.$$

Substituting the definition of the Pauli matrix ($\sigma_z = \left|0\right\rangle \left\langle 0\right| - \left|1\right\rangle \left\langle 1\right|$) yields the spin Hamiltonian of the undriven transmon qubit.

$$\hat{H}_{tr} = -\frac{1}{2}\omega_0 \sigma_z.$$

6.1.4 Spin Hamiltonian of a Driven Transmon

Control operations on transmon qubits are accomplished by driving transitions (ideally) within the computational subspace with an electric field of the form $\vec{E}(t) = \vec{E}_0(t)e^{-i\omega_d t} + \vec{E}_0(t)^* e^{i\omega_d t}$. Continuing to approximate the transmon as a two-level system allows the dipole operator $\vec{d} = \vec{d}_0\sigma_+ + \vec{d}_0^*\sigma_-$ for the transmon to be defined terms of the qubit raising and lowering operators $\sigma_\pm = (1/2)(\sigma_x \mp i\sigma_y)$, which act on states in the computational subspace as $\sigma_+ \left|0\right\rangle = \left|1\right\rangle$ and $\sigma_- \left|1\right\rangle = \left|0\right\rangle$. The Hamiltonian for the qubit-drive interaction in the stationary frame is

$$\hat{H}_E = -\vec{d}\cdot\vec{E}(t) = -\left(\vec{d}_0\cdot\vec{E}_0 e^{-i\omega_d t} + \vec{d}_0\cdot\vec{E}_0^* e^{i\omega_d t}\right)\sigma_+ - \left(\vec{d}_0^*\cdot\vec{E}_0 e^{-i\omega_d t} + \vec{d}_0^*\cdot\vec{E}_0^* e^{i\omega_d t}\right)\sigma_-$$

$$\equiv \left(\Omega e^{-i\omega_d t} + \tilde{\Omega} e^{i\omega_d t}\right)\sigma_+ + \left(\tilde{\Omega}^* e^{-i\omega_d t} + \Omega^* e^{i\omega_d t}\right)\sigma_-$$

where time dependence of the electric field amplitude $E_0(t)$ has been suppressed in the first line, and the substitutions $\Omega \equiv -\vec{d}_0\cdot\vec{E}_0(t)$ and $\tilde{\Omega} \equiv -\vec{d}_0\cdot\vec{E}_0(t)^*$ have been employed in the second.

The full Hamiltonian of the driven transmon in the stationary frame is the sum of the qubit and drive Hamiltonians $\hat{H} = \hat{H}_0 + \hat{H}_E$. Transforming the drive Hamiltonian into the interaction frame (a rotating frame at the qubit resonant frequency ω_0) using the propagator

$U = \exp(i\hat{H}_0 t)$ produces the interaction Hamiltonian $\hat{H}_{E,I} = U\hat{H}_E U^\dagger$

$$\hat{H}_{E,I} = U\hat{H}_E U^\dagger = \left(\Omega e^{-i(\omega_d-\omega_0)t} + \tilde{\Omega} e^{i(\omega_d+\omega_0)t}\right)\sigma_+ + \left(\tilde{\Omega}^* e^{-i(\omega_d+\omega_0)t} + \Omega^* e^{i(\omega_d-\omega_0)t}\right)\sigma_-.$$

Under the rotating wave approximation, the rapidly oscillating terms with frequencies $(\omega_0 + \omega_d)$ are dropped, and only terms with with frequencies $\Delta_q = \omega_d - \omega_0$ remain in the interaction Hamiltonian.

$$\hat{H}_E^{RWA} = \Omega e^{-i\Delta_q t}\sigma_+ + \Omega^* e^{i\Delta_q t}\sigma_-.$$

To obtain the drive Hamiltonian in the rotating frame of the microwave drive, it is necessary to first transform the rotating wave approximated Hamiltonian back to the stationary frame

$$\hat{H}_E^{RWA} = U^\dagger \hat{H}_{E,I}^{RWA} U = \Omega e^{-i\omega_d t}\sigma_+ + \Omega^* e^{i\omega_d t}\sigma_-$$

such that the total qubit and drive Hamiltonian is

$$\hat{H}^{RWA} = -\frac{1}{2}\omega_0\sigma_z + \Omega e^{-i\omega_d t}\sigma_+ + \Omega^* e^{i\omega_d t}\sigma_-.$$

Then, a final frame change using the propagator $U_E = \exp(-i\omega_d\sigma_z t/2)$ brings the Hamiltonian into the frame of the drive

$$\hat{H}_d = U_E\hat{H}^{RWA}U_E^\dagger - U_E\dot{U}_E^\dagger = -\frac{1}{2}\omega_0\sigma_z + \Omega\sigma_+ + \Omega^*\sigma_- + \frac{1}{2}\omega_d\sigma_z.$$

Algebraic rearrangements then produce the familiar form of the Hamiltonian for the transmon under the influence of a microwave drive used throughout this thesis

$$\hat{H}_d = \frac{\sigma_z}{2}\Delta_q + \Re[\Omega(t)]\frac{\sigma_x}{2} + \Im[\Omega(t)]\frac{\sigma_y}{2}.$$

6.2 Appendix B: Average Hamiltonian Theory

Average Hamiltonian theory (AHT) is a propagator expansion technique that aims to approximate the evolution of a system under a time dependent Hamiltonian by a series expansion of a time independent effective Hamiltonian. A systematic method of composite pulse design employing AHT can be developed by considering the interaction Hamiltonian of the qubit with the microwave drive.

$$\hat{H}_d(t) = \frac{\sigma_z}{2}\Delta_q(t) + \frac{|\Omega(t)|}{2}\left(\sigma_x \cos(\phi(t)) + \sigma_y \sin(\phi(t))\right).$$

Here, $\Delta_q = \omega_d - \omega_q$ is the angular frequency offset from resonance, and the drive field is described by an amplitude $\Omega(t)$ and phase $\phi(t)$. This form of the Hamiltonian in the rotating frame (resonant with ω_q) suggests the definition of an effective field and rotation axis such that

$$\hat{H}_d = \Omega_{\text{eff}}(t)\left(\frac{\sigma_x}{2}\sin(\theta(t))\cos(\phi(t)) + \frac{\sigma_y}{2}\sin(\theta(t))\sin(\phi(t)) + \frac{\sigma_z}{2}\cos(\theta(t))\right)$$

where

$$\Omega_{\text{eff}} = \sqrt{|\Omega(t)|^2 + \Delta_q(t)^2} \;\text{ and }\; \theta(t) = \arctan\left(\frac{|\Omega(t)|}{\Delta_q(t)}\right).$$

This allows the discrete propagator over a small timestep to be defined simply in terms of an angle-axis representation of the rotation

$$U_d(t+\Delta t, t) = \exp\left(-\frac{i\Omega_{\text{eff}}(t)}{2}(\mathbf{n}\cdot\boldsymbol{\sigma})\Delta t\right) = \exp\left(-i\beta_{\text{eff}}\,\mathbf{n}\cdot\boldsymbol{\sigma}/2\right).$$

Here $\mathbf{n}(t) = (\sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta))$ is the normalized direction vector at time t (the angles are understood to be a function of time), and $\boldsymbol{\sigma}$ is the vector of Pauli spin-matrices $(\sigma_x, \sigma_y, \sigma_z)$. This propagator generates an infinitesimal rotation $d\beta_{\text{eff}} = \Omega_{\text{eff}}\Delta t$ about $\mathbf{n}$ during the time step Δt that is related to the nominal rotation $d\beta$ about the on-

resonance rotation axis $\mathbf{n}_0 = (\cos(\phi), \sin(\phi), 0)$ as follows

$$d\beta_{\text{eff}} = \sqrt{\frac{|\Omega(t)|^2 + \Delta_q(t)^2}{|\Omega(t)|^2}} d\beta.$$

Since the interaction Hamiltonian is time dependent, and will therefore generally have at least one set of non-commuting time points during the duration of the pulse (t_a, t_b) $[H(t'), H(t'')] \neq 0$, an analytical expression is typically not obtainable. An approximate form of the propagator $U_d(t_a, t_b)$, is however, derivable using average Hamiltonian theory.

The general form of the propagator for a time dependent Hamiltonian may be written as follows

$$U(t_a, t_b) = \mathcal{T} \exp\left(\int_{t_a}^{t_b} \epsilon H(t) dt\right)$$

where ϵ is a small parameter with respect to which $U(t_a, t_b)$ has a convergent cumulant expansion.

$$\exp\left(\epsilon \overline{H}^{(1)} + \epsilon^2 \overline{H}^{(2)} + \ldots\right).$$

Applying the Wilcox formula [46] for the derivative of the matrix valued function,

$$\frac{d}{dx} \exp\left(A(x)\right) = \int_0^1 e^{-\lambda A(x)} \frac{d}{dx} A(x) e^{-\lambda A(x)} d\lambda \; e^{A(x)}$$

we obtain the Magnus expansion for the time independent effective Hamiltonian $\overline{H}$ that effects the same evolution as $H(t)$ over the full interval (t_a, t_b).

$$U(t_a, t_b) = \exp\left(-i\overline{H}(t_a - t_b)\right).$$

The first three orders of the Magnus expansion ($\overline{H} = \overline{H}^{(1)} + \overline{H}^{(2)} + \overline{H}^{(3)} + \ldots$) are

$$\overline{H}^{(1)} = \frac{1}{t_a - t_b} \int_{t_a}^{t_b} dt H(t)$$
$$\overline{H}^{(2)} = \frac{1}{2i(t_a - t_b)} \int_{t_a}^{t_b} dt \int_{t_a}^{t} dt' [H(t), H(t')]$$
$$\overline{H}^{(3)} = -\frac{1}{6(t_a - t_b)} \int_{t_a}^{t_b} dt \int_{t_a}^{t} dt' \int_{t_a}^{t'} dt'' \left\{ [H(t), [H(t'), H(t'')]] + [[H(t), H(t')], H(t'')] \right\}.$$

The Magnus expansion is only convergent to low order if the magnitude of the Hamiltonian is small, or more precisely $||H(t)(t_a - t_b)|| << 1$ for any time $t_a \leq t \leq t_b$. The Hamiltonian of the microwave pulse drive itself does not generally satisfy this condition, so it is necessary to separate the drive into a nominal component and a component containing the amplitude $\Omega_\epsilon(t) = \Omega_{\text{eff}}(t) - \Omega_0(t)$ and frequency Δ_q errors.

$$\hat{H}_d = \hat{H}_{\Omega_0(t)} + \hat{H}_{\Omega_\epsilon(t), \Delta_q(t)}.$$

Changing to the interaction frame of the nominal drive, we obtain a small error Hamiltonian, which can be expanded convergently

$$\widetilde{H}_{\Omega_\epsilon, \Delta_q}(t) = U^\dagger_{\Omega_0}(t, t_a) \hat{H}_{\Omega_\epsilon, \Delta_q}(t) U_{\Omega_0}(t, t_a).$$

Here the propagator for the nominal drive is

$$U_{\Omega_0} = \exp\left\{ -\int_{t_a}^{t} i\hat{H}_{\Omega_0}(t) dt \right\}.$$

The goal of composite pulse design using Average Hamiltonian Theory then becomes to design a pulse sequence such that the terms of the expansion

$$\widetilde{H}_{\Omega_\epsilon, \Delta_q}(t) = \widetilde{H}^{(1)} + \widetilde{H}^{(2)} + \widetilde{H}^{(3)} + \ldots$$

are zero up to some desired order. From the perspective of the resonance frame of the qubit

(ie. reversing the transformation to the frame of the drive), the full propagator is then,

$$U(t_b, t_a) = U_{\Omega_0}(t_b, t_a)\widetilde{U}_{\Omega_\epsilon, \Delta_q}(t_b, t_a) = U_{\Omega_0}(t_b, t_a) \exp\left\{-i \int_{t_a}^{t_b} \widetilde{H}^{(1)} + \widetilde{H}^{(2)} + \dots\right\}$$

$$\approx U_{\Omega_0}(t_b, t_a) * 1$$

which is exactly the propagator of the desired nominal rotation component. Specific examples of the use of this technique in composite pulse design can be found in [47] and [34]. Related techniques employing similar expansions of the error term have also been proposed for the design of general rotors to purportedly arbitrary precision [12].

6.3 Appendix C: Pulse Simulation Code

The pulse simulations described in chapter 4 were implemented using the following code. Each pulse in the composite pulse schedule was initialized as a member of a class corresponding to the appropriate single pulse shaping. The definition of the Gaussian class is included below.

```
class GaussianPulse():
    '''Gaussian pulse class that generates Qiskit Pulse object (a.u.) used for experimental schedules and pulse
    ↪  profile (rad/sec) used for simulation purposes
      -- for RB experiments, calibrated pi amplitude and desired pi multiple used to specify ideal flip angle
      -- for contour plot generation, pi_amp set to 1 and amplitude scaling performed by simulator'''
    def __init__(self, pi_amp, pi_mult, phase, duration_samples, sigma, name):
        self.type = 'Gaussian'
        self.name = name
        self.Tp = duration_samples*dt    #pulse time s
        self.pi_mult = pi_mult  #desired rotation angle specified as a multiple of pi
        self.phase = phase
        gaussian = pulse.Gaussian(duration=duration_samples, sigma = sigma, amp =
        ↪  pi_amp*pi_mult*(np.cos(phase)+1j*np.sin(phase)), name = f'{name} ideal', limit_amplitude = True)
        ↪  #Qiskit pulse object
        self.pulse_obj = gaussian
        self.waveform= gaussian.get_waveform()
        self.pulse_profile = omega1_conv*self.waveform.samples #multiply by rad/a.u to get amplitude in rad/sec
        ↪  #complex valued waveform (rad/sec)
```

Composite pulses were simulated using an array of single-pulse objects stored in an array, `Pulses`. For simulations, `pulse_profile` attribute that stores the complex-valued waveform in rad/sec was extracted from the pulse objects and concatenated to construct a composite pulse profile. The nominal drive amplitude in a.u scaling the complex-valued waveform was swept over an array of amplitudes, and a range of frequency offsets was specified in Hz.

```
def contour_plot_simulator(Pulses, amp_array, freqoffset_array, center_frequency, title=None, plotBool =
↪  False):
    rf= np.array([], dtype = complex)
    for p in Pulses:
        rf_temp = p.pulse_profile
        rf = np.append(rf, rf_temp) #append each pulse in the schedule

    f= freqoffset_array  #offsets from the resonant frequency (Hz)
    a= amp_array         #amplitudes to scale pulse profile by in a.u.
    t= np.linspace(0, len(rf)*dt,len(rf))

    rf_amp_func = np.abs(rf)
    rf_phase_func = np.angle(rf)

    #Define the Hamiltonian and Simulate the Pulse
    E = np.array([[1,0],[0,1]])
    Sx = np.array([[0,0.5],[0.5,0]])
    Sy = np.array([[0, -0.5*1j],[0.5*1j,0]],dtype =complex)
    Sz = np.array([[0.5,0],[0,-0.5]])

    prob_down = np.zeros([len(f), len(a)])  # array to store inversion probabilities for each amplitude,
    ↪  frequency pair
    for kf in range(len(f)):
        for ka in range(len(a)):
            U = E # initiate propagator as identity matrix
            for kt in range(len(rf)):
                H= Sz*2*np.pi*f[kf]
                ↪  +(a[ka]*rf_amp_func[kt])*(Sx*np.cos(rf_phase_func[kt])+Sy*np.sin(rf_phase_func[kt]))
                Udt = scipy.linalg.expm(-1j*H*dt)
                U = np.matmul(Udt,U)
            psi = np.matmul(U,[1,0])  #propage initial spin up state
            prob_down[kf,ka] = abs(np.matmul([0,1],psi)) #project onto final spin down state and store
            ↪  probability

    if plotBool ==True: #plot the microwave pulse schedule profile
        fig, axs= plt.subplots(1,3, constrained_layout=True)
```

```
        fig.set_size_inches(12,4)

        axs[0].plot(t, np.real(rf), label = 'RF Profile Real')
        axs[0].plot(t, np.imag(rf), label = 'RF Profile Imaginary' )
        axs[0].set_title('Real and Imaginary Components')
        axs[0].set_xlabel('time (s)')
        axs[0].set_ylabel('amplitude (rad/sec)')
        axs[0].legend()

        twiny = axs[1].twinx()
        axs[1].plot(t, np.abs(rf), label = 'RF Amplitude (left)')
        twiny.plot(t, np.angle(rf), color= 'orange', label = 'RF Phase (right)' )
        axs[1].set_title('Amplitude and Phase')
        axs[1].set_xlabel('time (s)')
        axs[1].set_ylabel('amplitude (rad/sec)')
        axs[1].legend()
        twiny.legend()

        # plot the simulated inversion probability contour
        simulated_contour_generator(axs[2],fig, amp_ax = a, freq_offsets = f, prb_down=
        ↪  prob_down,center_frequency=center_frequency, title =title)
    return prob_down
```

The inversion probability was stored for each amplitude, frequency offset pair and plotted with respect to the dimensionless parameters Ω/Ω_0 and Δ_q/Ω_0 by the `simulated_contour_generator`.

```
def simulated_contour_generator(pltax, fig, amp_ax, freq_offsets, title, prb_down, center_frequency):
    #Generate Contour Plot
    X,Y = np.meshgrid(amp_ax,(freq_offsets+center_frequency)*1e-9) # X of dim(N)=num cols in prob_down,
                                                                   #Y of dim(M) = num rows in prob_down
    cp = pltax.contourf(X,Y, prb_down, levels = np.arange(0,1.1,0.1))
    fig.colorbar(cp, ax = pltax, ticks = np.arange(0,1.1,0.1))
    pltax.set_title(title)
    pltax.set_yticks(np.round((freq_offsets[::4]+center_frequency)*1e-9,3),)
    pltax.set_yticklabels(np.round((freq_offsets[::4])*2*np.pi/(omega1_conv*x180_piamp),2)) #dimensionless
    ↪  parameters normalized by calibrated pi amplitude (x180_piamp) and a.u. to rad/sec conversion factor
    pltax.set_xticks(np.arange(0,max(amp_ax/x180_piamp),0.5)*x180_piamp)
    pltax.set_xticklabels(np.round(np.arange(0,max(amp_ax/x180_piamp),0.5),1))
    ax.set_xlabel(r'$\Omega/\Omega_0$')
    ax.set_ylabel(r'$\Delta_q/\Omega_0$')
```

6.4 Appendix D: Genetic Algorithm for Adiabatic Pulse Design

The cost function used to design adiabatic pulses was used to optimize the performance of solutions of the form [duration (dt), β, μ] based on several criteria including adiabaticity, inversion probability, pulse duration, and bandwidth. Since each criterion is of a different order of magnitude–inversion probability ranges from 0 to 1 while bandwidth is on the order of 10^8 Hz–terms in the sum were rescaled to contribute more equally to the cost function. The specific scaling factors used were adjusted based on simulated and experimental performance of the generated pulses.

An `HS1_pulse` class was defined to convert solution parameters [duration (dt), β, μ] into a complex-valued pulse profile used to simulate the inversion probability. Inversion probabilities were simulated using a modified version of the `contour_plot_simulator` (called `inversion_simulator`) that only performed simulations for a single user-specified amplitude on resonance ($\Delta_q = 0$) rather than an array of amplitudes and frequencies. The `HS1_pulse` class is defined below.

```
class HS1_pulse():
    '''Given a solution for the pulse parameters generates a Qiskit Pulse object (a.u.) used for experimental
    ↪  schedules and pulse profile (rad/sec) used for simulation purposes''
    def __init__(self, solution, name):
        self.type = 'HS1'
        duration_samples = int(solution[0])# length of the microwave pulse in samples
        t= np.linspace(0,duration_samples,duration_samples)*dt
        Tp= duration_samples*dt
        beta = solution[1]
        mu = solution[2]
        self.mu = mu
        self.beta=beta
        self.t = t           #time axis in seconds
        self.Tp = Tp         #pulse duration in seconds
        self.pulse_profile = omega1_conv*np.power(sech(beta*(2*t/Tp-1)),(1+1j*mu))
        self.wave = qiskit.pulse.Waveform(np.power(sech(beta*(2*np.linspace(0,1,pulse_length)-1)),(1+1j*mu)))
        self.name = name
        self.omega1 = omega1_conv*max_amp*sech(beta*(2*t/Tp-1)) #Omega(t) amplitude of drive in rad/sec
        self.deltaomega = -mu*beta*2/Tp*tanh(beta*(2*t/Tp-1))   #Delta_q(t) detuning of drive in rad/sec
```

The objective function maximized by the Pygad genetic algorithm was then implemented

by a sum of several terms, as shown below. Failure of the adiabaticity was penalized by a negative term proportional to the reciprocal minimum distance $\Omega_{\text{eff}} - |\frac{d\theta}{dt}|$ and by testing for flaws in the amplitude insensitivity of the inversion probability at several amplitudes using the `inversion_simulator`. High inversion fidelity was rewarded by a term proportional to the simulated inversion probability at 1 a.u. The bandwidth of the frequency modulation was rewarded by a positive term proportional to the amplitude of the frequency modulation. Longer pulse duration was penalized by a negative term proportional to the duration.

```
def fitness_func(solution, solution_idx):
    solution_pulse = HS1_pulse(solution= solution, Name = f'{solution}')

    #define large penalty for critical profile failures
    bigpenalty =  -1e18

    #define quantities for adiabaticity condition
    omega1 = solution_pulse.omega1          #amplitude
    deltaomega = solution_pulse.deltaomega #detuning
    omegaeff = np.sqrt(omega1**2+deltaomega**2)
    theta = np.arctan(omega1/deltaomega)
    for i in range(len(theta)):
        if theta[i]<0:
            theta[i]= theta[i]+np.pi
    dtheta_dt = (theta[1:]-theta[0:-1])/dt
    dtheta_dt = np.append(dtheta_dt, dtheta_dt[-1]) #derivative of rotation axis theta(t) change

    adicheck = True   #remains true if pulse adiabatic
    for index in range(len(dtheta_dt)):
        if dtheta_dt[index]>omegaeff[index]:
            adicheck = False
    if adicheck==False:
        #print('nonadiabatic')
        adiabatic_penalty = bigpenalty  #make huge to have very low fitness for non-adiabatic pulses
    else:
        adiabatic_penalty= -1/np.min(omegaeff-dtheta_dt)

    # Set up fidelity checks at various amplitudes to test for Rabi oscillations/check that the simulated pulse
    ↪  actually has an adiabatic profile
    fid1amp = inversion_simulator(solution_pulse, center_frequency, 0.2) #inversion probability on-resonance
    ↪  with amplitude scaling factor 0.2
    fid2amp = inversion_simulator(solution_pulse, center_frequency, 0.8)
    fid3amp = inversion_simulator(solution_pulse, center_frequency, 1)
```

```
    fid4amp = inversion_simulator(solution_pulse, center_frequency, 1.5)
    fid5amp = inversion_simulator(solution_pulse, center_frequency, 2)
    fid6amp = inversion_simulator(solution_pulse, center_frequency, 3)
    fid_diffs = np.array([np.abs(fid2amp-fid1amp),np.abs(fid3amp-fid2amp),
    np.abs(fid4amp-fid3amp),np.abs(fid5amp-fid4amp),np.abs(fid6amp-fid5amp)])

    if (fid_diffs< 0.08).all() == True: #inversion probability is amplitude-invariant, and profile is adiabatic
        fidelity = inversion_simulator(solution_pulse, center_frequency, 1)
    else:
        fidelity = bigpenalty

    bandwidth = solution[1]*solution[2]*2/solution_pulse.Tp  #amplitude of frequency modulation function
    time_penalty = - (solution[0]*dt)  #reciprocal duration in s

    #rescale terms in objective function
    fidelity_cost = fidelity*15
    bandwidth_objective = bandwidth*1e-7
    time_objective = time_penalty*1e8
    adiabatic_objective = adiabatic_penalty*2*1e8

    objective= fidelity_cost+bandwidth_objective+adiabatic_objective+time_objective #the function to be
    ↪  maximized by Pygad genetic algorithm
    return objective
```

The instance of the genetic algorithm was then generated with an initial population determined from combinations of parameters in the gene space (durations between 96 and 169 ns and β and μ parameters between 1 and 30) that generated simulated contours with amplitude-insensitive inversion.

```
def gaInst_init(init_pop, Tmin, Tmax): #Tmin in samples set to 432 (96 ns)  Tmax in samples set to 768 (169
↪  ns)
    initial_population = [[768. ,   5.2,   1.5],
                          [768. ,   4. ,   3. ],
                          [736. ,   4. ,   2.6],
                          [640. ,   5.2,   1.2],
                          [592. ,   4.5,   1.2],
                          [576. ,   4.8,   1.6],
                          [544. ,   5.2,   1.2]]
    gene_space = [{'low':Tmin, 'high':Tmax,'step':16}, {'low':1, 'high':30}, {'low':1, 'high':30}]
    num_generations = 200
    num_parents_mating = 4
    fitness_function =fitness_func
```

```
parent_selection_type = "tournament"
keep_parents = 2
crossover_type = "uniform"
mutation_type = "adaptive"
mutation_num_genes = [2,1]
ga_instance= pygad.GA(num_generations=num_generations,
                  num_parents_mating=num_parents_mating, initial_population = initial_population,
                  fitness_func=fitness_function,
                  parent_selection_type=parent_selection_type,
                  keep_parents=keep_parents,
                  crossover_type=crossover_type,
                  mutation_type=mutation_type,
                  mutation_num_genes=mutation_num_genes, gene_space =gene_space, on_generation =
                  ↪  on_generation, save_best_solutions=True)
return ga_instance
```

Bibliography

[1] K. Williams and L. S. Bouchard. Quantification of robustness, leakage and seepage errors for composite and adiabatic gates. submitted.

[2] P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26:1484–1509, 1997.

[3] R.P. Feynman. Simulating physics with computers. *Int. J. Theor. Phys.*, 21:467–488, 1982.

[4] J. Preskill. Quanutm computing in the nisq era and beyond. *Quantum*, 2:79, 2018.

[5] Y. Susuki, S. Uno, R. Raymond, T. Tanaka, T. Onodera, and N. Yamamoto. Amplitude estimation without phase estimation. *Quantum Information Processing*, 19(75), 2020.

[6] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver. A quantum engineer's guide to superconducting qubits. *Appl. Phys. Rev.*, 6:021318, 2019.

[7] Z. Chen, J. Kelly, C. Quintana, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A.G. Fowler, E. Lucero, A. Megrant E. Jeffrey, J. Mutus, M. Neeley, C. Neill, P.J.J. O'Malley, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T.C. White, A.N. Korotkov, and J. M. Martinis. Measuring and suppressing quantum state leakage in a superconducting qubit. *Phys. Rev. Lett.*, 116:020501, 2016.

[8] C. Wood and J. Gambetta. Quantification and characterization of leakage errors. *Phys. Rev. A.*, 97:032306, 2018.

[9] K. Brown, J. Harrow, and I. Chuang. Arbitrarily accurate composite pulse sequences. *Phys rev. A.*, 70:052318, 2004.

[10] N. Timoney, V. Elman, S. Glaser, C. Weiss, M. Johanning, W. Neuhauser, and C. Wunderlich. Error resistant single qubit gates with trapped ions. *Phys. Rev. A.*, 77:052334, 2018.

[11] G. Wolfowicz and J. Morton. Pulse techniques for quantum information processing. *eMagRes*, 5:1515–1528, 2016.

[12] W. Alway and J. Jones. Arbitrary precision composite pulses for NMR quantum computing. *J. Mag. Res*, 189:114–120, 2007.

[13] E. Collin, G. Ithier, A. Aassime, P. Joyez, D. Vion, and D. Esteve. NMR-like control of a quantum bit superconducting circuit. *Phys. Rev. Lett.*, 93(15):61–122, 2004.

[14] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, New York, NY, U.S.A., 10th anniversary edition, 2010.

[15] C. Dankert, R. Cleve, J. Emerson, , and E. Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A*, 80(012304), 2009.

[16] E. Magesan. Gaining information about a quantum channel via twirling. M.sc., University of Waterloo, Waterloo, Ontario, Canada, 2008.

[17] M. Barnhill. Extensions of randomized benchmarking. M.sc., University of Waterloo, Waterloo, Ontario, Canada, 2015.

[18] D. Manzano. A short introduction to the lindblad master equation. *AIP Advances*, 10:025106, 2020.

[19] J. Gyamfi. Fundamentals of quantum mechanics in lioville space. *Eur. J. Phys*, 41(063002), 2020.

[20] Bengtsson and K. Życzkowski. *Geometry of Quantum States: an Introduction to Quantum Entanglement*. Cambridge University Press, Cambridge, UK, 2006.

[21] R. Gilmore. *Lie Groups, Physics, and Geometry: an Introduction for Physicists, Engineers, and Chemists*. Cambridge University Press, New York, NY, U.S.A., 2008.

[22] J. Emerson, R. Alicki, and K. Życzkowski. Scalable noise estimation with random unitary operators. *Journal of Optics B: Quantum and Semiclassical Optics*, 7:S347, 2005.

[23] D. Gottesman. The heisenberg representation of quantum computers. In S. P. Corney, R. Delbourgo, and P. D. Jarvis, editors, *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, pages 32–43, Cambridge, MA, 1999. International Press.

[24] D. DiVincenzo. The physical implementation of quantum computation. *Fortschr. Phys.*, 48:771–783, 2000.

[25] E. Knill, R. Laflamme, and W. H. Zurek. Resilient quantum computation: error models and thresholds. *Proc. R. Soc. Lond. A*, 454:365–384, 1998.

[26] M. Suchara, A. Cross, and J. M. Gambetta. Leakage suppression in the toric code. *Quantum Inf. Comput.*, 15:0997–1016, 2015.

[27] I. L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *J. Mod. Opt.*, 44:2455–2467, 1997.

[28] J. F. Poyatos, J. I. Cirac, and P. Zoller. Complete characterization of a quantum process: the two-bit quantum gate. *Phys. Rev. Lett.*, 78:390–393, 1997.

[29] E. Magesan, J. M. Gambetta, and J. Emerson. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.*, 106:180504, 2011.

[30] J. Wallman and S. Flammia. Randomized benchmarking with confidence. *New J. Phys.*, 16:079501, 2014.

[31] E. Magesan et al. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.*, 109:080505, 2012.

[32] D. C. McKay, T. Alexander, L. Bello, M. J. Biercuk, L. Bishop, J. Chen, A. D. Córcoles J. M. Chow, D. Egger, S. Filipp, Juan Gomez, M. Hush, A. Javadi-Abhari, D. Moreda, P. Nation, B Paulovicks, E. Winston, C. J. Wood, J. Wootton, and J. M. Gambetta. Qiskit backend specifications for openqasm and openpulse experiments. e-print arXiv:1809.03452v1, Sep 2018.

[33] J. Wallman, J. Emerson, and M. Barnhill. Characterization of leakage errors via randomized benchmarking. *New J. Phys.*, 18:043021, 2016.

[34] M. Levitt. Composite pulses. *Prog. NMR Spec.*, 18:61–122, 1985.

[35] S. Wimperis. Broadband, narrowband, and passband composite pulses for use in advanced NMR experiments. *J. Mag. Res.*, 109:221–231, 1994.

[36] M. Saffman, I. I. Beterov, A. Dalal, E. J. Páez, and B. C. Sanders. Symmetric Rydberg controlled-Z gates with adiabatic pulses. *Phys. Rev. A.*, 101:062309, 2019.

[37] D. Li et al. Coherent state transfer between superconducting qubits via stimulated Raman adiabatic passage. *Applied Phys. Lett.*, 118:104003, 2021.

[38] M. Garwood and L DelaBarre. The return of the frequency sweep: Designing adiabatic pulses for contemporary NMR. *J. Mag. Res.*, 153:155–177, 2001.

[39] Ibm quantum. https://quantum-computing.ibm.com/.

[40] M. J. Peterer, S. J. Bader, X. Jin, F. Yan, A. Kamal, T. J. Gudmundsen, P. J. Leek, T. P. Orlando, W. D. Oliver, , and S. Gustavsson. Coherence and decay of higher energy levels of a superconducting transmon qubit. *Phys. Rev. Lett.*, 114:010501, 2015.

[41] F. Motzoi, J. M. Gambetta, P. Rebentrost, , and F. K. Wilhelm. Simple pulses for elimination of leakage in weakly nonlinear qubits. *Phys. Rev. Lett.*, 103:110501, 2016.

[42] M. Weringhaus, D. J. Egger, F. Roy, S. Machnes, F. K. Wilhelm, and S. Filipp. Leakage reduction in fast superconducting qubit gates via optimal control. *npj Quantum Inf.*, 7(14), 2021.

[43] X. Rong, J. Geng, F. Shi, Y. Liu, K. Xu, W. Ma, F. Kong, Z. Jiang, Y. Wu, and J. Du. Experimental fault-tolerant universal quantum gates with solid-state spins under ambient conditions. *Nature Communications*, 6(8748), 2015.

[44] U. Vool and M. Devoret. Introduction to quantum electromagnetic circuits. *International Journal of Circuit Theory and Applications*, 45:897–934, 2017.

[45] A. Shnirman, G. Sch on, and Z. Hermon. Quantum manipulations of small josephson junctions. *Phys. Rev. Lett*, 79(12):2371–2374, 1997.

[46] R.M. Wilcox. Exponential operators and parameter differentiation in quantum physics. *J. Math. Phys*, 8:962–982, 1966.

[47] A. Brinkman. Introduction to average hamiltonian theory. i. basics. *Concepts Magn. Res. Part A*, 2017.