Efficient Characterization of a Unitary Quantum Gate via a Projective Rabi Experiment

by

Anthony Chytros

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics (Quantum Information)

Waterloo, Ontario, Canada, 2022

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see the included Statement of Contributions. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

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

Statement of Contributions

This thesis has been authored by Anthony Chytros. Proofs for appendices A.1 and A.2 were provided by Matthew A. Graydon. Chapter 4 and 5 is based on work co-authored by Joel Wallman and Anthony Chytros.

Abstract

In the era of NISQ devices, having practical methods to characterize noisy processes is more important than ever. Characterization techniques allow us to determine performance of current experimental setups and aids in the development of error mitigation and correction protocols. Randomization-based techniques have become a key player in this area with techniques such as randomized benchmarking and cycle benchmarking providing efficient and robust methods of determining fidelity metrics. In this thesis, we introduce a novel randomization-based protocol for estimation of parameters in a sparse Hamiltonian. This is achieved via a projective Rabi experiment which interleaves projective channels between applications of a fixed quantum channel of interest to allow for a coherent amplification of some small user selectable subspace. An important result that we prove is the efficiency and robustness of the protocol making it a suitable for experimental systems. Specifically, we prove it is Heisenberg-limited meaning the uncertainty in the output reaches the best case scaling with experimental time restricted only by the fundamental limit of quantum physics. For this thesis, we outline the abstract algorithm and demonstrate how to construct approximate projectors using the character projection formula. Then, we prove the efficiency and robustness of the protocol. Finally, we walk through an example of a multi-qudit rotation gate demonstrating the functionality of our methodology and run numerical simulations to validate our results.

Acknowledgements

I would like to thank Joel Wallman for his supervision and mentoring throughout my time at the University of Waterloo. Additionally, I would like to thank Crystal Senko and Rajibul Islam for granting me the opportunity to be a part of their QuantumIon team. Finally, I would like to thank Matthew Graydon and Joshua Skanes-Norman for their enthusiastic assistance whenever I had a question, the useful discussions and their guidance toward new resources.

Dedication

This is dedicated to all those who helped me get this far.

Table of Contents

List of Figures

Chapter 1

Introduction to Quantum Computing

It is no secret that computers have become an essential part of everyday life. From scientific applications such as modelling complex systems to even daily use such as browsing the internet; computers are unavoidable in the modern world. This explosion of technology is in part thanks to Moore's Law, which states the number of components on a microchip will double every two years all while the cost of those components will decrease at a similar rate [17]. However, despite holding true since its inception in the mid 60s, Moore's Law is less of a physical law and more a goal post. One that is rapidly become difficult to reach, namely due to the introduction of quantum effects as components become smaller. This sets the stage to explore the union between classical computing and quantum mechanics leading to the concept of quantum computers.

Theories about the advantages quantum computers appeared not long after Paul Benioff proposed the first quantum mechanical model of the Turing machine in 1980 [6]. Physicists such as Richard Feynman quickly recognized the potential these quantum machines could possess. Although, it was not until the introduction of algorithms such as Shor's algorithm that the first concrete evidence of this claim was provided. Peter Shor's factorization algorithm shows that a quantum computer can factorize integers exponentially faster than the fastest known classical algorithm. This revelation raised many red flags for cryptography experts as the widely used RSA scheme depends on factorization being difficult. Developing a quantum computer from then on has become the goal of many physicists, engineers and cryptographers. Companies such as Google and IBM are investing heavily in the modern day and many of the greatest minds across several disciplines are deeply involved. Yet, despite all this interest, we have yet to see a fully functioning quantum computer. Fundamentally, the issue that persists is that quantum systems are very sensitive to noise

and prone to error. Any unwanted interactions, such as with the surrounding environment or amongst qubits, introduce errors that severely limit performance. Quantum computers capable of operating as intended in the presence of these errors are called 'fault-tolerant'.

As we work toward the first fault-tolerant quantum computer, we currently reside in the 'noisy intermediate scale quantum' (NISQ) era as originally termed by John Preskill. To achieve fault tolerance, quantum error correcting codes (ECCs) must be developed. For QECCs to work as intended, errors in a system must occur at some rate below a threshold set by that QECC. Literature suggests lower bound error rates as low as 10^{-6} [1] for generic local noise and $10^{-4} - 10^{-3}$ [2, 3] for stochastic Pauli noise. Achieving (and surpassing) these rates is a communal effort between experimentalists developing better physical implementations and theorists understanding the noise processes acting on these systems.

In this thesis, we study characterization techniques that allow us to extract important information from our system. Some characterization protocols intend to learn everything about a particular aspect of our system, such as gate set tomography (GST) [7, 8, 28] which shows exactly how a gateset acts on target qubits. Others give a single metric that is meant to correspond to the performance of a system. In this area, randomization-based protocols have become dominant. Techniques such as randomized benchmarking [10, 12, 14, 15, 18, 19, 27] and cycle benchmarking [16] output an average gate-set fidelity providing information as to how well implemented gates are. These randomization-based protocols are efficient in terms of experimental overhead, scalable in the number of qubits, and robust to both state preperation and measurement (SPAM) and gate dependent errors.

Here we introduce a new randomization-based characterization technique for efficiently learning parameters of a sparse Hamiltonian. We achieve this by designing a way to perform a projective Rabi experiment. In between repeated applications of a noisy quantum channel, a projective channel[1] is interleaved. This projective channel is carefully constructed from the weighted average of results for sequences with randomly sampled generalized Pauli operators and isolates the action of a noisy channel on a small subspace. Repeated applications coherently amplify the signal such that it can be robustly characterized with post-processing time that scales with $\mathcal{O}(\log(D))$ where D is the dimension of the Hilbert space.

[1] In the presence of Markovian noise this becomes an 'effective' projective channel. That is, the channel is not an exact projective channel but under relatively small errors it still behaves close to one.

This protocol is not only efficient and robust to small errors but also achieves Heisenberg scaling. That is, the standard deviation in the output scales according to $\mathcal{O}(1/T)$ where T is the experimental probing time [32]. The Heisenberg limit is a fundamental limit closely related to the Heisenberg uncertainty principal and is typically a topic of interest in quantum metrology as it establishes the best case scenario for learning properties of a quantum system. Quantum phase estimation (QPE) is one example of a protocol that can achieve this limit [13, 23, 26]. Typically, for QPE, Heisenberg scaling is achieved through the use of entangling states such as NOON states [23]. However, some protocols have been demonstrated to achieve this limit without the use of such states by instead utilizing coherent accumulation to amplify some signal of interest [26]. This technique can be included as another of these Heisenberg-limited protocols.

This thesis is structured as follows. Chapter 2 introduces all the necessary mathematics to understand the results of this thesis. Chapter 3 provides much needed context of characterization techniques such as Rabi oscillation experiments and randomized benchmarking. Chapter 4 introduces the algorithm in its entirety including a method for creating a projective channel, examination of robustness and the effect of finite sampling, and an example of a multi-qudit rotation gate. Finally, chapter 5 runs numerical simulations to validate our results.

Chapter 2

Preliminaries

In this chapter we will review the preliminary material necessary to understand the rest of this thesis. The state of a quantum system can be described by a an element in a Hilbert space $\mathcal{H}$ restricted to a d-dimensional complex system $\mathbb{C}^d$. Qudits ($d \geq 2$) are the building blocks of a quantum computer and are mathematically represented as vectors in the Hilbert space spanned by the orthonormal computational basis $\{|i\rangle; i \in \mathbb{Z}_d\}$ where $\mathbb{Z}_d := \mathbb{Z}/d\mathbb{Z} = \{0, 1, \ldots, d-1\}$. For a multi-qudit system the Hilbert space will be an n-tensor fold product $\mathcal{H}^{\otimes n}$ where n is the number of qudits and so we define the total system dimension as $D = d^n$.

In quantum information theory, we often deal with the quantum operator formalism. Suppose we have a pure quantum state (for a single qudit)

$$|\psi\rangle = \sum_{i=0}^{d-1} \alpha_i |i\rangle \tag{2.1}$$

such that $\sum_{i=0}^{d-1} |\alpha_i|^2 = 1$. We can then define a quantum density operator as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{2.2}$$

where $\sum_{i=0}^{d-1} p_i = 1$. A quantum density operator is an operator $\rho \in \mathbb{C}^{D\times D}$ that is positive semi-definite to ensure non-negative probabilities when measuring the eigenstate of an observable. That is, for all $x \in \mathbb{C}^D$ it holds true that $x^\dagger \rho x \geq 0$. Additionally, it is Hermitian ($A^\dagger = A$) and holds the property $\mathrm{Tr}\rho = 1$. A quantum density matrix is

a representation of a quantum density operator by choice of basis in the Hilbert space, however, the two terms are often used interchangeably. Quantum channels are linear maps that act on our density operators that then can represent ideal or noisy quantum gates. These channels are completely positive trace-preserving (CPTP) maps that take density operators to density operators. Choi's theorem allows quantum channels to be expressed in Kraus form

$$\mathcal{C}(\rho) = \sum_j C_j \rho C_j^\dagger \tag{2.3}$$

for fixed (but not unique) operators $\{C_j\}$. The composition of two channels $\mathcal{C}$ and $\mathcal{D}$ with operators $\{C_j\}$ and $\{D_k\}$ respectively is

$$\mathcal{C}(\mathcal{D}(\rho)) = (\mathcal{C} \circ \mathcal{D})(\rho) = \sum_{j,k} C_j D_k \rho D_k^\dagger C_j^\dagger. \tag{2.4}$$

The composition of channels can be simplified by expressing quantum channels as matrices.

Let $\{B_1, \ldots, B_{D^2}\}$ be a basis for $\mathbb{C}^{D\times D}$ that is orthonormal according to the normalized Hilbert-Schmidt inner product

$$\langle A, B\rangle = \frac{1}{D}\mathrm{Tr}A^\dagger B, \quad \forall A, B \in \mathbb{C}^{D\times D}. \tag{2.5}$$

Let $|M\rangle\rangle$ be a column vector obtained by vectorizing an operator $M \in \mathbb{C}^{D\times D}$. It can be expanded as the vector of expansion coefficients

$$|M\rangle\rangle = \sum_{j=1}^{D^2} \langle B_j, M\rangle \, |B_j\rangle\rangle \tag{2.6}$$

By linearity we have

$$\langle A, B\rangle = \langle\langle A| \, |B\rangle\rangle \tag{2.7}$$

with the natural dual $|A\rangle\rangle^\dagger = \langle\langle A|$. The action of any linear map (e.g. quantum channel) $\mathcal{L} : \mathbb{C}^{D\times D} \to \mathbb{C}^{D\times D}$ can be written as

$$|\mathcal{L}(M)\rangle\rangle = \left(\sum_{j=1}^{D^2} |\mathcal{L}(B_j)\rangle\rangle \, \langle\langle B_j|\right) |M\rangle\rangle \tag{2.8}$$

where notation is slightly abused by conflating the abstract channel $\mathcal{L}$ with its matrix representation via

$$\mathcal{L} = \sum_{j=1}^{D^2} |\mathcal{L}(B_j)\rangle\rangle \langle\langle B_j| \,. \tag{2.9}$$

The action of a channel $\mathcal{L}$ given a state and observable $\rho, Q \in \mathbb{C}^{D\times D}$ is now

$$\mathcal{L}|\rho\rangle = |\mathcal{L}(\rho)\rangle\rangle \tag{2.10}$$

$$\mathrm{Tr}(Q^\dagger \mathcal{L}(\rho)) = \langle\langle Q| \mathcal{L} |\rho\rangle\rangle \,. \tag{2.11}$$

The composition of channels $\mathcal{L}_1$ and $\mathcal{L}_2$ as well as the tensor product is

$$|\mathcal{L}_1 \circ \mathcal{L}_2(\rho)\rangle\rangle = \mathcal{L}_1\mathcal{L}_2 |\rho\rangle\rangle \tag{2.12}$$

$$\left|\mathcal{L}_1 \otimes \mathcal{L}_2(\rho^{\otimes 2})\right\rangle\rangle = \mathcal{L}_1 \otimes \mathcal{L}_2 \left|\rho^{\otimes 2}\right\rangle\rangle \,. \tag{2.13}$$

As an example, say we have some state $\rho \in \mathbb{C}^{D\times D}$ and want to apply a channel $\mathcal{G}$ which is a series of gates labeled G_1, G_2, G_3 with corresponding channels $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3$ in respective order. The Choi representation of that circuit would be

$$\mathcal{G}(\rho) = \sum_{i,j,k} G_{3,k}G_{2,j}G_{1,i}\rho G_{1,i}^\dagger G_{2,j}^\dagger G_{3,k}^\dagger \tag{2.14}$$

where $G_{i,j}$ is the ith gate and jth Kraus operator. The matrix representation on the other hand is

$$|\mathcal{G}(\rho)\rangle\rangle = \mathcal{G}_3\mathcal{G}_2\mathcal{G}_1 |\rho\rangle\rangle \tag{2.15}$$

thus greatly simplifying the representation of the circuit.

2.1 Group Theory

Groups play a fundamental role in quantum information theory as well as many other disciplines. This is because symmetric structures and groups are very closely related to one another. Many symmetric structures are encoded as groups such as Lie groups in geometry and the Poincaré group in special relativity. In relation to symmetries P.W. Anderson stated that "it is only slightly overstating the case to say that physics is the study of symmetry" [4]. Symmetry is perhaps the most powerful tool in all of physics and as such it

will be important to understand at least some group theory for the remainder of this thesis.

A group $(G, \cdot)$ is defined as a set G with a binary operator $\cdot$ acting on elements of G such that the following properties are satisfied:

1. Closure:
   - Given $a, b \in \mathbb{G}$ then $a \cdot b \in \mathbb{G}$
2. Associativity:
   - For all $a, b, c \in G$, we have $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
3. Identity:
   - There is an element $e \in G$ such that for every $a \in G$ we have $e \cdot a = a \cdot e = a$
4. Inverse:
   - For every $a \in G$ there exists $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$

Any set and binary operation that fulfills these properties can be considered a group. The following groups are relevant to this thesis.

The general linear group $GL_d(F)$ of dimension d and field F is the group of $d \times d$ invertible matrices with elements in F. An equivalent definition is $GL(V)$, the general linear group over a finite dimensional ($\mathbb{R}$ or $\mathbb{C}$) vector space V. This is the group of all automorphisms of V (i.e. the group of invertible linear transformations of V). The special linear group $SL(V)$ is a subgroup of the general linear group with the property that all matrices have determinant 1.

The unitary group $U(d)$ is the group of isometries of the d dimensional complex Hilbert space $\mathcal{H} = \mathbb{C}^d$. Naturally, $U(d) \subset GL_d(\mathbb{C})$ as the set of $d \times d$ unitary matrices is a subset of the $d \times d$ invertible matrices. The special unitary group $SU(d)$, similarly to the special linear group, is a subgroup of the unitary group of matrices with determinant 1.

The general Pauli group $\mathcal{P}_d$ (also known as the Weyl-Heisenberg group for $\mathcal{P}_d^n$) is the subgroup of $U(d)$ that is generated by $\mathcal{P}_d := \langle \tilde{\omega}\mathbb{I}_d; X_d; Z_d \rangle$ where

$$X_d = \sum_{j=0}^{d-1} |j \oplus 1\rangle\langle j|, \tag{2.16}$$
$$Z_d = \sum_{j=0}^{d-1} \omega^j |j\rangle\langle j|; \quad \omega = \exp(2\pi i/d)$$

are the shift and clock operators respectively and $\tilde{\omega} = \omega$ for odd d and $\tilde{\omega} = \omega^{1/2}$ for even d. X_d and Z_d have order d indicating $X^d = Z^d = \mathbb{I}_d$. Neglecting the phase $\tilde{\omega}$ produces a projective group where any two elements is a phase multiple of another element.

The Clifford group acts as the normalizer of the Pauli group and is defined as

$$\mathcal{C}_d^n := \{c \in U(d^n); c\tilde{\mathcal{P}}_d^n c^\dagger \in \mathcal{P}_d^n\}/\{\exp(i\theta)\mathbb{I}_d^n; \theta \in \mathbb{R}\} \tag{2.17}$$

where $\tilde{\mathcal{P}}_d^n := \mathcal{P}_d^n/\langle\tilde{\omega}\mathbb{I}_d\rangle$. The qudit Clifford group is generated by $\mathcal{C}_d^n = \langle CZ_d, F_d, P_d, Z_d\rangle$. Where the controlled-Z gate is:

$$CZ_d = \sum_{j=0}^{d-1}\sum_{j'=0}^{d-1} \omega^{jj'} |jj'\rangle\langle jj'| \tag{2.18}$$

F corresponding to the quantum Fourier transform:

$$F_d = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1}\sum_{j'=0}^{d-1} \omega^{jj'} |j'\rangle\langle j| \tag{2.19}$$

P the phase gate:

$$P_d = \sum_{j=0}^{d-1} \omega^{\frac{j(j+\varrho_d)}{2}} |j\rangle\langle j|, \quad \varrho_d = \begin{cases} 1, & \text{if d is odd,} \\ 0, & \text{otherwise.} \end{cases} \tag{2.20}$$

2.2 Representation Theory

Groups are abstract algebraic structures. In order to study them in the context of quantum information it is useful to work in the framework of matrices and linear operators. This

can be achieved through representation theory.

Given the general linear group $GL(V)$ over a finite-dimensional (real or complex) vector space V and group G, a representation φ_V of G on vector space V is a map:

$$\varphi_V : G \to GL(V) : g \mapsto \varphi(g) \tag{2.21}$$

with the property

$$\varphi_V(g)\varphi_V(h) = \varphi_V(gh) \quad \forall g, h \in G. \tag{2.22}$$

If a subspace W satisfies that for all $w \in W$ it holds true

$$\varphi_V(g)W \subset W \quad \forall g \in G \tag{2.23}$$

then W carries a subrepresentation of φ_V denoted as φ_W. If there is no non-trivial proper subspace fulfilling eq.(2.23) then, the representation φ_V is said to be irreducible. Two representations φ_V and $\varphi_{V'}$ are said to be equivalent ($\varphi \cong \varphi^{'}$) if and only if there exists an invertible linear map $T : V^{'} \to V$ such that

$$\varphi_V(g) = T\varphi_{V'}(g)T^{-1} \quad \forall g \in G. \tag{2.24}$$

Maschke's lemma allows every finite-dimensional representation of a group to be written as a direct sum of its irreducible representations

$$\varphi_V(g) \simeq \oplus_{\lambda \in R_G} \varphi_\lambda(g)^{\otimes m_\lambda} \quad \forall g \in G \tag{2.25}$$

where m_λ is the multiplicity integer of φ_λ or, in other words, the number of equivalent copies of φ_λ in φ. Representations are called multiplicity-free when $m_\lambda = 1$ for all φ_λ. One fundamental result of representation theory is Schur's lemma which pertains to irreducible representations. Let φ_V,$\varphi_{V'}$ be irreducible representations of a finite group G on vector spaces V,$V^{'}$ respectively. A linear map $A : V \to V^{'}$ satisfies

$$\varphi_{V'}(g)A = A\varphi_V(g) \quad \forall g \in G \tag{2.26}$$

if and only if

$$A = \begin{cases} 0 & \text{if } \varphi_V \ncong \varphi_{V'}, \\ \lambda \mathbb{I} & \text{if } \varphi_V \cong \varphi_{V'} \end{cases} \tag{2.27}$$

where $\lambda \in \mathbb{C}$. This means that there are no non-zero homomorphisms between distinct irreducible representations and that the only non-trivial endomorphisms of an irreducible representation are multiples of the identity. Setting $V^{'} = V$ and working with multiplicity free representation $\varphi_V(g) = \oplus_{\lambda \in R_G} \varphi_\lambda(g)$ then

$$A = \sum_i \lambda_i P_i \tag{2.28}$$

where P_i is a projector which is defined as a linear operator $P : V \to V$ that satisfies $P^2 = P$. Then we can see A is a sum over projectors onto subspaces of V carrying irreducible subrepresentations φ_λ. Given a representation ϕ on vector space V and a general linear map $A : V \to V$, the twirl of A with respect to ϕ, $\mathcal{T}_\phi$ can be defined as

$$\mathcal{T}_\phi(A) := \mathbb{E}_{G \in \mathbb{G}} \phi(G) A \phi(G)^\dagger \tag{2.29}$$

and further simplified for multiplicity free representations as

$$\mathcal{T}_\phi(A) = \frac{\mathrm{Tr}(A P_\lambda)}{\mathrm{Tr}(P_\lambda)} P_\lambda. \tag{2.30}$$

where P_λ is a projector onto the support of the representation φ_λ.

2.3 Character of a Representation

Let $\varphi : G \to V$ be a representation of a group G carried by (real or complex) vector space V. We define the character χ_φ as

$$\chi_\varphi : G \to \mathbb{C} : g \mapsto \chi_\varphi(g) = \mathrm{Tr}_V(\varphi(g)) \tag{2.31}$$

noting that $\mathrm{Tr}_V()$ is the trace over vector space V and the character is complex for complex vector spaces or real for real vector spaces. Given two representations φ, φ' we have the relations

$$\chi_{\phi \otimes \phi'} = \chi_\phi \chi_{\phi'} \tag{2.32}$$

$$\chi_{\phi \oplus \phi'} = \chi_\phi + \chi_{\phi'} \tag{2.33}$$

The inner product between the characters of two representations for a compact group is

$$\langle \chi_\phi, \chi_{\phi'} \rangle = \int_{\mathbb{G}} \chi_\phi(G)\overline{\chi_{\phi'}(G)}\mu(G). \tag{2.34}$$

The overline in the previous two expressions indicates complex conjugation and μ the normalized Haar measure. With this in mind, Schur's orthogonality relation states that for irreducible representations

$$\langle \chi_\phi, \chi_{\phi'} \rangle = \begin{cases} 0 \text{ if } \phi \ncong \phi', \\ 1 \text{ if } \phi \cong \phi'. \end{cases} \tag{2.35}$$

Lemma 1 (Character projection formula)**.** *Let $\mathbb{G}$ be a group with representation ϕ and normalized Haar measure μ. Let $\hat{\phi}$ be an irreducible subrepresentation of ϕ with character $\chi_{\hat{\phi}}$. Then we have the following*

$$|\hat{\phi}| \int_{\mathbb{G}} \chi_{\hat{\phi}}(G)\phi(G)\mu(G) = P_{\hat{\phi}} \tag{2.36}$$

where $P_{\hat{\phi}}$ is a projector onto all subrepresentations of ϕ that are equivalent to $\hat{\phi}$.

2.4 General Pauli Transfer Matrix Representation

Previously in this chapter we established the matrix representation of quantum channels given a basis $\{B_1, \ldots, B_{D^2}\}$. When selecting the basis as the projective Weyl-Heisenberg group (established in 2.1)

$$\mathbb{W}_{d,n} = \{X_d^x Z_d^z : x, z \in \mathbb{Z}_d^n\} \tag{2.37}$$

we obtain the general Pauli transfer matrix (PTM) representation. Given a quantum channel $\mathcal{E}$ its corresponding PTM representation will be a $D \times D$ matrix R with elements

$$R_{i,j} = \frac{1}{D}\mathrm{Tr}(W_i^\dagger \mathcal{E}(W_j)) \tag{2.38}$$

where $W_i \in \mathbb{W}_{d,n}$ are in a fixed but arbitrary order. A useful property of the PTM representation is that it is a true representation of any subgroup of $U(D)$. Given $U, V \in U(D)$ set $W = UV$. The PTM representation then shows

$$\mathcal{U}\mathcal{V}\left|X\right\rangle\!\right\rangle = \left|VUXU^\dagger V^\dagger\right\rangle\!\right\rangle = \left|WXW^\dagger\right\rangle\!\right\rangle = \mathcal{W}\left|X\right\rangle\!\right\rangle, \quad \forall X \in \mathbb{C}^{D\times D} \tag{2.39}$$

satisfying the definition of a representation in eq.(2.22). Figure 2.1 is an example of the PTM for a Pauli X gate and rotation gate generated by Pauli X and angle $\frac{\pi}{4}$ defined as:

$$R_G(\theta) = \exp(-i\theta G). \tag{2.40}$$

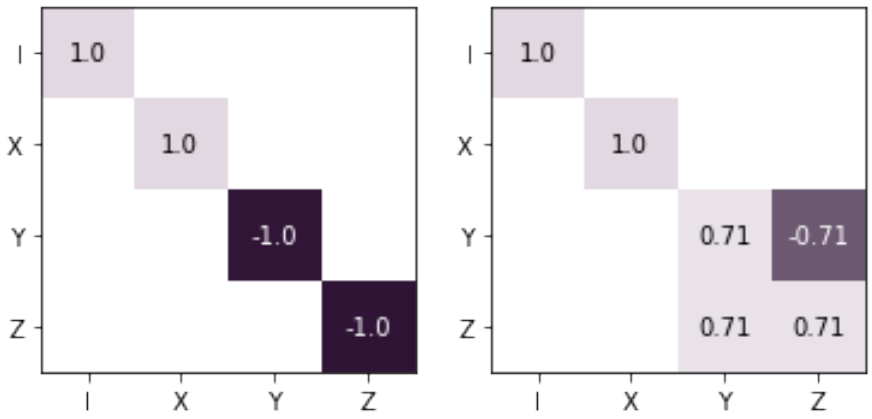

Figure 2.1: For $d = 2$, the PTM of a Pauli X gate (left) and rotation gate $R_X(\pi/4)$ (right) visualized. Blocks represent the actions of the gate on elements of the Weyl-Heisenberg group. E.g. the rotation gate generated by X maps Y to $0.71Y$ and $-0.71Z$.

Chapter 3

Review of Characterization Techniques

In this chapter we will review some notable characterization techniques. These techniques operate on different levels of abstraction in a quantum computer. We will focus on two of those levels for this chapter: the Hamiltonian level and the gate level. At the Hamiltonian level, the evolution of an isolated quantum system is described by the Schrödinger equation

$$\psi'(t) = -iH(t)\psi(t) \tag{3.1}$$

where $H(t) \in \mathbb{C}^{D\times D}$ is a time-dependent operator called the Hamiltonian that describes the energy of a system. Qudits are coupled with control Hamiltonians that are driven to manipulate the state of the system. At the gate level, time dependence is abstracted away by integrating time-dependent evolution over distinct intervals giving a unitary matrix which we call a gate. We primarily work within the gate level for this thesis, however, certain characterization/calibration techniques operate within the Hamiltonian level and so it is useful to introduce these levels of abstraction.

3.1 Hamiltonian Level

Here we review a notable Hamiltonian level characterization technique, Rabi oscillation experiments. A Rabi oscillation (sometimes known as a Rabi cycle or Rabi flop) refers to the cyclical behaviour of a few level quantum system (typically a two level system) in the presence of a periodically time-varying field. We can outline some of the foundational

physics behind these oscillations as follows. Let us consider a two level system with ground state $|g\rangle$ and excited state $|e\rangle$ with energies E_g and E_e respectively. The Hamiltonian of a classical field with time-varying amplitude is

$$H_R = E_g|g\rangle\langle g| + E_e|e\rangle\langle e| + (F_0 V_{eg} e^{-i\omega t}|e\rangle\langle g| + F_0^* V_{eg}^* e^{i\omega t}|g\rangle\langle e|) \tag{3.2}$$

where the amplitude of the classical field is denoted as F_0, V_{eg} is the matrix element for the interaction of the system with the field, and ω is the frequency the amplitude of the field varies with. If we have some initial state at $t = 0$ then its time evolution will give a solution of the form

$$|\psi(t)\rangle = e^{-iE_g t} C_g(t)|g\rangle + e^{-iE_e t} C_e(t)|e\rangle. \tag{3.3}$$

Setting the initial state to the ground state $|\psi(0)\rangle = |g\rangle$ we can solve for the coefficients $C_g(t), C_e(t)$ using Schrödinger equation to obtain

$$|\psi(t)\rangle = e^{i(\omega - E_e - E_g)t/2} \left[\left(\cos(\Omega t) + \frac{i\Theta}{2\Omega} \sin(\Omega t) \right) |g\rangle - i \frac{F_0 V_{eg}}{\Omega} e^{-i\omega t} \sin(\Omega t)|e\rangle \right] \tag{3.4}$$

where $\Theta = E_e - E_g - \omega$ and $\Omega = \sqrt{\Theta^2/4 + |F_0 V_{eg}|^2}$. Finally, measurement of the system being in the excited state gives the probability

$$P_e(t) = \left| \frac{F_0 V_{eg}}{\Omega} \right|^2 \sin^2(\Omega t) \tag{3.5}$$

Therefore the population of the excited state oscillates with a period of π/Ω. [29]

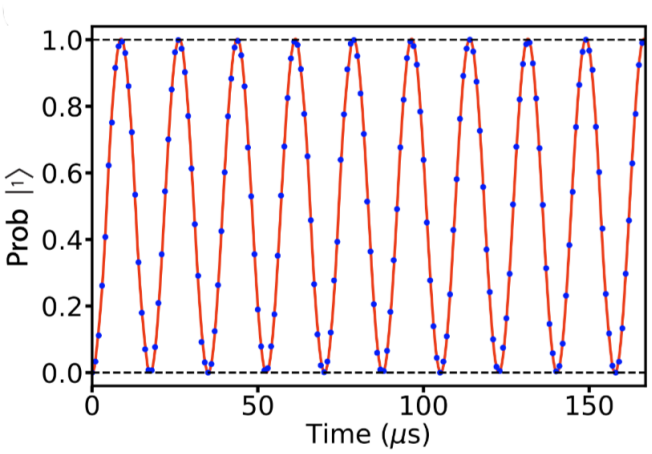

Figure 3.1: Example of Rabi oscillations in a $^{133}Ba^+$ trapped ion system obtained by applying a microwave rotation $R(\Omega_R t, 0)$ onto an ion in the $|0\rangle$ state. [11]

Rabi oscillation experiment is a ubiquitous term used to generally describe experiments that induce this oscillatory behaviour in a quantum system. The experiment itself can vary depending on context however all have the common characteristic of inducing Rabi oscillations in a quantum system. A Rabi oscillation experiment in the context of quantum information processing can be presented as follows:

1. Prepare a fixed state; for example the $|1\rangle$ state
2. Allow the state to evolve freely for time t under the Hamiltonian $H(t)$
3. Measure and obtain state $|i\rangle$
4. Repeat steps 1-3 k times
5. Report the probability as $P_i(t) = \frac{\text{Number of i's}}{k}$

This can then be repeated over a range of times and the output $P_i(t)$ can be fitted to a sinusoidal equation such as eq.(3.5) to obtain a figure similar to Figure 3.1. This procedure can be used, for example, to comparatively determine the performance of qubits based on how close experimental results are to the theoretical Rabi oscillations where the best performing qubits will be approximately close to the ideal.

3.2 Gate Level

Next, we turn our attention to a set of protocols at the gate level often referred to as benchmarking. Involving protocols capable of quantifying the performance of a quantum computer; benchmarking is one of the most useful tools we have towards developing a fault tolerant quantum computer as it provides us with the concept of fidelity.

Fidelity is the measure of 'closeness' between quantum operators. It is perhaps one of the most useful metrics that accompany benchmarking and we will examine protocols that provide us such a metric in this section.

In general, benchmarking is divided into two phases: data collection and data processing. The specifics of what each phase entails depends on the protocol but can generally be described as follows: [21]

- Data collection: Involves physically running an algorithm on a quantum computer. Steps are typically (1) state preparation, (2) application of a sequence of gates and (3) measurement.
- Data processing: Involves taking measurement data from the data collection phase and processing it to establish some information about the physical system.

When developing a benchmarking protocol there are two main considerations to be kept in mind. First is efficiency. As the size of our physical systems grow (i.e., as the number of qudits increases) the protocols we employ on these systems should run with reasonable overhead. That is, the resources required to run these protocols with some fixed precision should scale polynomially with the size of the system. This is a non-trivial problem as the complexity of a Hilbert space increases exponentially with the number of qudits as well as their dimensionality. Secondly is robustness to error. Benchmarking tools are intended to be run on devices prone to error. As such, these algorithms need to be designed to cope with errors that may occur experimentally such as gate or state preparation and measurement (SPAM) errors. If an algorithm can function with these errors then it is said to be robust. With these considerations in mind let us continue to some notable benchmarking techniques.

3.2.1 Standard Randomized Benchmarking

Perhaps the most prominent benchmarking technique is randomized benchmarking. In fact, all the benchmarking techniques that will be covered in this chapter can be considered as versions of randomized benchmarking. For this section however we will be covering what is referred to as standard randomized benchmarking.

The standard randomized benchmarking procedure is as follows. Given a finite group $\mathbb{G}$ and fixed positive integer m, run the following procedure: [12, 14, 22]

1. Choose a state ρ and POVM $\{Q, \mathbb{I} - Q\}$
2. Sample $\vec{G} = G_1, \ldots, G_m$ uniformly at random from $\mathbb{G}$
3. Prepare ρ and apply the gates $G_1, G_2, \ldots G_m$
4. Compute and apply inverse $G_{inv} = (G_m \ldots G_1)^\dagger$

5. Obtain the measurement $p_m(\vec{G}) := \langle\langle Q| \tilde{G}_{inv}\tilde{G}_m \dots \tilde{G}_1 |\rho\rangle\rangle$

6. Repeat steps 3-5 for many $\vec{G}$ and estimate the average $p_m := \mathbb{E}_{\vec{G}}(p_m(\vec{G}))$

The procedure is then repeated over different values of $m \in \mathbb{M}$ and a set of probability outputs $\{p_m\}_{m\in\mathbb{M}}$ are obtained. Intuitively, we can think of standard randomized benchmarking as applying a random sequence of gates such that (in the case of ideal noiseless gates) ρ is unchanged. Encountering noisy gates will result in deviations from the ideal measurement results in step 5. Thus, averaging these measurements over random sequences and fitting the results as a function of m can be used to identify the average fidelity of the gates being applied.

More formally, under the assumption of gate-independent CPTP noise and multiplicity-free representations, each output survival probability p_m can be shown to be

$$\begin{aligned} p_m &= \mathbb{E}_{G_1,\dots G_m} \langle\langle Q| \tilde{G}_{inv}\tilde{G}_m \dots \tilde{G}_1 |\rho\rangle\rangle && (3.6)\\ &= \langle\langle Q| \left(\mathbb{E}_{G\in\mathbb{G}}(\mathcal{G}^\dagger\mathcal{E}\mathcal{G})\right)^m |\rho\rangle\rangle && (3.7)\\ &= \sum_{\lambda\in R_\mathbb{G}} \langle\langle Q| \mathcal{P}_\lambda |\rho\rangle\rangle f_\lambda^m && (3.8) \end{aligned}$$

where $\tilde{G} = \mathcal{E}\mathcal{G}$ is a noisy implementation of $G \in \mathbb{G}$, $\mathcal{P}_\lambda$ is an orthogonal projector onto an irreducible representation of $\mathcal{G}$, $R_\mathbb{G}$ is an index set and $f_\lambda := \mathrm{Tr}(\mathcal{P}_\lambda\mathcal{E})/\mathrm{Tr}(\mathcal{P}_\lambda)$. Therefore, the averaged survival probability of the randomized circuit is equivalent not only to a $\mathbb{G}$-twirl of the error channel (eq (3.2)) but also a linear combination of quality parameters via Schur's lemma (eq 3.3). These parameters will, in general, be able to provide the average fidelity of the gate-set F_{avg}. In the data processing phase, fitting occurs such that

$$p_m \approx_{fit} \sum_{\lambda\in R_G} A_\lambda f_\lambda^m. \tag{3.9}$$

Then the average gate-set fidelity can be determined via

$$F_{avg} = \frac{2^{-q}\sum_{\lambda\in R_G}\mathrm{Tr}(\mathcal{P}_\lambda)f_\lambda}{2^q+1} \tag{3.10}$$

Eq.(3.9) requires the fitting of (potentially) many parameters which is a difficult task. However, selecting the finite group $\mathbb{G}$ as the Clifford group simplifies eq.(3.9) to

$$p_m \approx_{fit} A + Bf^m. \tag{3.11}$$

making Clifford based standard RB efficient and scalable. Another useful property is robustness to SPAM error. The parameters f_λ are solely dependent on gate implementations and SPAM dependency is concentrated into the A_λ terms.

3.2.2 Character Randomized Benchmarking

As mentioned in the standard randomized benchmarking section, non-Clifford RB can potentially require the fitting of many parameters. Thus, modifications would have to be introduced to circumvent this issue for non Clifford gatesets. This is where character randomized benchmarking comes in. By introducing concepts found in character theory (see section 2.3) a single parameter f_λ can be isolated regardless of gateset.

The character randomized benchmarking procedure is as follows. Given a finite group $\mathbb{G}$, a specific irreducible subrepresentation of that group $\phi_{\lambda'}$, a subgroup $\hat{\mathbb{G}} \subset \mathbb{G}$ and a irreducible subrepresentation of that group $\hat{\phi}$ run the following procedure for fixed integer m: [22]

1. Choose a state ρ and POVM $\{Q, \mathbb{I} - Q\}$
2. Sample $\vec{G} = G_1, \ldots, G_m$ uniformly at random from $\mathbb{G}$
3. Sample $\hat{G}$ uniformly at random from $\hat{\mathbb{G}}$
4. Prepare the state ρ and apply the gates $(G_1\hat{G}), G_2, \ldots G_m$
5. Compute the inverse $G_{inv} = (G_m \ldots G_1)^\dagger$ and apply
6. Measure $k_m^{\lambda'}(\vec{G}, \hat{G}) = |\hat{\phi}|\chi_{\hat{\phi}}(\hat{G}) \langle\langle Q| \tilde{\mathcal{G}}_{inv}\tilde{\mathcal{G}}_m \ldots \widetilde{\mathcal{G}_1\hat{\mathcal{G}}} |\rho\rangle\rangle$
7. Repeat for $\hat{G}$ and obtain the average $k_m^{\lambda'}(\vec{G}) = \mathbb{E}_{\hat{G}}(k_m^{\lambda'}(\vec{G}, \hat{G}))$
8. Repeat for $\vec{G}$ and obtain the average $k_m^{\lambda'} = \mathbb{E}_{\vec{G}}(k_m^{\lambda'}(\vec{G}))$

Then repeat for different values of $m \in \mathbb{M}$. The goal behind the procedure is to compile an additional gate $\hat{G}$ with its character (focused on subrepresentation $\hat{\phi}$) to construct the

character projection formula (eq.(2.36)). This projector will isolate a single component of eq.(3.9)

$$k_m^{\lambda'} = \sum_{\lambda \in R_{\mathbb{G}}} \langle\langle Q| \mathcal{P}_\lambda \mathcal{P}_{\hat{\phi}} |\rho\rangle\rangle f_\lambda^m \tag{3.12}$$

$$= \langle\langle Q| \mathcal{P}_{\hat{\phi}} |\rho\rangle\rangle f_{\lambda'}^m \tag{3.13}$$

$$\approx_{fit} A_{\lambda'} f_{\lambda'}^m \tag{3.14}$$

allowing for efficient characterization of non-Clifford gatesets. Repeat runs of the procedure on various subrepresentations allows for efficient learning of all parameters f_λ and gives full characterization of the gateset $\mathbb{G}$.

3.2.3 Cycle Benchmarking

Cycle benchmarking arises from further practical limitations that randomized benchmarking faces. Randomized benchmarking does a good job characterizing stochastic gate errors without conflating SPAM error however it fails to account for coherent errors such as crosstalk errors. Crosstalk can describe a large set of device-specific physical phenomena that occurs in many qubit systems. The term encompasses the phenomena of driving unintentional qubit interactions. For example, in a multi-qubit system applying a gate on your first qubit may interact with the second or third. Crosstalk error is the by-product of crosstalk that shows up in our computations typically violating either locality or independence (or both). [30] For single qubits, these coherent errors show up as unwanted rotations $U(\hat{n}, \epsilon) = exp(-i\epsilon\hat{n} \cdot \sigma/2)$ with $\hat{n}$ the axis of rotation and σ the Pauli vector [20]. Capturing these errors are crucial to the development of a fault tolerant quantum computer and cycle benchmarking aims to do just that. This is achieved by using the notion of cycles; a set of operations acting on a quantum register within a set period of time.

The cycle benchmarking procedure is outlined below and illustrated in Figure 3.2. [16]

1. Select K random Paulis P
2. Select two lengths m_1 and m_2 such that $\mathcal{G}^{m_1} = \mathcal{G}^{m_2} = \mathcal{I}$
3. Perform the following for each Pauli $P \in \mathbb{P}$, length $m \in (m_1, m_2)$ and $l \in (1, \ldots, L)$

(a) Select $m+1$ random Pauli cycles $\mathcal{R}_0, \mathcal{R}_1, \ldots, \mathcal{R}_m$ and define the randomized circuit

$$\mathcal{C}(P) = \mathcal{R}_m \mathcal{G} \mathcal{R}_{m-1} \mathcal{G} \ldots \mathcal{R}_1 \mathcal{G} \mathcal{R}_0 \tag{3.15}$$

(b) Calculate the expected outcome of $\mathcal{C}(P)$

(c) Apply $\mathcal{C}(P)$ to a state ρ and estimate the overlap

$$f_{P,m,l} = \mathrm{Tr}[\mathcal{C}(P)\widetilde{\mathcal{C}(\rho)}] \tag{3.16}$$

where the state ρ is a $+1$-eigenstate of P and operators $\tilde{\mathcal{B}}_P$ and $\tilde{\mathcal{B}}^\dagger_{\mathcal{C}(P)}$ are used to perform state preparation and measurement.

4. Estimate the composite process fidelity

$$F_{RC}(\tilde{\mathcal{G}}, \mathcal{G}) = \sum_{P \in \mathbb{P}} \frac{1}{|\mathbb{P}|} \left(\frac{\sum_{l=1}^{L} f_{P,m_2,l}}{\sum_{l=1}^{L} f_{P,m_1,l}} \right)^{\frac{1}{m_2 - m_1}} \tag{3.17}$$

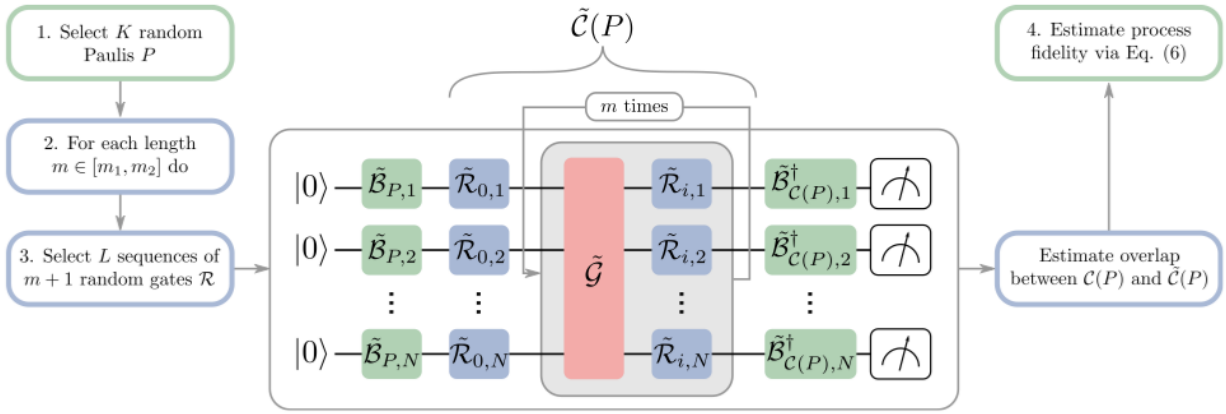

Figure 3.2: The cycle benchmarking procedure as introduced in [16]. The green $\tilde{\mathcal{B}}_{P,N}$ are basis changing operations, blue $\tilde{\mathcal{R}}_{i,N}$ are Pauli cycles, and red $\tilde{\mathcal{G}}$ is a noisy Clifford gate of interest.

The random Pauli cycles present in the procedure have the explicit purpose of transforming the coherent errors present in the circuit into stochastic Pauli noise processes via twirling. Twirling is the technique of conjugating a gate with randomly chosen gates from

a twirling set to tailor generic local noise into some other noise model [9]. Selecting the full set of Pauli gates as the twirling set will convert a general noise channel to a Pauli noise channel modeled by

$$\mathcal{E}(\rho) = \sum_{P \in \mathbb{P}^{\otimes n}} c_P P \rho P^\dagger \tag{3.18}$$

where c_P is the relative probability of Pauli error P occurring. One of the benefits of this noise tailoring is that stochastic Pauli errors only grow linearly with circuit depth as opposed to coherent errors worst case scenario of quadratic growth. Additionally, when comparing fault tolerant error rate thresholds for stochastic versus generic local noise one finds that stochastic noise are orders of magnitude higher than generic local noise.[20] In cycle benchmarking, these Pauli cycles will allow for the decay $\sum_{l=1}^{L} f_{P,m,l}/L$ to be extracted.

Process fidelity can be determined via

$$F(\tilde{\mathcal{G}}, \mathcal{G}) = \sum_{P \in \mathbb{P}^N} 4^{-N} F_P(\tilde{\mathcal{G}}, \mathcal{G}), \tag{3.19}$$

where

$$F_P(\tilde{\mathcal{G}}, \mathcal{G}) = 2^{-N} \mathrm{Tr}\left[\mathcal{G}(P)\tilde{\mathcal{G}}(P)\right]. \tag{3.20}$$

However, this method of process fidelity estimation is not robust to SPAM errors. Instead, extracting the decay of $F_P(\tilde{\mathcal{G}}^m, \mathcal{G}^m)$ will decouple SPAM error similar to randomized benchmarking. For generic noise channels this extraction is a non ideal task and so, as previously mentioned, Pauli cycles can be introduced to allow for the decay to be efficiently extracted giving us

$$F_{RC}(\tilde{\mathcal{G}}, \mathcal{G}) = \sum_{\mathcal{R} \in \mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}^{\otimes N}} 4^{-N} F(\tilde{\mathcal{G}}\tilde{\mathcal{R}}, \mathcal{G}\mathcal{R}). \tag{3.21}$$

In the procedure itself, composite fidelity is estimated by eq. (3.17). It can be shown that this fidelity is equivalent to the one in eq. (3.21) up to $\mathcal{O}([1 - F_{RC}(\tilde{\mathcal{G}}, \mathcal{G})]^2)$.

Chapter 4

Algorithm

In this chapter we outline our algorithm which we refer to as a 'projective Rabi experiment' due to its behaviour functioning similarly to the Rabi experiment outlined in Chapter 3 with the main difference being the interleaving of projective channels. Our protocol is inspired by cycle benchmarking [16] for its repeat applications of randomly sampled Pauli cycles and character RB [22] for its use of character projections; we aim to combine the two in order to coherently amplify the error in a subspace of our choosing. Say we have a fixed unitary quantum channel $\mathcal{A} \in \mathbb{U}(D)$ that we wish to characterize and a fixed positive integer m representing experimental time or, equivalently, the number of iterations of a quantum channel. Let $\mu, \nu : \mathbb{W}_{d,n} \to \mathbb{R}, \mathbb{C}$ be distributions where μ is a probability distribution we sample gates from and ν is some scalar function which we will typically choose to be a character of the Weyl group (introduced as the Weyl-Heisenberg group in 2.1). The projective Rabi experiment then is presented as follows:

1. Prepare ρ;
2. Apply a random gate U_0 sampled according to the measure μ;
3. Set $\chi = \nu(U_0)$
4. For each $j = 1, \ldots, m$:
   (a) Apply $\mathcal{A}$;
   (b) Apply a random gate U_j sampled according to the measure μ;
   (c) Set $\chi \to \chi\nu(U_j)$

5. Measure the expectation value μ of Q.

6. Return $\mu\chi$.

The output for the algorithm for a fixed choice of sampled Weyl operators $\{U_0, \ldots, U_m\}$ is then

$$\mu\chi = \langle\langle Q| \left(\prod_{j=m\to 1} \nu(U_j)\Theta(U_j)\Theta(A) \right) \nu(U_0)\Theta(U_0) |\rho\rangle\rangle \,. \tag{4.1}$$

where $\Theta : \mathbb{U}(D) \mapsto \mathbb{C}^{D^2\times D^2}$ is the implementation map of quantum channels and the directional notation shorthand is used

$$\prod_{j=a\to b} x_j = \begin{cases} x_a x_{a+1} \ldots x_{b-1} x_b, & \text{if } a \leq b \\ x_a x_{a-1} \ldots x_{b+1} x_b, & \text{if } a > b \end{cases} \tag{4.2}$$

to specify the order for products of non-commuting variables. As a simple example, say we are working with a single qubit system ($D = 2$) and we sample the Weyl operators X and Z in that order. Assuming implementation maps are ideal (i.e. $\Theta(U) = \mathcal{U}$) our single output would then be

$$\mu\chi = \nu(Z)\nu(X) \langle\langle Q| \mathcal{Z}\mathcal{A}\mathcal{X} |\rho\rangle\rangle \tag{4.3}$$

where we exploited the fact that $\nu(U)$ is a scalar bringing those terms to the front of the equation[1]. Returning to the general case, averaging the output over many samples of U_j gives

$$\langle\mu\chi\rangle = \mathbb{E}_{U_m \ldots U_0} \left(\langle\langle Q| \left(\prod_{j=m\to 1} \nu(U_j)\Theta(U_j)\Theta(A) \right) \nu(U_0)\Theta(U_0) |\rho\rangle\rangle \right) \tag{4.4}$$

$$= \int_{\mathbb{U}(D)} \left(\prod_{j=m\to 0} \mu(U_j) \right) \langle\langle Q| \left(\prod_{j=m\to 1} \nu(U_j)\Theta(U_j)\Theta(A) \right) \nu(U_0)\Theta(U_0) |\rho\rangle\rangle \tag{4.5}$$

Eq.(4.4) looks quite complex, however, we can simplify further by defining the quantum channel

$$\mathcal{B} = \int_{\mathbb{U}(D)} \mu(U)\nu(U)\Theta(U). \tag{4.6}$$

[1]There is not only significance in this decision aesthetically but in a physical sense as well as we can calculate these scalar functions in post-processing as long as we keep track of our sampled Weyls.

With this channel, our average output can be expressed as

$$\langle \mu\chi \rangle = \langle\langle Q| \left(\mathcal{B}\Theta(A)\right)^m \mathcal{B} |\rho\rangle\rangle . \tag{4.7}$$

So far our final output is quite abstract as our definition of the channel $\mathcal{B}$ holds a very general structure. However, we can carefully select our distributions to give it some meaningful structure. Specifically, we will demonstrate construction of a projection channel via the character projection formula introduced in section 2.3. Working in the PTM representation, choose the distributions μ, ν to be only non-zero over a subgroup $\mathbb{G} \subseteq \mathbb{W}_{d,n}$. Given a random gate $U \in \mathbb{G}$ we have the implementation map

$$\Theta(U) = \sum_{W \in \mathbb{W}_{d,n}} |UWU^\dagger\rangle\rangle\langle\langle W| = \sum_{W \in \mathbb{W}_{d,n}} \chi_{U,W} |W\rangle\rangle\langle\langle W| \tag{4.8}$$

where $\chi_{U,W} \in \{\omega^j : j \in \mathbb{Z}_d\}$ is the character of U on subrepresentation W and satisfies

$$UW = \chi_{U,W} WU \tag{4.9}$$

which holds for all $U \in \mathbb{W}_{d,n}$. Given some set $\mathbb{Q} \subseteq \mathbb{W}_{d,n}$ we choose ν such that

$$\nu(U) = \sum_{Q \in \mathbb{Q}} \bar{\chi}_{U,Q}. \tag{4.10}$$

Examining eq.(4.6) then gives

$$\begin{aligned}
\mathcal{B} &= \int_{\mathbb{U}(D)} \mu(U) \left(\sum_{Q \in \mathbb{Q}} \bar{\chi}_{U,Q} \right) \left(\sum_{W \in \mathbb{W}_{d,n}} \chi_{U,W} |W\rangle\rangle\langle\langle W| \right) && (4.11)\\
&= \sum_{W \in \mathbb{W}_{d,n}} |W\rangle\rangle\langle\langle W| \int_{\mathbb{U}(D)} \mu(U) \sum_{Q \in \mathbb{Q}} \chi_{U,W} \overline{\chi}_{U,Q} && (4.12)\\
&= \sum_{W \in \mathbb{W}_{d,n}} |W\rangle\rangle\langle\langle W| \langle \chi_W, \chi_Q \rangle && (4.13)\\
&= \sum_{W \in \mathbb{W}_{d,n}} |W\rangle\rangle\langle\langle W| \delta_{W,Q} && (4.14)\\
&= \sum_{Q \in \mathbb{Q}} |Q\rangle\rangle\langle\langle Q|. && (4.15)
\end{aligned}$$

Given $\mu(U)$ is the normalized Haar measure we can recognize the integral as the character inner product eq.(2.34) and use Schur's orthogonality relations to establish $W = Q$ (lines

4.12 - 4.14). Then $\mathcal{B}$ simplifies to a projector onto the elements of $\mathbb{Q}$ (4.15). Achieving this, we have the main result of our protocol. A small invariant subspace (determined by our selection of $\mathbb{Q}$) is isolated in $\Theta(A)$ and coherently amplified via m applications. Through our selections of ρ and Q we can either learn the eigenvalues of this subblock or perform full process tomography. We continue our analysis by ensuring this protocol functions as expected under conditions that more closely resemble experimental conditions.

4.1 Finite Sampling

As is the case with any protocol involving the calculation of means of probability distributions, the empirical mean will deviate from the theoretical due to finite sampling effects. As such, it is crucial to establish some metric that ensures our empirical mean is close to the theoretical within a reasonable number of samples. One way to achieve this is via Hoeffding's inequality [24]. Let $X_1, \ldots, X_N$ be N real independent random variables such that they are bounded as $a \leq X_i \leq b$ where $a, b \in \mathbb{R}$. Given $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ and $t \geq 0$ Hoeffding's inequality gives

$$P(|\bar{X} - \mathbb{E}[\bar{X}]| \geq t) \leq 2\exp(-\frac{2Nt^2}{(b-a)^2}) \tag{4.16}$$

an upper bound on the probability that the finite average deviates from the expectation by more than t. In the context of the projective Rabi algorithm, we are interested in establishing a bound on the empirical output $\langle \mu\chi \rangle_N$ obtained by sampled unitaries $U_1, \ldots, U_N$ drawn according to μ. Examining the empirical channel $\tilde{\mathcal{B}}_N$,

$$\tilde{\mathcal{B}}_N = \sum_{W \in \mathbb{W}_{d,n}} |W\rangle\!\rangle\langle\!\langle W| \left(\frac{1}{N} \sum_{i=1}^{N} \chi_{U_i,W} \sum_{Q \in \mathbb{Q}} \overline{\chi}_{U_i,Q} \right). \tag{4.17}$$

our independent random variables will be $X_i = \chi_{U_i,W} \sum_{Q\in\mathbb{Q}} \bar{\chi}_{U_i,Q}$. We can use eq.(4.16) for $d = 2$ as $\chi_{U,W} \in [-1, 1]$ however for larger dimension qudits $\chi_{U,W}$ can take complex values. Therefore, we must establish Hoeffding's inequality for complex independent random variables. Consider the complex variable Z as $Z = \Re(Z) + i\Im(Z)$ and applying Hoeffding's inequality to the real and imaginary components as well as use of Boole's inequality[2] gives us [31]

$$P(|\bar{X} - \mathbb{E}[\bar{X}]| \geq t) \leq 4\exp(-\frac{2Nt^2}{(b-a)^2})\exp(-\frac{2Nt^2}{(d-c)^2}). \tag{4.18}$$

[2]See A.3 for a proof that achieves similar scaling without the union bound.

Now we can establish the probability bound of empirical channel $\langle \mu\chi \rangle_N$ deviating from $\langle \mu\chi \rangle$ by $t \geq 0$ as

$$Pr(|\langle \mu\chi \rangle_N - \langle \mu\chi \rangle| \geq t) \leq 2\exp(-\frac{Nt^2}{2|\mathbb{Q}|^{2(m+1)}}) \tag{4.19}$$

$$Pr(|\langle \mu\chi \rangle_N - \langle \mu\chi \rangle| \geq t) \leq 4\exp(-\frac{Nt^2}{4|\mathbb{Q}|^{2(m+1)}}) \tag{4.20}$$

for qubits and qudits respectively. Given a confidence interval (t, δ) we also have

$$Pr(|\langle \mu\chi \rangle_N - \langle \mu\chi \rangle| \geq t) \leq \delta \tag{4.21}$$

which can be used to establish a lower bound on the number of samples

$$N \geq \frac{\log(2/\delta)(2|\mathbb{Q}|^{2(m+1)})}{t^2} \tag{4.22}$$

$$N \geq \frac{\log(4/\delta)(4|\mathbb{Q}|^{2(m+1)})}{t^2} \tag{4.23}$$

again for qubits and qudits respectively. As is evident from these results, the minimum number of samples we need to establish some confidence interval directly depends on the size of $\mathbb{Q}$ and the number of iterations m which can become intractable very rapidly. For example, if we have a confidence interval of $(0.02, 0.99)$ and we wish to isolate a 2×2 subspace (i.e. $|\mathbb{Q}| = 2$) over $m = 5$ iterations the minimum number of samples would be on the order of 10^6 for qubits and qudits. We can eliminate this dependence by normalizing our function $\nu(U)$ such that $\nu(U) \in [-1, 1]$. However, doing so naively would pickup a normalization term to the power of $m+1$ in our output effectively killing the signal. Instead we can readjust our probability distribution to compensate for this. This is achieved by rewriting eq.(4.11) as

$$\begin{aligned} \mathcal{B} &= \sum_{W \in \mathbb{W}_{d,n}} |W\rangle\!\rangle\langle\!\langle W| \int_{\mathbb{U}(D)} \mu(U) C \frac{1}{C} \sum_{Q \in \mathbb{Q}} \chi_{U,W} \overline{\chi}_{U,Q} \\ &= \sum_{W \in \mathbb{W}_{d,n}} |W\rangle\!\rangle\langle\!\langle W| \int_{\mathbb{U}(D)} \mu'(U) \frac{1}{C} \sum_{Q \in \mathbb{Q}} \chi_{U,W} \overline{\chi}_{U,Q} \end{aligned} \tag{4.24}$$

where we define a truncated distribution μ' by

$$\mu'(U) = \begin{cases} C\mu(U) & \text{if } \nu(U) \neq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{4.25}$$

such that C is defined to maintain the normalized Haar measure $\int \mu'(U) = 1$. If we select ν such that $\nu(U) = 0$ for enough terms to give $C = |\mathbb{Q}|$ the probability bounds become

$$Pr(|\langle \mu\chi \rangle_N - \langle \mu\chi \rangle| \geq t) \leq 2\exp(-\frac{Nt^2}{2}) \tag{4.26}$$

$$Pr(|\langle \mu\chi \rangle_N - \langle \mu\chi \rangle| \geq t) \leq 4\exp(-\frac{Nt^2}{4}) \tag{4.27}$$

and the number of samples is lower bounded as

$$N \geq \frac{\log(2/\delta)2}{t^2} \tag{4.28}$$

$$N \geq \frac{\log(4/\delta)4}{t^2} \tag{4.29}$$

eliminating any dependence on $|\mathbb{Q}|$ or m making finite sampling effects converge rapidly regardless of parameters (e.g. dimensionality, number of qubits, number of iterations, etc.). Once again, as an example examining $(0.02, 0.99)$, $|\mathbb{Q}| = 2$ and $m = 5$ we get 10^3 samples which is significantly more 'reasonable'. It is important to note that tighter bounds can be established however, since the probability scales nicely with N for this looser bound then it can be argued the same is true for the tighter bound.

4.2 Robustness

So far we have worked under the unrealistic assumption that our gates are perfectly implemented[3]. We now turn our attention to the case that our quantum gates undergo some noise model thus resulting in imperfect implementations. In particular, we deal with Markovian noise where noise is gate dependent but not time dependent. This gives us a model that is more realistic than gate independent noise but does not necessarily reflect non-Markovian experimental conditions. If we compare to randomized benchmarking, our analysis benefits from the absence of the application of inverse gates which typically introduces correlated errors. We can then greatly simplify our gate dependent noise analysis as the average noise factorizes. This is facilitated with a telescoping series argument. First, take the difference between two outputs eq. (4.7) of different implementations Θ and Ω

$$\epsilon = \langle\langle Q| \left(\mathcal{B}\Theta(A)\right)^m \mathcal{B} |\rho\rangle\rangle - \langle\langle Q| \left(\mathcal{B}'\Omega(A)\right)^m \mathcal{B}' |\rho\rangle\rangle . \tag{4.30}$$

[3]Perfectly implemented in the sense that we ignore errors in the phase as they do not affect final results.

Now set $a^m = (\mathcal{B}\Theta(A))^m$ and $b^m = (\mathcal{B}'\Omega(A))^m$ and using eq. (A.41) to isolate a^m and substituting into eq. (4.30) we have

$$\epsilon = \sum_{j=1}^{m} \langle\langle Q| \left(\mathcal{B}\Theta(A)\right)^{m-j} \left(\mathcal{B}\Theta(A) - \mathcal{B}'\Omega(A)\right) \left(\mathcal{B}'\Omega(A)\right)^{j-1} \mathcal{B} |\rho\rangle\rangle + \langle\langle Q| \left(\mathcal{B}'\Omega(A)\right)^{m} \left(\mathcal{B} - \mathcal{B}'\right) |\rho\rangle\rangle . \tag{4.31}$$

With this we can start examining our algorithm under gate dependent noise. The two sources we can look at are imperfect fixed channel implementations $\Theta(A)$ and projective channels $\mathcal{B}$. Firstly, let us examine the case that our projectors are perfectly implemented and our implementation of the fixed channel has a small perturbation $\Delta_{\mathcal{A}}$. In this case, we set $\Theta(A) = \mathcal{A} + \Delta_{\mathcal{A}}$, $\Omega(A) = \mathcal{A}$ and $\mathcal{B}$,$\mathcal{B}'$ are ideal projectors. Eq. (4.31) is then

$$\epsilon_{\mathcal{A}} = \sum_{j=1}^{m} \langle\langle Q| \left(\mathcal{B}(\mathcal{A} + \Delta_{\mathcal{A}})\right)^{m-j} \mathcal{B}\Delta_{\mathcal{A}} \left(\mathcal{B}\mathcal{A}\right)^{j-1} \mathcal{B} |\rho\rangle\rangle \tag{4.32}$$

$$= \sum_{j=1}^{m} \langle\langle Q| \left(\mathcal{B}\mathcal{A}\right)^{m-j} \mathcal{B}\Delta_{\mathcal{A}} (\mathcal{B}\mathcal{A})^{j-1} \mathcal{B} |\rho\rangle\rangle + \mathcal{O}(\Delta_{\mathcal{A}}^2). \tag{4.33}$$

If Δ is small in comparison to $\Theta(A)$ then our error will be small and we still isolate our signal of interest. Next, we want to ensure the orthogonality conditions of our noisy projectors. This orthogonality condition suggest we are still projecting onto orthogonal subblocks of our PTM and deviations from imperfect projectors are therefore small. Let us denote our imperfect projector as $\tilde{\mathcal{B}} = \mathcal{B} + \Delta_{\mathcal{B}}$ where $\Delta_{\mathcal{B}}$ is a small perturbative term as a result of imperfect implementations of our Weyl gates and finite sampling effects. Examining $m+1$ applications of this projector[4] we have

$$\begin{aligned} \tilde{\mathcal{B}}^{m+1} &= (\mathcal{B} + \Delta_{\mathcal{B}})^{m+1} \\ &= \mathcal{B}^{m+1} + \sum_{j=0}^{m} \mathcal{B}^{m-j} \Delta_{\mathcal{B}} \mathcal{B}^{j} + \mathcal{O}(\Delta_{\mathcal{B}}^2) \end{aligned} \tag{4.34}$$

[4]This is equivalent to setting $\Theta(A) = I$ in our primary algorithm. Our analysis still holds however for arbitrary $\Theta(A)$ under the condition $[\tilde{\mathcal{B}}, \Theta(A)] \simeq 0$.

where $\mathcal{O}(\Delta^2)$ denotes second order perturbation. If we introduce an orthogonal projector $\mathcal{B}^\perp$ such that $\mathcal{B}\mathcal{B}^\perp = \mathcal{B}^\perp\mathcal{B} = 0$ into eq. (4.34) we get

$$\mathcal{B}^\perp\tilde{\mathcal{B}}^{m+1} = \mathcal{B}^\perp\mathcal{B}^{m+1} + \sum_{j=0}^{m} \mathcal{B}^\perp\mathcal{B}^{m-j}\Delta_{\mathcal{B}}\mathcal{B}^j + \mathcal{O}(\Delta_{\mathcal{B}}^2) \tag{4.35}$$

$$= (\mathcal{B}^\perp\Delta_{\mathcal{B}}\mathcal{B})\mathcal{B}^{m-1} + \mathcal{O}(\Delta_{\mathcal{B}}^2) \tag{4.36}$$

$$\simeq 0 \tag{4.37}$$

where $\mathcal{B}\Delta\mathcal{B}^\perp = \mathcal{B}^\perp\Delta\mathcal{B} \simeq 0$ is assumed. Therefore, we see under repeat applications of our noisy projective channel orthogonality is maintained up to first order. With this, our projective gate can be defined as $\tilde{\mathcal{B}} = \mathcal{B}_{ID} + \mathcal{B}^\perp$ where $\mathcal{B}_{ID}\mathcal{B}^\perp = \mathcal{B}^\perp\mathcal{B}_{ID} = 0$. Then (4.31) gives

$$\epsilon_{\mathcal{B}} = \sum_{j=1}^{m} \langle\langle Q| \left((\mathcal{B}_{ID} + \mathcal{B}^\perp)\mathcal{A}\right)^{m-j} \mathcal{B}^\perp\mathcal{A} \left(\mathcal{B}_{ID}\mathcal{A}\right)^{j-1} \mathcal{B}_{ID} |\rho\rangle\rangle + \langle\langle Q| \left(\mathcal{B}_{ID}\mathcal{A}\right)^m \mathcal{B}^\perp |\rho\rangle\rangle \tag{4.38}$$

if $[\mathcal{B}_{ID}, \mathcal{A}] = 0$ we can exploit the orthogonality condition in both terms giving $\epsilon_{\mathcal{B}} = 0$. Finally, we examine both imperfect channels. Substituting $\Theta(\mathcal{A}) = \mathcal{A} + \Delta$, $\Omega(\mathcal{A}) = \mathcal{A}$, $\mathcal{B} = \mathcal{B}_{ID} + \mathcal{B}^\perp$ and $\mathcal{B}' = \mathcal{B}_{ID}$ we get

$$\epsilon_{\mathcal{A},\mathcal{B}} = \sum_{j=1}^{m} \langle\langle Q| \left(\mathcal{B}_{ID}(\mathcal{A} + \Delta_{\mathcal{A}})\right)^{m-j} \mathcal{B}_{ID}\Delta_{\mathcal{A}} \left(\mathcal{B}_{ID}\mathcal{A}\right)^{j-1} \mathcal{B}_{ID} |\rho\rangle\rangle \tag{4.39}$$

$$= \sum_{j=1}^{m} \langle\langle Q| \left(\mathcal{B}_{ID}\mathcal{A}\right)^{m-j} \mathcal{B}_{ID}\Delta_{\mathcal{A}} (\mathcal{B}_{ID}\mathcal{A})^{j-1} \mathcal{B}_{ID} |\rho\rangle\rangle + \mathcal{O}(\Delta_{\mathcal{A}}^2). \tag{4.40}$$

once again using the commutation between $\mathcal{B}_{ID}$ and $\mathcal{A}$ as well as the orthogonality between $\mathcal{B}_{ID}$ and $\mathcal{B}^\perp$. Thus, our final output will be of the form

$$\langle\widetilde{\mu\chi}\rangle = \langle\langle Q| \left(\mathcal{B}\mathcal{A}\right)^m \mathcal{B} |\rho\rangle\rangle + \epsilon. \tag{4.41}$$

giving a robust protocol for relatively small errors.

4.3 Example

As an example let us examine the case of an n-qudit rotation gate,

$$R = G^\theta \tag{4.42}$$

where $G \in \mathbb{W}_{d,n}$ is the generator of the rotation and $\theta \in \mathbb{R}$ is the angle of rotation. Say we wish to characterize the angle of rotation and notice that G can be diagonalized such that $G = CZ_0C^\dagger$ where C is a Clifford operator and $Z_0 = Z \otimes I^{\otimes n-1}$. Our rotation gate then becomes

$$R(\theta) = CZ_0^\theta C^\dagger. \tag{4.43}$$

Examining the case of $G = Z_0$ and generalizing is then simple as we can set $\mathbb{Q} \to C^\dagger \mathbb{Q} C$ to obtain rotations for other G. This is evident when we examine the PTM of these rotations. As an example, Figure 4.1 shows the PTM of a single qubit Z rotation versus a single qubit X rotation. Applying a Clifford to the Z rotation such that $X \to Y$, $Y \to Z$ and $Z \to X$ gives the PTM of the X rotation.

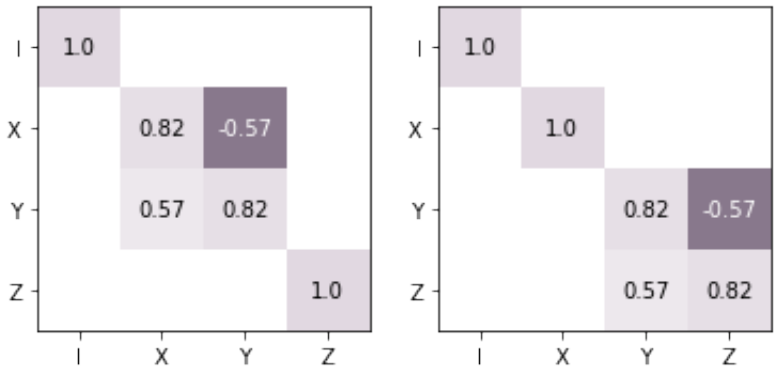

Figure 4.1: Pauli transfer matrix of a single qubit Z rotation (left) and X rotation (right) generated by $\theta = 0.609$ Rad. Diagonal blocks of 0.82 represent the $\cos(\theta)$ components of the rotation and off diagonal blocks of 0.57 represent the $\sin(\theta)$ components.

Continuing on, let $\mathbb{Q} = X_0\langle Z_0 \rangle$ where $X_0 = X \otimes I^{n-1}$. Our distribution ν then becomes

$$\begin{aligned} \nu(U) &= \bar{\chi}_{U,X_0} \sum_{j \in \mathbb{Z}_d} \bar{\chi}^j_{U,Z_0} \\ &= \begin{cases} d\bar{\chi}_{U,X_0} & [U, Z_0] = 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{4.44}$$

satisfying the condition $C = |\mathbb{Q}|$ as $|\mathbb{G}| = d^{2n}$, $|\mathbb{Q}| = d$ and d^{2n-1} terms commute with Z_0.

Now, examining the output of the circuit

$$\begin{aligned}
\langle \mu\chi \rangle &= \langle\langle Q| \left(\mathcal{B}\Theta(R)\right)^m \mathcal{B} |\rho\rangle\rangle \\
&= \langle\langle Q| \mathcal{B}(\Theta(R))^m |\rho\rangle\rangle \\
&= \langle\langle Q| \left(\sum_{Q\in\mathbb{Q}} |R^m Q (R^m)^\dagger\rangle\rangle\langle\langle Q| \right) |\rho\rangle\rangle \\
&= \langle\langle Q| \left(\sum_{i\in\mathbb{Z}_d} |Z_0^{m\theta} X_0 Z_0^i Z_0^{-m\theta}\rangle\rangle\langle\langle X_0 Z_0^i| \right) |\rho\rangle\rangle
\end{aligned} \tag{4.45}$$

where we used $[\mathcal{B}, \Theta(R)] = 0$. Note that the Z_0^i and $Z_0^{m\theta}$ terms also commute. We then want to solve for $Z_0^{m\theta} X_0 Z_0^{-m\theta}$. To achieve this we first can decompose $Z^{m\theta}$ into

$$Z_0^{m\theta} = \sum_{k\in\mathbb{Z}_d} a_k Z_0^k \tag{4.46}$$

Coefficients a_k can be solved for by using the Hilbert-Schmidt inner product and the fact that the Weyls form an orthonormal basis

$$a_k = \langle Z_0^k, Z_0^{m\theta} \rangle \tag{4.47}$$

$$= \frac{1}{D} \mathrm{Tr}((Z_0^\dagger)^k Z_0^{m\theta}) \tag{4.48}$$

$$= \frac{1}{D} \mathrm{Tr}(\sum_{j,j'\in\mathbb{Z}_d} \omega^{-jk} \omega^{j'm\theta} |j\rangle\langle j|j'\rangle\langle j'|) \mathrm{Tr}(I^{n-1}) \tag{4.49}$$

$$= \frac{1}{d} \sum_{j\in\mathbb{Z}_d} \omega^{j\theta} \omega^{-jk} \tag{4.50}$$

Examining $Z^{m\theta}XZ^{-m\theta}$ now

$$Z_0^{m\theta}X_0Z_0^{-m\theta} = \left(\frac{1}{d}\sum_{j,k\in\mathbb{Z}_d}\omega^{jm\theta}\omega^{-jk}Z_0^k\right)X_0\left(\frac{1}{d}\sum_{j',k'\in\mathbb{Z}_d}\omega^{-j'm\theta}\omega^{-j'k'}Z_0^{k'}\right) \tag{4.51}$$

$$= \frac{1}{d^2}\sum_{j,k,j',k'\in\mathbb{Z}_d}\omega^{jm\theta}\omega^{-jk}\omega^{-j'm\theta}\omega^{-j'k'}Z_0^kX_0Z_0^{k'} \tag{4.52}$$

$$= \frac{1}{d^2}\sum_{j,k,j',k'\in\mathbb{Z}_d}\omega^{jm\theta}\omega^{-jk}\omega^{-j'm\theta}\omega^{-j'k'}\omega^kX_0Z_0^{k+k'} \tag{4.53}$$

$$= \frac{1}{d^2}\sum_{j,j',n\in\mathbb{Z}_d}\omega^{jm\theta}\omega^{-j'm\theta}\omega^{-j'n'}\left(\sum_{k\in\mathbb{Z}_d}\omega^{k(-j+j'+1)}\right)X_0Z_0^n \tag{4.54}$$

$$= \frac{1}{d}\sum_{n\in\mathbb{Z}_d}\left(\omega^{-(d-1)m\theta}\omega^{-(d-1)n} + \omega^{m\theta}\sum_{j\in\mathbb{Z}_{d-1}}\omega^{-jn}\right)X_0Z_0^n \tag{4.55}$$

$$= \frac{\omega^{m\theta}}{d}\sum_{n\in\mathbb{Z}_d}\omega^n\left(\omega^{-dm\theta} + \sum_{j=1}^{d-1}\omega^{-jn}\right)X_0Z_0^n \tag{4.56}$$

The final circuit output is then

$$\langle\mu\chi\rangle = \frac{\omega^{m\theta}}{d}\sum_{n,i\in\mathbb{Z}_d}\omega^n\left(\omega^{-dm\theta} + \sum_{j=1}^{d-1}\omega^{-jn}\right)\langle\langle Q|\left|X_0Z_0^{n+i}\right\rangle\rangle\langle\langle X_0Z_0^i\right||\rho\rangle\rangle \tag{4.57}$$

Robust selection of ρ and Q allows for isolation of a single term that can be fitted and can be used to determine θ. For example, set $d = 2$ (i.e. the case of a multi-qubit Pauli

rotation gate) using eq. (4.57) we get

$$\langle \mu\chi \rangle = \frac{\omega^{m\theta}}{2} \sum_{n,i=0}^{1} \omega^{n} \left(\omega^{-2m\theta} + \omega^{-jn} \right) \langle\langle Q | \left| X_0 Z_0^{n+i} \right\rangle\rangle \langle\langle X_0 Z_0^{i} \right| | \rho \rangle\rangle \tag{4.58}$$

$$= \frac{1}{2}(\omega^{m\theta} + \omega^{-m\theta}) \sum_{i=0}^{1} \langle\langle Q | \left| X_0 Z_0^{i} \right\rangle\rangle \langle\langle X_0 Z_0^{i} \right| | \rho \rangle\rangle \tag{4.59}$$

$$+ \frac{1}{2}(\omega^{m\theta} - \omega^{-m\theta}) \sum_{i=0}^{1} \langle\langle Q | \left| X_0 Z_0^{1+i} \right\rangle\rangle \langle\langle X_0 Z_0^{i} \right| | \rho \rangle\rangle$$

$$= \cos(\pi m\theta) \sum_{i=0}^{1} \langle\langle Q | \left| X_0 Z_0^{i} \right\rangle\rangle \langle\langle X_0 Z_0^{i} \right| | \rho \rangle\rangle \tag{4.60}$$

$$+ i \sin(\pi m\theta) \sum_{i=0}^{1} \langle\langle Q | \left| X_0 Z_0^{1+i} \right\rangle\rangle \langle\langle X_0 Z_0^{i} \right| | \rho \rangle\rangle$$

with $\cos(\pi m\theta)$ on the diagonal of a 2×2 matrix and $\sin(\pi m\theta)$ on the off diagonal terms showing we have successfully isolated a single qubit rotational gate signal. Selections of ρ and Q can then be restricted to X_0 or $X_0 Z_0$ where if $\rho = Q$ then a $\cos(\pi m\theta)$ term is isolated and a $\sin(\pi m\theta)$ term otherwise. This result is essentially equivalent to a more generalized version of the QPE protocol outlined in *Kimmel et al.* [26]. As discussed in that paper, repeat applications of a single qubit rotation gate gives a Heisenberg-limited protocol for the determination of the parameter θ. We can conclude, if the time to perform an experimental run is proportional to m, the empirical estimate of θ has standard deviation $\sigma(\theta)$ that scales with $\mathcal{O}(1/m)$ achieving Heisenberg scaling.

Chapter 5

Numerical Simulations

We can now validate these results via numerical simulations. These simulations are performed via the True-Q™ software package [5] provided by Keysight Technologies Inc. This software package functions with Python 3.7 or above and provides a set of tools to run circuits on hardware and perform error diagnostics and error suppression. More importantly for us, it also provides a powerful and efficient simulator with the ability to introduce various realistic noise sources into our circuit. For our purposes, we use the built-in noise models which include over rotation, stochastic Pauli, and SPAM noise models.

We aim to simulate a full experimental protocol for the characterization of a multi-qubit rotation gate as presented in section 4.3. Here we set our system to be a two qubit device and the gate of interest to be a rotation generated by Weyl XX and angle of rotation $\pi/4$.[1] We run our protocol by sampling over the Weyl operators that commute with the generator XX, choose $\mathbb{Q} = \{IY, XZ\}$ and set $\rho = Q = IY$ where $Y = XZ$. Then, we run our algorithm for each $m = 1, \ldots, m_{max}$ where $m_{max} = 30$ with 5000 shots per circuit averaging over $N = 20$ random circuits. Assuming no noise sources present we demonstrate that we can successfully reproduce a Rabi oscillation experiment in a multi-qubit system under ideal conditions. Of course, as previously discussed these are unrealistic expectations to have when dealing with physical systems and so we can also introduce noise into our simulations. First we include over rotation errors. This is achieved in the software by replacing the ideal unitary U with $U_{err} = U^{1+\epsilon}$. For the purposes of our simulations we set ϵ as 0.1/0.2 for single qubit gates and multi-qubit gates, respectively. Next, we

[1] We can run these simulations on any arbitrary multi-qudit rotation given we adjust our sampling group and $\mathbb{Q}$ appropriately.

introduce a stochastic Pauli noise channel. These channels are an effective model for simulating noisy gates in physical implementations where the channel eq. 3.18 is interleaved for each cycle and we set $c_{X,Y,Z} = (0.01, 0.01, 0.01)$. Finally, we introduce some SPAM error in the form of a readout and state preparation error with a 1% chance of bit-flip during state preparation or readout. Simulating our experiment we find that we maintain our Rabi oscillations however, our system experiences decoherence with an increase in circuit depth that manifests itself as an amplitude damping channel. In both the ideal and noisy simulations, a least squares fitting function was used to estimate the parameter θ by fitting a cosine function to the simulated results. For noiseless circuits, the angle θ was determined to be 0.785 Rad $\pm$ 0.001 and for noisy circuits 0.805 Rad $\pm$ 0.010 agreeing with our expectations of 0.785 Rad.

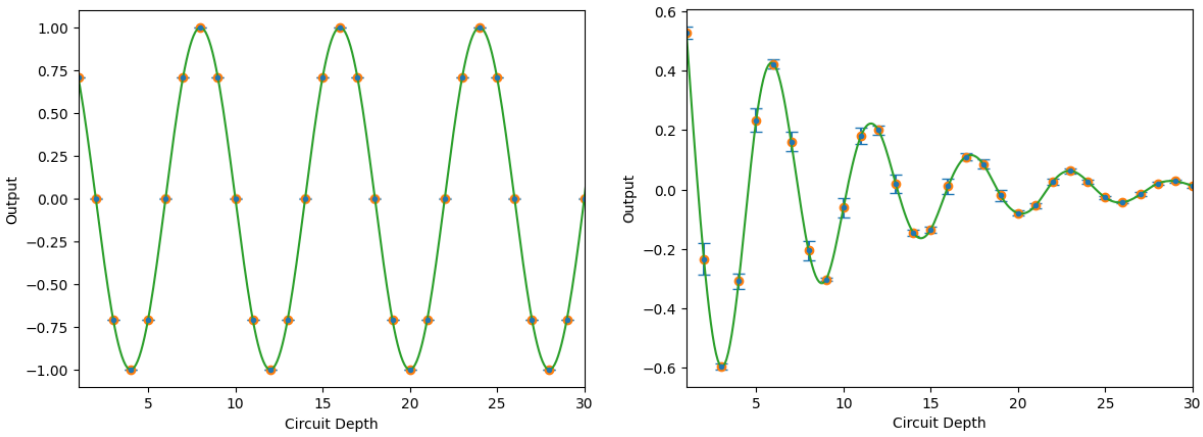

Figure 5.1: Experimental output for the two qubit $R_{XX}(\pi/4)$ gate in the case of an ideal system (left) and a system undergoing over-rotation, stochastic Pauli noise and SPAM errors (right). In both cases, a cosine function was able to be successfully fitted and θ determined.

In addition to the Rabi oscillations, we wish to verify our claims of rapid finite sampling convergence and Heisenberg-limited behaviour. To achieve this, we plot our standard deviation of θ versus number of samples and number of iterations, respectively.

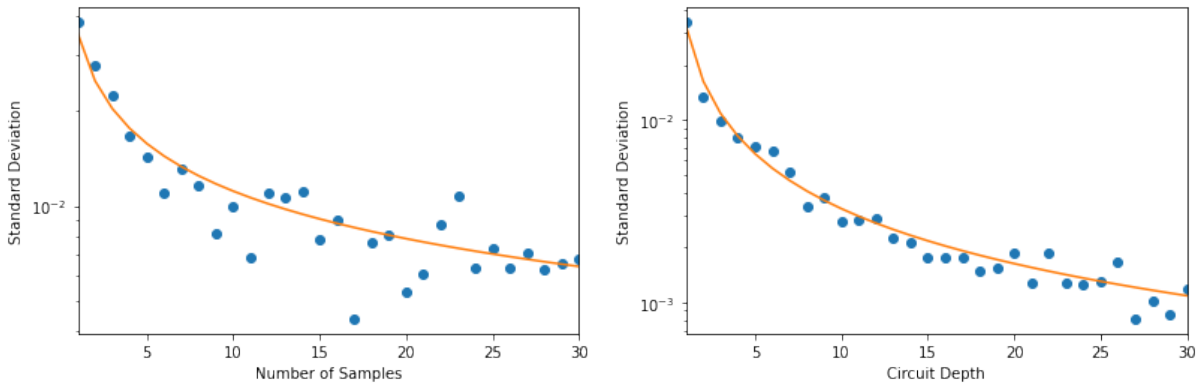

Figure 5.2: Standard deviation convergence under number of samples (left) and number of iterations (right).

On the left hand side, we plot our standard deviation versus sample size N. To do this, we set $m = 1$ to minimize its contribution in suppressing deviation. We simulate our noisy output using the same over rotation and readout errors and an inflated stochastic Pauli noise $c_{X,Y,Z} = (0.1, 0.1, 0.1)$ over different values of N. The function $f(x) = 1/\sqrt{x}$ is then fitted to the data. As we can see, our results match the expected scaling of $1/\sqrt{N}$. Thus, we can conclude that even under the effects of Markovian noise our protocol converges rapidly with finite samples.

Now, on the right hand side we plot standard deviation versus number of iterations (or equivalently circuit depth or number of cycles) m. We set $N = 1$ once again to minimize its contributions to the overall error. Fitting the function $f(x) = 1/x$ to our data we see our standard deviation scales accordingly with respect to m. This is what is expected given the Heisenberg-limited nature of our protocol.

To further validate this we can compare our results to ones obtained for a well known protocol, namely randomized benchmarking. Randomized benchmarking is simulated using the same parameters as before with the one exception being $N = 30$ to obtain well defined standard deviations (handled by True-Q). In figure 5.3, we can see a normalized standard deviation for both our projective Rabi experiment versus randomized benchmarking plotted against circuit depth. Comparatively, the projective Rabi experiment performs better than RB. This further validates the Heisenberg-limited nature of our protocol.

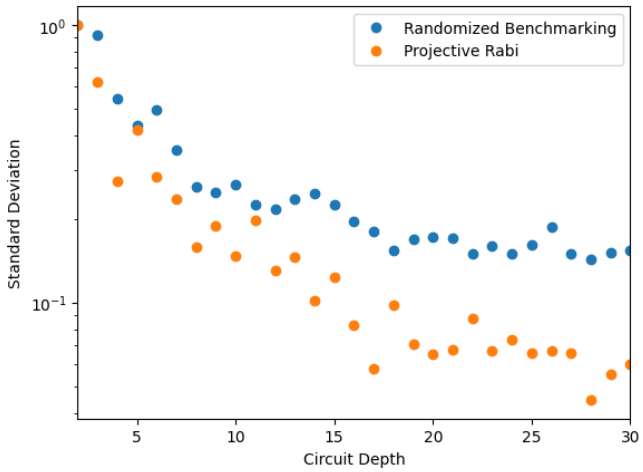

Figure 5.3: Comparison of standard deviation for the projective Rabi experiment versus randomized benchmarking. Standard deviations are normalized based off the maximum deviation for each protocol.

Chapter 6

Conclusion

In this thesis, we introduced a projective Rabi experiment. Functioning similarly to a Rabi oscillation experiment, the projective Rabi experiment's novelty stems from interleaved projection channels introduced between m applications of a quantum channel of interest $\mathcal{A}$ allowing for coherent amplification of some signal of interest. These projective channels are constructed using a fundamental result of character theory. Specifically, we chose sampling distributions to purposefully induce character transforms onto irreducible subrepresentations of the PTM via the character projection formula. This is an efficient task as it only requires the application of random Weyls and a fixed scalar function that can be efficiently computed.

We then examined the scalibility and robustness of the algorithm. It was demonstrated that, considering finite sampling effects, this algorithm is fully scalable given that the number of terms where $\nu(U) \neq 0$ is equal to $|\mathbb{G}|/|\mathbb{Q}|$ and a truncated sampling distribution is used. Additionally, it was shown that under gate dependent noise we maintain our intended output with a relatively small perturbative term.

Next, an example of a multi-qudit rotation gate was used to showcase the usefulness of the algorithm. We saw how projecting onto a small $d \times d$ subblock generated by the set $X_0\langle Z_0\rangle$ isolates a single signal in our gate that is coherently amplified through repeat applications. That signal could then be used to either isolate the rotation angle θ as a QPE technique or full process tomography on the subspace could be performed.

Finally, we validated our results with numerical simulations. We were able to demon-

strate that even with noise acting on our system we can obtain a Rabi oscillatory signal that when fitted to the appropriate cosine function can determine its rotation angle. We also demonstrated the standard deviation scaling as $\mathcal{O} \propto 1/m\sqrt{N}$ showing that both finite sampling errors converge rapidly and Heisenberg-limited scaling.

For future work, demonstrating the viability of this algorithm experimentally would be the next crucial step. Although robustness to Markovian experimental conditions such as gate dependent noise and SPAM error has been shown; experimental verification is still necessary as non-Markovian effects may affect results. Experimentally implementing our example of a multi-qudit rotation gate on a system that can perform high fidelity rotations is the recommendation. Additionally, exploring how to adapt the algorithm to different Hamiltonians of interest would be useful. As an example, the Ising Hamiltonian

$$H = -\sum_j J_{j,j+1}\sigma_j\sigma_{j+1} - \sum_j h\sigma_j \tag{6.1}$$

can be examined. Constructing projectors onto some subspace such that the coefficients $J_{j,j+1}$ can be determined would give an efficient and robust technique for full tomography of Ising systems.

References

[1] Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error rate. *SIAM Journal on Computing*, 38(4):1207–1282, 2008.

[2] Panos Aliferis and Andrew W. Cross. Subsystem fault tolerance with the Bacon-Shor code. *Physical Review Letters*, 98(22), may 2007.

[3] Panos Aliferis, Daniel Gottesman, and John Preskill. Accuracy threshold for postselected quantum computation. *Quantum information & computation*, 8, 04 2007.

[4] P.W. Anderson. More is different. *Science, New Series*, 177(4047):393–396, 1972.

[5] Stefanie J. Beale, Arnaud Carignan-Dugas, Dar Dahlen, Joseph Emerson, Ian Hincks, Pavithran Iyer, Aditya Jain, David Hufnagel, Egor Ospadov, Jordan Saunders, Andrew Stasiuk, Joel J. Wallman, and Adam Winick. True-q, June 2020.

[6] Paul Benioff. The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. *Journal of Statistical Physics*, 22:563–591, 05 1980.

[7] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Jonathan Mizrahi, Jonathan D. Sterk, and Peter Maunz. Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit, 2013.

[8] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Kenneth Rudinger, Jonathan Mizrahi, Kevin Fortier, and Peter Maunz. Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nature Communications*, 8(1), 2017.

[9] Zhenyu Cai and Simon C. Benjamin. Constructing smaller Pauli twirling sets for arbitrary error channels. *Scientific Reports*, 9(1), aug 2019.

[10] J. M. Chow, J. M. Gambetta, L. Tornberg, Jens Koch, Lev S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Randomized benchmarking and process tomography for gate errors in a solid-state qubit. *Phys. Rev. Lett.*, 102:090502, Mar 2009.

[11] Justin E. Christensen, David Hucul, Wesley C. Campbell, and Eric R. Hudson. High-fidelity manipulation of a qubit enabled by a manufactured nucleus. *npj Quantum Information*, 6(1), apr 2020.

[12] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Physical Review A*, 80(1), jul 2009.

[13] Alicja Dutkiewicz, Barbara M. Terhal, and Thomas E. O'Brien. Heisenberg-limited quantum phase estimation of multiple eigenvalues with up to two control qubits, 2021.

[14] Joseph Emerson, Robert Alicki, and Karol Życzkowski. Scalable noise estimation with random unitary operators. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(10):S347–S352, sep 2005.

[15] Jeffrey M. Epstein, Andrew W. Cross, Easwar Magesan, and Jay M. Gambetta. Investigating the limits of randomized benchmarking protocols. *Physical Review A*, 89(6), jun 2014.

[16] Alexander Erhard, Joel J. Wallman, Lukas Postler, Michael Meth, Roman Stricker, Esteban A. Martinez, Philipp Schindler, Thomas Monz, Joseph Emerson, and Rainer Blatt. Characterizing large-scale quantum computers via cycle benchmarking. *Nature Communications*, 10(1), nov 2019.

[17] Robert E. Fontana and Gary M. Decad. Moore's law realities for recording systems and memory storage components: Hdd, tape, nand, and optical. *AIP Advances*, 8(5):056506, 2018.

[18] J. P. Gaebler, A. M. Meier, T. R. Tan, R. Bowler, Y. Lin, D. Hanneke, J. D. Jost, J. P. Home, E. Knill, D. Leibfried, and D. J. Wineland. Randomized benchmarking of multiqubit gates. *Phys. Rev. Lett.*, 108:260503, Jun 2012.

[19] Christopher Granade, Christopher Ferrie, and D G Cory. Accelerated randomized benchmarking. *New Journal of Physics*, 17(1):013042, jan 2015.

[20] Akel Hashim, Ravi K. Naik, Alexis Morvan, Jean-Loup Ville, Bradley Mitchell, John Mark Kreikebaum, Marc Davis, Ethan Smith, Costin Iancu, Kevin P. O'Brien, Ian Hincks, Joel J. Wallman, Joseph Emerson, and Irfan Siddiqi. Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor. *Physical Review X*, 11(4), nov 2021.

[21] Jonas Helsen, Ingo Roth, Emilio Onorati, Albert H. Werner, and Jens Eisert. A general framework for randomized benchmarking, 2020.

[22] Jonas Helsen, Xiao Xue, Lieven M. K. Vandersypen, and Stephanie Wehner. A new class of efficient randomized benchmarking protocols, 2018.

[23] B L Higgins, D W Berry, S D Bartlett, M W Mitchell, H M Wiseman, and G J Pryde. Demonstrating Heisenberg-limited unambiguous phase estimation without adaptive measurements. *New Journal of Physics*, 11(7):073023, jul 2009.

[24] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[25] Mikhail Isaev and Brendan D. McKay. On a bound of Hoeffding in the complex case. *Electronic Communications in Probability*, 21, Jan 2016.

[26] Shelby Kimmel, Guang Hao Low, and Theodore J. Yoder. Robust calibration of a universal single-qubit gate set via robust phase estimation. *Physical Review A*, 92(6), dec 2015.

[27] Easwar Magesan, Jay M. Gambetta, and Joseph Emerson. Characterizing quantum gates via randomized benchmarking. *Physical Review A*, 85(4), apr 2012.

[28] Seth T. Merkel, Jay M. Gambetta, John A. Smolin, Stefano Poletto, Antonio D. Córcoles, Blake R. Johnson, Colm A. Ryan, and Matthias Steffen. Self-consistent quantum process tomography. *Phys. Rev. A*, 87:062119, Jun 2013.

[29] R. Merlin. Rabi oscillations, Floquet states, Fermi's golden rule, and all that: Insights from an exactly solvable two-level model. *American Journal of Physics*, 89(1):26–34, 2021.

[30] Mohan Sarovar, Timothy Proctor, Kenneth Rudinger, Kevin Young, Erik Nielsen, and Robin Blume-Kohout. Detecting crosstalk errors in quantum information processors. *Quantum*, 4:321, sep 2020.

[31] Dan Stahlke. Quantum interference as a resource for quantum speedup. *Physical Review A*, 90(2), Aug 2014.

[32] Sisi Zhou, Mengzhen Zhang, John Preskill, and Liang Jiang. Achieving the Heisenberg limit in quantum metrology using quantum error correction. *Nature Communications*, 9(1), 2018.

APPENDICES

Appendix A

Mathematical Proofs

A.1 Schur's Orthogonality Relations

Theorem 1. *Let $\mathbb{G}$ be a compact group with normalized Haar measure $\mu : \Sigma(\mathbb{G}) \to \mathbb{R}_{\geq 0}$. Let $\mathcal{V}$ and $\mathcal{Z}$ be finite dimensional complex vector spaces with inner product $\langle \cdot | \cdot \rangle : \mathcal{A} \times \mathcal{A} \to \mathbb{C}$ where $\mathcal{A}$ is the respective vector space. Let $\pi : \mathbb{G} \to U(\dim_{\mathbb{C}} \mathcal{V})$ and $\rho : \mathbb{G} \to U(\dim_{\mathbb{C}} \mathcal{Z})$ be inequivalent irreducible unitary representations and let $u, v, u_1, u_2, v_1, v_2 \in \mathcal{V}$ and $x, y \in \mathcal{Z}$. Schur's orthogonality relations are*

$$\int_{\mathbb{G}} \langle \pi(g)u|v\rangle \overline{\langle \rho(g)x|y\rangle} d\mu(g) = 0 \tag{A.1}$$

$$\int_{\mathbb{G}} \langle \pi(g)u_1|v_1\rangle \overline{\langle \pi(g)u_2|v_2\rangle} d\mu(g) = \frac{\langle u_1|u_2\rangle \overline{\langle v_1|v_2\rangle}}{\dim_{\mathbb{C}} \mathcal{V}}. \tag{A.2}$$

Proposition 1. *Let $L \in \mathrm{End}_{\mathbb{C}}(\mathcal{Z}, \mathcal{V})$ then*

$$\int_{\mathbb{G}} \pi(g) \circ L \circ \rho(g)^{\dagger} d\mu(g) = [\pi \cong \rho] \frac{\mathrm{Tr}(L)}{\dim_{\mathbb{C}} \mathcal{V}} \mathbb{I}_{\mathcal{V}} \tag{A.3}$$

Proof. First show $\tilde{L} \equiv \int_{\mathbb{G}} \pi(g) \circ L \circ \rho(g)^\dagger d\mu(g)$ is $\mathbb{G}$-invariant. Notice $\forall h \in \mathbb{G}$ we have

$$\pi(h) \circ \tilde{L} = \pi(h) \circ \int_{\mathbb{G}} \pi(g) \circ L \circ \rho(g)^\dagger d\mu(g) \tag{A.4}$$

$$= \int_{\mathbb{G}} \pi(hg) \circ L \circ \rho(g)^\dagger d\mu(g) \tag{A.5}$$

$$= \int_{\mathbb{G}} \pi(hg) \circ L \circ \rho(g)^\dagger d\mu(hg) \tag{A.6}$$

$$= \int_{\mathbb{G}} \pi(g) \circ L \circ \rho(h^{-1}g)^\dagger d\mu(g) \tag{A.7}$$

Since $\rho(h^{-1}g)^\dagger = \rho(g)^\dagger \rho(h)$ then $\tilde{L} \in \mathrm{Hom}_{\mathbb{G}}(\rho, \pi)$. From this Schur's Lemma implies $\tilde{L}$ must be $\lambda \mathbb{I}_{\mathcal{V}}$ with $\lambda \in \mathbb{C}$ determined by the trace when $\pi \cong \rho$ and $\tilde{L} = 0$ otherwise proving the proposition. Now the proof for Schur's orthogonality relations follows. Set $L = |v\rangle\langle y| \in \mathrm{End}_{\mathbb{C}}(\mathcal{Z}, \mathcal{V})$ then recognize $\overline{\langle \rho(g)x|y\rangle} = \langle y|\rho(g)x\rangle$ and $\langle \pi(g)u|v\rangle = \langle u|\pi(g)^\dagger v\rangle$ allowing us to use the proposition and thus proving Schur's orthogonality relations. □

A.2 General Character Projection Formula

Theorem 2. *Let $\mathbb{G}$ be a compact group with normalized Haar measure $\mu : \Sigma(\mathbb{G}) \to \mathbb{R}_{\geq 0}$. Let $\pi : \mathbb{G} \to U(\mathcal{H})$ be a unitary representation of $\mathbb{G}$ on a separable Hilbert space $\mathcal{H}$. Let π_k be a finite dimensional irreducible unitary subrepresentation of π on $\mathcal{H}_k$. Let $|e_1^{(k)}\rangle, |e_2^{(k)}\rangle, \ldots, |e_{n_k}^{(k)}\rangle$ be an orthonormal basis for $\mathcal{H}_k$ with $n_k \equiv \dim_{\mathbb{C}} \mathcal{H}_k$ and let $a_{j,l}^{(k)}(g) \equiv \langle e_j^{(k)}|\pi_k(g)|e_l^{(k)}\rangle$. Then the character projection formula is defined as*

$$P_{j,l}^{(k)} \equiv n_k \int_{\mathbb{G}} \overline{a_{j,l}^{(k)}(g)} \pi(g) d\mu(g) \tag{A.8}$$

Proposition 2.

$$\pi(h) P_{j,l}^{(k)} = \sum_{v=1}^{n_k} a_{v,j}^{(k)}(h) P_{v,l}^{(k)} \tag{A.9}$$

Proof.

$$\pi(h)P^{(k)}_{j,l} = n_k \int_{\mathbb{G}} \overline{a^{(k)}_{j,l}(g)}\pi(h)\pi(g)d\mu(g) \tag{A.10}$$

$$= n_k \int_{\mathbb{G}} \overline{a^{(k)}_{j,l}(g)}\pi(hg)d\mu(g) \tag{A.11}$$

$$= n_k \int_{\mathbb{G}} \overline{a^{(k)}_{j,l}(h^{-1}x)}\pi(x)d\mu(h^{-1}x) \tag{A.12}$$

$$= n_k \int_{\mathbb{G}} \sum_{v=1}^{n_k} \overline{\langle e^{(k)}_j|\pi(h^{-1}x)|e^{(k)}_v\rangle\langle e^{(k)}_v|\pi(x)|e^{(k)}_l\rangle}\pi(x)d\mu(h^{-1}x) \tag{A.13}$$

$$= n_k \sum_{v=1}^{n_k} \overline{\langle e^{(k)}_j|\pi(h^{-1})|e^{(k)}_v\rangle} \int_{\mathbb{G}} \overline{a^{(k)}_{v,l}(x)}\pi(x)d\mu(h^{-1}x) \tag{A.14}$$

$$= \sum_{v=1}^{n_k} \overline{a^{(k)}_{v,j}(h^{-1}}n_k \int_{\mathbb{G}} \overline{a_{v,l}(hy)}\pi(hy)d\mu(y) \tag{A.15}$$

$$= \sum_{v=1}^{n_k} \overline{a^{(k)}_{v,j}(h^{-1}}n_k \int_{\mathbb{G}} \overline{a_{v,l}(y)}\pi(y)d\mu(y) \tag{A.16}$$

$$= \sum_{v=1}^{n_k} a^{(k)}_{v,j}(h)P^{(k)}_{v,l} \tag{A.17}$$

□

Proposition 3.

$$P^{(k)}_{j,l}P^{(k^{'})}_{\mu,\nu} = \delta_{k,k'}\delta_{l,\mu}P^{(k)}_{j,\nu} \tag{A.18}$$

Proof.

$$P_{j,l}^{(k)} P_{\mu,\nu}^{(k')} = n_k \int_{\mathbb{G}} \overline{a_{j,l}^{(k)}(g)} \pi(g) P_{\mu,\nu}^{(k')} d\mu(g) \tag{A.19}$$

$$= n_k \int_{\mathbb{G}} \overline{a_{j,l}^{(k)}(g)} \sum_{t=1}^{n_k} a_{t,\nu}^{(k')}(g) P_{t,\nu}^{(k')} d\mu(g) \tag{A.20}$$

$$= n_k \sum_{t=1}^{n_k} \left[\int_{\mathbb{G}} \overline{a_{j,l}^{(k)}(g)} a_{t,\nu}^{(k')}(g) d\mu(g) \right] P_{t,\nu}^{(k')} \tag{A.21}$$

$$= n_k \sum_{t=1}^{n_k} \frac{\delta_{k,k'} \delta_{t,j} \delta_{\mu,l}}{n_k} P_{t,\nu}^{(k')} \tag{A.22}$$

$$= \delta_{k,k'} \delta_{l,\mu} P_{j,\nu}^{(k)} \tag{A.23}$$

□

Proposition 4. *Projector $P_{i,j}$ is self adjoint. That is,*

$$P_{j,l}^{(k)\dagger} = P_{j,l}^{(k)} \tag{A.24}$$

Proof.

$$P_{j,l}^{(k)\dagger} = n_k \int_{\mathbb{G}} a_{j,l}^{(k)}(g) \pi(g)^\dagger d\mu(g)^\dagger \tag{A.25}$$

$$= n_k \int_{\mathbb{G}} \overline{a_{j,l}^{(k)}(g^{-1})} \pi(g^{-1}) d\mu(g^1) \tag{A.26}$$

$$= n_k \int_{\mathbb{G}} \overline{a_{j,l}^{(k)}(g)} \pi(g) d\mu(g) \tag{A.27}$$

$$= P_{j,l}^{(k)} \tag{A.28}$$

Thus, we see the projectors onto the mutually inequivalent irreducible representations of π are a complete set of orthogonal projectors proving the general character projection formula. □

A.3 Complex Hoeffding Inequality

Theorem 3. *Let Z be a complex random variable. Now define the diameter of a complex variable as*

$$\mathrm{diam} Z = \inf\{c \in \mathbb{R} | P(|Z_1 - Z_2| > c) = 0\} \tag{A.29}$$

where Z_1, Z_2 are independent copies of Z. Let the complex random variable Z be bounded by $\text{diam}Z \leq d$. *Say we sample $Z_1, \ldots, Z_n$ then let $S_n = Z_1 + \ldots + Z_n$ with expectation value $\mathbb{E}[S_n]$. Hoeffding's inequality for complex random variables is then*

$$P(S_n - \mathbb{E}[S_n] \geq t) \leq \exp(-\frac{2t^2}{nd^2}) \tag{A.30}$$

Proof. From [25] Hoeffding's lemma for the case of a complex random variable is

$$\mathbb{E}[\exp(s(Z - \mathbb{E}[X])) \ \leq \exp(\frac{1}{8}s^2d^2) \tag{A.31}$$

From this the proof for Hoeffding's inequality follows from the case of a real random complex variable.

$$\begin{aligned} P(S_n - \mathbb{E}[S_n] \geq t) &= P(\exp(s(S_n - \mathbb{E}[S_n])) \geq \exp(st)) && \text{(A.32)} \\ &\leq \exp(-st)\mathbb{E}[\exp(s(S_n - \mathbb{E}[S_n]))] && \text{(A.33)} \\ &= \exp(-st)\prod_{i=1}^{n}\mathbb{E}[\exp(s(Z_i - \mathbb{E}[Z_i]))] && \text{(A.34)} \\ &\leq \exp(-st + \frac{1}{8}s^2nd^2) && \text{(A.35)} \end{aligned}$$

This upper bound is the best for the value of s minimizing the value inside the exponential. This value is found to be $s = \frac{4t}{nd^2}$ giving us Hoeffding's inequality. □

A.4 Telescoping series

Lemma 2. *For two arbitratry ordered lists of m elements $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ of an algebra with associative and distributive addition and multiplication we have,*

$$a_{m:1} - b_{m:1} = \sum_{j=1}^{m} a_{m:j+1}(a_j - b_j)b_{j-1:1} \tag{A.36}$$

where $a_{j:k}$ with $j \geq k$ is defined with respect to the list $\{a_1, \ldots, a_m\}$ as

$$a_{j:k} = a_j a_{j+1} \ldots a_{k-1} a_k. \tag{A.37}$$

Proof.

$$a_{m+1:1} - b_{m+1:1} = a_{m+1}a_{m:1} - a_{m+1}b_{m:1} + a_{m+1}b_{m:1} - b_{m+1}b_{m:1} \tag{A.38}$$

$$= a_{m+1}(a_{m:1} - b_{m:1}) + (a_{m+1} - b_{m+1})b_{m:1} \tag{A.39}$$

$$= \sum_{j=1}^{m+1} a_{m+1:j+1}(a_j - b_j)b_{j-1:1}. \tag{A.40}$$

□

One final note is that if all elements in both sets are identical then $a_{m:1} = a^m$ and $b_{m:1} = b^m$ and (A.36) becomes

$$a^m - b^m = \sum_{j=1}^{m} a^{m-j}(a-b)b^{j-1}. \tag{A.41}$$