國立臺灣大學電機資訊學院電子工程學研究所

碩士論文

Graduate Institute of Electronics Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master's Thesis

qFD: 針對量子機器中相干和去極化錯誤的診斷方法

qFD: Coherent and Depolarizing Fault Diagnosis for Quantum Processors

李彥緯

Yen-Wei Li

指導教授: 李建模 博士

Advisor: James Chien-Mo Li Ph.D.

中華民國 113 年 5 月

May, 2024

致謝

開頭先謝謝所有在碩士路上幫助我的人。首先是李建模老師，在我大學時對未來選組迷茫時，是聽了老師的專題說明會才加入實驗室的。當其他實驗室可能一個月才咪挺一次時，我每個禮拜都可以跟老師討論研究的方向，同時也督促自己要一直努力。偶爾犯錯時老師也會提醒我們該如何改正，很多都是以後職場上很有可能會犯的錯，謝謝老師一直以來的教導。第二要感謝的是承運，從碩一以來就給了我很多量子電路的知識，研究上的問題也是知無不言，彷彿我有兩個指導教授一樣。咪挺結束後也會跟我一起去活大買素食，讓我保持身體健康，謝謝承運在生活上和研究上都幫了我很多。

之後要謝謝的是學長們，尤其是育閔。謝謝他的研究讓我對量子電路診斷了解更多，投稿的論文也讓我可以去美國聖地牙哥報告，同時玩了好幾天。之後謝謝我的同屆們，兆和、哲嘉、秉翰、心屏在我課程上給了我很多幫助，還有在體育課時一起打桌球和羽球。謝謝之後進來的學弟學妹們，給了我很多報告和投影片上的建議，謝謝大家。

最後謝謝我的家人和朋友們，在我研究遇到瓶頸時給了我可以稍微放鬆的空間和時間，讓我可以撐過去，謝謝你們。

摘要

故障引起的錯誤會嚴重影響噪聲中等規模量子（NISQ）電路的正確性。在這項工作中，我們提出了一種診斷 NISQ 電路中相干和非相干故障的技術。所提出的技術包括三個階段：粗略診斷、精細診斷和去極化診斷。粗略診斷通過在布洛赫球上進行網格搜索來定位相干故障的大致範圍。然後，精細診斷根據縮小的搜索空間精確定位相干故障。最後，去極化診斷測量極化故障的大小。我們使用 Qiskit 模擬器進行了技術演示，其中包括無噪聲和有噪聲的後端。我們診斷了具有 27 個單量子位元、6 個雙量子位元相干故障和 7 個注入的去極化故障的量子電路。診斷出的故障量子閘與注入的故障量子閘之間的平均量子閘保真度超過 99.95%，優於相同條件下的傳統量子過程層析。還在 IBM Q 設備上進行了實驗，結果顯示，超過 99.83% 的平均量子閘保真度表明我們的技術仍然在實際量子電路設備上保持良好的分辨率。

關鍵字：雜訊中等規模量子、量子電路、錯誤診斷

Abstract

Errors caused by faults would strongly affect the correctness of noisy intermediate-scale quantum (NISQ) circuits. In this work, we propose a technique for diagnosing coherent and incoherent faults for NISQ circuits. The proposed technique contains three phases: rough diagnosis, fine diagnosis, and depolarizing diagnosis. Rough diagnosis grid searches the Bloch sphere to locate an approximate range of a coherent fault. Fine diagnosis then precisely locates the coherent fault size based on the narrowed-down search space. At last, depolarizing diagnosis measures the depolarizing fault size. We demonstrate our technique using the Qiskit simulator with noise-free and noisy backends. The diagnosis accuracy between the diagnosed faulty gates and the injected faulty gates is over 99.95%, which is better than traditional quantum process tomography under the same conditions. Our results show that the diagnosis error of coherent faults does not affect the diagnosis accuracy of depolarizing faults. Experiments on the IBM Q devices have also been performed, and results of over 99.83% diagnosis accuracy show that our technique still preserves good resolution on real quantum circuit devices.

Keywords: noisy intermediate-scale quantum, quantum circuit, fault diagnosis

Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Motivation

Quantum circuits are one of the most anticipated technologies in recent years. Applications of quantum circuits have been studied in various fields, including chemistry, machine learning, and cryptography. Many companies have built numerous quantum processors, including Google's 70-qubit processor (*Sycamore* [1]), IonQ's 32-qubit processor (*Forte* [8]), and IBM's 133-qubit processor (*Heron* [9]). IBM also announced that their 127-qubit quantum processor had beaten supercomputers when performing a classical brute-force computation [10]. Figure 1.1 is the quantum processor developed by the *Google Quantum AI team*.

While the size of quantum processors kept increasing, the reason quantum processors are not commercialized is that they are vulnerable to *errors* in the noisy intermediate-scale quantum (NISQ) era [11]. Figure 1.2 shows the correlation between the number of qubits and the limiting error rate. We have to both increase the number of qubits and keep the limiting error rate low enough to achieve a useful error corrected QC. There are two
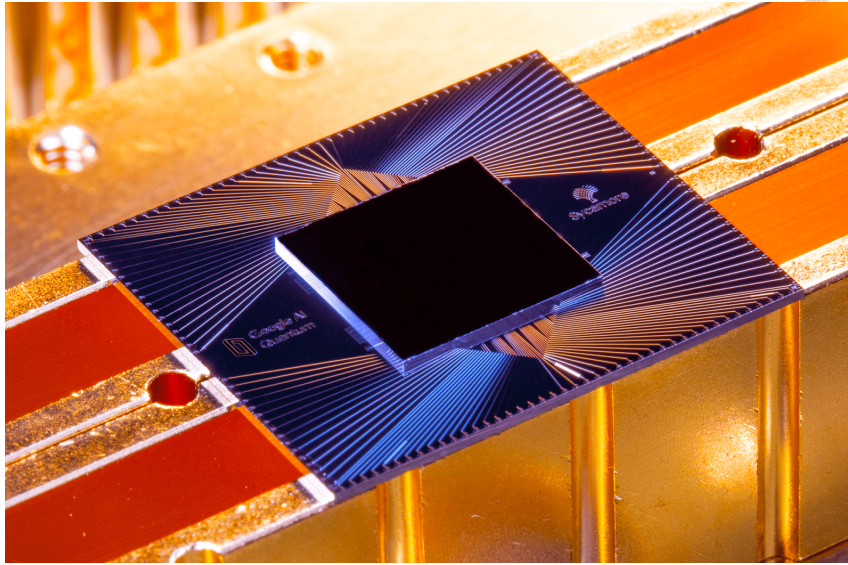
Figure 1.1: Google's quantum processor Sycamore [1]

sources of errors. *Noises* are randomly induced by the environment, and we cannot predict the error effects from noises, such as dephasing noises. *Faults*, however, are induced by defects of the quantum processor hardware. Unlike noises, the effects of faults remain time-independent, meaning the same fault-induced errors occur every time we execute a quantum circuit. Hence, fault diagnosis is crucial for the success of quantum processors. Our research aims to diagnose the faults in quantum processors. For manufacturers, diagnosis can help to re-calibrate their quantum processors or repair the root cause of faults. For users, diagnosis can help error mitigation or error suppression on their circuits to obtain more accurate results than the original circuits. Fig.1.4 is the future roadmap of quantum computing released by IBM [12]. We can see that they focus on increasing the number of qubits and canceling errors caused by noise.
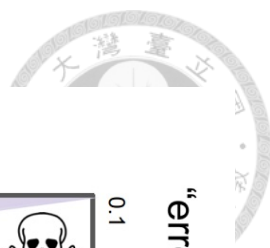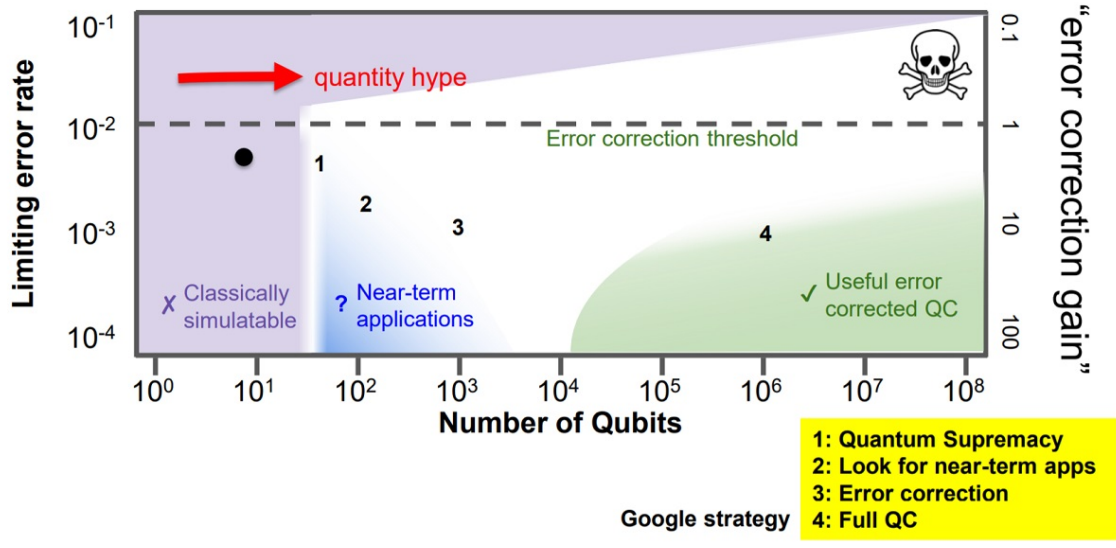
Figure 1.2: Characters of quantum processor technology
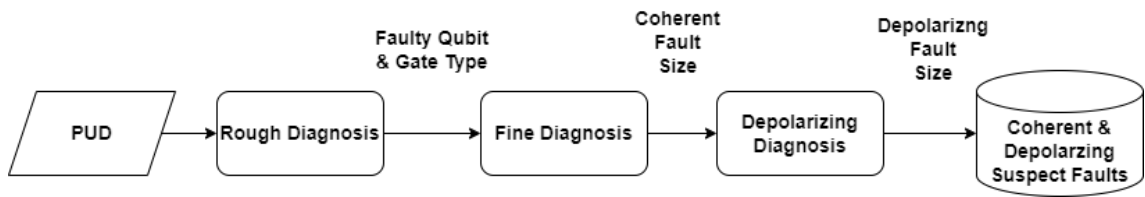
Figure 1.3: Our proposed diagnosis flow
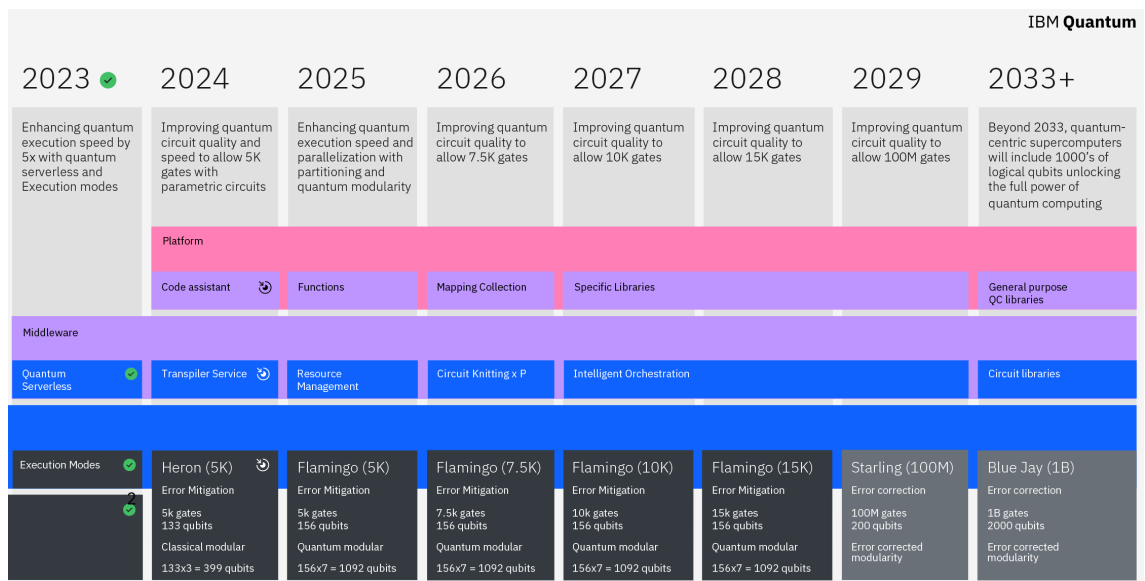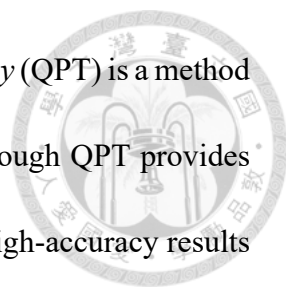
Figure 1.4: Roadmap of IBM quantum computing

1.2 Proposed Techniques

In this paper, our technique diagnoses the *faulty qubit*, *faulty gate type* and *fault size* with high accuracy. The faulty qubit and gate type represent which qubit and gate type a fault acts on. The fault size represents how large the fault effect is. We can diagnose both *coherent faults* and *depolarizing faults*. Coherent faults maintain phase and coherence but may alter the amplitude or orientation of quantum states. Depolarizing faults cause a loss of coherence and shrink the quantum states towards the center of a Bloch sphere. Our overall flow is shown in Fig. 1.3. Our proposed technique contains three stages: *rough diagnosis*, *fine diagnosis*, and *depolarizing diagnosis*. We take the *processor under diagnosis*(PUD) as input. Rough diagnosis locates the approximate range of the coherent fault's size by grid search and reports the faulty qubit and faulty gate type. Fine diagnosis generates QC to distinguish the *diagnosed coherent fault* from others and reports the coherent fault size. Finally, depolarizing diagnosis uses curve fitting to diagnose the *diagnosed depolarizing fault* and report the depolarizing fault size. At the end of our qFD flow, we report the faulty qubit, faulty gate type, coherent fault size, and depolarizing fault size.

Plenty of past research has been conducted on the diagnosis of quantum circuits (QCs) [13], [3], and [4]. These works aimed to diagnose quantum circuits. They are given a quantum circuit and find faulty qubits and faulty gates. Our work, on the other hand, aims to diagnose the quantum processor. We generate diagnosis quantum circuits to find faulty

qubits, faulty gate types, and fault sizes. *Quantum Process Tomography* (QPT) is a method to characterize errors in QPs by reconstructing the faulty gate. Although QPT provides detailed information about a QP, the number of shots needed to get high-accuracy results is too high. *Randomized benchmarking* (RB) is another method to characterize errors in QPs by inserting random gates into an empty QC [5]. Although RB can calculate *error per gate*, it cannot diagnose the fault size, making it difficult to mitigate the effect of an error.

The advantages of our proposed diagnosis technique are as follows: First, we can diagnose both coherent and depolarizing faults, which cause severe errors in today's NISQ processors. Second, our qFD flow can achieve 16 times lower diagnosis error (defined later) compared with quantum process tomography within the same number of shots. Third, users can specify their confidence level (CL) and our diagnosis tool can find a coherent fault within a minimum number of shots. Last, we found an important fact, that the diagnosis error of coherent fault has little impact on the diagnosis accuracy of depolarizing fault.

1.3 Contributions

Our work has five main contributions.

1. We propose a fault diagnosis flow for quantum processors. All steps in the flow are automatic. The flow can diagnose noisy quantum processors in the presence of

faults, which causes quantum processors to malfunction.
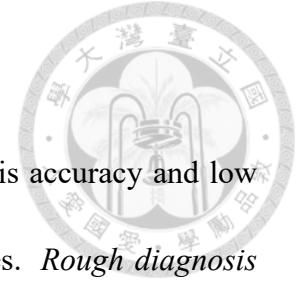
2. We propose three diagnosis stages that achieve good diagnosis accuracy and low diagnosis error. The proposed technique contains three stages. *Rough diagnosis* searches the Bloch sphere and reports the faulty qubit and faulty gate type. *Fine diagnosis* reports the coherent fault size. *Depolarizing diagnosis* reports the depolarizing fault size.

3. We can diagnose both coherent and incoherent (depolarizing) faults. Both faults cause severe errors in modern quantum processors. We report the faulty qubit, faulty gate type, and fault size of both coherent and incoherent (depolarizing) faults.

4. Our diagnosis flow is compared with traditional error characterize method: quantum process tomography. We can achieve better diagnosis accuracy with the same number of shots. We require less shots than quantum process tomography to achieve the same diagnosis accuracy.

5. The proposed technique is applicable on real quantum processors. We perform our diagnosis flow on a real IBM quantum processor. The results also achieve good diagnosis accuracy, indicating that our flow is feasible on real quantum processors. Our work focuses on diagnosing faults whose fault sizes are large.

1.4 Organization

The rest of this paper is organized as follows. Chapter 2 provides information about quantum circuit concepts, past research for quantum circuit diagnosis, the methodology that inspired our work, and fault modeling of quantum processors. Chapter 3 explains the details of our qFD flow. Chapter 4 shows the experimental results on a classical simulator and an IBM quantum processor. Chapter 6 concludes this paper.

Chapter 2 Background

2.1 Quantum Circuit Concepts

We introduce some basic concepts of QCs and the metrics used in this paper. The quantum state $|\psi\rangle$ of an n-qubit QC is expressed as [14]:

$$|\psi\rangle = \sum_{i=0}^{2^n - 1} a_i |i\rangle$$

, where a_i represents the amplitude of the i_{th} basis state $|i\rangle$. Measurement in QCs would collapse the quantum state into one of the 2^n states. The probability of observing $|i\rangle$ after measurement is $|a_i|^2$, while the summation of $|a_i|^2$ for all i equals 1. Measurement results are generated by Monte Carlo method and represented as an *output probability distribution* (OPD). The number of repetitions used in the Monte Carlo method is defined as *shots*. We then introduce a quantum circuit (QC). A quantum circuit is analogous to a classical circuit in classical computing. It consists of a sequence of quantum gates applied to qubits, the fundamental units of quantum information. Each gate manipulates the quantum state of the qubits in a specific way, such as flipping the state, entangling qubits, or performing
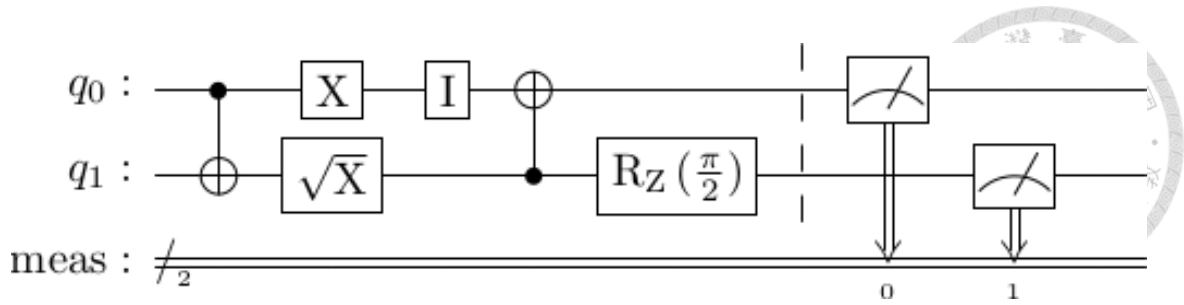
Figure 2.1: Example of a quantum circuit

rotations in the Bloch sphere. Quantum circuits are designed to perform specific quantum computations or algorithms. The execution of a quantum circuit involves initializing qubits in a certain state, applying the sequence of gates according to the circuit design, and measuring the final state of the qubits to obtain the desired output.

Figure 2.1 is an example of a quantum circuit. In this figure, a horizontal line represents a qubit in a quantum circuit, and every element on the line before the vertical dotted line represents a quantum gate. The two elements after the vertical dotted line are measurement operations. They perform measurement and transport the classical bits to the bit lines that are marked $meas$. We then introduce all quantum gates used in Fig.2.1. A X gate is a single-qubit rotation through π radians around the x-axis, which is expressed as follows:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

A $\sqrt{X}$ gate performs the rotation about the X-axis by $\pi/2$ radians, which is expressed as follows:

$$\sqrt{X} = \frac{1}{2}\begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}$$

An I gate is a single-qubit operation that leaves the basis state unchanged, which is ex-

pressed as follows:

$$ID = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

A R_Z gate is a single-qubit rotation through angle θ (radians) around the z-axis, which is expressed as follows:

$$RZ(\theta) = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}$$

A $CNOT$ gate is expressed as follows:

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

A CNOT gate leaves the control qubit unchanged and performs a Pauli-X gate on the target qubit when the control qubit is in state $|1\rangle$. When the control qubit is in state $|0\rangle$, the CNOT gate leaves the target qubit unchanged.

2.2 Fault Modeling and Diagnosis Metrics of QP

Before we diagnose the PUD, we must define our fault model. In [13] and [3], they apply the missing gate fault model. However, QCs have a much more complicated and unpredictable behavior than conventional circuits, so it is unrealistic to apply a missing gate fault model. As a result, in both [6] and [4], they applied three kinds of behavior

10

fault models, which are *ratio*, *bias* and *truncation*. Ratio fault degrades the parameters of a gate by a fixed ratio. Bias fault adds or minuses a fixed number of gate parameters. Truncation fault truncates gate parameters to a fixed value. However, since they choose $U3(\theta, \phi, \lambda)$ as their only basis gate, their fault behavior is limited. Also, they focus on diagnosing QCs but not QPs, while we aim to diagnose the faulty effect that appears on hardware such as qubits and resonators in QPs.

There are two important fault models: coherent fault and incoherent fault. In thesis, we are able to diagnose rotation fault and depolarizing fault. Figure 2.2 is an example of the effect of a coherent fault. The green quantum state is the original fault-free state. The coherent fault causes the fault-free state state to rotate to the faulty state (red). The faulty state differs from the fault-free state by an angle of $pi/4$. The state fidelity decreases from 1 to 0.85. Figure 2.3 is an example of the effect of a depolarizing fault [2]. A depolarizing fault causes the quantum state to shrink towards the center of the Bloch sphere uniformly. The smaller red sphere inside the Bloch sphere is the quantum state affected by a depolarizing fault.

All fault models apply qATG's [6] assumption that a fault will affect the same gate on a qubit. We define a faulty gate for coherent faults, $rf(\theta, \phi)$, which is expressed as:

$$rf(\theta, \phi) = RZ(\frac{\phi}{2} + \frac{\pi}{2}) \cdot RX(\frac{\theta}{2}) \cdot RZ(\frac{\phi}{2} - \frac{\pi}{2}) \tag{2.1}$$

$$\theta \in [0, 2\pi], \quad \phi \in [0, 2\pi]$$

11

We define θ and ϕ as the fault size for a coherent fault. A $rf(\theta, \phi)$ applies an arbitrary rotation after the fault-free gate, causing the quantum state on the Bloch sphere to rotate to another position. The rotation consists of RZ, RX, and RZ gates sequentially, also known as *ZXZ decomposition* [15].
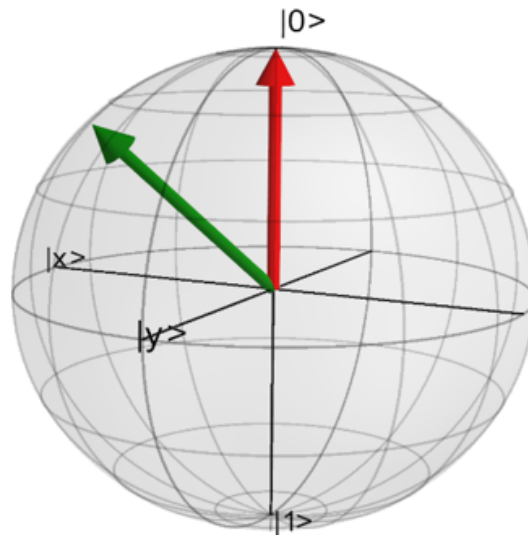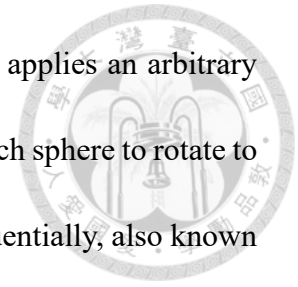
Figure 2.2: Demonstration of a coherent fault
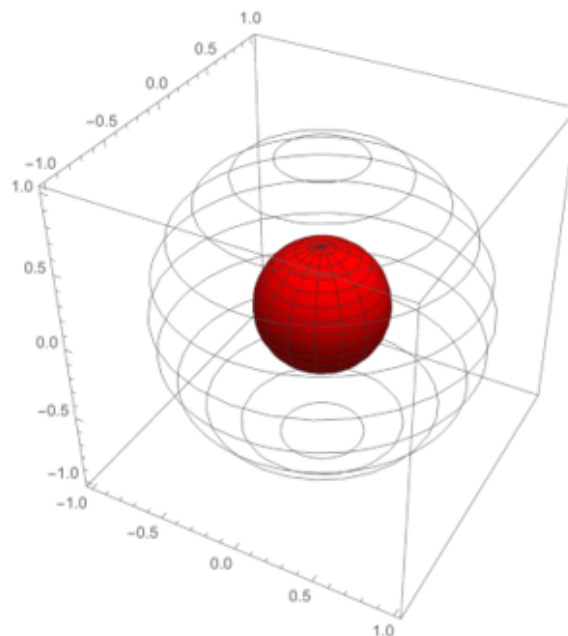
Figure 2.3: Demonstration of a depolarizing fault [2]

Here we demonstrate how we implement ZXZ decomposition in Fig.2.4. Assume we

12
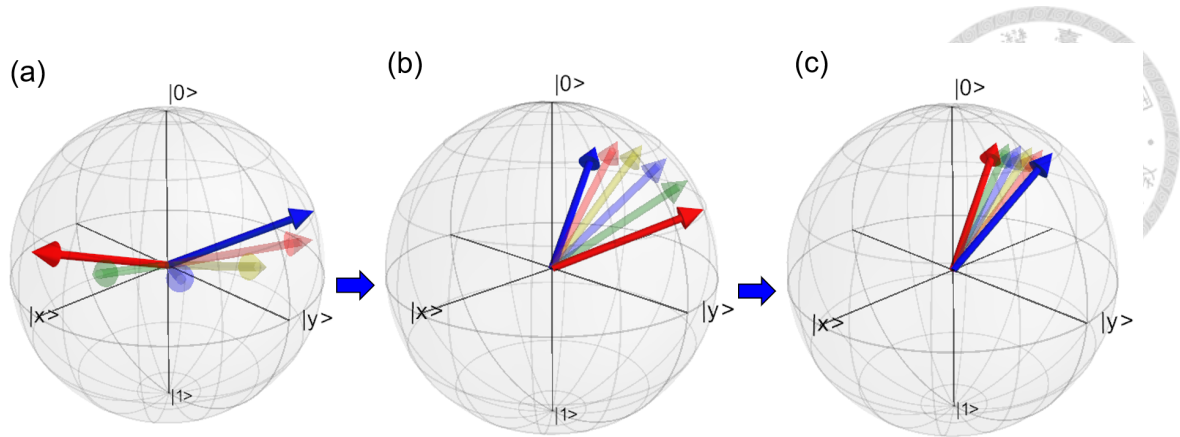
Figure 2.4: Demonstration of $rf(1,3)$

have a $rf(1,3)$. Note that the red-colored state represents the state before rotation, while the blue-colored state represents the state after rotation. In Fig.2.4(a), the initial state is rotated by the $RZ(1/2 + \pi/2)$ gate. Next in Fig.2.4(b), the state is then rotated by the $RX(3)$ gate. Finally in Fig.2.4(c), the state is rotated by the $RZ(1/2 - \pi/2)$ gate.

As for coherent faults in a two-qubit gate, we modify the fault model from [6]. A $rf(\theta, \phi)$ and its conjugate are applied before and after a CNOT gate, respectively. There are three kinds of fault models, which are $control$, $target$, and $both$. The name of the fault model represents which qubit the fault injects into. When a fault is injected into a control qubit of a CNOT gate, it reflects an imperfect control capability, whereas injecting a fault into the target qubit of a CNOT gate represents the imperfect influence of the control qubit. Finally, both scenarios may happen simultaneously, so the fault may exist in both the control and target qubits. A demonstration is shown in Fig. 2.5. In the figure, $U1$ and $U2$ are different $rf(\theta, \phi)$, while $U1$ and $U1^\dagger$ represent the control fault model. $U2$ and $U2^\dagger$ represent the target fault model. At last, both fault model consists of $U1$, $U1^\dagger$, $U2$, and $U2^\dagger$.

13   doi:10.6342/NTU202400956
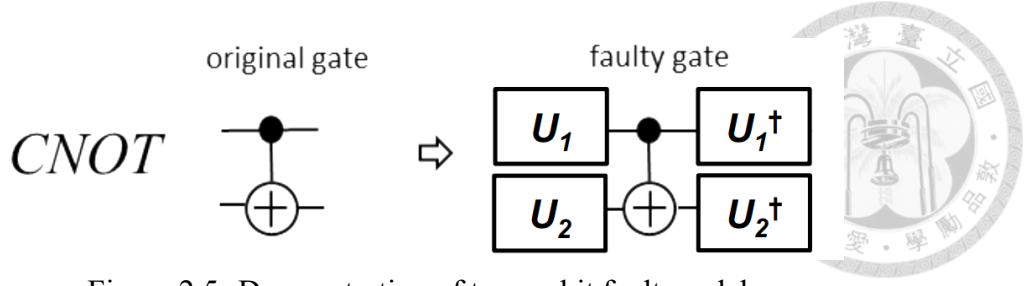
Figure 2.5: Demonstration of two-qubit fault model

To evaluate the effectiveness of our diagnosis, we introduce two key metrics: *diagnosis accuracy* and *diagnosis error*. Diagnosis accuracy is measured by the *average gate fidelity* between the diagnosed faulty gate and the actual faulty gate in a QP. Ranging from 0 to 1, a diagnosis accuracy of 1 denotes an exact match between the diagnosed and actual faulty gate. The equation of diagnosis accuracy is shown below:

$$DA(\mathcal{E}, U) = \frac{dF_{\mathrm{pro}}(\mathcal{E}, U) + 1}{d + 1} \tag{2.2}$$

In the equation, $F_{\mathrm{pro}}(\mathcal{E}, U)$ is the *process fidelity* between the diagnosed faulty gate $\mathcal{E}$ and the actual faulty gate U. d is the dimension of the channel. For example, the actual faulty gate is $rf(1, 0)$, and the diagnosed coherent fault is $rf(1.1, 0)$. We first calculate the process fidelity of $rf(1, 0)$ and $rf(1.1, 0)$. Since both $rf(1, 0)$ and $rf(1.1, 0)$ are unitary matrices, the process fidelity can be expressed as:

$$F_{\mathrm{pro}}(\mathcal{E}, U) = \frac{Tr[S_U^\dagger S_{\mathcal{E}}]}{d^2} \tag{2.3}$$

Note that the Tr is the trace operator, $S_U^\dagger$ is the conjugate transpose of the actual faulty gate *superoperator matrix*, and $S_{\mathcal{E}}$ is the diagnosed coherent fault superoperator matrix. We calculate the conjugate transpose matrix of $rf(1.1, 0)$, which is:

$$\begin{bmatrix} 0.73 & 0.45j & -0.45j & 0.27 \\ 0.45j & 0.73 & 0.27 & -0.45j \\ -0.45j & 0.27 & 0.73 & 0.45j \\ 0.27 & -0.45j & 0.45j & 0.73 \end{bmatrix}$$

Then we multiply the conjugate transpose superoperator matrix of $rf(1.1, 0)$ with $rf(1, 0)$ superoperator matrix and calculate the trace.

$$F_{\text{pro}}(rf(1.1, 0), rf(1, 0)) = \frac{3.99}{2^2} = 0.9975$$

Finally, we calculate the diagnosis accuracy equals to 99.83% :

$$DA(rf(1.1, 0), rf(1, 0)) = \frac{2 * 0.9975 + 1}{2 + 1} = 99.83\%$$

Diagnosis error represents the inaccuracy between the diagnosed faulty gate and the actual faulty gate. We define diagnosis error as 1 minus the diagnosis accuracy.

$$DE(\mathcal{E}, U) = 1 - DA(\mathcal{E}, U) \tag{2.4}$$

The diagnosis error of the example is $1 - 99.83\% = 0.17\%$.

Unlike average gate fidelity, *state fidelity* measures the similarity between two quan-

15

tum states. The equation of state fidelity is shown as follows:

$$F = |\langle \psi | \phi \rangle|^2 \tag{2.5}$$

Note that ψ and ϕ are the two quantum states being compared.

For the depolarizing fault model, we define a parameter λ as the fault size. λ is the ratio at which a state shrinks towards the center of the sphere. It ranges from 0 to 1, where 0 represents the fault-free state, and 1 represents the *maximally mixed state*. The equation of depolarizing fault is:

$$E(\rho) = (1 - \lambda)\rho + \lambda \mathrm{Tr}[\rho]\frac{I}{2^n} \tag{2.6}$$

Note that Tr is the trace operator, $E(\rho)$ is a function that takes fault-free state (ρ) as input and outputs the quantum state injected with depolarizing fault, and n is the number of qubits. We then define the diagnosis error of depolarizing faults as follows:

$$DE(\mathcal{E}, U) = \|\lambda_{\mathcal{E}} - \lambda_U\| \tag{2.7}$$

The $\lambda_{\mathcal{E}}$ and λ_U are the fault sizes of diagnosed faulty gate and actual faulty gate. For example, the fault size of the actual faulty gate is 0.90, and the fault size of the diagnosed coherent fault is 0.88. Then the diagnosis error would be $\|0.90 - 0.88\| = 0.02$.

2.3 Past Research for QC Diagnosis
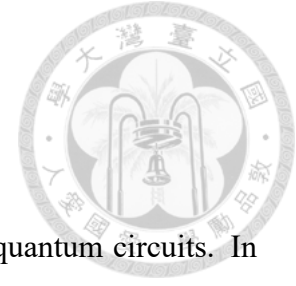
We analyze the pros and cons of past research for diagnosing quantum circuits. In [13], they proposed a *binary tomographic test* method to diagnose faults on phase gates. A binary tomographic test consists of the number of shots, the quantum circuit, and the input vectors. They slice the quantum circuit and keep the part they aim to diagnose. Then they apply gates before and after their target sliced quantum circuit. The gates applied before the target sliced quantum circuit represent manipulated input vectors. The gates applied after the target sliced quantum circuit represent measurement operators. After measuring the target sliced quantum circuit, they can diagnose phase gate faults in the QC.

In [3], they apply a similar binary tomographic test method. Different from [13], they calculate the optimal input vectors according to the quantum state before and after the target sliced quantum circuit. They conclude their paper by applying their strategy to benchmark circuits like the one in Fig.2.6. They can diagnose all six faults in Fig.2.6 same as [13]. Moreover, their method reports all six faults as their only suspect fault. While in [13], they report two suspect faults in five of the six faults.

However, these two methods cannot operate in noisy environments. As a result, [4] proposed a diagnosis method to locate a fault's location and behavior in noisy environments. They executed QCs that were injected with different faults separately. Then, they can report the fault whose OPD is the most similar to the OPD of the circuit under diagnosis. An example can be seen in Fig.2.7 [4]. The figure demonstrates an example of the
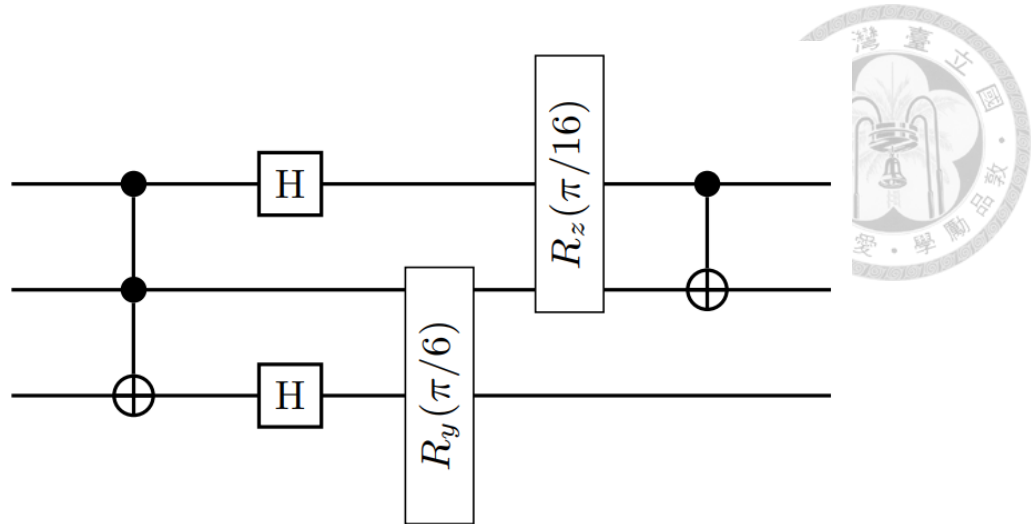
Figure 2.6: Benchmark circuit for diagnosis (figure taken from [3])

OPD$_{f4}$
(0.499, 0.475, 0.013, 0.013)

OPD$_{f1}$
(0.490, 0.490, 0.010, 0.010)

OPD$_{CUD}$
(0.499, 0.476, 0.009, 0.016)

OPD$_{f3}$
(0.500, 0.473, 0.014, 0.013)

distance

OPD$_{f2}$
(0.474, 0.499, 0.014, 0.013)
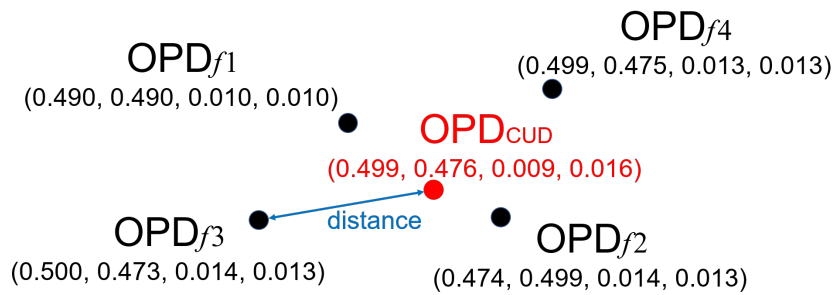
Figure 2.7: Distance between OPD of the circuit under diagnosis and OPD of fault f_3 proposed in [4]

distance between OPD of the circuit under diagnosis and OPD of fault f_3. The smaller the distance between two OPDs, the higher their similarity.

In summary, [13], [3], and [4]'s work diagnoses whether there is a faulty gate in a QC. However, QCs are configurable software in superconducting quantum processors. Suppose there is a hardware defect in the quantum processor, then the hardware-induced fault location might not be the same in different QCs. [13], [3], and [4] are not enough to identify the hardware defect. On the other hand, our work diagnoses quantum processors. We only need to diagnose the QP once to find the hardware-induced fault location. We do not have to diagnose many different QC.

18

2.4 Past Research for QP Diagnosis

A well-known method to characterize errors in QPs is QPT [16]. QPT involves systematically preparing a set of input states, applying the target quantum operation to each input state, and then performing measurements to determine the resulting output states. By collecting and analyzing the measurement data, they can reconstruct the complete description of the quantum process, represented by a *Choi matrix* [17]. Although QPT is the most accurate method to characterize error in QPs, it often requires too many shots to achieve high diagnosis accuracy results. For example, QPT requires more than 10^7 shots to achieve 99.95% diagnosis accuracy.

Another common technique to characterize errors in QPs is RB [5]. RB involves repetitively inserting random quantum gates into a QC while ensuring the final output state remains $|0\rangle$. RB measures the error rate by creating a curve correlating the QC's depth with the probability of measuring the state $|1\rangle$. From the curve, RB can report the error per gate. Fig. 2.8 shows the probability of measuring $|0\rangle$ as a function of the number of steps for each randomized gate [5]. The state fidelity is plotted on a logarithmic scale. By measuring the slope of the line we can get the size of the error per gate. RB is much faster compared with QPT. However, RB cannot report the fault size of a faulty quantum gate, preventing direct error mitigation techniques.
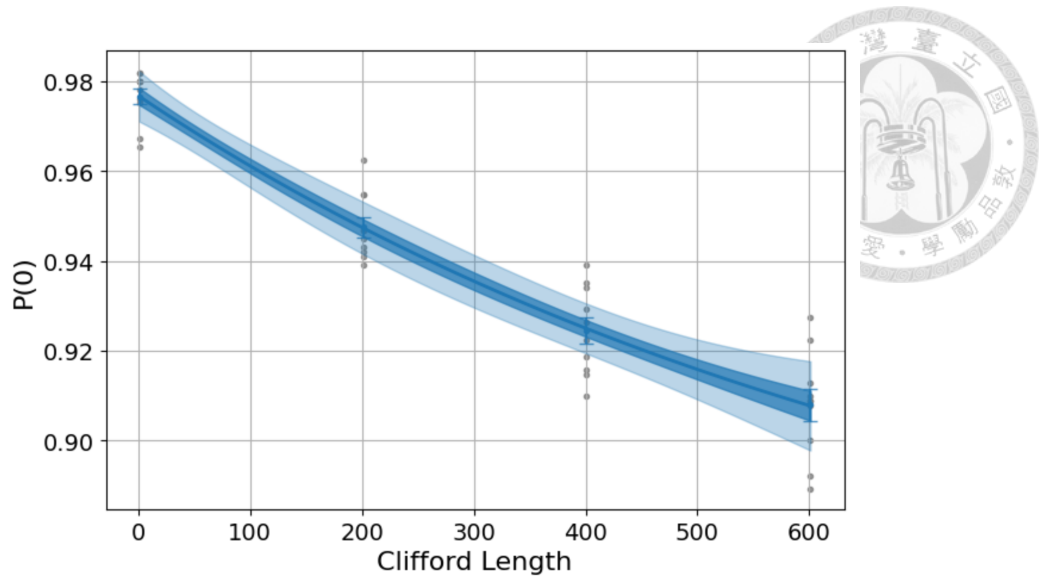
Figure 2.8: Demonstration of Randomized Benchmarking [5]

2.4.1 qATG

We now introduce the method upon which our work is based: Automatic Test Generation for Quantum Circuits (qATG) [6]. The main goal of qATG is creating *test configurations* to test if a QC is faulty or not. The flow of qATG is shown in Fig.2.9. Given a fault, they recursively insert gates into an empty QC *test template* to maximize the state fidelity between the faulty and fault-free states. While the faulty effect might not accumulate if they kept stacking up the faulty gate, they added an *activation gate* before a faulty gate. A faulty gate and an activation gate combined is called a *test element*. They optimize the parameters of an activation gate using gradient descent. In our work, we apply a similar methodology in our rough diagnosis phase to create a QC that maximizes the state fidelity between the faulty and fault-free states. The usage of the QC created in the rough diagnosis stage and test configuration in qATG is different. In qATG, they apply statistics methods (non-central chi-square distribution) to test if the QP is faulty. In our work qFD,
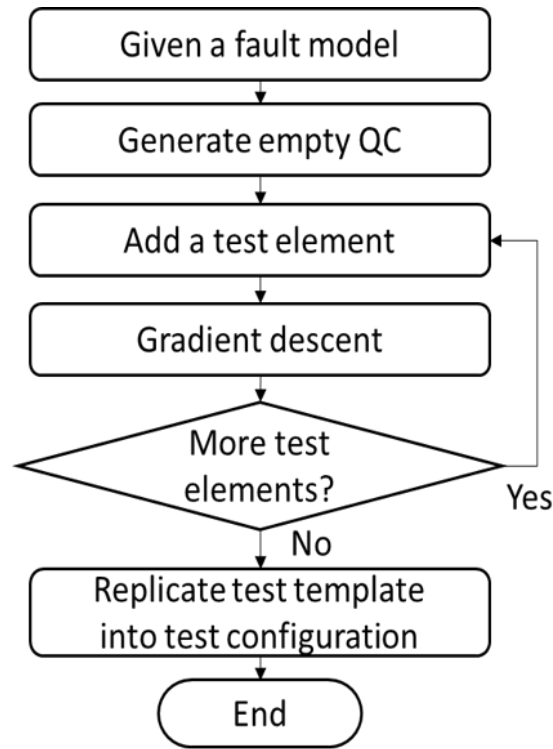
Figure 2.9: Flow of qATG

we aim to diagnose faults in the QP.

Figure 2.10 demonstrates an example test configuration (QC) on a 5-qubit quantum processor. In the figure, there are five test templates, each of which tests a qubit. Each test template has four test elements, each of which activates and propagates the fault effects so that we can easily observe them at outputs.
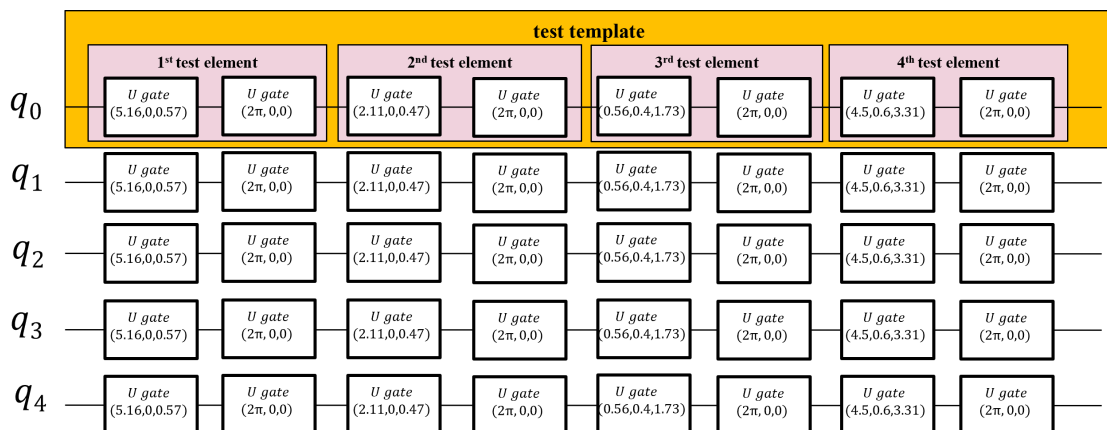
Figure 2.10: Example QC generated by qATG [6]

21

Chapter 3 Proposed Technique

Figure 3.1 is our overall flow. Given a PUD and a user-defined search space of coherent fault model, our diagnosis flow can report two diagnosed fault size for each qubit and each gate type of the PUD. The main stage of our diagnosis flow contains three phases: rough diagnosis, fine diagnosis, and depolarizing diagnosis. Rough diagnosis first generates the corresponding *rough diagnosis QC* for each coherent fault. Then rough diagnosis performs fault simulation and dropped non-suspect faults, then locates an approximate range of the suspect coherent fault's size. Since we diagnose each qubit and gate type separately, if the suspect coherent fault size is not zero, we report the faulty qubit and faulty gate type. Then, if the resolution of the grid search is larger than the minimum resolution we defined, we restrict the search space and generate new rough diagnosis QCs again. We then run fine diagnosis to search for the fault size of the coherent fault. After we finish diagnosing the coherent fault, we run depolarizing diagnosis and report the depolarizing fault size. The outcomes are the faulty qubit, faulty gate type, coherent fault size, and depolarizing fault size. We then repeat the main stage of our diagnosis flow until all qubits and gate types are diagnosed.
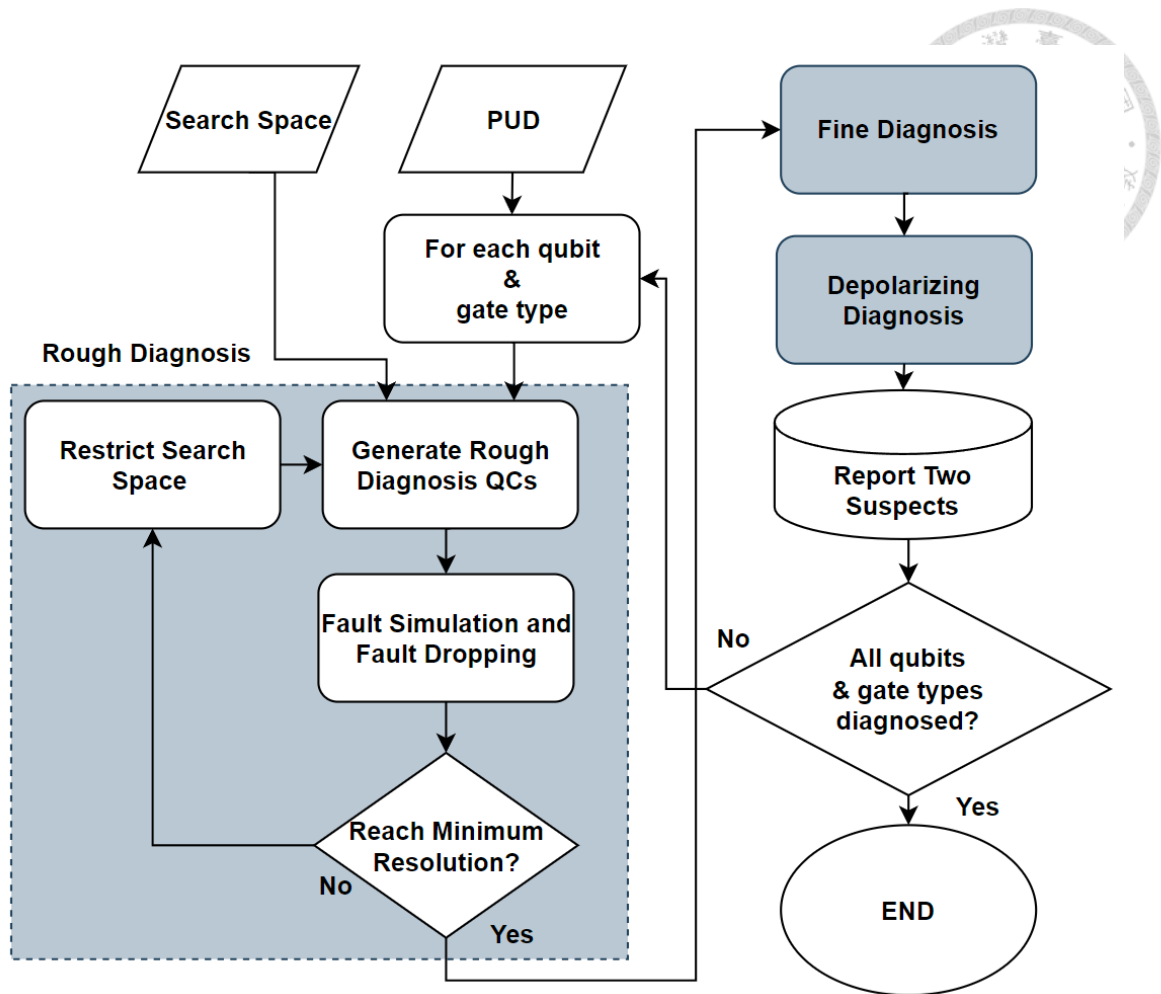
Figure 3.1: Overall Flow

3.1 Rough Diagnosis

In rough diagnosis, we conduct fault simulation and filter faults whose OPD differs significantly from the PUD's OPD. Before diagnosis, we are given a search space for our diagnosed coherent fault. We then divide the range of coherent fault sizes into uniform grids, where each corner of a grid is a fault with its distinct pair of (θ, ϕ). The grid size is defined as our *resolution* and is decided by the total number of faults. To achieve the best results with good resolution and a short time, we construct a fault list of around 100 faults. A demonstration of the search space can be seen in Fig.3.2. A part of the Bloch sphere
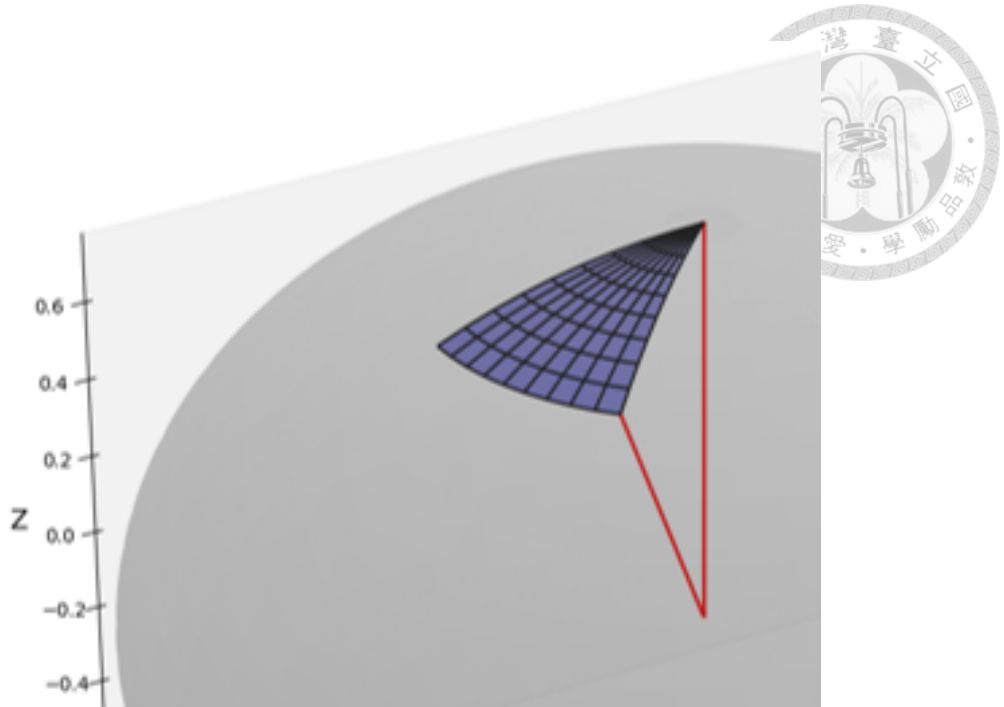
Figure 3.2: Demonstration of search space

is shown in the figure. Each intersection of the grid lines on the Bloch sphere's surface represents a coherent fault of different fault sizes.

We first generate rough diagnosis QCs using qATG [6] for every fault. A rough diagnosis QC can maximize the difference between the fault-free and faulty state, which is identical to the test configuration in qATG [6]. We then execute every rough diagnosis QC on the PUD. To gain more information, we measure the OPD on three axes rather than the Z-axis only. We denote the result OPD of the PUD as OPD_{PUD}. Next, we perform fault simulation for each coherent fault. Starting from a rough diagnosis QC, we inject every fault into it and measure the OPD on three axes as well. To distinguish if the OPD of an injected fault f OPD_f is close enough compared with the OPD_{PUD}, we introduce the *Hellinger Distance* [18] as our metric. The Hellinger distance is defined as follows. P and Q are two discrete probability distributions where $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n)$.

The Hellinger distance $HD(P, Q)$ between P and Q is shown as equation (3.1).

$$HD(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{n} (\sqrt{p_i} - \sqrt{q_i})^2} \qquad (3.1)$$

State Fidelity $F(P, Q)$ of its corresponding Hellinger distance is shown as equation (3.2).

$$F(P, Q) = (1 - HD^2(P, Q))^2 \qquad (3.2)$$

We define a threshold TH_R so that a rough diagnosis QC's Hellinger Distance between its faulty and fault-free state must exceed TH_R. Since we divide the search space into uniform grids, we can set a fixed threshold $thres_R$ based on the resolution. We then calculate the Hellinger distance between OPD_{PUD} and OPD_f. If the difference between them is larger than $thres_R$, the fault f is dropped from the fault list. The rough diagnosis QC generated by fault f is also removed. We then repeatedly perform fault simulation on the remaining rough diagnosis QCs. After we finish fault simulation on every rough diagnosis QC, we determine if the resolution is smaller than our predefined threshold. Otherwise, we restrict the search space by setting the new range based on the minimum and maximum angles from the last step of fault simulation and fault dropping. Then, we set the resolution smaller and generate a fault list with the new search space again. During the process of fault dropping, if all faults are dropped, we increase TH_R until the fault list is not empty. At the end of rough diagnosis, we report multiple faults whose fault sizes are within the smaller search space. Also, if the fault size is not 0 (not fault-free), we report the faulty qubit and faulty gate type.
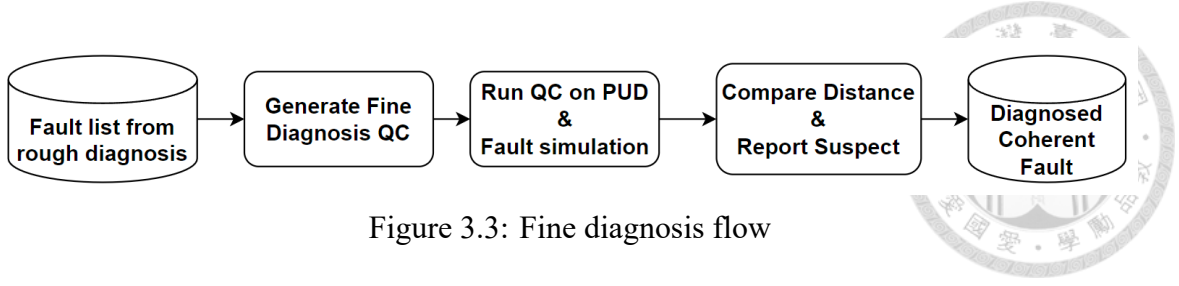
Figure 3.3: Fine diagnosis flow

3.2 Fine Diagnosis

In fine diagnosis, we want to report the diagnosed coherent fault whose OPD is most similar to the PUD's OPD. Fig. 3.3 shows the details of our fine diagnosis flow. The output of rough diagnosis is a fault list with n faults. Fine diagnosis aims to distinguish the only diagnosed coherent fault from the others. We first generate pairs from the fault list, with a total of $\binom{n}{2}$ exhaustive pairs. We generate a *fine diagnosis QC* for each pair of faults. Similar to qATG [6], we apply activation gates to maximize the faulty effect between two faults. The qATG process is iterated until the state fidelity between two faulty states is lower than our defined threshold TH_F.

We then conduct fault simulation on all fine diagnosis QCs and locate the diagnosed coherent fault. We first execute all the fine diagnosis QCs on the PUD. Next, we perform fault simulation of a fine diagnosis QC. Different from rough diagnosis, we compare $HD(OPD_{\text{PUD}}, OPD_{f_1})$ and $HD(OPD_{\text{PUD}}, OPD_{f_2})$, where (f_1, f_2) is the pair of faults that generates the fine diagnosis QC. The fault with a larger Hellinger distance indicates that it is less likely to be the suspect than the other fault. We conduct fault simulations on all fine diagnosis QCs. We identify one fault ($f_{suspect}$) whose OPD is closest to the PUD's OPD than the other $n-1$ fault. Finally, we report $f_{suspect}$ as the diagnosed coherent fault.
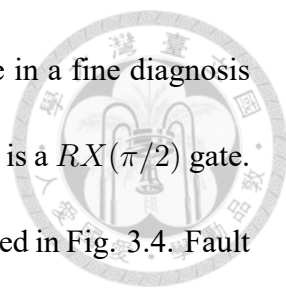
Fig.3.4 shows an example that illustrates how an activation gate in a fine diagnosis QC can distinguish two faults. Suppose the original fault-free gate R_O is a $RX(\pi/2)$ gate. Fault f_1 consists of a R_O plus a $RZ(\pi/4)$ gate (R_{f_1}), which is colored red in Fig. 3.4. Fault f_2 consists of a R_O plus a $RZ(-\pi/4)$ gate (R_{f_2}), which is colored green in Fig. 3.4. Both faulty states initially start in $|0\rangle$, shown in Fig. 3.4(a). We apply the two faulty operations to the two states separately. The resulting two states then have a $\pi/2$ difference shown in Fig. 3.4(b). However, after we apply the faulty operations to the two states again in Fig. 3.4(c), the difference between the two states becomes smaller than in Fig. 3.4(b).

To improve the method above, we added an activation gate before the two faulty gates to extend the differences between the two faulty effects. The activation gate R_{a_d} between Fig. 3.4(d) and Fig. 3.4(e) is an identity gate with no effects. The activation gate R_{a_e} between Fig. 3.4(e) and Fig. 3.4(f) is a $RX(3\pi/2)$ gate, which can maximize the difference between the two states in Fig. 3.4(e). The final states in Fig. 3.4(f) are thus rotated to orthogonal. The example demonstrates activation gates in our fine diagnosis QC can maximize the faulty effect difference between two faults.

To achieve a user-specified diagnosis $confidence level$ (CL), we need to find the minimum number of shots needed. Suppose f_1 is our suspect fault, and we generate a fine diagnosis QC with f_1 and another fault f_2. For a single shot, if $HD(OPD_{\text{PUD}}, OPD_{f_1}) < HD(OPD_{\text{PUD}}, OPD_{f_2})$, we denote this as a *successful shot*. We consider a diagnosis successful when the number of successful shots $(x+1)$ is more than half of the number of total shots $(2x+1)/2$. The probability of a successful diagnosis when we perform fault
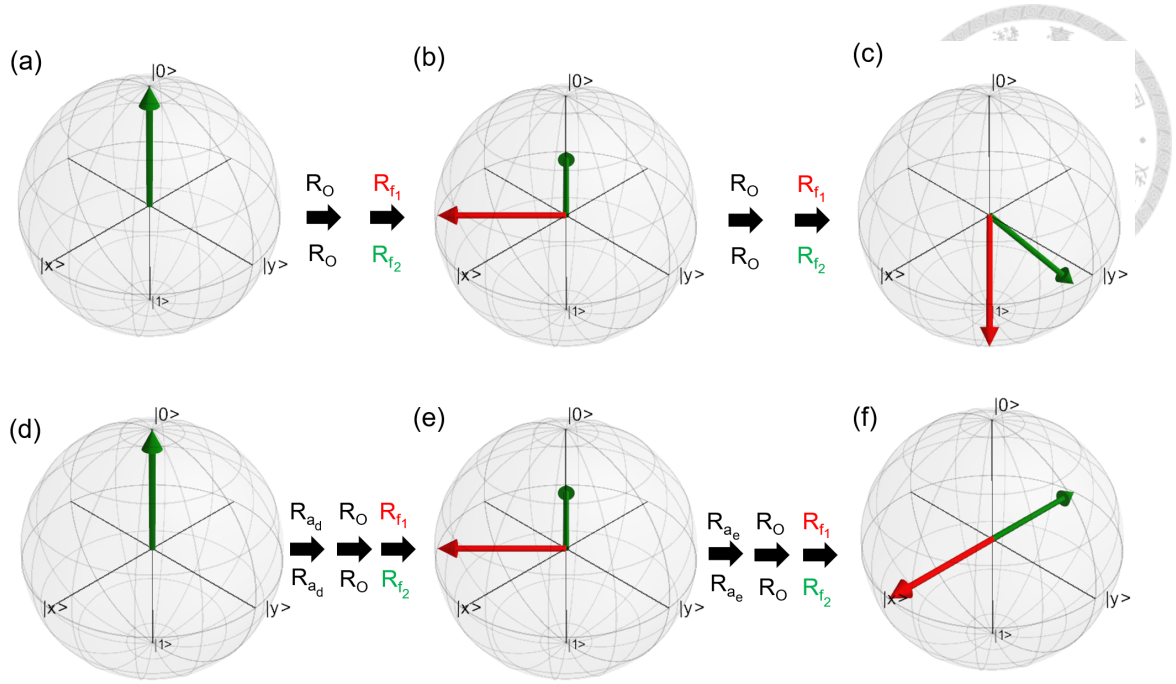
Figure 3.4: Distinguish two faults with activation gate

simulation on a fine diagnosis QC is $\sum_{k=x+1}^{2x+1} \binom{2x+1}{k}(TH_F)^{2x+1-k}(1-TH_F)^k$. For the total

success probability to be above CL, the formula is as follows:

$$ CL \leq \left(\sum_{k=x+1}^{2x+1} \binom{2x+1}{k}(TH_F)^{2x+1-k}(1-TH_F)^k \right)^{n-1} \tag{3.3} $$

In this paper, we set our $CL = 99\%$. According to equation (5), we can obtain the number

of shots $2x+1 = 17$, as long as $n < 200$.

3.3 Depolarizing Diagnosis

Depolarizing diagnosis aims to measure the depolarizing fault size α of the *target*

gate. The target gate is the faulty gate diagnosed by the fine diagnosis. Fig. 3.5 shows the

details of our depolarizing diagnosis flow. We generate a quantum circuit c by an iterative

process. I) Initially, c is an empty circuit. II) We insert a target gate, followed by a gate
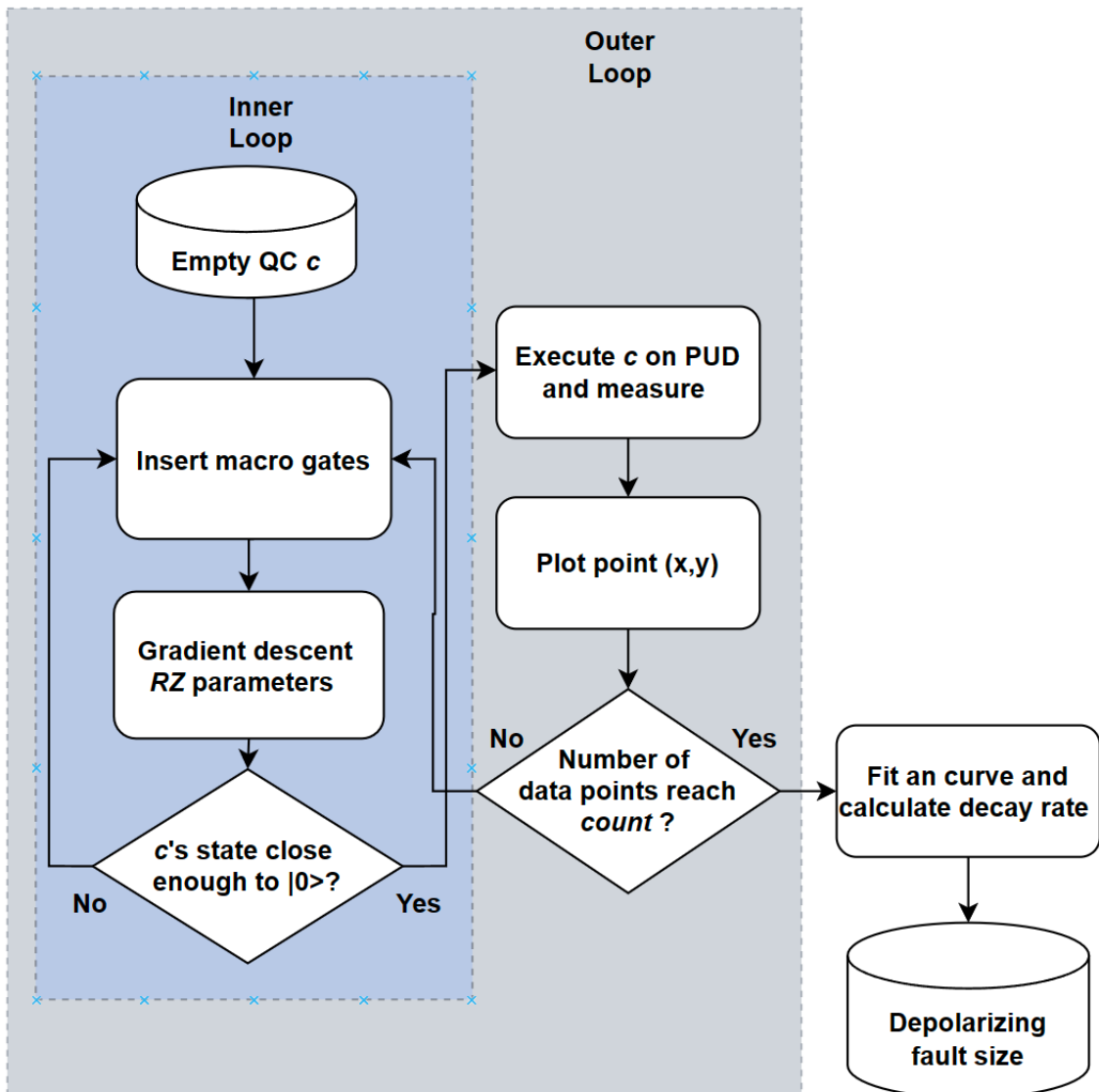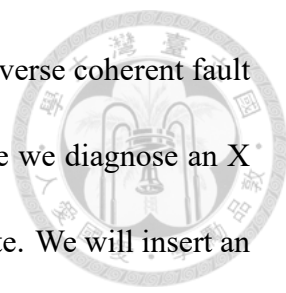
Figure 3.5: Depolarizing diagnosis flow

modeling the inverse diagnosed coherent fault. The purpose of the inverse coherent fault

is to cancel the faulty effect of the target gate. For example, suppose we diagnose an X

gate with a $rf(0.5, 0)$ after fine diagnosis. The target gate is the X gate. We will insert an

X gate into c, followed by an inverse coherent fault (modeled by a $rf(-0.5, 0)$). We then

insert a RZ gate into c. We define the target gate, inverse coherent fault modeled gate, and

the RZ gate together as a *macro gate*. III) On the simulator, we perform gradient descent

on the parameters of the RZ gates. The goal of the RZ gates is to rotate the final state

back to the basis state $|0\rangle$. IV) We repeatedly insert a constant number of macro gates into

c until the state fidelity between the QC's final state and basis state $|0\rangle$ is above a very

high threshold, which is 0.99 in our experiment.

We move on to the outer loop step in Fig.3.5. V) We then execute the QC on the

PUD and measure the OPD. VI) We plot a data point on the *depolarizing diagnosis plot*,

where the x-axis represents the number of gates, and the y-axis represents the probability

of measuring $|0\rangle$. VII) We return to step II until the count of data points reaches the

user-specified number $count$ in the depolarizing diagnosis plot. In our experiment, we set

$count$ as 100. VIII)We can fit a curve: $y = a \cdot \alpha^x + b$ that describes the correlation between

the numbers of a fault-injected gate (x) and the probability of measuring $|1\rangle$ (y). Note that

the a, b, and α are the variables we search to fit the curve, while α is the decay rate of

measuring $|0\rangle$, both a and b are coefficients whose values are around 0.5. The decay rate

α is the fault size of the diagnosed depolarizing fault, which is the goal of our depolarizing

diagnosis. IX) Finally, we report a depolarizing fault from depolarizing diagnosis.

Figure 3.6 shows an exponential decay curve of our depolarizing diagnosis plot. In this figure, we inject a depolarizing fault whose fault size is 0.004. As we add more gates into c, the probability of measuring $|0\rangle$ keeps decreasing because of the depolarizing fault. We can observe that the y value of the curve converges to 0.5 when we add more than 700 gates into c. This is because the quantum state reaches the maximally mixed state.
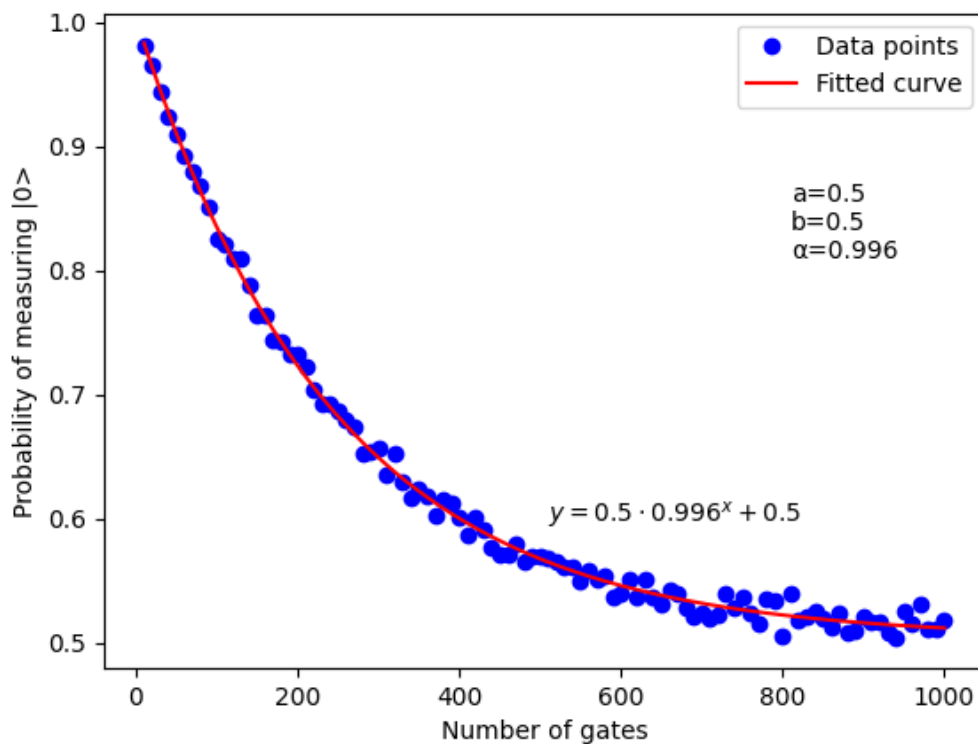
Figure 3.6: Fidelity decay curve of the depolarizing diagnosis

Chapter 4   Experimental Results

4.1   Experiment Setup

We now show the experiment settings. The number of shots to execute the PUD and to perform fault simulation on classical simulators are 1,024 and 4,096, respectively. For rough diagnosis, the first iteration's resolution is 0.5 radius, and the second iteration's resolution is 0.05 radius. We only have two iterations because an angle difference of 0.05 radius is equivalent to 99.99% diagnosis accuracy, which is about the noise size of current NISQ processors. Although we can diagnose larger resolutions, noises mask the effect of a fault. We can diagnose smaller resolutions, but it's important to note that the diagnosis error may increase. However, the overhead, measured in terms of the number of shots and gates required, will decrease. We generate rough diagnosis QC until the state fidelity between fault-free and faulty states (TH_R) is below 0.3. We generate fine diagnosis QC until the state fidelity between two faulty states (TH_F) is below 0.1.

For the gate types we aim to diagnose, we choose X, SX, ID for single-qubit and CX for two-qubit. For most IBM quantum processors, their basis gates are X, SX, ID,

RZ, and CX. IBM QPs implement RZ gates by adding a virtual phase rather than inserting the actual gate into the QC, so there should be no faults on RZ gates. Our diagnosis experiment covers all basis gates most IBM QPs use.
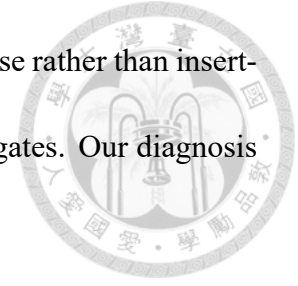
There are four backends on which we run our experiments. The first three backends are run on classical simulators, and the fourth backend is run on a real IBM quantum processor. The first three backends are noise-free, artificial noise model, and *IBM noise model* [19]. We select *Aer simulator* in Qiskit [20] as the classical simulator. Table 4.1 lists the version for Qiskit.

qiskit-terra	0.21.2
qiskit-aer	0.11.0
qiskit-ignis	0.7.1
qiskit-ibmq-provider	0.19.2
qiskit-ibm-provider	0.7.0
qiskit	0.38.0

Table 4.1: Qiskit version

For the artificial noise model backend, we use the depolarizing error as the noise model of gate error. The error rates of single-qubit and two-qubit gates are 0.1% and 1%, respectively. The readout error rate is 1.5%, representing the probability that output is inverted. For the IBM noise model backend, we select the *ibm_hanoi* as the noise model. For the real IBM quantum processor backend, we choose the *ibm_hanoi* in our experiment. Figure 4.1 is the topology of *ibm_hanoi*. The color on the qubit indicates the error rate of it. The lighter the color, the larger the error rate. Note that the Fig.4.1 here is not the condition when we perform out experiment. For the IBM noise model backend, we choose physical qubits 11 and 12 as the PUD to perform our flow, since the noise is
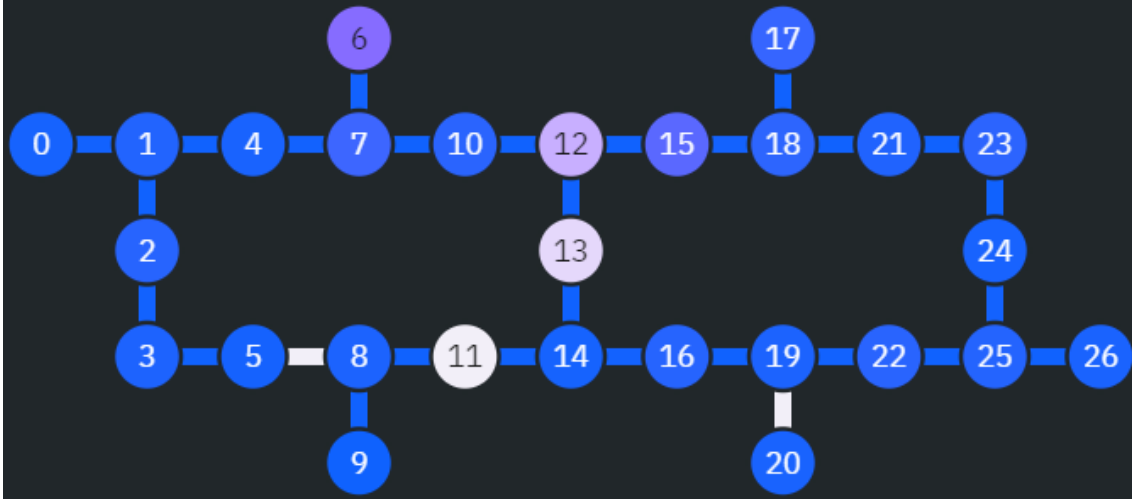
Figure 4.1: Topology of real IBM quantum processor ibm_hanoi (figure taken from [7])

smallest on the qubits. For the real IBM quantum processor, we choose qubits 22 and 25 as the PUD to perform our flow, since the noise is smallest on the qubits.

4.2 Coherent Fault Diagnosis

In this simulation, we inject coherent faults that have been observed in real QPs: *phase drifting faults* [21] and *over-rotation faults* [22]. Phase drifting faults are faults with arbitrary ϕ and $\theta = 0$. Phase drifting faults in quantum processors refer to the gradual change or fluctuation in the relative phase of qubits during quantum operations. It arises due to various factors such as environmental interference, imperfections in control mechanisms, and interactions with neighboring qubits or external fields. Tables 4.2, 4.3 and 4.4 are the diagnosis results of injecting phase drifting faults whose gate type are all ID gates.

Over-rotation faults are faults with arbitrary θ and $\phi = 0$. When an over-rotation error occurs, the rotation angle applied to the qubit exceeds the intended value, leading

Table 4.2: Diagnosis results of phase drifting faults on noise-free backend

Injected(ϕ)	Sus.(θ,ϕ)	DA(%)	Shots	Gates
0.00	(0.00, 0.00)	100.00	448,529	9,258
0.05	(0.04, 0.03)	99.92	492,765	9,535
0.10	(0.00, 0.10)	100.00	522,257	9,720
0.20	(0.00, 0.20)	100.00	1,075,217	14,730
0.40	(0.00, 0.40)	100.00	522,257	9,972
0.79	(0.00, 0.80)	100.00	559,134	6,975
1.57	(0.07, 1.51)	99.78	589,895	8,604
3.14	(0.00, 3.15)	100.00	691,302	10,148
4.71	(0.00, 4.70)	100.00	1,603,713	16,678
Average	**N/A**	**99.97**	**722,785**	**10,624**

Table 4.3: Diagnosis results of phase drifting faults on artificial noise model backend

Injected(ϕ)	Sus.(θ,ϕ)	DA(%)	Shots	Gates
0.00	(0.00, 0.00)	100.00	448,529	9,258
0.05	(0.10, 0.00)	99.79	448,529	9,258
0.10	(0.00, 0.10)	100.00	522,257	9,720
0.20	(0.00, 0.20)	100.00	522,257	9,972
0.40	(0.00, 0.40)	100.00	522,257	9,972
0.79	(0.00, 0.80)	100.00	595,998	7,315
1.57	(0.00, 1.55)	99.99	731,333	15,030
3.14	(0.00, 3.15)	100.00	927,869	12,610
4.71	(0.00, 4.70)	100.00	2,042,971	18,529
Average	**N/A**	**99.98**	**751,333**	**11,296**

to deviations from the desired quantum operation. These errors can arise due to various factors, including imperfections in the hardware and environmental noise. Tables 4.5, 4.6 and 4.7 are the diagnosis results of injecting over-rotation faults whose gate type are all X gates. We perform five diagnoses for each fault and show the average result. We explain the results of the experiments from the tables. Column 1 shows the diagnosed suspect size we reported; column 2 shows the diagnosis accuracy; columns 3 and 4 show the number of shots and basis gates used, respectively. The results show that our method achieves a very high diagnosis accuracy of over 99.97% in this simulation. We also discover that the number of shots is proportion to the number of gates used. This is because when we execute a rough diagnosis QCs (total number of gates increase), the number of shots used

35

Table 4.4: Diagnosis results of phase drifting faults on IBM noise model backend

Injected(ϕ)	Sus.(θ,ϕ)	DA(%)	Shots	Gates
0.00	(0.00, 0.00)	100.00	448,529	9,366
0.05	(0.03, 0.04)	99.95	503,859	4,955
0.10	(0.00, 0.10)	100.00	522,257	9,720
0.20	(0.00, 0.20)	100.00	522,257	9,972
0.40	(0.00, 0.40)	100.00	522,257	9,972
0.79	(0.00, 0.80)	100.00	559,175	8,836
1.57	(0.02, 1.54)	99.93	626,841	12,096
3.14	(0.00, 3.15)	100.00	995,436	12,248
4.71	(0.00, 4.70)	100.00	1,569,934	17,041
Average	**N/A**	**99.98**	**693,856**	**10,966**

is also increased.

Table 4.5: Diagnosis results of over-rotation faults on noise-free backend

Injected(θ)	Sus.(θ,ϕ)	DA(%)	Shots	Gates
0.00	(0.00, 0.00)	100.00	448,532	9,553
0.05	(0.05, 0.00)	100.00	463,281	9,537
0.10	(0.10, 0.00)	100.00	507,518	9,912
0.20	(0.20, 0.00)	100.00	691,217	13,590
0.40	(0.40, 0.00)	100.00	691,217	12,210
0.79	(0.80, 0.00)	100.00	711,179	13,408
1.57	(1.56, 0.01)	99.99	691,322	11,191
2.35	(2.35, 0.00)	100.00	826,643	16,434
3.09	(3.10, 0.00)	100.00	405,653	13,068
Average	**N/A**	**100.00**	**604,062**	**12,100**

Next, we inject faults with random θ and ϕ to show our flow can work successfully in our simulation. We inject 9 single-qubit faults and run our flow on the three backends. For single qubit faults, we show the results in Tables 4.8, 4.9 and 4.10. In Tables 4.8, 4.9 and 4.10, the header column is the gate type we inject a fault into. Columns 1 and 2 are the injected fault sizes θ and ϕ. The rest of the table follows the same format as Tables 4.2. The diagnosis accuracies of the 9 single-qubit results are 99.95%, with the 6 two-qubit results being 99.98%.

Then we inject 6 two-qubit faults and run our flow on the three backends. For two

Table 4.6: Diagnosis results of over-rotation faults on artificial noise model backend

Injected(θ)	Sus.(θ,ϕ)	DA(%)	Shots	Gates
0.00	(0.00, 0.00)	100.00	448,529	9,366
0.05	(0.05, 0.00)	100.00	478,020	9,345
0.10	(0.10, 0.00)	100.00	890,897	13,038
0.20	(0.20, 0.00)	100.00	1,271,825	17,181
0.40	(0.40, 0.00)	100.00	691,217	12,210
0.79	(0.80, 0.00)	100.00	585,825	11,165
1.57	(1.55, 0.00)	99.99	691,315	10,730
2.35	(2.35, 0.00)	100.00	691,519	17,384
3.09	(3.10, 0.00)	100.00	405,776	29,332
Average	**N/A**	**100.00**	**683,880**	**14,417**

Table 4.7: Diagnosis results of over-rotation faults on IBM noise model backend

Injected(θ)	Sus.(θ,ϕ)	DA(%)	Shots	Gates
0.00	(0.00, 0.00)	100.00	448,532	9,661
0.05	(0.05, 0.00)	100.00	478,027	9,727
0.10	(0.10, 0.00)	100.00	890,897	13,038
0.20	(0.20, 0.00)	100.00	1,367,057	18,150
0.40	(0.40, 0.00)	100.00	691,217	12,210
0.79	(0.80, 0.00)	100.00	1,084,443	13,508
1.57	(1.56, 0.00)	99.99	691,353	11,608
2.35	(2.35, 0.00)	100.00	894,173	15,458
3.09	(3.10, 0.00)	100.00	405,762	17,733
Average	**N/A**	**100.00**	**772,385**	**13,455**

qubit faults, we show the results in Tables 4.11, 4.12, and 4.13. In Tables 4.11, 4.12, and 4.13, the header column is the gate type. Note that [0,1] indicates the control qubit is 0, and the target qubit is 1. Relatively, [1,0] indicates the control qubit is 1, and the target qubit is 0. Column 1 is the qubits into which we inject a fault. Columns 2 and 3 are the injected fault sizes. The rest of the table follows the same format as Table 4.2. The average gate fidelities between the injected and diagnosed suspect faults of the 6 two-qubit results are 99.98%. As a result, our method is still feasible with the whole Bloch sphere as our search space.

Table 4.8: Results of random single-qubit faults on noise-free backend

Fault injected			Diagnosis results			
Gate type	θ	ϕ	Sus.(θ,ϕ)	DA(%)	Shots	Gates
X	2.41	4.86	(2.40, 4.85)	100.00	1,708,672	158,965
X	1.57	2.11	(1.54, 2.10)	99.98	894,523	30,679
X	0.94	3.28	(0.95, 3.26)	99.99	691,818	27,986
I	0.78	4.54	(0.79, 4.55)	99.99	1,029,283	16,180
I	0.45	5.41	(0.45, 5.40)	100.00	691,380	13,219
I	2.23	3.13	(2.21, 3.14)	99.99	962,586	62,323
SX	2.81	5.88	(2.80, 5.86)	100.00	1,029,579	94,719
SX	1.11	1.49	(1.11, 1.49)	99.99	691,591	193,797
SX	0.92	4.62	(1.03, 4.70)	99.71	691,336	54,460
Average			N/A	99.96	932,308	72,481

Table 4.9: Results of random single-qubit faults on artificial noise model backend

Fault injected			Diagnosis results			
Gate type	θ	ϕ	Sus.(θ,ϕ)	DA(%)	Shots	Gates
X	2.41	4.86	(2.40, 4.85)	100.00	1,709,196	174,430
X	1.57	2.11	(1.57, 2.04)	99.93	1,748,127	20,347
X	0.94	3.28	(0.95, 3.28)	99.99	691,703	24,490
I	0.78	4.54	(0.78, 4.55)	99.99	961,661	14,200
I	0.45	5.41	(0.45, 5.40)	100.00	691,285	9,969
I	2.23	3.13	(2.23, 3.13)	99.99	759,763	61,099
SX	2.81	5.88	(2.84, 5.90)	99.88	929,238	69,453
SX	1.11	1.49	(1.10, 1.50)	100.00	691,563	162,596
SX	0.92	4.62	(1.05, 4.71)	99.61	691,308	59,961
Average			N/A	99.93	985,983	66,283

4.3 Depolarizing Fault Diagnosis

We demonstrate in this section that our flow can diagnose different sizes of depolarizing faults. We inject 7 depolarizing faults whose sizes are no larger than 0.032. This is a reasonable number because typical depolarizing noise in real QPs is smaller than 0.01. The error caused by depolarizing faults would not affect the rough and fine diagnosis results. Since we need the coherent fault information before diagnosing depolarizing faults, we injected 6 coherent faults from Section 4.2. The results are shown in Fig. 4.2. The

38

Table 4.10: Results of random single-qubit faults on IBM noise model backend

Fault injected			Diagnosis results			
Gate type	θ	ϕ	Sus.(θ,ϕ)	DA(%)	Shots	Gates
X	2.41	4.86	(2.40, 4.85)	100.00	1,167,504	140,421
X	1.57	2.11	(1.57, 2.10)	99.99	759,589	39,304
X	0.94	3.28	(0.95, 3.29)	99.99	691,587	20,340
I	0.78	4.54	(0.78, 4.54)	99.98	691,376	14,056
I	0.45	5.41	(0.45, 5.40)	100.00	691,298	10,405
I	2.23	3.13	(2.23, 3.15)	99.99	692,138	56,725
SX	2.81	5.88	(2.80, 5.85)	100.00	1,029,640	115,242
SX	1.11	1.49	(1.10, 1.50)	100.00	691,618	212,591
SX	0.92	4.62	(1.03, 4.70)	99.71	691,380	85,454
Average			N/A	99.96	789,570	77,171

Table 4.11: Diagnosis results of two-qubit gate faults on noise-free backend

Fault injected				Noise-free			
Gate type	Qubit	θ	ϕ	Sus.(θ,ϕ)	DA(%)	Shots	Gates
CX[0,1]	0	1.84	2.45	(1.85, 2.45)	100.00	7,980,067	320,068
CX[0,1]	1	0.37	1.19	(0.35, 1.21)	99.95	2,761,819	95,025
CX[0,1]	0&1	2.02	5.59	(2.01, 5.58)	99.97	2,289,146	97,101
CX[1,0]	0	2.70	0.64	(2.70, 0.52)	99.95	2,491,823	109,785
CX[1,0]	1	1.29	0.70	(1.30, 0.71)	99.98	2,491,548	91,328
CX[1,0]	0&1	0.11	2.02	(0.10, 2.00)	99.98	2,322,513	79,362
Average				N/A	99.97	3,389,486	132,112

x-axis is the depolarizing fault size we injected. The y-axis is the diagnosis error between the depolarizing fault size we reported and the depolarizing fault size we injected. Every line represents the same coherent fault of different depolarizing fault sizes. The parameters in the legend of each line are the gate type of injected fault and the coherent fault size. The two fault sizes are (0.4, 0) and (0, 1.57), which represent cases of the highest coherent diagnosis accuracy (100%) and the lowest coherent diagnosis accuracy (99.78%), respectively. The coherent fault diagnosis results are shown in subsection 4.2. We can see that the diagnosis error is less than 0.1% when the fault size is smaller than 0.04.

From the above experiments, we found an important fact, that the diagnosis error of coherent fault has little impact on the diagnosis accuracy of depolarizing fault. In Fig.4.2,

Table 4.12: Diagnosis results of two-qubit gate faults on artificial noise model backend

Fault injected				artificial noise model			
Gate type	Qubit	θ	ϕ	Sus.(θ,ϕ)	DA(%)	Shots	Gates
CX[0,1]	0	1.84	2.45	(1.85, 2.45)	100.00	7,036,268	313,423
CX[0,1]	1	0.37	1.19	(0.35, 1.20)	99.95	2,761,819	95,025
CX[0,1]	0&1	2.02	5.59	(2.01, 5.61)	99.97	2,185,434	91,661
CX[1,0]	0	2.70	0.64	(2.71, 0.53)	99.95	2,352,344	109,785
CX[1,0]	1	1.29	0.70	(1.30, 0.70)	100.00	2,452,054	89,971
CX[1,0]	0&1	0.11	2.02	(0.09, 2.01)	99.98	2,345,795	79,362
Average				N/A	99.98	3,188,952	129,871

Table 4.13: Diagnosis results of two-qubit gate faults on IBM noise model backend

Fault injected				IBM noise model			
Gate type	Qubit	θ	ϕ	Sus.(θ,ϕ)	DA(%)	Shots	Gates
CX[0,1]	0	1.84	2.45	(1.85, 2.45)	99.74	4,394,120	429,348
CX[0,1]	1	0.37	1.19	(0.35, 1.37)	99.95	2,491,715	111,643
CX[0,1]	0&1	2.02	5.59	(2.01, 5.60)	99.98	2,289,024	91,661
CX[1,0]	0	2.70	0.64	(2.70, 0.57)	99.97	2,424,069	98,249
CX[1,0]	1	1.29	0.70	(1.30, 0.70)	100.00	2,423,957	89,971
CX[1,0]	0&1	0.11	2.02	(0.10, 2.02)	99.97	2,390,090	79,817
Average				N/A	99.98	3,098,541	130,174

for the two lines (brown and green) with the lowest coherent diagnosis accuracy, we can also diagnose the depolarizing fault size accurately. This demonstrates that the diagnosis error of coherent faults does not affect the diagnosis accuracy of depolarizing faults. This is because the effect of the depolarizing fault masks the effect caused by the diagnosis error of coherent faults [23]. We demonstrate the results in Fig.4.3. In this figure, the x-axis is the diagnosis error of coherent faults. The y-axis is the diagnosis error of depolarizing faults. We fix the depolarizing fault size to 0.4%. As predicted, the diagnosis error of coherent faults increases as the diagnosis error of depolarizing faults increases. We discover that even though the diagnosis error of coherent fault reaches 15%, the diagnosis error of depolarizing fault is only 0.15%.

Likewise, a depolarizing fault does not affect the diagnosis accuracy of rough and
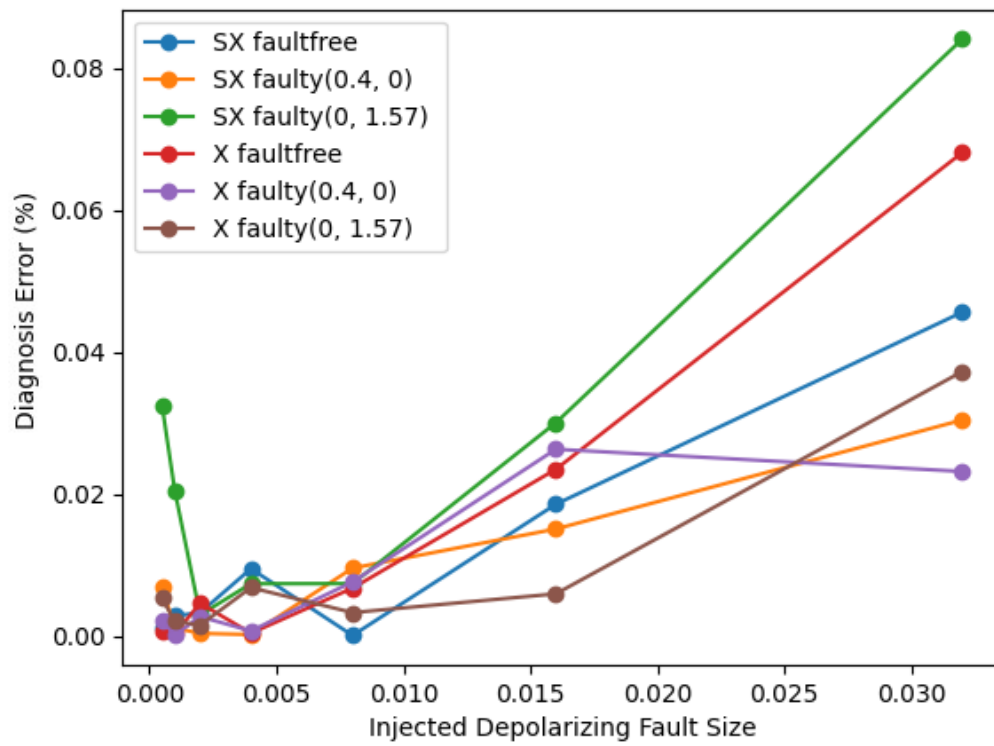
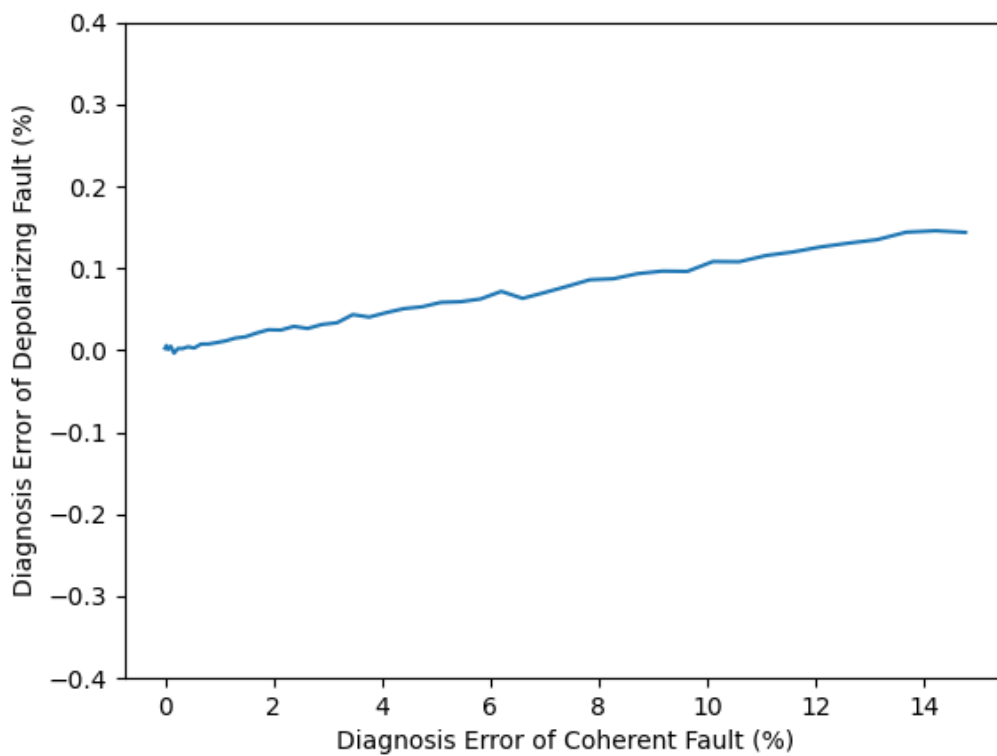Figure 4.2: Results of depolarizing diagnosis

Figure 4.3: Effect of diagnosis error of coherent faults

41

fine diagnosis. In Tables 4.9,4.3, and 4.6, we can see that the diagnosis accuracies of the noisy backends are about the same compared with the diagnosis accuracy of the noise-free backend. This is because the depolarizing noise in the noisy backends only causes the quantum state to shrink towards the Bloch sphere, but not rotate the quantum state. Thus, the depolarizing noise does not affect the diagnosis accuracy while we perform rough and fine diagnosis.
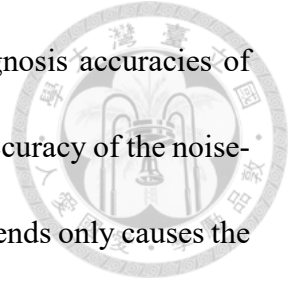
4.4 Compare qFD with Quantum Process Tomography

Next, to demonstrate our method is better than past research, we compare our flow with QPT [16]. Two features for comparison are diagnosis accuracy and the number of shots required. We don't compare the number of gates because it only affects the noise size. In previous tables, we demonstrated that our method can still work in noisy environments. We first perform QPT with all the single-qubit faults from Tables 4.2 and 4.5. Note that we perform QPT using the noise-free backend. QPT calculates the average gate fidelity between the reconstructed and injected faulty gates, which is defined as diagnosis accuracy in our paper. Then, we perform our method on two search spaces: one is the whole Bloch sphere, and the other is a quarter Bloch sphere. We selected a quarter of the Bloch sphere because the sizes of coherent faults are not that large in real QPs. A smaller search space is more realistic than a larger search space.

We plot the results in Fig.4.4. The x-axis is the number of shots. The y-axis is the diagnosis accuracy. Note that in this experiment, we only compare our results of coherent

fault with those of QPT. From the plot, we can see that our method with the whole Bloch sphere as search space can achieve 99.97% diagnosis accuracy, indicating that the diagnosis error is 16 times smaller than QPT with the same number of shots. Moreover, our method achieves the same diagnosis accuracy with just around 230,000 shots when the search space is limited to a quarter of the Bloch sphere, which indicates that the diagnosis error is 32 times smaller than QPT.

Next, we demonstrate how many shots QPT needs to achieve the results of qFD. The results are plotted in Fig.4.5. The format of the figure is identical to that of Fig.4.4. We discover that to achieve the same diagnosis accuracy of qFD, QPT requires more than 10,000,000 shots, which is 14 times more than qFD. As a result, our qFD flow performs better than QPT.

We then compare the results of two-qubit faults with QPT's result. We first perform QPT with all the faults from Table 4.11. Then we plot the results of our flow and the results from QPT in Fig.4.6. Similar to the results in Fig. 4.4, our method still our-performs QPT.

4.5 Experiments on Real IBM Quantum Processors

Finally, we demonstrate our method is feasible on real IBM QPs. Based on IBMQ's data, we chose the least noisy qubits to run our flow. We injected both single and two-qubit coherent faults into the QP. The results are shown in Tables 4.14, Table 4.15, and
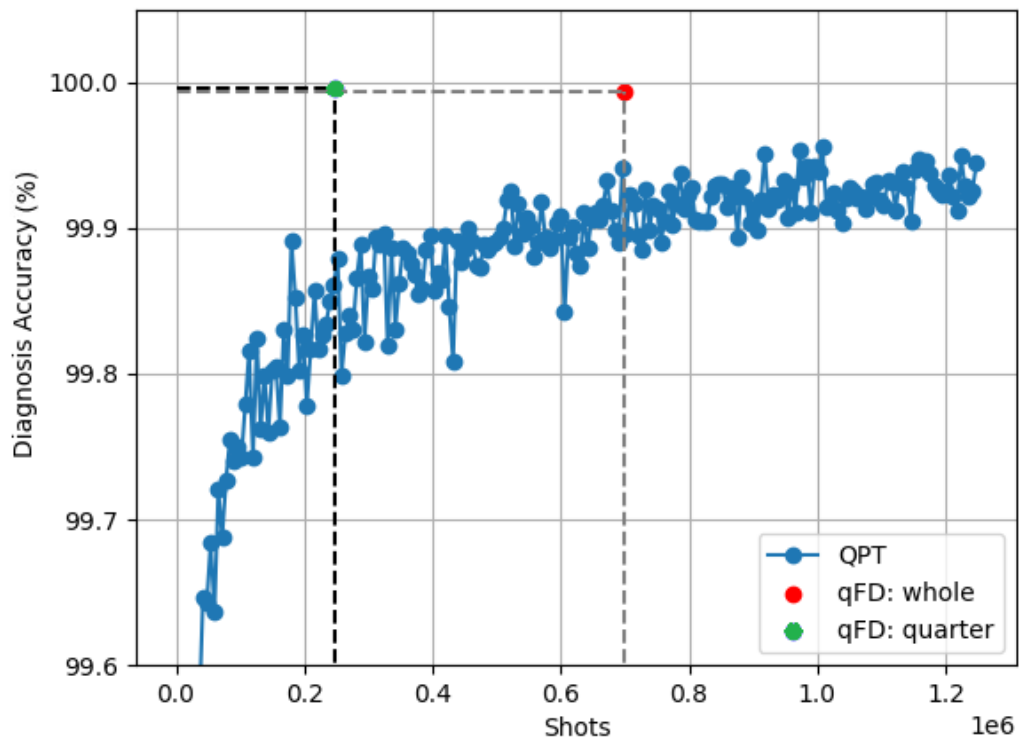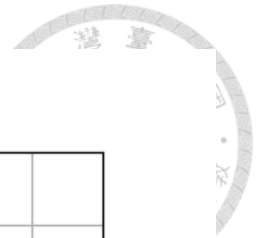
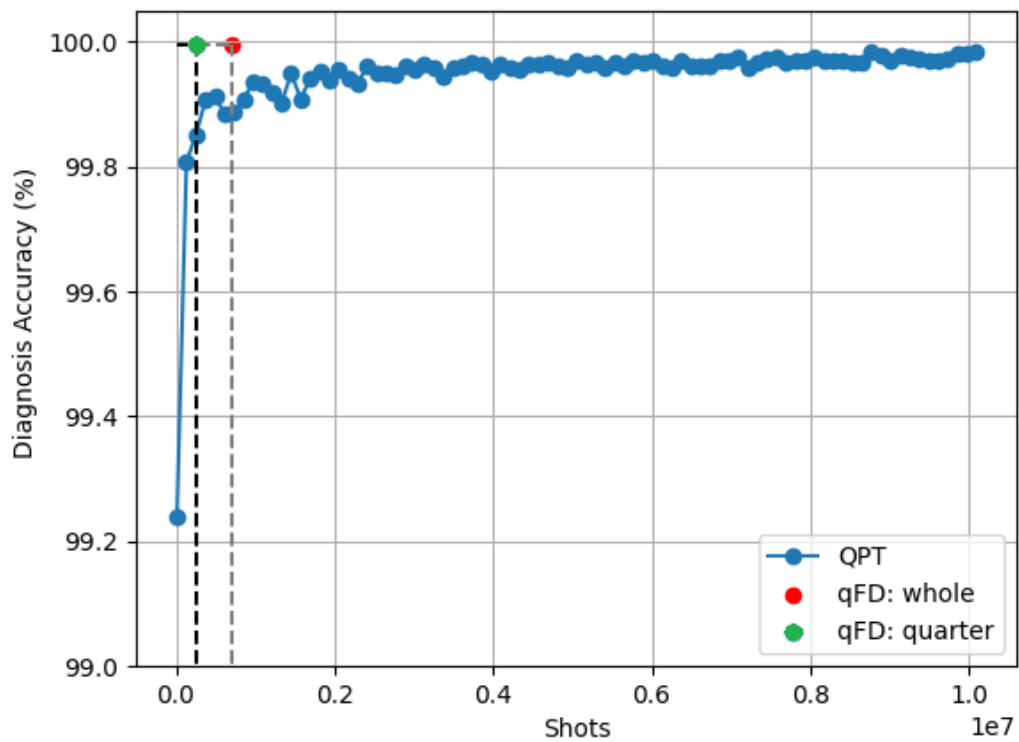Figure 4.4: Compare the diagnosis accuracy of qFD and QPT

Figure 4.5: Compare shots required of qFD and QPT
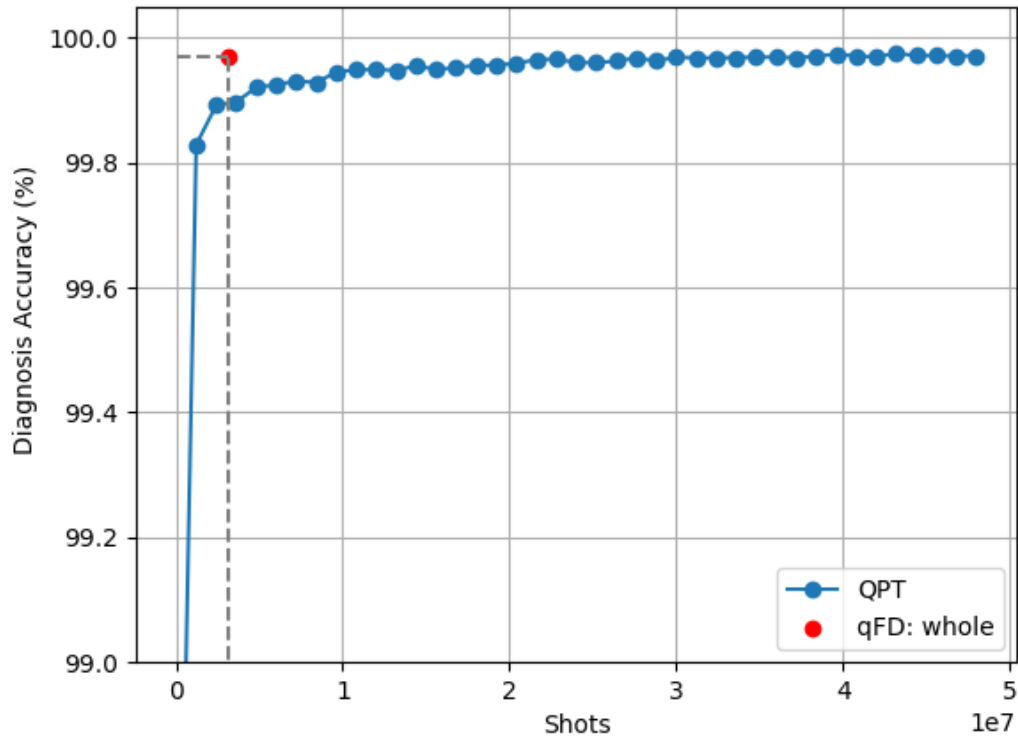
Figure 4.6: Compare qFD and QPT injected with two-qubit faults

Table 4.16. For single-qubit results, the diagnosis accuracy is 99.99%, which is within 0.01% inaccuracy from the injected faults. For two-qubit results, the diagnosis accuracy of 99.83% is only 0.17% from the injected faults. We can discover that there is only a small inaccuracy in diagnosis accuracy, while the number of shots is small. This experiment demonstrates that if a coherent fault exists on real QPs, our method can diagnose its precise behavior within a feasible number of shots.

Table 4.14: Diagnosis results of phase drifting faults on real IBM QP

Injected(ϕ)	Sus.(θ,ϕ)	DA(%)	Shots	Gates
0.00	(0.00, 0.00)	100.00	448,563	4,407
0.05	(0.03, 0.04)	99.95	503,859	4,955
0.10	(0.00, 0.10)	100.00	522,291	5,130
0.20	(0.00, 0.20)	100.00	1,075,251	10,530
0.40	(0.00, 0.40)	100.00	691,217	12,210
0.79	(0.00, 0.80)	100.00	798,860	7,883
1.57	(0.00, 1.56)	99.99	699,581	7,238
3.14	(0.00, 3.15)	100.00	691,544	6,953
4.71	(0.00, 4.73)	99.99	1,874,379	18,570
Average	**N/A**	**99.99**	**792,958**	**7,866**

Table 4.15: Diagnosis results of over-rotation faults on real IBM QP

Injected(θ)	Sus.(θ,ϕ)	DA(%)	Shots	Gates
0.00	(0.00, 0.00)	100.00	448,563	4,407
0.05	(0.05, 0.00)	100.00	448,600	4,440
0.10	(0.10, 0.00)	100.00	503,859	4,950
0.20	(0.20, 0.00)	100.00	691,251	6,780
0.40	(0.40, 0.00)	100.00	860,211	8,430
0.79	(0.80, 0.00)	100.00	691,251	6,780
1.57	(1.58, 0.00)	99.99	691,557	6,960
2.35	(2.35, 0.00)	100.00	699,920	11,468
3.09	(3.14, 0.00)	99.98	406,397	4,485
Average	**N/A**	**99.99**	**604,623**	**6,522**

Table 4.16: Diagnosis results of two-qubit gate faults on real IBM QP

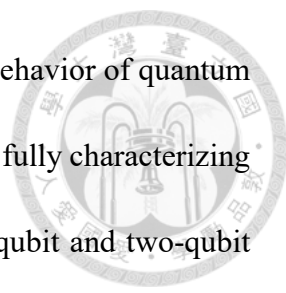| Fault injected				Real IBM QP			
Gate type	Qubit	θ	ϕ	Sus.(θ,ϕ)	DA(%)	Shots	Gates
CX[0,1]	0	1.84	2.45	(1.91, 2.44)	99.74	4,394,120	429,348
CX[0,1]	1	0.37	1.19	(0.35, 1.25)	99.70	2,374,127	316,540
CX[0,1]	0&1	2.02	5.59	(1.98, 5.50)	99.83	2,289,137	284,318
CX[1,0]	0	2.70	0.64	(2.65, 0.59)	99.87	2,352,344	109,785
CX[1,0]	1	1.29	0.70	(1.28, 0.68)	99.93	2,452,054	89,971
CX[1,0]	0&1	0.11	2.02	(0.14, 2.01)	99.93	2,345,795	79,362
Average				**N/A**	**99.83**	**2,759,479**	**336,849**

Chapter 5   Discussion

In this chapter, we discuss some features which can be improved. The first part is the scope of our two-qubit fault model. Unlike the one-qubit fault model, the two-qubit fault model is not universal. However, if we inject our $rf(\theta, \phi)$ before and after into the two qubits to simulate a universal two-qubit fault, the search complexity might be too high. A smaller search space such as a quarter of a Bloch sphere might be the solution.

The second part is about measuring a QC on three axes. The purpose of measuring a QC on three axes is that we want more information on the quantum state. Then we can remove the suspect faults whose Hellinger Distance between the suspect fault and the PUD is larger than the threshold TH_R on either axis. However, at the last step when generating a rough diagnosis QC, we maximize the distance between the fault-free and faulty state on the z-axis. As a result, measuring on the x and y axis would be redundant, and the number of shots required would also decrease by one-third.

The third part is a comparison with *gate set tomography*(GST) [24]. In our work, we only compare qFD with quantum process tomography. However, GST is more applicable in the real world. GST and quantum process tomography are both techniques used in

quantum information processing to characterize and understand the behavior of quantum systems, particularly quantum gates. Gate set tomography focuses on fully characterizing a set of quantum gates used in quantum computing, such as single-qubit and two-qubit gates. It aims to determine the properties and behavior of these gates, including their fidelity and error rates, providing insight into how accurately they perform desired operations. Quantum process tomography aims to characterize the overall quantum evolution or transformation of a quantum system, considering the combined effect of all the gates and operations applied to it. It involves reconstructing the quantum process performed on a quantum state, capturing the entire sequence of operations from input to output. The main reason our qFD cannot compare with GST is that we do not perform calibration of state preparation and measurement. For future work, we can apply techniques to diagnose faults that exist in the state preparation and measurement stage.
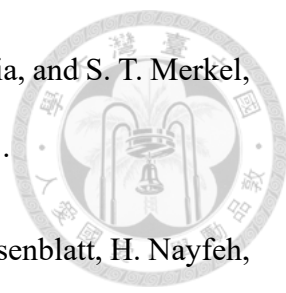
Chapter 6   Conclusion

In this work, we proposed a diagnosis flow for QPs in the NISQ era. Our flow contains rough diagnosis, fine diagnosis, and depolarizing diagnosis. Rough diagnosis grid searches the Bloch sphere and locates an approximate range of the coherent fault's size. Fine diagnosis distinguishes the diagnosed coherent fault from others. At last, depolarizing diagnosis identifies the size of the depolarizing fault. To demonstrate our diagnosis flow, we injected 27 single-qubit, 6 two-qubit coherent faults, and 7 depolarizing faults. For coherent faults, the diagnosis accuracy of single-qubit and two-qubit are over 99.95% and 99.98%, respectively. For depolarizing faults, the diagnosis error between our diagnosed fault size and the actual fault size is smaller than 0.1%. Compared with QPT, our qFD flow can achieve 16 times lower diagnosis error with the same number of shots. At last, we show our flow can be conducted on IBM quantum processors with nearly 0% and 0.17% diagnosis errors for single and two-qubit faults, respectively. Our results show that the diagnosis error of coherent faults has a small impact on the diagnosis accuracy of depolarizing faults.

References

[1] A. Morvan *et al.*, "Phase transition in random circuit sampling," 2023.

[2] V. Jagadish and F. Petruccione, "An invitation to quantum channels," *Quanta*, vol. 7, p. 54, July 2018.

[3] D. Bera, "Detection and diagnosis of single faults in quantum circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 3, pp. 587–600, 2017.

[4] Y.-M. Li, C.-Y. Hsieh, Y.-W. Li, and J. C.-M. Li, "Diagnosis of quantum circuits in the nisq era," in *2023 IEEE 41st VLSI Test Symposium (VTS)*, pp. 1–7, 2023.

[5] E. Knill *et al.*, "Randomized benchmarking of quantum gates," *Physical Review A*, vol. 77, no. 1, p. 012307, 2008.

[6] C.-H. Wu *et al.*, "qatg: Automatic test generation for quantum circuits," in *2020 IEEE International Test Conference (ITC)*, pp. 1–10, IEEE, 2020.

[7] IBM Quantum, "27q exploratory falcon r5 systems ibm_hanoi." Retrieved from https://quantum-computing.ibm.com.

[8] IonQ Forte. `https://ionq.com/quantum-systems/forte`, 2022.

[9] D. C. McKay, I. Hincks, E. J. Pritchett, M. Carroll, L. C. G. Govia, and S. T. Merkel, "Benchmarking quantum processor performance at scale," 2023.

[10] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, "Evidence for the utility of quantum computing before fault tolerance," *Nature*, vol. 618, pp. 500–505, Jun 2023.

[11] J. Preskill, "Quantum computing in the nisq era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[12] D. Gil, "The 2023 ibm research annual letter," *IBM Research Blog*, 2024. Available at `https://research.ibm.com/blog/research-annual-letter-2023`.

[13] A. Paler, I. Polian, and J. P. Hayes, "Detection and diagnosis of faulty quantum circuits," in *17th Asia and South Pacific Design Automation Conference*, pp. 181–186, IEEE, 2012.

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

[15] A. De Vos and S. De Baerdemacker, "Block-zxz synthesis of an arbitrary quantum circuit," *Phys. Rev. A*, vol. 94, p. 052317, Nov 2016.

[16] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, "Quantum-process tomography: Resource analysis of different strategies," *Phys. Rev. A*, vol. 77, p. 032322, Mar 2008.

[17] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear Algebra and its Applications*, vol. 10, no. 3, pp. 285–290, 1975.

[18] E. Hellinger, "Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen.," *Journal für die reine und angewandte Mathematik*, vol. 1909, no. 136, pp. 210–271, 1909.

[19] IBM Quantum. `https://quantum-computing.ibm.com/`, 2021.

[20] M. S. Anis *et al.*, "Qiskit: An open-source framework for quantum computing," 2021.

[21] C. Miquel, J. P. Paz, and W. H. Zurek, "Quantum computation with phase drift errors," *Phys. Rev. Lett.*, vol. 78, pp. 3971–3974, May 1997.

[22] J.-M. Reiner, S. Zanker, I. Schwenk, J. Leppäkangas, F. Wilhelm-Mauch, G. Schön, and M. Marthaler, "Effects of gate errors in digital quantum simulations of fermionic systems," *Quantum Science and Technology*, vol. 3, p. 045008, aug 2018.

[23] Z. Cai, X. Xu, and S. C. Benjamin, "Mitigating coherent noise using pauli conjugation," *npj Quantum Information*, vol. 6, p. 17, Feb 2020.

[24] E. Nielsen, J. K. Gamble, K. Rudinger, T. Scholten, K. Young, and R. Blume-Kohout, "Gate set tomography," *Quantum*, vol. 5, p. 557, Oct. 2021.