



國立臺灣大學理學院物理學系  
碩士論文

Department of Physics  
College of Science  
National Taiwan University  
Master Thesis

基於原始閘用來刻劃量子閘保真度的隨機標竿分析法  
Primitive-gate based randomized benchmarking method for  
characterizing quantum gate fidelity

徐義竣

I-Chun Hsu

指導教授：管希聖 博士

Advisor: Hsi-Sheng Goan, Ph.D.

中華民國 110 年 2 月

February 2021



# Acknowledgements

Thank you all very much.



## 摘要

隨著越來越多的量子位元能使用，量子邏輯閘的錯誤率變成急迫的議題。現今有許多刻劃量子邏輯閘保真度的協議，像是量子過程解析。然而，因為量子過程解析需要的資源成指數成長隨著量子位元數目變大，所以量子過程解析不在適用。此外，量子過程解析無法區分態準備錯誤與測量錯誤與量子閘錯誤。隨機標竿分析法能夠區分態準備錯誤與測量錯誤與量子閘錯誤且所需的資源少於量子過程解析。現今，隨機標竿分析法比起量子過程解析被更頻繁的使用。然而，隨機標竿分析法在實驗中有些不同於理論中的要求。也就是說，理論上的隨機標竿分析法假設在實際實驗中是不實際的。我們在這篇論文中分析理論以及實驗的隨機標竿分析法差別且介紹新的方法叫做基於原始閘的隨機標竿分析法。我們的基於原始閘的隨機標竿分析法能夠更貼近實驗上的結果。透過數值模擬，我們展現我們的基於原始閘的隨機標竿分析法比起理論隨機標竿分析法以及實驗隨機標竿分析法有更多優勢。 **關鍵字：量子過程解析，隨機標竿分析法，量子邏輯閘操作，量子閘保真度**



# Abstract

Recently with more and more qubits becoming available and functional, the errors on qubit operations become an urgent issue. There are many protocols for characterizing quantum gate operations such as quantum process tomography (QPT). However, when the number of qubits become larger, QPT is no longer appropriate due to its required resources exponentially growing with qubit size. In addition, QPT mixes the state preparation and measurement (SPAM) errors with the gate errors. Instead, randomized benchmarking (RB), another method for estimating the gate error rate, can separate the SPAM errors from the gate errors. Moreover, the resource requirement by RB is less than that by QPT. Nowadays, RB are frequently used to characterize quantum gate operations. However, the situations to perform RB in experiments is a little bit different from the RB theory requirements. That is, the assumptions in some RB theory are impractical in the real experiments. We will analyse the discrepancies and introduce our new method based RB, called primitive randomized benchmarking (PRB). Our new method, PRB, can be made closer to the real RB experiments. Through numerical simulations, we show that our proposed PRB has several appealing advantages over RB and RB with decomposition into primitive gates.

**Keywords:** Quantum process tomography, Randomized benchmarking, Quantum gate operations, Gate fidelity



# Contents

Acknowledgments	I
摘要	II
Abstract	III
List of Figures	VI
List of Tables	VII
<b>1 Introduction</b>	<b>1</b>
<b>2 Mathematical background</b>	<b>4</b>
2.1 Pauli group and Clifford group . . . . .	4
2.2 Average gate fidelity . . . . .	7
2.3 Unitary 2-Design . . . . .	9
2.4 Quantum channel . . . . .	10
2.4.1 Depolarizing Channel . . . . .	11
2.4.2 Dephasing Channel . . . . .	12
<b>3 Randomized Benchmarking</b>	<b>13</b>
3.1 Randomized Benchmarking . . . . .	13
3.1.1 Advantage of using RB . . . . .	14
3.2 PRB . . . . .	16
3.2.1 Advantages of PRB over RB . . . . .	17
3.3 DRB . . . . .	19

3.3.1	A comparison between PRB and DRB . . . . .	19
<b>4</b>	<b>Protocols</b>	<b>21</b>
4.1	Protocols . . . . .	21
4.1.1	Protocol of RB . . . . .	21
4.1.2	Protocol of RB with decomposition . . . . .	22
4.1.3	Protocol of PRB . . . . .	23
4.2	Effect of the noise . . . . .	24
4.2.1	Noisy channel as a map . . . . .	25
4.2.2	Noise in the Hamiltonian . . . . .	25
<b>5</b>	<b>Simulation results</b>	<b>29</b>
5.1	Results of rescaling . . . . .	29
5.1.1	Depolarizing channel . . . . .	29
5.1.2	Dephasing channel . . . . .	40
5.2	Noise from the Hamiltonian . . . . .	41
5.2.1	Good model verified . . . . .	41
5.2.2	PRB work . . . . .	41
5.2.3	PRB under small repetition . . . . .	42
<b>6</b>	<b>Conclusions</b>	<b>46</b>
<b>A</b>	<b>Appendix Title</b>	<b>47</b>
	<b>Bibliography</b>	<b>48</b>





# List of Figures

2.2	<b>Bloch sphere.</b> The outer surface represents all the pure states on a Bloch sphere, and the origin is a maximally mixed state. . . . .	9
2.3	<b>The schematic diagram of Haar unitary</b> [1]. (a) all the Haar random unitary elements in the red region. (b) a distribution of the unitary design of elements at the red dots. (c) We can design a system containing the specific Hamiltonian. (d) If the unitary design set is a group, each element have its inverse one. That is, each element can go back to identity. . . . .	10
2.4	<b>Single qubit depolarizing channel</b> [2]. It shows the effect of the depolarizing channel on the Bloch sphere of a qubit system. . . . .	11
2.5	<b>Single qubit dephasing channel</b> [2]. It shows the effect of the dephasing channel on the Bloch sphere of a qubit system. . . . .	12
3.1	<b>Randomized benchmarking sequence</b> . . . . .	13
4.1	<b>Silicon qubit device</b> [3] . . . . .	25
4.2	<b>Primitive gate implantation</b> [3] . . . . .	26



# List of Tables

2.1	<b>1-qubit Clifford group [4]</b> $\theta$ is the rotation angle. $n_x$ , $n_y$ and $n_z$ correspond to the weighted direction of rotation around $x$ -axis, $y$ -axis and $z$ -axis, respectively. . . . .	6
2.2	<b>Number of elements in the Clifford group <math>C_n</math></b> . . . . .	7
3.1	<b>The decomposition table of the single qubit Clifford gates [5]</b>	18
5.1	<b>Simulation results of the value of <math>r</math>, <math>p</math>, <math>A</math> and <math>\alpha</math>.</b> . . . . .	41



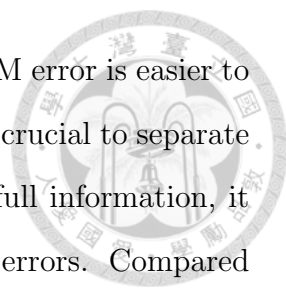




# Chapter 1

## Introduction

There is no perfect thing in the world, neither do the computers. Classical computers have errors, but they seem to work well. The secret is the error corrections. The first step to correct the errors is to characterize the errors. Therefore, we follow the classical procedure to first characterize the errors when we build quantum computers. However, quantum computers are affected by noise much more easily than classical computers because quantum states are much more fragile than classical one, and even the heat fluctuation can destroy them at low temperature. One idea is characterizing the noise as detailed as possible so that we can correct the tiny errors caused by a small fluctuation. To follow the idea, quantum process tomography (QPT) [2] is a method which can obtain the full information about the noise on a process or a quantum gate. However, more details always come with more resources, which makes QPT not friendly to experiments, not to mention that its resource requirements grow exponentially with qubit size. Another way to characterize the effect of quantum noise on a process or a quantum gate operation but not to specify all the details of noise is the so-called randomized benchmarking (RB) [6]. Because RB uses averaging to evaluate the effect of the noise, it requires less resources than QPT. Another issue is that we may sometimes want to know only the gate errors. However, running an experiment of a quantum circuit consisting of quantum gates usually requires three steps: state preparations, gate operations and measurements. The errors, occurring on the first and last steps, are called the state



preparations and measurement (SPAM) errors. Because the SPAM error is easier to characterize, we only discuss issues on the gate errors. Thus, it is crucial to separate the SPAM errors from the gate errors. Although QPT can get full information, it mixes these errors. That is to say, it is sensitive to the SPAM errors. Compared to QPT, RB is free from the SPAM errors due to its protocol design, which we will discuss later. To visualize QPT, we can imagine the noise as a black box. For different input states, the black box give us different corresponding output states. The idea of QPT is that we characterize the black box by observing the output states for different input states which were sent into the black box. Once we know the details of the black box, we can predict the output state for any known input state. To achieve the goal, we must send a lot of input states into the black box and observe the corresponding output states. Knowing the effect of the black box on the input states, we can calculate its effect on any other states. However, QPT is a hard experimental procedure. If we have an  $N$ -dimensional Hilbert space, the number of the set  $\{\rho_i\}$  of the required input states is  $N^2$ . For each  $\rho_i$ , we have to perform quantum state tomography to obtain the output state, i.e. determine the density matrix consisting of  $N^2$  parameters. This implies that full quantum process tomography needs to determine  $N^4$  numbers. For example, a process on single qubit ( $N = 2$ ) requires acquiring 16 data sets! This is why quantum process tomography has so far been limited to very few qubits.

Another way to characterize the effect of the noise is RB, which is thought to be more promising when facing more qubits. Because it uses average error value rather the whole information of the noise, RB has only one relevant parameter to estimate for the gate error rate, in contrast to  $N^4$  parameters of QPT. Another advantage of RB is that it can separate the gate error from the SPAM error, whereas QPT cannot. The RB protocol assumes that the preparation and measurement errors are the same for every experiment. When the number of gates increases, the error rate also increases. Thus, we can contribute this increasing error rate to the gate operations. RB has only one parameter, called the depolarizing strength, which

can be obtained from the decay of the survival probability with the Clifford gate number. If the survival probability decays too fast, the uncertainty in fitting will become too large to get the exact depolarizing strength. The experiment of two-qubit gates in [7] reported the above situation. To deal with the problem, a method, called character randomized benchmarking, has been proposed [8].

Many protocols based on RB have been developed to solve different problems, such as interleaved RB [9] and dihedral benchmarking [10]. We propose a new method, called primitive RB (PRB), to deal with the rescaling problem. When the number of qubits become larger, the number of Clifford gates also becomes substantially large. This may cause the finite size bias or effect in a real experiment. Moreover, each gate is composed of many primitive gates in a real experiment. When the number of qubits becomes larger, the number of primitive gates to compose each Clifford gate increases. Thus, the survival will decay fast using the original RB. Our PRB tries to solve the two problems mentioned above.

This thesis is organized as follows. In Chap. 2, we introduce some mathematical concepts and tools that are useful for describing and understanding the technique of RB. In Chap. 3, we describe the main ideas and advantages of RB and PRB. In Chap. 4, we present the protocols of how we perform simulations and describe the noise models we use. In Chap. 5, we describe two problems which the original RB face and present simulation results to show that our PRB can work better. In Chap. 6, we summarize what we have done and discuss the possible future works.



## Chapter 2

# Mathematical background

In this chapter, some useful mathematical concepts and background essential for RB are introduced.

### 2.1 Pauli group and Clifford group

The simplest Pauli group  $\mathbf{P}_1$  on single qubit contains 16-element matrix group consisting of the  $2 \times 2$  identity matrix and all of the Pauli matrices

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{X} = \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (2.1)$$

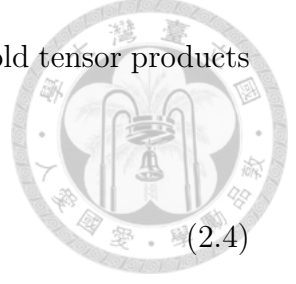
$$\mathbf{Y} = \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \mathbf{Z} = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (2.2)$$

combined with four multiplicative factors  $\{+1, -1, +i, -i\}$  with matrix multiplication as a group operation. Thus, the total 16-element  $\mathbf{P}_1$  are

$$\mathbf{P}_1 = \{\pm\mathbb{I}, \pm i\mathbb{I}, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}. \quad (2.3)$$

The general Pauli group  $\mathbf{P}_n$  on  $n$  qubits is generated from all  $n$ -fold tensor products of Pauli matrices with overall phase  $\{+1, -1, +i, -i\}$

$$\mathbf{P}_n = \{\pm 1, \pm i\} \times \{\mathbb{1}, X, Y, Z\}^{\otimes n}. \quad (2.4)$$



Moreover, the commutator and anticommutator of Pauli matrices are

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{i,j,k}\sigma_k, \quad (2.5)$$

$$\{\sigma_i, \sigma_j\} = 2i\delta_{i,j}\mathbb{1}, \quad (2.6)$$

where  $\varepsilon_{i,j,k}$  is the Levi-Civita antisymmetric tensor and  $\delta_{i,j}$  is the Kronecker delta. In this paper, we define Clifford group on  $n$  qubits  $\mathbf{C}_n$  [11] different from the usual one. Namely, we ignore the 8 different elements with different in each Clifford gate global phases which make difference when operated on quantum computer because the statistics of the measurement outcomes of  $U$  and  $e^{i\phi}U$  are the same. This definition is more convenient because the total number of group elements will be less eight times than the usual one. The  $\mathbf{C}_n$  is any unitary that take Pauli into Pauli via conjugation.

$$\mathbf{C}_n = \{U \in U(2^n) | \sigma \in \pm \mathbf{P}_n^* \Rightarrow (U\sigma U^\dagger) \in \pm \mathbf{P}_n^*\} / U(1) \quad (2.7)$$

$$\mathbf{P}_n^* = \mathbf{P}_n \setminus \{\mathbb{1}^{\otimes n}\} \quad (2.8)$$

Take single qubit Clifford group  $\mathbf{C}_1$  for example. We know that  $\pm \mathbf{P}_1^* = \{\pm X, \pm Y, \pm Z\}$ . The structure of  $\mathbf{P}_1^*$  remains the same under conjugation. For example,

$$Y = -iXZ \Rightarrow UYU^\dagger = -iUXU^\dagger UZU^\dagger \quad (2.9)$$

Therefore, we don't keep track of  $Y$  in all of this because if we know how  $X$  and  $Z$  behave, we also get  $Y$ 's behavior from Eq.(2.9). Moreover, from Eq.(2.5)  $X$  and  $Z$  anticommute and so do the  $UXU^\dagger$  and  $UZU^\dagger$ . The Clifford elements can map  $X$  to any element of  $\pm\mathbf{P}_1^*$ , which have 6 options and map  $Z$  to any element of  $\pm\mathbf{P}_1^* \setminus \{\pm UXU^\dagger\} = \{\pm Y, \pm Z\}$ , which have 4 options. So, the number of elements in  $\mathbf{C}_1$  is  $6 \times 4 = 24$ . Table 2.1 shows the 24 gates in a single-qubit Clifford group, which is 8 times less than the usual Clifford group.

Gate name	$\theta$	$n_x$	$n_y$	$n_z$
$\mathbb{I}$	0	0	0	0
<b>V</b>	$\frac{1}{2}\pi$	1	0	0
<b>X</b>	$\pi$	1	0	0
<b>V</b> $^\dagger$	$-\frac{1}{2}\pi$	1	0	0
<b>h</b> $^\dagger$	$\frac{1}{2}\pi$	0	1	0
<b>Y</b>	$\pi$	0	1	0
<b>h</b>	$-\frac{1}{2}\pi$	0	1	0
<b>S</b>	$\frac{1}{2}\pi$	0	0	1
<b>Z</b>	$\pi$	0	0	1
<b>S</b> $^\dagger$	$-\frac{1}{2}\pi$	0	0	1
<b>H</b>	$\pi$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	0
	$\pi$	$\frac{1}{\sqrt{2}}$	0	$\frac{1}{\sqrt{2}}$
	$\pi$	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
	$\pi$	$-\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	0
	$\pi$	$\frac{1}{\sqrt{2}}$	0	$-\frac{1}{\sqrt{2}}$
	$\pi$	0	$-\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
	$\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$\frac{\sqrt{3}}{1}$	$\frac{\sqrt{3}}{1}$
	$-\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{\sqrt{3}}{1}$
	$\frac{2}{3}\pi$	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{\sqrt{3}}{1}$
	$-\frac{2}{3}\pi$	$-\frac{1}{\sqrt{3}}$	$\frac{\sqrt{3}}{1}$	$\frac{\sqrt{3}}{1}$
	$\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$\frac{\sqrt{3}}{1}$
	$-\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$\frac{\sqrt{3}}{1}$
	$\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$\frac{\sqrt{3}}{1}$	$-\frac{1}{\sqrt{3}}$
	$-\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
	$\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$\frac{\sqrt{3}}{1}$	$-\frac{1}{\sqrt{3}}$
	$-\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
	$\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$\frac{\sqrt{3}}{1}$	$-\frac{1}{\sqrt{3}}$
	$-\frac{2}{3}\pi$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$

Table 2.1: **1-qubit Clifford group** [4]  $\theta$  is the rotation angle.  $n_x$ ,  $n_y$  and  $n_z$  correspond to the weighted direction of rotation around  $x$ -axis,  $y$ -axis and  $z$ -axis, respectively.

Indeed, the number of elements in the Clifford group  $\mathbf{C}_n$  is [12]

$$|\mathbf{C}_n| = \prod_{i=1}^n 2(4^i - 1) \cdot 4^i. \quad (2.10)$$

$n$	$ C_n $
1	24
2	11250
3	92897280
4	12128668876800
5	25410822678459187200



Table 2.2: Number of elements in the Clifford group  $C_n$

As you can see from Fig.2.2 the number of Clifford group grows with  $n$  very quickly. It is not wise to produce each of the all Clifford elements by individual pulse or way. In fact, we can decompose each Clifford element into specific subgroup [12], called generators, e.g., H, S and CNOT

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (2.11)$$

However, there still are other choice to generate Clifford group. In this thesis, we simulate different generator sets discussed later.

## 2.2 Average gate fidelity

To characterize the performance of gate operation, one needs to define a criterion. First, we consider that gate performance may be affected by the input state. For example, in Fig.2.1 (a), the input state is the zero state  $|0\rangle$ , the ideal gate is taken to be the identity gate  $\mathbb{I}$ , and the noise can be imagined as  $\mathbf{Z}$ . The noisy output state in this case is still the zero state  $|0\rangle$ , which is the same as noiseless case. This means that zero state  $|0\rangle$  is robust under the  $\mathbf{Z}$  noise. However, in Fig.2.1 (b), the set up is all the same as in Fig.2.1 (a) expect that the input state is the plus state  $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ . In this case, the output state becomes the minus state  $|-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$ , which is different from the noiseless case. Thus, we say that the plus state  $|+\rangle$  is not robust under the  $\mathbf{Z}$  noise. Because the input state affects the performance of



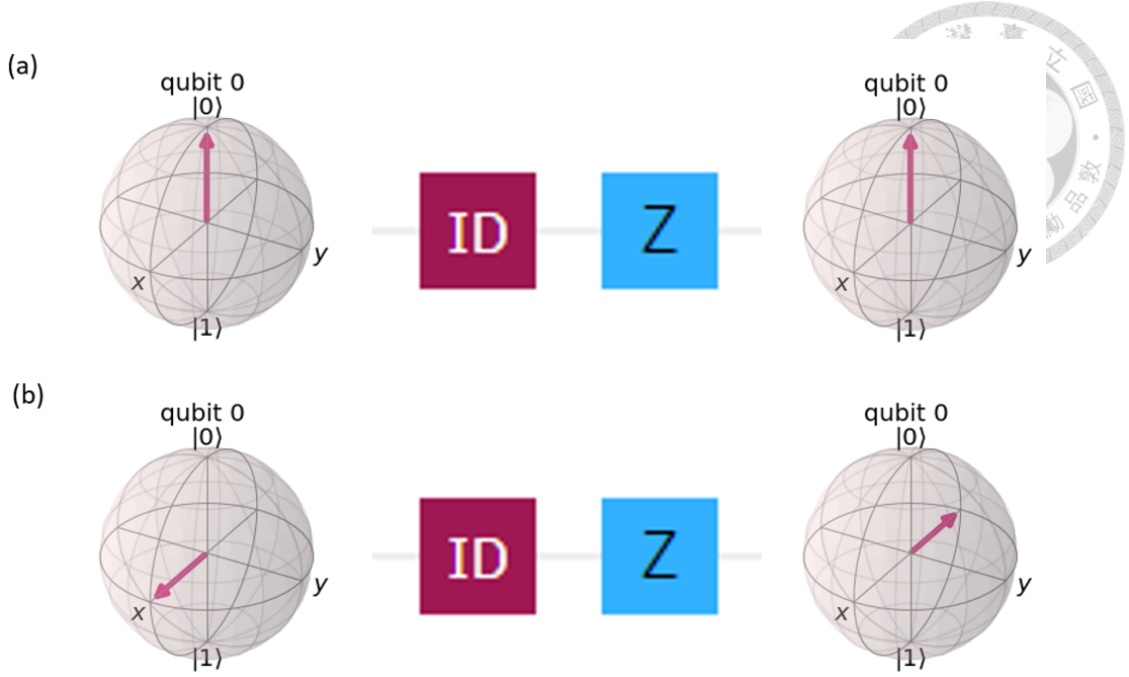


Figure 2.1: **The performance of gate under different input states.** ID is identity gate which is our target gate and Z is Pauli-Z which is our noise. The input state is (a) the zero state  $|0\rangle$ . We see the output state is the same at that of noiseless case. The input state in (b) is the plus state  $|+\rangle$ . We see the output state changes to the  $|-\rangle$  state. That is, the  $|+\rangle$  state is affected by the Z noise.

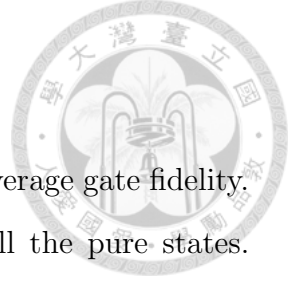
the gate, we have to take all the different input state into consideration. However, because there are infinite number of input states, inputting them one by one will be impossible and inefficient. A better way is averaging all of the effect caused by all the input state. The definition of average fidelity of gate (channel)  $\mathbf{F}_{\text{avg}}$  [2] is

$$\mathbf{F}_{\text{avg}}(\mathcal{E}) = \int d\psi \langle \psi | \mathcal{E}(\psi) | \psi \rangle, \quad (2.12)$$

where the integral is over the uniform Haar measure  $d\psi$  on whole state space with normalization  $\int d\psi = 1$ . In RB, we usually extend Eq.(2.12) to

$$\mathbf{F}_{\text{avg}}(\mathcal{E}, \mathcal{U}) = \int d\psi \langle \psi | \mathcal{U}^\dagger \mathcal{E} \mathcal{U}(\psi) | \psi \rangle, \quad (2.13)$$

where  $\mathcal{U}$  represents ideal gate and  $\mathcal{E}$  represents noisy implementation gate. Thus, Eq.(2.13) shows how well  $\mathcal{E}$  approximates  $\mathcal{U}$ . If  $\mathbf{F}_{\text{avg}}(\mathcal{E}, \mathcal{U}) = 1$ ,  $\mathcal{E}$  is implemented perfectly like ideal gate  $\mathcal{U}$ . However, if  $\mathbf{F}_{\text{avg}}(\mathcal{E}, \mathcal{U}) < 1$ , it implies that  $\mathcal{E}$  is implemented with errors.



## 2.3 Unitary 2-Design

Now, we have to integrate all the input states to obtain the average gate fidelity. For a qubit, one may take the the whole input states to be all the pure states. For example, the surface of the Bloch sphere represents all possible pure states as shown in Fig.2.2. However, the number of pure states is infinite as the points on the surface. In experiment, it would be impossible to test all the pure states. Thus, we define a discrete set  $U_k$  with probability  $p_k$  such that

$$\sum_{k=1}^K p_k U_k^\dagger \mathcal{E}(U_k \rho U_k^\dagger) U_k = \int U^\dagger \mathcal{E}(U \rho U^\dagger) U dU. \quad (2.14)$$

A discrete set  $U_k$  with this property is called unitary 2-design. Usually, we find that such discrete set is uniformly distributed, that is  $p_i = \frac{1}{K}$ . The main idea of unitary

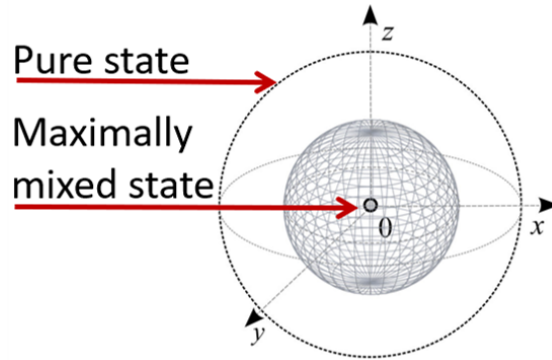


Figure 2.2: **Bloch sphere.** The outer surface represents all the pure states on a Bloch sphere, and the origin is a maximally mixed state.

2-design is that using finite elements to represent infinite elements' behavior. Thus, in experiment, we can design the system Hamiltonian to operate the finite elements. In other words, averaging over a unitary 2-design set (uniform distribution over the Clifford group) is equivalent to averaging over the uniform Haar distribution. Moreover, randomly choosing a unitary element from the Clifford group is indistinguishable from choosing from the uniform Haar distribution as shown in Fig.2.3. We say that Clifford group is unitary 2-design [13].

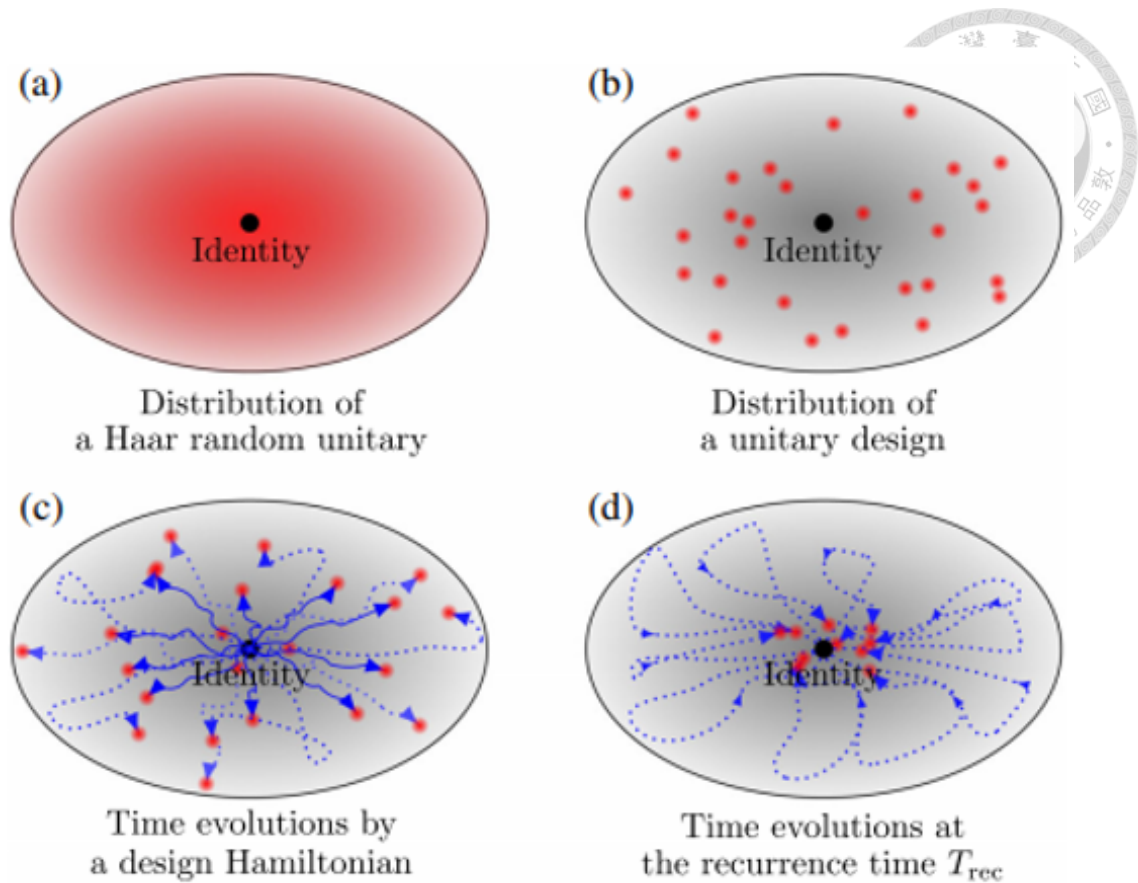


Figure 2.3: **The schematic diagram of Haar unitary** [1]. (a) all the Haar random unitary elements in the red region. (b) a distribution of the unitary design of elements at the red dots. (c) We can design a system containing the specific Hamiltonian. (d) If the unitary design set is a group, each element have its inverse one. That is, each element can go back to identity.

## 2.4 Quantum channel

When we talk about the Schrodinger's equation, we write down the state evolution  $|\psi_f\rangle = U |\psi_i\rangle$ . We say that the evolution  $U$  map initial state  $|\psi_i\rangle$  to final state  $|\psi_f\rangle$ . Generally, the state may not be pure state. Thus, we replace pure state  $|\psi\rangle$  with density operator  $\rho$ . Then, the mapping of density operator  $\rho_i$  to another density operator  $\rho_f$  is called quantum channel. Usually, one uses  $\mathcal{E}(\rho)$  to describe the effect of a system of Hilbert space dimension  $d$  through a quantum channel in operator-sum representation as

$$\mathcal{E}(\rho) = \sum_{k=1}^M E_k \rho E_k^\dagger, \quad (2.15)$$

with the quantum (Kraus) operators  $E_k$  satisfying

$$\sum_{k=1}^M E_k^\dagger E_k \leq I, \quad (2.16)$$

and  $1 \leq M \leq d^2$ .



### 2.4.1 Depolarizing Channel

The depolarizing channel Fig.2.4 maps uniformly pure state into mixed state and it is defined as follows,

$$\mathcal{E}(\rho) = p\rho + (1-p)\frac{\mathbb{1}}{2} \quad (2.17)$$

$$= \left(\frac{1+3p}{4}\right)\rho + \frac{1-p}{4}(X\rho X + Y\rho Y + Z\rho Z), \quad (2.18)$$

where  $p$  is the depolarizing parameter. It implies that input state with probability  $p$  remains unchanged and completely lose information with probability  $1-p$ . The Kraus operators  $E_k$  are  $\sqrt{1+3p}\mathbb{1}/2$ ,  $\sqrt{1-p}X/2$ ,  $\sqrt{1-p}Y/2$ ,  $\sqrt{1-p}Z/2$ .

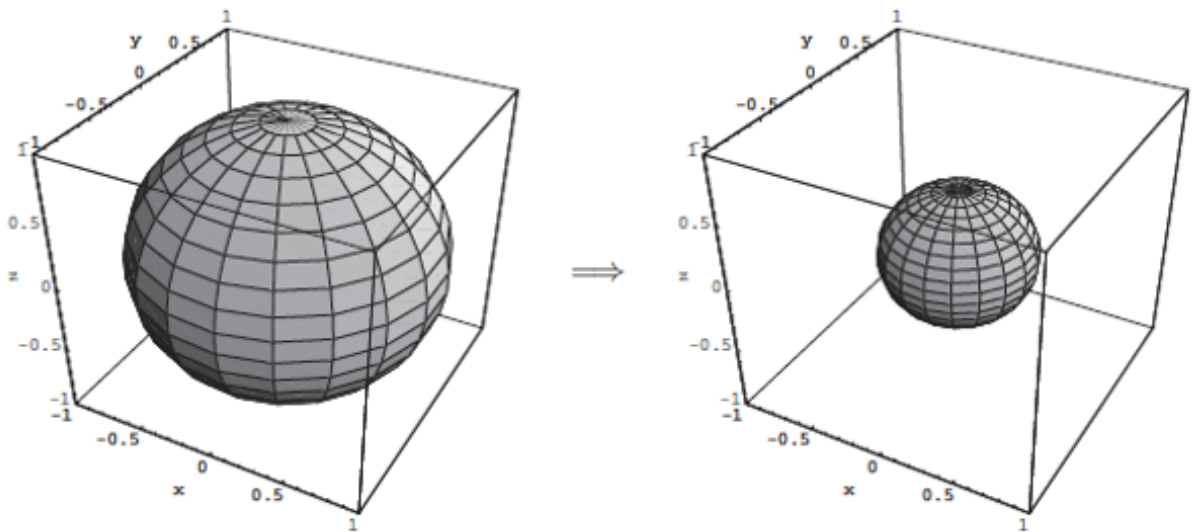


Figure 2.4: **Single qubit depolarizing channel** [2]. It shows the effect of the depolarizing channel on the Bloch sphere of a qubit system.

## 2.4.2 Dephasing Channel

The dephasing quantum channel can be described by

$$\mathcal{E}(\rho) = \left(\frac{1+p}{2}\right)\rho + \left(\frac{1-p}{2}\right)Z\rho Z. \quad (2.19)$$

Dephasing channel implies that input state will lose phase information with probability  $1 - p$  and remains unchanged with probability  $p$ . The Kraus operators are  $\sqrt{\frac{1+p}{2}}\mathbb{1}$ ,  $\sqrt{\frac{1-p}{2}}Z$ . The dephasing channel shrink equator of pure states on the Bloch sphere into elliptical ball, as shown in Fig.2.5.

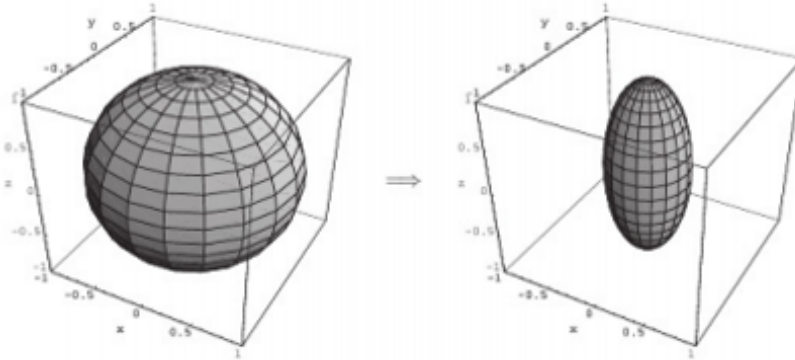


Figure 2.5: **Single qubit dephasing channel** [2]. It shows the effect of the dephasing channel on the Bloch sphere of a qubit system.





## Chapter 3

# Randomized Benchmarking

We introduce randomized benchmarking (RB) to characterize the gate performance. Moreover, we propose a new method based on RB, called primitive randomized benchmarking (PRB).

### 3.1 Randomized Benchmarking

RB [6] was proposed for its less resource requirement compared to QPT. It uses unitary 2-design (Clifford group) to obtain the average gate fidelity. The RB procedure is as follows.

1. prepare an initial state  $\rho$ , which is usually the ground state  $|0\rangle\langle 0|$ .
2. uniformly and randomly choose gate  $C_i$  from the Clifford group gate set.
3. apply the inverse gate  $C_{\text{inv}} = (\prod_{i=1}^m C_i)^\dagger$  which will reverse the whole sequence and the whole sequence become identity gate  $\mathbb{I}$ .
4. repeatedly execute for the sequence, choose different sequences in length  $m$  and vary length  $m$
5. fit the survival probability  $= Ap^m + B$ , where A and B absorb the state preparation and measurement (SPAM) error, and  $p$  is the depolarizing parameter.



Figure 3.1: Randomized benchmarking sequence



### 3.1.1 Advantage of using RB

There are several advantages of using RB to characterize the gate performance. The first advantage is that RB can separate the gate error from the SPAM error. The main idea is that the experiment on the RB sequence is composed of 3 parts: state preparation, gate operation, and measurement. For different gate length  $m$ , the survival probability will decay. However, the state preparation and measurement are fixed for every sequence. Thus, the reason why the survival probability will drop with the gate length is due to the gate error in each gate. With a longer gate length, the survival probability becomes smaller. The second advantage is that it needs only one input state which usually is computational basis  $|0\rangle$ . That's because its following random Clifford operations will take the input state into a uniformly distribution on the whole unitary group. Here, we take one input state and randomized Clifford operations rather than many input states to average over the whole input states. The third advantage is that RB has a simple formula  $Ap^m + B$  for fitting [6] and the average gate fidelity  $F_{\text{avg}}$  is related to  $p$  through [8].

$$F_{\text{avg}} = \int U^\dagger \mathcal{E}(U\rho U^\dagger) U dU. \quad (3.1)$$

$$= \frac{1}{K} \sum_{k=1}^K C_k^\dagger \mathcal{E}(C_k \rho C_k^\dagger) C_k \quad (3.2)$$

$$= p + \frac{1-p}{2^n} \quad (3.3)$$

The equality of Eq (3.2) is because Clifford group  $C$  is a unitary 2-design. Thus, we can directly focus on Clifford group gates  $C_k$  to get the whole unitary gates. Eq (3.3) comes from the reason that every quantum channel become depolarizing channel after the randomized Clifford group gate operations  $C_k$ . The noiseless case process is

$$C_{\text{inv}} \circ C_m \cdots \circ C_2 \circ C_1 \circ \rho.$$

With noise, the RB sequence becomes

$$C_{\text{inv}} \circ \Lambda \circ C_m \circ \cdots \Lambda \circ C_2 \circ \Lambda \circ C_1 \circ \rho.$$



where  $\Lambda$  represents a noisy channel. After inserting  $C_1^\dagger C_1$  between  $\Lambda$  and  $C_2$ , the sequence becomes

$$C_{\text{inv}} \circ \Lambda \circ C_m \circ \cdots \Lambda \circ C_2 \circ C_1 \circ C_1^\dagger \circ \Lambda \circ C_1 \circ \rho.$$

Noting that the insertion will not change the original sequence because of  $C_1^\dagger C_1 = \mathbb{I}$ .

Moreover, we change notation  $D_1 \equiv C_1$  such that

$$C_{\text{inv}} \circ C_m \circ \Lambda \circ \cdots \Lambda \circ C_2 \circ C_1 \circ D_1^\dagger \circ \Lambda \circ D_1 \circ \rho.$$

To here, it is a repeating procedure. We insert  $C_2^\dagger C_1^\dagger C_1 C_2$  between  $\Lambda$  and  $C_2$ ;  $D_2 \equiv C_1 C_2$  such that

$$C_{\text{inv}} \circ \Lambda \circ C_m \cdots D_2^\dagger \circ \Lambda \circ D_2 \circ D_1^\dagger \circ \Lambda \circ D_1 \circ \rho.$$

Repeat and repeat, we can get

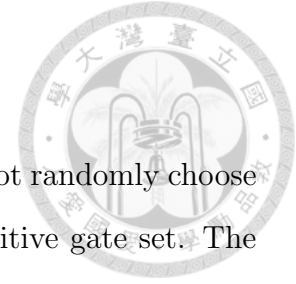
$$\begin{aligned} C_{\text{inv}} \circ D_m^\dagger \circ \Lambda \circ D_m \cdots D_2^\dagger \circ \Lambda \circ D_2 \circ D_1^\dagger \circ \Lambda \circ D_1 \circ \rho \\ = C_{\text{inv}} \circ \{D_i^\dagger \circ \Lambda \circ D_i\}^m \circ \rho, \end{aligned}$$

where  $D_i = C_1 C_2 \dots C_i$ . Therefore, single gate and noise  $\{D_i \circ \Lambda \circ D_i^\dagger\}$  is similar to Eq. (3.2) and the single noise channel  $\Lambda$  becomes the depolarizing channel [8].

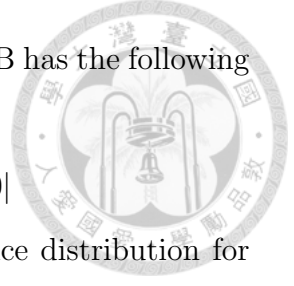
Many gates and noise channels become  $A\rho^m + B$  [6].



## 3.2 PRB



We propose PRB based on RB. The difference is that we do not randomly choose from the Clifford group. Instead, we choose gate from the primitive gate set. The motivation is that, in experiment, we actually use some specific gate set, called primitive gates, to generate the gates in the Clifford group. The Gottesman-Knill theorem [14][15][16] states that the Clifford group can be generated by H, S and CNOT. Moreover, for single-qubit case the generator become H and S. Therefore, in experiment, people use different gate set that is easier to implement in their experiment, called primitive gate, to construct the Clifford gates. For example, in IBM transmon qubit system, people used the single-qubit gate set  $X-\frac{\pi}{2}$  and virtual Z [17] and the two qubit gate CNOT [18]. The UNSW team used in there silicon qubit system the single-qubit gate set (I, X, Z,  $2\pi/3$ , X/2, Z/2, H) [3], and the two-qubit gate set (X/2, X/2 + CROT, Z-CROT, CROT and the virtual Z) [19]. Because different Clifford gates can be generated by different combination of the primitive gates, which means different physical processes, different Clifford gates may have been through different noise channels. For example, in virtual-Z gate set [17], the Clifford Z gate is noiseless due to its design. But for X gate and Y gate, they used real control pulses to implement in reality, and these gates unavoidably have noise. However, the original RB required every Clifford gate to have the same or close enough error. Because there is a gap between the RB theory and experiments toward noise for each Clifford gate, we propose a new method of PRB, trying to connect the theory with experiments more effectively. PRB has the same procedure as RB, but does not choose randomly the gates from the Clifford group. Instead, PRB calculate the times and the probability distribution the primitive gates appear when constructing each Clifford gate and choose the primitive gates randomly but with appearance according to their probability distribution. For example, single qubit Clifford gates in Table 3.1 are generated by the primitive gates ( $\mathbb{I}, \pm\frac{X}{2}, \pm\frac{Y}{2}, X, Y$ ). Each Clifford gate contains an average of 1.875 primitive gates. The distribution of primitive gate appearance ( $\mathbb{I}, X, \frac{X}{2}, -\frac{X}{2}, Y, \frac{Y}{2}, -\frac{Y}{2}$ ) is (1,4,12,8,4,8,8). Thus, we



choose primitive gates from the appearance distribution. The PRB has the following procedure:

1. prepare initial state  $\rho$ , which is usually the ground state  $|0\rangle\langle 0|$
2. randomly choose primitive gate  $P_i$  according to its appearance distribution for the Clifford group
3. apply the inverse gate  $P_{\text{inv}} = (\prod_{i=1}^m P_i)^\dagger$  which will reverse the whole sequence back to the original initial state and the whole sequence become identity gate  $\mathbb{I}$ .
4. repeatedly execute the sequence in length  $m_p$  for some certain times, choose different sequences in the same length  $m_p$  then vary the length  $m_p$  and repeat the same procedure for the new length  $m_p$ . Noting that here length  $m = \frac{m_p}{\kappa}$  is Clifford gate number,  $m_p$  is primitive gate number and  $\kappa$  is the average number of primitive gate per Clifford gate on average.
5. fit the survival probability  $= Ap^m + B$ , where coefficients  $A$  and  $B$  absorb state preparation and measurement (SPAM) error and  $p$  is depolarizing parameter.

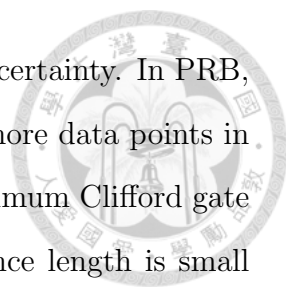
### 3.2.1 Advantages of PRB over RB

Compared with RB, PRB is more flexible and has more data points for a given sequence length. For example, RB sequence length is [1,2,3,4,5]. If each Clifford gate is composed by 2 primitive gates on average, we need to apply 2 pulses to achieve one Clifford gate on average. When we do the RPB, the primitive gate sequence length will be converted to Clifford gate length. Thus, the PRB primitive gate sequence length [1,2,3,4,5] is converted to Clifford gate sequence length as [0.5, 1, 1.5, 2, 2.5]. We can see that PRB can have more data points for a given  $m$  value than RB because PRB can have the data on rational numbers, but RB only have data on integer numbers. PRB can be more reliable than RB when gate fidelity is low. Low gate fidelity will cause the RB sequence curve to drop too fast to have small deviation. The low gate fidelity usually happens when the number of qubits becomes large. Vandersypen's group reported that [7] they did find the gate fidelity due to the large uncertainty when performing the fitting. Therefore, they proposed the



Clifford gate	Physical gates
1	$\mathbb{I}$
2	$Y/2, X/2$
3	$-X/2, -Y/2$
4	$X$
5	$-Y/2, -X/2$
6	$X/2, -Y/2$
7	$Y$
8	$-Y/2, X/2$
9	$X/2, Y/2$
10	$X, Y$
11	$Y/2, -X/2$
12	$-X/2, Y/2$
13	$Y/2, X$
14	$-X/2$
15	$X/2, -Y/2, -X/2$
16	$-Y/2$
17	$X/2$
18	$X/2, Y/2, X/2$
19	$-Y/2, X$
20	$X/2, Y$
21	$X/2, -Y/2, X/2$
22	$Y/2$
23	$-X/2, Y$
24	$X/2, Y/2, -X/2$

Table 3.1: The decomposition table of the single qubit Clifford gates [5]



character randomized benchmarking [8] to deal with the large uncertainty. In PRB, we handle this problem by more intuitively way that is adding more data points in between. PRB can have acceptable accuracy even when the maximum Clifford gate sequence length is small. That the number of maximum sequence length is small means people can perform fewer physical operations, or pulses. Thus, PRB is more friendly for the experiment. The reason is that the RB decay curve constructed by the PRB method can provide have more information when sequence length is small. Once more information about the decay behavior when the sequence length is small is known, it helps to fit the function  $Ap^m + B$  more accurately.

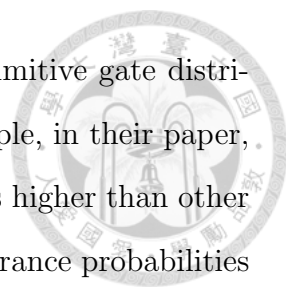
### 3.3 DRB

Direct randomized benchmarking (DRB) is another method for characterizing gate fidelity [20]. The DRB procedure is as follows.

1. prepare an initial state  $\rho$ , which is usually the ground state  $|0\rangle\langle 0|$ .
2. uniformly and randomly choose a gate  $C_i$  from the Clifford group gate set and decompose the Clifford gate into its corresponding primitive gates.
3. randomly choose primitive gate  $P_i$  according to a user-defined distribution
4. apply a gate  $D_t$  which transforms the whole circuit into a multi-qubit computational basis state.
5. repeatedly execute the sequence in length  $m_p$  for some certain times, choose different sequences in the same length  $m_p$ , then vary the length  $m_p$  and repeat the same procedure for the new length  $m_p$ .
6. fit the survival probability  $= Ap^{m_p} + B$ , where coefficients  $A$  and  $B$  absorb state preparation and measurement (SPAM) error and  $p$  is depolarizing parameter.

#### 3.3.1 A comparison between PRB and DRB

The similarity between PRB and DRB is that both use the primitive gate distribution to benchmarking the gate fidelity. However, the goal of DRB is to evaluate



the gate fidelity of the primitive gate set by the user-defined primitive gate distribution. This is useful for strong gate-dependent error. For example, in their paper, they want to evaluate the error of CNOT gate whose error rate is higher than other single qubit gates. They use two distributions with CNOT appearance probabilities of  $p_{CNOT} = 0.75$  and  $p_{CNOT} = 0.25$  within one primitive gate, respectively. That is, at step 3, we have probability  $p_{CNOT} = 0.75$  (or  $p_{CNOT} = 0.25$ ) of choosing a CNOT gate. If the error rate of CNOT gate dominates, then the error rate for  $p_{CNOT} = 0.75$  distribution is three times larger than that for  $p_{CNOT} = 0.25$  distribution. However, PRB try to evaluate the average error rate per Clifford gate which is with the same goal as the original RB. PRB's goal is to solve the gap of the primitive gate decomposition of the Clifford gate between the theory and the experiment. Thus, we have to transform the primitive gate length into the Clifford gate length for PRB (c.f. DRB) and thus the distribution is not user-defined for PRB (c.f. DRB). We observe that using a wrong distribution of choosing primitive gates for PRB fails to obtain a result consistent with the experiment value, which will be discussed in section 5.1.



# Chapter 4

## Protocols

In this Chapter, we simulate RB, RB with decomposition and PRB behaviors under different types of noise.

### 4.1 Protocols

#### 4.1.1 Protocol of RB

The protocol of RB is as follows.

1. Prepare an initial state  $\rho$
2. Uniformly and randomly choose Clifford gates  $C$  from Clifford group
3. Apply the inverse gate  $C_{\text{inv}} = \prod_{i=1}^m (C_i)^\dagger$  and perform the measurements.
4. Repeatedly execute the sequence of length  $m$ , choose different sequences of length  $m$  and repeat step 3. Vary the length  $m$  and repeat steps 3 and 4.
5. Fit  $A\rho^m + B$  by using the weighted least squares method and set  $B = \frac{1}{2}$ .

At step 1, we usually let  $\rho = |0\rangle\langle 0|$  because  $|0\rangle\langle 0|$  is a common computational basis state. After step 3, noiseless RB sequence becomes

$$C_{\text{inv}} \circ C_m \cdots \circ C_2 \circ C_1 = \mathbb{I}, \quad (4.1)$$

and we measure the RB sequence and always get survival probability equal 1. That is, the initial state doesn't change after the RB sequence. However, in real world,

because noise is everywhere, the noisy RB sequence becomes

$$C_{\text{inv}}^{\sim} \circ C_m^{\sim} \cdots \circ C_2^{\sim} \circ C_1^{\sim} \neq \mathbb{I}, \quad (4.2)$$



where  $C_i^{\sim}$  with  $i = 1, 2, \dots, m$  is noisy a Clifford gate. In contrast to the ideal RB sequence, the resultant state after the noisy RB sequence cannot maintain at the same initial state. The survival probability drops when the sequence length  $m$  increases.

### 4.1.2 Protocol of RB with decomposition

In realistic experiment, we usually have to decompose each Clifford gate. That is, Clifford gates are not operated individually directly. Some specific gates, called the primitive gates, are used or combined to construct the Clifford gates. Thus, RB with decomposition by the primitive gates has the following steps

1. Prepare an initial state  $\rho$
2. Uniformly and randomly choose Clifford gates  $C$  from Clifford group
3. Apply the inverse gate  $C_{\text{inv}} = \prod_{i=1}^m (C_i)^\dagger$ .
4. Decompose each Clifford gate into primitive gates and execute and measure the sequence with  $m$  Clifford gates.
5. Repeatedly execute the sequence of the length  $m$ , choose different sequences of length  $m$  and repeat step 3 and 4. Vary length  $m$  and repeat step 3, 4, and 5.
6. Fit  $Ap^m + B$  by using the weighted least squares method and set  $B = \frac{1}{2}$ .

We see that the difference between the RB theory and realistic RB experiment implementation is the decomposition of Clifford gates. The simple RB theory assumes that every Clifford gate has the same error. However, when we execute each single Clifford gate, we decompose the Clifford gate into the combination of the primitive gates which are often also Clifford gates. For example, a Hadamard gate can be achieved by a  $\frac{\pi}{2}$  rotation around the Y-axis, followed by a  $\pi$  rotation around the X-axis. That is,  $H = X_\pi Y_{\pi/2}$ . Here,  $X_\pi$  and  $Y_{\pi/2}$  are the primitive gates and also

the Clifford gates. In simple RB theory, we consider this Hadamard gate have an error, but, in experiment, due to decomposition, it has twice the errors if assuming that each Clifford gates has the same error. We will see the contradiction in the simulation results in Chapter 5. To solve this problem, we propose that RB with decomposition has to rescale the number of gate length. That is, we do not view a Clifford gate with decomposition as one single gate length. Instead, we evaluate it by how many primitive gates on average it needs. The procedure of the rescaled RB with decomposition is as follows.

Steps 1-5 are the same as RB with decomposition described above.

Step 6. Fit  $Ap^{\hat{m}} + B$  by using weighted least squares and set  $B = \frac{1}{2}$ , where  $\hat{m} = m \cdot \kappa$  and  $\kappa$  is the average number of primitive gates needed to construct a Clifford gate. For example, the primitive gate set  $\{\mathbb{I}, \pm X_{\pi/2}, \pm Y_{\pi/2}, X_{\pi}, Y_{\pi}\}$  can construct the whole single-qubit Clifford group Table 3.1. We add up all the primitive gates used in decomposition table and divide by 24, the number of single-qubit Clifford group elements. The average number of the primitive gates per Clifford gate used is 1.875. In this case,  $\kappa = 1.875$ , we rescale the number of single Clifford gate length 1 to 1.875.

### 4.1.3 Protocol of PRB

Although RB with decomposition can be modified by rescaling, we find that the rescale method fails when face large  $\kappa$ . To solve this problem, we propose another protocol, called PRB, which has the following steps:

1. Prepare an initial state  $\rho$ .
2. Randomly choose primitive gates  $P_i$  not uniformly but weighted according to the decomposition table or the appearance distribution.
3. Apply the inverse gate  $P_{\text{inv}} = \prod_{i=1}^m (P_i)^\dagger$ , and perform the measurements.
4. Repeatedly execute the sequence of length  $\hat{m}$ , choose different sequences of length  $\hat{m}$  and repeat step 3. Vary length  $\hat{m}$  and repeat steps 3 and 4. Here  $\hat{m}$  is the primitive gate length not the Clifford gate length.

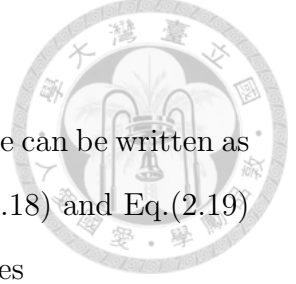


6. Fit  $Ap^m + B$  by using the weighted least squares methods and set  $B = \frac{1}{2}$ , where  $m = \frac{\dot{m}}{\kappa}$ .

Noting that step 2 is different from the above protocols, in which we do not choose the primitive gate uniformly. Instead, we calculate the number of occurrences on the decomposition table, and use their appearance distribution to choose the primitive gates. The idea comes from the experiment. For experimentalists, they want to know what physical operations (primitive gates) to compose the corresponding Clifford gates. They experience that some operations (primitive gates) appear more frequently than the others. Since the Clifford gates are chosen randomly and uniformly, why don't we randomly choose the primitive gates according to their occurrences frequency to construct the Clifford gates? The first immediate advantage of PRB is that it can provide more data points for give  $m$  value due to  $m = \frac{\dot{m}}{\kappa}$ . RB and RB with decomposition have only integer Clifford gate numbers  $m$ . However, PRB can have fraction Clifford gate number  $m = \frac{\dot{m}}{\kappa}$  from  $\dot{m}$ . For example,  $\dot{m} = \{1, 2, 3\}$  and  $\kappa = 1.875$ , the resultant Clifford gate number  $m = \{\frac{1}{1.875}, \frac{2}{1.875}, \frac{3}{1.875}\}$ . The second advantage is that for each data point (each Clifford gate number), the standard deviation is smaller than RB with decomposition. The reason is that each data point provided by RB with decomposition is an average value. For example, some Clifford gates have 3 primitive gates, some have 1 and so on. So we use the average number  $\kappa$  to represent. Situations worsen when  $\kappa$  increases. Despite such elegant advantages, whether the behavior of PRB is the same as RB is still not very clear. Although numerical simulations seem to support that, shown in chapter 5. We still need an analytical proof on justification.

## 4.2 Effect of the noise

In this section, we construct the effect of noise on the gate operations from two different ways.



### 4.2.1 Noisy channel as a map

In the simple RB theory, the noise assumption is that each noise can be written as a quantum channel (map). Thus, we use quantum channel Eq.(2.18) and Eq.(2.19) as the noise channels. The noisy RB sequence of Eq.(4.2) becomes

$$C_{\text{inv}} \circ \mathcal{E} \circ C_m \cdots \mathcal{E} \circ C_2 \circ \mathcal{E} \circ C_1 \circ \rho, \quad (4.3)$$

where  $\mathcal{E}$  is either Eq.(2.18) or Eq.(2.19). For RB with decomposition, the noisy sequence Eq.(4.2) becomes

$$C_{\text{inv}} \circ \mathcal{E}^\kappa \circ C_m \cdots \mathcal{E}^\kappa \circ C_2 \circ \mathcal{E}^\kappa \circ C_1 \circ \rho, \quad (4.4)$$

where  $\kappa$  is the average number of the primitive gates per Clifford gate. For PRB, the noisy sequence Eq.(4.4) become

$$C_{\text{inv}} \circ \mathcal{E} \circ P_m \cdots \mathcal{E} \circ P_2 \circ \mathcal{E} \circ P_1 \circ \rho. \quad (4.5)$$

where  $P$  is a primitive gate.

### 4.2.2 Noise in the Hamiltonian

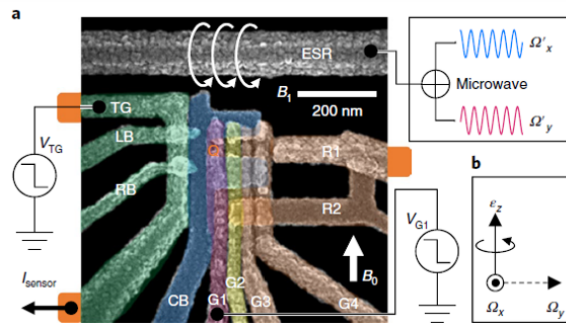


Figure 4.1: Silicon qubit device [3]

We simulate the gate errors of the silicon quantum dot qubits made by UNSW's group [3]. They made silicon qubits by MOS technique. They use a DC magnetic field to produce the Zeeman effect to split the energy levels of the group state  $|0\rangle$

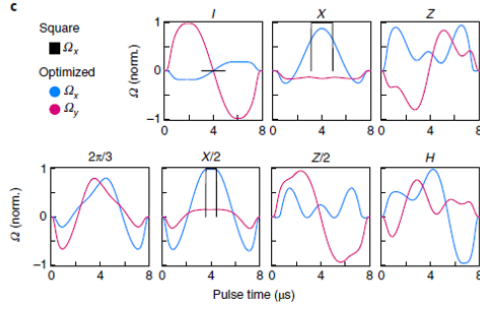


Figure 4.2: **Primitive gate implantation** [3]

and the excited state  $|1\rangle$ , and use the AC magnetic field to perform the electron spin resonance (ESR) to control qubit rotation (see Fig.4.1). In the rotating frame, the effective Hamiltonian can be written as

$$H = \Omega_x \sigma_x + \Omega_y \sigma_y + \epsilon_z \sigma_z \quad (4.6)$$

with  $\Omega_x = \Omega_y = 285.7(\text{kHz})$  and  $\epsilon_z \sim 16.7(\text{kHz})$ .

The basis states is

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

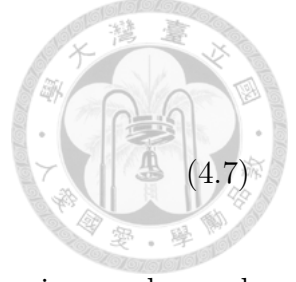
The equation of motion in Schrodinger picture is

$$i\hbar \frac{d}{dt} |\psi_s\rangle = H |\psi_s\rangle.$$

We choose a value of  $\epsilon_z$  every  $400 \mu\text{s}$  from a Gaussian distribution with standard deviation being  $16.7 \text{ kHz}$  to simulate the Z-detuning noise in the experiment. Since the Hamiltonian Eq. (4.6) is time-independent in each gate time, the time evolution operator can be written as  $U = e^{-iHt}$ . Moreover, the control ESR pulses are taken as square pulses. To get the evolution operator, we first define Hamiltonian with

square pulse scheme as

$$H = e_x \Omega_x \sigma_x + e_y \Omega_y \sigma_y + \epsilon_z \sigma_z, \quad (4.7)$$



where  $e_x$  and  $e_y$  is either 0 or 1. If the value is 0, it means that there is no pulse, and if the value is 1, it means that a pulse is activated. The evolution time is decided by the angle  $\theta$  we want,

$$t = \frac{\theta}{2 \cdot 285.7 \text{ kHz}}. \quad (4.8)$$

To expand  $U = e^{-iHt}$ , we normalize  $H$  and  $t$ ,

$$\hat{H} = \frac{H}{\sqrt{(e_x \Omega_x)^2 + (e_y \Omega_y)^2 + (\epsilon_z)^2}} \quad (4.9)$$

$$\hat{t} = t \cdot \sqrt{(e_x \Omega_x)^2 + (e_y \Omega_y)^2 + (\epsilon_z)^2}. \quad (4.10)$$

Thus, the time evolution operator is

$$U = \exp\{-iHt\} \quad (4.11)$$

$$= \cos(\hat{t}) \mathbb{1} - i \sin(\hat{t}) \hat{H}. \quad (4.12)$$

Moreover, the noise effect from the Hamiltonian in the setup breaks the assumption of the simple theory of RB. First, the noise is non-Markovian because we change the noise strength every 400  $\mu s$ , while RB assumes the noise being Markovian. Thus, the simple RB fitting formula does not work. An empirical formula to deal with the

problem is used in [21] and [3] to add a parameter  $\alpha$  into original RB formula

$$\begin{aligned} F &= Ap^m + B \\ &= Ap^m + \frac{1}{2} \\ &= \frac{A}{2}(1 - 2r)^m + \frac{1}{2} \\ &= \frac{A}{2} \exp\{m \ln(1 - 2r)\} + \frac{1}{2} \\ &\simeq \frac{A}{2} \exp\{-2mr\} + \frac{1}{2} \end{aligned} \tag{4.13}$$

as

$$F \simeq \frac{A}{2} \exp\{-2(mr)^\alpha\} + \frac{1}{2}. \tag{4.14}$$

We will also use Eq. (4.14) as our fitting formula for RB and PRB simulations.





# Chapter 5

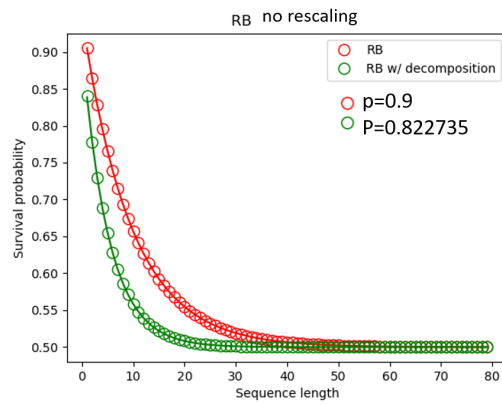
## Simulation results

### 5.1 Results of rescaling

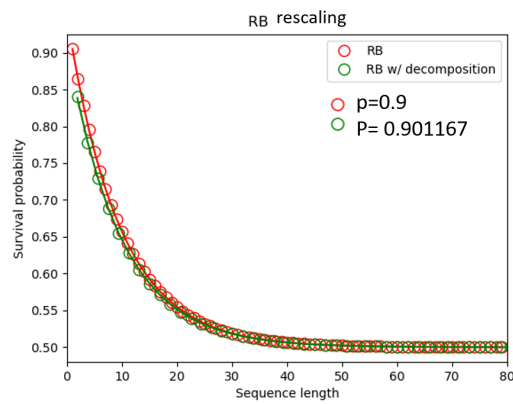
In this section, we consider noise channel as a map, the same assumption used in the simple theory of RB [6] and verify that the results of RB with decomposition need to be rescaled.

#### 5.1.1 Depolarizing channel

We set depolarizing strength  $p = 0.9$  in Eq. (2.18) as our theoretical (true) value for the depolarizing channel. First, we show that RB with decomposition needs rescaling. We see from Fig.5.1(a) that the RB simulations give the exact theoretical value  $p=0.9$  but the simulations of RB with decomposition give the wrong value  $p=0.82$ . The effect of decomposition really exists. Moreover, in a real experiment, the decomposition is unavoidable. Thus, we suggest that the simulation result of RB with decomposition needs rescaling. After rescaling, the  $p$  is corrected to approach the right value as shown in Fig.5.1(b). Second, we show that when  $\kappa$  is large, the rescaling does not work that well. We see from Fig.5.2(b) that for the case of  $\kappa = 3.5$  the difference in value  $p$  after rescaling is larger than that of  $\kappa = 1.875$ . The reason is that when  $\kappa$  is large, the standard deviation of the simulation result becomes large. Thus, we introduce PRB to solve this problem. We see from Fig.5.5(b) that

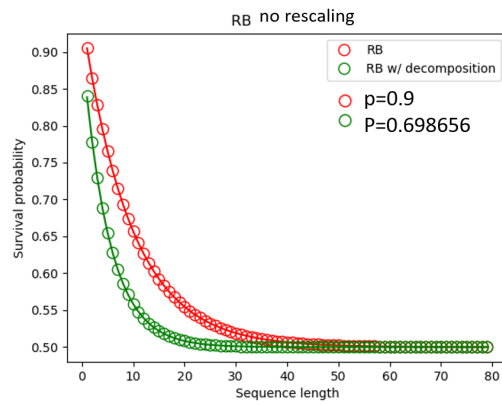


(a)

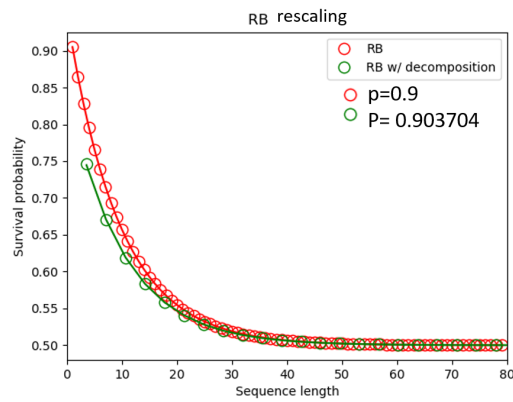


(b)

Figure 5.1: Simulation results of RB (red line) and RB with decomposition (green line) for  $\kappa = 1.875$  under the depolarizing channel. (a) Simulation results without rescaling. (b) Simulation results with rescaling.



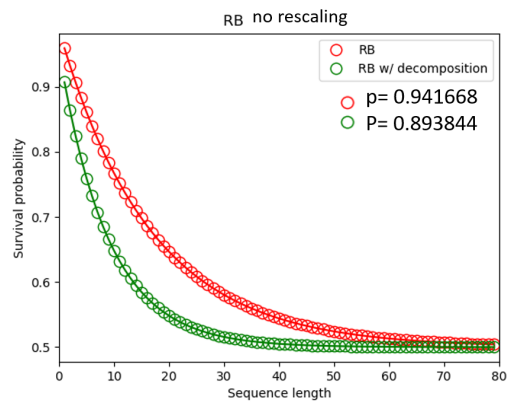
(a)



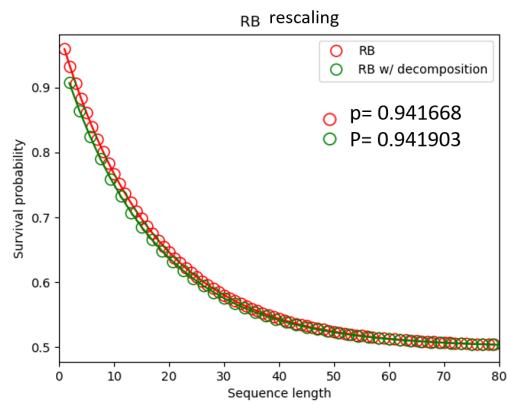
(b)

Figure 5.2: Simulation results of RB (red line) and RB with decomposition (green line) for  $\kappa = 3.5$  under the depolarizing channel. (a) Simulation results without rescaling. (b) Simulation results with rescaling.



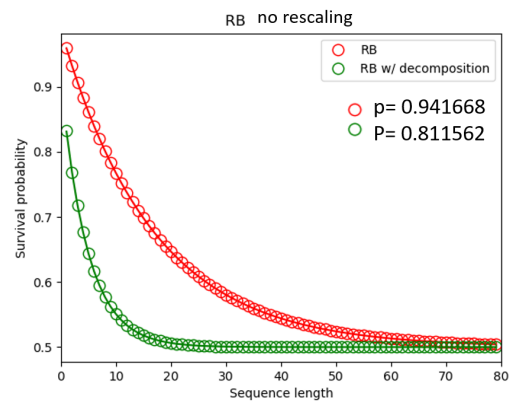


(a)

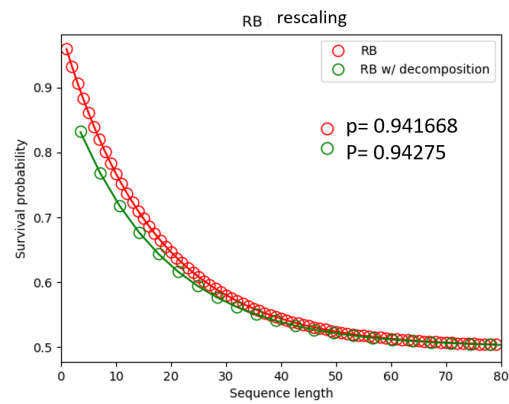


(b)

Figure 5.3: Simulation results of RB (red line) and RB with decomposition (green line) for  $\kappa = 1.875$  under the dephasing channel. (a) Simulation results without rescaling. (b) Simulation results with rescaling.

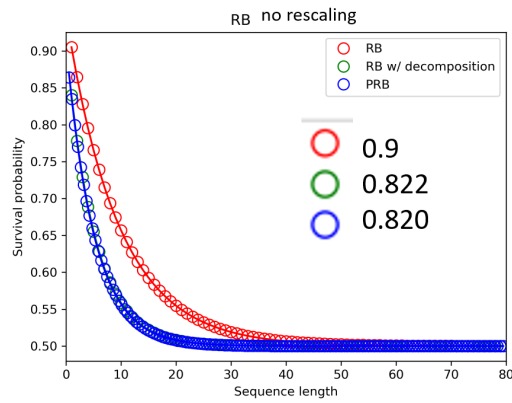


(a)

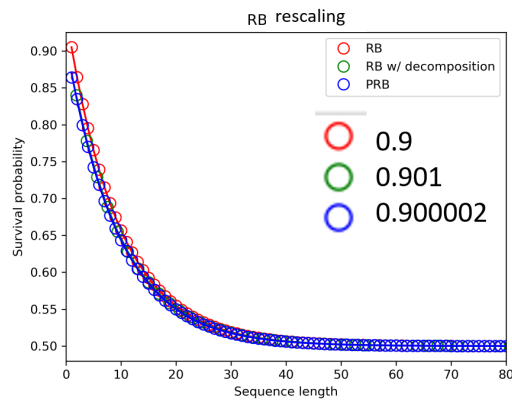


(b)

Figure 5.4: **Simulation results of RB (red line) and RB with decomposition (green line) for  $\kappa = 3.5$  under the dephasing channel.** (a) Simulation results without rescaling. (b) Simulation results with rescaling.

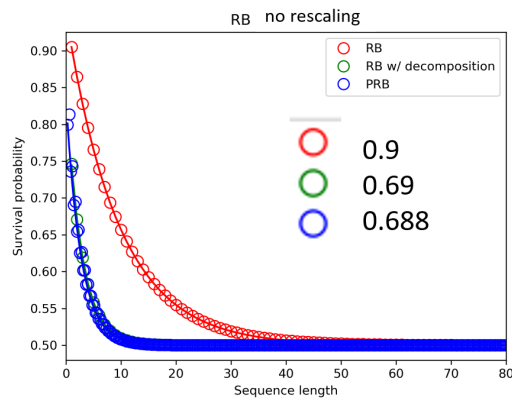


(a)

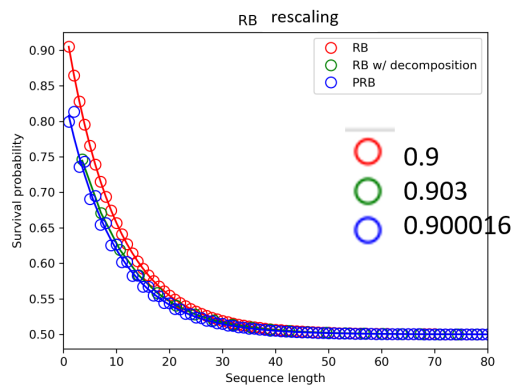


(b)

Figure 5.5: Simulation results of RB (red line), RB with decomposition (green line) and PRB (blue line) for  $\kappa = 1.875$  under the depolarizing channel. (a) Simulation results without rescaling. (b) Simulation results with rescaling.

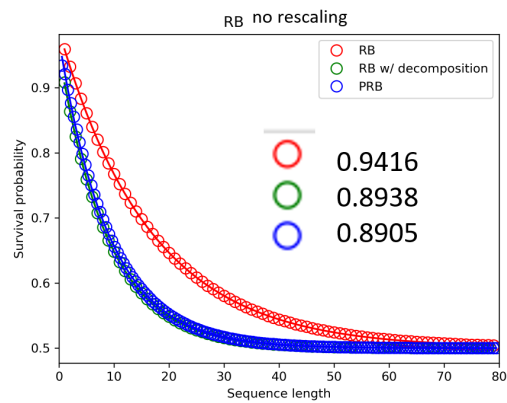


(a)

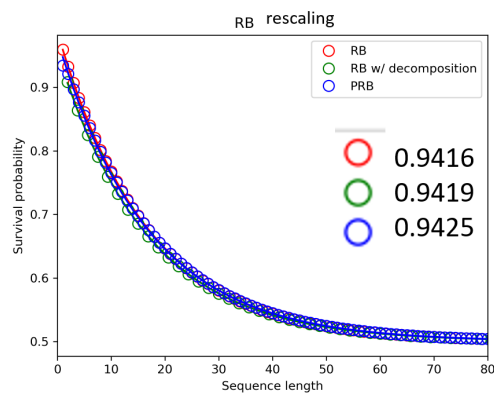


(b)

Figure 5.6: **Simulation results of RB (red line), RB with decomposition (green line) and PRB (blue line) for  $\kappa = 3.5$  under the depolarizing channel.** (a) Simulation results without rescaling. (b) Simulation results with rescaling.

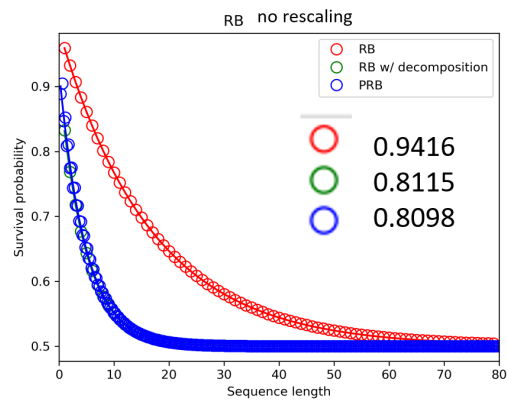


(a)

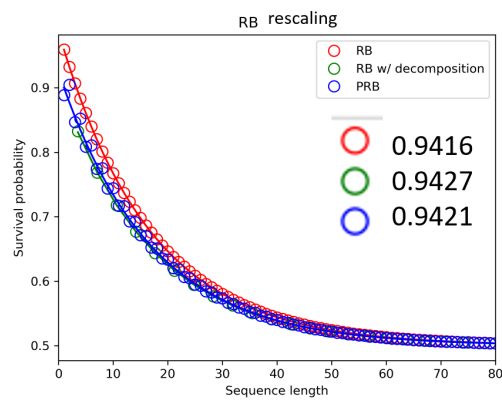


(b)

Figure 5.7: Simulation results of RB (red line), RB with decomposition (green line) and PRB (blue line) for  $\kappa = 1.875$  under the dephasing channel. (a) Simulation results without rescaling. (b) Simulation results with rescaling.

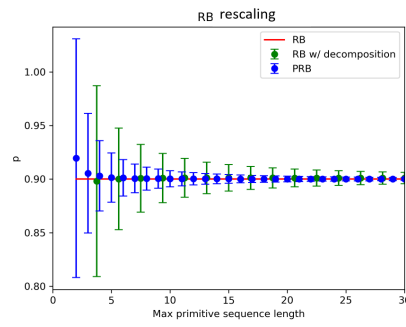


(a)

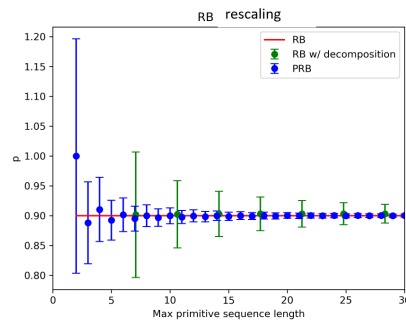


(b)

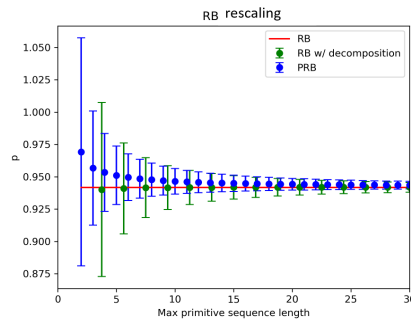
Figure 5.8: Simulation results of RB (red line), RB with decomposition (green line) and PRB (blue line) for  $\kappa = 3.5$  under the dephasing channel. (a) Simulation results without rescaling. (b) Simulation results with rescaling.



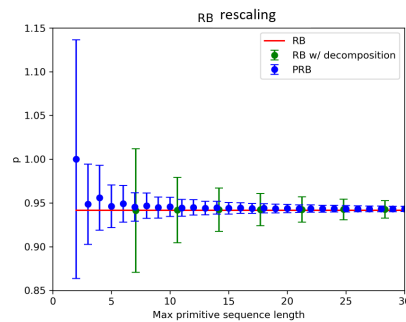
(a)



(b)

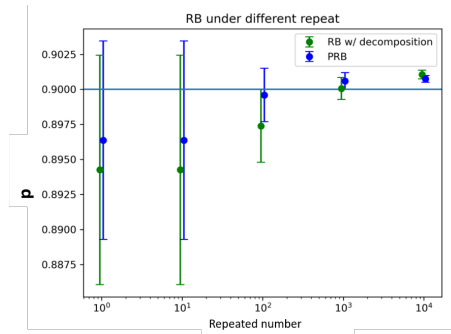


(c)

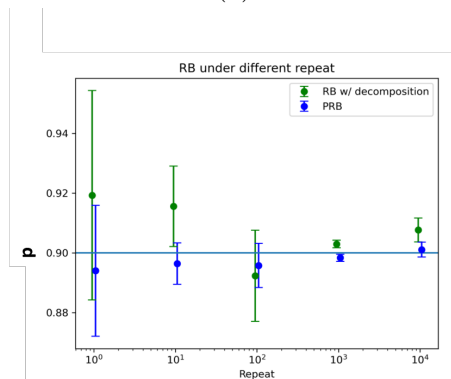


(d)

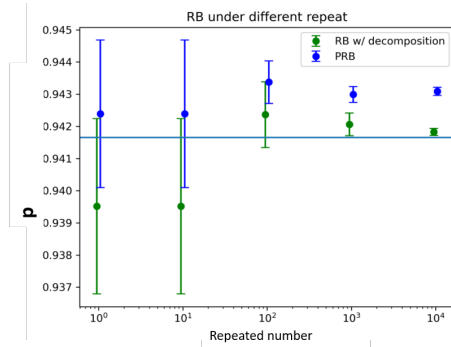
Figure 5.9: Comparison between the standard deviations of the simulation resultant value  $p$  for different RB schemes: RB (red line), RB with decomposition (green line) and PRB (blue line). (a) Standard deviation for  $\kappa = 1.875$  under the depolarizing channel. (b) Standard deviation for  $\kappa = 3.5$  under the depolarizing channel. (c) Standard deviation for  $\kappa = 1.875$  under the dephasing channel. (d) Standard deviation for  $\kappa = 3.5$  under the dephasing channel.



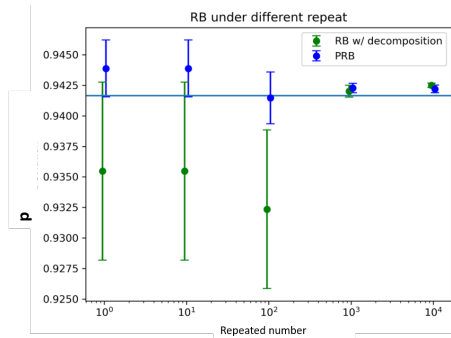
(a)



(b)



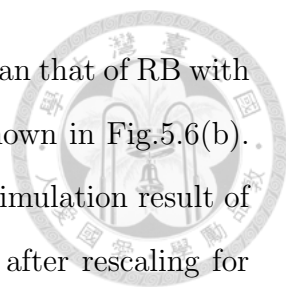
(c)



(d)

Figure 5.10: Comparison between the deviations of the simulation resultant value  $p$  for different RB schemes: RB (red line), RB with decomposition (green line) and PRB (blue line). (a) Deviation for  $\kappa = 1.875$  under the depolarizing channel. (b) Deviation for  $\kappa = 3.5$  under the depolarizing channel. (c) Deviation for  $\kappa = 1.875$  under dephasing channel. (d) Deviation for  $\kappa = 3.5$  under the dephasing channel.





simulation result of PRB after rescaling gives a closer value of  $p$  than that of RB with decomposition for  $\kappa = 1.875$ . This is also evident for  $\kappa = 3.5$  shown in Fig.5.6(b). Third, we show in Fig.5.9(a) that the standard deviation of the simulation result of  $p$  value of PRB is smaller than that of RB with decomposition after rescaling for  $\kappa = 1.875$ . This is even so for  $\kappa = 3.5$  as shown in Fig.5.9(b). With small standard deviation, we can reduce the repeated number to obtain the  $p$  value with the same accuracy. To emphasize this propriety, and mimic a more realistic case. We add the SPAM error and choose  $p$  for each gate set from a normal distribution with mean = 0.9 and standard deviation = 0.00999 rather than a fixed  $p$  value. By adding above two factors, we can approximate our simulations closer to real experiments than before. Moreover, the uncertainty become larger. It is impossible to get the exact value because of finite sample size effect in an experiment. With a limited resource, experimentists can do only finite repetition on RB sequences. Thus, PRB with a smaller standard deviation than others has its advantage. This is, we can just perform a few repeated experiments using PRB, and we can get value  $p$  closer to the exact theoretical value than other RB schemes. Under the depolarizing channel, the simulation result of value  $p$  of PRB has smaller deviations than those of RB with decomposition as shown in Figs.5.10(a) and 5.10(b). Moreover, the  $p$  value obtained by PRB is closer to exact value than that obtained by RB with decomposition as shown in Fig.5.11(a) and 5.11(b). That is, PRB is more accurate when the repeated number is small, and thus is friendly to experiments because experimentists can not do the same experiments unlimited times.

## 5.1.2 Dephasing channel

We set dephasing strength  $p = 0.9$  in Eq. (2.19) for the dephasing channel. First, we show in Fig.5.3 that rescaling is needed under the dephasing channel. Second, when  $\kappa$  is large, rescaling works less well as shown in Fig.5.4(b). Third, the standard deviation of value  $p$  obtained by PRB is also smaller for the dephasing channel. But different from depolarizing channel is that when  $\kappa (= 1.875)$  is small,

	UNSW group	RB	PRB	PRB w/ half sequence	PRB w/ uniform distribution
1000r	1.7(3)	1.76(2)	1.70(8)	1.70(0)	1.405
$p$		0.997	0.997	0.997	0.997
$A$		0.538	0.538	0.54	0.54
$\alpha$	0.8(1)	0.817	0.826	0.795	0.812

Table 5.1: **Simulation results of the value of  $r$ ,  $p$ ,  $A$  and  $\alpha$ .**

PRB for the dephasing channel does not work better than RB with decomposition. The reason is that RB sequence from Clifford group can transform any map into a depolarizing channel, but PRB sequence from an appearance distribution can not exactly transform any map into depolarizing channel, especially when  $\kappa$  is small. But when  $\kappa = (3.5)$  is big, PRB works better. We think that the uncertainty in value  $p$  at  $\kappa = 3.5$  is larger than PRB's approximation. Moreover, PRB has more data points for fitting for a given  $m$  value when  $\kappa(= 3.5)$  is larger.

## 5.2 Noise from the Hamiltonian

### 5.2.1 Good model verified

In this section, we show that our model is good enough to explain noise in silicon qubit [3].

We construct our Hamiltonian in Chapter 4 and compare our simulation result with the experiment result in Fig.5.12(a). We observe that our RB simulation curve in Fig.5.12(b) and  $p$  value in Table are very close to the real experiment. We can state with confidence that our model is good enough to explain the experiment results.

### 5.2.2 PRB work

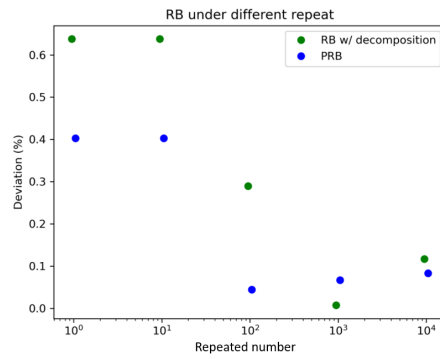
We note that the simulation result of PRB in Fig.5.12(c) can give us the same result as RB. Moreover, if we simulate PRB's distribution not from decomposition table but from a uniform distribution, we see that the  $p$  value obtained by this way

from Fig.5.12(d) is not right.

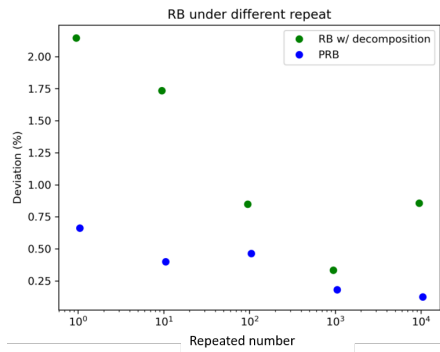
### 5.2.3 PRB under small repetition

With a small standard deviation in  $p$  value of PRB, we reduce the number of repetitions. We see that a small number of repetitions, PRB is closer to exact value with a smaller deviation as shown in Fig.5.13(a). We also test the case when PRB is performed with a half maximum sequence length, the result is shown in Fig.5.13(b). For both the situations, we can state that PRB work better than RB with decomposition when the number of repetitions is small.

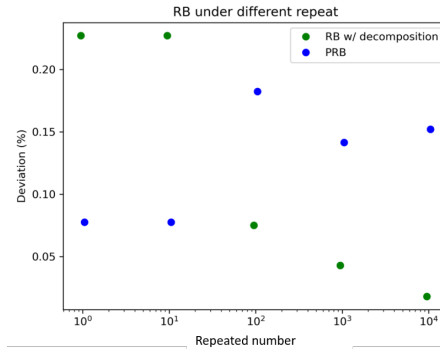




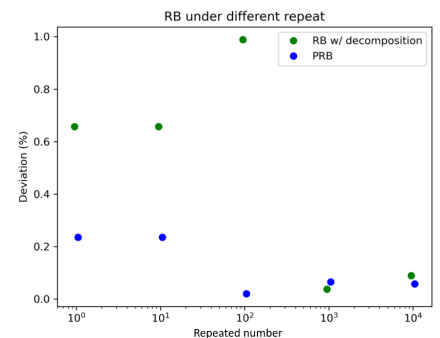
(a)



(b)

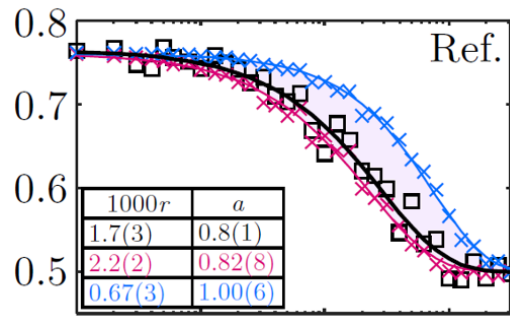


(c)

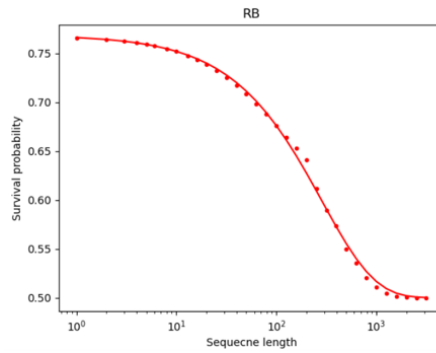


(d)

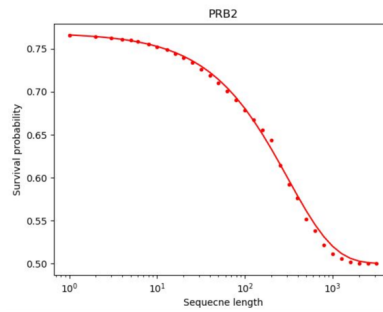
Figure 5.11: Comparison of deviation in  $p$  value v.s. repeated number between RB with decomposition (green dots) and PRB (blue dots). (a) Standard deviation  $\kappa = 1.875$  in depolarizing channel. (b) Standard deviation  $\kappa = 3.5$  in depolarizing channel. (c) Standard deviation  $\kappa = 1.875$  in dephasing channel (d) Standard deviation  $\kappa = 3.5$  in dephasing channel



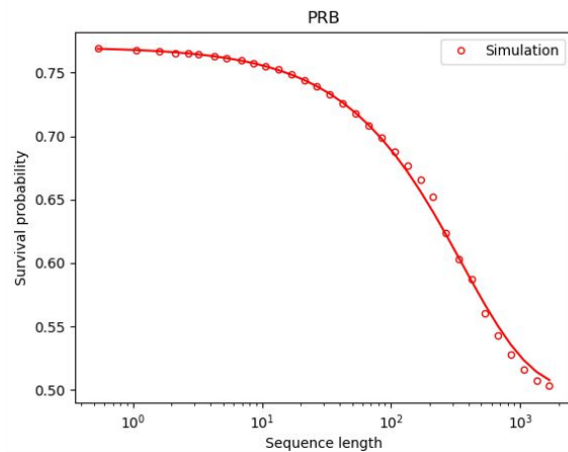
(a)



(b)

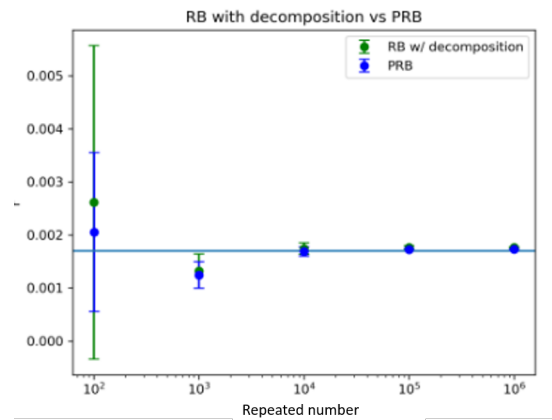


(c)

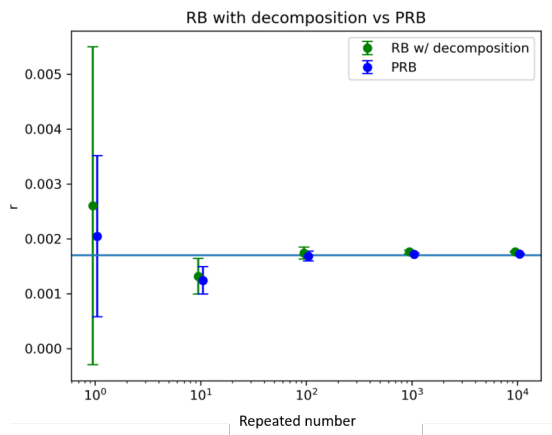


(d)

Figure 5.12: **Simulation results of the noise from the Hamiltonian to mimic the real experimental result.**(a) Experimental result from the UNSW's group [3]. (b) RB simulation result. (c) PRB simulation result. (d) PRB under uniform distribution.



(a)



(b)

Figure 5.13: Comparison of error rate  $r$  v.s. repeated number between RB with decomposition and PRB with the number of repetition.(a) the same as [3].(b) a half of that in (a).



## Chapter 6

# Conclusions

We have pointed out that there is a gap between the RB theory and real experimental implementation. The results of RB with decomposition of primitive gates obtained in an experiment needs rescaling. The rescaling works well for small values of average number of primitive gates per Clifford gate  $\kappa$ . When the value of  $\kappa$  becomes large, the standard deviation of  $p$  value obtained becomes big, which in turn causes rescaling not working that well. The solution we propose is PRB. PRB has the following advantages.

First, PRB can have more data points than RB for a given Clifford gate length, especially when  $\kappa$  is large. Second, PRB is more accurate when the number of repetitions is small because of its small standard deviation, especially when  $\kappa$  is large. Third, PRB can still work well with a reduced maximum gate length in a sequence. For large qubit system, each Clifford gate need more primitive gates. It implies large  $\kappa$ . The RB curve for multi-qubit gate may decay too fast to get a reliable result [7]. Also, limited numbers of measurements and repetitions cause finite sample size effect. PRB can help because of its nice feature of small standard deviation. We believe that PRB can work well toward a large qubit system. We have performed numerical simulation to demonstrate our proposed PRB can work better than RB and RB with decomposition scheme. It would be desired to show analytically that PRB actually work. This will be one of our major future works.



# Appendix A

## Appendix Title

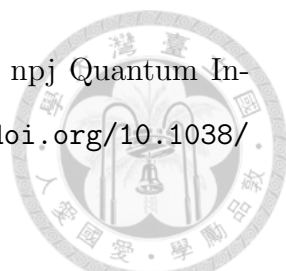
This is the appendix.

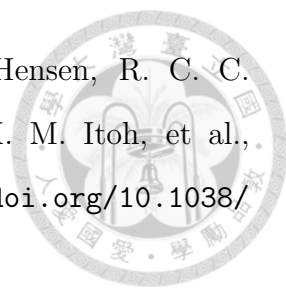




# Bibliography

- [1] Y. Nakata, C. Hirche, M. Koashi, and A. Winter, *Phys. Rev. X* **7**, 021006 (2017), URL <https://link.aps.org/doi/10.1103/PhysRevX.7.021006>.
- [2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, USA, 2011), 10th ed., ISBN 1107002176.
- [3] C. H. Yang, K. W. Chan, R. Harper, W. Huang, T. Evans, J. C. C. Hwang, B. Hensen, A. Laucht, T. Tanttu, F. E. Hudson, et al., *Nature Electronics* **2**, 151 (2019), ISSN 2520-1131, URL <https://doi.org/10.1038/s41928-019-0234-1>.
- [4] G. E. Crooks (2020), URL [https://threeplusone.com/pubs/on\\_gates.pdf](https://threeplusone.com/pubs/on_gates.pdf).
- [5] J. T. Muhonen, A. Laucht, S. Simmons, J. P. Dehollain, R. Kalra, F. E. Hudson, S. Freer, K. M. Itoh, D. N. Jamieson, J. C. McCallum, et al., *Journal of Physics: Condensed Matter* **27**, 154205 (2015), URL <https://doi.org/10.1088/0953-8984/27/15/154205>.
- [6] E. Magesan, J. M. Gambetta, and J. Emerson, *Phys. Rev. A* **85**, 042311 (2012), URL <https://link.aps.org/doi/10.1103/PhysRevA.85.042311>.
- [7] X. Xue, T. F. Watson, J. Helsen, D. R. Ward, D. E. Savage, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, S. Wehner, and L. M. K. Vandersypen, *Phys. Rev. X* **9**, 021011 (2019), URL <https://link.aps.org/doi/10.1103/PhysRevX.9.021011>.

- 
- [8] J. Helsen, X. Xue, L. M. K. Vandersypen, and S. Wehner, *npj Quantum Information* **5**, 71 (2019), ISSN 2056-6387, URL <https://doi.org/10.1038/s41534-019-0182-7>.
- [9] 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, et al., *Phys. Rev. Lett.* **109**, 080505 (2012), URL <https://link.aps.org/doi/10.1103/PhysRevLett.109.080505>.
- [10] A. Carignan-Dugas, J. J. Wallman, and J. Emerson, *Phys. Rev. A* **92**, 060302 (2015), URL <https://link.aps.org/doi/10.1103/PhysRevA.92.060302>.
- [11] M. Ozols (2008), URL <http://home.lu.lv/~sd20008/papers/essays/Clifford%20group%20%5Bpaper%5D.pdf>.
- [12] P. Selinger, *Logical Methods in Computer Science* **Volume 11, Issue 2** (2015), URL <https://lmcs.episciences.org/1570>.
- [13] C. Dankert, R. Cleve, J. Emerson, and E. Livine, *Phys. Rev. A* **80**, 012304 (2009), URL <https://link.aps.org/doi/10.1103/PhysRevA.80.012304>.
- [14] D. Gottesman, in *22nd International Colloquium on Group Theoretical Methods in Physics* (1998), pp. 32–43, [quant-ph/9807006](https://arxiv.org/abs/quant-ph/9807006).
- [15] M. Van Den Nes, *Quantum Info. Comput.* **10**, 258–271 (2010), ISSN 1533-7146.
- [16] P. Selinger, *Logical Methods in Computer Science* **Volume 11, Issue 2** (2015), URL <https://lmcs.episciences.org/1570>.
- [17] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, *Phys. Rev. A* **96**, 022330 (2017), URL <https://link.aps.org/doi/10.1103/PhysRevA.96.022330>.
- [18] H. Abraham, AduOffei, R. Agarwal, I. Y. Akhalwaya, G. Aleksandrowicz, T. Alexander, M. Amy, E. Arbel, Arijit02, A. Asfaw, et al., *Qiskit: An open-source framework for quantum computing* (2019).

- 
- [19] W. Huang, C. H. Yang, K. W. Chan, T. Tantt, B. Hensen, R. C. C. Leon, M. A. Fogarty, J. C. C. Hwang, F. E. Hudson, K. M. Itoh, et al., Nature **569**, 532 (2019), ISSN 1476-4687, URL <https://doi.org/10.1038/s41586-019-1197-0>.
- [20] T. J. Proctor, A. Carignan-Dugas, K. Rudinger, E. Nielsen, R. Blume-Kohout, and K. Young, Phys. Rev. Lett. **123**, 030503 (2019), URL <https://link.aps.org/doi/10.1103/PhysRevLett.123.030503>.
- [21] M. A. Fogarty, M. Veldhorst, R. Harper, C. H. Yang, S. D. Bartlett, S. T. Flammia, and A. S. Dzurak, Phys. Rev. A **92**, 022326 (2015), URL <https://link.aps.org/doi/10.1103/PhysRevA.92.022326>.