# Extensions of Randomized Benchmarking

by

Marie Barnhill

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics

Waterloo, Ontario, Canada, 2015

© Marie Barnhill 2015

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

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

**Abstract**

The characterization of noisy quantum circuits is an important step in the development of large-scale quantum computers. As experimental quantum architectures approach the threshold for fault-tolerant quantum computing, obtaining a benchmark of process fidelities is needed to verify the integrity of the system and ensure functionality. Additionally, the ability to successfully implement a quantum code of choice may depend on the type of noise that is present in the system. It is therefore necessary to characterize not only the fidelity of the process, but also the structure of the noise.

Randomized benchmarking is a characterization protocol which extracts partial information from a noisy quantum process. This thesis gives a detailed review of the randomized benchmarking protocol, and presents the results of three collaborative projects which use techniques similar to randomized benchmarking to obtain previously hidden information. The first protocol is designed to obtain a benchmark on loss errors. This protocol also provides information on the detector efficiency and an additional constraint on estimated parameters from randomized benchmarking. The second protocol is an extension of the first, under the special case of coherent leakage errors, for which we obtain an estimate of the leakage rate. The third protocol provides a method of testing for spatial correlations in noisy processes. This test may be used to determine the correlation structure of errors, and the error correcting codes needed for successful experiments in a particular architecture. All protocols presented in this thesis are scalable, platform-independent and insensitive to state preparation and measurement errors. We also provide results of numerical tests of these protocols for simulated noise, which show their accuracy and robustness.

**Acknowledgements**

I would like to thank my supervisor, Joseph Emerson for taking me into his group of graduate students and providing guidance and invaluable feedback throughout my two years at IQC. My thanks also go to Joel Wallman for his willingness to answer my questions and include me in his research, and all members of the Emerson group for their insightful comments and their company in the office.

I would also like to that my partner, Taylor, for moving to Waterloo to be with me, and providing support and encouragement on a daily basis.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

The prospect of a robust universal quantum computer has generated considerable interest. The discovery of an algorithm by Shor [64] for efficiently factoring natural numbers with a quantum computer (an impossible task for a classical computer) has motivated rapid advance in quantum technologies. However, a roadblock to a universal quantum computer is the large number of possible error processes that arise naturally in a quantum system. Decoherence of the system, imperfect implementations of quantum gates, and coupling to the environment are examples of potential threats to a quantum computer. Fortunately, fault-tolerant quantum computation (FTQC) in combination with the threshold theorem [1, 42, 44], tells us that a certain amount of noise in the system can be accounted for and corrected, given we know the noise model, and are assured that the noise level is below a certain threshold. It is therefore vital that as experimental platforms improve and approach the error threshold we are able to characterize errors reliably. Because a quantum computer would require many - potentially thousands - of quantum bits (qubits), error characterization methods must be efficiently scalable in the dimension of the system, which poses a problem for traditional characterization tools.

This thesis expands on a robust and scalable error characterization called randomized benchmarking (RB). The standard RB protocol will be described, with both step-by-step instructions for its experimental implementation, and the mathematical justification for its success. Its merits and its limitations will also be explored, leading to ways to extend the RB protocols to benchmark error models not accounted for by standard RB. Changes to the RB protocol in order to flag and characterize three types of important, but often neglected, types of errors are presented and mathematically justified, along with numerical simulations showing their functionality in experimentally relevant conditions. A major advantage of the standard RB protocol and its variations is that they may be used to

benchmark any quantum information processing (QIP) device. Like RB, the protocols introduced in this thesis are platform-independent and they are described generally, for use by any quantum information processor.

The three contributions of this work are a result of the author's collaborations with Joel Wallman and Joseph Emerson. The material in Chapter 4 is based off the result for characterizing loss rates in [70], for which the author provided the material in section 4.6, the numerical simulation data and analysis in section 4.4, and contributed to the fit model derivations and the writing of the paper. The material in Chapter 5 is expanded as a special case of Chapter 4, and is based on the work in Ref. [69] for characterizing coherent leakage, for which the author participated in the description of the experimental protocol and derivation of fit models, provided the numerical data and analysis, and contributed to the writing of the paper. Note that what is called incoherent leakage in Ref. [69], is now termed loss, to remain consistent with standard terminology. Chapter 6 contains current progress on a joint collaboration for expanding and generalizing the simultaneous randomized benchmarking protocol in Ref. [30], for which the author participated in the model derivations, and contributed the numerical simulations and analysis. Chapters 2, 3 consist of appropriate background information on randomized benchmarking, and the author does not claim any originality of this work.

# Chapter 2

# Mathematical Preliminaries

This chapter will provide a brief overview of the notation, background, and mathematical tools used in this thesis. For a more complete overview of quantum information and representation theory see Refs. [58] and [32], respectively.

## 2.1 Quantum Information Basics

### 2.1.1 Quantum States

Quantum mechanical systems are represented mathematically as existing on complex Hilbert spaces. A complex Hilbert space, $\mathcal{H}$ is defined as a complex vector space, in combination with an inner product structure. Using Dirac notation, a vector on $\mathcal{H}$ is represented by a ket $|\psi\rangle$. The inner product of $|\psi\rangle$ with $|\phi\rangle$ is given by $\langle\phi|\psi\rangle$, where $\langle\phi|$ is a called a bra and is defined by the designated inner produce structure. The set of linear operators on $\mathcal{H}$, denoted $L(\mathcal{H})$, and the trace inner product

$$\langle\phi|\psi\rangle = \mathrm{Tr}(\psi^\dagger\phi) \tag{2.1}$$

define a Hilbert space. As per standard notation, the dagger operation $\psi^\dagger = \bar{\psi}^T$ is equal to the conjugate transpose of $\psi$. In the remainder of this thesis, all Hilbert spaces will be assumed to have the trace inner product structure. Any state vector can be written as a discrete linear combination of orthonormal (ON) basis elements of $\mathcal{H}$

$$|\psi\rangle = \sum_j c_j \, |\phi_j\rangle \,, \tag{2.2}$$

where $\{|\phi_j\rangle\}$ are the orthonormal basis elements. This thesis will deal with only finite-dimensional quantum systems, for which states are written as finite linear combinations of ON basis elements.

The state space for a single quantum bit (qubit) has basis $\{|0\rangle, |1\rangle\}$. Unlike classical bits, qubits can be found in a superposition of states $|0\rangle$ and $|1\rangle$, described by a linear combination as in Eq. (2.2).

States $\psi$ can also be represented by density operators,

$$\rho = \sum_j p_j |\psi_j\rangle \langle\psi_j|, \tag{2.3}$$

where $|\phi\rangle \langle\psi|$ denotes the outer product of $\psi$ and $\phi$, and $\sum_j p_j = 1$. States for which

$$\rho = \rho^2 \tag{2.4}$$

are called pure states. If a state is not a pure state it is called a mixed state.

### 2.1.2 Composite Systems

The full Hilbert space of a composite system composed of separable Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ is

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B \tag{2.5}$$

where $\otimes$ denotes the tensor product operation. The dimension of the composite system is $\dim(\mathcal{H}_{AB}) = \dim(\mathcal{H}_A) \times \dim(\mathcal{H}_B)$. $\mathcal{H}_{AB}$ is spanned by the independent set of states $\{|\psi_l\rangle \otimes |\phi_k\rangle\}, \forall l, k$ where $\{|\psi_l\rangle\}$ and $\{|\phi_k\rangle\}$ form orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively.

### 2.1.3 Evolution

A quantum computer does not consist of states alone. Implementing computations requires manipulating states, which is accomplished by applying linear transformations in the Hilbert space. Ideal transformations in quantum mechanics are given by unitary operators.

**Definition 2.1.1.** A unitary operator $U$ is a bounded linear operator on a Hilbert space that satisfies

$$UU^\dagger = U^\dagger U = \mathbb{I}. \tag{2.6}$$

Unitary operators represent evolution in a closed quantum system. Since by definition of a unitary $U^\dagger = U^{-1}$, i.e., $U$ is invertible, therefore all unitary operations are reversible.

Consider a general (not necessarily unitary) map on $\mathcal{H}$, $\mathcal{E} : L(\mathcal{H}) \to L(\mathcal{H})$. The following definitions provide useful classifications related to the properties of $\mathcal{E}$

**Definition 2.1.2.** Positivity: $\mathcal{E}$ is a positive map if $\mathcal{E}(\rho)$ is positive whenever $\rho$ is positive.

**Definition 2.1.3.** Complete positivity: Let $\mathcal{H}_{AB}$ be the composite system $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ for arbitrary $B$, where $\mathcal{E} : L(\mathcal{H}_A) \to L(\mathcal{H}_A)$ . $\mathcal{E}$ is completely positive if $\mathcal{E} \otimes \mathbb{I}_k$ is positive, where $\mathbb{I}_k$ is the identity operator acting on $\mathcal{H}_B$

In Def. 2.1.3, the system is extended to ensure that the output state of any composite system satisfies the positivity condition. Positive, but not completely positive, maps do not guarantee that the transformation of a positive state on a composite system will be positive.

**Definition 2.1.4.** Trace preserving: $\mathcal{E}$ is trace-preserving if $\mathrm{Tr}(\mathcal{E}(\rho)) = \mathrm{Tr}(\rho)$

A map that is both trace preserving and completely positive (CPTP) is called a *quantum channel*. CPTP maps are typically considered to represent the most general type of quantum evolution. In experimental quantum computation, ideal transformations are usually unitary, and noisy transformations are CPTP maps. However, we will explore in Chapter 4 the case when noisy implementations are non-trace-preserving

### 2.1.4 Quantum Circuits

To perform a computation, a quantum computer applies sequences of (unitary) transformations to prepared states. The choice of sequence is determined by the objective of the computation, and the individual unitary transformations can be referred to equivalently as *gates*, similarly to a classical computer. (Note that in some cases, i.e., noisy computations, open quantum systems, gates may also refer to CPTP maps.) In this section we will give definitions of several often-used quantum gates, and give an example of a graphical representation of quantum circuits.

The Pauli matrices are the set of quantum gates defined as the generators of rotation on a qubit, and are given by

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.7}$$

in the $\{|0\rangle, |1\rangle\}$ basis. Note that $\mathbb{I}$ is the identity operator, $X$ corresponds to a bit flip, $Z$ corresponds to a phase flip, and $Y$ corresponds to a both a bit and phase flip, since $Y = iXZ$. The Pauli matrices form the Pauli group $\mathcal{P} = \{\mathbb{I}, X, Y, Z\}$.

Other important gates are the Hadamard gate ($H$), the phase gate ($P$), and the controlled-NOT (C-NOT) gate.

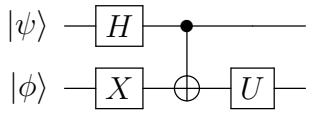
$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \tag{2.8}$$

The C-NOT gate is a unitary transformation on two qubits. The C-NOT gate may be called an entangling gate because it works by flipping the state of the second qubit, dependent on the state of the first qubit. In the two qubit basis $\{|0\rangle, |1\rangle\} \otimes \{|0\rangle, |1\rangle\} = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, C-NOT is written

$$\text{C-NOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{2.9}$$

The C-NOT gate performs a bit flip ($X$ gate) on the second (target) qubit if the first (control) qubit is in the $|1\rangle$ state. If the first qubit is the $|0\rangle$ state, the C-NOT, applies the identity to the second qubit (the second qubit is unaffected).

Sequences of quantum gates can be represented by circuit diagrams, which keep track of to which qubit(s) each gate is applied, and the order with which gates are applied. The example in Fig. 2.1 shows two qubits, initally prepared in states $|\psi\rangle$ and $|\phi\rangle$. The applied gates are: a Hadamard gate on the first qubit and a Pauli-$X$ gate on the second qubit, followed by a C-NOT gate with control on qubit one, and target on qubit two. The final step is an unspecified unitary on the second qubit, while the first is unmodified.




(2.10)

Figure 2.1: Example of a circuit diagram.

### 2.1.5 Measurements

Ideal measurement procedures of quantum systems are represented by observables, or self-adjoint operators. Like a density operator, a self-adjoint operator with a set of eigenvalues

$\{a_m\}$ admits a spectral decomposition of orthogonal projection operators

$$A = \sum_m a_m \left|a_m\right\rangle \left\langle a_m\right| . \tag{2.11}$$

The set of physically observable outcomes is given by the eigenvalues of $A$. The probability of measuring outcome $a_m$, given a preparation $\rho$ is $\mathrm{Tr}(\left|a_m\right\rangle \left\langle a_m\right| \rho)$, and this result is known as the Born rule. After measurement, the state is projected onto the state of $\left|a_m\right\rangle \left\langle a_m\right|$ associated with $a_m$ by the von Neumann projection postulate. Measurements of this type are called projection valued measures (PVMs).

A more general type of measurement is called a positive-operator valued measure (POVM). In the discrete case, which is the important case for this thesis, a measurement is comprised of a set of linear operators $\{M_m\}$ acting on $\mathcal{H}$ that must satisfy

$$\sum_m M_m^\dagger M_m = \mathbb{I}. \tag{2.12}$$

measurements are indexed by outcomes $m$. If the state is initially prepared in $\rho$, then the probability of obtaining outcome $m$ is

$$\Pr(m) = \mathrm{Tr}(M_m \rho M_m^\dagger) \tag{2.13}$$

where $\sum_m \Pr(m) = 1$. After the state is measured, it collapses to

$$\frac{M_m \rho M_m^\dagger}{\mathrm{Tr}(M_m \rho M_m^\dagger)}. \tag{2.14}$$

By convention, POVM elements are defined as

$$Q_m = M_m^\dagger M_m \tag{2.15}$$

and the set $\{Q_m\}$ is the POVM.

## 2.2 Representations of quantum channels

Here we present three useful representations of quantum channels, adapted from their descriptions in Refs. [22, 37].

### 2.2.1 Kraus Representation

The evolution of a quantum system for a given time is described by a superoperator, $\mathcal{E}$, acting on the system preparation $\rho$. Consider the construction of an operation where a preparation is subjected to a transformation, given by the unitary $U_k$, with probability $p_k$. The total transformation is then given by the convex combination

$$\mathcal{E}(\rho) = \sum_k p_k U_k \rho U_k^\dagger, \tag{2.16}$$

which is not necessarily a unitary transformation. If we define the operator $A_k = \sqrt{p_k} U_k$, then Eq. (2.16) be written

$$\mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger, \tag{2.17}$$

and the $A_k$ are called the *Kraus operators*. The constraint

$$\sum_k A_k^\dagger A_k = \mathbb{I} \tag{2.18}$$

ensures $\mathcal{E}$ is trace-preserving. By the Kraus representation theorem, a superoperator $\mathcal{E}$ is a CPTP map if and only if there is an operator sum decomposition of $\mathcal{E}$ subject to constraints (2.16) and (2.18) [22]. Therefore all CPTP maps can be represented by Kraus operators. If $\mathcal{E}$ is a CPTP map, it is also a *unital* map if the condition

$$\sum_k A_k A_k^\dagger = \mathbb{I} \tag{2.19}$$

on the Kraus operators holds.

**Definition 2.2.1.** Unital maps: $\mathcal{E}$ is unital if it is a CPTP map that maps the identity operator to the identity operator.

A CPTP map on a $d$-dimensional space can be represented by at most $d^2$ Kraus operators. The Kraus representation theorem also requires the assumption that the system and the environment are initially uncorrelated. Failure of this condition can result in non-completely postive maps [59].

### 2.2.2 $\chi$-matrix

The $\chi$-matrix is an object which fully represents a quantum channel, $\mathcal{E}$. It is defined relative to a trace-orthonormal operator basis $\mathcal{B} = \{B_1, \ldots, B_{d^2}\}$ for the operator space $\mathcal{H}_{d^2}$ (i.e., $\mathrm{Tr}B_i^\dagger B_j = \delta_{i,j}$), which is often chosen to be the Pauli operators, and it can be calculated by expanding the Kraus operators in terms of $\mathcal{B}$. Given $\mathcal{B}$ is an orthonormal basis for operators on dimension $d$, any Kraus operator $A_k$ of $\mathcal{E}$ can be written

$$A_k = \sum_\alpha B_\alpha \mathrm{Tr}(B_\alpha^\dagger A_k) \tag{2.20}$$

then the action of $\mathcal{E}$ is equivalent to

$$\mathcal{E}(\rho) = \sum_{i,j} \sum_k \alpha_{ik} \alpha_{jk}^* B_i \rho B_j^\dagger \tag{2.21}$$

where $\alpha_{ik} = \mathrm{Tr}(B_i^\dagger A_k)$. The $\chi$-matrix is then defined for a channel $\mathcal{E}$ by

$$\mathcal{E}(\rho) = \sum_{i,j} \chi_{i,j} B_i \rho B_j^\dagger \tag{2.22}$$

and has matrix elements

$$\chi_{ij} = \sum_k \alpha_{ik} \alpha_{jk}^*. \tag{2.23}$$

The $\chi$-matrix is independent of the choice of Kraus operators for a given $\mathcal{E}$, and is a Hermitian matrix by definition. It is positive semi-definite with $\mathrm{Tr}(\chi) = d$ if and only if $\mathcal{E}$ CPTP map [22].

### 2.2.3 Liouville Representation

In the remainder of this thesis we will often work in the Liouville (sometimes called super-operator) representation of quantum channels. The Liouville representation is also defined relative to a trace-orthonormal operator basis $\mathcal{B} = \{B_1, \ldots, B_{d^2}\}$ for the operator space $\mathcal{H}_{d^2}$. We refer to the special case where $\mathcal{B}$ is the (normalized) set of Pauli operators as the Pauli-Liouville representation.

Density operators $\rho$ and measurement outcomes $M$ are represented by column and row vectors $|\rho)$ and $(M|$ whose $i$th elements are $\mathrm{Tr}(B_i^\dagger \rho)$ and $\mathrm{Tr}(M^\dagger B_i)$ respectively. The Born

rule can then be expressed as $\mathrm{Tr}M\rho = (M|\rho)$. Completely positive maps $\mathcal{E} : \mathcal{H}_{d^2} \to \mathcal{H}_{d^2}$ are represented by matrices $\boldsymbol{\mathcal{E}}$ such that

$$\boldsymbol{\mathcal{E}}_{ij} = \mathrm{Tr}[B_i^\dagger \mathcal{E}(B_j)], \tag{2.24}$$

where $\boldsymbol{\mathcal{E}}|\rho) = |\mathcal{E}[\rho])$ for all $\rho$. Note that bolded notation is used to denote the Liouville representation of the channel, and round-bracket Dirac notation is used to denote the Liouville representation of quantum states and measurements.

One advantage of using the Liouville representation is that channels compose via matrix multiplication. This allows probabilities for long sequences of channels to be easily calculated.

Given the choice of $B_1 = d^{-1/2}\mathcal{I}_d$, it is convenient to express $\boldsymbol{\mathcal{E}}$ in block-matrix form as

$$\boldsymbol{\mathcal{E}} = \begin{pmatrix} S(\mathcal{E}) & \vec{m}(\mathcal{E}) \\ \vec{n}(\mathcal{E}) & \mathbf{R}(\mathcal{E}) \end{pmatrix}, \tag{2.25}$$

where $\mathbf{R}$ is a $d^2 - 1 \times d^2 - 1$ matrix. We will sometimes omit the argument when it is clear from the context. For trace-preserving (TP) maps, $S(\mathcal{E}) = 1$ and $\vec{m} = \vec{0}$, while for unital maps, $S(\mathcal{E}) = 1$ and $\vec{m} = \vec{n} = 0$ [71]. We show this using the more well-known Kraus operator representation.

*Proof.* We start by proving the trace-preserving condition on $S(\mathcal{E})$. Let $\mathcal{E}$ be a quantum channel with operator sum decomposition $\mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger$. Then in the Pauli-Liouville representation

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{ij} &= \mathrm{Tr}[B_i^\dagger \mathcal{E}(B_j)] \\ &= \mathrm{Tr}\left( B_i^\dagger \sum_k A_k B_j A_k^\dagger \right). \end{aligned} \tag{2.26}$$

By the linearity and cyclic properties of the trace we find:

$$\boldsymbol{\mathcal{E}}_{ij} = \mathrm{Tr}\left( B_j \sum_k A_k^\dagger B_i^\dagger A_k \right). \tag{2.27}$$

Since $S(\mathcal{E}) = \boldsymbol{\mathcal{E}}_{11}$, the choice of $B_1 = d^{-1/2}\mathcal{I}_d$ implies

$$\begin{aligned} S(\mathcal{E}) &= \mathrm{Tr}\left( B_1 \sum_k A_k^\dagger B_1^\dagger A_k \right) \\ &= \mathrm{Tr}\left( \frac{1}{d} \sum_k A_k^\dagger A_k \right) \end{aligned} \tag{2.28}$$

and by the trace-preserving condition on the Kraus operators in Eq. (2.18)

$$S(\mathcal{E}) = \mathrm{Tr}\left(\frac{1}{d}\mathbb{I}\right) = 1. \tag{2.29}$$

All unital maps are trace-preserving, so $S(\mathcal{E}) = 1$ for unital maps as well, although this is also easy to show using the unital condition in Eq (2.19). By recalling the trace-orthonormal condition on the operator basis, $\mathrm{Tr}B_i^\dagger B_j = \delta_{i,j}$, a similar calculation can be performed to show that $\vec{m} = \vec{0}$ for trace-preserving maps, and $\vec{n} = \vec{0}$ for unital maps. □

We therefore refer to $\vec{n}$ and $\mathbf{R}$ as the non-unital and unital components respectively.

## 2.3 Unitary $t$-designs

A unitary $t$-design is a finite set of unitary operations with special properties.

**Definition 2.3.1.** A unitary $t$-design for $d$ dimensions is a finite set $\{U_1, \ldots U_K\} \subseteq U(D)$ of unitary matrices on $\mathbb{C}^d$ such that for every homogeneous complex valued polynomial $p_{(t,t)}(U)$ of degree at most $t$ in both the matrix elements of $U$, and the complex conjugates of those elements,

$$\frac{1}{K}\sum_{j=1}^{K} p_{(t,t)}(U_j) = \int_{U(d)} \mathrm{d}U p_{(t,t)}(U) \tag{2.30}$$

where the integral is taken over the Haar measure, $U(d)$ [17, 21].

By this definition it is easy to see every $t$-design is a $(t-1)$-design. Examples of 2- and 1-designs are the Clifford group (see section 2.4) and the Pauli group (see section 2.1.4) respectively.

### 2.3.1 2-design Twirling

The twirl of a quantum channel $\mathcal{E}$ is defined as the average of $\mathcal{E}$ under the composition $\mathcal{U}^\dagger \circ \mathcal{E} \circ \mathcal{U}$ where $\circ$ denotes composition (i.e. apply $\mathcal{E}$ then apply $\mathcal{U}$), and $\mathcal{U}$ is a unitary

operation, $\mathcal{U}(\rho) = U\rho U^\dagger$. The averaged channel is

$$\begin{aligned}\overline{\mathcal{E}}(\rho) &= \int_{U(d)} \mathcal{U}^\dagger \circ \mathcal{E} \circ \mathcal{U}(\rho)\ \mathrm{d}\mu(\mathcal{U}) \\ &= \int_{U(d)} U\mathcal{E}(U^\dagger \rho U)U^\dagger\ \mathrm{d}\mu(U). \end{aligned} \tag{2.31}$$

where $\mu(U)$ is a probability distribution for $\mathcal{U}$. Of particular interest is the case where the measure $\mu(U) = \mu_H(U)$ on $U(d)$ is the unique bi-invariant normalized Haar measure. In this case

$$\begin{aligned}\overline{\mathcal{E}}(\rho) &= \int_{U(d)} U\mathcal{E}(U^\dagger \rho U)U^\dagger\ \mathrm{d}\mu_H(U) \\ &= p\rho + (1-p)\frac{\mathbb{I}}{d}. \end{aligned} \tag{2.32}$$

That is, the twirl of $\mathcal{E}$ over the Haar measure is a one-parameter depolarizing channel [23]. This result may be used to extract some important parameters which describe a quantum channel. As we will show later (and as is shown in various randomized benchmarking papers), the depolarizing parameter $p$ is directly related to the channel fidelity of $\mathcal{E}$.

An important feature of 2-designs arises by considering the set of unitaries $\mathcal{U}_j \in \mathcal{U}$ to be a discrete set. The discrete twirl over $\mathcal{U}$ is given by

$$\begin{aligned}\mathcal{E}^{\mathcal{U}}(\rho) &= \sum_j \mu_j \mathcal{U}_j^\dagger \circ \mathcal{E} \circ \mathcal{U}_j(\rho) \\ &= \sum_j \mu_j (U_j^\dagger \mathcal{E}(U_j \rho U_j^\dagger) U_j^\dagger) \end{aligned} \tag{2.33}$$

where $\{\mu_j\}$ is a probability distribution over the $\mathcal{U}_j$, and we denote the discrete twirled channel by $\mathcal{E}^{\mathcal{U}}$. The twirling condition for 2-designs gives an alternate definition of 2-designs.

**Definition 2.3.2.** The set of unitaries $\{U_j\}$ forms a unitary 2-design if and only if, the twirl of $\mathcal{E}$ over $\{U_j\}$ is equal to the full Haar twirl of $\mathcal{E}$, for any channel $\mathcal{E}$ and any state $\rho$ [8, 51].

## 2.4 The Clifford Group

**Definition 2.4.1.** A group $\mathcal{G}$ is defined as an algebraic structure consisting of a set of elements $\{g_i\}$, and a group operation " $\cdot$ " obeying four axioms:

1. Closure: $g_1 \cdot g_2 = g_3$, $g_3 \in \mathcal{G}$.
2. Associativity: $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$
3. Identity: $\exists I \in \mathcal{G}$ such that $I \cdot g = g, \forall g \in \mathcal{G}$.
4. Inversibility: $\forall g \in \mathcal{G},\ \exists g^{-1} \in \mathcal{G}$ such that $g \cdot g^{-1} = g^{-1} \cdot g = I$.

There are two groups of particular importance for this thesis. First is the Pauli group, and second is the Clifford group, described in this section.

The $n$-qubit Clifford group $\mathcal{C}_n$ is a set of unitary operators, defined as the normalizer of the Pauli group (up to overall phases), $\mathcal{P}_n = \{\mathbb{I}, X, Y, Z\}^{\otimes n} \times \{\pm 1, \pm i\}$.

**Definition 2.4.2.** A unitary operator $C$ of dimension $d = 2^n$ is a member of $\mathcal{C}_n$ if and only if $CPC^\dagger \in \mathcal{P}_n$ for all $P \in \mathcal{P}_n$.

The Clifford group for $n = 1$ can be fully generated by the hadamard and phase gate, defined in Eq. (2.8). For $n \geq 2$, the generating set of $\mathcal{C}_n$ requires the addition of the C-NOT gate. For arbitrary $n$, $\mathcal{C}_n$ is generated by tensor products of the three generating gates.

The Clifford group alone does not provide any additional "quantum power", it can be efficiently simulated on a classical computer as proved by the Gottesmann-Knill theorem [34, 45]. To achieve universal quantum computation, additional operations outside the Clifford group are required. However, only a single gate outside the Clifford group need be added for universal quantum computation. Examples of gates outside $\mathcal{C}_n$ are the $\pi/8$ rotation gate, and the Toffoli gate. Additionally, a technique called magic state distillation, in combination with the Clifford group, is sufficient for universal quantum computation [43, 61].

**Theorem 1.** *The uniform distribution over the Clifford group on n qubits is a unitary 2-design.*

The twirling condition for 2-designs therefore applies to the Clifford group, and the $\mathcal{C}_n$-twirl is therefore a depolarizing channel. There are implementations of robust quantum computing that have a close relationship with the Clifford group, which makes it an interesting and important group in the study of quantum information.

## 2.5 Schur's Lemma

To understand Schur's Lemma we define two key concepts: irreducible representations and intertwining operators. Proof of Schur's lemma requires supplemental background information in group theory. We omit the proof here because our work uses mainly consequences of Schur's lemma, but for the interest of the reader it can be found in Appendix D or Ref. [65].

A representation (rep) of a group $\mathcal{G}$ on a vector space $V$ is a homomorphism $\phi : \mathcal{G} \to \mathrm{GL(V)}$, where GL(V) is the group of automorphisms of $V$. Given finite dimension $d$, $\mathrm{GL(V)} \equiv \mathrm{GL(d, \mathbb{C})}$ is the group of invertible $d \times d$ matrices. A map $\phi : \mathcal{G} \to \mathcal{F}$ between groups is a homomorphism if it preserves group structure. This implies that $\phi(I_{\mathcal{G}}) = I_{\mathcal{F}}$ and $\phi(g^{-1}) = \phi(g)^{-1}$. A subrepresentation (subrep) $\phi_W$ of $\phi$ is the restriction of $\phi$ to a subspace $W \subseteq V$ that is invariant under the action of $\mathcal{G}$, that is, $\phi(g)W \subseteq W$ for all $g \in \mathcal{G}$.

**Definition 2.5.1.** An irreducible representation (irrep) is a representation $\phi$ whose only subreps are $\phi_0$ (i.e., the null rep $\phi(g) = 0$ for all $g \in \mathcal{G}$) and $\phi$ itself [29].

Since the reps we consider are unitary reps of compact groups, if $W$ is a subrep of $V$ then the orthogonal complement $W^{\perp}$ is a subrep as well. Therefore any rep can be decomposed into a direct sum of irreps, which may occur with some multiplicity. Any basis that decomposes a rep into a direct sum of irreps is called a Schur basis. The simplest rep is the trivial rep $\phi(g) = 1$ (which is also an irrep) for all $\mathcal{G}$, which is defined for any group $\mathcal{G}$.

**Definition 2.5.2.** A map $T : V \to W$ is an intertwining operator if for reps $\phi : \mathcal{G} \to GL(V)$ and $\phi' : \mathcal{G} \to GL(W)$, $T(\phi(g)) = \phi'(T(g))$.

The set of all intertwining operators forms a complex vector space.

Now that we have defined irreps and intertwining operators, we give the second form of Schur's lemma, which is the form most applicable to this work.

**Theorem 2.** *Schur's Lemma (of the second form): Let $\phi : \mathcal{G} \to \mathrm{GL(V)}$ be an irrep, and $T : V \to V$ be a intertwining operator for $\phi$, where $V = \mathbb{C}^n$. Then $T = \lambda \mathbb{I}$ for some $\lambda$.*

Combining knowledge of twirls with Schur's lemma provides an alternate intuition for twirling. Recall that the twirl of $A$ is denoted by $A^{\mathcal{G}}$. Let $\phi$ be a nontrivial $d$-dimensional irreducible representation of a group $\mathcal{G}$ and $A \in \mathrm{GL(d, \mathbb{C})}, \mathrm{B} \in \mathrm{GL(d+1, \mathbb{C})}$. Then all three of the following statements follow directly from Schur's Lemma [32].

- $A^{\mathcal{G}} = (\mathrm{Tr}A/d)\mathbb{I}_d$;
- $B^{\mathcal{G}} = B_{11} \oplus b\mathbb{I}_d$ where the representation of $\mathcal{G}$ is $1 \oplus \phi$ and $b = (\mathrm{Tr}B - B_{11})/d$; and
- $\sum_{g\in\mathcal{G}} \phi(g) = 0$.

# Chapter 3

# Randomized Benchmarking

## 3.1 The Quantum Error Characterization Problem

The accuracy threshold theorem promises that for error rates below a certain threshold value, quantum computation is possible in the presence of errors. Completely characterizing the noise affecting a quantum system (using quantum process tomography and others) is not scalable in the number of qubits and is therefore infeasible for larger systems [60]. In response to this difficulty, methods of scalable quantum gate characterization have emerged, which have the cost of obtaining only partial information about the system in question. However, many of these characterization methods make assumptions about the noise and available resources, such as: negligible or no state preparation and measurement (SPAM) errors, and noisy processes are close to ideal operations. In recent years a toolkit of characterization protocols called randomized benchmarking has been developed to reliably solve these issues [50, 51]. RB methods are scalable and robust characterization protocols which are insensitive to SPAM errors, and have been used with success in various experimental implementations such as NMR, superconducting, and ion trap qubits [12, 15, 62].

When considering time- and gate-independent, Markovian noise, it is standard and convenient to represent the ideal operation by a unitary, $\mathcal{U}$, and the implemented gate by quantum operation, $\mathcal{E}_{\mathcal{U}}$. The error is then defined as

$$\mathcal{E} = \mathcal{U}^{\dagger} \circ \mathcal{E}_{\mathcal{U}}, \tag{3.1}$$

where $\mathcal{E}$ can be either the error on the gate, or the noise due to free evolution of the system. Equivalently, $\mathcal{E}$ can be written as $\mathcal{E}'_{\mathcal{U}} \circ \mathcal{U}^{\dagger}$, where in general $\mathcal{E}'_{\mathcal{U}} \neq \mathcal{E}_{\mathcal{U}}$.

In this thesis we will focus on gate- and time-independent errors, and furthermore, we will make the assumption of Markovian errors, unless otherwise stated. If such errors are of interest to the reader, approaches which give partial RB results for gate- and time-dependent noise, and a model of non-Markovian noise can be found in Refs. [4, 51, 71]. These approaches can be applied to all of the RB-variant protocols outlined in this thesis, but we note that full (scalable) characterization of non-Markovian noise remains a difficult open problem.

## 3.2 Average Fidelity

One often used measure for distinguishing how alike two CPTP maps $\mathcal{E}_1$ and $\mathcal{E}_2$ are to one another is the channel fidelity,

$$\mathcal{F}_{\mathcal{E}_1,\mathcal{E}_2}(\rho) = F\left(\mathcal{E}_1(\rho), \mathcal{E}_2(\rho)\right) \tag{3.2}$$

$$= \left(tr\sqrt{\sqrt{\mathcal{E}_1(\rho)}\mathcal{E}_2(\rho)\sqrt{\mathcal{E}_1(\rho)}}\right)^2. \tag{3.3}$$

Here $F$ is the usual fidelity between quantum states as in Nielsen and Chuang [58]. Often it is useful to compare an ideal implementation of a gate with its actual, noisy implementation. Consider the ideal operation as being a unitary, $\mathcal{U}$, and the actual gate a quantum operation, $\mathcal{E}_\mathcal{U}$, as in Eq. (3.1). The fidelity of $\mathcal{E}_\mathcal{U}$ to $\mathcal{U}$ is then equal to the fidelity of $\mathcal{E}$ to $\mathcal{I}$,

$$\mathcal{F}_{\mathcal{E}_\mathcal{U},\mathcal{U}}(\phi) = \mathcal{F}_{\mathcal{E},\mathcal{I}}(\phi) = tr\left(|\phi\rangle\langle\phi|\mathcal{E}(|\phi\rangle\langle\phi|)\right) \tag{3.4}$$

for $\phi \leftrightarrow |\phi\rangle\langle\phi| \in \mathbb{CP}^{d-1}$, a pure state in the complex projective space. The expression for the average gate fidelity (over states) is obtained by integrating $\mathcal{F}_{\mathcal{E}_\mathcal{U},\mathcal{U}}$ over $\mathbb{CP}^{d-1}$ using the unique unitarily invariant measure on pure states, the Fubini-Study measure $\mu_{FS}$ [7],

$$\overline{\mathcal{F}_{\mathcal{E}_\mathcal{U},\mathcal{U}}} = \overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}} = \int_{\mathbb{CP}^{d-1}} tr\left(|\phi\rangle\langle\phi|\mathcal{E}(|\phi\rangle\langle\phi|)\right) \mathrm{d}\mu_{FS}(\phi). \tag{3.5}$$

The average fidelity of an operation, or group of operations gives a good idea of the performance of this group in practice. The basic idea behind RB is to partially characterize a set of noisy gates acting on an arbitrary number of qubits. In the case of standard RB, information related to the average gate fidelity is obtained. The protocol in section 3.3 is the standard RB protocol for benchmarking a unitary 2-design $\mathcal{G} \subseteq U(d)$, but variations of the protocol have been proposed to characterize individual and non-2-design gates [14, 52], as well as other interesting parameters besides the average gate fidelity, which will be the contribution of this thesis.

## 3.3 Experimental Protocol

To implement standard randomized benchmarking, a unitary 2-design $\mathcal{G}$ with implementable elements $g \in \mathcal{G}$ must be designated. Noisy implementations of the $g$ are given by $\mathcal{E}_g = g \circ \mathcal{E}$ for some fixed $\mathcal{E}$. Ideally, $\mathcal{G}$ should be a group of interest, and an experimentalist should want to characterize the accuracy of its implementation by their apparatus. Often $\mathcal{G}$ is chosen as the Clifford group, because of its applications in FTQC, and it conveniently is a 2-design. The RB protocol for a chosen $\mathcal{G}$ is as follows:

1. Choose a sequence length $m \in \mathbb{N}$.
2. Choose a sequence $\mathbf{k} = (k_1, \ldots, k_m)$ of $m$ integers uniformly at random where $k_j \in \{1, \ldots, |\mathcal{G}|\}$.
3. Prepare an initial $d$-dimensional system in some state $\rho$ (usually $\rho \approx |0\rangle\langle 0|$).
4. Apply the sequence of operations $g_{k_m} \circ \ldots \circ g_{k_0}$, where $g_{k_0} = \prod_{i=1}^{m} g_{k_i}^{-1}$, is the inversion operator of the sequence.
5. Perform a POVM $\{Q, \mathcal{I} - Q\}$.
6. Repeat steps $3 - 5$ enough times to estimate
$$Q_{\mathbf{k}} = \mathrm{Tr}[Q g_{k_m} \circ \mathcal{E} \circ \ldots \circ g_{k_0} \circ \mathcal{E}(\rho)], \tag{3.6}$$
the probability of the outcome $Q$ occurring, to a desired precision.
7. Repeat steps $2 - 6$ enough times to estimate the expectation value of $Q_{\mathbf{k}}$:
$$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = |\mathcal{G}|^{-m} \sum_{\mathbf{k}} Q_{\mathbf{k}} \tag{3.7}$$
to a desired precision.
8. Repeat steps $1 - 7$ for various choices of $m$, and fit the results to the exponential decay curve
$$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = A(\mathcal{E}) p^m + B(\mathcal{E}) \tag{3.8}$$
to obtain partial information on the noise affecting $\mathcal{G}$. A diagram of a single trial of the experimental protocol is illustrated in Fig. 3.1.

## 3.4 Derivations of the Fit Models

The full derivation of the fit model can be found in [50]. We will show here that the decay parameter $p$ in Eq.(3.8) is related to the average gate fidelity, and state preparation and measurement errors are absorbed by the constants $A(\mathcal{E})$ and $B(\mathcal{E})$.

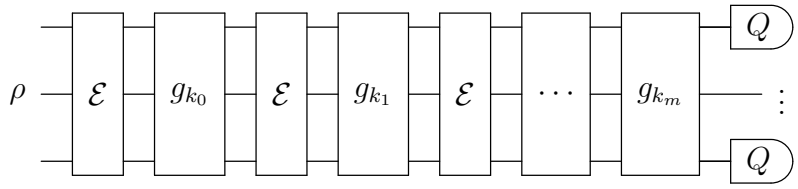



Figure 3.1: Circuit diagram of the randomized benchmarking procedure.

The choice of $g_{k_0}$ in step 3 can be calculated using a classical computer according to the Gottesmann-Knill theorem [34, 45]. This specifically-selected gate allows that each group gate can be thought of as the composition $g_{k_i} = c_i \circ c_{i-1}^\dagger$ with $g_{k_m} = c_m^\dagger$, and $c_i \in \mathcal{G}$ by the properties of groups. By considering this rearrangement, we see that randomized benchmarking produces a sequence of conjugations, $c_j^\dagger \circ \mathcal{E} \circ c_j$, as shown in Fig. 3.2.

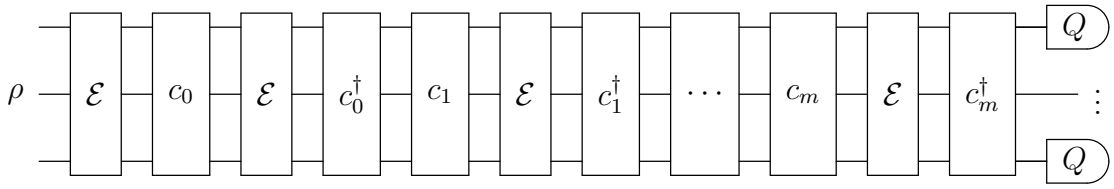



Figure 3.2: Circuit diagram of the twirling "trick" in standard RB.

Randomized benchmarking protocols make use of the fact that randomly sampling channels $g$ from the Clifford group (or any unitary 2-design), and averaging over a number of selections of these channels, closely approximates the full discrete $\mathcal{G}$-twirl of $\mathcal{E}$ [17], denoted by $\mathcal{E}^{\mathcal{G}}$. The discrete $\mathcal{G}$-twirl of $\mathcal{E}$ is given by

$$\mathcal{E}^{\mathcal{G}} = \frac{1}{|\mathcal{G}|} \sum_{j=1}^{|\mathcal{G}|} c_j^\dagger \circ \mathcal{E} \circ c_j. \tag{3.9}$$

Because $\mathcal{G}$ is a unitary 2-design, the discrete twirl is equal to the twirl over the the full Haar measure by the twirling condition for 2-designs. The average fidelity is invariant

under the twirling operation, therefore the twirled channel, $\mathcal{E}^{\mathcal{G}}$ is a depolarizing channel, $\mathcal{D}_p$, with average fidelity identical to that of $\mathcal{E}$ [23],

$$\mathcal{E}^{\mathcal{G}}(\rho) = p\rho + \frac{1-p}{d}\mathcal{I}_d \tag{3.10}$$

for dimension $d$. The decay parameter $p$ in Eq. (3.8) and Eq. (3.10) is related to the average gate fidelity of the group $\mathcal{G}$ in the following way

$$\overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}} = p + \frac{(1-p)}{d}. \tag{3.11}$$

The sequence of twirls in a length-$m$ randomized benchmarking sequence is the $m$-fold composition of $\mathcal{E}^{\mathcal{G}}$ with itself,

$$(\mathcal{E}^{\mathcal{G}})^m(\rho) = p^m\rho + \frac{1-p^m}{d}\mathcal{I}_d, \tag{3.12}$$

therefore, the sequence fidelity decays exponentially with the length of the sequence, giving us the exponential model in Eq. (3.8). To derive this, note that after averaging over sequences

$$\begin{aligned}\mathbb{E}_{\mathbf{k}}Q_{\mathbf{K}} &= \mathrm{Tr}[Q\mathcal{E}^{\mathcal{G}} \circ \ldots \circ \mathcal{E}^{\mathcal{G}} \circ \mathcal{E}(\rho)] \\ &= \mathrm{Tr}[Q(\mathcal{E}^{\mathcal{G}})^m \circ \mathcal{E}(\rho)].\end{aligned} \tag{3.13}$$

Replacing $(\mathcal{E}^{\mathcal{G}})^m$ with the expression in Eq. (3.12) gives Eq. (3.8), with $A(\mathcal{E}) = \mathrm{Tr}[Q(\rho - \frac{\mathbb{I}_d}{d})\mathcal{E}(\rho)]$ and $B(\mathcal{E}) = \frac{1}{d}\mathrm{Tr}[Q\mathcal{E}(\rho)]$.

We may also consider the average error rate (or infidelity) of $\mathcal{G}$, which is directly related to the average fidelity. We define the average error rate $r$ for the group $\mathcal{G}$ as

$$\begin{aligned}r &= 1 - \overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}} \\ &= \frac{(d-1)(1-p)}{d}.\end{aligned} \tag{3.14}$$

## 3.5 RB in the Pauli-Liouville Representation

The Liouville picture provides an equivalent, intuitive view of randomized benchmarking. Any quantum channel $\mathcal{E}$, i.e. a completely positive trace-preserving map, can be written

$$\boldsymbol{\mathcal{E}} = \begin{pmatrix} 1 & 0 \\ \vec{n}(\mathcal{E}) & \mathbf{R}(\mathcal{E}) \end{pmatrix} \tag{3.15}$$

where $\boldsymbol{\mathcal{E}}$ is a $d^2 \times d^2$ square matrix, and represents the Liouville representation of $\mathcal{E}$. The $d^2 - 1$ dimensional column vector $\vec{n}$ is the non-unital block and the $(d^2 - 1) \times (d^2 - 1)$ matrix $\mathbf{R}$ is the unital block. Given that $\mathcal{G}$ is a 2-design, the $\mathcal{G}$-twirl in this representation, by Schur's lemma is

$$\boldsymbol{\mathcal{E}}^{\mathcal{G}} = \begin{pmatrix} 1 & 0 \\ 0 & p\mathbb{I}_{d^2-1} \end{pmatrix} \tag{3.16}$$

where $p$ is the decay parameter in Eq. (3.8), and

$$p = \frac{\mathrm{tr}(\mathbf{R}(\mathcal{E}))}{d^2 - 1}. \tag{3.17}$$

For sequences of length $m$, it is obvious that $(\boldsymbol{\mathcal{E}}^{\mathcal{G}})^m$ is a single-parameter diagonal matrix of the same block-structure as $\boldsymbol{\mathcal{E}}^{\mathcal{G}}$, with parameter $p^m$. Clearly the sequence of operations produces the decay curve in Eq. 3.8, with constants $A$ and $B$ depending on the preparation and measurement.

## 3.6 Relation to the $\chi$-matrix

The $\chi$ matrix, as described in section 2.2.2, is an important object in experimental quantum computation, and relates to methods in process tomography. It may therefore be useful to show the particular information given by RB on the $\chi$-matrix. Calculating the average fidelity of $\mathcal{E}$, using the representation in Eq. (2.22) gives

$$\overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}} = \frac{\chi_{0,0} d + 1}{d + 1} \tag{3.18}$$

rearranging, we have

$$\chi_{0,0} = \frac{\overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}}(d+1) - 1}{d} = p\left(1 - \frac{1}{d^2}\right) + \frac{1}{d^2}. \tag{3.19}$$

Therefore, by estimating the parameter $p$ in RB, one can directly estimate the $(0,0)$ entry of the $\chi$-matrix. $\chi_{0,0}$ is also the probability that the identity operation occurs for Pauli channels [51].

## 3.7 Bound on the Worst-Case Error

The decay parameter also provides an upper- and lower-bound on the worst-case error rate, which is given by the diamond distance of $\mathcal{E}$ from the identity operation; $\frac{1}{2}\|\mathcal{E} - \mathcal{I}_d\|_\diamond$, where the diamond distance of a map $\Delta$ is

$$\|\Delta\|_\diamond = \sup_\psi \|\mathcal{I} \otimes \Delta(\psi)\|_1. \tag{3.20}$$

As shown in Ref. [71], the worst case error of $\mathcal{E}$ is bounded by

$$\frac{r(d+1)}{d} \leq \frac{1}{2}\|\mathcal{E} - \mathcal{I}_d\|_\diamond \leq \sqrt{d^{-1}(d+1)r}. \tag{3.21}$$

## 3.8 Limits of Randomized Benchmarking

Although randomized benchmarking has proved to be a powerful tool for estimating the average gate fidelity, there are limitations to the technique. Firstly, there are instances where assigning a fidelity to individual quantum gates may be useful, or required, instead of the average fidelity of the Clifford group (or any 2-design) obtained through standard RB. Quantum process tomography offers full characterization of individual quantum gates, but as previously mentioned, there are scalability issues associated with this technique. A variation of RB in Ref. [52] called interleaved randomized benchmarking addresses this concern by slightly altering the RB protocol to provide a bound on the average fidelity of a single gate within the benchmarking group. This thesis will not focus on this technique, but additional literature on this method can be found in Ref. [41].

Secondly, standard RB gives a single piece of information, the average gate fidelity. Gates are not fully characterized by their average fidelity, and we may require additional information to ensure that an implementation of a desired algorithm will perform without issue. For example, certain architectures are vulnerable to probabilistic and irretrievable loss of a qubit, which cannot be distinguished from other types of noise if only the average fidelity is known. Other architectures are prone to errors which exhibit coherences between the qubit subspace, and a second, disjoint subspace of the Hilbert space. This type of noise is called leakage, and is trace-preserving on the full Hilbert space, but not within the desired subspace. The standard RB method simply assumes the errors are trace-preserving in the qubit subspace, and therefore does not provide characterization of loss or leakage. Loss and leakage processes are problematic in quantum computation because many quantum error correction codes [13, 33, 66] rely on the assumption that the qubits remain in the code

space. Noise models for qubit loss and leakage, and potential solutions for characterizing them are discussed in detail in Chapters 4 and 5. Additionally, RB cannot detect if the noise on two or more qubits is comprised of single qubit errors, or if the noise performs an entangling (or correlated) operation on multiple qubits. Correlated errors are discussed in detail in Chapter 6, along with a variation of the RB protocol which provides a test for these types of errors.

We also point out for standard RB, (and the protocols presented in this thesis) there are limitations on the accuracy of the estimation of $p$ in RB due to finite sampling, and the number of experiments that are possible to perform in a timely manner. Obtaining rigorous confidence intervals on the parameters obtained from our protocol is still an open problem. However, these types of limitations are addressed in Ref. [51], and techniques bounding the number of sequences to be sampled [71] and using Bayesian methods to refine prior information [35] have recently been described, and should also be applicable to our protocol. We also provide numerical simulations to gauge the effects of finite sampling on the new protocols presented in this thesis.

# Chapter 4

# Characterizing Loss

Many experimental implementations of qubits—such as ion traps [10, 75], optical lattices [67] and linear optics [26]—suffer from irretrievable loss, that is, there is a nonzero probability of the qubit vanishing (as opposed to leaking to other energy levels). Such loss of normalization can be a substantial obstacle to many quantum information protocols, requiring different error-correction techniques to achieve fault-tolerance [68, 26, 74]. For example, the surface code may not be used directly if there is any probability of losing a qubit, while for the topological cluster states, loss rates of less than 1% are required to avoid impractical overheads [74].

However, there are two substantial challenges in characterizing loss. Firstly, the loss rate may depend on the state of the qubit, such as when a qubit is encoded in a superposition of vacuum and single-photon states. Secondly, the loss due to imperfect operations has to be distinguished from the inefficiency of the detector [38]. Quantum process tomography [16, 60] could be used to characterize loss, however, it is inefficient in the number of qubits and is sensitive to SPAM errors [54] and so cannot distinguish between loss due to imperfect operations and inefficient detectors.

This chapter covers results which are published as Ref. [70]. We present a robust and efficient protocol that characterizes the loss rate due to imperfect operations averaged over input states. Our protocol is platform-independent, simple to implement and analyze, and only assumes that the noise is Markovian. We begin by defining survival rates and then present our protocol and derive the associated analytical decay curve under the assumption of Markovian noise. We then prove that the average loss rate estimated via our protocol provides a practical bound on the loss rate for any state. Since our protocol is robust to SPAM errors, the choice of state and measurement in our protocol is unconstrained.

However, we discuss two particularly suitable choices. The first of these allows one of the parameters in randomized benchmarking [50] to be independently estimated and leads to a new test for non-Markovian effects. The first choice also allows for an estimate of the unitarity metric introduced in Ref. [72] with no additional experimental overhead. The second choice maximizes the signal, reducing the resources required to obtain a reliable fit. In addition, we demonstrate that our protocol produces reliable estimates of loss rates through a numerical simulation under an error model that has the greatest variation in loss over states. Finally we illustrate how the analytical model breaks down when applied to systems that have reversible (coherent) leakage to an ancillary level.

## 4.1 Average Survival Rates

In order to distinguish between inefficient detectors and lossy processes, we now define survival rates. Many methods for characterizing a process $\mathcal{E}$ (including randomized benchmarking [23, 45, 50]) assume it is trace-preserving. However, many experimental processes are not trace-preserving, but instead a state $\rho$ has a survival rate under $\mathcal{E}$

$$S(\rho|\mathcal{E}) = \frac{\mathrm{Tr}[\mathcal{E}(\rho)]}{\mathrm{Tr}\rho} \tag{4.1}$$

that is less than 1, or, equivalently, a nonzero loss rate $L(\rho|\mathcal{E}) = 1 - S(\rho|\mathcal{E})$. Since the trace is linear and any unnormalized density matrix is proportional to a unit-trace density matrix, the survival rate averaged over all states (hereafter the average survival rate) is simply the survival rate of the maximally mixed state, that is, $S(\mathcal{E}) := S(\frac{1}{d}\mathbb{I}|\mathcal{E})$. Correspondingly, the average loss rate is $L(\mathcal{E}) = 1 - S(\mathcal{E})$.

In the Liouville representation for basis $\mathcal{B} = \{B_1, \ldots, B_{d^2}\}$, typically $B_1$ is chosen to equal the normalized identity operator. It is therefore easy to see

$$S(\mathcal{E}) = \frac{\mathrm{Tr}[\mathcal{E}(\frac{1}{d}\mathbb{I})]}{\mathrm{Tr}(\frac{1}{d}\mathbb{I})} = \mathrm{Tr}\left[\tfrac{1}{\sqrt{d}}\mathbb{I}\mathcal{E}\left(\tfrac{1}{\sqrt{d}}\mathbb{I}\right)\right] = \boldsymbol{\mathcal{E}}_{11}. \tag{4.2}$$

So the average survival rate is equal to the identity component, or the $(1,1)$ component, of the Liouville representation of $\mathcal{E}$.

## 4.2 Experimental Protocol

We now present a protocol for characterizing the average survival rate $S(\mathcal{E})$ in the experimental implementations $\{\mathcal{E}_g\}$ of a set of gates $\mathcal{G} = \{g_1, \ldots, g_{|\mathcal{G}|}\}$ that are at least a unitary

1-design (e.g., the Pauli or Clifford groups) [17]. For simplicity, we assume the noise is time- and gate-independent Markovian noise, so that $\mathcal{E}_g = g \circ \mathcal{E}$ for some fixed map $\mathcal{E}$. This approach can be extended to accommodate time- and gate-dependent noise and a model of non-Markovian noise by applying the approaches of Refs. [51, 71, 4].

Our protocol for estimating $S(\mathcal{E})$ is as follows.

1. Choose a sequence length $m \in \mathbb{N}$.
2. Choose a random sequence $\mathbf{k} = (k_1, \ldots, k_m)$ of $m$ integers uniformly at random, where $k_j \in \{1, \ldots, |\mathcal{G}|\}$.
3. Prepare a state $\rho$.
4. Apply the sequence of gates $g_{k_m} \circ \ldots \circ g_{k_1}$.
5. Measure some operator $Q$ (e.g., a self-adjoint operator or POVM element).
6. Repeat steps 3–5 to estimate
$$Q_{\mathbf{k}} = \mathrm{Tr}\left[Q g_{k_m} \circ \mathcal{E} \circ \ldots \circ g_{k_1} \circ \mathcal{E}(\rho)\right] \tag{4.3}$$
to a desired precision.
7. Repeat steps 2–6 to estimate the expectation value of $Q_{\mathbf{k}}$:
$$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = |\mathcal{G}|^{-m} \sum_{\mathbf{k}} Q_{\mathbf{k}} \tag{4.4}$$
to a desired precision (see, e.g., Ref. [71] for methods to bound the number of sequences required to obtain a given precision).
8. Repeat steps 1–7 for multiple $m$ and fit to the decay curve
$$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = D(Q) S(\rho|\mathcal{E}) S^{m-1}(\mathcal{E}), \tag{4.5}$$
derived below, to obtain estimates of $S(\mathcal{E})$ and $S(\rho|\mathcal{E})D(Q)$ (assuming $\mathrm{Tr}\rho = 1$) where $D(Q) = \mathrm{Tr}\, Q/d$ is the detector efficiency.

(Note that the above protocol differs from the randomized benchmarking protocol of Ref. [50] in that no inversion gate is applied prior to the measurement.) The assumption $\mathrm{Tr}\rho = 1$ is needed, since for $m = 1$, $\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = D(Q) S(\rho|\mathcal{E})$ only under this assumption. If $\mathrm{Tr}\rho < 1$, then the experimental value of $\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}}$ for $m = 1$ will differ from this expression.

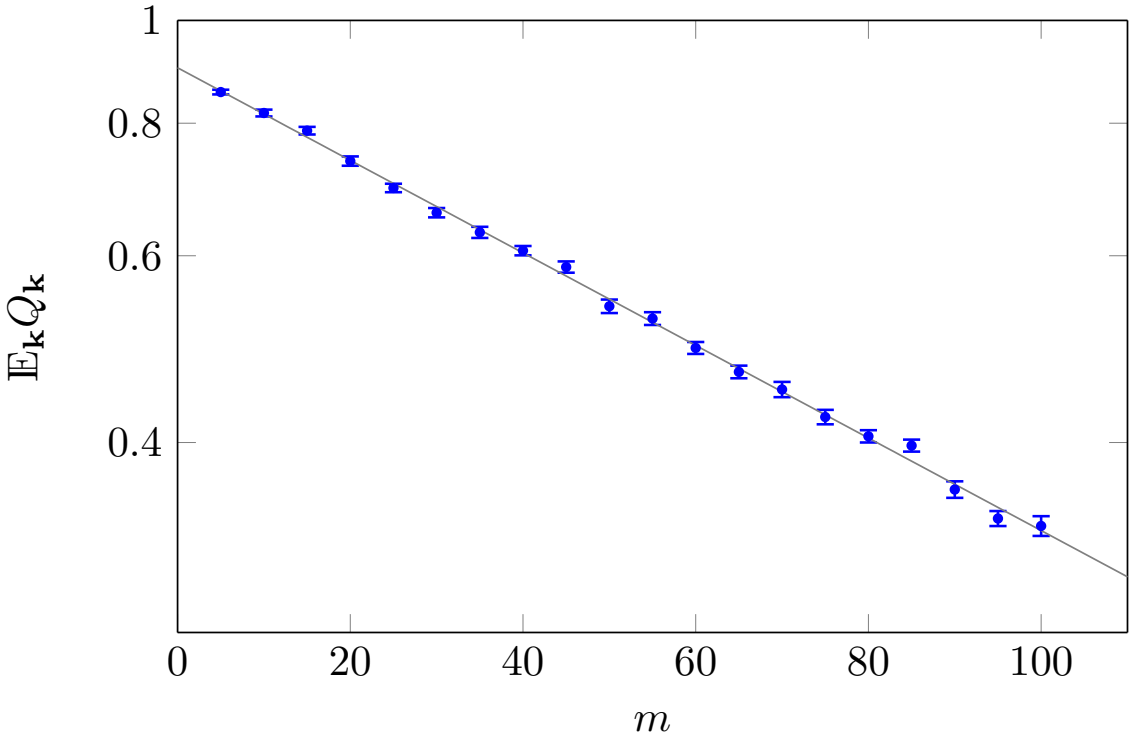



Figure 4.1: Semilog plot of numerical data for our protocol demonstrating robust identification of the average loss rate. The numerical data is obtained for loss model described by Eq. (4.9). Fitting the numerical results gave $S(\mathcal{E}) = 0.9900(2)$ and $D(Q) = 0.902(8)$, compared to the theoretical values $S(\mathcal{E}) = 0.9901$ and $D(Q) = 0.910$ respectively.

Results of a numerical simulation of our protocol for a specific loss model are illustrated in Fig. 4.1, demonstrating the robust performance of our protocol.

For the numerical simulation, the set of operations $\mathcal{G}$ is the set of single-qubit Paulis, and we modeled the error as $\mathcal{E}$ as

$$\mathcal{E}(\rho) = (\alpha \left|0\right\rangle \left\langle 0\right| + \left|1\right\rangle \left\langle 1\right|)\rho(\alpha \left|0\right\rangle \left\langle 0\right| + \left|1\right\rangle \left\langle 1\right|), \tag{4.6}$$

where $\alpha = 0.99$. The channel $\mathcal{E}$ corresponds to loss from the $|0\rangle$ state and, as proven in Proposition 1 below, has the greatest variation of loss over states. The measurement was set to $0.87 \left|\phi\right\rangle \left\langle \phi\right| + 0.95 \left|\phi^{\perp}\right\rangle \left\langle \phi^{\perp}\right|$ where $\{\left|\phi\right\rangle, \left|\phi^{\perp}\right\rangle\}$ is a randomly-chosen orthonormal basis, to model an imperfect detector. The data points are the estimates of $\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}}$ for

$m = 5, 10, ..., 100$ obtained by sampling 30 random sequences of single-qubit Pauli operators (unitary 1-design) and the error bars are the standard errors of the mean. The grey line is the fit to the model in Eq. (4.5), obtained using MATLAB's nlinfit package.

## 4.3 Analysis

To derive the decay curve in Eq. (4.5), note that averaging over all sequences corresponds to independently averaging over all group gates $g_{k_j}$, so that

$$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = \mathrm{Tr}\left[Q\overline{\mathcal{G}} \circ \mathcal{E} \circ \ldots \circ \overline{\mathcal{G}} \circ \mathcal{E}(\rho)\right] \tag{4.7}$$

where $\overline{\mathcal{G}}(\rho) = |\mathcal{G}|^{-1} \sum_{g\in\mathcal{G}} g\rho g^\dagger$ (noting that a unitary channel corresponds to unitary conjugation). Since $\mathcal{G}$ is a unitary 1-design (and a linear map), $\overline{\mathcal{G}}(A) = \mathrm{Tr}(A)\mathbb{I}/d$ for all $d \times d$ matrices $A$ [17, 3]. Therefore, assuming $\mathrm{Tr}\rho = 1$, $\overline{\mathcal{G}} \circ \mathcal{E}(\rho) = S(\rho|\mathcal{E})\mathbb{I}/d$ and $\overline{\mathcal{G}} \circ \mathcal{E}(\mathbb{I}/d) = S(\mathcal{E})\mathbb{I}/d$ and so Eq. (4.7) simplifies to Eq. (4.5).

The average survival rate obtained via our protocol is one possible figure of merit that could be used to characterize loss, with an important alternative being the worst-case loss. However, as we now prove, the average loss provides a practical bound for the worst-case loss:

**Proposition 1.** *For any quantum channel $\mathcal{E}$ and state $\rho$ for a $d$-dimensional system,*

$$L(\rho|\mathcal{E}) \leq L(\mathcal{E})d.$$

*Moreover, for all $d$ there exist channels $\mathcal{E}$ and states $\rho$ that saturate this bound.*

*Proof.* Let $\rho$ and $\mathcal{E}$ be arbitrary states of and channels for a $d$-dimensional system. Since the trace is linear and any valid state can be written as $\rho = \tau\mathrm{Tr}\rho$ where $\tau$ is a unit-trace density matrix, the survival rate is independent of $\mathrm{Tr}\rho$, so we assume $\mathrm{Tr}\rho = 1$ without loss of generality.

Let $\rho' = (\mathbb{I} - \rho)/(d-1)$, which is a valid quantum state since it is Hermitian and positive-semidefinite by construction and has unit trace. Since $\rho'$ is a valid quantum state, the probability of detecting a system in the state $\rho'$ after applying $\mathcal{E}$ is a true probability and thus

$$\frac{dS(\mathcal{E}) - S(\rho|\mathcal{E})}{d-1} = \frac{\mathrm{Tr}[\mathcal{E}(\mathbb{I}) - \mathcal{E}(\rho)]}{d-1} = \mathrm{Tr}\mathcal{E}(\rho') \leq 1, \tag{4.8}$$

where we have used the fact that quantum channels and the trace are linear. Rearranging and substituting $L = 1 - S$ gives the desired bound.

To see that the bound is saturated, fix $d$ and consider the channel

$$\mathcal{E}(\rho) = [\mathbb{I} + (\alpha - 1)\,|0\rangle\langle 0|]\rho[\mathbb{I} + (\alpha - 1)\,|0\rangle\langle 0|] \tag{4.9}$$

for $\alpha \in [0, 1]$. For this channel,

$$\mathcal{E}(|j\rangle\langle j|) = \begin{cases} \alpha^2\,|0\rangle\langle 0| & j = 0 \\ |j\rangle\langle j| & j \neq 0, \end{cases} \tag{4.10}$$

so $L(|j\rangle\langle j|\,|\mathcal{E}) = \delta_j(1 - \alpha^2)$ and

$$L(\mathcal{E}) = \frac{1}{d}\sum_j L(|j\rangle\langle j|\,|\mathcal{E}) = \frac{1 - \alpha^2}{d}. \tag{4.11}$$

Therefore there exists a channel $\mathcal{E}$ and a state $\rho$ such that $L(\rho|\mathcal{E}) = L(\mathcal{E})d$. □

For average survival rates close to 1, the estimate of $S(\rho|\mathcal{E})D(Q)$ can be used to directly estimate the detector efficiency $D(Q)$, since

$$S(\rho|\mathcal{E})D(Q) \in [(1 - d[1 - S(\mathcal{E})])D(Q), D(Q)] \tag{4.12}$$

by Proposition 1. Consequently, $S(\rho|\mathcal{E})D(Q)/S(\mathcal{E})$ will give an estimate of $D(Q)$ that is accurate to within a factor of $(d-1)L(\mathcal{E})$. Estimating $D(Q)$ can be used to estimate the efficiency of the detector as

$$\eta = \frac{D(Q)}{D(Q_{\text{ideal}})}, \tag{4.13}$$

where $Q_{\text{ideal}}$ and $Q$ are the ideal and actual measurement operators. That is, $\eta$ is the ratio of observed to expected detector "clicks", averaged (independently) over all states.

## 4.4 Model of Loss for Numerical Simulations

For the numerical simulation of loss in Fig. 4.1, the error was modelled by $\mathcal{E}$ as in Eq. (4.9) for a single qubit, where $\alpha = 0.99$. Therefore, the survival rate is $S(\mathcal{E}) = 1 - L(\mathcal{E})$, where $L(\mathcal{E})$ is defined in Eq. (4.11). This gives,

$$S(\mathcal{E}) = 1 - \frac{1 - \alpha^2}{d} = \frac{1 + \alpha^2}{2} \tag{4.14}$$

for $d = 2$.

Alternatively, in the Pauli-Liouville representation the error channel is given by

$$\boldsymbol{\mathcal{E}} = \begin{pmatrix} \frac{\alpha^2+1}{2} & 0 & 0 & \frac{\alpha^2+1}{2} - 1 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ \frac{\alpha^2+1}{2} - 1 & 0 & 0 & \frac{\alpha^2+1}{2} \end{pmatrix}, \tag{4.15}$$

and again $S(\mathcal{E}) = \boldsymbol{\mathcal{E}}_{11} = \frac{\alpha^2+1}{2}$. Replacing $\alpha$ with 0.99 gives the theoretical value $S(\mathcal{E}) = 0.9901$.

The measurement was modeled by $Q = x\left|\phi\right\rangle\left\langle\phi\right| + y\left|\phi^{\perp}\right\rangle\left\langle\phi^{\perp}\right|$ where $x = 0.87$, $y = 0.95$, and $\{\left|\phi\right\rangle, \left|\phi^{\perp}\right\rangle\}$ was a randomly-chosen orthonormal basis to model an imperfect detector. Independent of basis, $\mathrm{Tr}Q = x + y = 1.82$, therefore $D(Q) = \mathrm{Tr}Q/d = 0.91$.

## 4.5 Choosing Initial States and Measurements

Our protocol is robust to SPAM errors, in that the choice of the state $\rho$ and measurement operator $Q$ only effect the value of the constant $S(\rho|\mathcal{E})D(Q)$. However, there are two choices of $Q$ and $\rho$ that have particular benefits.

(i) The most useful scenario corresponds to choosing $\mathcal{G}$ to be a unitary 2-design [17] and choosing $\rho, Q \approx \left|0\right\rangle\left\langle 0\right|$ as in randomized benchmarking [50]. There are two major advantages to this choice. Firstly, with this choice the same data can also be used to estimate the unitarity of $\mathcal{E}$, a quantitative measure of how the noise $\mathcal{E}$ effects the purity of input states [72]. Secondly, estimating the constant prefactor in Eq. (4.5) with this choice is particularly useful because it allows an additional and vital constraint to be imposed when fitting randomized benchmarking data to the fidelity decay curve. In Ref. [50], it was shown that the fidelity decay curve is

$$A(\mathcal{E}')p^m + B(\mathcal{E}') \tag{4.16}$$

where $p$ is related to the average gate fidelity, $\mathcal{E}'$ is the average error under the convention that the experimental implementation of $g$ is written as $\mathcal{E}_g = \mathcal{E}' \circ g$ (in contrast to our choice of $\mathcal{E}_g = g \circ \mathcal{E}$) and

$$\begin{aligned} A(\mathcal{E}') &= \mathrm{Tr}[Q\mathcal{E}'(\rho - \tfrac{\mathbb{I}}{d})] \\ B(\mathcal{E}') &= \mathrm{Tr}[Q\mathcal{E}'(\tfrac{\mathbb{I}}{d})]. \end{aligned} \tag{4.17}$$

If the alternative convention of writing errors as $\mathcal{E}' \circ g$ is applied to Eq. (4.7), then the constant prefactor $S(\rho|\mathcal{E})D(Q)$ in Eq. (4.5) becomes $B(\mathcal{E}')$. Since the fidelity decay curve is in terms of observable properties, it is independent of the choice of convention and so $B(\mathcal{E}') = S(\rho|\mathcal{E})D(Q)$. Obtaining a precise estimate of the constant term for randomized benchmarking is important for two reasons. First, underestimating the constant term $B(\mathcal{E}')$ [and hence overestimating the coefficient $A(\mathcal{E}')$] will lead to an overestimate of the decay parameter $p$, or, equivalently, an underestimate of the average gate infidelity. That is, *underestimating the constant term will falsely indicate that the gates are performing better than they actually are.* Second, the values of the constants $A$ and $B$ in randomized benchmarking are not completely independent: they must satisfy particular constraints in order to correspond to physical Markovian noise processes. In particular, for qubits, note that

$$B(\mathcal{E}) - A(\mathcal{E}) = \mathrm{Tr}[Q\mathcal{E}(\rho^{\perp})] \tag{4.18}$$

where $\rho^{\perp}$ is the state whose Bloch vector is anti-parallel to that of $\rho$. Therefore $B(\mathcal{E}) - A(\mathcal{E})$ is a probability and so must be nonnegative if the noise is truly Markovian. Consequently, if $B(\mathcal{E}) - A(\mathcal{E})$ is (strongly) negative, then either the noise is non-Markovian or strongly gate dependent and so the estimate of the average gate infidelity in randomized benchmarking is not known to be accurate. Moreover, if the prefactor $S(\rho|\mathcal{E})D(Q)$ in Eq. (4.5) is estimated by setting $m = 1$, then the resulting estimate is unaffected by the presence of non-Markovian effects between sequential operations (since there is only one operation applied). Therefore if the estimate obtained by setting $m = 1$ differs from the estimate obtained from fitting the randomized benchmarking data under the protocol of Ref. [50], then this disagreement indicates that non-Markovian effects are present in the data for the latter.

(ii) Alternatively, given any allowed choice of $\mathcal{G}$, choosing $Q \approx \mathbb{I}$ and $\rho$ to be any unit-trace density matrix will maximize the value of the constant prefactor in Eq. (4.5), reducing the number of experiments required to obtain a desired precision (since $\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}}$ is close to one for sufficiently small $m$). Note that this data can be collected under the same experimental configuration as case (i), where $Q = |0\rangle\langle 0|$ and $\mathcal{G}$ is a unitary 2-design, by simply re-incorporating the outcomes associated with $\mathbb{I} - Q$ that are discarded in case (i). This data gives independent information because by assumption the probabilities of these two outcomes are not constrained to add to 1 due to presence of loss.

## 4.6 Relation to the $\chi$-matrix

We can also consider the situation in the $\chi$-matrix representation.

**Proposition 2.** *The parameter $S(\mathcal{E})$ is related to the $\chi$-matrix in the following way:*

$$S(\mathcal{E}) = \frac{\chi_{0,0} + \chi_{1,1} + \chi_{2,2} + \chi_{3,3}}{d} = \frac{\mathrm{Tr}(\chi)}{d} \tag{4.19}$$

*so $\mathrm{Tr}(\chi) = d \cdot S(\mathcal{E})$ for a trace-decreasing map, where $S(\mathcal{E})$ quantifies the percent of information not lost to the environment.*

*Proof.* In the normalized Pauli-Liouville basis $\boldsymbol{\mathcal{E}}_{\alpha\beta} = \mathrm{Tr}\left(\frac{P_\alpha^\dagger}{\sqrt{d}}\mathcal{E}(\frac{P_\beta}{\sqrt{d}})\right)$, and in the $\chi$ matrix representation $\mathcal{E}^{(\chi)} = \sum_{\gamma,\delta}\chi_{\gamma,\delta}\left(\frac{P_\alpha}{\sqrt{d}}\rho\frac{P_\beta^\dagger}{\sqrt{d}}\right)$, for unnormalized Pauli operators $P_i$. Combining the two, we find:

$$\boldsymbol{\mathcal{E}}_{\alpha\beta} = \frac{1}{d^2}\sum_{\gamma,\delta}\chi_{\gamma,\delta}\mathrm{Tr}(P_\alpha P_\gamma P_\beta P_\delta). \tag{4.20}$$

Then the element $S(\mathcal{E})$ can be found by letting $P_\alpha = P_\beta = I$.

$$S(\mathcal{E}) = \frac{1}{d^2}\sum_{\gamma,\delta}\chi_{\gamma,\delta}\mathrm{Tr}(P_\gamma P_\delta) \tag{4.21}$$

$$= \frac{1}{d^2}\sum_{\gamma,\delta} d\cdot\chi_{\gamma,\delta}\delta_{\gamma,\delta} \tag{4.22}$$

$$= \frac{\mathrm{Tr}(\chi)}{d} \tag{4.23}$$

Note: In the case where $\mathcal{E}$ is a CPTP map, $S(\mathcal{E}) = 1 \to \mathrm{Tr}(\chi) = d$, as expected. □

## 4.7 Failure of the Loss Protocol for Coherent Leakage

A distinct, but closely related error to loss is (coherent) leakage, wherein the system is "leaked" from the qubit subspace to other energy levels. Leakage errors are non-Markovian errors on the qubit subspace, since the system can return to the qubit subspace. Coherent leakage is a known consequence of control imperfections in certain implementations of the coupling gate in ion traps [63] and the controlled-phase gate in superconducting qubits [19,

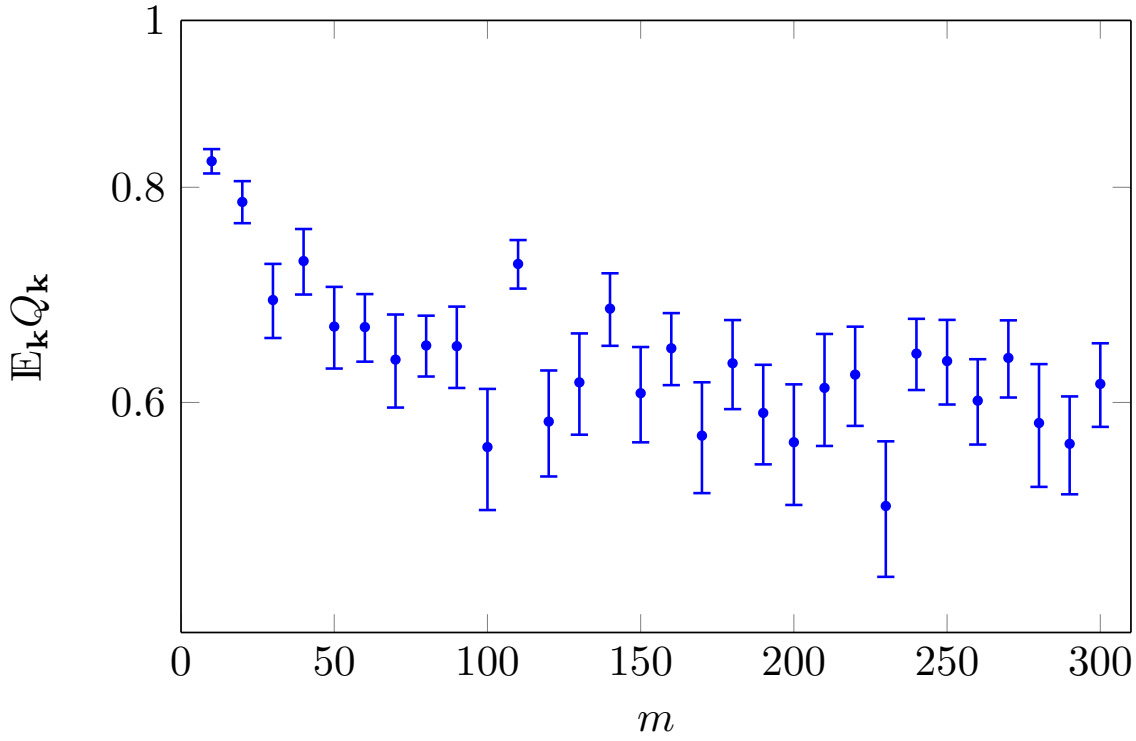



Figure 4.2: Semi-log plot demonstrating the signature of non-Markovian leakage under our protocol. Numerical results were obtained for a model of coherent leakage from a qubit subspace to a third level under a small random unitary on the full qutrit space.

20]. Fig. 4.2 shows the results of our protocol given a model of coherent leakage, in particular, an error model for a random (fixed) unitary acting on a qutrit with a random relative phase between the leakage level and the qubit levels. The data points are the estimates of $\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}}$ for $m = 10, 20, ..., 300$ obtained by sampling 30 random sequences of single-qubit Pauli operators and the error bars are the standard errors of the mean. The results initially appear to fit a single exponential decay, but then quickly converge to a constant, similar to the behavior observed in Ref. [25]. Consequently, if experimental data for our protocol does not neatly fit a single exponential, one explanation would be that there is a leakage level that has not been accounted for. A simple protocol for estimating rates of coherent leakage has been provided in Ref. [69] and in Chapter 5 of this thesis.

## 4.8 Summary

In this chapter, we have presented a platform-independent and robust protocol for characterizing the average loss rate due to noisy implementations of operations. Our protocol can also be used to estimate the detector efficiency, provided the loss rate due to noisy operations is sufficiently small. Since our protocol is easy to implement, it is also a promising technique for experimentally optimizing quantum control, as done, e.g. in Ref. [40] using randomized benchmarking experiments.

Experimentally implementing our protocol yields a single exponential decay curve which can be fitted to our analytical expression to obtain the average loss rate. If the experimental data deviates significantly from a single decay curve, the experimental noise is either strongly gate-dependent or non-Markovian. We have illustrated that the decay can be observed and fitted in practice through numerical simulations of loss for a specific error model and also that non-Markovian leakage to an ancillary level results in a deviation from a single exponential. However, fully characterizing how the present protocol (and other randomization-based protocols) behave in the presence of non-Markovian noise remains an open problem.

Our protocol is scalable and robust against state-preparation and measurement errors. However, particular choices of preparations and measurements give extra information. If the set of gates is chosen to be a unitary 2-design and the preparation and measurement are the same as those used in standard randomized benchmarking, then our current protocol can be applied to directly estimate one of the parameters in randomized benchmarking and thus provides a test to indicate non-Markovian noise. Furthermore, with this choice of preparation and measurement, the same data obtained via our protocol can be used to estimate the unitarity presented in Ref. [72] and thus to estimate how close the noise is to depolarizing noise.

# Chapter 5

# Characterizing Leakage

An important error mechanism in many experimental implementations of quantum information is leakage outside of the Hilbert space under consideration. Such leakage errors can be a substantial obstacle to fault-tolerant computation [2, 27, 31, 74].

Standard RB only provides limited information about leakage rates [25]. There are platform-dependent methods for characterizing leakage in many of the leading experimental approaches to quantum computation, such as ion trap qubits [18], superconducting qubits [76, 56] and quantum dots [53]. However, these approaches do not have all the advantages of RB, in particular, scalability with the number of qubits, robustness to SPAM errors and no assumptions about the underlying error process beyond the assumption of Markovianity.

Furthermore, leakage is distinct from loss, which we emphasize by using the term coherent leakage. In many physical implementations, coherent leakage can be considered as a coherent transition to an extra dimension (e.g., an electron excitation to an energy level outside the Hilbert space being considered) which later transitions back to the Hilbert space under consideration. These transitions back to the Hilbert space make coherent leakage a fundamentally non-Markovian process.

We present a protocol that provides an estimate of the average leakage rate for coherent leakage over a given set of quantum gates. We consider computational and leakage spaces of arbitrary dimensions, so that our protocol can be applied to both physical and logical qudit systems. We demonstrate that our protocol produces reliable estimates of leakage rates through a numerical simulation of our protocol for a specific error model.

## 5.1 Defining Leakage Rates

Leakage is a non-Markovian process, which we model as a Markovian process on a higher-dimensional space. We can therefore again consider experimental implementations of unitaries $g$ to be written as $\mathcal{E}_g = g \circ \mathcal{E}$ for some completely positive map $\mathcal{E} : L(\mathcal{H}) \to L(\mathcal{H})$. Note that if, as is often the case, $g$ acts on a subspace $\mathcal{H}_1$ where $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$, then we implicitly extend $g$ to $g \oplus \mathcal{I}_{\mathcal{H}_2}$ where $\mathcal{I}_{\mathcal{H}_j}$ denotes the identity on $\mathcal{H}_j$ and $\oplus$ denotes the direct sum, such that for two matrices $A$ and $B$,

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}. \tag{5.1}$$

The dimension of $\mathcal{H}_j$ is given by $d_j$ and $d = d_1 + d_2$.

Many methods for characterizing noisy channels $\mathcal{E}_g$ assume $\mathcal{E}$ remains (is trace-preserving) in the designated subspace. However, an important limitation of many experimental implementations is that errors are not trace-preserving on the desired space. Meaning, generally the probability $\mathrm{Tr}[\mathcal{P}_{\mathcal{H}_1}\rho]$ of the system being in a Hilbert space $\mathcal{H}_1 \subseteq \mathcal{H}$ can decrease, and the probability $\mathrm{Tr}[\mathcal{P}_{\mathcal{H}_2}\rho]$ of the system being in a separate Hilbert space $\mathcal{H}_2 \subseteq \mathcal{H}$ can increase upon applying an operation, where

$$\mathcal{P}_{\mathcal{H}_j} = \mathcal{I}_{\mathcal{H}_j} \oplus 0 \tag{5.2}$$

is the projector onto $\mathcal{H}_1$ [5]. (Conversely, $\mathrm{Tr}[\mathcal{P}_{\mathcal{H}_1}\rho]$ may increase, and $\mathrm{Tr}[\mathcal{P}_{\mathcal{H}_2}\rho]$ decrease.) If the noise is trace-decreasing on the full Hilbert space of the system then the error is characteristic of loss as in Chapter 4. However, if the noise is trace-preserving on the full space, but trace-decreasing within the desired subspace, the error is characteristic of coherent leakage to extra level(s).

We define the survival rate of a state in an arbitrary subspace $\rho \in \mathcal{H}_j$ under a CP map $\mathcal{E}$ to be

$$S(\rho|\mathcal{E}, \mathcal{H}_j) = \frac{\mathrm{Tr}[\mathcal{P}_{\mathcal{H}_j}\mathcal{E}(\rho)]}{\mathrm{Tr}\rho}. \tag{5.3}$$

We will consider survival rates averaged over states in a subspace $\mathcal{H}_1$ of $\mathcal{H}$. In order to define these averages, note that any $\rho \in L(\mathcal{H}_1)$ can be written as $p\tau$ for some $p \in [0, 1]$ and $\tau \in L(\mathcal{H})$ such that $\mathrm{Tr}\tau = 1$. Substituting this into Eq. (5.3) gives

$$S(\rho|\mathcal{E}, \mathcal{H}_1) = \mathrm{Tr}[\mathcal{P}_{\mathcal{H}_1}\mathcal{E}(\tau)], \tag{5.4}$$

which is a linear function of $\tau$. Consequently, the average survival rate in $\mathcal{H}_1$ over any measure $\mathrm{d}\tau$ over mixed states that is invariant under unitaries acting on $\mathcal{H}_1$ is

$$\begin{aligned} S(\mathcal{E}, \mathcal{H}_1) &= \int \mathrm{d}\tau \mathrm{Tr}[\mathcal{P}_{\mathcal{H}_1} \mathcal{E}(\tau)] \\ &= \mathrm{Tr}[\mathcal{P}_{\mathcal{H}_1} \mathcal{E}(d_1^{-1} \mathcal{P}_{\mathcal{H}_1})], \end{aligned} \tag{5.5}$$

where we have used the fact that $\int \mathrm{d}U U \tau U^\dagger = d_1^{-1} \mathcal{P}_{\mathcal{H}_1}$ for any density operator $\tau$, where $\mathrm{d}U$ is the Haar measure over unitaries acting on $\mathcal{H}_1$.

Since CP maps are linear and all quantum states can be written as $p\rho$ for some $p \in [0, 1]$ and $\rho \in L(\mathcal{H})$ such that $\mathrm{Tr}\rho = 1$, the survival rate for $\mathcal{H}_1 = \mathcal{H}$ is strictly nonincreasing under composition, that is, $S(\rho|\mathcal{E}' \circ \mathcal{E}, \mathcal{H}_1) \leq S(\rho|\mathcal{E}, \mathcal{H}_1)$ for all CP maps $\mathcal{E}'$. In contrast, if $\mathcal{H}_1 \subsetneq \mathcal{H}$, the survival rate can increase if $\mathcal{E}$ has coherences between $\mathcal{H}_1$ and $\mathcal{H}_2$. We therefore define the coherent survival rate to be

$$\begin{aligned} S_{\mathrm{coh.}}(\mathcal{E}) &= (\mathrm{Tr}[\mathcal{P}_{\mathcal{H}_1} \mathcal{E}(d_1^{-1} \mathcal{P}_{\mathcal{H}_1})] + \mathrm{Tr}[\mathcal{P}_{\mathcal{H}_2} \mathcal{E}(d_2^{-1} \mathcal{P}_{\mathcal{H}_2})])/2 \\ &= [S(\mathcal{E}, \mathcal{H}_1) + S(\mathcal{E}, \mathcal{H}_2)]/2. \end{aligned} \tag{5.6}$$

That is, the coherent survival rate is the average of the average survival probabilities in the two subspaces. Coherent leakage rates can then be defined as $L_{\mathrm{coh.}}(\mathcal{E}) = S(\mathcal{E}) - S_{\mathrm{coh.}}(\mathcal{E})$ where $S(\mathcal{E})$ is the survival rate on the full Hilbert space from the loss protocol, see Def. 4.1. If $\mathcal{E}$ is trace-preserving on the full space, which can be determined by the method in Chapter 4, then $S(\mathcal{E}) = 1$ and $L_{\mathrm{coh.}}(\mathcal{E}) = 1 - S_{\mathrm{coh.}}(\mathcal{E})$.

## 5.2 Experimental Protocol

We now present a protocol for characterizing the average survival rate $S_{\mathrm{coh.}}(\mathcal{E})$ over a set of operations $\mathcal{G} = \{g = v \oplus (\pm 1, \pm i)w : v \in \mathcal{V}, w \in \mathcal{W}\}$, where $\mathcal{V}$ and $\mathcal{W}$ are unitary 1-designs [17] on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Note that standard RB requires a unitary 2-design, which is a strictly stronger requirement, and that our protocol does require some experimental control over the auxiliary space $\mathcal{H}_2$.

For brevity we assume again that the noise is time- and gate-independent, though results for time- and gate-dependent noise can be obtained by applying the approaches of Ref. [71]. Our protocol works well when leakage errors exist on a much larger scale than individual gate errors, but may require some signal amplification if this is not the case.

1. Choose a sequence length $m \in \mathbb{N}$.

2. Choose a random sequence $\mathbf{k} = (k_1, \ldots, k_m)$ of $m$ integers uniformly at random, where $k_j \in \{1, \ldots, |\mathcal{G}|\}$.

3. Prepare a state $\rho$.

4. Apply the sequence of gates $g_{k_m} \circ \ldots \circ g_{k_1}$.

5. Measure some operator $Q$ (e.g., a self-adjoint operator or POVM element).

6. Repeat steps 3–5 to estimate

$$Q_{\mathbf{k}} = \mathrm{Tr}\left[Q g_{k_m} \circ \mathcal{E} \circ \ldots \circ g_{k_1} \circ \mathcal{E}(\rho)\right] \tag{5.7}$$

   to a desired precision.

7. Repeat steps 2–6 to estimate the expectation value of $Q_{\mathbf{k}}$:

$$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = |\mathcal{G}|^{-m} \sum_{\mathbf{k}} Q_{\mathbf{k}} \tag{5.8}$$

   to a desired precision (see, e.g., Ref. [71] for methods to bound the number of sequences required to obtain a given precision).

8. Repeat steps 1–7 for multiple $m$ and fit to the decay curve in Eq. (5.9), derived below, to obtain estimates of $S_{\text{coh.}}(\mathcal{E})$.

(Note that the protocol is identical to the loss protocol in section 4.2, except for the choice of group $\mathcal{G}$ and fit model, and that this protocol differs from the standard RB protocol in that no inversion gate is applied immediately prior to the measurement.)

Averaging the results over a number of random sequences with fixed $m$ will give an estimate of

$$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = A(\mathcal{E})\lambda_+^{m-1} + B(\mathcal{E})\lambda_-^{m-1} \tag{5.9}$$

for $\mathcal{H}_1 \subsetneq \mathcal{H}$, where the constants $A$ and $B$ relate to state-preparation and measurement errors and the $\lambda_\pm$ are fit parameters that give the coherent survival probability through $S_{\text{coh.}}(\mathcal{E}) = (\lambda_+ + \lambda_-)/2$. If the noise is trace-preserving on $\mathcal{H}$ (that is, if $S(\mathcal{E}) = 1$), then Eq. (5.9) simplifies to

$$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = A(\mathcal{E}) p_{\text{coh.}}^{m-1} + B(\mathcal{E}) \tag{5.10}$$

where the fit parameter $p_{\text{coh.}}$ is related to the coherent survival rate by $p_{\text{coh.}} = 2S_{\text{coh.}}(\mathcal{E}) - 1$. Fitting the relevant decay curves gives an estimate of the survival rates.

## 5.3 Derivation of the Fit Model

To derive the fit model we will work in the Liouville representation as outlined in section 2.2.3. Due to the property that channels compose via matrix multiplication in the Liouville representation, the probability for a sequence $\mathbf{k}$ is

$$Q_{\mathbf{k}} = (Q|\boldsymbol{g}_{k_m}\boldsymbol{\mathcal{E}}\dots\boldsymbol{g}_{k_1}\boldsymbol{\mathcal{E}}|\rho), \tag{5.11}$$

where $Q$ and $\rho$ are the experimental POVM elements and density matrices respectively. The average probability over all sequences of length $m$ is

$$\begin{aligned}\mathbb{E}_{\mathbf{k}}Q_{\mathbf{k}} &= |\mathcal{G}|^{-m}\sum_{k\in\mathbb{N}^m_{|\mathcal{G}|}}(Q|\boldsymbol{g}_{k_m}\boldsymbol{\mathcal{E}}\dots\boldsymbol{g}_{k_1}\boldsymbol{\mathcal{E}}|\rho)\\ &= (Q|\left[\bar{\boldsymbol{\mathcal{G}}}\boldsymbol{\mathcal{E}}\right]^m|\rho)\,.\end{aligned} \tag{5.12}$$

Since $\mathcal{G}$ is a group,

$$\bar{\boldsymbol{\mathcal{G}}}^2 = |\mathcal{G}|^{-2}\sum_{g,h\in\mathcal{G}}\boldsymbol{gh} = |\mathcal{G}|^{-2}\sum_{g',h'\in\mathcal{G}}\boldsymbol{g}' = \bar{\boldsymbol{\mathcal{G}}}, \tag{5.13}$$

so the average probability simplifies to

$$\mathbb{E}_{\mathbf{k}}Q_{\mathbf{k}} = (Q|\left[\bar{\boldsymbol{\mathcal{G}}}\boldsymbol{\mathcal{E}}\bar{\boldsymbol{\mathcal{G}}}\right]^{m-1}|\rho'), \tag{5.14}$$

where $\rho' = \mathcal{E}(\rho)$.

In order to complete the derivations, we now appeal to special properties of the group $\mathcal{G}$ chosen to characterize leakage rates.

To characterize coherent leakage, $\mathcal{G}$ is chosen so that any element $g \in \mathcal{G}$ can be written as $g = v \oplus \mu w$ for $\mu \in \{\pm 1, \pm i\}$, where $v$ and $w$ are elements of unitary 1-designs $\mathcal{V}$ and $\mathcal{W}$ on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Then, using the matrix basis $|i\rangle\langle j|$ for the operator space, so that $\boldsymbol{U} = U \otimes U^*$, where $^*$ denotes complex conjugation, we have (by consequences of Schur's Lemma)

$$\begin{aligned}\bar{\boldsymbol{\mathcal{G}}} &= |\mathcal{G}|^{-1}\sum_{\mu,v,w}(v\oplus\mu w)\otimes(v\oplus\mu w)^*\\ &= \left(|\mathcal{V}|^{-1}\sum_v v\otimes v^*\right)\oplus 0\oplus 0\oplus\left(|\mathcal{W}|^{-1}\sum_w w\otimes w^*\right)\\ &= |d_1^{-1/2}\mathcal{P}_{\mathcal{H}_1})(d_1^{-1/2}\mathcal{P}_{\mathcal{H}_1}| + |d_2^{-1/2}\mathcal{P}_{\mathcal{H}_2})(d_2^{-1/2}\mathcal{P}_{\mathcal{H}_2}|.\end{aligned} \tag{5.15}$$

Setting $B_1 = d_1^{-1/2}\mathcal{P}_{\mathcal{H}_1}$ and $B_2 = d_2^{-1/2}\mathcal{P}_{\mathcal{H}_2}$, we then have

$$\bar{\mathcal{G}}\boldsymbol{\mathcal{E}}\bar{\mathcal{G}} = s \oplus 0, \tag{5.16}$$

where $s$ is a $2 \times 2$ matrix. We can easily take powers of $s$ by putting it in lower-triangular form, so that

$$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}} = A(\mathcal{E})\lambda_+^{m-1} + B(\mathcal{E})\lambda_-^{m-1} \tag{5.17}$$

where

$$\lambda_\pm = \frac{s_{11} + s_{22}}{2} \pm \frac{1}{2}\sqrt{(s_{11} - s_{22})^2 + 4s_{12}s_{21}}. \tag{5.18}$$

are the eigenvalues of $s$, and $A$ and $B$ are constants for a given $\mathcal{E}$ (which absorb both the SPAM and the unitary that makes $s$ lower-triangular).

The sum of the eigenvalues is equal to $s_{11} + s_{22} = 2S_{\text{coh.}}(\mathcal{E})$ since

$$\begin{aligned} s_{11} &= (B_1|\boldsymbol{\mathcal{E}}|B_1) = \text{Tr}\left[\mathcal{P}_{\mathcal{H}_1}\mathcal{E}(\frac{1}{d_1}\mathcal{P}_{\mathcal{H}_1})\right] \\ s_{22} &= (B_2|\boldsymbol{\mathcal{E}}|B_2) = \text{Tr}\left[\mathcal{P}_{\mathcal{H}_2}\mathcal{E}(\frac{1}{d_2}\mathcal{P}_{\mathcal{H}_2})\right]. \end{aligned} \tag{5.19}$$

If the noise is trace-preserving on $\mathcal{H}$, then one of the eigenvalues must be one (corresponding to $\mathbb{I}_{d_1+d_2}$), and the other must then be $2S_{\text{coh.}}(\mathcal{E}) - 1$.

## 5.4 Simulation Results

Results of a numerical simulation of our protocol for a relevant model of leakage are illustrated in Fig. 5.1, demonstrating robust performance with a specific error model.

For numerical simulations of our protocol for characterizing coherent leakage, we adopted a noise model that is motivated by experimental techniques that use an auxiliary level (e.g., "shelving" in ion trap experiments [11, 63]) to protect certain states while performing another operation. The ideal shelving gate is a Pauli $X$ rotation between the second and third level, that is, $V_{ideal} = 1 \oplus X$. The group $\mathcal{G}$ of operations can be chosen as $\{P \oplus \pm 1 : P = \mathbb{I}, X, Y, Z\}$. Our model of a coherent leakage error at each time step is

$$\mathcal{E}_X = V_{\gamma_2} \circ \delta U_2 \circ V_{\gamma_1} \circ \delta U_1, \tag{5.20}$$

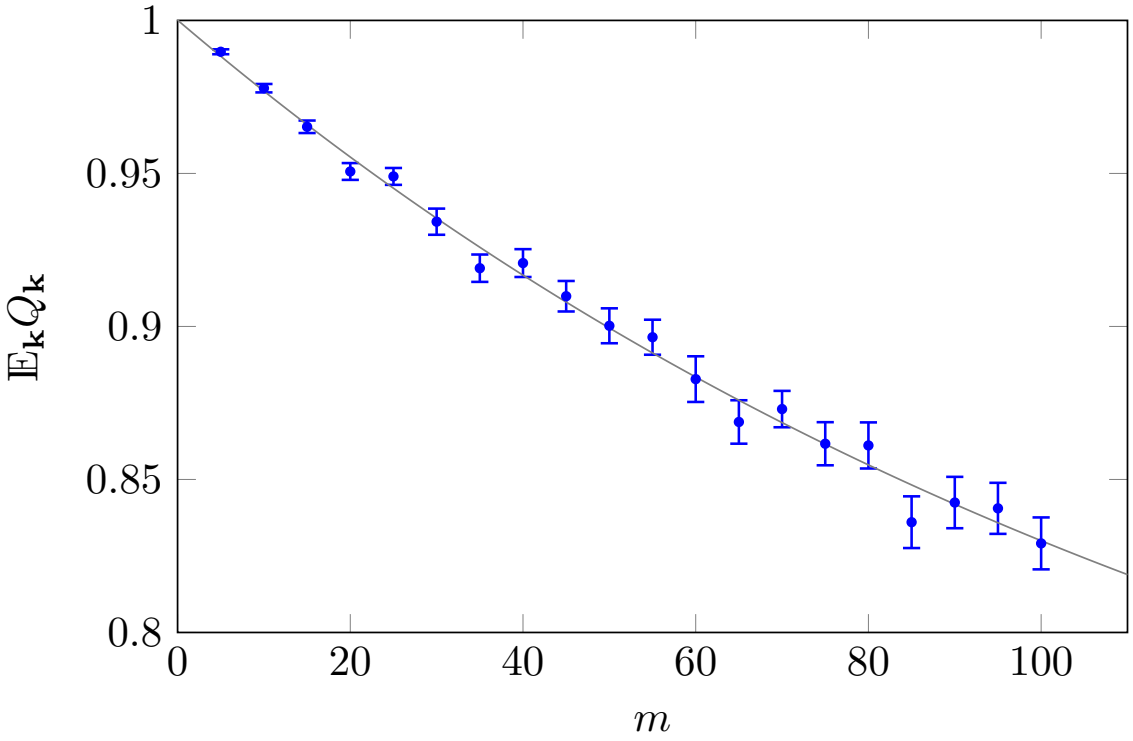



Figure 5.1: Numerical data for our protocol demonstrating robust indentification of the average leakage rate. The numerical data is obtained for the leakage model described by the composite noise channel in Eq. (5.20). Fitting the results gave $S_{\text{coh.}}(\mathcal{E}_X) = 0.996(2)$ with an $r^2$ value of 0.991, compared to the theoretical value $S_{\text{coh.}}(\mathcal{E}_X) = 0.996$.

where

$$V_\gamma = 1 \oplus \begin{pmatrix} i\sin\gamma & \cos\gamma \\ \cos\gamma & i\sin\gamma \end{pmatrix}$$
$$\delta U = e^{i\phi UXU^\dagger} \oplus 1. \tag{5.21}$$

That is, our noise model consists of imperfect shelving ($V_{\gamma_1}$) and unshelving ($V_{\gamma_2}$) gates, together with some small unitary on the code space ($\delta U_1$ and $\delta U_2$) that represents some operation or noise that takes place while the qubit is shelved. The channel $\mathcal{E}_X$ is trace-preserving on the combined code and leakage space, but is trace-decreasing when restricted to the code space. We chose $U_1$ and $U_2$ to be random, fixed unitaries on the Haar measure, and $\phi = 0.01$, $\gamma_1 = 0.04$ and $\gamma_2 = 0.03$ were fixed. The data points are the estimates of

$\mathbb{E}_{\mathbf{k}} Q_{\mathbf{k}}$ for $m = 5, 10, ..., 100$ obtained by sampling 300 random sequences of operators in $\mathcal{G}$ and the error bars are the standard errors of the mean. The grey line represents the fit to the model obtained using MATLAB's nlinfit package.

To obtain the theoretical value of the coherent survival rate, we consider the Gell-Mann basis as our operator basis, $\mathcal{B}$, for defining the Liouville representation. That is,

$$B_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad B_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$B_4 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B_5 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B_6 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$
$$B_7 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad B_8 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad B_9 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{5.22}$$

The third, fourth and fifth elements correspond to the standard (nontrivial) Pauli basis elements. Using this basis to define the Liouville representation, we find by the linearity of the trace,

$$\begin{aligned} s_{11} + s_{22} &= \mathrm{Tr}[B_1 \mathcal{E}(B_1)] + \mathrm{Tr}[B_2 \mathcal{E}(B_2)] \\ &= \mathrm{Tr}[\mathcal{P}_{\mathcal{H}_1} \mathcal{E}(\frac{1}{2} \mathcal{P}_{\mathcal{H}_1})] + \mathrm{Tr}[\mathcal{P}_{\mathcal{H}_2} \mathcal{E}(\mathcal{P}_{\mathcal{H}_2})] \\ &= 2 S_{\mathrm{coh.}}(\mathcal{E}) \end{aligned} \tag{5.23}$$

for any channel $\mathcal{E}$ acting on $\mathcal{H}$. In our model of coherent leakage, the subspace $\mathcal{H}_1$ is the qubit subspace, and the subspace $\mathcal{H}_2$ is the shelving level and has dimension one. The full Hilbert space $\mathcal{H}$ is a three-dimensional quantum system, or a *qutrit* space.

Consider now the protocol applied with the qubit shelving gate applied twice (once to shelve, and once to de-shelve the qubit) after each application of the group gates, each followed by a fixed unitary noise, as in Fig. 5.1. The output is then

$$(Q| \left[ \bar{\boldsymbol{\mathcal{G}}} \boldsymbol{\mathcal{E}}_X \bar{\boldsymbol{\mathcal{G}}} \right]^{m-1} \boldsymbol{\mathcal{E}}_X |\rho) \tag{5.24}$$

where $\mathcal{E}_X$ is as defined in Eq. (5.20). This error model is trace-preserving on the full Hilbert space, that is, $S(\mathcal{E}_X) = s_{11}(\mathcal{E}_X) = 1$, and the decay may therefore be fit to the model in Eq. (5.10).

Given our choice of random unitaries, and parameters $\gamma_1$ and $\gamma_2$, fit parameter was

$$\begin{aligned} p_{\text{coh.}} &= 2S_{\text{coh.}}(\mathcal{E}_X) - 1 \\ &= s_{11}(\mathcal{E}_X) + s_{22}(\mathcal{E}_X) - 1 \\ &= s_{22}(\mathcal{E}_X) \\ &= 0.9926 \end{aligned} \tag{5.25}$$

by direct calculation in the Liouville representation with the above basis. The theoretical coherent survival rate was then $S_{\text{coh.}}(\mathcal{E}_X) = 0.996$.

## 5.5 Summary

In this section, we have presented a protocol for characterizing average survival rates under coherent leakage to an orthogonal subspace. Experimentally implementing our protocol yields a decay curve which can be fitted to our analytical expressions to obtain the average probability of a leakage event occurring. We have also demonstrated that the decay can be observed and fitted in practice through numerical simulations of leakage for specific error models.

Our protocol is scalable and robust against state-preparation and measurement errors. Our current protocol can also be applied in conjunction with standard RB to determine both the average leakage rate and the average gate infidelity over a unitary 2-design such as the Clifford group.

# Chapter 6

# Characterizing Correlations

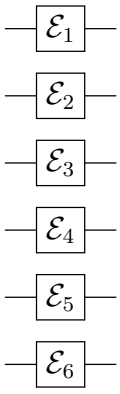
Because some FTQC codes assume errors in the computation are uncorrelated, it is important to determine whether there will be spatial correlations between qubits in a given implementation of quantum information processing. As in previous sections, we consider an ideal implementation of a quantum channel $\mathcal{U}$, which has experimental implementation $\mathcal{U} \circ \mathcal{E}$, where $\mathcal{E}$ is an error channel, which may represent an error in implementing the gate or the noise due to free evolution of the system.

The symmetrization method described in [24] for a general noise model $\mathcal{E}$ is based on identifying a symmetry associated with interesting features, and symmetrizing the noise process to obtain the twirled channel $\mathcal{E}^{\mathcal{G}}$ which contains a reduced number of parameters that relate to the interesting properties. This process consists of conjugating $\mathcal{E}$ with an operator drawn from the single-qubit Clifford group, and averaging over the group, that is, twirling $\mathcal{E}$ over the $\mathcal{C}_1^{\otimes n}$ group. This symmetrization method was adapted to the randomized benchmarking framework for two qubits in [30]. Here we adapt that 2-qubit example to arbitrary $n$ qubits, and add a step which makes the fitting procedure of the protocol robust to SPAM errors.

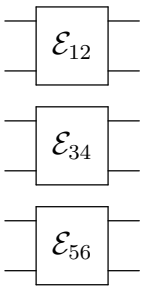
## 6.1 Definition of a Correlated Channel

Before describing symmetrization protocols in detail, an important question to answer is: "What is a correlated (or uncorrelated) channel?" We are considering correlations in space (between qubits), and define an uncorrelated channel as $\mathcal{E}$ such that $\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \ldots \otimes \mathcal{E}_n$. That is, an uncorrelated channel on $n$ qubits is a channel which can be written as a tensor

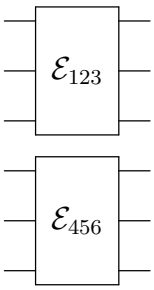
product of $n$ single-qubit channels. Conversely, a correlated channel is simply any channel which is not uncorrelated. One can imagine correlations between two individual qubits, but also between many qubits, or sets of qubits, see Fig. 6.1 for examples. Our protocol allows identification of the correlated error structure. That is, whether the errors resemble the type in Fig. 6.1a, 6.1b, 6.1c etc., or a combination of various types. (Note that correlations are not limited to existing between specific qubits, i.e. channels of the form $\mathcal{E}_{23}$, or $\mathcal{E}_{45}$, etc. are equally valid.)




(a) Example of uncorrelated errors.




(b) Example of correlated 2-qubit errors.




(c) Example of correlated 3-qubit errors.

Figure 6.1: Circuit diagrams showing uncorrelated and correlated errors.

Consider a quantum channel which rotates all states in a register by a global phase. This channel is not independent, and it may also be tempting to say such a channel is correlated, when in fact it can be written as $n$ single qubit channels, $\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \ldots \otimes \mathcal{E}_n$, and by our definition is therefore uncorrelated. Conceptually, the notions of correlation and independence have many subtleties. For this thesis we will model correlated and uncorrelated channels by the simple description above.

## 6.2 Two Qubit Example

Let us first consider the characterization of correlations in a bipartite system. The symmetrization protocol for two qubits consists of twirling over the $\mathcal{C}_1 \otimes \mathcal{C}_1$ group. In the case of full two-qubit Clifford twirl, as in standard randomized benchmarking, there are two irreducible subspaces: $\phi_0$, the projector onto the identity operator, and $\phi_1$, the projector

onto the remaining Pauli group members. In the case of the $\mathcal{C}_1 \otimes \mathcal{C}_1$ twirl, there are four irreducible subspaces: $\phi_0 = \mathcal{I} \otimes \mathcal{I}$, $\phi_2 = \mathcal{I} \otimes \mathcal{P}$, $\phi_1 = \mathcal{P} \otimes \mathcal{I}$, $\phi_{12} = \mathcal{P} \otimes \mathcal{P}$, where $\mathcal{P}$ is the vector of Pauli operators for each qubit subsystem. Therefore by Schur's lemma we obtain a Pauli-Liouville representation matrix of four blocks proportional to the identity by twirling $\mathcal{E}$

$$\boldsymbol{\mathcal{E}}^{\mathcal{C}_1 \otimes \mathcal{C}_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & q_2 I_3 & 0 & 0 \\ 0 & 0 & q_1 I_3 & 0 \\ 0 & 0 & 0 & q_{12} I_9 \end{pmatrix} \tag{6.1}$$

where $q_k = \mathrm{Tr}(\phi_k \mathcal{E})/\mathrm{Tr}(\phi_k)$. Performing sequences of Clifford gates, of various lengths $m$, puts the symmetrization protocol into the randomized benchmarking framework. The sequence fidelity has three exponential terms,

$$F_{seq}(m, Q, \rho) = A_1 q_1^m + A_2 q_2^m + A_{12} q_{12}^m + B \tag{6.2}$$

for constants $A_k$ and $B$. The individual $q_k$ may be extracted by preparing initial states, or measurements, which have support on only one of the irreducible subspaces.

The Clifford twirl produces a depolarizing channel, and in this case if $\mathcal{E}$ is uncorrelated, the twirled channel is a tensor product of two one-qubit depolarizing channels,

$$\begin{aligned} \mathcal{E}^{\mathcal{C}_1 \otimes \mathcal{C}_1} &= \mathcal{E}_1^{\mathcal{C}_1} \otimes \mathcal{E}_2^{\mathcal{C}_1} \\ &= \mathcal{D}_{p_1} \otimes \mathcal{D}_{p_2}, \end{aligned} \tag{6.3}$$

where $\mathcal{D}_p = p\rho + (1-p)\mathbb{I}/d$ is the depolarizing channel with depolarizing parameter $p$ [17]. We can therefore impose a test to flag correlations by examining $\Delta(\mathcal{E}) = q_{12}(\mathcal{E}) - q_1(\mathcal{E}) q_2(\mathcal{E})$. If the noise does not contain any correlations then, $q_{12}(\mathcal{E}) = q_1(\mathcal{E}) q_2(\mathcal{E})$. Therefore, any deviation from $\Delta(\mathcal{E}) = 0$ indicates the presence of undesired correlations in the noise.

There are several issues with this protocol. First, there are contrived examples of correlated channels for which $\Delta(\mathcal{E}) = 0$. Second, this protocol is sensitive to SPAM errors, as it requires preparing or measuring states with support in only one of the irreducible subspaces. SPAM errors are detectible, by noting deviations from a single exponential in the results of the protocol, but it would be preferable to have a protocol which is certainly insensitive to these errors. Third, this protocol should be made to check for correlations generally between any two disjoint subsets of the full quantum register. The following sections address these concerns by providing a generalization of the protocol that is fully scalable and robust to spam, and examining values of $\Delta(\mathcal{E})$ for various error models.

## 6.3 Experimental Protocol

The protocol presented here expands on previous work by generalizing the 2-qubit simultaneous RB protocol outlined in the previous section and in Ref. [30] to $n$ qubits, and providing a version of the protocol that is fully robust to SPAM errors. Our protocol consists of performing the standard randomized benchmarking procedure, except the group elements are tensor products of the single-qubit Clifford group, $\mathcal{C}_1^{\otimes n}$, as opposed to the full Clifford group on the $n$-qubit space. To make the protocol robust to SPAM errors, an additional random Pauli gate is applied at the final time-step of each independent trial.

First, choose some subset, $\mathcal{S} \subseteq \{1, \ldots, n\}$, of the full set of $n$ qudits, for which characterizing characterizing correlations is of interest. Let $\mathcal{G}$ be an $n$-fold tensor product of a single-qudit unitary 2-design, and $\mathcal{F} \subseteq \mathcal{G}$ be a single-qudit unitary 1-design. Going forward, we will assume $\mathcal{G} = \mathcal{C}_1^{\otimes n}$, the tensor product of the single-qubit Clifford group, and $\mathcal{F} = \mathcal{P}$, the single-qubit Pauli group. However, this protocol is not restricted to qubits, or any particular unitary 2- and 1-designs.

All elements $g \in \mathcal{G}$ can be written $g_i = C_{i_1} \otimes C_{i_2} \otimes \ldots \otimes C_{i_n}$, where the $C$ are the single-qubit Clifford gate. All elements $f \in \mathcal{F}$ are single qubit Paulis, $f_i = P_i$. As in standard RB, noisy implementations are again given by $\mathcal{E}_g = g \circ \mathcal{E}$, for an error $\mathcal{E}$. Our protocol is as follows:

1. Choose a sequence length $m \in \mathbb{N}$.
2. Choose a sequence $\mathbf{k} = (k_1, \ldots, k_m)$ of $m$ integers uniformly at random where $k_j \in \{1, \ldots, |\mathcal{G}|\}$.
3. Prepare an initial $n$ qubits in some state $\rho$ (usually $\rho \approx |0\rangle\langle 0|^{\otimes n}$).
4. Apply the sequence of operations $g_{k_m} \circ \ldots \circ g_{k_0}$, where $g_{k_0} = \prod_{i=1}^{m} g_{k_i}^{-1}$, is the inversion operator of the sequence.

This is the standard $\mathcal{C}_1^{\otimes n}$ RB protocol. To make this protocol robust to SPAM errors, we add the following steps:

5. Choose a strictly upper-triangular, $|\mathcal{S}| \times |\mathcal{S}|$ matrix $\mathbf{J}$, with elements $J_{i,j} \in \{1, \ldots, |\mathcal{F}|\}$; and a vector $\mathbf{L}$, of length $(n - |\mathcal{S}|)$, with elements $L_i \in \{1, \ldots, |\mathcal{F}|\}$.

6. At time step $m+1$, apply the gate $f_{\mathbf{J},\mathbf{L},\mathcal{S}}$ where

$$f_{\mathbf{J},\mathbf{L},\mathcal{S}} = \prod_{s' \notin \mathcal{S}} f^{(s')}_{L_{s'}} \prod_{s,t \in \mathcal{S}} f^{(s)}_{J_{s,t}} f^{(t)}_{J_{s,t}} \tag{6.4}$$

and $f^{(s)} = \mathbb{I}^{\otimes s-1} \otimes f \otimes \mathbb{I}^{\otimes n-s}$.

When preparation and measurement procedures are relatively very efficient, steps 5 and 6 may be omitted from the experimental protocol.

7. Measure some operator $Q$ (e.g., a self-adjoint operator or POVM element) with support in $\mathcal{S}$.

8. For a fixed $\mathbf{k}$, $\mathbf{J}$, $\mathbf{L}$, and $\mathcal{S}$, repeat steps 2–7 to estimate

$$Q_{\mathbf{k},\mathbf{J},\mathcal{S}} = \mathrm{Tr}[Q f_{\mathbf{J},\mathbf{L},\mathcal{S}} \circ g_{k_m} \circ \mathcal{E} \circ \ldots \circ g_{k_0} \circ \mathcal{E}(\rho)], \tag{6.5}$$

to a desired precision.

9. Repeat steps 1–8, choosing $\mathbf{k}$, $\mathbf{J}$, and $\mathbf{L}$ uniformly at random, but keeping $\mathcal{S}$ fixed, to estimate the expectation value

$$\mathbb{E}_{\mathbf{k},\mathbf{J},\mathbf{L}}(Q_{\mathbf{k},\mathbf{J},\mathcal{S}}) = \mathrm{Tr}\left[Q \mathcal{R}_{\mathcal{S}} \circ \left(\mathcal{E}^{\mathcal{G}}\right)^m \circ \mathcal{E}(\rho)\right] \tag{6.6}$$

to a desired precision, where

$$\mathcal{E}^{\mathcal{G}} = |\mathcal{G}|^{-1} \sum_{g \in \mathcal{G}} g^\dagger \circ \mathcal{E} \circ g \tag{6.7}$$

is the subsystem twirl of $\mathcal{E}$, and

$$\mathcal{R}_{\mathcal{S}} = |\mathcal{F}|^{-2} \sum_{\mathbf{J},\mathbf{L}} f_{\mathbf{J},\mathbf{L},\mathcal{S}}. \tag{6.8}$$

10. Repeat all steps for various $m$, and fit to the decay curve in Eq. (6.18) derived below.

## 6.4 Derivation of the Fit Model

To derive the decay curve for this protocol we will work in the Pauli-Liouville representation. The single-qubit Clifford group, $\mathcal{C}_1$ contains 2 irreducible subspaces: the projector onto the identity, and the projector on the Pauli group. Therefore, $\mathcal{C}_1^{\otimes n}$ contains $2^n$ irreducible subspaces, so by Schur's lemma the twirled channel can then be written as $2^n$ blocks, all proportional to the identity, i.e.

$$\boldsymbol{\mathcal{E}}^{\mathcal{G}} = 1 \bigoplus_{j=1}^{2^n-1} q_{\mathcal{S}_j} \mathbb{I}_{d_j} \tag{6.9}$$

for some $q_{\mathcal{S}_j}, d_j$. The constants $q_{\mathcal{S}}$ are related to the average gate fidelity. See section 6.2 or Ref. [30] for the two-qubit example.

To eliminate SPAM errors in the fitting procedure, our choice of $f_{\mathbf{J},\mathbf{L},\mathcal{S}}$ is very specially selected to get rid of all $q_{\mathcal{S}'}$ in $\mathcal{E}^{\mathcal{G}}$, for which $\mathcal{S}' \neq \mathcal{S}$. To see how our choice of $f_{\mathbf{J},\mathbf{L},\mathcal{S}}$ accomplishes this, we will obtain a general expression for $\mathcal{R}_{\mathcal{S}}$, and show how it applies to the two qubit example. Let

$$U_{\mathbf{L}} = \prod_{s' \notin \mathcal{S}} f_{L_{s'}}^{(s')} \tag{6.10}$$

and

$$U_{\mathbf{J}} = \prod_{s,t \in \mathcal{S}} f_{J_{s,t}}^{(s)} f_{J_{s,t}}^{(t)} \tag{6.11}$$

so that $f_{\mathbf{J},\mathbf{L},\mathcal{S}} = U_{\mathbf{L}} U_{\mathbf{J}}$. For simplicity, assume that the subset $\mathcal{S}$ consists of the first $|\mathcal{S}|$ qubits. This is not necessary but will greatly simplify the derivation. Under this assumption, $U_{\mathbf{L}}$ can be written

$$\boldsymbol{U}_{\mathbf{L}} = \mathbb{I}^{\otimes |\mathcal{S}|} \otimes \boldsymbol{f}_{L_1} \otimes \boldsymbol{f}_{L_2} \otimes \ldots \otimes \boldsymbol{f}_{L_{n-|\mathcal{S}|}}. \tag{6.12}$$

Here we assume $\mathbb{I}$ is the single-qubit identity, $\mathbb{I}_4$ in the Pauli-Liouville representation. Given that tensor products of $f \in \mathcal{F}$ form a unitary 1-design, namely, the $n$-qubit Pauli group, the averaging over the group gives

$$|\mathcal{F}|^{-1} \sum_{\mathbf{L}} \boldsymbol{U}_{\mathbf{L}} = \mathbb{I}^{\otimes |\mathcal{S}|} \otimes (1 \oplus 0). \tag{6.13}$$

Next, recall that the matrix $\mathbf{J}$ is chosen to be strictly upper-triangular, and designate $f_0 = \mathbb{I}$. Writing only the non-identity terms, $U_{\mathbf{J}}$ can be expanded as

$$\begin{aligned} \boldsymbol{U}_{\mathbf{J}} = (\boldsymbol{f}_{J_{1,2}} \otimes \boldsymbol{f}_{J_{1,2}} \otimes \mathbb{I}^{\otimes|\mathcal{S}|-2} \otimes \mathbb{I}^{\otimes n-|\mathcal{S}|})(\boldsymbol{f}_{J_{1,3}} \otimes \mathbb{I} \otimes \boldsymbol{f}_{J_{1,3}} \otimes \mathbb{I}^{\otimes|\mathcal{S}|-3} \otimes \mathbb{I}^{\otimes n-|\mathcal{S}|}) \\ \ldots (\mathbb{I}^{\otimes|\mathcal{S}|-2} \otimes \boldsymbol{f}_{J_{|\mathcal{S}|-1,|\mathcal{S}|}} \otimes \boldsymbol{f}_{J_{|\mathcal{S}|-1,|\mathcal{S}|}} \otimes \mathbb{I}^{\otimes n-|\mathcal{S}|}). \end{aligned} \tag{6.14}$$

Averaging over the group $\mathcal{F}$ gives,

$$|\mathcal{F}|^{-1} \sum_{\mathbf{J}} \boldsymbol{U}_{\mathbf{J}} = |\mathcal{F}|^{-1} \left[ \sum_{J_{1,2}} (\boldsymbol{f}_{J_{1,2}} \otimes \boldsymbol{f}_{J_{1,2}} \otimes \mathbb{I}^{\otimes n-2}) \ldots \sum_{J_{|\mathcal{S}|-1,|\mathcal{S}|}} (\mathbb{I}^{\otimes|\mathcal{S}|-2} \otimes \boldsymbol{f}_{J_{|\mathcal{S}|-1,|\mathcal{S}|}} \otimes \boldsymbol{f}_{J_{|\mathcal{S}|-1,|\mathcal{S}|}} \mathbb{I}^{\otimes n-|\mathcal{S}|}) \right], \tag{6.15}$$

which can also be written as a matrix composed of blocks of identity matrices, thanks to the special properties of Paulis. We find after multiplying and averaging,

$$\begin{aligned} \boldsymbol{\mathcal{R}}_{\mathcal{S}} &= |\mathcal{F}|^{-2} \sum_{\mathbf{L}} \boldsymbol{U}_{\mathbf{L}} \sum_{\mathbf{J}} \boldsymbol{U}_{\mathbf{J}} \\ &= 1 \oplus 0 \oplus \mathbb{I}_3 \oplus 0. \end{aligned} \tag{6.16}$$

The location of the non-zero, $\mathbb{I}_3$ block depends on the ordering of the Pauli basis. It corresponds to non-identity, Pauli basis elements $P_i^{\otimes|\mathcal{S}|} \otimes \mathbb{I}^{\otimes n-|\mathcal{S}|}$, where $P_i^{\otimes|\mathcal{S}|} \in \{X^{\otimes|\mathcal{S}|}, Y^{\otimes|\mathcal{S}|}, Z^{\otimes|\mathcal{S}|}\}$.

To demonstrate the implications of this, we attest that the action of $\mathcal{R}_{\mathcal{S}}$ takes all blocks of the subsystem twirl to zero, except for the block corresponding to the set $\mathcal{S}$, and the indentity block, $S(\mathcal{E})$, which is always equal to 1 for CPTP maps. We show this for our 2 qubit example, but it is easy to see how it will apply generally, once a specific set of qubits is chosen. That is,

$$\mathcal{R}_{\mathcal{S}} \circ \left(\boldsymbol{\mathcal{E}}^{\mathcal{G}}\right)^m = 1 \oplus 0 \oplus q_S^m \mathbb{I}_3 \oplus 0. \tag{6.17}$$

Preparation and measurement errors will therefore introduce no unwanted dependence on $m$, and averaging over uniformly-random choices of $\mathbf{k}$, $\mathbf{J}$, and $\mathbf{L}$ for a fixed set $\mathcal{S}$ gives an estimate of

$$\mathbb{E}_{\mathbf{k},\mathbf{J}}(Q_{\mathbf{k},\mathbf{J},\mathcal{S}}) = A_{\mathcal{S}}(\mathcal{E}) q_{\mathcal{S}}^m + B_{\mathcal{S}}(\mathcal{E}), \tag{6.18}$$

where $A_{\mathcal{S}}$ and $B_{\mathcal{S}}$ are constants that incorporate SPAM errors.

## 6.5 Two Qubit Example, Continued

Continuing with the two qubit example from Section 6.2, we describe how to extract a particular $q_{\mathcal{S}}$. First, let $q_{\mathcal{S}} = q_{\mathcal{S}_{12}}$, the parameter corresponding to the set of both qubits. There are no qubits $s' \notin \mathcal{S}$ to consider, so we need only to calculate $|\mathcal{F}|^{-1} \sum_{\mathbf{J}} \boldsymbol{U}_{\mathbf{J}}$. The matrix $\mathbf{J}$ has only one non-zero entry so,

$$\boldsymbol{U}_{\mathbf{J}} = (\boldsymbol{f}_{J_{1,2}} \otimes \boldsymbol{f}_{J_{1,2}}). \tag{6.19}$$

Which we can easily take the average of, given that $\mathcal{F}$ is the Pauli group,

$$\begin{aligned} \boldsymbol{\mathcal{R}}_{\mathcal{S}_{12}} &= |\mathcal{F}|^{-1} \sum_{\mathbf{J}} \boldsymbol{U}_{\mathbf{J}} \\ &= |\mathcal{F}|^{-1} (\mathbb{I} \otimes \mathbb{I} + X \otimes X + Y \otimes Y + Z \otimes Z) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_3 \end{pmatrix}. \end{aligned} \tag{6.20}$$

Taking $\boldsymbol{\mathcal{E}}^{\mathcal{C}_1 \otimes \mathcal{C}_1}$ as in Eq. (6.1), the modified $\mathcal{C}_1 \otimes \mathcal{C}_1$-twirled channel is then

$$\mathcal{R}_{\mathcal{S}_{12}} \circ \boldsymbol{\mathcal{E}}^{\mathcal{C}_1 \otimes \mathcal{C}_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q_{12} I_3 \end{pmatrix}. \tag{6.21}$$

By preparing or measuring with support in $\mathcal{S}$, we can extract $q_{12}$, but any error in preparation or measurement will be absorbed by the constants $A_{\mathcal{S}}$ and $B_{\mathcal{S}}$, and will not effect the decay parameter.

Next, say we want to calculate the parameter for just one qubit in our two-qubit system. Let $q_{\mathcal{S}} = q_{\mathcal{S}_1}$, the parameter corresponding to qubit 1. There is an $s' \notin \mathcal{S}$ (qubit 2) so we will first calculate $|\mathcal{F}|^{-1} \sum_{\mathbf{L}} \boldsymbol{U}_{\mathbf{L}}$.

$$\boldsymbol{U}_{\mathbf{L}} = \mathbb{I} \otimes \boldsymbol{f}_{L_1}. \tag{6.22}$$

Averaging gives,

$$\begin{aligned} |\mathcal{F}|^{-1} \sum_{\mathbf{L}} \boldsymbol{U}_{\mathbf{L}} &= |\mathcal{F}_1|^{-1} (\mathbb{I} \otimes \mathbb{I} + \mathbb{I} \otimes X + \mathbb{I} \otimes Y + \mathbb{I} \otimes Z) \\ &= \mathbb{I} \otimes (1 \oplus 0_3). \end{aligned} \tag{6.23}$$

For a single qubit $\boldsymbol{U}_{\mathbf{J}}$ is the identity so,

$$\mathcal{R}_{\mathcal{S}_1} \circ \boldsymbol{\mathcal{E}}^{\mathcal{C}_1 \otimes \mathcal{C}_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & q_1 I_3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{6.24}$$

A similar calculation can be done for $q_{\mathcal{S}} = q_{\mathcal{S}_2}$.

## 6.6 Testing for Correlations

In analogy with the 2-qubit example, the values $q_{\mathcal{S}}$ can be used to test whether the experimental noise $\mathcal{E}$ is correlated as follows. The Clifford subsystem twirl $\mathcal{E}^{\mathcal{C}_1^{\otimes n}}$ of an uncorrelated channel $\mathcal{E} = \mathcal{E}_1 \otimes \ldots \otimes \mathcal{E}_n$, is simply

$$\mathcal{E}_1^{\mathcal{C}_1} \otimes \ldots \otimes \mathcal{E}_n^{\mathcal{C}_1} = \mathcal{D}_{p_1} \otimes \ldots \otimes \mathcal{D}_{p_n} \tag{6.25}$$

where $\mathcal{D}_p$ is the depolarizing channel with depolarizing parameter $p$. Therefore for any uncorrelated channel, set $\mathcal{S}$, and partitioning of $\mathcal{S}$ into $L$ disjoint sets $\mathcal{S}_1, \ldots, \mathcal{S}_L$,

$$q_{\mathcal{S}} - q_{\mathcal{S}_1} q_{\mathcal{S}_2} \ldots q_{\mathcal{S}_L} = 0. \tag{6.26}$$

If the estimates $q_{\mathcal{S}}$ for multiple sets $\mathcal{S}$, obtained by fitting the fidelity decay to Eq. (6.18), do not satisfy Eq. (6.26), then the experimental noise is correlated.

In the following section we will examine numerically simulated values of

$$\Delta(\mathcal{E}) = q_{\mathcal{S}}(\mathcal{E}) - q_{\mathcal{S}_1}(\mathcal{E}) q_{\mathcal{S}_2}(\mathcal{E}) \tag{6.27}$$

for various noise models to show how our protocol can be used to identify $n$-qubit errors in high-fidelity gates.

## 6.7 Simulation Results

Although uncorrelated noise by nature produces a test result of $\Delta(\mathcal{E}) = 0$, there are certainly possible cases of errors which will result in $\Delta(\mathcal{E}) = 0$, but where correlations are present. For example, the controlled-NOT gate, which is highly correlated, will test as uncorrelated. Fortunately, the counter-examples that we know of correspond to a very

high rate of error, and typically we are interested in implementations with high fidelity, or low error rate. However, as the gate fidelity approaches 1, the $q_{\mathcal{S}}$ also tend close to 1, and $\Delta(\mathcal{E})$, therefore becomes very small. By numerically and analytically examining several examples of error models, we demonstrate how the protocol will work in experiment, given a finite number of sequences, and demonstrate that for these error models, high fidelity counter-examples do not exist. We also give insight on the precision needed to detect non-zero $\Delta(\mathcal{E})$ for high-fidelity, correlated errors, based on the gate fidelity of $\mathcal{E}$. For this section we will set $\mathcal{S}_1$ and $\mathcal{S}_2$ to be single-qubit subspaces, and $\mathcal{S}$ to be the two-qubit space.

### 6.7.1 Example: Correlated Pauli Errors

Pauli channel errors are an important class of errors in quantum error correction, so we are interested in testing the protocol for this type of error. Consider a two-qubit noise model that with probability $p$, is a two-qubit Pauli channel without an identity component, and with probability $1-p$, is an uncorrelated noise channel, with probabilities $s_1$ and $s_2$ of Pauli errors on the individual qubits. This is an example of a correlated channel, provided $p \neq 0$ and $p \neq 1$.

The results of simulations of the symmetrization experiment for Pauli channel errors are displayed in Fig. 6.2. The resulting test parameter, $\Delta(\mathcal{E})$ is plotted in terms of $p$ in the stated Pauli error model for $p = [0, 0.05]$, since we are typically more concerned with small $p$. The number of repetitions was $k = 50$ per sequence length. The sequence lengths used were $m = 5, 10, 15, ..., 100$ (larger plot) and $m = 200, 400, ..., 4000$ (inset). The grey line is a curve showing the analytical values of $\Delta(\mathcal{E})$, and the error bars were determined from the 95% confidence intervals of the parameters. The numerically simulated data aligns well with the analytical curve, showing both the high sensitivity of the test, and the ability of the protocol to accurately estimate the test parameters.

We now derive an expression for the analytical curve in Fig. 6.2 based on the Pauli error model, and show how $\Delta(\mathcal{E})$ may be used to estimate the probability of two-qubit errors. For correlated Pauli errors, the error channel $\mathcal{E}$ takes the form

$$\mathcal{E}(\rho) = (1-p)\Lambda(\rho) + p\mathcal{F}(\rho) \tag{6.28}$$

where $\Lambda$ denotes Pauli-$X$ errors on the individual qubits,

$$\Lambda = ((1-s_1)\mathcal{I} + s_1 X) \otimes ((1-s_2)\mathcal{I} + s_2 X)\,. \tag{6.29}$$

In Fig. 6.2 the probabilities of Pauli errors on the individual qubits were constant and chosen as $s_1 = s_2 = 10^{-5}$. We introduce the uncorrelated, 2-qubit error channel $\mathcal{F}$, and

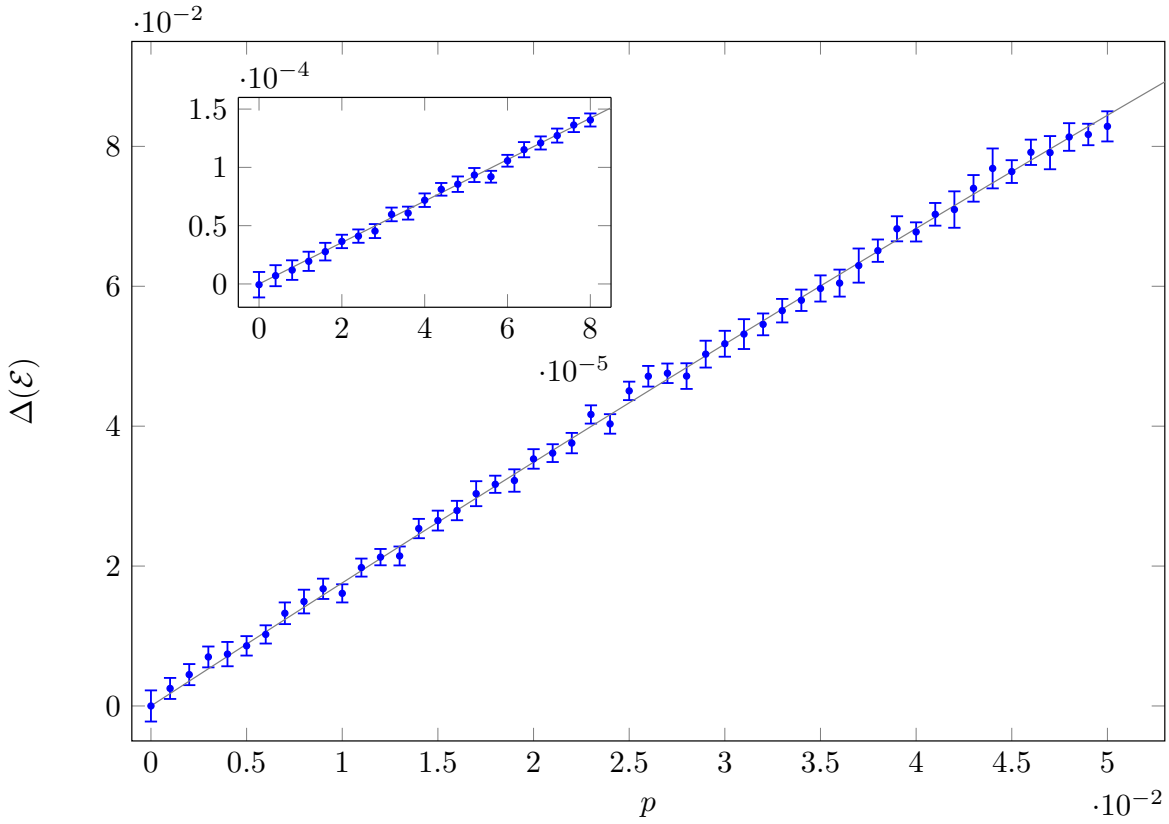



Figure 6.2: Plot showing the identification of correlations for classically-correlated Pauli errors by the symmetrized randomized benchmarking protocol, for up to error rates of $\sim 5\%$ (larger plot) and $\sim 0.01\%$ (inset).

since we are interested in Pauli errors, we set

$$\mathcal{F} = X \otimes X. \tag{6.30}$$

The channel $\mathcal{F}$ has $q_{\mathcal{S}_1}(\mathcal{F}) = -1/3$, $q_{\mathcal{S}_2}(\mathcal{F}) = -1/3$, and $q_{\mathcal{S}}(\mathcal{F}) = 1/9$. Equivalently, $\mathcal{F}$ could be any tensor product of two non-indentity Pauli operators.

Notice that the full channel $\mathcal{E}(\rho) = (1-p)\Lambda(\rho)+p\mathcal{F}(\rho)$ can be put into Kraus form with only uncorrelated Kraus operators $\sqrt{p_i}A_i$ as indicated in Table 6.1. This indicates that $\mathcal{E}$ is *classically-correlated*, that is, $\mathcal{E}$ can be written as a convex combination of uncorrelated channels.

| $A_i$ | $p_i$ |
|---|---|
| $\mathcal{I} \otimes \mathcal{I}$ | $(1-p)(1-s_1)(1-s_2)$ |
| $\mathcal{I} \otimes X$ | $(1-p)(1-s_1)s_2$ |
| $X \otimes \mathcal{I}$ | $(1-p)s_1(1-s_2)$ |
| $X \otimes X$ | $p+(1-p)s_1s_2$ |

Table 6.1: Kraus operators for correlated Pauli errors.

Using the two qubit Pauli basis, we can write the full error channel in the Pauli-Liouville representation, from which we can directly extract $q_{\mathcal{S}_1}(\mathcal{E})$, $q_{\mathcal{S}_2}(\mathcal{E})$, and $q_{\mathcal{S}}(\mathcal{E})$. We find that

$$\Delta(\mathcal{E}) = p_0 \left(p - p^2\right) (1 - q_{\mathcal{S}_1}(\mathcal{F}))(1 - q_{\mathcal{S}_2}(\mathcal{F})). \tag{6.31}$$

Without loss of generality we can assume the probability of no errors, $p_0 = 1 - O(r)$ where $r$ is the infidelity of $\Lambda$ to the identity, to show that $\Delta(\mathcal{E})$ can then be approximated by $p$ multiplied by a constant factor,

$$\begin{aligned} \Delta(\mathcal{E}) &= \left(p - p^2 - pO(r) + p^2O(r)\right) (1 - q_{\mathcal{S}_1}(\mathcal{F}))(1 - q_{\mathcal{S}_2}(\mathcal{F})) \\ &= p(1 - q_{\mathcal{S}_1}(\mathcal{F}))(1 - q_{\mathcal{S}_2}(\mathcal{F})) + O(p^2, pr, r^2). \end{aligned} \tag{6.32}$$

Given that $\mathcal{F}$ is a Pauli channel, $(1 - q_{\mathcal{S}_1}(\mathcal{F}))(1 - q_{\mathcal{S}_2}(\mathcal{F})) = 16/9$, and for high fidelity channels $p + (1-p)s_1s_2 \approx p$, so $\frac{9}{16}\Delta(\mathcal{E})$ gives a close estimate of the probability of a two-qubit error. Generally, for Pauli errors on $n$ qubits, it is analogous to see that $\Delta(\mathcal{E})$ can be used to approximate the probability of an error occurring on both partitions, $\mathcal{S}_1$ and $\mathcal{S}_2$.

### 6.7.2 Example: Correlated Dephasing

Single qubit dephasing noise is described by a channel which for a $Z$-error occurs with some probability $s$, and has Kraus operators $\{\sqrt{1-s}\mathbb{I}, \sqrt{s}Z\}$. We can express dephasing noise on two qubits by taking the tensor product of two the single qubit dephasing operators,

$$\Lambda = ((1-s_1)\mathcal{I} + s_1Z) \otimes ((1-s_2)\mathcal{I} + s_2Z)\,. \tag{6.33}$$

where $s_1$ and $s_2$ are the individual dephasing parameters for each qubit. This is another case of the description of uncorrelated Pauli errors on single-qubits in Eq. (6.29), where the Pauli channel is chosen as phase flip gate, $Z$.

To express correlated dephasing noise, we require additional Kraus terms which represent the correlations. From Ref. [46], the correlated dephasing channel $\mathcal{E}$ has the following (unnormalized) Kraus operators $\{\mathcal{I}\otimes\mathcal{I}, Z\otimes\mathcal{I}, \mathcal{I}\otimes Z, Z\otimes Z, \frac{1}{2}(iZ+\mathcal{I})\otimes(iZ+\mathcal{I}), \frac{1}{2}(i\mathcal{I}+Z)\otimes(i\mathcal{I}+Z)\}$. To model correlated dephasing, we can therefore add the last two Kraus operators to the tensor product of dephasing channels and write the full channel in the form: $\mathcal{E}(\rho) = (1-p)\Lambda(\rho) + p\mathcal{F}(\rho)$ where $\mathcal{F}$ corresponds to the sum of the final two Kraus operators. The full set of Kraus operators and coefficients is then given by Table 6.2.

| $A_i$ | $p_i$ |
|---|---|
| $\mathcal{I}\otimes\mathcal{I}$ | $(1-p)(1-s_1)(1-s_2)$ |
| $\mathcal{I}\otimes Z$ | $(1-p)(1-s_1)s_2$ |
| $Z\otimes\mathcal{I}$ | $(1-p)s_1(1-s_2)$ |
| $Z\otimes Z$ | $(1-p)s_1s_2$ |
| $(iZ+\mathcal{I})\otimes(iZ+\mathcal{I})$ | $p/2$ |
| $(i\mathcal{I}+Z)\otimes(i\mathcal{I}+Z)$ | $p/2$ |

Table 6.2: Kraus operators for correlated dephasing errors.

The channel $\mathcal{F}$ has $q_{\mathcal{S}_1}(\mathcal{F}) = 1/3$, $q_{\mathcal{S}_2}(\mathcal{F}) = 1/3$, and $q_{\mathcal{S}}(\mathcal{F}) = 1/9$. The test parameters of the full channel $\mathcal{E}$, in terms of the probabilities $p$, $s_1$, and $s_2$, is given by:

$$\Delta(\mathcal{E}) = (p-p^2)(1-2s_1)(1-2s_2)(1-q_{\mathcal{S}_1}(\mathcal{F}))(1-q_{\mathcal{S}_2}(\mathcal{F})). \tag{6.34}$$

Similarly to the Pauli error case, $\Delta(\mathcal{E})$ can be written as a constant times $p$, with a small perturbation, given high fidelity computations,

$$\Delta(\mathcal{E}) = p(1-q_{\mathcal{S}_1}(\mathcal{F}))(1-q_{\mathcal{S}_2}(\mathcal{F})) + O(p^2, pr, r^2). \tag{6.35}$$

Fig. 6.3 demonstrates the results of the test in a numerically simulated experiment, with slow dephasing. The value of the test parameter, $\Delta(\mathcal{E})$ is shown in terms of various probabilities $p$. The data points are the result of a simulated symmetrization experiment where the number of repetitions was $k = 50$ per sequence length. The sequence lengths chosen were $m = 5, 10, ..., 100$. The grey line is a curve showing the analytical values of $\Delta(\mathcal{E})$, and the error bars were determined from the 95% confidence intervals of the parameters. We also see that for small errors, $\Delta(\mathcal{E})$ may be used to estimate the probability $p$, which relates to the "amount" of correlation, given it is known that correlations are only present as correlated dephasing noise.

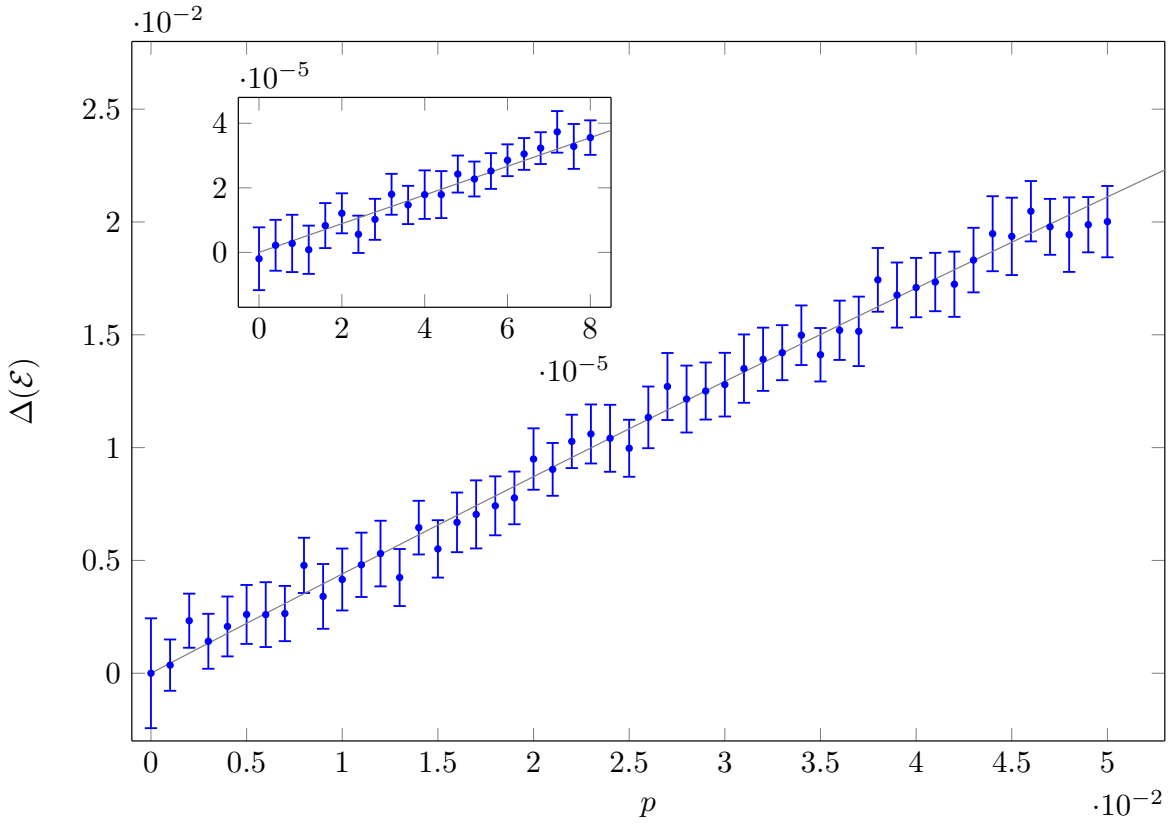



Figure 6.3: Plot showing the identification of correlations for correlated dephasing errors by the symmetrized randomized benchmarking protocol, for error rates up to $\sim 4\%$ (larger plot) and $\sim 0.009\%$ (inset).

### 6.7.3 Example: Small C-NOT Errors

As mentioned earlier, if the error on $\mathcal{G}$ is a controlled-NOT gate then the test for spatial correlations will fail. However, a more realistic scenario, which we explore in this section, is a small probability of controlled-NOT error.

Fig. 6.4 shows the value of the test parameter, $\Delta(\mathcal{E})$ in terms of various probabilities of C-NOT errors, $p$. The data points are the result of a simulated symmetrization experiment where the number of repetitions was $k = 50$ per sequence length. The sequence lengths chosen were $m = 5, 10, ..., 400$. The grey line is a curve showing the analytical values of $\Delta(\mathcal{E})$, and the error bars were determined from the 95% confidence intervals of the

parameters.

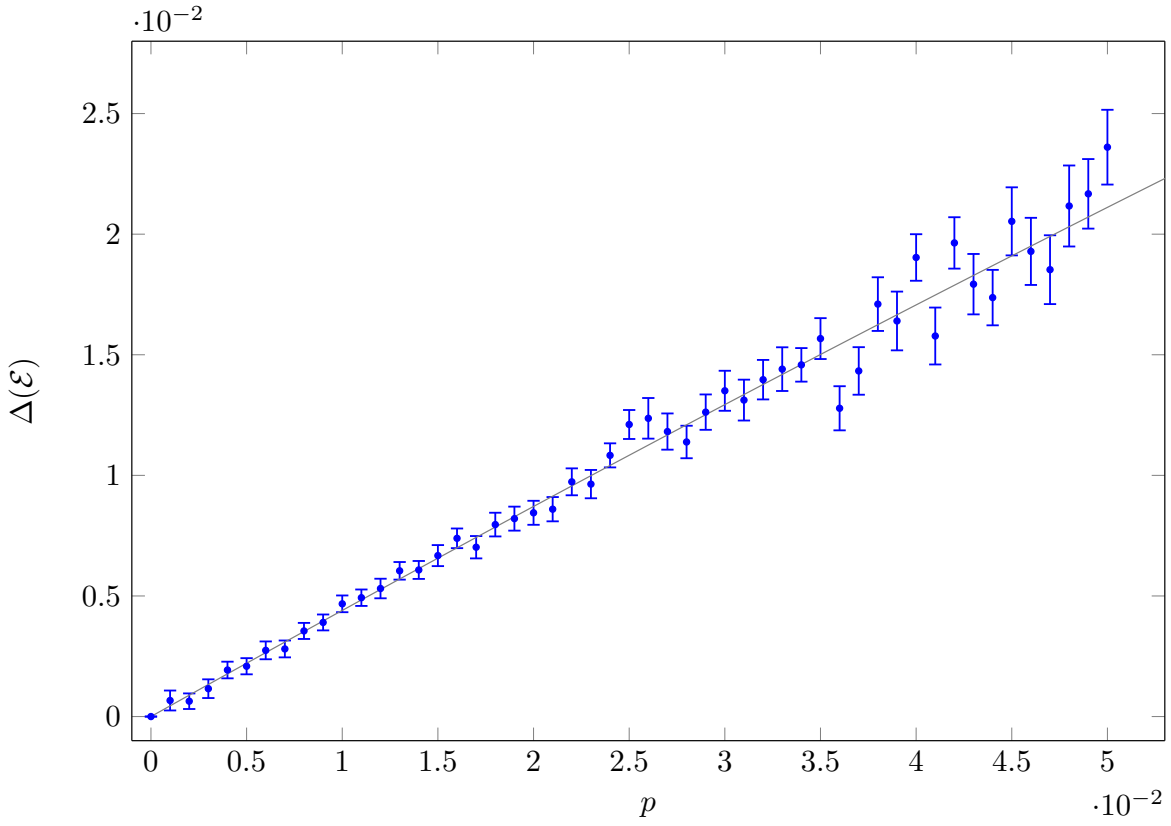



Figure 6.4: Plot showing identification of correlations for small probabilities of C-NOT, for probabilities $p$ of 0 to 0.05, by the symmetrized randomized benchmarking protocol.

We will model small C-NOT errors as a probability $p$ of a C-NOT gate being applied, otherwise, with probability $1-p$, the identity is applied. Based on this, the error channel can be written as the operator sum decomposition

$$\mathcal{E}(\rho) = (1-p)\rho + pCX\rho CX^{\dagger} \tag{6.36}$$

where $CX$ is the controlled-NOT or controlled-X gate. The channel $CX$ has $q_{\mathcal{S}_1}(CX) = 1/3$, $q_{\mathcal{S}_2}(CX) = 1/3$, and $q_{\mathcal{S}}(CX) = 1/9$.

By direct calculation, the test parameter is

$$\Delta(\mathcal{E}) = (p - p^2)(1 - q_{\mathcal{S}_1}(CX))(1 - q_{\mathcal{S}_2}(CX)), \tag{6.37}$$

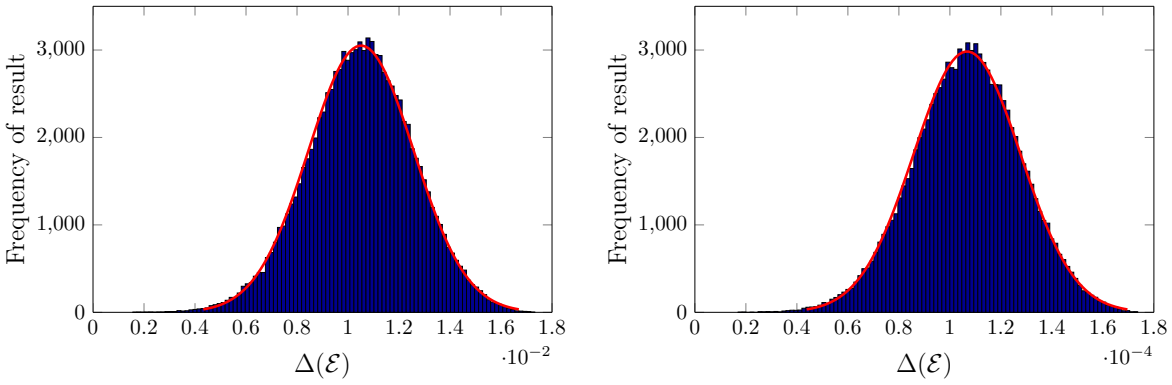
and given that $p$ is small,

$$\Delta(\mathcal{E}) = p(1 - q_{\mathcal{S}_1}(CX))(1 - q_{\mathcal{S}_2}(CX)) + O(p^2). \tag{6.38}$$

We see that if $p = 0$, then $\Delta(\mathcal{E}) = 0$, as expected, and if $p = 1$ we observe the failure of the test. For any non-zero $p \neq 1$, however $\Delta(\mathcal{E}) \neq 0$, meaning the test will be accurate. We, somewhat surprisingly, also see that for small $p$, given prior knowledge about the error model, estimating the test parameter allows a reasonable estimation of the probability of a $CX$ error, to within a constant factor.

### 6.7.4 Example: Correlated random small unitaries

Consider correlated two-qubit unitaries, which are close to identity. Such errors are small, but should produce $\Delta(\mathcal{E}) \neq 0$ since they contain correlations. To check the sensitivity of the test for this type of error, we generated a large number of random unitaries with equal fidelity to the identity, and compiled the values of $\Delta(\mathcal{E})$ taken directly from the individual unitaries in Fig. 6.5. The results approximate Gaussian distributions with means (a) $\mu = 0.0105$, (b) $\mu = 1.07\times10^{-4}$, and standard deviations (a) $\sigma = 2.06^{-3}$, (b) $\sigma = 2.10\times10^{-5}$ as found using MATLAB's histfit package, and demonstrated by the red fit curves.




(a) (b)

Figure 6.5: Histogram of test results from 100,000 randomly generated two-qubit small unitaries for two error rates; (a) $8.0 \times 10^{-3}$ and (b) $8.0 \times 10^{-5}$ from identity.

The results show that $\Delta(\mathcal{E})$ must be estimated with higher precision as the size of the error decreases. To further show this, in Fig. 6.6 we compiled the means of the approximate

Gaussian distributions obtained from the fitting to the data from 1000 random unitaries with error rates from the identity, $r = 1 - \overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}}$, where $\overline{\mathcal{F}_{\mathcal{E},\mathcal{I}}}$ is the average channel fidelity of $\mathcal{E}$ to the identity.

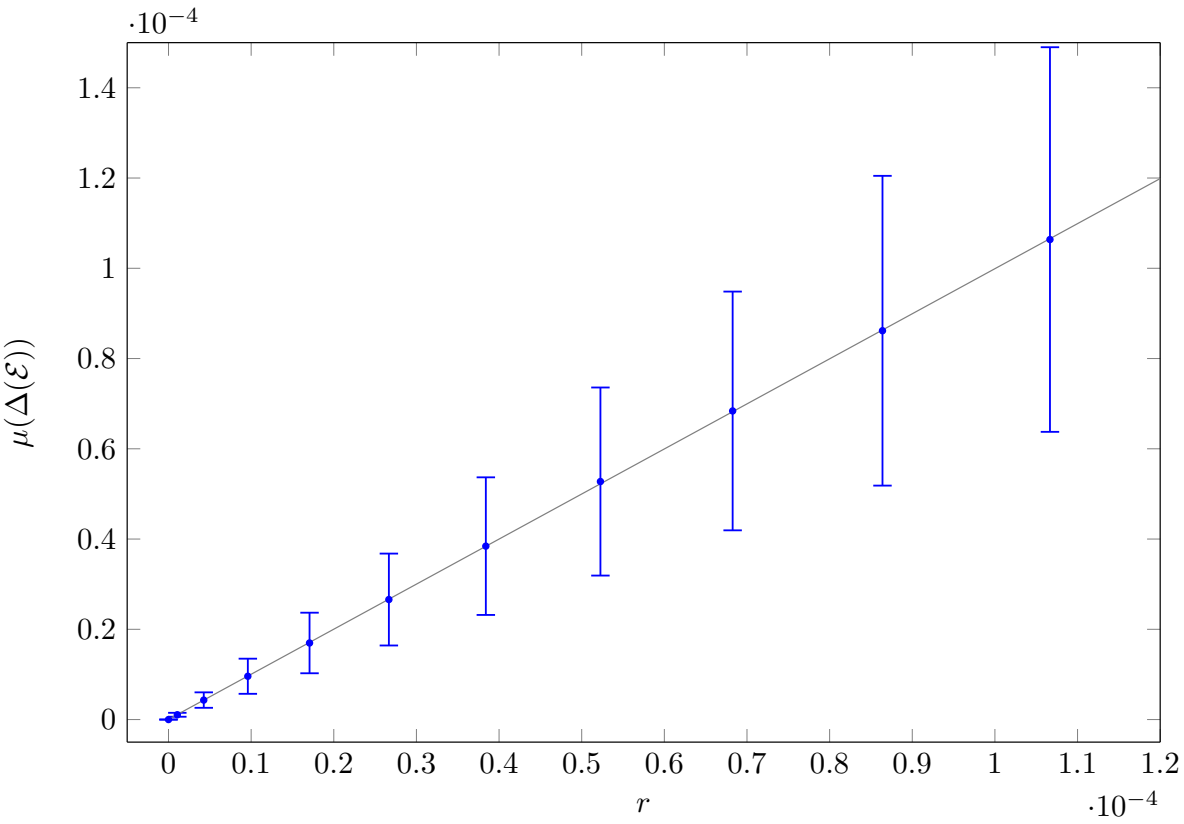



Figure 6.6: Plot of the mean $\Delta(\mathcal{E})$ for randomly generated small unitaries with error rates from identity $r$ ranging from 0 to $1.1 \times 10^{-4}$. Frequency data for 1000 random unitaries was used to determine $\mu(\Delta(\mathcal{E}))$ for each choice of $r$. The error bars represent the range within which 95% of $\Delta(\mathcal{E})$'s will fall for random unitary errors.

The error bars in Fig. 6.6 are two standard deviations, $\sigma$, and were also obtained by fitting to a Gaussian distributions. The data was fit using MATLAB's linear regression and gave $\mu(\Delta(\mathcal{E})) = 0.996r$ (grey line). The results therefore show numerically that $\Delta(\mathcal{E})$ gives a fair estimation of the infidelity for random unitary error, since for 95% of unitaries, $\Delta(\mathcal{E})$ will fall within $2\sigma$. More importantly, the result of this analysis indicates that information about the average fidelity (which can be found via RB) can be used to determine the a

threshold $\epsilon$ below which measuring $\Delta(\mathcal{E}) \leq \epsilon$ will indicate uncorrelated noise.

## 6.8 Summary

In this chapter, we have given a description of a scalable process which is robust to SPAM and combines a generalized version of symmetrization and randomized benchmarking to perform a check for correlated noise. We outlined the example in previous work [30] of the protocol for two qubits and resolved issues associated with this protocol. We have provided numerics to demonstrate the robustness of the test and the experimental protocol. We also state that for Pauli errors, this protocol may be used to determine the probability of two disjoint subsets of qubits both experiencing errors, in addition to the test of whether or not the errors are correlated in a particular computation, which suggests that this protocol may be very useful for quantum error correcting codes.

# Chapter 7

# Conclusion and Future Work

Randomization methods are known to provide important information on noisy quantum channels at relatively minimal cost. These methods exploit random sampling and properties of special groups in order to achieve this. The information that can be gained from these methods, at low cost is clearly not limited to the average gate fidelity. This thesis has shown three examples where randomization, group averaging and twirling can be used to obtain additional characteristics of noisy gates. We first introduced the concept of loss and loss rates, which correspond to probabilities of irretrievable loss of a qubit. We outlined step by step instructions for implementing a protocol to estimate loss rates by fitting a decay curve. We derived the fit model, and showed that the fit results provide additional information on detector efficiencies. We then introduced coherent leakage as a case where the randomized loss estimation protocol fails. Coherent leakage rates were derived, using steps analogous to the loss rate derivation. Instructions for estimating coherent leakage rates were presented as a slightly modified version of the instructions for estimating loss. The protocol requires some control over leakage levels. Finally, we introduced a generalized and robust protocol which tests for spatially correlated errors, using single-qubit Clifford twirls. Our protocol addresses issues and limitations on previous work on symmetrization schemes. We suggested that the revised version of this protocol may be very useful in practice, as some quantum error correcting codes require prior information on levels of correlations between qubits. For each protocol, we provided a full analysis of our models and numerical simulations of experiments, showing robust performance for realistic error models. These methods have arisen in a timely fashion, as many implementations of quantum processing are approaching the fault-tolerant threshold (though with a limited number of qubits), and will continue to provide a useful characterization toolkit as the number of realizable qubits increases.

The aim for the future is that these protocols will prove useful for experimentalists. The first two methods may be used to certify that leakage and/or loss does not exist (or only exists to a insignificant extent) in a particular implementation. If loss and/or leakage is present, the protocol can also be used by experimentalists to determine the rate of loss/leakage (and by extension the overhead required to correct it). For all three protocols we have assumed any errors are gate- and time-independent. Future work may include a proof of robustness of these protocols, in particular the leakage and correlations protocol, under gate- and time-dependent errors. Additionally, comparing our protocol for characterizing correlations to that by Laforest in Ref. [46] to determine for what situations which of these protocols is optimal, is a goal for the future. We also suggest that modifications on the protocol for benchmarking leakage may be useful as a future project to benchmark non-Markovian noise in general, which is currently an open problem. Future work related to the protocol for characterizing correlations could be to explore in depth how it can be used to select an optimal quantum error correcting code depending on assumptions on the noise model, and the results of the protocol. Simulations and analysis of the protocol for three or more qubits could provide initial insights into this problem. As is the case with RB, all three protocols can be used in the tuning of experimental instruments to minimize or eliminate errors, by performing repetitions of the experiments at various settings. The usefulness of randomization techniques is certainly not limited to these examples, other related protocols include Refs. [14, 72] and there are almost certainly many more applications of these for quantum error characterization yet to be realized.

# APPENDICES

# Appendix A

# Properties of the Trace

**Definition A.0.1.** The trace operation for an $N \times N$ matrix $A$ is defined by

$$\mathrm{Tr}A = \sum_{j}^{N} A_{jj}, \tag{A.1}$$

where $A_{jk}$ is the element in the $j^{th}$ row and $k^{th}$ column of $A$.

One useful property of the trace is its *linearity*, that is $\mathrm{Tr}(A+B) = \mathrm{Tr}A + \mathrm{Tr}B$.

*Proof.*

$$\begin{aligned} \mathrm{Tr}(A+B) &= \sum_{j,k}^{N} (A_{jj} + B_{kk}) \\ &= \sum_{j}^{N} A_{jj} + \sum_{k}^{N} B_{kk} \\ &= \mathrm{Tr}A + \mathrm{Tr}B. \end{aligned} \tag{A.2}$$

□

Also notice that

$$\begin{aligned}\mathrm{Tr}(\alpha A) &= \sum_{j}^{N} \alpha A_{jj} \\ &= \alpha \sum_{j}^{N} A_{jj} \\ &= \alpha \mathrm{Tr} A \end{aligned} \tag{A.3}$$

so, more generally,

$$\mathrm{Tr}\left(\sum_i \alpha_i A_i\right) = \sum_i \alpha_i \mathrm{Tr} A_i. \tag{A.4}$$

The trace also has a cyclic property. Suppose $A$ is an $n \times m$ matrix and $B$ is an $m \times n$ matrix then $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

*Proof.*

$$\begin{aligned}\mathrm{Tr}(AB) &= \sum_{k}^{n} (AB)_{kk}, \\ &= \sum_{i}^{n} \sum_{j}^{m} A_{ij} B_{ji} \\ &= \sum_{j}^{m} \sum_{i}^{n} B_{ji} A_{ij} \\ &= \mathrm{Tr}(BA). \end{aligned} \tag{A.5}$$

□

# Appendix B

# Tensor Products

The tensor product is an operation which combines two vector spaces to form one larger space. For a general $m \times n$ matrix $A$, and a general $p \times q$ matrix $B$ the tensor product $A \otimes B$ is defined as

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{pmatrix} \tag{B.1}$$

where

$$A_{ij}B = \begin{pmatrix} A_{ij}B_{11} & A_{ij}B_{12} & \dots & A_{ij}B_{1q} \\ A_{ij}B_{21} & A_{ij}B_{22} & \dots & A_{ij}B_{2q} \\ \vdots & \vdots & \vdots & \vdots \\ A_{ij}B_{p1} & A_{ij}B_{p2} & \dots & A_{ij}B_{pq} \end{pmatrix}. \tag{B.2}$$

Tensor products can be taken of quantum states or operators. The notation $A^{\otimes k}$ represents the tensor product of $A$ with itself $k$ times. That is $A^{\otimes k} = A \otimes A \otimes \ldots \otimes A$. For example, $A^{\otimes 3} = A \otimes A \otimes A$.

# Appendix C

# Proof of Theorem 1

To prove that the Clifford group on $n$ qubits $\mathcal{C}_n$ forms a 2-design, we prove the 2-design twirling condition, that is, for all $\rho$,

$$\frac{1}{|\mathcal{C}_n|}\sum_{C\in\mathcal{C}_n} C^\dagger\mathcal{E}(C\rho C^\dagger)C = \int_{U(d)} \mathrm{d}U U^\dagger\mathcal{E}(U\rho U^\dagger)U. \tag{C.1}$$

The original proof can be found in Ref. [17].

*Proof.* Let the map $\mathcal{E}$ be a linear map, and have the form

$$\mathcal{E}(\rho) = A\rho B \tag{C.2}$$

for $A, B \in L(\mathbb{C}^d)$. The 2-design twirling condition can then be alternatively expressed as

$$\frac{1}{|\mathcal{C}_n|}\sum_{C\in\mathcal{C}_n} C^\dagger AC\rho C^\dagger BC = \int_{U(d)} \mathrm{d}U U^\dagger AU\rho U^\dagger BU. \tag{C.3}$$

The RHS is the twirl over the bi-invariant Haar Measure, and from Ref. [23] can be expressed as

$$\int_{U(d)} \mathrm{d}U U^\dagger AU\rho U^\dagger BU = \frac{\mathrm{Tr}(AB)\mathrm{Tr}(\rho)}{d}\frac{\mathbb{I}}{d} + \frac{d\mathrm{Tr}(A)\mathrm{Tr}(B) - \mathrm{Tr}(AB)}{d(d^2-1)}\left(\rho - \mathrm{Tr}(\rho)\frac{\mathbb{I}}{d}\right). \tag{C.4}$$

This is also the depolarizing channel discussed in Eq. (2.32).

To evaluate the LHS, we first observe that the discrete Pauli-twirl of $\mathcal{E}$ maps the state $\rho$ to

$$\mathcal{E}^{\mathcal{P}_n}(\rho) = \frac{1}{d^2}\sum_j P_j A P_j \rho P_j B P_j. \tag{C.5}$$

The $P_j \in \mathcal{P}_n$ form a basis, so all $A$ and $B$ can be expressed by $A = \sum_{a=1}^{d^2} \alpha_a P_a$ and $B = \sum_{b=1}^{d^2} \beta_b P_b$. Then the Pauli-twirl can be expressed

$$\begin{aligned}
\mathcal{E}^{\mathcal{P}_n}(\rho) &= \frac{1}{d^2}\sum_{a=1}^{d^2}\sum_{b=1}^{d^2}\sum_j \alpha_a\beta_b P_j P_a P_j \rho P_j P_b P_j \\
&= \sum_{a=1} \alpha_a\beta_a P_a \rho P_a \\
&= \sum_{k=1} r_k P_k \rho P_k
\end{aligned} \tag{C.6}$$

by the special properties of Paulis, and where $r_1 = \mathrm{Tr}(A)\mathrm{Tr}(B)/d^2$ and $\sum_{k=1}^{d^2} r_k = \mathrm{Tr}(AB)/d$.

Every Clifford element $C \in \mathcal{C}_n$ can be expressed as $C = P_j Q_i$, where $P_j \in \mathcal{P}_n$ and the $Q_i$ are elements of $\mathcal{C}_n/\mathcal{P}_n$. the full Clifford twirl $\mathcal{E}^{\mathcal{C}_n}$ can then be expressed

$$\begin{aligned}
\frac{1}{|\mathcal{C}_n|}\sum_{i=1}^{|\mathcal{C}_n|/|\mathcal{P}_n|}\sum_{j=1}^{d^2} Q_i^\dagger P_j A P_j Q_i \rho Q_i^\dagger P_j B P_j Q_i &= \frac{1}{|\mathcal{C}_n|}\sum_{i=1}^{|\mathcal{C}_n|/|\mathcal{P}_n|}\sum_{j=1}^{d^2}\sum_{a=1}^{d^2}\sum_{b=1}^{d^2}\alpha_a\beta_b Q_i^\dagger P_j P_a P_j Q_i \rho Q_i^\dagger P_j P_b P_j Q_i \\
&= \frac{|\mathcal{P}_n|}{|\mathcal{C}_n|}\sum_{i=1}^{|\mathcal{C}_n|/|\mathcal{P}_n|}\sum_{k=1}^{d^2} r_k Q_i^\dagger P_k Q_i \rho Q_i^\dagger P_k Q_i.
\end{aligned} \tag{C.7}$$

To see that the RHS is equal to the LHS of Eq. (C.3), set apart the identity element $P_1 = \mathbb{I}$, and recall that the Clifford group is defined as the normalized of the $n$-qubit Pauli group, so that $Q_j^\dagger P_k Q_j \in \mathcal{P}_n, \forall j,k$, where non-identity Pauli elements are mapped with equal frequency to other non-identity Pauli elements under conjugation by a Clifford group map.

Therefore,

$$\begin{aligned}\mathcal{E}^{\mathcal{C}_n} &= r_1\rho + \frac{|\mathcal{P}_n|}{|\mathcal{C}_n|}\sum_{i=1}^{|\mathcal{C}_n|/|\mathcal{P}_n|}\sum_{k=2}^{d^2} r_k Q_i P_k Q_i^\dagger \rho Q_i P_k Q_i^\dagger \\ &= r_1\rho + \frac{1}{d^2-1}\left(\sum_{k=2}^{d^2} r_k\right)\sum_{l=2}^{d^2} P_l\rho P_l \\ &= r_1\rho + \frac{1}{d^2-1}\left(\sum_{k=1}^{d^2}(r_k) - r_1\right)\left(\sum_{l=1}^{d^2}(P_l\rho P_l) - \rho\right). \end{aligned} \tag{C.8}$$

Substituting the known expressions for $r_k$, and $\sum_{l=1}^{d^2} P_l\rho P_l = d\mathrm{Tr}(\rho)\mathbb{I}$ we find

$$\left(\frac{d\mathrm{Tr}(A)\mathrm{Tr}(B) - \mathrm{Tr}(AB)}{d(d^2-1)}\right)\rho + \left(\frac{d\mathrm{Tr}(AB) - \mathrm{Tr}(A)\mathrm{Tr}(B)}{d^2-1}\right)\mathrm{Tr}(\rho)\frac{\mathbb{I}}{d}. \tag{C.9}$$

which, with a bit of rearranging, is equal to the RHS of Eq. (C.4). □

# Appendix D

# Proof of Theorem 2

To prove Schur's lemma of the second form, we first state and prove Schur's lemma of the first form.

**Theorem 3.** *Schur's lemma (of the first form): Suppose we have two irreducible representations:*

$$\begin{aligned} \phi &: \mathcal{G} \to GL(V) \\ \phi' &: \mathcal{G} \to GL(W) \end{aligned} \tag{D.1}$$

*for a finite group $\mathcal{G}$. If the map $T : V \to W$ is an intertwining operator for $\phi$ and $\phi'$, then either $T = 0$ (the zero map), or $T$ is invertible.*

*Proof.* Suppose $T \neq 0$, then Ker$T$ is a proper subspace of $V$ and $\phi$ is an irrep, which implies Ker$T = \{0\}$. Therefore $T$ is injective.

And invertible map is injective and surjective, so it follows that if $T$ is not invertible, it must not be surjective. Suppose $T$ is not surjective, then Im$T$ is a proper subspace of $W$. However, because $\phi'$ is an irrep, Im$T = \{0\}$, which violates the $T \neq 0$ assumption, since Ker$T = \{0\}$. Therefore, $T$ must be surjective.

Combining these results we see that either $T$ is injective and surjective, and therefore invertible, or $T = 0$. □

We now prove theorem 2.

*Proof.* Suppose $T : V \to V$ is an intertwining operator for irrep $\phi : \mathcal{G} \to GL(V)$. $V$ is a complex vector space, therefore, $T$ has at least one eigenvalue $\lambda$. Let

$$T' = T - \lambda \mathbb{I}. \tag{D.2}$$

Since $T$ is an intertwining operator, $T'$ is also an intertwining operator;

$$\begin{aligned} \phi(T') &= \phi(T - \lambda \mathbb{I}) \\ &= \phi(T) - \phi(\lambda \mathbb{I}) \\ &= (T - \lambda \mathbb{I})(\phi) \\ &= T'(\phi). \end{aligned} \tag{D.3}$$

Note that Ker$T' \neq \{0\}$ since $T$ has at least one eigenvalue. Therefore $T'$ is not invertible so by Schur's lemma of the first form, $T' = 0$, and it follows that $T = \lambda \mathbb{I}$. □

# Appendix E

# List of abbreviations

| Abbreviation | Full Name |
|---|---|
| FTQC | Fault-Tolerant Quantum Computation |
| RB | Randomized Benchmarking |
| QIP | Quantum Information Processing |
| ON | Orthonormal |
| CPTP | Completely Positive, Trace-Preserving |
| PVM | Projection Valued Measurement |
| POVM | Positive-Operator Valued Measurement |
| SPAM | State Preparation and Measurement |
| NMR | Nuclear Magnetic Resonance |

Table E.1: List of abbreviations used and their full meanings.

# References


[1] D. Aharonov and M. Ben-Or. Fault-Tolerant Quantum Computation with Constant Error Rate. *SIAM Journal on Computing*, 38(4):12071282, 1999.

[2] P. Aliferis and B. M. Terhal. Fault-tolerant quantum computation for local leakage faults. *Quantum Information & Computation*, 7:139, 2007.

[3] A. Ambainis, M. Mosca, A. Tapp, and R. De Wolf. Private quantum channels. *IEEE 54th Annual Symposium on Foundations of Computer Science*, 2000.

[4] H. Ball, T. M. Stace, S. T. Flammia, and M. J. Biercuk. The effect of noise correlations on randomized benchmarking. *eprint: arXiv:quant-ph/1504.05307v1*, 2015.

[5] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and John M. Martinis. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature*, 508(7497):500–3, 2014.

[6] S. D. Barrett and T. M. Stace. Fault tolerant quantum computation with very high threshold for loss errors. *Physical Review Letters*, 105(20):1–4, 2010.

[7] I. Bengtsson and K. Zyczkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, Cambridge, UK, 2006.

[8] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824, 1996.

[9] R. Bhatia. *Matrix Analysis*. Springer, 1997.

[10] R. B. Blakestad, C. Ospelkaus, A. P. Vandevender, J. H. Wesenberg, M. J. Biercuk, D. Leibfried, and D. J. Wineland. Near-ground-state transport of trapped-ion qubits through a multidimensional array. *Physical Review A - Atomic, Molecular, and Optical Physics*, 84(3):032314, 2011.

[11] R. Blatt, H. Häffner, C. F. Roos, C. Becher, and F. Schmidt-Kaler. Ion Trap Quantum Computing with Ca+ Ions. *Quantum Information Processing*, 3(1-5):61–73, October 2004.

[12] K. R. Brown, A. C. Wilson, Y. Colombe, C. Ospelkaus, A. M. Meier, E. Knill, D. Leibfried, and D. J. Wineland. Single-qubit-gate error below 10-4 in a trapped ion. *Physical Review A*, 84(4):030303, 2011.

[13] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098–1105, 1996.

[14] A. Carignan-Dugas, J. J. Wallman, and J. Emerson. Using randomized benchmarking outside 2-designs. *In progress*, 2015.

[15] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Randomized benchmarking and process tomography for gate errors in a solid-state qubit. *Physical Review Letters*, 102(4):090502, 2009.

[16] I. L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455, 1997.

[17] C. Dankert, R. Cleve, J. Emerson, and E. Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Physical Review A*, 80(1):012304, 2009.

[18] S. J. Devitt, S. G. Schirmer, D. K. L. Oi, J. H. Cole, and L. C. L. Hollenberg. Subspace confinement: how good is your qubit? *New Journal of Physics*, 9(10):384–384, 2007.

[19] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Demonstration of two-qubit algorithms with a superconducting quantum processor. *Nature*, 460(7252):240, 2009.

[20] L. DiCarlo, M. D. Reed, L. Sun, B. R. Johnson, J. M. Chow, J. M. Gambetta, L. Frunzio, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf. Preparation and measurement of three-qubit entanglement in a superconducting circuit. *Nature*, 467(7315):574, 2010.

[21] J. Emerson. Random and pseudo-random states and operators: Constructions and applications. University Lecture, 2004.

[22] J. Emerson. Theory of open quantum systems with applications. University Lecture, 2015.

[23] J. Emerson, R. Alicki, and K. Życzkowski. Scalable noise estimation with random unitary operators. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(10):S347, 2005.

[24] J. Emerson, M. Silva, O. Moussa, C. A. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme. Symmetrized characterization of noisy quantum processes. *Science*, 317(5846):1893–6, 2007.

[25] J. M. Epstein, A. W. Cross, E. Magesan, and J. M. Gambetta. Investigating the limits of randomized benchmarking protocols. *Physical Review A*, 89(6):062321, 2014.

[26] B. Fortescue, S. Nawaf, and M. Byrd. Fault-tolerance against loss for photonic FTQEC. *e-print: arXiv:quant-ph/1405.1766v1*, 2014.

[27] A. G. Fowler. Coping with qubit leakage in topological codes. *Physical Review A*, 88(4):042308, 2013.

[28] C. A. Fuchs and J. Van De Graaf. Cryptographic Distinguishability Measures for Quantum-Mechanical States. *Information Theory, IEEE Transactions on*, 45(4):1216, 1999.

[29] W. Fulton and J. Harris. *Representation theory. A first course.* Graduate Texts in Mathematics. Springer.

[30] J. M. Gambetta, A. D. Córcoles, S. T. Merkel, B. R. Johnson, J. A. Smolin, J. M. Chow, C. A. Ryan, C. Rigetti, S. Poletto, T. A. Ohki, M. B. Ketchen, and M. Steffen. Characterization of addressability by simultaneous randomized benchmarking. *Physical Review Letters*, 109(24):240504, 2012.

[31] J. Ghosh, A. G. Fowler, J. M. Martinis, and M. R. Geller. Understanding the effects of leakage in superconducting quantum-error-detection circuits. *Physical Review A*, 88(6):062329, 2013.

[32] R. Goodman and N. R. Wallach. *Symmetry, Representations, and Invariants.* Graduate Texts in Mathematics. Springer, 2009.

[33] D Gottesman. Stabilizer codes and quantum error correction. 2008.

[34] D. Gottesman. The Heisenberg Representation of Quantum Computers. *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics, eds. S. P. Corney, R. Delbourgo, and P. D. Jarvis*, pages 32–43, Cambridge, MA, International Press, 1999.

[35] C. Granade, C. Ferrie, and D. G. Cory. Accelerated Randomized Benchmarking. *New J. Phys.*, 17(1):013042, 2014.

[36] D. Gross, K. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. *Journal of Mathematical Physics*, 48(5):052104, 2007.

[37] T. F. Havel. Robust procedures for converting among Lindblad, Kraus and matrix representations of quantum dynamical semigroups. *Journal of Mathematical Physics*, 44(2):534 – 557, 2014.

[38] D. Hogg, D. W. Berry, and A. I. Lvovsky. Efficiencies of Quantum Optical Detectors. *Physical Review A*, 90(5):053846, 2014.

[39] M. Horodecki, P. Horodecki, and R. Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Physical Review A*, 60(3):1888, 1999.

[40] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I. C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis. Optimal quantum control using randomized benchmarking. *Physical Review Letters*, 112:240504, 2014.

[41] S. Kimmel, M. P. da Silva, C. A. Ryan, B. R. Johnson, and T. Ohki. Robust extraction of tomographic information via randomized benchmarking. *Physical Review X*, 4:011050, 2014.

[42] A. Y. Kitaev. Quantum error correction with imperfect gates. *Quantum Communication, Computing, and Measurement (Proc. 3rd Int. Conf. of Quantum Communication and Measurement) (Plenum Press, New York, 1997)*, pages 181 – 188, 1997.

[43] E. Knill. Fault-tolerant postselected quantum computation: Schemes. *e-print: arXiv:quant-ph/0402171*, 2004.

[44] E. Knill, R. Laflamme, and W. H. Zurek. Resilient quantum computation: Error models and thresholds. *Science*, 279:342345, 1998.

[45] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. *Physical Review A*, 77(1):012307, 2008.

[46] M. Laforest. *Error characterization and quantum control benchmarking in liquid state NMR using quantum information processing techniques*. PhD thesis, University of Waterloo, 2008.

[47] C. C. López, A. Bendersky, J. P. Paz, and D. G. Cory. Progress toward scalable tomography of quantum maps using twirling-based methods and information hierarchies. *Physical Review A*, 81(6):062113, 2010.

[48] E. Magesan. Gaining information about a quantum channel via twirling. Master's thesis, University of Waterloo, 2008.

[49] E. Magesan, R. Blume-Kohout, and J. Emerson. Gate fidelity fluctuations and quantum process invariants. *Physical Review A*, 84(1):012309, 2011.

[50] E. Magesan, J. M. Gambetta, and J. Emerson. Scalable and Robust Randomized Benchmarking of Quantum Processes. *Physical Review Letters*, 106(18):180504, 2011.

[51] E. Magesan, J. M. Gambetta, and J. Emerson. Characterizing quantum gates via randomized benchmarking. *Physical Review A*, 85(4):042311, 2012.

[52] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen. Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking. *Physical Review Letters*, 109(8):080505, 2012.

[53] J. Medford, J. Beil, J. M. Taylor, S. D. Bartlett, A. C. Doherty, E. I. Rashba, D. P. DiVincenzo, H. Lu, A. C. Gossard, and C. M. Marcus. Self-consistent measurement and state tomography of an exchange-only spin qubit. *Nature nanotechnology*, 8(9):654–9, 2013.

[54] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen. Self-consistent quantum process tomography. *Physical Review A*, 87(6):062119, 2013.

[55] M. Mičuda, I. Straka, M. Miková, M. Dušek, N. J. Cerf, J. Fiurášek, and M. Ježek. Noiseless loss suppression in quantum optical communication. *Physical Review Letters*, 109(18), 2012.

[56] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. Wilhelm. Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits. *Physical Review Letters*, 103(11):110501, 2009.

[57] M. A. Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, 303(4):249, 2002.

[58] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, New York, 1st edition, 2000.

[59] P. Pechukas. Reduced Dynamics Need Not Be Completely Positive. *Physical Review Letters*, 73(8):1060, 1994.

[60] J. Poyatos, J. I. Cirac, and P. Zoller. Complete Characterization of a Quantum Process: The Two-Bit Quantum Gate. *Physical Review Letters*, 78(2):390, 1997.

[61] B. W. Reichardt. Improved magic states distillation for quantum universality. *e-print: arXiv:quant-ph/0411036*, 2004.

[62] C. A. Ryan, M. Laforest, and R. Laflamme. Randomized benchmarking of single- and multi-qubit control in liquid-state nmr quantum information processing. *New Journal of Physics*, (18):013034, 2009.

[63] P. Schindler, D. Nigg, T. Monz, J. T. Barreiro, E. Martinez, S. X. Wang, S. Quint, M. F. Brandl, V. Nebendahl, C. F. Roos, M. Chwalla, M. Hennrich, and R. Blatt. A quantum information processor with trapped ions. *New Journal of Physics*, 15, 2013.

[64] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, (5):1484–1509, 1997.

[65] B. Simon. *Representations of Finite and Compact Groups (Graduate Studies in Mathematics ; V. 10)*. American Mathematical Society, Providence, RI, 1996.

[66] A. M. Steane. Error Correcting Codes in Quantum Theory. *Physical Review Letters*, 77:793–797, 1996.

[67] J. Vala, K. B. Whaley, and D. S. Weiss. Quantum error correction of a qubit loss in an addressable atomic system. *Physical Review A - Atomic, Molecular, and Optical Physics*, 72(5):052318, 2005.

[68] M. Varnava, D. E. Browne, and T. Rudolph. Loss tolerance in one-way quantum computation via counterfactual error correction. *Physical Review Letters*, 97(12):120501, 2006.

[69] J. J. Wallman, M. Barnhill, and J. Emerson. Robust Characterization of Leakage Errors. *e-print: arXiv:quant-ph/1412.4126v2*, 2015.

[70] J. J. Wallman, M. Barnhill, and J. Emerson. Robust characterization of loss rates. *Physical Review Letters*, 115:060501, 2015.

[71] J. J. Wallman and S. T. Flammia. Randomized benchmarking with confidence. *New Journal of Physics*, 16(10):103032, 2014.

[72] J. J. Wallman, C. Granade, R. Harper, and S. T. Flammia. Estimating the Coherence of Noise. *e-print: arXiv:quant-ph/1503.07865v1*, 2015.

[73] J. Watrous. Simpler semidefinite programs for completely bounded norms. *e-print: arXiv:quant-ph/1207.5726v2*, 2012.

[74] A. C. Whiteside and A. G. Fowler. Upper bound for loss in practical topological-cluster-state quantum computing. *Physical Review A*, 90(5):052316, 2014.

[75] K. Wright, J. M. Amini, D. L. Faircloth, C. Volin, S. C. Doret, H. Hayden, C. S. Pai, D. W. Landgren, D. Denison, T. Killian, R. E. Slusher, and A. W. Harter. Reliable transport through a microfabricated X-junction surface-electrode ion trap. *New Journal of Physics*, 15:033004, 2013.

[76] Z. Zhou, S.-I. Chu, and S. Han. Rapid Optimization of Working Parameters of Microwave-Driven Multilevel Qubits for Minimal Gate Leakage. *Physical Review Letters*, 95(12):120501, 2005.