Implementation of variational quantum algorithms on superconducting qudits

Shuxiang Cao

Exeter College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2023

Acknowledgements

I would like to first extend my deepest gratitude to my family for their support throughout my DPhil journey. Your love and encouragement have been driving me toward success and I am forever grateful.

My thanks go to Dr. Peter Leek for his supervision and mentorship during my DPhil studies. Thank you for allowing me to start this journey within the group, and for not blaming me for making silly mistakes and breaking equipment. I am also grateful for your patience and kindness which allows me to explore most of my own interests during my DPhil. I enjoyed the time I spent here in Oxford, and this journey could not have been completed without you!

I want to extend a special thanks to Professor John Morton and Professor Simone Severini for introducing me to my supervisor Peter during the UCL summer school. Your kindness and support made this journey possible.

I thank Giulio Campanaro for his collaboration and significant contributions to the fabrication process. I spent the most time with you and you are so supportive and have a great sense of humor. I not only enjoyed but had so much fun working with you. Also, your dedication and hard work ensured our project's success!

I am also grateful to Simone Fasciati for being a mindful person and providing help with fabrication. You are always the most serious and

responsible for the work. Your attention to detail had a significant impact on me, and I learned a lot from you.

I also want to thank James Wills for his valuable insights and support throughout the process. I was intimidated by your poker face at first, but I soon realized how kind and helpful you are. It is really a pleasure to work with you!

Special appreciation goes to Mustafa Bakr for your excellent work on the multiplexer and for supporting me during my DPhil. Your expertise and guidance were crucial in understanding the project's complexities. And you are very supportive and helpful.

I want to thank Brian Vlastakis for your help and guidance during my early DPhil. You acted as my unofficial second supervisor and helped me set off the right path. I am grateful for your patience in teaching me all the details and helping me revise my first paper. Your support is instrumental for me to obtain a successful DPhil.

I want to thank Andy Patterson, Joseph Rahamin, Takahiro Tsunoda, and Peter Spring for their help during the early stages of my academic journey. Your expertise and support were instrumental, and I am grateful for your contributions.

Finally, I thank everyone else from the Leeklab family and beyond who helped me throughout this journey. Your support and encouragement were greatly appreciated and I will always be grateful for your help in my success. Thank you all for being a part of this journey and helping me reach this point in my academic career.

Abstract

Quantum computing is considered an emerging technology with promising applications in chemistry, materials, medicine, and cryptography. Superconducting circuits are a leading candidate hardware platform for the realisation of quantum computing, and superconducting devices have now been demonstrated at a scale of hundreds of qubits. Further scale-up faces challenges in wiring, frequency crowding, and the high cost of control electronics. Complementary to increasing the number of qubits, using qutrits (3-level systems) or qudits (d-level systems, d>3) as the basic building block for quantum processors can also increase their computational capability. A commonly used superconducting qubit design, the transmon, has more than two levels. It is a good candidate for a qutrit or qudit processor. Variational quantum algorithms are a type of quantum algorithm that can be implemented on near-term devices. They have been proposed to have a higher tolerance to noise in near-term devices, making them promising for near-term applications of quantum computing. The difference between qubits and qudits makes it non-trivial to translate a variational algorithm designed for qubits onto a qudit quantum processor. The algorithm needs to be either rewritten into a qudit version or an emulator needs to be developed to emulate a qubit processor with a qudit processor.

This thesis describes research on the implementation of variational quantum algorithms, with a particular focus on utilising more than two com-

putational levels of transmons. The work comprises building a two-qubit transmon device and a multi-level transmon device that is used as a qutrit or a qudit ($d = 4$). We fully benchmarked the two-qubit and the single qudit devices with randomised benchmarking and gate-set tomography, and found good agreement between the two approaches. The qutrit Hadamard gate is reported to have an infidelity of $3.22 \pm 0.11 \times 10^{-3}$, which is comparable to state-of-the-art results. We use the qudit to implement a two-qubit emulator and report that the two-qubit Clifford gate randomised benchmarking result on the emulator (infidelity $9.5 \pm 0.7 \times 10^{-2}$) is worse than the physical two-qubit (infidelity $4.0 \pm 0.3 \times 10^{-2}$) result. We also implemented active reset for the qudit transmon to demonstrate preparing high-fidelity initial states with active feedback. We found the initial state fidelity improved from 0.900 ± 0.011 to 0.9932 ± 0.0013 from gate set tomography.

We finally utilised the single qudit device to implement quantum algorithms. First, a single qutrit classifier for the iris dataset was implemented. We report a successful demonstration of the iris classifier, which yields the training accuracy of the qutrit classifier as 0.96 ± 0.03 and the testing accuracy as 0.94 ± 0.04 among multiple trials. Second, we implemented a two-qubit emulator with a 4-level qudit and used the emulator to demonstrate a variational quantum eigensolver for hydrogen molecules. The solved energy versus the hydrogen bond distance is within 1.5×10^{-2} Hartree, below the chemical accuracy threshold.

From the characterisation, benchmarking results, and successful demonstration of two quantum algorithms, we conclude that higher levels of a transmon can be used to increase the size of the Hilbert space used for quantum computation with minimal extra cost.

Contents

Chapter 1

Introduction

1.1 Computability and the Church-Turing Thesis

The efficient transmission and processing of information with machines have underpinned the fast pace of technological advancement over the last few decades. Computers, as the enabling technology for information processing, have led the revolution. The discipline of Computer Science has emerged from this revolution and involves the study of the capabilities of computers. Mathematicians define problems that can be solved on a computer as *computable functions* [260]. Three equivalent computational models describe computable functions: Turing machines, λ calculus, and recursive functions [244, 260, 245]. Informally speaking, these models can be considered different ways to write computer programs with finite and precise instructions. The programs described with these models can be executed simply with paper and pen under predefined rules, without any human intelligence involved. If the program can always produce a result in a finite number of steps, it is then called *effectively calculable* [260, 245, 12]. A conjecture is then stated that if the function can be written in one of the computing models with a finite number of instructions, it should be executable on a computer and deliver the result in finite steps. This is the famous Church-Turing conjecture proposed in the 1930s, which formally states that: any effectively calculable function is computable [260].

The conjecture captures the very nature of computability and has not been for-

mally proven. However, the conjecture is believed to be true in the mathematical and philosophical communities. Later a stronger version of the conjecture that tries to capture the difficulty of executing the programs was proposed: a model of computation can be simulated by another model efficiently. "Efficiently" in this context means that the time required to convert the original problem into a simulatable one, and then translate the simulated solution back into the original problem, is bounded by a polynomial function of the input size. Such limits are sometimes referred to as a *polynomial overhead* in complexity theory. This stronger version of the conjecture is called the extended Church-Turing (ECT) conjecture [65]. The ECT conjecture shows the power of computers: any real-world computation problem can be translated into an equivalent computation problem on a Turing machine and executed. ECT works well for classical problems; however, it is invalidated when a classical computer tries to simulate a quantum system [86, 66].

1.2 Quantum Computing

Consider using a classical computer to simulate a game of flipping quantum coins, which can have superpositions of heads and tails. Quantum coins can be entangled such that the probability of a coin ending up with heads or tails may depend on the state of other coins. For N quantum coins, there are 2^N possible outcomes, therefore requiring 2^N classical computation resources to represent the result. Generally speaking, for a quantum system with N elements, each with d orthogonal states, the resources required to store the information would be d^N. Although manipulations of a quantum system can be described precisely in steps, we have only found methods to calculate the function on a classical computer with exponential overhead [194].

The difficulty of simulating quantum processes with classical computers leads to the idea of quantum computing. Instead of using a classical machine, one could utilize a controllable quantum machine to simulate other quantum systems. The very origi-

nal proposal of quantum computing came from Richard Feynman to simulate physical systems for many-body physics with quantum machines [86]. In 1981, David Deutsch extended the ECT conjecture into a quantum version [66]. The whole statement becomes that a universal quantum computer could efficiently simulate any phenomena in the physical world.

Since then, quantum computing has been carefully studied, even beyond simulating other quantum systems. The quantum Fourier transform (QFT) [194] was discovered to be exponentially faster than the classical Fourier transform. Soon, a few quantum algorithms, such as quantum phase estimation [109], Shor's algorithm [239], and the HHL algorithm [109], were developed based on QFT that harness the exponential speed-up. Among all, Shor's algorithm is the most famous, leading to calculating discrete logarithms, which can break the foundation of most of the currently widely used asymmetric cryptography systems [228, 128].

1.3 Development of Near-Term Quantum Computers

To execute the quantum algorithms that inherit exponential speed-up from the Quantum Fourier transform, the quantum computer needs to consist of enough qubits, execute quantum gates, and be able to suppress the error rate to arbitrarily low levels. The error rate can be reduced by applying quantum error correction schemes, and these quantum devices are called *fault-tolerant* [39, 92]. As yet, no quantum device in the world satisfies these requirements, but efforts are being made towards fault-tolerant quantum computing [149, 196, 117, 90, 38, 3]. Some recent milestones include encoding quantum information into binomial bosonic logical qubits for error correction [117], implementation of square and hexagonal GKP codes using superconducting microwave cavities [38], the implementation of the distance-3 surface code [149], and a demonstration of the scaling of the surface code [3].

The more mature quantum systems we have now are known as noisy intermediate-scale quantum (NISQ) [217] devices. In 2019, Google claimed to achieve "quantum supremacy" on random circuits with 53 qubits [2], a milestone that a quantum computer can perform calculations beyond the capability of any classical computer [10]. In 2022, IBM announced their 128 qubit quantum computer [1]. These quantum processors are noisy as they are not fault-tolerant and typically have tens to hundreds of qubits. Researchers have found some use cases for these NISQ devices: variational quantum algorithms (VQA) have emerged as a promising solution that utilises the advantage of NISQ devices but also has a larger tolerance to the noise of the NISQ hardware. The first VQA was proposed and demonstrated in 2013 for solving molecule energy [211]. Later, more applications emerged for machine learning [103, 18], combinatorial optimization [134, 84, 8], and finance-related optimizations [32, 81, 114, 113]. A more comprehensive discussion of VQAs can be found in Chapter 2.

1.4 Using Qudits for Quantum Information Processing

A quantity to characterize the computational capability of a quantum computer is the maximum size of its Hilbert space [194]. A machine consisting of N elementary building blocks, each of which has d quantum states (i.e., a Hilbert space of dimension d), has a Hilbert space of dimension d^N. Since the computational capability scales as d^N, we can maximise it by using the largest d and N that we can. Most research to date focuses on using $d = 2$ systems and scaling up the number of qubits N, and soon runs into difficulties such as wiring [251], control electronics [248], and frequency crowding [257, 237]. A complementary strategy is increasing d by replacing the basic building block qubit with a qutrit (3-level system) or qudit (d-level system, d>3). This method does not introduce extra hardware complexity; however, it requires more effort to implement the control scheme for individual building blocks and could suffer

larger noise. Also, operating a qudit-based quantum computer requires extra effort to design quantum algorithms to utilize the entire Hilbert space. The algorithms must be compiled using the qudit gate set, which is different from the qubit gate set.

Quantum Algorithms for Qudit-Based Quantum Computing

The study of qudit-based quantum algorithms has developed to now have a solid theoretical foundation [280, 160, 35, 67]. One major difference is that qudit and qubit gate sets are different. Unitaries representing the single qudit operations are in the $SU(d)$ group. The qudit Clifford group is defined to normalize the generalized Pauli group; see Section 5.1 for more details. To fully utilize the Hilbert space provided by the qudit-based quantum computer, algorithms designed for qubit-based systems need to be modified and adapted. Some of the well-known quantum algorithms have direct generalizations of the qubit counterparts, such as the Deutsch-Jozsa algorithm [83, 185], Bernstein-Vazirani algorithm [150], Grover's algorithm [121], QFT [43, 209], and Shor's algorithm [30]. These generations keep the same principle but change the positional notation from the base-2 numeral system to the base-d numeral system.

On the other hand, it is not always possible to directly generalize the quantum algorithm into a qudit version by changing the positional notation. Some applications, especially near-term applications, are designed to run on qubit-based quantum processors [46, 259]. For example, the variational quantum algorithms for chemistry applications decompose the molecule Hamiltonian into tensor products of qubit Pauli groups, which originates from the fact that the spin of an electron is a two-level system. Generalizing the qudit version for these variational algorithms requires an innovative encoding mechanism and cannot be implemented by simply changing the positional notation. Emulating a qubit system with qudits can be an alternative way of implementing these algorithms. When the Hilbert space of a qudit equals that of an n-qubit system, a qudit can be used to emulate the n-qubit system directly.

For example, a qudit with $d = 4$ can be used to emulate a two-qubit system (i.e., $d = 2^n$), as was first proposed in [137]. In Chapters 5 and 7 of this thesis, we show an experimental implementation of this emulation.

It is also worth mentioning that hybrid systems using qubits and qudits together for quantum computation have also been investigated and shown to have an advantage over qubit-only processors. The qubit Toffoli gate can be decomposed into shorter circuits with an intermediate qutrit [220], and a similar approach can be applied to reduce the circuit depth of larger arithmetic circuits [13].

Realisation of Qudit-Based Quantum Computing

Experimentally utilising qudits as the basic building blocks for quantum processors has been researched in the past, and achievements have been made on different physical platforms. Here we review the major experimental achievements to date in qudit quantum information processing.

Photonics Photonic systems have been considered a promising candidate for quantum information processing [241], and implementing photonic qudits is an emerging research field. Research from [87] demonstrated photonic qudits encoded in a photon's orbital angular momentum (OAM). The following works of OAM photonic qudits include demonstrating multi-qudit entanglement [169] and on-chip large-scale integration [278]. The work from [159] demonstrated quantum algorithms on a two-qudit OAM photonic quantum processor with $d = 4$. It also benchmarked the quantum processor with the Deutsch-Jozsa algorithm, the Bernstein-Vazirani algorithm, and quantum phase estimation. This research also reports that using the qudit system significantly improves the photon detection rate. Another line of research [151] demonstrated an on-chip photonic qudit system that encodes information into different frequency modes. This experiment successfully demonstrated entanglements between two qudits where $d = 10$ for each qudit, and the setup can scale to qudits

with even larger dimensions. The photonic qudit has also been used for quantum communications [59].

Ion Traps Ions confined in radio-frequency traps can be used to process quantum information, and it has been considered one of the most promising platforms for quantum information processing [34]. Recently a universal qudit trapped-ion quantum processor was demonstrated using the native multilevel structure of chains of $^{40}Ca^{+}$ ions [227]. This research demonstrated entangling gates for qudits with $d = 3$ and $d = 5$ using an 8-ion quantum register.

Nuclear Magnetic Resonance Nuclear magnetic resonance (NMR) enables the manipulation and detection of molecules' nuclear spin states using radio-frequency electromagnetic waves, and it has been one of the successful platforms for early demonstrations of quantum information processing [263, 129]. Using a deuterium (spin-1 particle) nucleus as the NMR qutrit, previous research has implemented applications of determining the parity of a given permutation of three objects [70], implementing quantum Fourier transform [71], and contextuality tests [72]. Another line of research [95] uses a spin $-\frac{3}{2}$ nuclei to implement a qudit ($d = 4$) from the Zeeman splitting using a strong static magnetic field.

Superconducting Circuits Superconducting circuits are one of the leading candidates for implementing universal quantum computers [148, 271]. The most widely used building block in superconducting quantum circuits is the transmon qubit [143], and while most research focuses on using its lowest two energy levels as a qubit, the transmon naturally has higher levels that can be utilised, albeit with increased susceptibility to charge noise decoherence [212, 23, 74]. Superconducting qutrits were used to simulate the topological Maxwell metal bands [252], simulate the dynamics of quantum information in strongly interacting systems [27], and simulate Tensor

Monopoles [253].

The research reviewed above has shown that qudit-based quantum systems are promising for quantum information processing. While qudits have been experimentally implemented on various physical systems for quantum simulation and well-known quantum algorithms, they have not yet been used to demonstrate variational quantum algorithms. To demonstrate variational quantum algorithms on qudit processors, algorithms designed for qubits must be adapted for qudits or use qudits to emulate qubits' behaviour. This thesis research focuses on implementing variational quantum algorithms on a superconducting qudit that emulates a two-qubit quantum processor.

1.5 Synopsis

This thesis describes research on the implementation of variational quantum algorithms, with a particular focus on utilising more than 2 computational levels of transmons. The research includes building a two-qubit transmon device and a multi-level transmon device that emulates two qubits. Then the performance of the two devices is characterized and compared. Two variational quantum algorithms are demonstrated on the multi-level transmon device. The remainder of the thesis is structured into the following chapters:

Chapter 2 introduces variational quantum algorithms, including the variational quantum eigensolver for calculating molecule eigenstates and quantum machine learning. It reviews the mathematical background and the optimization approach for optimizing quantum ansatz parameters. This chapter also includes a proposal for encoding the training label into the input data for a variational quantum classification task.

Chapter 3 introduces the principles of superconducting circuits and how to implement quantum operations on superconducting circuits. This chapter first describes

quantum operations that need to be implemented for a universal quantum computer and then describes the way to quantize superconducting circuits and how to engineer the Hamiltonian to implement quantum gates.

Chapter 4 introduces the particular circuit architecture used in this work and describes the electronic setup. It also reports the results of the basic characterization of the fabricated two-qubit device and multi-level transmon device.

Chapter 5 describes methods for benchmarking the performance of the quantum processor. First, the mathematical background of quantum channels and superoperator formalism is introduced. State and process tomography are then described. Randomized benchmarking and gate-set tomography for characterizing gate errors are also explained. In the end, it discusses the experimental implementation of these methods and makes a comparison between the two-qubit and the multi-level transmon device.

Chapter 6 describes the implementation of active reset on our multi-level quantum system in order to prepare high-fidelity initial qudit states.

Chapter 7 presents the demonstration of quantum algorithms on the device. Two variational algorithms are presented in this chapter. One is for implementing a quantum classifier with a qutrit, and the other is for implementing a variational quantum eigensolver for hydrogen molecules on a 4-level qudit emulating two qubits.

Chapter 8 concludes the thesis and presents an outlook for future work.

Chapter 2

Quantum Algorithms on Near-term devices

In this chapter, I start by reviewing variational quantum algorithms. Section 2.1 covers the mathematics behind quantum variational algorithms, while Section 2.2 delves into their applications in quantum chemistry and machine learning. Section 2.3 discusses optimising these algorithms, which includes some of the review materials I contributed to in Ref. [259]. In Section 2.4, I introduce a new concept involving quantum-encoded labels and data for machine learning, though the benefits appear limited. The idea and analysis were completed by me and were published in Ref. [42] with revisions from coauthors: Leonard Wossnig, Brian Vlastakis, Peter Leek, and Edward Grant.

The purpose of developing quantum computers is to implement quantum information processing for practical applications. In this chapter, the concepts of quantum information processing are reviewed, with a focus on near-term quantum algorithms that may be implemented on imperfect noisy devices without error correction. The particular example of quantum machine learning is discussed in detail, including a proposal for data encoding and cost-function embedding directly on a variational quantum circuit. This proposal enables quantum information processing on the cost function value from a variational quantum circuit, without converting back to classical information. A simulated experiment is demonstrated to prove the correctness of the encoding method and its performance under a depolarisation noise channel is investigated.

2.1 Quantum information processing

Quantum information processing was proposed to utilise the quantum nature of superposition and entanglement to represent information and to generalise the concept of information processing to the quantum world [66, 86]. There are many ways to implement quantum information processing, such as gate-based quantum computing [240, 14, 284], measurement-based quantum computing [63, 222, 275, 223], quantum annealing [9, 132], and quantum random walks [51, 49, 50]. Here, the gate-based quantum computing approach is focused on and explained. Gate-based quantum computation uses quantum registers to store the quantum information and perform calculations by manipulating the quantum state stored in the quantum register. The calculated result is retrieved into classical data by measuring the quantum register at the end of the calculation. In this chapter, we mainly discuss the use of pure states and coherent quantum operations. Mixed states and noisy gates are discussed in chapter 5.

Quantum registers A quantum register is a quantum system comprising multiple elementary quantum systems. Each elementary quantum system can be described by a Hilbert space with d orthogonal states. The state of such a quantum system can be written as

$$|\psi\rangle = \sum_{i=0}^{d} c_i |i\rangle \tag{2.1}$$

where c_i is the complex coefficient of each state and satisfies $\sum_{i=0}^{d} |c_i|^2 = 1$. Suppose the i-th elementary quantum system can represent the state in Hilbert space $\mathcal{H}_i$. The Hilbert space $\mathcal{H}$ of a quantum register is given by $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 ... \otimes \mathcal{H}_n$, where $\otimes$ denotes the tensor product and n is the number of elementary quantum systems in the quantum register.

A quantum state can also be written in the density matrix representation to capture the classical uncertainty. The density matrix ρ of state $|\psi\rangle$ is given by $\rho = |\psi\rangle \langle\psi|$. The tensor product notation between $|\psi\rangle$ and $\langle\psi|$ is omitted.

Quantum gates A quantum gate is a unitary operation that transforms the quantum state in a quantum register. The transformed state $|\psi'\rangle$ after applying gate U can be written as

$$|\psi'\rangle = U |\psi\rangle \tag{2.2}$$

And the corresponding density matrix ρ' is given as

$$\rho' = U\rho U^\dagger = \Phi(\rho) \tag{2.3}$$

where $\Phi(\rho)$ is a superoperator representing the unitary transformation of U to the density matrix ρ. More detailed discussion about superoperator formalisms can be found in chapter 5.

Quantum measurements The information obtained from the quantum system is classical information. This information is characterised by the probabilistic expectation value of a quantum observable or measurement operator. The physical measurement observable needs to be Hermitian. The expectation value $E(O)$ of an observable $\hat{O}$ on state $|\psi\rangle$ is given by

$$E(O) = \langle\psi|\,\hat{O}\,|\psi\rangle = Tr(\hat{O}\rho) \tag{2.4}$$

2.2 Variational quantum algorithms

Decades of research have yielded multiple quantum applications such as Shor's algorithm, Quantum Fourier transform, and Quantum phase estimation. In order to be practically useful, these algorithms require the quantum processor to have a few thousand qubits and be fault-tolerant to be put into practice. Quantum information processing devices available nowadays contain up to a few hundred qubits with considerable error and are called noisy intermediate-scale quantum (NISQ) devices [217]. In 2013 a variational quantum algorithm (VQA), was proposed that utilises a classical optimiser to train a parameterised quantum circuit (PQC) [211], which could overcome the restrictions of NISQ devices while still obtaining considerable quantum advantage.

A VQA defines an objective function and solves the problem by minimising the objective function. Usually, the objective function is constructed conventionally based on the measurement outcome. Denote $\mathbf{O}(\boldsymbol{\theta}) = (\hat{O}_1(\boldsymbol{\theta}^{(1)}), \hat{O}_2^{(2)}(\boldsymbol{\theta}^{(2)}), \ldots, \hat{O}_a(\boldsymbol{\theta}^{(a)}))$ as the observables used to compose the objective function and a as the number of observables. The objective function is given by

$$\mathcal{L}(\boldsymbol{\theta}) = C(\mathbf{O}(\boldsymbol{\theta})) \tag{2.5}$$

where C is a function that maps the observed expectation value to the objective

function and usually has a simple linear form

$$C(\mathbf{X}) = \sum_i c_i X_i \tag{2.6}$$

where c_i is a constant defined by the problem as the coefficient of each measurement expectation value, X_i is the i-th component of $\mathbf{X}$. Note that such a linear form preserves the analytical properties and it is essential for using the analytical methods to directly calculate the gradient or implement an analytical gradient-free optimization strategy [190, 201, 145, 233].

A measurement expectation value is given by

$$\langle \hat{O}_k(\boldsymbol{\theta}^{(k)}) \rangle = \langle \psi_0 | \, U^{(k)\dagger}(\boldsymbol{\theta}^{(k)}) \hat{M}^k U^{(k)}(\boldsymbol{\theta}^{(k)}) \, | \psi_0 \rangle \, , \tag{2.7}$$

where k is the index of the parameterised variable $\boldsymbol{\theta}$, $|\psi_0\rangle$ is the initial state on the quantum computer. $\hat{M}^{(k)}$ is a Hermitian measurement operator, usually chosen to be the tensor product of Pauli operators to match the physical measurement implementation on quantum hardware. $U_k(\boldsymbol{\theta}^{(k)})$ is the variational ansatz defined as

$$U^{(k)}(\boldsymbol{\theta}^{(k)}) = \prod_j U_j^{(k)}(\theta_j^{(k)}) \tag{2.8}$$

and each U_j is a quantum gate, which is generalized as

$$U_j^{(k)}(\boldsymbol{\theta}_j^{(k)}) = \exp(i\theta_j^{(k)} P_j^{(k)}) \tag{2.9}$$

where $P_j^{(k)}$ is a Hermitian matrix, usually a tensor product of Pauli operators. It's worth noting that this type of decomposition is also applicable to two-qubit gates. In fact, any arbitrary two-qubit gate can be expressed as the exponential of a two-qubit Pauli operator with parameterised coefficient, sandwiched by some single-qubit rotations.

It is sometimes convenient to utilise the superoperator formalism and consider the noise in the optimisation process. The expectation value can then be written as

$$\langle \hat{O}_k(\boldsymbol{\theta}^{(k)}) \rangle = \mathrm{Tr}[\hat{M}_k \Phi^{(k)}(\boldsymbol{\theta}^{(k)}) \rho_0], \tag{2.10}$$

where ρ_0 is the initial state density operator and $\Phi^{(k)}(\boldsymbol{\theta}^{(k)})$ denotes the transformation matrix, which is given by

$$\Phi^{(k)}(\boldsymbol{\theta}^{(k)}) = \prod_j \Phi_j^{(k)}(\theta_j^{(k)}). \tag{2.11}$$

For devices that support single-shot readout, for each shot the quantum device would yield a bit string $\mathbf{s}$. For each string $\mathbf{s}$, a measurement value could be calculated with $M_i(\mathbf{s})$, and the expectation value is the average of each $M_k(\mathbf{s})$.

$$\langle \hat{O}_k(\boldsymbol{\theta}^{(k)}) \rangle = \sum_j \text{Prob}(\mathbf{s}(\boldsymbol{\theta}^{(k)}) = b_j) M_k(b_j) \tag{2.12}$$

where $b_j \in B$, B covers all possible single-shot bit strings (all binary numbers from 0 to 2^{n-1}) of the measurement outcome.

Due to the physical implementation from the quantum hardware, not all quantum computing systems support single-shot readout. Some systems can only yield an expectation value by averaging the signal from the readout. For example, some NMR systems use an ensemble of molecules to implement quantum computing and cannot read the state of every single molecule [78, 221]. In practice, the quantum hardware system may directly yield an expectation value, and the measurement approaches vary from different physical systems.

The following section discusses a few types of VQA. The upper label (k) is simplified without ambiguity.

The variational quantum eigensolver (VQE)

The variational quantum eigensolver (VQE) is an approach that utilises the Rayleigh-Ritz variational method [155, 291] to find the upper bound of the lowest eigenvalue of a given Hamiltonian [211, 176]. The VQE seeks the state $|\psi\rangle$ such that the expectation of the Hamiltonian $\hat{H}$ is minimised,

$$E_0 \leq \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \tag{2.13}$$

To parameterise the state $|\psi\rangle$, the formula is usually written as a state transformed by a parameterised gate from the ground state $|0\rangle$, and the optimisation problem becomes

$$E_{VQE} = \min_{\vec{\theta}} \langle\psi_0| U(\vec{\theta}) \hat{H} U^{\dagger}(\vec{\theta}) |\psi\rangle \tag{2.14}$$

The optimised expectation value gives the upper bound of the eigenstate energy. If the $|\psi\rangle$ is well parameterised, it should be accurate enough to be indistinguishable from the actual ground state energy.

While the principle of VQE is straightforward, a few technical details need to be clarified for implementing VQE on quantum hardware. Here we describe the theory of obtaining the configuration of our experiment in later chapters.

Hamiltonian construction First, the Hamiltonian needs to be defined from the physical system whose ground state we want to find. Usually, we start from a physical Hamiltonian derived from an existing theory. In this thesis, we focus on the electronic structure Hamiltonian for small molecules [170, 250]. From non-relativistic quantum mechanics with the Born-Oppenheimer approximation (the motion of the nuclei is neglected as they are much heavier than electrons), the electronic Hamiltonian depends parametrically on the nuclear positions R_k and can be written as

$$\hat{H} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} \tag{2.15}$$

where

$$\hat{T}_e = -\sum_i \frac{\hbar^2}{2M_i} \nabla_i^2 \tag{2.16}$$

$$\hat{V}_{ne} = -\sum_{i,k} \frac{e^2}{4\pi\epsilon_0} \frac{Z_k}{|\mathbf{r_i} - \mathbf{R_k}|} \tag{2.17}$$

$$\hat{V}_{ee} = -\sum_{i,k} \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r_i} - \mathbf{r_j}|} \tag{2.18}$$

$\mathbf{r_i}$ is the position of the i-th electron, M_i is the mass of the electron, Z_k is the atomic number of the k-th nucleus, e is the elementary charge, $\hbar$ is the reduced Planck constant, and ∇_i^2 is the Laplace operator for the i-th electron.

The next step is to define a basis to represent the electronic wave function. One option is to use the STO-nG basis sets for every single atom [211, 133, 152] where it defines n primitive Gaussian orbitals and fits them into a single Slater-type Orbital (STO). One of the common choices is the STO-3G basis, which is constructed from three Gaussians, and the orbital function is given by

$$\chi(r) = c_1\gamma_1(r) + c_2\gamma_2(r) + c_3\gamma_3(r) \tag{2.19}$$

where

$$\gamma_i(r) = (\frac{2\alpha_i}{\pi})^{\frac{3}{4}} e^{-\alpha_1 r^2} \tag{2.20}$$

The STO-nG basis is denoted as a minimal basis because it only includes orbitals necessary to represent the valence shell of an atom. Some more precise models include higher energy orbitals, allowing for additional polarisation and diffuse and higher angular momentum functions, such as the correlation-consistent polarised Valence n-Zeta (cc-pVnZ) basis sets [77]. However, implementing this model requires a much larger Hilbert space, so has not yet been widely used.

Once the single atom basis function has been selected, the many-body basis of the whole molecule is constructed from the Hartree-Fock method as the Slater determinants

$$\psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \cdots & \phi_n(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \cdots & \phi_n(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_n) & \phi_2(\mathbf{x}_n) & \cdots & \phi_n(\mathbf{x}_n) \end{vmatrix}, \tag{2.21}$$

where $\phi_j(\mathbf{x}_j)$ denotes a spin-orbital of the chosen basis, parameterized by $\mathbf{x}_j = (\mathbf{r}_j, \sigma_j)$ where $\mathbf{r}_j$ and σ_j representing the spatial coordinates and the spin configuration for the j-th electron. The N_e-electron many-body basis function can then be written as

$$|\psi\rangle = |\phi_1 \phi_2 \dots \phi_{N_e}\rangle \,, \tag{2.22}$$

where $\phi_j = 1$ means that the j-th basis function $\phi_j(\mathbf{x})$ is occupied, and $\phi_j = 0$ is unoccupied.

Once the basis of the wave function is given, the Hamiltonian can be constructed by simple projection onto the single-particle basis function. Now we use $|\phi_i\rangle$ to represent the i-th basis, and the matrix elements of one-electron terms h_{pq} are given by

$$\begin{aligned} h_{pq} &= \langle \phi_p | \, \hat{T}_e + \hat{V}_{ne} \, | \phi_q \rangle \\ &= \int \mathrm{d}\mathbf{x} \; \phi_p^*(\mathbf{x}) \left(-\frac{\hbar^2}{2m_e} \nabla^2 - \sum_k \frac{e^2}{4\pi\epsilon_0} \frac{Z_k}{|\mathbf{r} - \mathbf{R}_k|} \right) \phi_q(\mathbf{x}) \end{aligned} \tag{2.23}$$

The two-electron interaction terms h_{pqrs} are given by

$$\begin{aligned} h_{pqrs} &= \langle \phi_p \phi_q | \, \hat{V}_{ee} \, | \phi_r \phi_s \rangle \\ &= \frac{e^2}{4\pi\epsilon_0} \int \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 \frac{\phi_p^*(\mathbf{x}_1) \, \phi_q^*(\mathbf{x}_2) \, \phi_r(\mathbf{x}_2) \, \phi_s(\mathbf{x}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \end{aligned} \tag{2.24}$$

The quantized form of the Hamiltonian is then written as

$$\hat{H} = \sum_{i=1}^{m} \sum_{p,q=1}^{n} h_{pq} \, |\phi_p^{(i)}\rangle \, \langle \phi_q^{(i)}| + \frac{1}{2} \sum_{i \neq j}^{m} \sum_{p,q,r,s=1}^{n} h_{pqrs} \, |\phi_p^{(i)} \phi_q^{(j)}\rangle \, \langle \phi_r^{(i)} \phi_s^{(j)}| \,. \tag{2.25}$$

Hamiltonian encoding The goal is to measure the expectation value of the given Hamiltonian. Therefore, the following step is to map the quantized form into a measurement operator and the electron wavefunction into qubit space. First, the n single-particle basis functions can have a one-to-one mapping to the Hilbert space basis of $\log_2(n)$ qubits.

There are a few ways to implement the encoding, such as Jordan-Wigner encoding, parity encoding, Bravyi-Kitaev encoding, and Ternary trees encoding. Here, we describe the principle of the Jordan-Wigner encoding, which we used in our experiment [174].

The Jordan-Wigner transformation rewrites each operator into a tensor product of four types of single-qubit operators: $|0\rangle\langle 0|$, $|0\rangle\langle 1|$, $|1\rangle\langle 0|$, and $|1\rangle\langle 1|$, which can be mapped to Pauli operators as follows:

$$\begin{aligned} |0\rangle\langle 0| &= \frac{1}{2}(I+Z), \quad & |0\rangle\langle 1| &= \frac{1}{2}(X+iY) \\ |1\rangle\langle 0| &= \frac{1}{2}(X-iY), \quad & |1\rangle\langle 1| &= \frac{1}{2}(I-Z). \end{aligned} \tag{2.26}$$

After the transformation, the Hamiltonian is mapped into a weighted sum of a set of directly measurable operators, usually Pauli operators for qubit-based systems. In practice, these operators' string representation is referred to as the *Pauli string*.

Ansatz selection The ansatz refers to the parametrised quantum circuit to implement the state preparation unitary $U(\vec{\theta})$. In principle, the ansatz can be chosen arbitrarily as long as it covers the ground state solution of the problem. Therefore from the hardware point of view, the hardware-efficient ansatz, which gives the easiest way to implement parametrisation on hardware, is very popular. However, optimising the random ansatz can be problematic. For example, a too well-parameterised circuit may suffer the so-called 'barren-plateau' problem [175], where the gradient information would mostly vanish before the optimisation converges. Therefore, some other ansatz based on prior knowledge, for example, the Unitary Coupled Cluster (UCC) ansatz [211], are widely used.

Coupled Cluster (CC) theory is a post-Hartree-Fock method that recovers a portion of electron correlation energy by evolving an initial wave function (usually the Hartree-Fock wave function) under the action of parametrised excitation operators [16, 214, 204, 203, 127, 57]. When the excitation operators are chosen to have only single-electron excitations and double-electron excitations, the method is named the CC Single and Double (CCSD) method. Usually, CC uses excitation operators to represent only transitions from an occupied orbital to an unoccupied orbital. The new orbital is constructed by applying the exponentiation of cluster operator T on

the initial state. When using the Hartree-Fock state as the initial state, the CC gives wave function $|\psi\rangle = e^{\hat{T}} |\psi_{HF}\rangle$. For CCSD, the cluster operator is given by $\hat{T} = \hat{T}_1 + \hat{T}_2$, where $\hat{T}_1$ and $\hat{T}_2$ are the single and double excitation operators:

$$\hat{T}_1 = \sum_{ia} t_i^a \hat{a}_a^\dagger \hat{a}_i \tag{2.27}$$

$$\hat{T}_2 = \sum_{ijab} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i. \tag{2.28}$$

The original CCSD method is not variational and has failed to approximate systems with strong correlation, and the improved solution for CCSD usually scales exponentially in system size [202, 213, 242, 243, 153, 154]). Also, the operator $e^{\hat{T}}$ is not unitary, and therefore the CC ansatz cannot be implemented as a series of quantum gates.

The UCC theory [88, 254, 15] is adapted from the CC theory and developed to be implemented on quantum computers. For any linear operator $\hat{T}$, the expression $(\hat{T} - \hat{T}^\dagger)$ is an anti-Hermitian operator, and its exponential is a unitary operator. Moreover, the elements of the truncated cluster operator for UCC are identical to that of CC, as shown in equation (2.27) and equation (2.28).

$$|\psi\rangle = e^{\hat{T} - \hat{T}^\dagger} |\psi_{HF}\rangle . \tag{2.29}$$

For convenience, the excitation operators in $\hat{T}$ are grouped together with their corresponding conjugate in $\hat{T}^\dagger$, and we denote the grouped operator as τ. The parametrised version of the UCC ansatz can then be written as:

$$U(\vec{t}) = e^{\sum_j t_j (\tau_j - \tau_j^\dagger)}, \tag{2.30}$$

where t_j correspond to the amplitude weights presented in equation (2.27) and equation (2.28), and with j spanning all the excitation operators included.

Using similar mappings as the Hamiltonian, each of the fermionic operators in the ansatz can be written as a weighted sum of Pauli operators, and the ansatz can be written as:

$$U(\vec{t}) = e^{\sum_j \sum_k i t_j \hat{P}_{k,j}}, \tag{2.31}$$

Now to implement the UCC ansatz on a quantum computer it is only necessary to decompose $U(\vec{t})$ into a series of parametrised quantum gates that is physically implementable. For more complicated chemistry problems, the circuit can be approximated with Trotter-Suzuki decomposition (also called *Trotterization*), which separates the sum in the exponent of equation (2.31) into a product of exponentials. For the experiments implemented in this thesis, the ansatz unitary has an explicit form given.

This section gives a walk-through of the principles of VQE. More technical details for VQE implementation can be found in [259].

Quantum machine learning

Another popular research direction is utilising parametric quantum circuits for quantum machine learning. The main reason for their popularity is that the machine learning models can tolerate more noise in the intermediate layers and outputs, making them appealing candidates for the first applications in quantum computing [103, 120, 210, 4].

One exciting aspect of parameterised quantum circuit (PQC) based machine learning is its equivalence to tensor-network-based machine learning algorithms [120, 157]. Promising results indicate that certain types of these algorithms can be executed efficiently on a quantum computer but cannot be efficiently evaluated classically, implying that they have stronger expressive power [75] and a more flexible feature space [111, 235]. Based on these results, there is hope that such quantum models will yield a practical advantage in machine learning.

In this thesis, we demonstrated two machine learning algorithms, one with simulation and one with actual hardware implementation. Both experiments focused on the classification problem. Here we discuss the theory of quantum classifiers.

Quantum classifiers Classification belongs to the category of supervised learning. In supervised learning, we are given a dataset $\mathcal{S} := \{x_i, y_i\}_{i=1}^m$ of m points, where $x_i \in \mathcal{X}$ are *i.i.d.* samples drawn from a fixed but unknown probability distribution, and $y_i \in \mathcal{Y}$ are the corresponding labels, i.e., the ideal output for a given input x_i. In a classification task, the labels are obtained from a finite set of possible outcomes, for example, a binary outcome, i.e., $\mathcal{Y} = \{0, 1\}$. The goal of the classification task is then to find a function, the so-called model f, such that $f(x_i)$ returns a result that matches the label y_i for all $i \in [m]$ with the highest possible accuracy on a subset of $\mathcal{S}$, called the training set $\mathcal{S}_t$. Additionally, we require that the function f also performs well on samples $x_j \notin \mathcal{S}_t$, which we haven't used to train, i.e., find the model f, the so-called unseen data or test set $\mathcal{S}_p = \mathcal{S} \setminus \mathcal{S}_t$. A learning algorithm, or training algorithm $\mathcal{A}$, then takes as input the dataset $\mathcal{S}$, and possibly additional hyperparameters, and returns a model $f = \mathcal{A}(\mathcal{S}, \cdot)$ which minimises some specified measure of the discrepancy between the desired output and the model output. The measure of this discrepancy is called the cost function $\ell : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$, also called the objective of the learning problem, and is an integral part of the solution.

In the PQC setting, the model $f(x_i)$ is given by a large unitary matrix $U \in \mathbb{C}^{2^n \times 2^n}, U^\dagger U = \mathbb{I}$ acting on an n-qubit system, and the input x_i is given by a state $|\psi_i\rangle$. A prediction of the model is then given by the expectation value for the measurement of one qubit after U is applied to some input state $|\psi_i\rangle \in \mathbb{C}^{2^n}$, see figure 2.2a. Here we denote the qubits that store $|\psi_i\rangle$ as *data qubits*. U can be decomposed in several different ways into elementary logical elements, which are called quantum gates. One such decomposition results in a sequence of single-qubit rotation gates and CNOT gates, which can then be parameterised by the rotation angles θ of the single-qubit gates [265, 69]. After obtaining the gate sequence, we arbitrarily choose a subset of the qubits as the *output qubits*. The state of the output qubits is given by the density matrix ρ_{output}. We then consider their measurement output as the prediction

of the ansatz. We choose only one of the qubits for binary classification problems as the output and ignore the information stored in all others. Once we have the output of the measurement of $|\psi_i\rangle$, i.e., the prediction of our model, the cost function can be evaluated by measurement and applying some post-processing to evaluate the discrepancy $\ell(f(|\psi_i\rangle), y_i)$. The training task then consists of adjusting the parameters θ of the model $U(\theta)$ to match the expectation value of a measurement of the output qubit to the corresponding label in the dataset. Throughout this thesis, we let this measurement be on the Z-basis.

Relationship to semidefinite programming

Semidefinite programming (SDP) is an essential subfield of convex optimization and has a close connection with VQAs [20, 33, 206]. Here we denote the space of all $n \times n$ real symmetric matrices as $\mathbb{S}^n$, and the semidefinite programming is defined as follows [262, 64]:

$$\begin{array}{rl} \min_{X \in \mathbb{S}^n} & \langle C, X \rangle_{\mathbb{S}^n} \\ \text{subject to} & \langle A_k, X \rangle_{\mathbb{S}^n} \leq b_k, \quad k = 1, \ldots, m \\ & X \succeq 0 \end{array} \tag{2.32}$$

where the matrix X is the unknown variable that requires optimization, and C, A_k are symmetric matrices. $\langle A, B \rangle$ denotes the inner product between A and B. One of the biggest difficulties in solving SDP is finding solutions that fulfil the positive-semidefinite constraint $X \succeq 0$.

For example, the SDP is solved by first defining the associated *barrier problem* [112], which creates a virtual barrier with a loss function tending to infinity at the boundary where X is no longer positive-semidefinite. Such a numerical approach is computationally difficult to confirm if X is still positive-semidefinite.

Now consider that a density matrix of a quantum system ρ plays the role of X. In order to be physically meaningful, ρ must be positive-semidefinite since we cannot observe negative probability. Under this configuration, there is no need to spend

extra resources to maintain the positive-semidefinite constraints.

An arbitrary SDP problem can be encoded into a quantum computer using a mixed-state density operator. We can write the unitary X parametrised by an ansatz parameter vector θ_j as:

$$X = U(\theta)\Lambda(\lambda)U(\theta)^\dagger \tag{2.33}$$

where $\Lambda = \mathrm{diag}(\lambda_i)$. As long as all $\lambda_i \geq 0$, X is positive semidefinite. If Λ is the density matrix of a physical mixed state, then $\sum \lambda_i = 1$. An arbitrary SDP problem does not have this constraint, therefore to map an arbitrary SDP problem onto a quantum computing problem, we need to release this constraint with additional construction. Now consider the density matrix X as an unnormalised mixed state and define Γ_i as a density matrix with only the i-th diagonal entry equal to 1, and all other entries are 0. Then we can write the expectation value:

$$\mathrm{Tr}(CX) = \mathrm{Tr}(CU(\theta)\Lambda(\lambda)U(\theta)^\dagger) \tag{2.34}$$

$$= \mathrm{Tr}(C\sum_i U(\theta)\lambda_i\Gamma_i U(\theta)^\dagger) \tag{2.35}$$

$$= \sum_i \lambda_i \mathrm{Tr}(CU(\theta)\Gamma_i U(\theta)^\dagger) \tag{2.36}$$

The state $U(\theta)\Gamma_i U(\theta)^\dagger$ can be prepared on a quantum computer by selecting a different initial state (flip qubits) for different i. Then the objective function $\langle C, X\rangle$ is a measurement expectation value with respect to the measurement operator C, and we decompose C and A_k into Pauli matrices as the measurement observables:

$$C = \sum_i c_i P_i \tag{2.37}$$

$$A_k = \sum_j c_{kj} P_j \tag{2.38}$$

This construction shows that all variational quantum algorithms can be viewed as semidefinite programming. Since the initial state is prepared to be a pure state, VQE restricts the parameter X to be rank-1 and trace-preserving or simply implies that X is an idempotent matrix ($X^2 = X$). This constraint forces the density matrix to represent a pure state. Similarly, the quantum approximate optimisation algorithm (QAOA) is the quantum version of semidefinite programming for optimisation combinations.

2.3 Optimisation strategy

The VQA is, in essence, an optimisation problem; it aims at heuristically constructing an approximation of an electronic wavefunction through iterative learning of ansatz parameters. For the algorithm to be viable, it must be that it can learn a good enough approximation of the solution within a tractable number of learning steps. It was already demonstrated that optimisation of the variational quantum ansatz is NP-hard [24], meaning that there exist at least some problems in which finding an exact solution for the VQA problem is intractable. As such, efficient optimisation strategies that provide a well-approximated solution within an acceptable number of iterations are essential for any variational algorithms to be put into practice. Studies from recent years show the expectation value landscape has some analytical properties that are useful to extract information, such as evaluating gradients directly on quantum devices [190, 201, 233]. In addition, the ansatz landscape can be efficiently approximated to accelerate the convergence [145, 205]. Such prior knowledge helps develop efficient optimisation strategies for variational quantum algorithms. This section presents the most relevant optimisers and strategies developed specifically for parametric circuits.

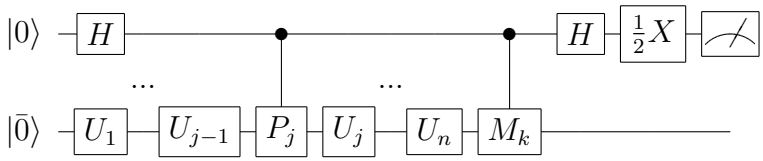


Figure 2.1: Quantum circuit that evaluates $\mathrm{Im}(\langle\phi_0|V_k^{j\dagger}(\boldsymbol{\theta})\hat{M}_k U(\boldsymbol{\theta})|\phi_0\rangle)$.

Analytical gradient calculation

The gradient of measurement observables can be evaluated on quantum computers directly by utilising the analytical property of the ansatz. For ansätze made from a sequence of parametrised quantum gates, the partial derivative is

$$\frac{\partial \langle \hat{O}_k(\boldsymbol{\theta})\rangle}{\partial \theta_j} = 2\,\mathrm{Im}(\langle\phi_0|V_k^{j\dagger}(\boldsymbol{\theta})\hat{M}_k U(\boldsymbol{\theta})|\phi_0\rangle), \tag{2.39}$$

where the operator $V_k^j(\boldsymbol{\theta})$ is defined as inserting P_k^j between the $(j-1)$-th term and j-th term, which comes from the derivative of the j-th term:

$$V_k^j(\boldsymbol{\theta}) = e^{i\theta_{N_k} P_k^{N_k}} \cdots e^{i\theta_j P_k^j} P_k^j e^{i\theta_{j-1} P_k^{j-1}} \cdots e^{i\theta^1 P_k^1}. \tag{2.40}$$

The difficulty of evaluating the gradient directly is that $V_k^{j\dagger}(\boldsymbol{\theta})\hat{M}_k U(\boldsymbol{\theta})$ is usually not Hermitian, therefore there is no known method that can directly convert it to a quantum circuit. Two methods to analytically measure the gradient have been developed, and it is worth noticing that these two methods are measuring the gradient with the exact same mechanism, but with direct and indirect measurements [183]. Here, direct measurement means the observed quantity is encoded into a single quantum observable, while indirect measurement calculates the quantity with multiple quantum observables.

Direct analytical gradient measurement

The imaginary part of $\langle\phi_0|V_k^{j\dagger}(\boldsymbol{\theta})\hat{M}_k U(\boldsymbol{\theta})|\phi_0\rangle$ can be directly evaluated by introducing an ancillary qubit [229]. The quantum circuit that implements this evaluation is shown in Fig 2.1. First, we prepare the ancillary qubit into the $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ state, and the rest of the quantum register into $|\phi(\boldsymbol{\theta})\rangle$ state. Now the system has the state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |\psi(\boldsymbol{\theta})\rangle\,. \tag{2.41}$$

We can then apply the P_j gate controlled by the ancillary qubit and apply the remaining unitary. Finally, we apply the measurement observable controlled by the ancillary qubit. This gives the new state:

$$|\psi\rangle = \frac{(|0\rangle \otimes U(\boldsymbol{\theta})\,|\phi_0\rangle + |1\rangle \otimes \hat{M}_k V_k^j(\boldsymbol{\theta})\,|\phi_0\rangle)}{\sqrt{2}}. \tag{2.42}$$

Now we apply the Hadamard gate to the ancillary qubit, which gives:

$$|\psi\rangle = \frac{|0\rangle \otimes (U\,|\phi_0\rangle + \hat{M}_k V_k^j(\boldsymbol{\theta})\,|\phi_0\rangle) + |1\rangle \otimes (U\,|\phi_0\rangle - \hat{M}_k V_k^j(\boldsymbol{\theta})\,|\phi_0\rangle)}{2}. \tag{2.43}$$

The imaginary part of $\langle\phi_0|V_k^{j\dagger}(\boldsymbol{\theta})\hat{O}_i U(\boldsymbol{\theta})|\phi_0\rangle$ is now encoded in the ancillary qubit in the Y-basis.

Indirect analytical gradient measurement To illustrate this method first we make some modifications to our notation. We now absorb the common terms of U and V into the state and measurement operator, and define:

$$\begin{aligned} W_{k,1}(\boldsymbol{\theta}) &= \exp(i\theta_{j-1}P_k^{j-1}) \cdots \exp(i\theta^1 P_k^1) \\ W_{k,2}(\boldsymbol{\theta}) &= \exp(i\theta_n P_k^n) \cdots \exp(i\theta_j P_k^j) \\ |\phi(\boldsymbol{\theta})\rangle &= W_{k,1}(\boldsymbol{\theta})\,|\phi_0\rangle \\ Q_k(\boldsymbol{\theta}) &= W_{k,2}^\dagger \hat{M}_i W_{k,2}. \end{aligned} \tag{2.44}$$

Then we have:

$$\frac{\partial \langle \hat{O}_k(\boldsymbol{\theta}) \rangle}{\partial \theta_j} = 2\,\mathrm{Im}(\langle \phi(\boldsymbol{\theta}) | Q_k(\boldsymbol{\theta}) P_k | \phi(\boldsymbol{\theta}) \rangle). \tag{2.45}$$

We use the following construct which is related to $\langle \phi(\boldsymbol{\theta}) |\, P_k Q_k(\boldsymbol{\theta})\, | \phi(\boldsymbol{\theta}) \rangle$:

$$\langle \phi(\boldsymbol{\theta}) | (\mp i P_k) Q_k(\boldsymbol{\theta}) (\pm i P_k) | \phi(\boldsymbol{\theta}) \rangle = \mp\,\mathrm{Im}(\langle \phi(\boldsymbol{\theta}) | P_k Q_k(\boldsymbol{\theta}) | \phi(\boldsymbol{\theta}) \rangle) + C, \tag{2.46}$$

where $C = \langle \phi(\boldsymbol{\theta}) |\, Q_k(\boldsymbol{\theta})\, | \phi(\boldsymbol{\theta}) \rangle + \langle \phi(\boldsymbol{\theta}) |\, P_k Q_k(\boldsymbol{\theta}) P_k\, | \phi(\boldsymbol{\theta}) \rangle$ is a constant value.

For single-qubit gates given by $U_j(\theta_j) = \exp(i\theta_j P_k)$, $P_k \in \{X, Y, Z\}$ where $\{X, Y, Z\}$ are Pauli matrices and with exactly two eigenvalues, we have [233]:

$$\pm i P_k = \exp(\mp \frac{\pi}{4} i P_k), \tag{2.47}$$

therefore:

$$\langle \phi(\boldsymbol{\theta}) | (\mp i P_k) Q_k(\boldsymbol{\theta}) (\pm i P_k) | \phi(\boldsymbol{\theta}) \rangle = \langle \phi(\boldsymbol{\theta} \pm \frac{\pi}{2} \mathbf{e}_j) |\, Q_k(\boldsymbol{\theta} \pm \frac{\pi}{2} \mathbf{e}_j)\, | \phi(\boldsymbol{\theta} \pm \frac{\pi}{2} \mathbf{e}_j) \rangle\,, \tag{2.48}$$

and:

$$\frac{\partial \langle \hat{O}_k(\boldsymbol{\theta}) \rangle}{\partial \theta_j} = \langle \hat{O}_k(\boldsymbol{\theta} + \frac{\pi}{2} \mathbf{e}_j) \rangle - \langle \hat{O}_k(\boldsymbol{\theta} - \frac{\pi}{2} \mathbf{e}_j) \rangle\,. \tag{2.49}$$

The parameter shift rule can be generalised to a general parameter shift rule and give gradient g_{gPSR} to obtain higher-order derivatives [233, 119]:

$$g_{gPSR}(\theta, \gamma_1, \gamma_2) := r[L(\theta + \gamma_1) + L(\theta + \gamma_2)] \tag{2.50}$$

where $2r$ is the difference of the eigenvalue of the gate generator G. For single-qubit gates with Pauli operators as the generator, $r = 1$.

$$g_{gPSR}(\theta, \gamma_1, \gamma_2) = \frac{\sin(2r\gamma_1) + \sin(2r\gamma_2)}{2} g^{(1)}(\theta) - \frac{\cos(2r\gamma_1) - \cos(2r\gamma_2)}{4r} g^{(2)}(\theta), \quad (2.51)$$

and because of the sinusoidal property of the expectation value, the k-th order derivative has a constant multiplier different from the $(k-2)$-th order derivative:

$$g^{(k+2)} = -\frac{1}{4r} g^{(k)}. \quad (2.52)$$

So far the analytical gradient methods only apply to parametrised single-qubit gates. The relation presented in equation 2.47 holds only when P_k has exactly two different eigenvalues. To use the parameter shift rule for arbitrary two-qubit gates, one can decompose the multi-qubit gate into a sequence of single-qubit gates and products of the same Pauli matrices, which always have two different eigenvalues [60]. When the operators have 3 different eigenvalues, the gradient can be evaluated with a modified 4-value shift rule [122]. For other operators, the value P can be polynomially expanded into a linear combination of low-rank (2 or 3 eigenvalues) operators [122]. Alternatively, the objective function can be decomposed into trigonometric polynomials, and the gradient can be evaluated with trigonometric interpolation methods [283, 256].

It is worth mentioning that for VQE, the parameter shift rule can be implemented before encoding the fermionic excitation to the qubits, which reduces the required number of measurements [147].

Analytical optimisation

In this section, we review optimisation methods that take advantage of the analytical property of the objective function landscape. Consider the ansatz circuit as a CPTP mapping as a product of individual quantum super-operators:

$$\Phi(\boldsymbol{\theta}) = \Phi_k(\theta_k) \dots \Phi_2(\theta_2)\Phi_1(\theta_1). \quad (2.53)$$

Here $\Phi_k(\theta_k)$ are k-th parameterised quantum gates where $\Phi_k(\theta_k)\rho := U_k \rho U_k^\dagger$ and $U_k = \exp(-i\frac{\theta_k}{2}P_k)$. Here P_k are tensor products of single-qubit Pauli operators as $P_k \in \{I, X, Y, Z\}^{\otimes N}$. Each superoperator can be expanded around $\boldsymbol{\theta}_0$ as:

$$\Phi_k(\theta_0 + \theta) = a(\theta)\Phi_{ak} + b(\theta)\Phi_{bk} + c(\theta)\Phi_{ck}, \tag{2.54}$$

where $a(\theta), b(\theta) = 1 \pm \cos(\theta)$ and $c(\theta) = \frac{1}{2}\sin(\theta)$.

Now we only allow n parameters to be variables and fix all the others, the whole ansatz can be expanded into:

$$\Phi(\boldsymbol{\theta}_0+\boldsymbol{\theta}) = \prod_{k=1}^{\nu}[a(\theta_k)\Phi_{ak} + b(\theta_k)\Phi_{bk} + c(\theta_k)\Phi_{ck}], \tag{2.55}$$

where $\boldsymbol{\theta}$ is the displacement around the reference point $\theta_{\mathbf{0}}$.

Sequential optimisation with sinusoidal fitting (Rotosolve) A variety of quantum variational algorithms deal with linear objective functions. For example, the energy value for VQE is a linear objective function. For these problems, the objective function landscape of every single variable is a sinusoidal function. The full information of such a function can be obtained by measuring three points of the objective function. The sequential optimisation method [266, 190, 201] utilises this feature and calculates the minima of the sinusoidal function directly. By iteratively finding the minima of every single parameter while fixing the other parameters, a greedy method can be derived to optimise the parameters of the ansatz efficiently.

From equation (2.55), we can simplify $a(\theta), b(\theta)$ and $b(\theta)$ into a simpler form if we only allow one single parameter to change. When all the parameters are independent, the objective function can be transformed as:

$$\mathcal{L}(\boldsymbol{\theta}_0+\theta\cdot e_i)=\mathcal{L}_i(\theta)_{\boldsymbol{\theta}_0}=\mathrm{Tr}[M\Phi_i(\boldsymbol{\theta}_0+\theta\cdot e_i)] \tag{2.56}$$

$$=\mathrm{Tr}[M(a(\theta)\Phi_a+b(\theta)\Phi_b+c(\theta)\Phi_c)] \tag{2.57}$$

$$=A\sin(\theta+\phi)+C. \tag{2.58}$$

The parameter can be evaluated analytically with the following equation when the Hermitian generator has exactly two eigenvalues.

$$\theta^*=\phi-\tfrac{\pi}{2}-\arctan2\left(\phi_1,\ \phi_2\right)+2\pi k, \tag{2.59}$$

where:

$$\phi_1=2\mathcal{L}_i(\theta)_{\boldsymbol{\theta}_0}-\mathcal{L}_i(\theta+\frac{\pi}{2})_{\boldsymbol{\theta}_0}-\mathcal{L}_i(\theta-\frac{\pi}{2})_{\boldsymbol{\theta}_0} \tag{2.60}$$

$$\phi_2=\mathcal{L}_i(\theta+\frac{\pi}{2})_{\boldsymbol{\theta}_0}-\mathcal{L}_i(\theta-\frac{\pi}{2})_{\boldsymbol{\theta}_0}. \tag{2.61}$$

All the single-qubit gates have Pauli operators as their generators, which satisfy the above equations. The cost function can also sum sinusoidal functions with different periods [283, 266] when the generator has more than two distinct eigenvalues.

When several gates share the same parameter, the objective function can be transformed into a similar form but with different oscillation periods:

$$\mathcal{L}(\boldsymbol{\theta}_0+\theta\cdot e_i)=A\sin(k\theta+\phi)+C, \tag{2.62}$$

where k is the number of appearances of θ_i in the ansatz. The same approach can be applied to find the analytical solution.

It is important to note that the conditions for validity of this optimisation method imply that it cannot be applied to some ansätze (for instance, in the case where a

parametrised controlled unitary is used). In general terms, Rotosolve and its subsequent extension require that all parametrised gates subject to optimisation have 2π-periodicity and full rank for the Hermitian matrix generating the rotation. The work presented by Wierisch *et al.* [283], relying on alternative general parameter-shift rules, demonstrates a generalisation of Rotosolve which allows for it to be used on all quantum gates with arbitrary frequencies.

2.4 Simulating Quantum Neural Network Training with Label Qubits

This section discusses a theoretical proposal to encode the cost function into the parameterised quantum circuits (PQCs) for classification tasks. Just as with classical deep learning, parameter training is the most time-consuming process of building a PQC model. Finding efficient ways to obtain optimal or near-optimal parameters is a major objective of current research. Several interesting methods have been proposed to optimise parameters. Firstly, there are gradient-based optimisation methods [108]; the partial derivative of the expectation value can be directly evaluated with a Hadamard Test [61, 229], or the *shift-rule* can be used to do a Hadamard Test with indirect measurements [184, 233]. Alternatively, the gradient-free optimisation method *rotosolve* mentioned in the previous section can be used [201, 190].

Motivation

The target to be optimised (in machine learning, the *cost function*) is typically calculated from measured expectation values. Therefore, the value of the cost function is no longer encoded in the quantum state. This brings us some inconvenience. For example, the Hadamard Test and shift-rule methods can be used together with the chain rule to calculate the cost-function gradient [17], but this requires several function

evaluations[1]. In addition, the gradient-free method *rotosolve*, which was developed for the specific value landscape of Hermitian expectation values, cannot be applied to classical non-linear cost functions [201, 190].

Therefore, an open question is whether we can find a method which embeds the cost function into a quantum circuit. Embedding the cost function into the quantum circuit directly will not only allow PQC-based ML models to be trained with the optimisation techniques we mentioned above but also opens the possibility for performing further quantum operations after the cost function is evaluated, such as quantum-enhanced measurement [98] and reduction of measurement times by phase estimation [277]. Here we present a method which achieves the cost function embedding for PQC-based ML models.

Embedding Cost Functions and Training Labels

I propose two cost functions that can be encoded directly through a quantum circuit, i.e., the outcome of the quantum circuit corresponds to the evaluation of the cost function for a given input. We then propose a corresponding data encoding method that allows calculating averaged cost function values among the encoded training dataset. Our approach could be used in combination with existing algorithms [85, 234, 48, 103] by replacing the part of data encoding and cost function evaluating in these existing frameworks.

Cost Function Embedding For the concrete implementation, we add another qubit $|\phi\rangle$, which holds the desired output for the supervised learning task, which we call the *label qubit*.

For the binary classification problem, the label $\beta \in \{0, 1\}$ can be encoded as states $|0\rangle$ or $|1\rangle$ of the label qubit.

[1]Suppose our cost function is given by $f(x)$ and the measurement expectation value is given by $g(x)$, the derivative is then given by $f(g(x))' = f'(g(x))g'(x)$. To calculate this derivative, both the intermediate gradient $g(x)'$ and the expectation value $g(x)$ need to be evaluated

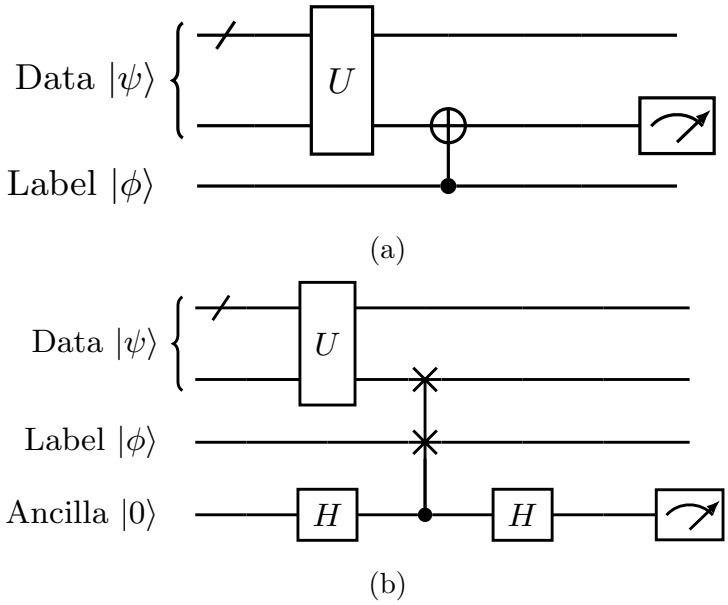


Figure 2.2: Two ways to embed cost functions into quantum circuits. Reproduced from Ref. [42]. The data qubits store the input data and the label qubit stores the label (expected output). U denotes an arbitrary ansatz, while we consider the output from the ansatz is encoded in one qubit (here it is the qubit represented as the second line of the circuit). In Figure (a) we apply a CNOT gate on the output qubit, and flip the state based on the label qubit. This will give us the CNOT cost function and encode the output into the Z-axis expectation value of the output qubit. In Figure (b) we add an ancillary qubit and apply a Fredkin gate between the ansatz output and the label, controlled by the ancillary qubit. This will encode the overlap between the ansatz output and the label into the ancillary qubit.

A simple way to implement a cost function for binary classification is to use a CNOT gate to flip the output state based on the input training label, see figure 2.2a. Notably, under this operation, the output qubit remains $|1\rangle$ if the prediction and the expectation value are unequal, and will be $|0\rangle$ if they are the same.

Suppose the expectation value of the measurement $\langle M \rangle = \alpha$, and recall that the label is $\beta \in \{0, 1\}$, then the CNOT cost function can be shown to be

$$E_{cost(\psi)} = (1 - 2\beta)\alpha + \beta \tag{2.63}$$

Note that this can be transformed into a well-studied loss function in machine learning, the so-called hinge loss [76].

In general, we want to treat continuous as well as binary outcomes, i.e., $\mathcal{Y} = \mathbb{R}$. In

the quantum setting, this means that the ancillary qubit would no longer be in one of the states $\{|0\rangle, |1\rangle\}$, and the above CNOT trick can no longer be applied. To handle this scenario, we can fall back to a well-known method that has been proposed for quantum classification, the Swap test [36, 5]. The Swap Test is a method to encode the overlap between two unknown states into an ancillary qubit. A swap test consists of a Hadamard gate followed by a controlled swap gate between two states and a final Hadamard gate. If the two states that we aim to compare, let them be $|\psi\rangle$ and $|\phi\rangle$, are the same, i.e., $\langle\psi|\phi\rangle = 1$, then the Swap Test acts as the identity. This means that the output state remains the same. However, if the two states are different, i.e., $\langle\psi|\phi\rangle \neq 1$, then the difference will be reflected in the phase of the control qubit. A swap test between the label state and the prediction state can hence be used to estimate the overlap. We can then use the overlap as the output of our cost function, see figure 2.2b.

Suppose the input state is $|\psi\rangle$, then the Z-axis expectation value of the measurement of the ancillary qubits can be written as [36, 5]

$$E_{cost(\psi)} = \frac{1 - \langle\phi| \rho_{output} |\phi\rangle}{2} \tag{2.64}$$

The controlled swap gate is also referred to as the quantum Fredkin gate, and several physical realizations using photonics or superconducting systems have been realized [207, 199, 158]. We hence have a single-qubit encoding of the cost function.

Data Encoding Now let's consider the data encoding for the cost functions discussed above. In classical deep learning, the cost function is evaluated individually for each sampled data point from the training dataset. Then the cost function value is averaged across all calculated values. This trivially costs $O(N)$ repetitions of the cost function evaluation for N data points. Since we are only interested in the averaged cost function value instead of the individual value, it is natural to ask whether

quantum parallelism can speed-up its evaluation. As we will now show in detail, we find interestingly that there is in fact *no quantum speed-up that can be gained*.

Consider a single data point. The goal of this procedure is to transfer each data point into a quantum state. For fully digitised input data, i.e. if all the input values are binary (for example, 0 or 1), the encoding step can be done by mapping the classical "1" to the $|1\rangle$ state and classical "0" to the $|0\rangle$ state. Alternatively, if we would like to prepare non-digital input for training, we first normalise it to the range $\left(0, \frac{\pi}{2}\right)$, and then encode the data in the amplitude of the input qubits. This is done via the rotation angles, and we obtain for a single data point j the angle set $\gamma^j := \{\gamma_0^j, \gamma_1^j ... \gamma_n^j\}$ such that

$$|\psi_i^j\rangle = \cos(\gamma_i^j)\,|0\rangle + \sin(\gamma_i^j)\,|1\rangle \tag{2.65}$$

Where i is the dimension index of the data point. Taking a data point in N dimensions, the entire state is then given by the tensor product

$$|\psi^j\rangle = \bigotimes_{i=1}^{N} |\psi_i^j\rangle \tag{2.66}$$

This method is often referred to as *qubit encoding* [249]. The qubit encoding method stores each dimension of the data in the amplitude on an individual qubit. It hence maps the data to an exponentially large feature space. This allows the use of simpler models in the classifier task, such as linear classifiers. Additionally, this method is well suited for NISQ devices as it requires only single-qubit gates to implement.

Next, we need to consider how to represent our entire dataset and let the quantum circuit work out the averaged cost function naturally. Suppose the circuit giving the cost function is represented by the unitary C, the measurement by the operator O, and the initial state for each data point be $|\chi_i\rangle = |\psi_i\rangle \otimes |\phi_i\rangle$. The cost function with respect to the i-th datapoint x_i (encoded in ψ_i) can then be represented by

$$[E_{cost}]_{x_i} = \langle \chi_i | C^\dagger O C | \chi_i \rangle = \langle \chi_i | D | \chi_i \rangle , \tag{2.67}$$

where $D = C^\dagger O C$. The average of the cost function is then given by

$$\overline{E_{cost}} = \frac{1}{N} \sum_{i=0}^{N} \langle \chi_i | D | \chi_i \rangle \tag{2.68}$$

where N is the number of data points in the dataset.

Here we can use a mixed state, which is the classical average of all possible states of our datasets. To see this, recall that the desired cost function value is the average of the cost function values across the entire dataset. If we hence input a mixed state,

$$\rho_{mix} = \frac{1}{n} \sum_{n} |\phi_i\rangle |\psi_i\rangle \langle \psi_i | \langle \phi_i | , \tag{2.69}$$

which is the average of all states in our dataset, the outcome of the calculation is similarly the average over the dataset. This mixed state can be constructed by randomly selecting one of the samples in every single run of the algorithm, and then averaging the outcome over all runs. Since all the qubits are separable, we need single-qubit rotations for each qubit to prepare each state. Notably, the accuracy of this will depend on the concrete distribution and the number of repetitions.

Rather than using the mixed state described above, it is natural to think if we can prepare data in a pure superposition state. We find that it is possible, but that there is no speed-up compared to the mixed-state preparation method. Consider a uniform superposition state over the entire dataset. After evaluating the circuit, we will then obtain the expectation value as:

$$\begin{aligned} \overline{E_{cost}} =& \sqrt{\frac{1}{N}} \sum_{i=0}^{N} \langle \chi_i | D \sqrt{\frac{1}{N}} \sum_{i=0}^{N} | \chi_i \rangle \\ =& \frac{1}{N} \sum_{i=0}^{N} \langle \chi_i | D | \chi_i \rangle + \frac{1}{N} \sum_{i=0}^{N} \sum_{j \neq i}^{N} \langle \chi_i | D | \chi_j \rangle \end{aligned} \tag{2.70}$$

The first term of equation 2.70 would give us then exactly what we desire if $i = j$. However, it also includes the terms for $i \neq j$. Although the input states $|\chi_i\rangle$ are always orthogonal to each other for fully digitised input data, we cannot guarantee that $\langle\chi_j| D |\chi_i\rangle = 0$ holds. To fix this problem, we can introduce an additional register $|\epsilon_i\rangle$ which we call the *index qubits*. The new initial state is then given by $|\chi_i\rangle = |\psi_i\rangle \otimes |\phi_i\rangle \otimes |\epsilon_i\rangle$. We put the index of each data point into these qubits and don't apply any operation on them throughout the circuit. Since $\{|\epsilon_i\rangle\}$ is an orthogonal set, the entire state after running the circuit will also be orthogonal, and we will obtain $\langle\chi_i| D |\chi_j\rangle = 0$ for any $i \neq j$.

It is well known that such a superposition state can be prepared using variants of Grover's state preparation [105, 231]. However, these approaches require relatively deep circuits, and are therefore not immediately applicable for NISQ devices. Moreover, the index qubits we introduced here do not participate in any computation of the PQC. We can measure these index qubits at the very beginning, resulting in a state collapse into a specific data point. This superposition-state encoding is therefore equivalent to the mixed-state encoding we proposed above, which randomly selects a data point at the beginning and averages the outcome at the end. This encoding will have the same outcome as long as the difference between the random number generator and the quantum randomness is negligible.

By simply preparing the input data into a superposition the expected value of the quantum circuit is not the same as evaluating each data point individually and taking the average. This is because of the quantum interference between each data point. Once the interference is removed by adding index qubits, there is no quantum speed-up from encoding the data in superposition.

Simulation Result

We use the gradient-finding method using the Hadamard Test [61, 229] and the Adam optimiser [138]. First, we give a brief review of the Hadamard Test gradient estimation, and then demonstrate how to use Hadamard Test gradient estimation together with the proposed method in this section. Here we show two tasks, to train a classifier for an XOR gate and to train a classifier for a more realistic problem, the canonical Iris flower dataset.

Gradient Estimation with the Hadamard Test Gradient estimation with the Hadamard Test has previously been used to calculate the partial derivative of the eigenenergy of a molecule [61, 229]. The Hamiltonian Pauli terms used to approximate molecular energies are typically multi-qubit terms. For this reason, they can require a separate control operation on a large number of qubits, which in NISQ devices has the potential to introduce a prohibitive level of noise [194]. Here, by encoding the cost function valuation into a single qubit, the control operation is reduced to a simple controlled-Z gate which is acceptable for NISQ applications.

A circuit that implements this Hadamard Test gradient estimation together with cost function embedding is shown in figure 2.3. Using this method, it is possible to evaluate the partial derivative of the expectation value of the output qubit with respect to the rotation angle θ. Here θ parametrises the rotation generated by some Pauli operator P. For example, for P being the Pauli X operator, we will have a single qubit rotation gate with angle θ about the X axis. Typically there will be multiple single-qubit gates inside the ansatz. We can perform this method on each one of them to get the partial derivative of our cost function with respect to each rotation angle. The whole process is given by the following: We begin by preparing an ancillary qubit in the $|0\rangle$ state, and then apply a Hadamard gate to obtain the plus state, $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$. Next, we apply a controlled-P gate right after the single-qubit rotation

gate $e^{i\theta P}$, where the control is on the ancillary qubit. Finally, because we encoded our cost function value in the Z-axis of the second qubit, we add a controlled-Z gate between the ancillary qubit and the second qubit, followed by another Hadamard gate on the ancillary qubit which rotates the state back to the $\{0, 1\}$-basis. The resulting ancillary qubit now contains the gradient encoded in its phase. Finally, we apply a $\frac{\pi}{2}$ rotation about the X-axis on the ancillary qubit so that the gradient is encoded in the Z-axis of the ancilla and can be determined by measuring in the $\{|0\rangle, |1\rangle\}$ basis.

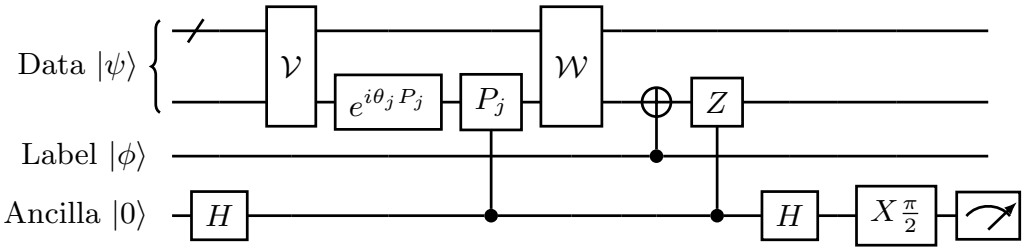


Figure 2.3: Estimating the gradient value by measuring the ancillary qubit. Reproduced from Ref. [42]. The circuit $U = V(e^{i\theta_j P_j} \otimes \mathbb{I}_k^{(n-1)})W$ is the parametric quantum circuit and the entry of the gradient corresponding to parameter θ_j is encoded into the expectation value of the ancillary qubit.

We can therefore use the circuit in figure 2.3 to estimate the gradient by simply measuring the single ancillary qubit. The input data is encoded into data register ψ represented as the first two lines, and the expected output (label) is encoded in the label register $|\phi\rangle$. When we optimise the parameter θ_j we decompose the ansatz U into $U = V(e^{i\theta_j P_j} \otimes \mathbb{I}_k^{(n-1)})W$ where $\mathbb{I}_k^{(n-1)}$ is the identity operation on all the other qubits except the k-th qubit (to which we applied the single qubit rotation), and $e^{i\theta_j P_j}$ is the single qubit rotation gate parametrised by θ_j where P_j is a Pauli operator. We first insert a controlled-P_j gate right after the single qubit rotation, then add a controlled-Z gate after the output of the cost function. Both these gates are controlled by an ancillary qubit prepared in state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. In the end, we swap the Z and X axes of the ancillary qubit with a Hadamard gate and then do a $\pi/2$ rotation along the X axes. The partial derivative of the cost function with

respect to θ_j is stored in the ancillary qubit.

XOR Experiment We can now proceed to test the above methods for training PQCs. The first example task consists of training the circuit to perform the classical exclusive or (XOR) operation. Here we require the circuit to implement the truth table of XOR, i.e., yield true if and only if the input bits are different, and false otherwise. We use $|0\rangle$ to denote the FALSE value, and $|1\rangle$ to denote TRUE. Here we demonstrate the method involving encoding data in superposition. The circuit ansatz is shown in figure 2.4.

The input bits are $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, the corresponding labels, i.e., evaluations of XOR on the respective input, are $|0\rangle$, $|1\rangle$, $|1\rangle$, $|0\rangle$, and the indices are given by $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$. Including an extra qubit as the ancillary qubit, the final input state is given by

$$|\chi\rangle = \frac{1}{2}(|000000\rangle + |010110\rangle + |101010\rangle + |111100\rangle) \tag{2.71}$$

After the gradients are determined by evaluating the circuit they are used to update the parameters using the Adam optimiser [138].

The simulation result is shown in figure 2.5. During the experiment, we were able to see that circuits with CNOT or Fredkin cost functions can successfully give us the correct gradient direction. We observed that training converges for both fully digitised data encoding and mixed-state encoding, and the convergence rates are the same.

Iris Experiment Next, we investigate a more realistic scenario in which we implement a classifier for the Iris dataset [89]. The Iris dataset contains 150 labelled examples in total, with three different types of Iris flowers. Each example is described by four features. We prepare a mixed state with the method from equation 2.69, and

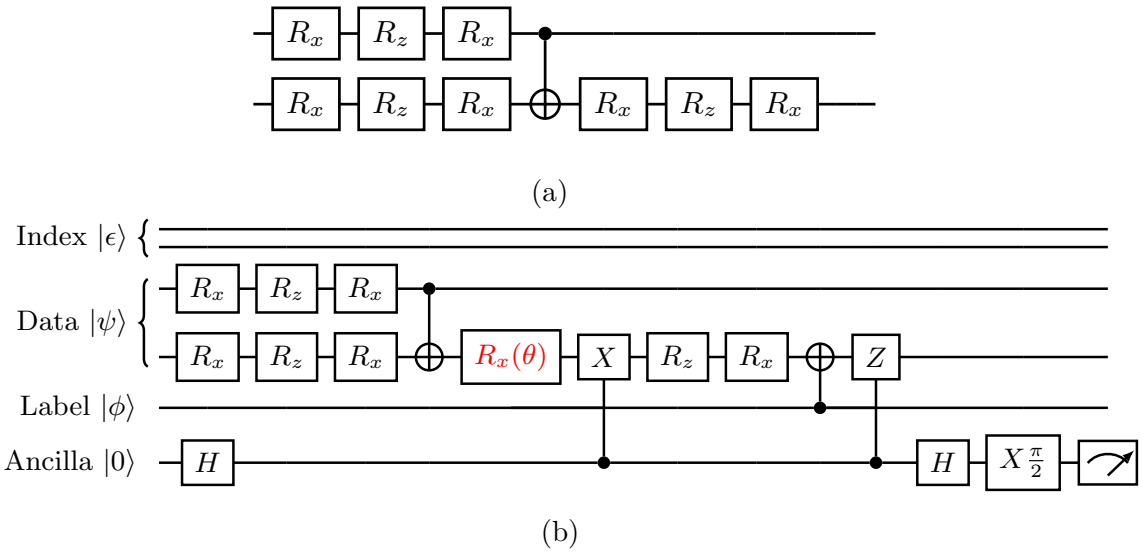


Figure 2.4: (a) The ansatz used for training the XOR circuit. (b) Evaluation of the partial derivative of theta for training the XOR circuit with CNOT cost function. Reproduced from Ref. [42]. This circuit is used to find the partial derivative of the CNOT cost function with respect to the rotation angle θ. Here we choose an ansatz with arbitrary single qubit rotation applied on each input qubit first, then use a CNOT gate to exchange the information between two qubits, and again apply an arbitrary single-qubit rotation on the output qubit. We use three parametrised single-qubit rotations, Rx, Rz, Rx, to construct an arbitrary single-qubit rotation. The input state is encoded with equation 2.71.

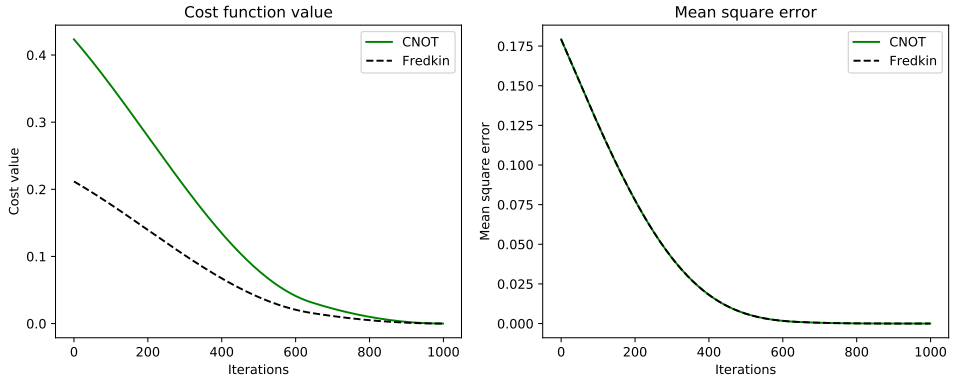


Figure 2.5: Numerical simulation result for circuit shown in figure 2.4. Reproduced from Ref. [42]. Here we tested the cost function value and the mean square error for the same training dataset prepared as equation 2.71. During the training process, the circuits for evaluating the partial derivative of each of the parameters θ_i are generated. For each iteration, all these circuits are evaluated and give the derivative of each parameter, which is the gradient of the cost function. Then the gradient value is passed to the Adam optimiser. Here the learning rate of the optimiser is 10^{-3}. At the end of each iteration, the parameters are updated by the Adam optimiser. The result here shows both CNOT and Fredkin cost functions can be used for training, and they give similar convergence speeds.

the training circuit is shown in figure 2.6. Several interesting results were found during the training process. See figure 2.7. First, we report that by using the Adam optimiser, the PQCs can converge. The CNOT cost function and the Fredkin cost function were able to achieve similar training performance and convergence rates.

In order to investigate the robustness of the circuit, we applied the depolarisation channel to the system. The channel is described as $\Delta_\lambda(\rho) = \lambda\rho + (1-\lambda)/2^n\mathbb{I}$ where n is the number of qubits. We observe that the algorithm still converges to a region close to the optimal point when $\lambda = 0.999$ is applied. When $\lambda = 0.99$, the cost function no longer converges.

We can see for the circuit in figure 2.6 that we can tolerate depolarisation noise when $\lambda = 0.999$, which is quite close to gate fidelities of state-of-art NISQ devices [141, 156]. This method could also be combined with error mitigation [82], which can suppress some errors and make this algorithm even more robust to the noise on a

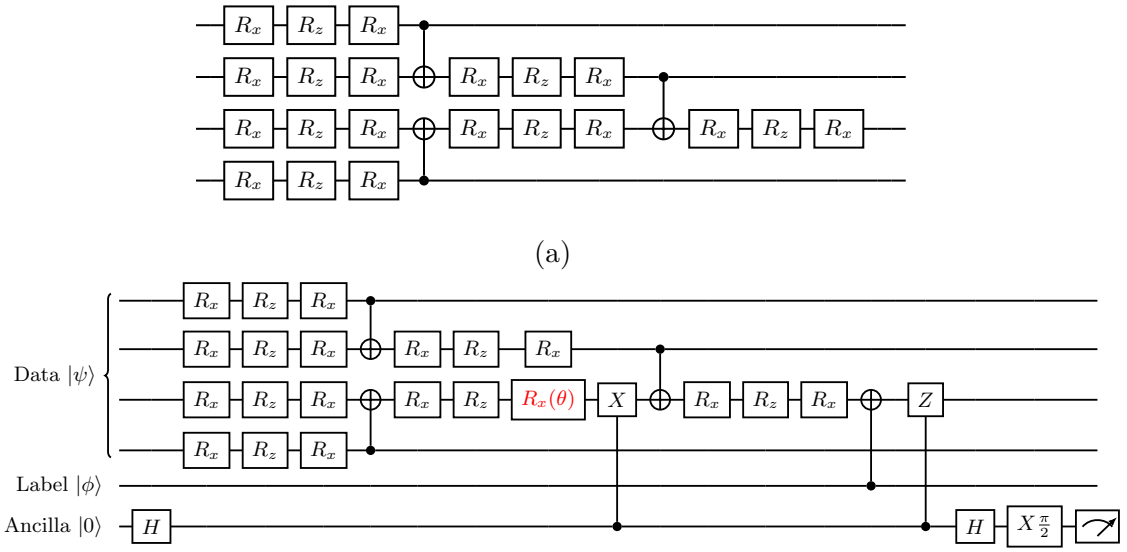


Figure 2.6: (a) The ansatz of the hierarchical classifiers. (b) Training circuit for quantum hierarchical classifiers. The input state is encoded with equation 2.69. Reproduced from Ref. [42]. Here we demonstrate that the circuit gives the partial derivative of the cost function against the parameter θ, which parametrises one of the single qubit rotation gates (highlighted in red). A controlled-X gate (CNOT) is applied right after the parametrised gate [103].

NISQ machine.

Conclusion

We now summarise the advantages of our proposed method. Firstly, the encoding of the value of the cost function into a measurement expectation value enables the ability to do further quantum information processing. For example, the direct usage of advanced optimisation methods such as the Hadamard Test, *shift-rule*, and *rotosolve*. Since the cost function value is encoded in the state of a single qubit, the implementation of a full optimisation algorithm making use of this method can be straightforward.

Secondly, the proposed data encoding method allows for efficiently estimating expectation values across the training dataset. It is classically inefficient to calculate the cost function for each example and then average them across the entire dataset.

Instead, the proposed method indicates that the expectation value can be obtained via random sampling from the training dataset.

We show that simply encoding the dataset into a superposition state will not give us the average of the cost function values. To fix this issue, we introduced the index qubit. However, we showed that index qubit would make it equivalent to doing a mixed-state preparation and that such superposition state preparation has no speed-up compared to mixed-state preparation.

In conclusion, we have proposed a method to encode cost functions into quantum circuits and a corresponding method for preparing the input data. This method, therefore, enables quantum information processing on the cost function value. The averaged cost value can be calculated across the entire dataset with both mixed-state data preparation and superposition data preparation. We demonstrated gradient evaluation of the cost function with the Hadamard Test and investigated its performance under a depolarisation noise channel.

2.5 Discussion

In this chapter, we have reviewed the principles of variational quantum algorithms (VQAs). The focus is given to VQE for solving molecular ground energies and implementing quantum machine learning with variational circuits. As a preliminary study, a simulation experiment for embedding the cost function and dataset into the circuit has also been demonstrated. The VQAs must be carefully chosen to fit in the physics setup to demonstrate the advantage of a multi-level system. Here two experiments are chosen.

In the remainder of this thesis, we will explore the use of quantum systems with more than 2 levels to implement VQAs. In order to quantify the advantage from higher levels, we here carefully choose the VQA to implement, such that it can make direct comparisons with a two-level system. First, we use our device as a 3-level

system and implement a variational quantum classifier for the Iris dataset. The Iris dataset requires 4 degrees of freedom to encode it; therefore, it cannot be directly implemented into a two-level system without dimension reduction. The device is also used as a 4-level system, emulating the behaviour of a two-qubit system. The emulator implements a variational quantum eigensolver for hydrogen molecules. This experiment reduces the required qubit number for the same algorithm and demonstrates the advantage of a multi-level quantum system. The details are presented in chapter 7.

Before demonstrating these implementations of VQAs on multi-level quantum systems, we will describe in detail the experimental hardware system that will be used. The process of device design, fabrication, calibration and characterization is introduced in the following chapters.

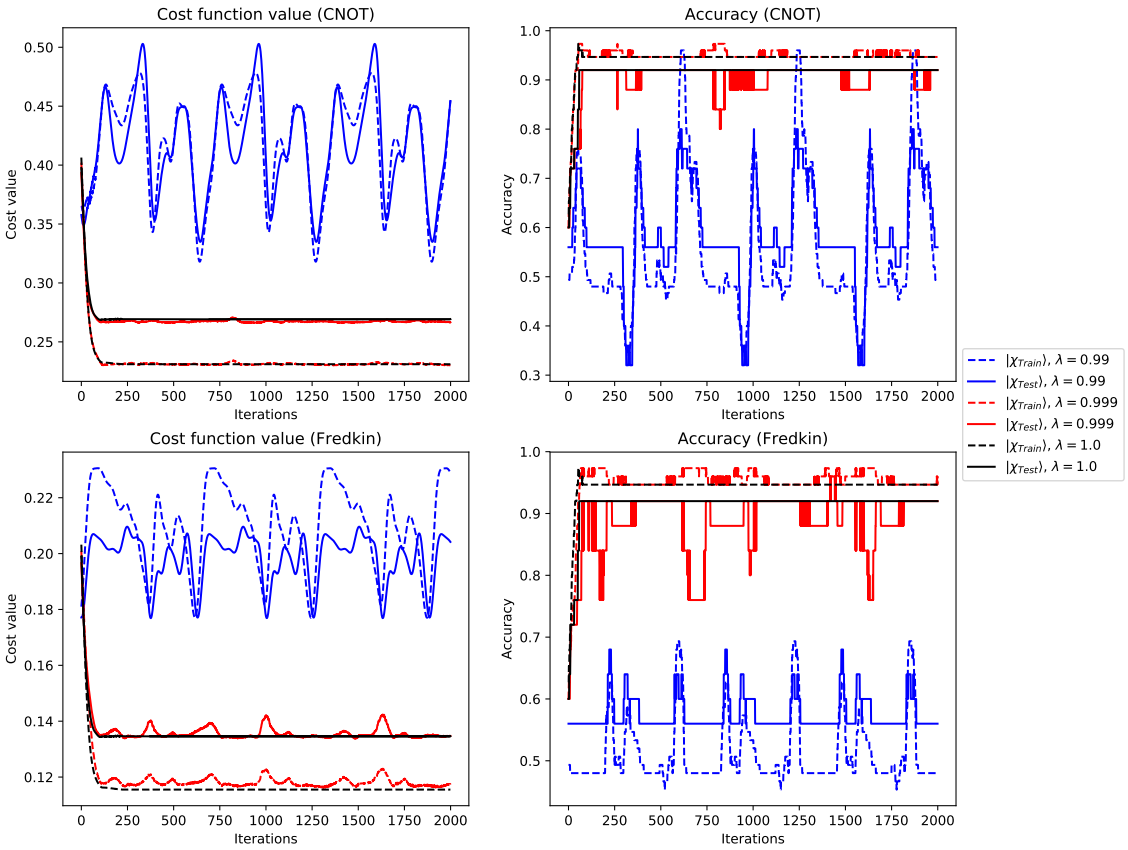


Figure 2.7: Simulation of training the hierarchical classifier for the Iris dataset. Reproduced from Ref. [42]. Here we chose two flowers (labels 1 and 2) from the Iris dataset. The training process is the same as figure 2.5. Here the learning rate is 10^{-2}. We consider the prediction of the flower label as 2 if the probability to measure and get state $|1\rangle$ is greater than 0.5, otherwise, the prediction will be considered to be label 1. The left two plots depict the cost function values over time, under different levels of noise showing in different colours. The training converges without noise or under a small amount of noise, but it fails to converge when subjected to significant noise. The right-side two plots show accuracy measurements, supporting that excessive noise can prevent the training process from converging. The accuracy is the ratio of the correct prediction among the entire training dataset $|\chi_{Train}\rangle$ (dashed lines) or the test dataset $|\chi_{Test}\rangle$ (solid lines). Then we add the depolarisation channel $\Delta_\lambda(\rho) = \lambda\rho + (1-\lambda)/2^n\mathbb{I}$ where n is the number of qubits to the circuit and compare the simulation between results with or without noise.

Chapter 3

Superconducting quantum circuits

This chapter presents a review of superconducting quantum circuits, summarised from textbooks and previously published review papers. At the end of this chapter, I summarise the parameter requirements of our quantum processor and propose the design of the resonator that is optimal for distinguishing the lowest four-level states of a transmon.

This chapter discusses the theory of superconducting quantum circuits, focusing on the concepts needed to understand the research presented in the remainder of the thesis. We first introduce the Hamiltonian of a lumped element circuit with circuit quantisation techniques. We proceed to describe the key physics used to operate the commonly-used transmon qubit, including the single transmon dynamics with an external electromagnetic field, the interaction between the transmon and a linear oscillator for qubit readout, and the interaction between two transmons for entanglement generation. We give a specific focus on operating more than two energy levels of a transmon and discuss reading out these higher-level transmon states. Then we discuss the device parameters for an operable transmon.

3.1 Introduction to superconducting quantum circuits

Superconducting circuits [189, 273, 31, 268, 167, 282, 140] are electrical circuits made of superconducting materials and operated at cryogenic temperatures. The property of superconductivity makes the current dissipationless and allows superconducting circuits to exhibit quantum behaviour when weakly driven by external electromagnetic fields. The temperature at tens of millikelvin keeps the superconducting circuits in their ground state for most of the time when idle.

Circuit quantum electrodynamics (Circuit QED) [26, 290, 172, 52, 19, 25] aims to model the fundamental interactions between electromagnetic fields and on-chip artificial atoms, usually made from mesoscopic superconducting circuits. These artificial atoms have an atom-like energy spectrum and exhibit quantum behaviour at cryogenic temperatures. Utilizing the nanofabrication technology developed in the past century from the semiconductor industry, superconducting circuits can be fabricated to meet the engineered parameters for quantum state manipulation [148]. These advantages make superconducting circuits a promising candidate as a physical system

for quantum information processing.

The research to develop quantum processors using superconducting circuits has continued for a few decades, and various superconducting qubits have been experimentally studied [140]. The relaxation time (T_1) and dephasing time (T_2) describe how long a qubit retains its information, which are key characteristic parameters of the quality of qubits. The relaxation and dephasing time details are described in chapter 4. Figure 3.1 describes the evolution of qubit lifetimes and coherence times with different designs. In the past two decades, the coherence times in superconducting qubits have increased exponentially and are now in the range $10^2\ \mu s - 10^3\ \mu s$. Recently, active error correction has been introduced to qubit designs to increase the encoded qubit lifetime [56, 182, 181].

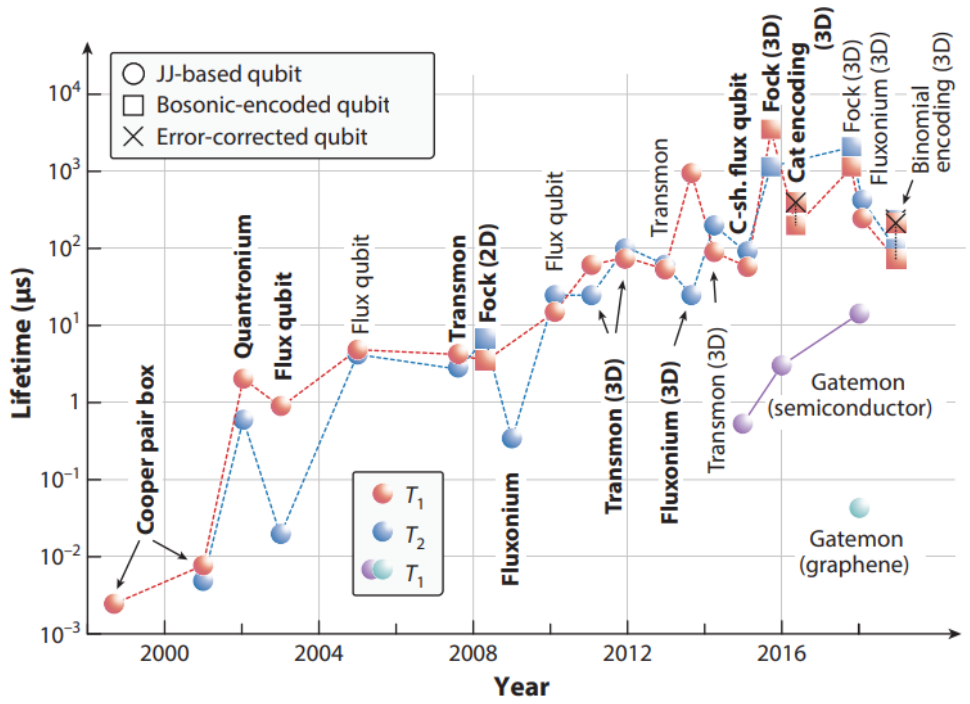


Figure 3.1: Evolution of lifetimes and coherence times in superconducting qubits adapted from [140] with permission of Annual Reviews. Bold font indicates the first demonstration of a given modality. Error-corrected qubits represent qubit encodings in which a layer of active error correction has been implemented to increase the encoded qubit lifetime. For encoded qubits, the non-error-corrected T_1 and T_2 times used in this figure are for the encoded, but not error-corrected, version of the logical qubit.

We now focus on reviewing the theory of Circuit QED with a specific focus on the design, readout, and operation of a transmon, which is used later in the experiments.

Lumped elements

Superconducting circuits typically can be modelled as lumped elements, since their operating frequency is $4-10$ GHz, at which light has a free-space wavelength ($\lambda \approx 1$ cm) significantly larger than the device's physical length ($s \approx 1$ mm). The typical elements of superconducting circuits include capacitors, inductors, and Josephson junctions. These two-port elements exhibit simple voltage and current relationships between the ports. To form a Hamiltonian describing the dynamics of superconducting circuits, a change of variables technique is applied and the current and voltage are replaced by charge Q and flux Φ as follows:

$$\Phi(t) = \int_{-\infty}^{t} v(t^{'})dt^{'} \tag{3.1}$$

$$Q(t) = \int_{-\infty}^{t} i(t^{'})dt^{'} \tag{3.2}$$

where $i(t)$ and $v(t)$ are the current and the voltage as a function of time t.

Capacitor The energy stored in the capacitor of capacitance C can be written as

$$E_C = \frac{1}{2C}(Q - Q_{offset})^2 \tag{3.3}$$

where Q_{offset} is the charge offset that can be present in the system.

Inductor The energy stored in the inductor of inductance L can be written as

$$E_L = \frac{1}{2L}(\Phi - \Phi_{offset})^2 \tag{3.4}$$

where Φ_{offset} is the flux offset that can be present in the system.

Josephson junction Apart from capacitors and inductors, the Josephson junction (JJ) is another key component for superconducting circuits, which provides non-linearity to the Hamiltonian [173, 144]. The Josephson junction is a non-dissipative component made from a thin insulating barrier sandwiched by two superconducting electrodes [130]. The JJ exhibits the Josephson effect; the current I passing through the JJ depends on the phase difference θ between two superconducting electrodes, given by the first Josephson relation $I = I_c \sin(\theta)$ where I_c is the critical current of JJ. The phase difference θ is further related to the flux stored across the junction ϕ, given by $\theta = 2\pi\phi/\Phi_0$. The inductive energy of a JJ, therefore, exhibits a non-linear relationship to phase, and the energy increase when the supercurrent flows through it can be calculated as:

$$\Delta E_J = \int_{t1}^{t2} IV dt = \int_{t1}^{t2} Id\Phi = \int_{\phi_1}^{\phi_2} I_c \sin\phi d\left(\Phi_0 \frac{\phi}{2\pi}\right) = -\frac{\Phi_0 I_c}{2\pi}\Delta\cos\phi \tag{3.5}$$

And the Josephson junction energy is then defined as:

$$E_J(\phi) = -\frac{\Phi_0 I_c}{2\pi}\cos(\phi) = -E_J\cos(\phi) \tag{3.6}$$

where $E_J = \Phi_0 I_c/2\pi$.

By applying the canonical quantisation with commutation relation $[\hat{\Phi}, \hat{Q}] = i\hbar$, the variable Q and Φ can be replaced with operator $\hat{Q}$ and $\hat{\Phi}$. Second quantisation is applied to transform the system into the Fock-basis (photon-number basis) description to model the discrete quantum system for quantum information processing. The coordinate operators are replaced with creation and annihilation operator $\hat{a}^\dagger$ and $\hat{a}$ as

$$\hat{\Phi} = \Phi_{ZPF}(\hat{a}^\dagger + \hat{a}), \qquad \hat{Q} = iQ_{ZPF}(\hat{a}^\dagger - \hat{a}) \tag{3.7}$$

where Φ_{ZPF} and Q_{ZPF} are the flux and charge zero-point fluctuations respectively.

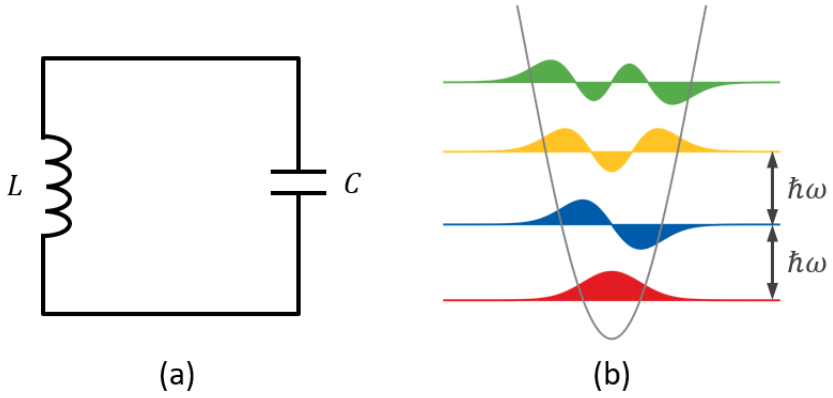


Figure 3.2: LC resonator. (a) The lumped element circuit diagram of an LC resonator, which is composed of an inductor L and capacitance C. (b) The energy landscape of an LC resonator. After quantisation, the neighbouring energy levels have the same energy gap $\hbar\omega$, where $\omega = 1/\sqrt{LC}$ is the oscillation frequency of the resonator.

Harmonic oscillators

Superconducting circuits behave similarly to room temperature electric circuits, and an *ab initio* theory for superconducting circuits can begin with the Hamiltonian of a standard LC circuit, which would have an energy oscillation between a capacitor and an inductor. The Hamiltonian dynamics can be formulated by [148]

$$\mathcal{H} = T + V,\ T = \frac{Q^2}{2C},\ V = \frac{\Phi^2}{2L} \tag{3.8}$$

where T and V are the kinetic and the potential energy in standard Hamiltonian dynamics, corresponding to the electrical and magnetic energy in our context. Q is the charge on the capacitor, and Φ is the flux passing through the inductor. L and C are the inductance and capacitance, respectively. The Hamiltonian can satisfy

$$\frac{\partial \mathcal{H}}{\partial Q} = -\dot{\Phi}, \qquad \frac{\partial \mathcal{H}}{\partial \Phi} = \dot{Q} \tag{3.9}$$

where Q‘ and Φ can be used as the system’s canonical variables, and the quantised Hamiltonian can be written with the canonical quantisation, which gives

$$\hat{\mathcal{H}} = \frac{\hat{Q}^2}{2C} + \frac{\hat{\Phi}^2}{2L} \tag{3.10}$$

with commutation relation $[\hat{\Phi}, \hat{Q}] = i\hbar$. The Hamiltonian is usually written in a more convenient form

$$\hat{\mathcal{H}} = 4E_C\hat{n}^2 + \frac{1}{2}E_L\hat{\phi}^2 \tag{3.11}$$

where $E_C = e^2/2C$ is the charging energy, $E_L = (\Phi_0/2\pi)^2/L$ is the inductive energy, $\hat{n} = \hat{Q}/2e$ the operator for the number of Cooper pairs and $\hat{\phi} = 2\pi\hat{\Phi}/\Phi_0$ the phase across the Josephson junction. $\Phi_0 = h/2e$ is the magnetic flux quantum, h is the Planck constant and e is the electron charge. The Hamiltonian of a quantised resonator is then given by

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2}) \tag{3.12}$$

$$\hat{\phi} = \sqrt{\frac{\hbar Z_C}{2}}(\hat{a}^\dagger + \hat{a}), \qquad \hat{q} = i\sqrt{\frac{\hbar Z_C}{2}}(\hat{a}^\dagger - \hat{a}) \tag{3.13}$$

where $Z_C = \sqrt{\frac{L}{C}}$ is the characteristic impedance.

3.2 Superconducting transmons

The transmon is a design of superconducting qubit that is adapted from the Cooper-pair boxes but less sensitive to charge noise [143]. The circuit diagram consists of a Josephson junction and a shunting capacitance, engineered so that the capacitive energy is significantly lower than the Josephson energy ($E_C \ll E_J$, the transmon regime), as shown in Figure 3.3.

Multilevel Transmon Hamiltonian

The energy of a transmon is the summation of the Josephson energy and the capacitance energy. From Equations 3.11 and 3.6, the Hamiltonian of a transmon can be

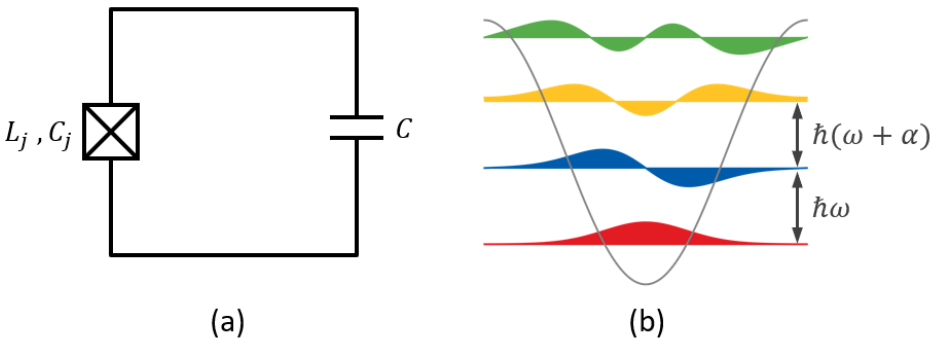


Figure 3.3: Transmon qubit. (a) The lumped element circuit diagram for a transmon qubit. A transmon qubit is composed of a Josephson junction with junction inductance L_j, junction capacitance C_j and an extra shunting capacitance C. (b) The energy landscape of a transmon qubit. Compared to a resonator, the energy gap between neighbouring levels are different with an anharmonicity $\hbar\alpha$.

written as [143, 148]

$$\mathcal{H}_q = 4E_C(\hat{n} - n_g)^2 - E_J \cos(\hat{\phi}) \tag{3.14}$$

where n_g denotes the effective offset charge of the device, measured in units of the Cooper pair charge $2e$, and it is noise on this variable that is known as charge noise. The eigenenergy of Hamiltonian 3.14 can be exactly solved as [143]:

$$E_m(n_g) = E_C a_{2[n_g + k(m,n_g)]}(-E_J/2E_C) \tag{3.15}$$

where $a_\nu(q)$ is Mathieu's characteristic value, and $k(m, n_g)$ is a function appropriately sorting the eigenvalues. Ideally, the energy would be designed to be independent of the charge noise; however, one can only suppress the dependence in practice. Here the charge dispersion ϵ_m is introduced to describe the variation of the energy level m to environmental offset charge n_g. The energy E_m can be approximated [143] in the transmon regime as:

$$E_m(n_g) = E_m(n_g = 1/4) - \frac{\epsilon_m}{2}\cos(2\pi n_g) \tag{3.16}$$

and the charge dispersion relation for the m-th level is given by

$$\Delta f_m = \frac{1}{\hbar}\epsilon_m \cos(2\pi n_g) \tag{3.17}$$

where the peak-to-peak value for the charge dispersion of the m-th energy level ϵ_m is approximated by

$$\epsilon_m \approx (-1)^m E_C \frac{2^{4m+5}}{m!}\sqrt{\frac{2}{\pi}}\left(\frac{E_J}{2E_C}\right)^{\frac{m}{2}+\frac{3}{4}} e^{-\sqrt{8\frac{E_J}{E_C}}} \tag{3.18}$$

An immediate result is ϵ_m is exponentially suppressed with the increasing value of E_J/E_C. Also, the higher transmon energy level is more sensitive to the charge noise. Here we list $|\epsilon_i/\epsilon_0|$ to have an intuitive look at the significance of charge noise at different energy levels.

$\lvert\epsilon_1/\epsilon_0\rvert$	$\lvert\epsilon_2/\epsilon_0\rvert$	$\lvert\epsilon_3/\epsilon_0\rvert$
$32\left(\frac{E_J}{2E_C}\right)^{\frac{1}{2}}$	$512\left(\frac{E_J}{2E_C}\right)$	$\frac{2048}{3}\left(\frac{E_J}{2E_C}\right)^{\frac{3}{2}}$

Since in practice $E_J/E_C \gg 1$, it can clearly be seen that operating on higher levels of the transmon device would experience greater charge noise, which we have characterised on our device and will be discussed in chapter 4.

Similar to quantising the LC circuits, the second quantisation can be applied to the transmon Hamiltonian by introducing the creation and annihilation operator $\hat{a}^\dagger$ and $\hat{a}$ as

$$\hat{n} = i n_0(\hat{a}^\dagger - \hat{a}) \tag{3.19}$$

$$\hat{\phi} = \phi_0(\hat{a}^\dagger + \hat{a}), \tag{3.20}$$

where (3.21)

$$n_0 = \left(\frac{E_J}{32E_C}\right)^{1/4} \tag{3.22}$$

$$\phi_0 = \left(\frac{2E_C}{E_J}\right)^{1/4} \tag{3.23}$$

and taking the Taylor expansion for the transmon Hamiltonian gives

$$\begin{aligned}\hat{\mathcal{H}}_q &= -4E_C n_0^2(\hat{a}^\dagger - \hat{a})^2 - E_C\left(1 - \frac{1}{2}\phi_0^2(\hat{a}^\dagger + \hat{a})^2 + \frac{1}{24}\phi_0^4(\hat{a}^\dagger + \hat{a})^4 + \dots\right) \\ &\approx \sqrt{8E_C E_J}\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) - E_C - \frac{E_C}{12}(\hat{a}^\dagger + \hat{a})^4\end{aligned} \tag{3.24}$$

The Hamiltonian can be further simplified by removing the fast-rotating terms (the terms that don't conserve the particle number, having an uneven number of a and $a^\dagger$).

$$\begin{aligned}\mathcal{H}_q &\approx \omega_0\hat{a}^\dagger\hat{a} + \frac{\alpha}{2}\left((\hat{a}^\dagger\hat{a})^2 + \hat{a}^\dagger\hat{a}\right) \\ &= \left(\omega_0 + \frac{\alpha}{2}\right)\hat{a}^\dagger\hat{a} + \frac{\alpha}{2}(\hat{a}^\dagger\hat{a})^2\end{aligned} \tag{3.25}$$

where we define the transmon base frequency $\omega_0 = \sqrt{8E_C E_J} - E_C$. The approximated Hamiltonian describes a Duffing oscillator exhibiting second-order non-linear dynamics. The quantised Duffing oscillator would have energy spacings different by the anharmonicity, approximated as $\alpha = -E_C$. Denote $\omega = \omega_0 + \alpha$, and rewrite $a^\dagger a = \sum k\,|k\rangle\langle k|$, the transmon Hamiltonian can be expressed as

$$\begin{aligned}\hat{\mathcal{H}}_q &= \omega\hat{a}^\dagger\hat{a} + \frac{\alpha}{2}\hat{a}^\dagger\hat{a}(\hat{a}^\dagger\hat{a} - 1) \\ &= \sum_k\left(\left(\omega - \frac{\alpha}{2}\right)k + \frac{\alpha}{2}k^2\right)|k\rangle\langle k| \\ &= \sum_k \omega_k|k\rangle\langle k|\end{aligned} \tag{3.26}$$

where $\omega_k = \left(\omega - \frac{\alpha}{2}\right)k + \frac{\alpha}{2}k^2$.

Single transmon gates

Operating a single transmon gate is achieved by sending a microwave drive pulse to the device. A semi-classical approach is usually used to describe microwave-atom behaviour by adding a drive term $\hat{\mathcal{H}}_d$ resulting from the presence of an external field with a time-dependent voltage $V_d(t)$ [148].

$$\hat{\mathcal{H}}_d = 2en_0 \frac{C_d}{C_\Sigma} V_d(t)(\hat{a}^\dagger - \hat{a}) \tag{3.27}$$

where C_d is the capacitance to ground and $C_\Sigma = C_d + C$. Define $\Omega = 2en_0 \frac{C_d}{C_\Sigma}$, the drive term can be simplified as

$$\hat{\mathcal{H}}_d = \Omega V(t)(\hat{a}^\dagger - \hat{a}) \tag{3.28}$$

A typical method to move forward and analyse the transmon behaviour is to move into a rotating frame at the same frequency as the drive.

$$U(t) = e^{-i(\omega_d t + \phi)\hat{a}^\dagger \hat{a}} \tag{3.29}$$

and define the new observation frame $|\psi\rangle^R = U(t)\,|\psi(t)\rangle$ which is static when it's idle. The Hamiltonian in this frame can be evaluated by taking the derivative of the evolution, given as [148]:

$$\hat{\mathcal{H}}^R = U^\dagger \hat{\mathcal{H}} U - iU^\dagger \frac{\partial U}{\partial t} \tag{3.30}$$

which gives the transmon term

$$\hat{\mathcal{H}}^R = \hat{\mathcal{H}}_q^R + \hat{\mathcal{H}}_d^R \tag{3.31}$$

where

$$\hat{\mathcal{H}}_q^R = \hbar[-\Delta \hat{a}^\dagger \hat{a} + \frac{\alpha}{2} \hat{a}^\dagger \hat{a}(\hat{a}^\dagger \hat{a} - 1)] \tag{3.32}$$

$$\hat{\mathcal{H}}_d^R = \hbar\omega(t)(e^{-i\omega_d t + \phi}\hat{a} + e^{i\omega_d t + \phi}\hat{a}^\dagger) \tag{3.33}$$

Now consider the subspace of $\{|k-1\rangle, |k\rangle\}$. Note that

$$\langle k-1|\,\hat{\mathcal{H}}\,|k\rangle = \hbar\sqrt{k}V(t)e^{i\omega_d t + \phi} \tag{3.34}$$

$$\langle k|\,\hat{\mathcal{H}}\,|k-1\rangle = \hbar\sqrt{k}V(t)e^{-i\omega_d t+\phi} \tag{3.35}$$

For simplicity, we will use notation $X=\sigma_x, Y=\sigma_y, Z=\sigma_z$ and $ZX=\sigma_z\otimes\sigma_x$ etc. The Hamiltonian $\hat{\mathcal{H}}$ in the subspace can be written as:

$$\hat{\mathcal{H}} = -\frac{1}{2}\left(\Omega_I(t)\sigma_x + \Omega_Q(t)\sigma_y\right) \tag{3.36}$$

where $\Omega_I(t)$ and $\Omega_Q(t)$ denote the I (in-phase, $\phi=0$) component and the Q (out-of-phase, $\phi=\pi/2$) component of the pulse $V(t)$ [148].

Now we define the in-phase drive as rotating the X axis on the Bloch-sphere, and then the out-of-phase is naturally rotating the Y axis. Then we put this result back into the Hamiltonian in the rotating frame, we have:

$$\frac{1}{\hbar}\hat{\mathcal{H}}_k^R = \frac{\sqrt{k}\Omega_I}{2}\hat{X} + \frac{\sqrt{k}\Omega_Q}{2}\hat{Y} + \frac{\omega-\omega_{k,k-1}}{2}\hat{Z} \tag{3.37}$$

where $\omega_{k,k-1}=\omega_k-\omega_{k-1}$ is the qubit transition frequency between the $|k-1\rangle$ and $|k\rangle$ state. The analysis above provides a way to operate a single-photon transition gate on a multi-level transmon system. Compared to the qubit system, the major difference is that the required amplitude to drive a full π pulse is lower and scaled as $1/\sqrt{k}$ for higher-level transitions.

Virtual Z gate The physical drive implements the X or Y rotation in a qubit-like subspace. The Z rotation can be implemented virtually by shifting the phase of all gates in the rest of the sequence [180]. The implementation of the Z rotation for a multi-level system is an extension of the qubit virtual Z gate. Suppose we would like to implement a Z gate in the following gate sequence:

$$U = G_n G_{n-1}...G_k Z_m(\theta) G_{k-1}...G_0 \tag{3.38}$$

where G_k denotes a gate. Now insert an identity gate sequence $Z_m(\theta)Z_m^{-1}(\theta)$ between all the following gates, and we get

$$U = Z_m(\theta)Z_m^{-1}(\theta)G_nZ_m(\theta)Z_m^{-1}(\theta)...Z_m(\theta)Z_m^{-1}(\theta)G_kZ_m(\theta)G_{k-1}...G_0 \tag{3.39}$$

which is equivalent to the original gate sequence. Now we rewrite the $G'_k = Z_m^{-1}(\theta)G_kZ_m(\theta)$ and get a new sequence:

$$U = Z_m(\theta)G'_nG'_{n-1}...G'_kG_{k-1}...G_0 \tag{3.40}$$

where G'_k is implemented by shifting the phase of driving pulses. The last gate $Z_m^{-1}(\theta)$ will not make any difference if the state is measured with an operator that commutes with $Z_m^{-1}(\theta)$, which is always the case for the standard dispersive readout. For a four-level qudit system, the virtual Z gate can be implemented as shown in Table 3.1.

	$\{\|0\rangle, \|1\rangle\}$ subspace	$\{\|1\rangle, \|2\rangle\}$ subspace	$\{\|2\rangle, \|3\rangle\}$ subspace
$Z^1(\theta)$	$-\theta$	θ	0
$Z^2(\theta)$	0	$-\theta$	θ
$Z^3(\theta)$	0	0	$-\theta$

Table 3.1: Shifting the drive pulse sequence to implement multilevel transmon virtual Z gate

Decomposition of single qudit gates Given by Equation 3.37, the X and Y rotation can be implemented with the same driving pulse with a different initial phase. Our setup chooses the X rotation to be implemented with $\phi = 0$ and the Y rotation to be implemented with $\phi = \pi/2$. The qubit Hadamard gate H is implemented with the following sequence:

$$H = Z(\pi)Y(\pi/2) \tag{3.41}$$

where $Z(\pi)$ is implemented by the virtual Z gate.

The qutrit Hadamard gate H is implemented with the following sequence of gates [186]:

$$H = H_{12}Z_1(\pi)Z_2(\pi)Y_{01}(\theta_{\mathrm{m}})H_{12}, \tag{3.42}$$

where $\theta_{\mathrm{m}} = 2\arccos(1/\sqrt{3})$ is the magic angle [37], and $Y_{01}(\theta_{\mathrm{m}})$ is implemented with a single pulse modulating the amplitude. The virtual gates $Z_1(\theta)$ and $Z_2(\theta)$ add phases on the $|1\rangle$ and $|2\rangle$ states, respectively. H_{12} is the Hadamard gate in the $\{|1\rangle, |2\rangle\}$ subspace, which is synthesised by $H_{12} = Y_{12}(\pi/2)Z_2(\pi)$, where $Y_{ij}(\theta)$ is the Pauli Y rotation with rotation angle θ applied to the $\{|i\rangle, |j\rangle\}$ subspace.

3.3 Transmon Resonator Coupling

While the transmon stores quantum information and provides manipulability for quantum control, the information still needs to be measured and extracted into classical information. A standard approach to implementing such measurement is with a quantum harmonic oscillator coupled to the transmon as a "probe" of the transmon state. Instead of measuring the transmon directly, the resonator exhibits a frequency shift which depends on the state of the transmon [274]. Therefore, the transmon state can be extracted by observing the transmission or reflection signal from the coupled resonator.

On the other hand, measured quantum states would "collapse" and lose some information. The probe needs to be carefully engineered to minimise its interaction with the transmon so that it is not detrimental to the information while the transmon is not being measured. The *dispersive regime* [274], where the detuning between the qubit transition frequency and the resonator frequency is much larger than the interaction strength, is where the coupling resonator is engineered for circuit QED. In the dispersive regime, there is no longer a direct exchange of energy between the two systems. Instead, the qubit and resonator shift each other's frequencies.

In this section, we discuss the theory of transmon-resonator coupling and dispersive readout for multilevel transmon systems.

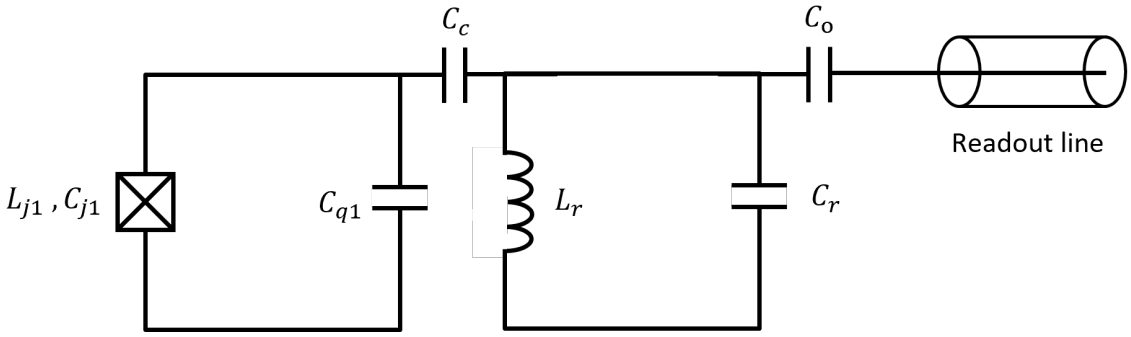


Figure 3.4: Lumped element circuit diagram describing transmon-resonator coupling. The circuit is composed of a transmon qubit with junction inductance and capacitance L_j,C_j and qubit shunting capacitance C_q, and a resonator with a linear inductance L_r and a capacitance C_r. The two components are coupled via a capacitance C_g. The resonator is then coupled to a readout line with capacitance C_o.

The Transmon-Resonator Hamiltonian

The circuit QED community has developed a theory to analyse the qubit-resonator Hamiltonian utilising the dispersive approximation. The resonator is capacitively coupled to the transmon qubit. We can look at the Jaynes-Cummings Hamiltonian in the dispersive limit, which can describe the Hamiltonian of the qubit-resonator system [22, 23]:

$$\mathcal{H} = \hbar\omega_r a^\dagger a + \sum_{k=0}^{n} \hbar\omega_k \left|k\right\rangle \left\langle k\right| + \sum_{k=1}^{n-1} \hbar\chi_{k-1} \left|k\right\rangle \left\langle k\right| + \sum_{j=0}^{n-1} \hbar s_j \left|j\right\rangle \left\langle j\right| a^\dagger a \tag{3.43}$$

where $\chi_j = g_j^2/\Delta_j$ and g_j is the coupling strength to the transmon transition j and $j+1$, Δ_j is the detuning between the resonator frequency and the j and $j+1$ transition frequency. The resonator exhibits a dispersive shift $s_j = -(\chi_j - \chi_{j-1})$ from interaction with the transmon in state $|j\rangle$. Here we define $\chi = \chi_1 = -s_1$ to keep our notation the same as the conventions in most of the transmon literature. With a similar argument, the transmon frequency also picks up a "Lamb shift" dependent on the photon number within the resonator.

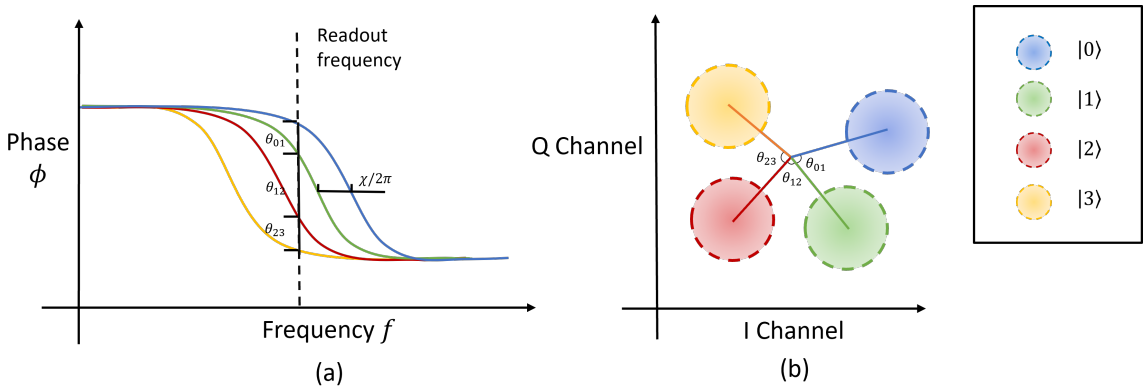


Figure 3.5: (a) Spectroscopy phase response of a resonator coupled to a qubit. The resonator oscillation frequency shifts depending on the state of the qubit, giving different responses in different colours. For reading 4 levels simultaneously, a readout signal is sent to the resonator with a frequency that maximally distinguishes the phase response of all 4 different states. (b) The scatter plot of the I channel signal and Q channel signal, often referred to as the "IQ plane". The integrated sum of the demodulated signal over time gives a point on the IQ plane, where four different states correspond to four different regions on the IQ plane, whose size is determined by the measurement noise. The separation angle θ_{ij} between region i and j is related to the frequency shift shown in (a).

Dispersive Readout

The above analysis shows that the coupled resonator resonance frequency is dependent on the transmon state. Therefore, the transmon state can be inferred by probing the resonator's frequency. Since the dispersive shift is usually small, it is easier to distinguish the transmon state by the phase of the probing signal. A typical reflection response from a resonator is shown in Figure 3.5. The resonator phase response is shifted in the left figure; by choosing a centre frequency that can distinguish more signals, a single pulse could distinguish multiple states. Usually, the device is designed to have $\chi = \kappa$ so that a readout frequency can be chosen to make $\theta_{01} = \pi$ to maximise the readout fidelity. For reading out 4 different states simultaneously, we would like to have $\theta_{01} \approx \theta_{12} \approx \theta_{23} \approx \pi/2$, which can be approximately configured by setting $\chi = \kappa/2$.

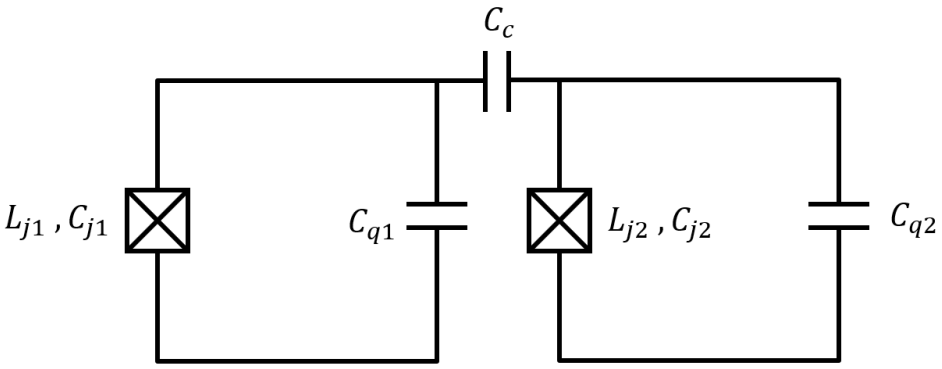


Figure 3.6: Lumped element circuit diagram describing Transmon-transmon coupling. The circuit is composed of two transmon qubits with junction inductance and capacitance L_{ji}, C_{ji} and qubit shunting capacitance C_{qi} where i denotes the index of the qubit. Two qubits are coupled via a capacitance C_c.

3.4 Transmon-transmon coupling and two-qubit gates

Interactions between two transmons are required on a quantum processor to implement universal quantum computation. Various coupling mechanisms allow the interaction between two transmons, and the easiest way is to introduce a capacitance between two transmons, as shown in Figure 3.6. Other proposals include using an inductance to couple two qubits [68, 225] and carefully designing the qubits to generate a longitudinal coupling [285]. Additionally, the coupling can be mediated by a coupler [195, 272] or a quantum bus that couples to multiple qubits [166, 118]. In this section, we focus on the capacitive coupling between two fixed-frequency transmons and implement entanglement with an all-microwave activated interaction, cross-resonance [238, 168, 55, 226]. The following discussion treats the transmon only as a two-level system, while theoretical analysis and experimental demonstration of multi-level transmon entangling gates have been demonstrated in Ref. [100].

Transmon-transmon Hamiltonian

The two capacitively coupled transmons are modelled with the following Hamiltonian:

$$\frac{1}{\hbar}\hat{\mathcal{H}} = \left(\omega_{01}^{(1)} + \frac{\alpha^{(1)}}{2}(\hat{a}^\dagger\hat{a} - 1)\right)\hat{a}^\dagger\hat{a} + \left(\omega_{01}^{(2)} + \frac{\alpha^{(2)}}{2}(\hat{b}^\dagger\hat{b} - 1)\right)\hat{b}^\dagger\hat{b} + J(\hat{a}^\dagger\hat{b} + \hat{b}^\dagger\hat{a}) \quad (3.44)$$

where J is the capacitance coupling strength. An effective Hamiltonian in the rotating frame is given by:

$$\frac{1}{\hbar}\hat{\mathcal{H}} = \left(\omega_{01}^{(1)} + \frac{\alpha^{(1)}}{2}(\hat{a}^\dagger\hat{a} - 1)\right)\hat{a}^\dagger\hat{a} + \left(\omega_{01}^{(2)} + \frac{\alpha^{(2)}}{2}(\hat{b}^\dagger\hat{b} - 1)\right)\hat{b}^\dagger\hat{b} + 2\Omega_{ZZ}\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b} \quad (3.45)$$

where the frequency of each qubit will be slightly modified, $\tilde{\omega}_{01}^{(1)} = \omega_{01}^{(1)} + \frac{J^2}{\Delta_{12}} + \Omega_{ZZ}$ and $\tilde{\omega}_{01}^{(2)} = \omega_{01}^{(2)} - \frac{J^2}{\Delta} + \Omega_{ZZ}$. The coefficient of the coupling term is then given by:

$$\Omega_{ZZ} = \frac{J^2(\alpha^{(1)} + \alpha^{(2)})}{(\Delta + \alpha^{(1)})(\Delta_{12} - \alpha^{(2)})} \quad (3.46)$$

α is the anharmonicity of the qubit, $\Delta = \omega^{(1)} - \omega^{(2)}$ is the frequency difference between the two qubits.

Cross-resonance gate

The cross-resonance gate is one of the two-qubit gates for fixed-frequency transmons [238, 168, 55, 226]. The cross-resonance interaction is induced by driving one qubit (the control qubit) at the frequency of the other qubit (the target qubit). While driving the control qubit, a cancellation pulse is usually sent to the target qubit to compensate for the crosstalk between the two qubits. This section outlines how the cross-resonance gate works based on reference [226] where a more detailed description can be found.

Driving cross-resonance For simplicity, we will use notation $X = \sigma_x, Y = \sigma_y, Z = \sigma_z$ and $ZX = \sigma_z \otimes \sigma_x$ etc. When an external electric field is applied to the system, an extra term should be added to the Hamiltonian:

$$\hat{\mathcal{H}}_d = \frac{1}{\hbar}\Omega\cos((\omega_{01}^{(2)} + \delta)t)(\hat{a} + \hat{a}^\dagger) \quad (3.47)$$

where Ω describes the Rabi frequency in the presence of the drive. Note that this term is still in the initial frame instead of the rotating frame, and needs to be moved to the rotating frame with the Schrieffer-Wolff transformation with a unitary $\hat{U} = e^{-i\hat{S}}$ given by:

$$\mathcal{H} = e^{iS}\mathcal{H}e^{-iS} \tag{3.48}$$

where

$$\hat{U} = \exp\left(-(\omega_{01}^{(2)} + \delta)(\hat{a}^\dagger\hat{a} + \hat{b}^\dagger\hat{b})t\right) \tag{3.49}$$

and S can be evaluated with perturbation theory. This results in a Hamiltonian [125]:

$$\frac{1}{\hbar}\hat{\mathcal{H}}^R = \left(\Delta - \delta + \frac{\alpha^{(1)}}{2}(\hat{a}^\dagger\hat{a} - 1)\right)\hat{a}^\dagger\hat{a} + \left(-\delta + \frac{\alpha^{(2)}}{2}(\hat{b}^\dagger\hat{b} - 1)\right)\hat{b}^\dagger\hat{b} \tag{3.50}$$

where $\Delta = \omega^{(1)} - \omega^{(2)}$. With the approximation $e^{iS} \approx I + iS$, the effective Hamiltonian becomes:

$$\frac{1}{\hbar}\hat{\mathcal{H}}^R \approx \frac{\Omega}{2}\left(-\frac{J}{\Delta}\frac{\alpha^{(1)}}{\Delta + \alpha^{(1)}}ZX\right) - \frac{\Omega}{2}\left(-\frac{J}{\Delta}\frac{\alpha^{(1)}}{\Delta + \alpha^{(1)}}IX + XI\right) \tag{3.51}$$

which gives the ZX term for generating the entanglement. In practice, the driving phase determines the ratio of coefficient ZX between ZY terms, and the exact coefficient of each term is experimentally measured. The effective Hamiltonian is given by:

$$\hat{\mathcal{H}}^R = \mu_{ZZ}ZZ + \mu_{ZX}ZX + \mu_{ZY}ZY + \nu_{IX}IX + \nu_{IY}IY + \nu_{IZ}IZ \tag{3.52}$$

where μ_{ZZ}, μ_{ZX}, and μ_{ZY} are the cross-resonance term coefficients and the ν_{IZ}, ν_{IX}, and ν_{IY} are the crosstalk term coefficients.

CNOT from cross-resonance From equation 3.52, the Hamiltonian contains cross-resonance terms μ_{ZZ}, μ_{ZX}, and μ_{ZY} and crosstalk terms ν_{IZ}, ν_{IX}, and ν_{IY}. First, the initial phase of the cross-resonance drive is calibrated so that $\mu_{ZY} = 0$. Then an external drive to the target qubit is applied to cancel the effect of the ν_{IY} and ν_{IX} terms. For our implementation, the correction for μ_{ZZ} and ν_{IZ} are ignored.

By carefully choosing the pulse width and amplitude, the $ZX(-\pi/2) = \exp(-i\frac{\pi}{4}ZX)$ gate is implemented from the Hamiltonian.

The CNOT gate can be decomposed into the following sequence:

$$CNOT = X_t(-\pi/2)Z_c(-\pi/2)ZX(-\pi/2) \tag{3.53}$$

where subscript t denotes the gate acts on the target qubit and c denotes the control qubit.

3.5 Discussion

This section reviewed the circuit-QED theory for constructing the quantum system. To have an operable processor, we would like the device to have the following properties:

1. The transmon should be designed to have $E_c \ll E_J$ in order to minimise charge-noise decoherence, particularly important when using the transmon as a qudit since charge noise dispersion increases at higher energy levels.

2. The anharmonicity $\alpha \approx E_c$ needs to be engineered large enough to address each transition separately. The drive frequencies required to drive the different transitions need to be separated by more than approximately the inverse pulse length.

3. The anharmonicity needs to follow $2\alpha < 500$ MHz to use one DAC system to operate at least three transitions since the bandwidth of our control electronics is restricted to about 500 MHz.

4. The dispersive shift χ_i needs to be large enough to distinguish two states but not too large so that at least four states can be discriminated with one single-shot pulse. For an ideal design, four regions on the IQ scatter plot have a phase difference of $\pi/2$, which can be implemented by configuring $\chi = \kappa/2$.

5. For the implementation of the cross resonance gate between two transmons as qubits, the qubit-qubit coupling strength J is defined so that the cross resonance drive enables a significant ZX term but has a small enough ZZ term.

From points (1) to (3), the designed device aims to have qubit frequency 5 GHz $>$ $\omega_{01} > 4$ GHz and anharmonicity $\alpha \approx E_c = 200$ MHz, which gives $E_J/E_C > 50$ that suppresses the charge noise. The α provides a good separation between neighbouring transition levels and also requires 400 MHz bandwidth of the DAC system to drive transitions in $\{|0\rangle, |1\rangle\}$ and $\{|2\rangle, |3\rangle\}$ subspaces, which is within the operating range of our instrument.

From point (4), the resonator is designed to have a linewidth $\kappa \approx 1$ MHz. The resonator frequency is designed to be 11 GHz $> \omega_r >$ 8 GHz, and the coupling strength between the qubit and resonator $g \approx 200$ MHz, which gives a $\chi \approx 0.25$ MHz.

From point (5), we aim the two-qubit detuning Δ at the swept point to generate the theoretical lowest error rate for cross-resonance interaction, given by $\Delta = -0.3\alpha \approx 60$ MHz [168]. The qubit-qubit coupling strength aims at $J \approx 1$ MHz. The device fabrication process is described in Appendix B.

Chapter 4

Device Design and Basic Characterisation

In this chapter, I first review the coaxmon architecture used in constructing superconducting circuits, followed by experimental results showcasing the devices' basic properties. The review content in section 4.1 is summarised from publications from my research group [219, 247]. The devices mentioned in this chapter were fabricated by me with assistance from Giulio Campanaro and Simone Fasciati, and then installed in the fridge for testing. Vivek Chidambaram, Boris Shteynas, and Giulio Campanaro prepared the dilution fridge's electronics in section 4.2, while I set up room temperature electronics with Giulio's help. I wrote the experimental code and collected and analysed all data presented in sections 4.3 and 4.4.

This chapter introduces the particular circuit architecture used in this work and details the control electronic system used in the experiment. Basic characterisation of the devices used in the later experiments is then presented, in particular, a coupled two-qubit device and an uncoupled transmon device as a qudit, and we refer to these two devices as "2QB" and "1QD" devices, respectively.

4.1 The Coaxmon Architecture

The coaxmon qubit is a variant of the transmon qubit where qubits and resonators are coupled vertically to the coaxial control lines wired above and below the chip [219]. This design does not require wire bonding to connect the control lines, which simplifies the fabrication process. Also, the off-chip wiring design provides a dense arrangement of the control and readout lines compared to the 2D circuit designs that require all components to share one side of the substrate. See figure 4.1.

The coaxmon architecture builds the qubit and the resonator on the front and back of a double-sided polished silicon wafer. A fixed-frequency coaxmon qubit consists of an inner circular electrode and an outer ring-shaped electrode that is connected by a Josephson junction. The circular shape of the qubits means that they have coaxial electromagnetic field symmetry that is well-matched to the coaxial control pin. The resonator, with a similar circle-and-ring shape, is printed on the other side of the substrate, capacitively coupled to the qubit. The reflected signal from the resonator is used for the readout. The inductance of the resonator is provided by a thin wire connecting the inner and outer electrodes. The size of the coaxmon qubits is designed to be similar to a coaxial cable; the coupling strength between the qubit and resonator is typically in the range 150-250 MHz, which is relatively large compared to other circuit QED designs. The qubit and the resonator are far detuned to stay in the dispersive regime, usually with a choice of frequency ω_{01} at 4-7 GHz and ω_r at 8-11 GHz for qubits and resonators, respectively. This circuit architecture

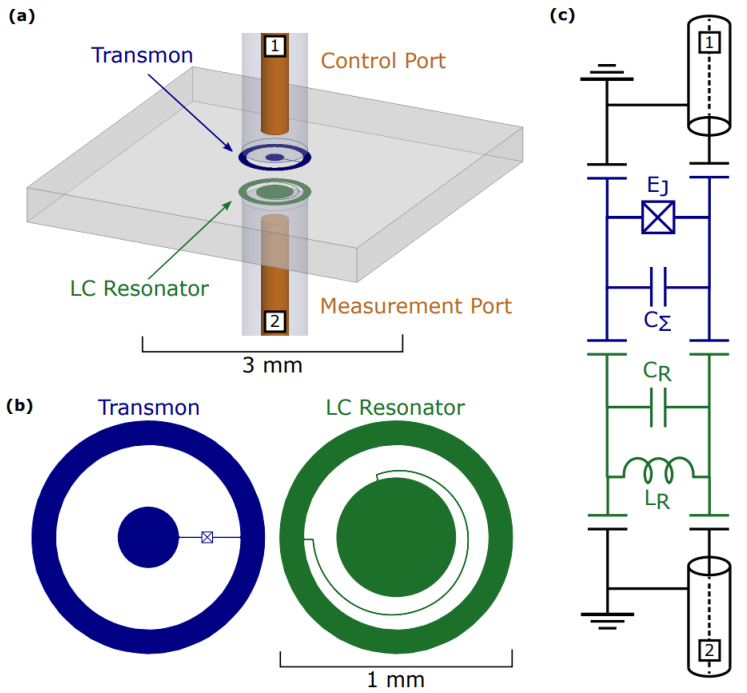


Figure 4.1: Figures are adapted from [219]. (a) The architecture diagram cartoon. The transmon is fabricated on the front of the chip, and the resonator is fabricated on the backside of the wafer. The resonator and the transmon couple to the off-plane measurement port and the control port, respectively. (b) The geometry of the transmon and the resonator. The transmon and the resonator contain centre and outer ring electrodes. A Josephson junction connects the two electrodes of the transmon, and a thin wire inductance connects the resonator. (c) The lumped element model of the circuit. The transmon and the resonator are capacitively coupled, and each of them is also capacitively coupled to the control line and the measurement line.

was first demonstrated in reference [219], a two-qubit gate based on cross-resonance was demonstrated in [208], and a high-coherence uncoupled 4-qubit device was also recently demonstrated [247].

4.2 Device Fabrication

The fabrication process of the quantum processor chip involves four stages.

First, metal deposition is carried out by processing the double-sided polished wafer in the BOE etchant to remove the surface oxide. The wafer is then immediately transferred to a Plassys MEB550S evaporator, which has a specially designed wafer holder, where it is baked overnight and aluminium deposition is carried out. Removing the oxide reduces dielectric loss from the silicon and improves relaxation time. This process is repeated for both sides of the substrate.

Second, the wafer is coated with photoresist using a non-contact spin-coater chuck, baked, and then flipped using a contact chuck for spin coating. The non-contact chuck vibrates during spinning and may not form a uniform resist profile for pattern writing, but it is sufficient for a protection layer. The coated wafer is then transferred to a mask aligner to print the large pattern, including coaxmon pads and coupling arms. After photoresist development, an aluminium wet etchant is used to remove the unwanted aluminium on the wafer. This process is repeated for the other side of the substrate, and the optical tool from the mask-aligner is used to implement alignment between the front and back sides of the wafer.

Third, the wafer is cleaned and spin-coated with the protection resist layer using a non-contact chuck. Then, PMMA is applied to the qubit side for junction fabrication, using the traditional Dolan-bridge method.

Finally, the whole wafer is coated with photoresist using a non-contact chuck and sent to our collaborator at the University of Southampton for wafer dicing. A more detailed recipe can be found in appendix B.

4.3 Experiment Setup

In this section, we describe the laboratory setup used for our experiments and how it is used to characterise the performance of the 2QB and 1QD devices. Both devices use the same cryostat and control electronics.

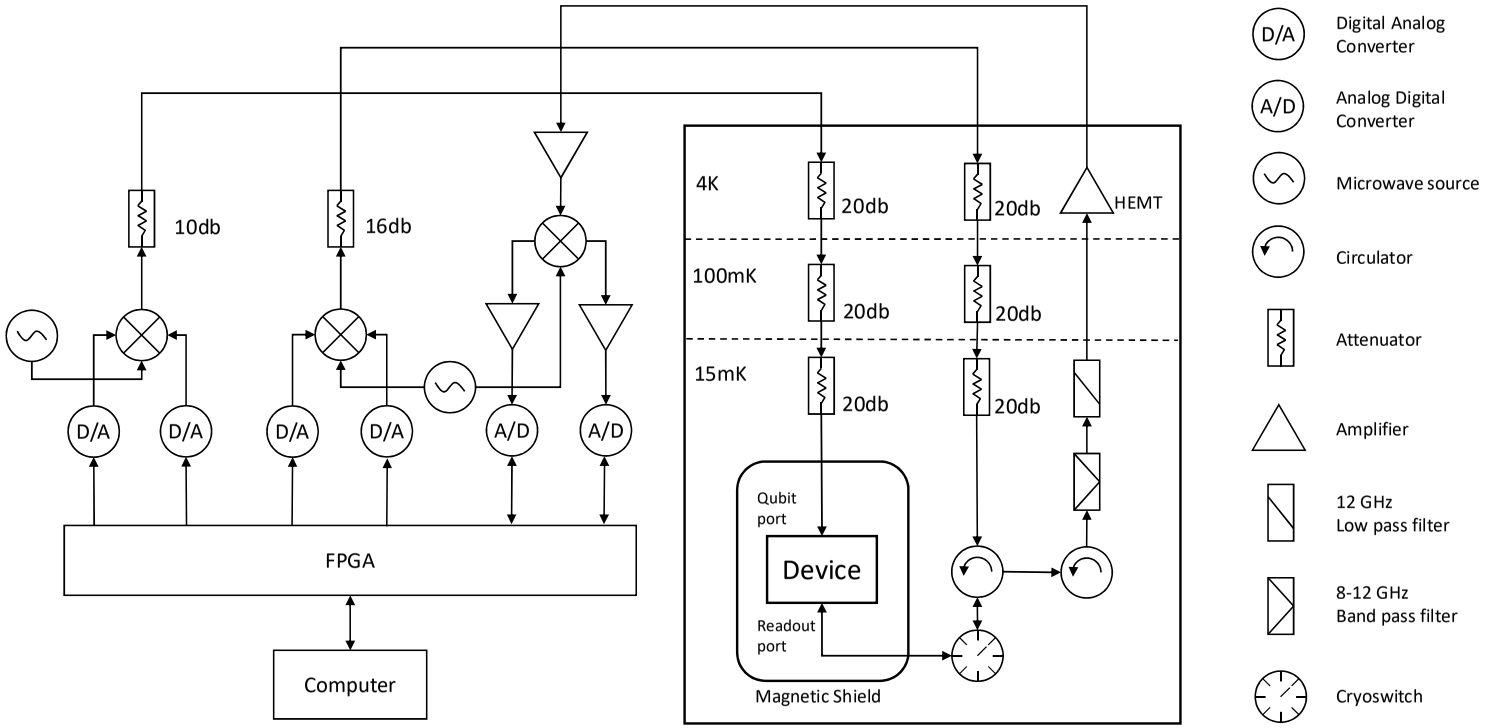


Figure 4.2: The electronics and fridge setup for the experiment in this thesis.

In order to operate the circuits in the quantum regime, the circuits need to be cooled to cryogenic temperatures in the millikelvin range such that $\hbar\omega_{01} \ll k_B T$. The sample is placed inside a Triton XL dilution refrigerator from Oxford Instruments. We then need to control the circuits using signals in the microwave frequency regime from 3 GHz to 11 GHz. Our setup uses the IQ mixing method to generate microwave pulses, which comprises a custom Radio Frequency (RF) signal Arbitrary Waveform Generator (AWG) and Analogue to Digital Conversion (ADC) system for intermediate frequency (IF) signal generation and microwave sources generating microwaves in the gigahertz frequency range. The IF signal is generated by Abaco FMC230 and Abaco FMC110 boards, mounted on an Abaco PC821 PCI-e board with a Xilinx Kintex UltraScale FPGA. The signal waveform is uploaded to the FPGA on-chip memory and then played by the DACs. The IF signal is mixed with a high-frequency local oscillator signal from a Rohde & Schwarz SMA110 microwave source and converted to qubit control pulses and resonator readout pulses. The electronics and fridge setup for the experiment is illustrated in Figure 4.2.

The generated signal from the electronics system needs to be delivered to the device inside the fridge, and the readout signal needs to be delivered out of the fridge to the ADC system. These signals are delivered with coaxial lines. The qubit control input and readout input lines are wired to a 20 dB attenuator at the 4K plate, still plate, and the cold plate before being wired to the device, giving 60 dB attenuation. The reflected readout signal goes through two circulators and is amplified by a HEMT amplifier by 30 to 35 dB. Then the readout signal is further amplified by another two-stage amplification setup at room temperature.

Once the readout signal comes back from the fridge, it is wired to the AD converter on the FMC110 board, which digitises the signal and delivers the digital signal to the FPGA board for further analysis. The FPGA board demodulates the signal on-chip and reports the signal's position on the scattering plot of the in-phase (I) channel

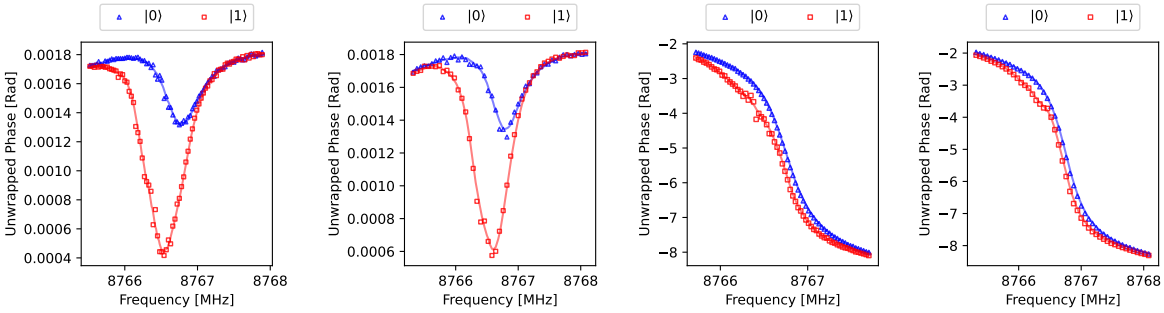


Figure 4.3: Resonator spectroscopy of the 2QB device. (a) and (b) are the magnitude response and (c) and (d) are the phase response for Q_A and Q_B, respectively. Each point is averaged from 1000 shots, and each shot has an interval of 1000 μs to wait until the qubit relaxes to the ground state.

and quadrature (Q) channel. The FPGA is programmed to implement onboard state discrimination and pulse feedback control. The details are discussed in Chapter 6.

4.4 Characterising the 2QB Device

The 2QB device chip contains two pairs of coupled qubits. For the experiments in this thesis, we selected one of these two pairs, which had superior coherence and qubit frequencies in the straddling regime, optimal for implementing the cross-resonance gate.

Readout Characterisation To measure the qubit states, we will use circuit QED dispersive readout as described in Section 3.3. First, the resonator parameters need to be characterised to implement the dispersive readout. Here we implement the resonator spectroscopy experiment that measures the reflected signal from the readout resonator. The readout signal is a 10 μs square pulse. To find the resonator response with respect to the $|0\rangle$ and $|1\rangle$ state of the qubit, a π pulse is applied to the qubit before the measurement. The experiment is performed on each resonator separately, and the parameters of the resonators are found in Table 4.1.

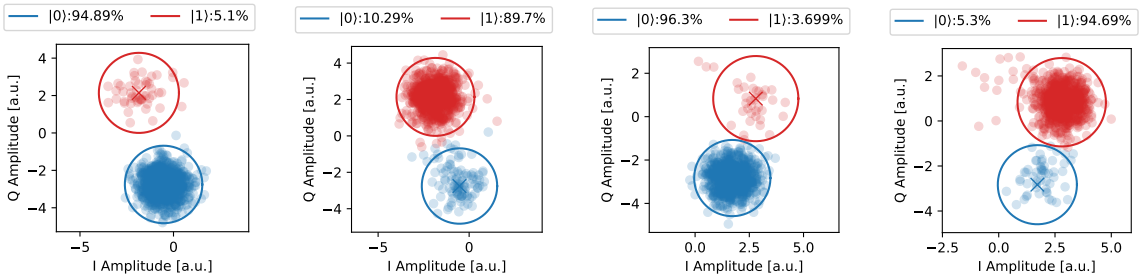
We now move on and use pulsed signals to implement single-shot readout. Single-

shot readout gives the discrete state of the system with the signal collected from a single measurement pulse. The collected data usually contains two traces: the voltage versus time for the I channel and the Q channel. As we expect, the resonator responses depend on the qubit state. We aggregate the voltage signal demodulated at the measurement frequency and convert two traces into a point on the 2D plane: the IQ plane. Then single-shot readout requires a discriminator to decide the transmon state with the measured point on the IQ plane. A more detailed discussion of the single-shot readout is presented in Chapter 6.

The signal amplitude, frequency, and width must be carefully chosen to optimise the readout fidelity. Higher signal amplitude gives a better signal-to-noise ratio (SNR), but a readout pulse that is too strong may populate too many photons into the resonator, exceeding the critical photon number, and causing non-linear effects to dominate the resonator's behaviour, which makes the process no longer a valid dispersive readout [274, 94]. The longer pulse width also increases the SNR; however, the population decay during the measurement would introduce more measurement errors. Also, the choice of frequency changes the relative position of the integrated signal distribution in the IQ plane, affecting the SNR. For this device, the single-shot readout is implemented on this device with a 3 microseconds square readout pulse. The state is discriminated by a spherical parameterised Gaussian Mixture Model (GMM) [224] with SNR given by

$$SNR = \frac{2||\mathbf{x_g} - \mathbf{x_e}||}{\sigma_g + \sigma_e} \tag{4.1}$$

where x_g and x_e are the centres of the Gaussian distribution coordinates on the IQ plane, σ_g and σ_e are the standard deviations of the ground and excited Gaussian distribution, see Figure 4.4. We find the signal-to-noise ratio for qubit A and B readout are $SNR_A = 7.22$, $SNR_B = 6.16$ respectively.



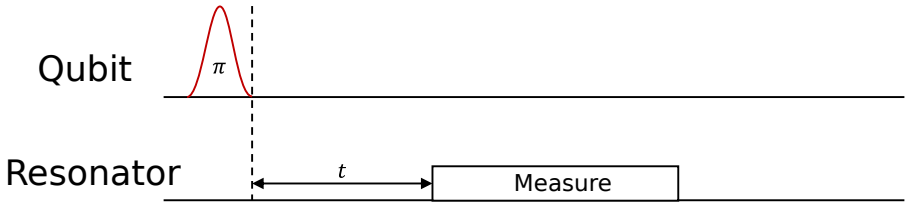
(a) (b) (c) (d)

Figure 4.4: Single-shot readout calibration for the 2QB device. (a) and (b) are the scatter plot of the demodulated signal of Q_A, while prepared at $|0\rangle$ and $|1\rangle$ state, respectively. (c) and (d) are for Q_B. The circle's radius represents the 3σ distance of the Gaussian distribution, where σ is the standard deviation of the Gaussian distribution. The spherical GMM model used in this experiment assumes the standard deviation is equivalent on both the I and Q axis and has no correlation.

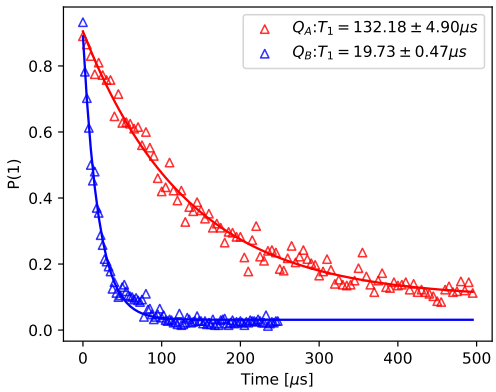
Energy Relaxation There are a variety of mechanisms that can cause energy relaxation of the excited states of qubits back towards the ground state, including radiation through the control and readout ports [116], dielectric loss in the chip substrate [215, 276, 188] as well as the dissipation from quasiparticles [45]. The stochastic process from the dissipation makes the transmon move towards a mixed-state equilibrium distribution, which follows the Boltzmann distribution. The energy relaxation time T_1 is a characteristic metric to quantify the speed of the energy relaxation process. The T_1 is defined by

$$P_{|1\rangle}(t) = P_{|1\rangle}(0)e^{-t/T_1} \tag{4.2}$$

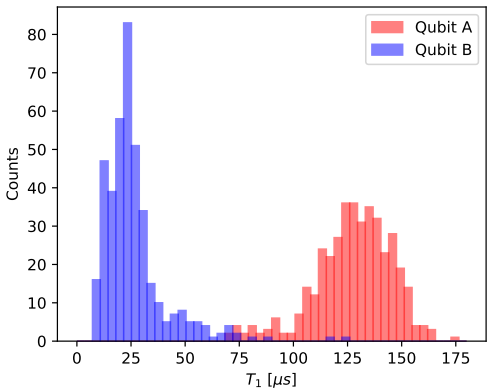
where $P_{|i\rangle}(t)$ denotes the population of state $|i\rangle$ at time t. $T_1 = 1/\gamma_1$, and γ_1 is the energy relaxation rate. To determine T_1 for the 2QB device, each qubit is excited to its $|1\rangle$ state and $P_{|1\rangle}(t)$ is measured. The experiment is repeated 400 times over a period of 13 hours to obtain the statistics. The results are shown in Figure 4.5. The T_1 times extracted for qubits A and B are $T_{1A} = 127 \pm 19$ μs and $T_{1B} = 26 \pm 14$ μs respectively. I suspect that Q_B has much lower coherence because I damaged the pad of the transmon during the junction resistance measurement before mounting it to



(a)



(b)



(c)

Figure 4.5: Relaxation time of the 2 qubit device. (a) is the pulse scheme of the experiment, where t is the delay time between the state preparation pulse and the measurement. (b) is a fitting example. Each T_1 time measurement samples 100 points with 5 μs interval for qubit A and 2.5 μs interval for qubit B. (c) is the distribution of the fitted T_1 time.

the fridge.

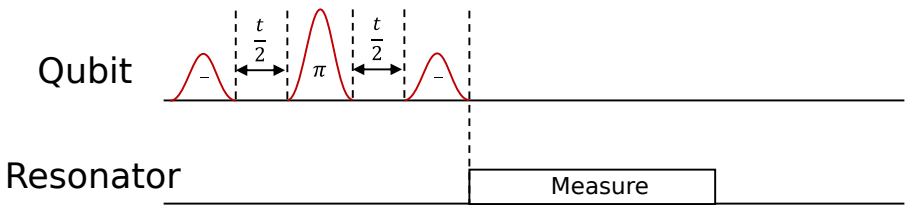
Dephasing The interaction with the environment may perturb the transmon energy levels, resulting in a stochastic Z interaction from noise in transmon frequencies. The information stored in the qubit phase would be affected by this noise and eventually lost. This decoherence mechanism is called dephasing [25].

The dephasing rate $\gamma_\phi = 1/T_\phi$ is characterised by pure dephasing time T_ϕ. The energy relaxation would also contribute to the dephasing process, which gives the total dephasing rate

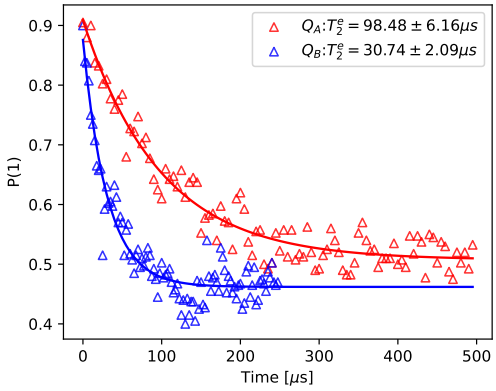
$$\gamma_2 = \frac{1}{T_2} = \frac{\gamma_1}{2} + \gamma_\phi \tag{4.3}$$

Experimentally, the T_2 value can be measured with the Ramsey interferometry experiment or the Hahn spin echo experiment [107]. The experiment pulse sequence for implementing Hahn spin echo is shown in 4.6 (a). A π pulse intermediates two $\pi/2$ pulses, and the interval between the π pulse and each $\pi/2$ pulse is the same. Consider a constant Z rotation is applied to the qubits caused by the constant frequency detuning, discussed in Equation 3.37. After the $\pi/2$ pulse, the state is prepared to a superposition and represented by a point on the equator of the Bloch sphere. As the state undergoes constant rotation around the Z-axis on the Bloch sphere, it accumulates phase during a period of free precession. A π pulse is then applied, inverting the qubit's phase. As the qubit continues to precess, the accumulated phase before and after the π pulse effectively cancels each other. Therefore, the Hahn spin echo experiments are robust to a constant frequency detuning during the experiment. This detuning could be caused by low-frequency noise, and the effect of the constant frequency detuning can be cancelled by introducing a similar echo scheme into the gate sequence [110, 267]. Therefore, the T_2 value extracted from the Hahn spin echo is the most useful metric and the reported T_2 values are extracted by Hahn spin echo experiments. The spin echo experiment is performed 400 times to capture the distribution of the T_2 time for both qubits, and the experiment result is shown in qubit 4.6 (b) and (c). The T_2 time from the spin echo experiment for qubit A is found to be $T_{2A}^{(e)} = 103 \pm 13$ μs and $T_{2B}^{(e)} = 38 \pm 17$ μs. Qubit B exhibits much lower T_1 and T_2 times, which I suspect may be due to damage caused to the qubit pads during the room temperature junction resistance measurement. We accidentally scratched the surface of Qubit B with the measurement pin during the process.

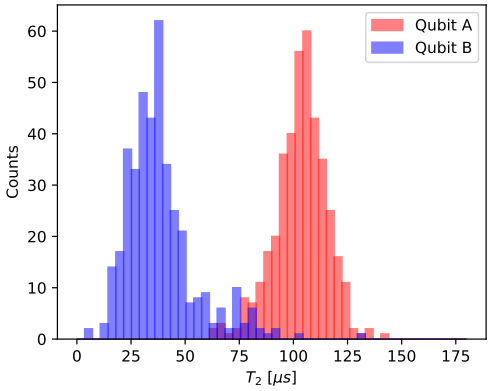
Coupling Characterisation The last parameter that needs to be determined is the qubit-qubit coupling strength. The Ramsey interferometry experiment is performed on each qubit, detuned with 0.05 MHz, with and without the other qubit



(a)

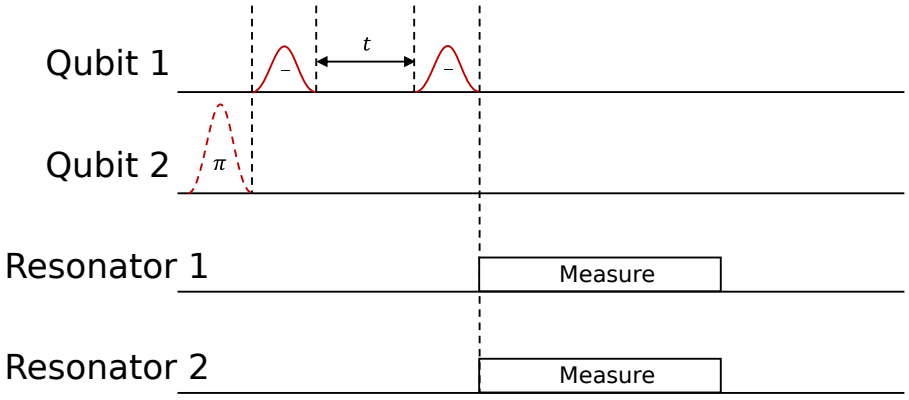


(b)

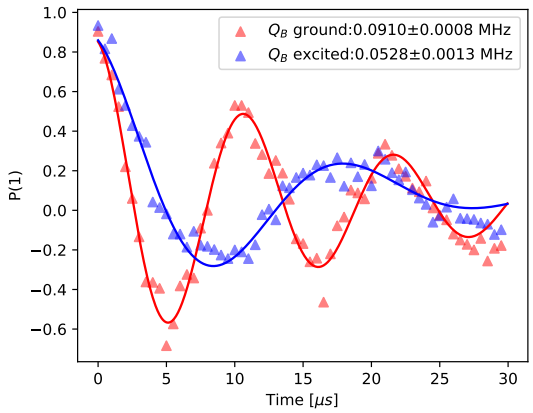


(c)

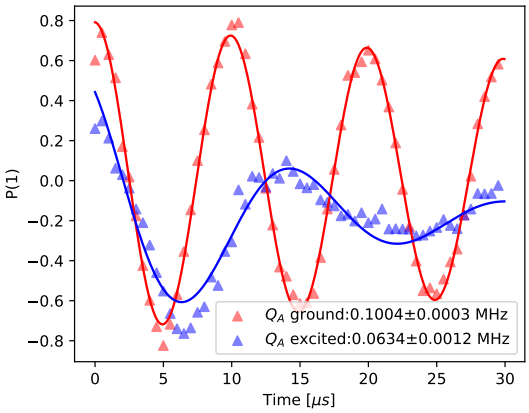
Figure 4.6: Measurement of the energy relaxation times (T_1) on the 2QB device. (a) Pulse scheme for the Hahn spin echo experiment. (b) example data set and fit curve for the experiment, and (c) is the distribution of extracted $T_2^{(e)}$ value.



(a)



(b)



(c)

Figure 4.7: Characterising the qubit coupling strength. (a) is the pulse scheme for the state-dependent Ramsey interferometry experiment. The pulse in dashed lines denotes two experiments performed, with and without the pulse. (b) and (c) is the Ramsey interferometry experiment result measuring Q_A and Q_B respectively, with and without the other qubit excited.

excited. The pulse scheme and the experimental result are shown in 4.7. From the Hamiltonian described in 3.46, the frequency difference between the two Ramsey interferometry experiments with and without the other qubit excited gives $2\Omega_{ZZ}$. The J value can be approximated from Equation 3.46, which gives $J \approx 1.31$ MHz from experimental results.

4.5 Characterising the 1QD Device

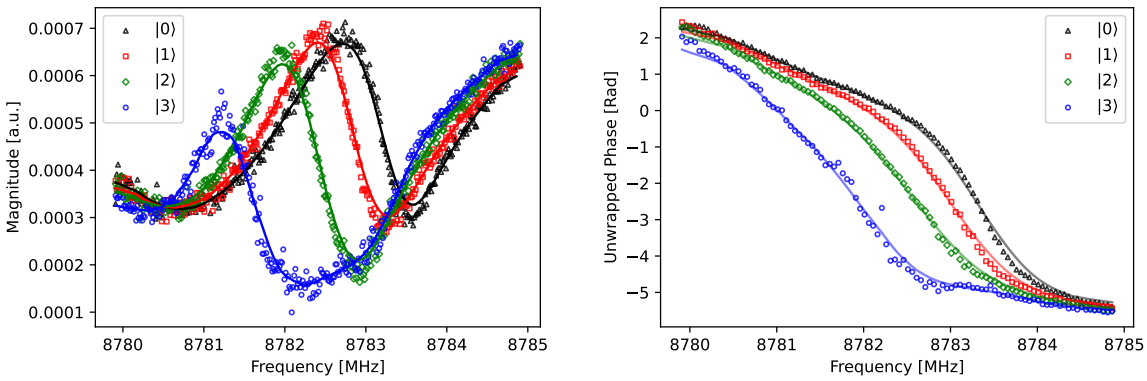
A second device, "1QD," that will be used in experiments later in this thesis is a single transmon device on a 4-qubit chip without coupling to other transmons. The device is mounted to a readout multiplexer. The readout multiplexer is a cavity that couples to all four resonators on the chip, and therefore all qubits can be read by the same readout line with frequency multiplexing.

Readout Characterisation Using dispersive readout to distinguish higher levels of the transmon has been experimentally demonstrated in [212, 79]. In the spectroscopy experiment, the resonator exhibits a dispersive shift $s_j = -(\chi_j - \chi_{j-1})$ from interaction with the transmon in state $|j\rangle$, where $\chi_j = g_j^2/\Delta_j$ and g_j is the coupling strength to the transmon transition $|j\rangle \leftrightarrow |j+1\rangle$ and Δ_j is the detuning between the resonator frequency and the $|j\rangle \leftrightarrow |j+1\rangle$ transition frequency. See discussions in section 3.37 for more details.

A 10 μs square pulse is used for the readout. Each point is averaged from 1000 shots, and between each shot a delay of 1000 μs is used to let the qubit relax to the ground state. The result of the resonator spectroscopy experiment is shown in figure 4.8.

The single-shot state discrimination on this device is implemented with a 7 μs square pulse. The reflected signal depends on the transmon state, and by carefully choosing the readout frequency, four states can be distinguished with only one measurement. Just as with the experiment for 2QB, the discriminator for different states is now a 4-component GMM model; see figure 4.9. The SNR ratio between each pair of states is calculated with formula 4.1 and listed here:

$\|0\rangle \leftrightarrow \|1\rangle$	$\|0\rangle \leftrightarrow \|2\rangle$	$\|0\rangle \leftrightarrow \|3\rangle$	$\|1\rangle \leftrightarrow \|2\rangle$	$\|1\rangle \leftrightarrow \|3\rangle$	$\|2\rangle \leftrightarrow \|3\rangle$
6.63	12.31	10.74	9.31	11.59	6.50



(a) (b)

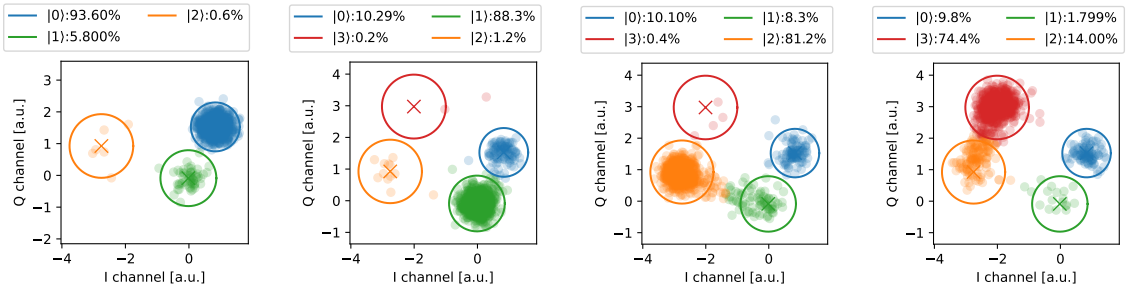
Figure 4.8: Resonator spectroscopy of the 1QD device. (a) is the magnitude response when the transmon is at a different state, and (b) is the phase response. Each point is averaged from 1000 shots, and each shot has an interval of 1000 μs to wait until the transmon relaxes to the ground state. The solid line is the moving average of the experimental data.

Energy Relaxation Compared to the two-level system, the decay dynamics are now modelled by a decay matrix Γ, which captures the decay rate between all energy levels [212]. The state population versus the waiting time is given by

$$\vec{P}(t) = (I - \Gamma^{\top})^{t}\vec{P}_0 \tag{4.4}$$

where $P(t)$ is the population of each state observed at time t, corresponding to the initial state population $\vec{P}_0$. The Γ matrix is assumed to have only diagonal elements, and the Γ_{10}, Γ_{21}, and Γ_{32} are non-zero. Then the qudit is prepared to $|1\rangle$, $|2\rangle$, $|3\rangle$ states and its relaxation is observed. An example of a fitted plot is shown in figure 4.10 (a). The distribution is given in figure 4.10 (b). The effective $T_1^{(ij)} = 1/\Gamma_{ji}$ is then defined from the Γ matrix, where i and j are the neighbouring energy levels. The distribution gives $T_1^{(01)} = 189 \pm 41$ μs, $T_1^{(12)} = 119 \pm 21$ μs, $T_1^{(23)} = 87 \pm 23$ μs.

Dephasing Similar to the two-level system, the spin echo experiment is applied to each neighbouring transition to give a T_2 describing the phase coherence of that transition. To remove the effect of energy relaxation, the normalized survival population



(a) (b) (c) (d)

Figure 4.9: Single-shot measurement result of the 1QD device and the calibrated GMM model for state discrimination. (a) to (d) are the scatter plots of the single-shot measurement result after preparation to the $|0\rangle$ to $|3\rangle$ states, respectively. The discriminator is the same model as the 2QB device; however, we now identify 4 separate Gaussian distributions corresponding to the 4 different levels. The percentage value in the legend gives the population of each state of the distribution shown in the figure.

$P(|i\rangle)_N$ is defined as follows:

$$P(|i\rangle)_N = \frac{P(i)}{P(i) + P(i-1)} \tag{4.5}$$

where $P(i)$ is the population ratio measured in state $|i\rangle$ and $i > 0$. The normalized survival population is then independent of the energy relaxation decay. Then we fit $P(|i\rangle)_N$ to

$$P(|i\rangle)_N(t) = e^{-t/T_2^{(i)}} P(|i\rangle)_N(0) \tag{4.6}$$

where $T_2^{(i)}$ is the effective T_2 value for the $|i\rangle, |i-1\rangle$ subspace. The experimental result is shown in figure 4.11. From the experiment, the characterised parameters are $T_2^{(01)} = 118 \pm 21$ μs, $T_2^{(12)} = 76 \pm 27$ μs, $T_2^{(23)} = 35 \pm 14$ μs.

Charge Noise From the theory presented in chapter 3, we expect that the higher energy levels of a transmon are more sensitive to charge noise. For a 4-level system working as a qudit, the error on the higher levels induced by the charge noise can be detrimental to executing quantum algorithms. We implement a Ramsey interferome-

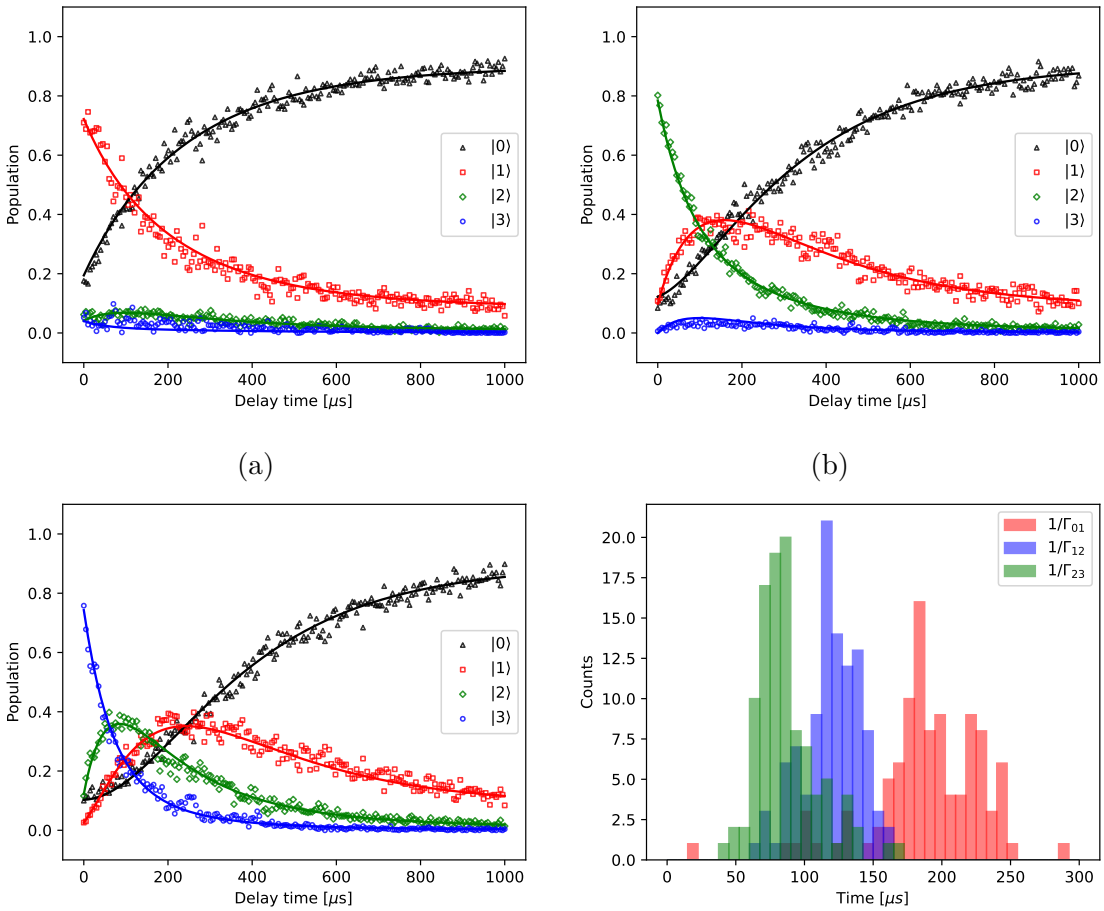


Figure 4.10: Characterisation of the T_1 coherence time of the 1QD device for different levels. (a) to (c) shows the population versus the delay time before measurement when the qudit is prepared to $|1\rangle$ to $|3\rangle$ states, respectively. (d) shows the distribution of the $1/\Gamma_{ij}$ of different levels.

try experiment to characterise the charge sensitivity of the transmon at higher levels [255, 171].

In this experiment, we focus on the characterisation of the charge sensitivity of the $\{|2\rangle, |3\rangle\}$ subspace, for it would be the most sensitive subspace among all three subspaces containing neighbouring levels of the four lowest transmon levels. The pulse sequence of Ramsey interferometry is shown in figure 4.12 (a). The Ramsey interferometry experiment takes 100 points per trace. The expectation value is estimated with 1000 shots per point. The maximum time interval t is 100 μs, which

gives a spectral resolution of approximately 2.5 kHz. The charge noise coupled to the transmon results in random individual Ramsey curves and, when superimposed, yields a sinusoidal beating pattern, which is observed in figure 4.12 (b). The two sets of extracted qudit transition frequencies correspond to the two charge-parity states (See discussions about equation 3.17).

The Ramsey interferometry experiments were repeated, running for 90 minutes. The spectrum versus the experiment time is shown in figure 4.12 (c). We found the frequency shift due to charge noise is approximately 20 kHz and observed the frequency drift over time. The frequency shift is significantly lower than the single qudit pulse Rabi rate (10 MHz for a 50 ns long π pulse), which won't cause a significant error when implementing quantum algorithms.

4.6 Discussion

We have described the characterization of transmon devices, focusing on the two devices that will be used later in this thesis. The 2QB device parameters are listed in Table 4.1. The 1QD device parameters are listed in Table 4.2. The 1QD device has long T_1 and T_2 times, 189$\pm$40 μs and 118$\pm$22 μs, respectively. The higher levels suffer from larger noise, and the $\{|2\rangle, |3\rangle\}$ subspace of the 1QD device has the lowest T_1 and T_2 times, 87 μs and 35 μs, respectively, which is still a few hundred times the gate duration. The single-shot readout can discriminate four states simultaneously. The significantly larger T_1 and T_2 times allow executing near-term quantum algorithms and implementing a two-qubit emulator on the 1QD device. The frequency shift due to the charge noise of the $\{|2\rangle, |3\rangle\}$ subspace is approximately 20 kHz, which is significantly lower than the single qudit gate Rabi rate and therefore won't contribute significant error for implementing algorithms on the 1QD device.

Parameter	Q_A	Q_B
f_{01} (MHz)	4144.41	4211.28
f_{Res} (MHz)	8766.72	8716.9
χ (MHz)	0.05	0.15
g (MHz)	104.56	174.19
κ (MHz)	0.276	0.433
T_1 (μs)	127 ± 19	26 ± 14
$T_2^{(e)}$ (μs)	103 ± 13	38 ± 17

Table 4.1: Basic parameters of the 2QB device.

Parameter	Value
f_{01} (MHz)	4134.33
f_{12} (MHz)	3937.66
f_{23} (MHz)	3721.58
f_{Res} (MHz)	8782.41
χ (MHz)	0.288
g (MHz)	177.08
κ (MHz)	0.524
$1/\Gamma_{01}$ (us)	189 ± 40
$1/\Gamma_{12}$ (us)	120 ± 21
$1/\Gamma_{23}$ (us)	87 ± 23
T_2 echo 01 (us)	118 ± 22
T_2 echo 12 (us)	76 ± 27
T_2 echo 23 (us)	35 ± 14

Table 4.2: Basic parameters of the 1QD device.

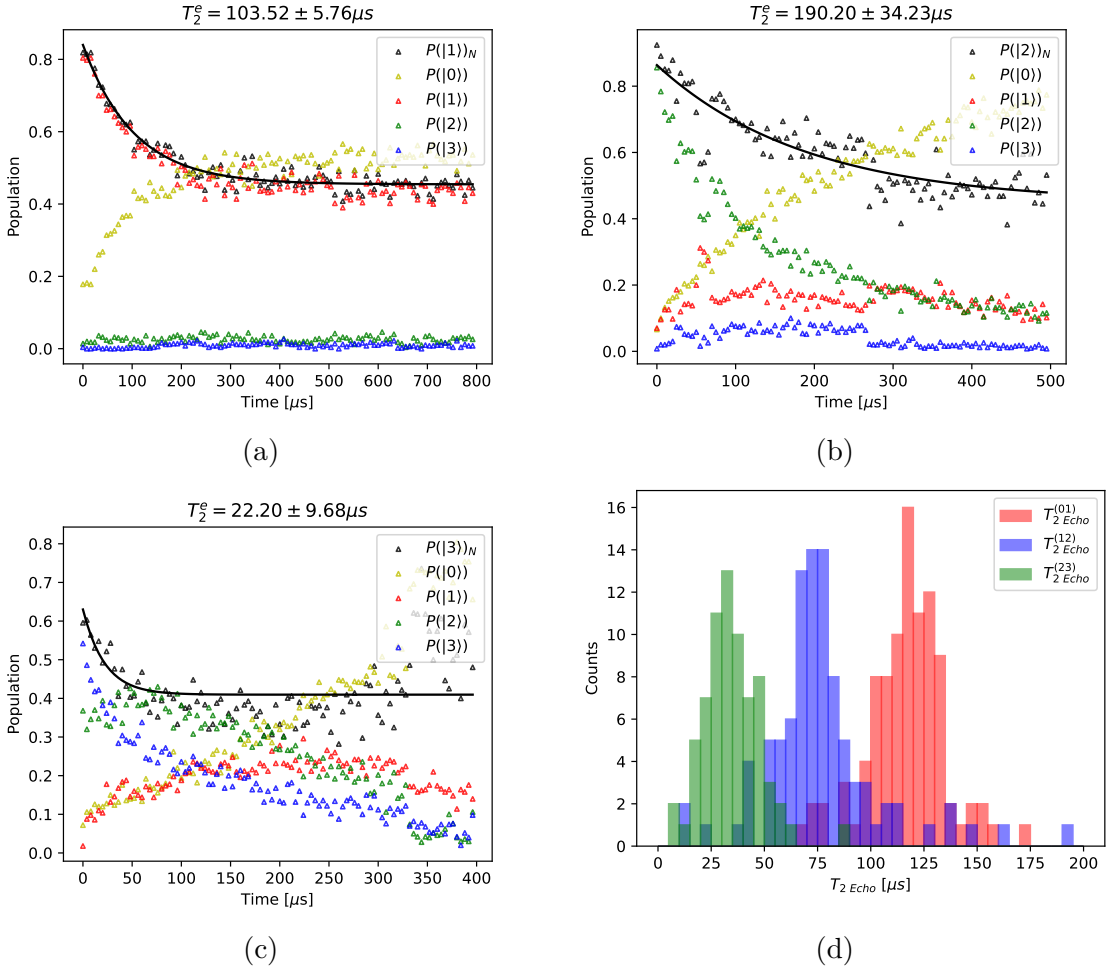
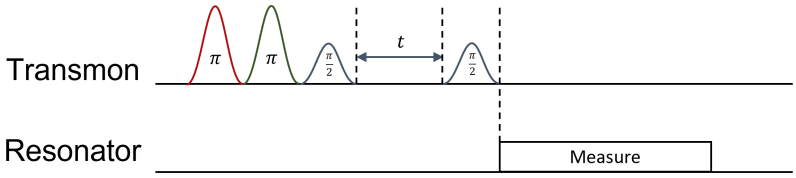
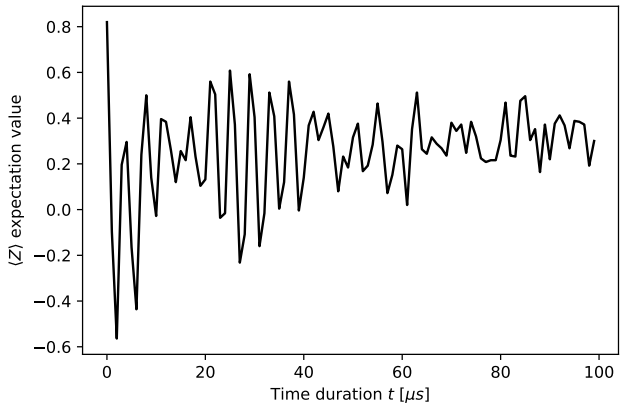


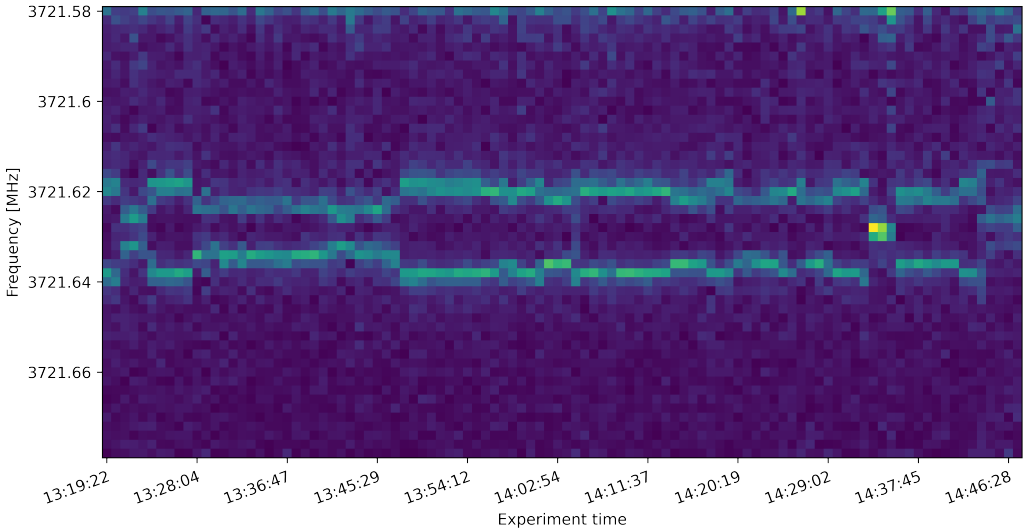
Figure 4.11: Characterisation of the T_2 coherence time of the 1QD device with the Hahn spin echo experiment. (a) to (c) show the population versus the delay time when the Hahn spin echo is implemented on the $\{|0\rangle, |1\rangle\}$, $\{|1\rangle, |2\rangle\}$, and $\{|2\rangle, |3\rangle\}$ subspaces, respectively. The T_2^e value is extracted by fitting the renormalised survival population (the black points) where unused levels are excluded. (d) shows the extracted T_2 echo distribution extracted from fitting the normalized survival population.



(a)



(b)



(c)

Figure 4.12: Characterising charge sensitivity on the $\{|2\rangle, |3\rangle\}$ subspace of the 1QD device. (a) Pulse sequence implementing the Ramsey interferometry experiment. First, the transmon is prepared to the $|2\rangle$ state by two consecutive π pulses in the $\{|0\rangle, |1\rangle\}$ and $\{|1\rangle, |2\rangle\}$ subspaces. Then two $\{|2\rangle, |3\rangle\}$ subspace $\pi/2$ pulses with interval t are applied and followed by a measurement. (b) A typical trace of the Ramsey interferometry experiment. By sweeping the interval t, a trace of the expectation value is collected to show the phase accumulation over time. A beating pattern caused by the charge noise can be observed in the trace. (c) The Ramsey spectrum of the transition frequency of the $\{|2\rangle, |3\rangle\}$ subspace. The X axis denotes the starting time of the Ramsey experiment. The Y axis denotes the Ramsey oscillation speed plus the drive frequency, representing the predicted transition frequency.

Chapter 5

Benchmarking Quantum Processor Performance

This chapter begins with a review of quantum benchmarking's mathematical formalism, summarised from [191, 104]. I developed a graphical language in Section 5.1 to illustrate superoperator relationships, inspired by similar applications of graphical language in Ref. [261]. Section 5.2 provides a summary of existing benchmarking's mathematical background. In Section 5.3, we present experimental results evaluating quantum logic gate fidelities, including a novel set of measurement fiducials I devised to minimise the number of measurements needed for qutrit state tomography. The idea of this minimised basis originated from me, and I wrote the code for experimentation and data analysis and collected the data through experiments. The results of benchmarking qutrit and qudit have been published in Ref. [41] and [40], respectively.

Characterising the performance of quantum logic gates is an essential task in building a quantum information processor. If we want to explore replacing qubits with multilevel qudits, it is important to determine how well the qudit-based circuits perform compared to those based on qubits. One of the most commonly used quantum benchmarking protocols, randomised benchmarking (RB), can produce average gate fidelity with confidence estimation [142, 54, 93, 137, 91, 135, 197, 186, 179, 218]. However, RB does not describe the exact error of each gate. Gate set tomography (GST) can instead be used to estimate the complete representation of the quantum process of the entire gate set, as well as the state preparation and measurement (SPAM) error [191, 104, 192, 106]. GST has been demonstrated on superconducting qubits, ion-trap systems, and silicon spin qubits [29, 289, 187], and it is an emerging technique for small-scale system characterization.

In this chapter, we first review the essential mathematical background for describing noisy quantum operations. It includes superoperator formalism for the manipulation of quantum channels and density matrices. Additionally, we introduce tensor networks to intuitively describe the mathematics behind superoperator formalism. Following this introduction, we describe the RB and GST protocols used to characterise our device gate errors and present the characterisation results from the two measured devices, 2QB and 1QD.

5.1 Mathematical Background

The Pauli and Clifford Groups

The Pauli group (or the generalised Pauli group for a system with more than 2 levels) is usually used as the basis for quantum operators. Within each qudit, the elements of the Pauli group can be generated by the following two generators [102, 99, 216]:

$$Z_d = \sum_{j=0}^{d-1} \omega^j \left|j\right\rangle \left\langle j\right|, \quad X_d = \sum_{j=0}^{d-1} \left|j\right\rangle \left\langle j+1\right| \tag{5.1}$$

where the addition of integers in $|j+1\rangle$ is considered modulo d, and d is the dimension of the qudit. $\omega = e^{2\pi i/d}$. The Pauli group for a single qudit can then be written as:

$$\mathbf{P}_d = \{\omega^\lambda X_d^x Z_d^z | \lambda \in \{0...d-1\}, x \in \{0...d-1\}, z \in \{0...d-1\}\} \tag{5.2}$$

The Pauli group for a system composed of multiple qudits can be represented as:

$$\mathbf{P}_d^{(n)} = \{\omega^\lambda P_1 \otimes P_2 ... \otimes P_n | \lambda \in \{0...d-1\}, P_j \in \mathbf{P}_d\} \tag{5.3}$$

The Clifford group is defined to be the group of all unitaries that normalise the Pauli group, given by

$$\mathbf{C}_d^{(n)} = \{V \in U_{d^n} | V \mathbf{P}_d^{(n)} V^\dagger \in \mathbf{P}_d^{(n)}\} \tag{5.4}$$

And the Clifford gates are defined to implement the rotation of an element of the Clifford group.

Qubit For a single qubit, the Pauli group comprises Pauli operators, given by [101]

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.5}$$

The qubit Clifford group is generated by

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, S = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \tag{5.6}$$

which results in a rotation of multiples of $\pi/2$ across the Pauli basis.

Qutrit For a single qutrit, the Pauli group is formed by generalised Pauli operators. One choice of these is the Gell-Mann matrices [97, 232]:

$$\lambda_1 = X_{01} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_2 = X_{02} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\lambda_3 = X_{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \lambda_4 = Y_{01} = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda_5 = Y_{02} = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}$$

$$\lambda_6 = Y_{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad \lambda_7 = Z_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda_8 = Z_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{5.7}$$

The qutrit Clifford group is generated by [99]

$$H = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \omega \end{pmatrix} \tag{5.8}$$

where $\omega = e^{i\frac{2}{3}\pi}$, which results in a rotation of multiples of $\pi/3$ across the Gell-Mann basis. The Clifford group has 216 elements when ignoring the global phases. Note that, unlike the qubit Hadamard gate, the qutrit Hadamard gate is not self-inverse. $HH = X_{23}(\pi)$ and $HHH = H^{-1}$. For convenience, we define a qutrit Z rotation to be

$$Z(a,b) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega^a & 0 \\ 0 & 0 & \omega^b \end{pmatrix} \tag{5.9}$$

The qutrit Clifford gates can be constructed with the following three forms [279]:

$$HZ(a_3,a_4)H^{-1}Z(a_1,a_2) \tag{5.10}$$

$$HZ(a_5,a_6)H^{-1}Z(q_1,q_2)HZ(p_1,p_2)H^{-1} \tag{5.11}$$

$$H^{-1}Z(a_7,a_8)H^{-1}Z(m_1,m_2) \tag{5.12}$$

where $(a_i,a_{i+1}) \in \mathcal{A}$, $(p_i,p_{i+1}) \in \mathcal{A}$, $(q_i,q_{i+1}) \in \mathcal{Q}$, $(m_i,m_{i+1}) \in \mathcal{M}$. and $\mathcal{P} = \{(1,1),(2,2)\}$, $\mathcal{N} = \{(1,0),(0,1),(2,0),(0,2)\}$, $\mathcal{M} = \{(0,0),(2,1),(1,2)\}$, $\mathcal{Q} = \mathcal{P} \cup \mathcal{N}$, $\mathcal{A} = \mathcal{Q} \cup \mathcal{M}$.

Tensor networks

Tensor networks are a class of graphical reasoning approaches to investigating multilinear maps and are particularly useful in representing quantum processes. This section reviews the basics of tensor network notations and their application to quantum information. For further details, see references [200, 21].

Tensor Tensors are the generalisation of scalars, vectors, and matrices. An intuitive illustration of a rank d tensor is a d-dimensional cube filled with numbers at each unit volume. Specifically, a rank-0 tensor is a scalar; a rank-1 tensor is a vector; and a rank-2 tensor is a matrix. Just as a matrix does not have to be square, the number of entries (width) in each tensor dimension is not necessarily equal. This width is called the bond dimension. For each tensor dimension, an index is assigned to denote the particular axis. For example, a rank-3 tensor can be written as $T^{i,j,k}$ where i, j, k are the indices.

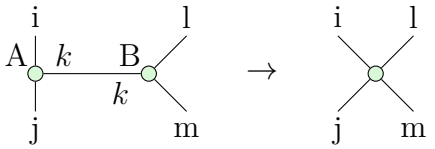
Tensor network notations use a graph to represent tensors. Each tensor is represented by a node. Each edge or *leg* represents a dimension of the tensor. For example, a scalar can be represented as a node without edges. A vector is a node with only one edge, and a rank-d tensor has d edges.

Scalar Vector Matrix Tensor

Like a matrix product, a tensor contraction is an operation summing over a repeated index of two tensors. For example,

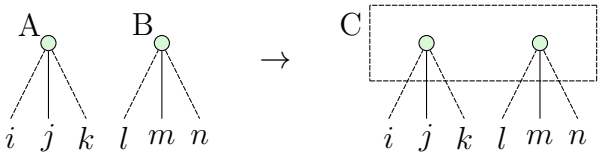
$$A^{i,j,k}B^{m,n}_{k} = C^{i,j,m,n} \tag{5.13}$$

In tensor network notation, contracting two tensors is represented by connecting two legs.



Putting two tensors together without contraction is equivalent to doing the Kronecker product of two tensors.

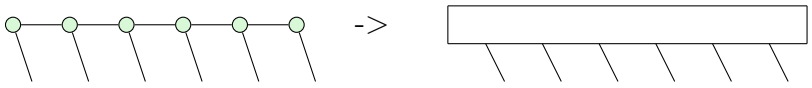
$$A^{i,j,k}B^{l,m,n} = C^{i,j,k,l,m,n} \tag{5.14}$$



The trace of a rank-2 tensor (a matrix) can be considered as contracting the tensor with itself.

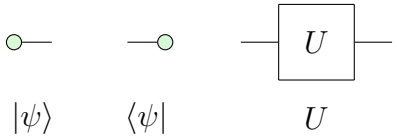
$$\mathrm{Tr}(A) = A^i_i \tag{5.15}$$

Tensor networks In tensor network notation, equations can instead be represented as a graph. The contraction of a dimension of two tensors is represented by connecting two legs. These legs are referred to as *closed legs*. Legs not connected are referred to as *open legs*.

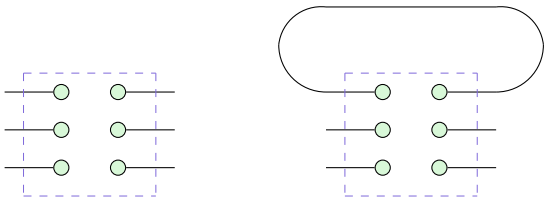


Using tensor network notation, a large tensor can be broken into contractions of smaller tensors, which form a tensor network. The advantage of tensor network representation is that it captures the relationship within the tensor and reduces the Hilbert space dimensions required to represent the contracted tensor.

Quantum information processing with tensor networks The basic concepts in quantum information processing can be represented with tensor networks. Quantum states and unitary quantum operations can be represented with rank-1 and rank-2 tensors.



The pure state density matrix can be represented as the Kronecker product of the state vector and its conjugate. The partial trace of the density matrix can be represented by contracting the legs of the subsystem.



$|\psi\rangle\langle\psi|$ Partial traced density matrix

In this section, we have reviewed the basics of tensor network methods and demonstrated how to use tensor networks to do basic quantum information processing calculations. Tensor network methods are more useful when generalised to non-unitary evaluations and describe the superoperator formalism.

Superoperator Formalism

In a perfect decoherence-free quantum processor, a quantum process is a unitary evolution of a pure quantum state. However, for NISQ devices with non-negligible noise, we need to consider the statistics of the noise and represent our quantum

state using density operators and quantum processes as linear transformations of the density operator. Now a quantum process is no longer a linear transformation acting on a vector in Hilbert space but acting on an operator in the Hilbert space. The linear transformation acting on an operator is called a *superoperator*. To simplify the representation, calculations can be moved to another space, called *Schmidt-Hilbert space* [58]. In this new space, the original operators in the Hilbert space become vectors, and superoperators become operators and can be represented as matrices. This method of describing quantum states and processes is called the *superoperator formalism* [139]. Note that in this section, the tensor network representation is used to illustrate the intuitions of the linear transformations. The representation does not distinguish the transpose, adjoint, and conjugate operators from the original operator, and constraints arising from the inner product inside the Schmidt-Hilbert space are ignored. For more details about this graphical method, please refer to [286].

Quantum States A state of the system is represented by a density operator $\rho \in \mathcal{H}$ in some Hilbert space $\mathcal{H}$. The matrix representation of the density operator is given by the inner product between each computational basis. For example, the density matrix of an N-qubit system is

$$\rho = |\psi\rangle \langle\psi| = \begin{pmatrix} \langle 0|\psi\rangle \langle\psi|0\rangle & \dots & \langle 0|\psi\rangle \langle\psi|d\rangle \\ \dots & & \dots \\ \langle d|\psi\rangle \langle\psi|0\rangle & \dots & \langle d|\psi\rangle \langle\psi|d\rangle \end{pmatrix} \tag{5.16}$$

where $|i\rangle, 0 \leq i \leq d = 2^N - 1$ is the i-th basis state.

Since the density operator is a Hermitian operator, the density operator can also be represented as a weighted summation of Pauli operators $P_j \in \mathcal{P}^{\otimes n}$ where $\mathcal{P} = \{I, X, Y, Z\}$:

$$\rho = \sum_{i=0}^{d} \lambda_i P_i, \ \ \lambda_i \in [-1, 1] \tag{5.17}$$

Here the quantum operator is no longer represented by a matrix but a vector

$\{\lambda_i\}$. Define the *super-ket* $|\rho\rangle\rangle = \{\lambda_i\}$, as the vector in the Hilbert-Schmidt space. The inner product in Hilbert-Schmidt space is defined as

$$\langle\langle A|B\rangle\rangle = \frac{1}{d}A^\dagger B \tag{5.18}$$

where A and B are density operators. The expectation value p of measurement operator E for a state ρ is

$$p = \mathrm{Tr}\{E\rho\} \tag{5.19}$$

For simplicity, the Pauli operator $\bar{P}_i$ is introduced as

$$\bar{P}_i = \frac{1}{\sqrt{d}}P_i \tag{5.20}$$

to avoid the factors of $\sqrt{d}$ in the equation. Now the state ρ can be represented as a state vector:

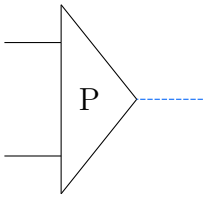
$$|\rho\rangle\rangle = \sum_i |i\rangle\rangle \langle\langle i|\rho\rangle\rangle \tag{5.21}$$

where each element of the state vector is given by

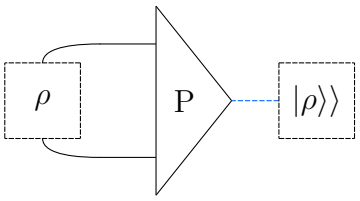
$$|\rho\rangle\rangle_i = \langle\langle i|\rho\rangle\rangle = \mathrm{Tr}\{\bar{P}_i\rho\} \tag{5.22}$$

$$\bar{P}_i \in \{\frac{1}{\sqrt{2}}I, \frac{1}{\sqrt{2}}X, \frac{1}{\sqrt{2}}Y, \frac{1}{\sqrt{2}}Z\}^{\otimes n} \tag{5.23}$$

In the tensor network representation, we define a rank 3 tensor $\mathcal{P}^i_{kj}$. Here k and j denote the two dimensions of each Pauli matrix, and i denotes the index of the Pauli matrix. It can be considered as all Pauli matrices stacked together. Here we use the solid black line to denote the legs of k and j, since they represent quantities in the Hilbert space, and the contraction follows the rule of inner products in the Hilbert space. The blue dashed lines represent the quantities in the Schmidt-Hilbert space, and the rule of inner product is slightly different up to a constant to the Hilbert space.



The relation between the super-ket and the density operator can be represented as follows. The dashed box means it is a slot and we could plug in the matrices. If both ρ and $|\rho\rangle\rangle$ are plugged in, the contracted result should be 1. If only the ρ is plugged in, after contraction the result would be the $|\rho\rangle\rangle$ and vice versa.



Quantum Operations Quantum operations are defined as linear maps on the density matrix as $\rho \to \Lambda(\rho)$ where $\Lambda(\rho) \in \mathcal{H}$. Restrictions on map Λ are required to ensure it describes a physical process in our universe. Suppose $\rho' = \Lambda(\rho)$ for any $\rho \in \mathcal{H}$:

1. Complete Positivity (CP): The probability must be non-negative. ρ' must have nonnegative probabilities for measuring the eigenstates of any observables. This means the density matrix in our universe must be semi-definite. 2. Trace Preservation (TP): The total probability must be conserved[1]. The probability of measuring system ρ in its eigenstates must sum up to 1, therefore $\mathrm{Tr}\{\rho'\} = 1$.

Any map that is physical must be a CPTP map. To fulfil the CP requirement, the map Λ can be written as [194]:

$$\Lambda(\rho) = \sum_{i=1}^{d^2} K_i \rho K_i^\dagger \tag{5.24}$$

for some $N \leq d^2$ where $d = 2^n$ is the dimension of the Hilbert space. Matrices K_i are the so-called *Kraus* operators, and this representation is called the *Kraus*

[1]This requirement can be relaxed when there are leakage errors, which transform the qubit state outside of the computation space.

representation. The map fulfils the TP requirement when $\sum_i K_i^\dagger K_i = I$ where I denotes the identity. This is called the *Completeness Condition* [194].

Process Matrix A convenient method to manipulate quantum operations can be derived immediately from the Pauli basis representation of a quantum state and the Pauli Transfer Matrix (PTM) representation of a quantum operation. Quantum operators are represented as matrices of dimension $d^2 \times d^2$.

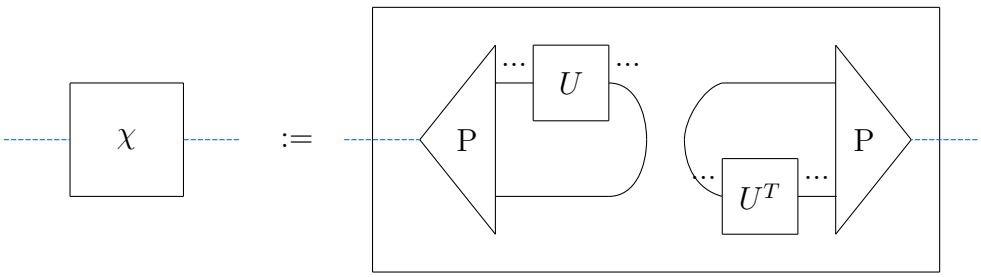
The Kraus operator is not unique for a given quantum operation. A convenient choice is to expand the Kraus operator into a linear combination of Pauli operators.

$$K_i = \sum_{j=1}^{d^2} a_{ij} P_j, P_j \in \mathcal{P}^{\otimes n} \tag{5.25}$$

Inserting this into Eq. 5.24 gives

$$\Lambda(\rho) = \sum_{j,k=1}^{d^2} \chi_{jk} P_j \rho P_k \tag{5.26}$$

where $\chi_{jk} = \sum_i a_{ij} a_{ik}^*$. χ is a $d^2 \times d^2$ complex-valued matrix, and completely determines the map Λ. The χ matrix is called the *process matrix*. The χ matrix is Hermitian and positive semidefinite. For TP map, $\sum_{jk} \chi_{jk} P_k P_j = I$[2]. In the tensor network representation, we consider how to relate the process matrix with the unitary U in the Hilbert space. The following diagram shows its relationship.



The process matrix can be considered to represent unitary U and its transpose into the basis in the Hilbert-Schmidt space.

[2] $I = \sum_i K_i^\dagger K_i = \sum_{ijk} a_{ik}^* P_k a_{ij} P_j = \sum_{jk} (\sum_i a_{ij} a_{ik}^*) P_k P_j = \sum_{jk} \chi_{jk} P_k P_j$

Pauli Transformation Matrix (PTM) Quantum map Λ can also be represented as Pauli transfer matrix (PTM):

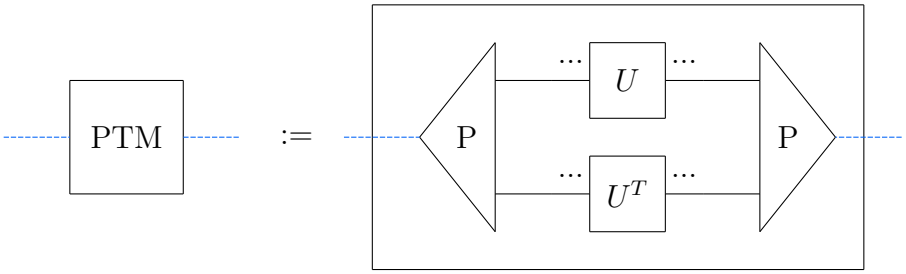
$$(R_\Lambda)_{ij} = \frac{1}{d}\mathrm{Tr}\{P_i \Lambda(P_j)\} \tag{5.27}$$

The $(R_\Lambda)_{ij}$ denotes the expectation value when the state is prepared at P_j basis and measured at P_i basis after the quantum map. Note $\Lambda(P_j)$ is a density matrix, and from Eq. 5.17, $(R_\Lambda)_{ij}$ also must be real and in the interval $(-1, 1)$. From these definitions, manipulating operators can be transformed into manipulating state vectors and matrices defined above.

$$|\Lambda(\rho)\rangle\rangle = R_\Lambda |\rho\rangle\rangle \tag{5.28}$$

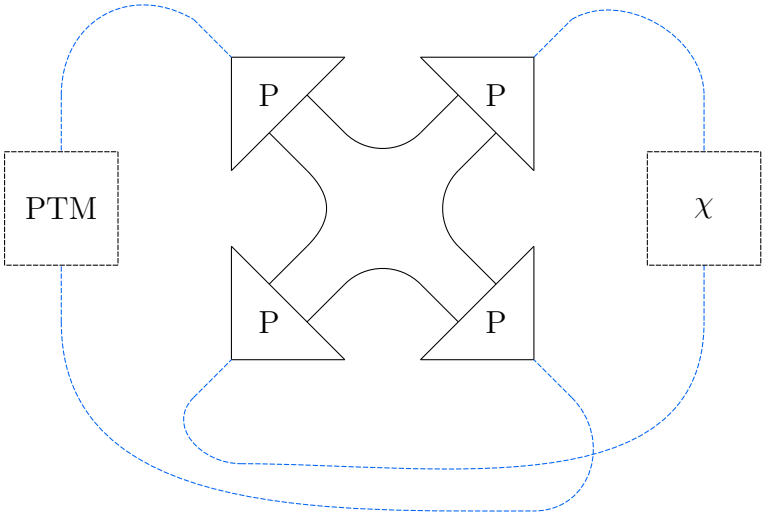
$$R_{\Lambda_1 \cdot \Lambda_2} = R_{\Lambda_1} R_{\Lambda_2} \tag{5.29}$$

The relationship between the PTM and the unitary operator is shown as follows.



The diagram above shows that the PTM for a unitary operation is the Kronecker product of U and U^T, representing them in the Pauli basis in the Schmidt-Hilbert space.

From the definition of the process matrix and PTM, we can show that the process matrix and PTM can be transferred from one to another with the following relationship.



Choi-Jamiołkowski Isomorphism The Choi-Jamiołkowski isomorphism (or channel-state duality) refers to the correspondence between quantum channels and quantum states [53, 124, 126]. More generally, Choi-Jamiołkowski isomorphism provides a method to investigate quantum channels by investigating quantum states.

Here we first consider the scenario where we only have pure states and a fully coherent quantum channel. Each quantum channel can be described by a unitary operation U, which transforms the state $|\psi_i\rangle$ to $|\psi_i'\rangle$. To fully describe the operation U, we need to describe $|\psi_i\rangle$, $|\psi_i'\rangle$ and their one-to-one relationship. A quantum state can be constructed to describe all this information.

$$|\Psi\rangle = \frac{1}{\sqrt{d}} \sum |\psi_i\rangle \otimes |\psi_i'\rangle \tag{5.30}$$

This quantum state $|\Psi\rangle$ is a superposition of all orthogonal input states $|\psi_i\rangle$ and output states $|\psi_i'\rangle$ from operator U. This state can be prepared by first having a simple maximum entanglement state $|\Psi^+\rangle$:

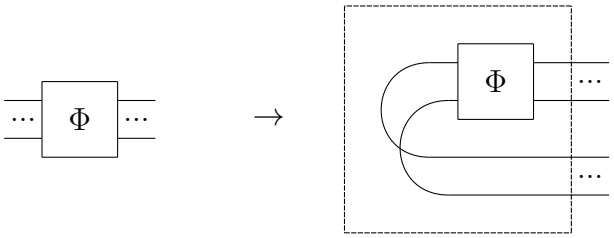
$$|\Psi^+\rangle = \frac{1}{\sqrt{d}} \sum (|i\rangle \otimes |i\rangle) \tag{5.31}$$

$$|\Psi\rangle = U \otimes I(|\Psi^+\rangle) \tag{5.32}$$

The above notation can be extended into superoperator formalism. Consider a quantum channel represented by a CPTP map Φ. The Choi-Jamiołkowski state is defined as:

$$\rho_{CJ}(\Phi) = (\Phi \otimes I)(\Psi) \tag{5.33}$$

The tensor network notation can illustrate this idea more readily. Say we have a quantum map Φ, with two legs on the left to denote the input and two on the right to denote the output. For a quantum state representing the complete information of Φ, the input leg can be bent to the right side, represented by a maximally entangled two quantum system and applying Φ to one of them.



5.2 Quantum Tomography and Benchmarking

Metrics for Quantum State and Process

State Fidelity The state fidelity measures the closeness of two quantum states. The fidelity of two quantum states is defined as [194, 131]

$$F_s(\rho, \tilde{\rho}) = \left(\mathrm{Tr} \left[\sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}} \right] \right)^2 \tag{5.34}$$

where ρ and $\tilde{\rho}$ are the density matrices of the two states we are comparing.

For most practical cases, we are comparing a noisy mixed state with an ideal pure state. The state fidelity can be simplified in this case as

$$F_s(|\psi\rangle, \rho) = \langle\psi|\rho|\psi\rangle \tag{5.35}$$

where ρ is the density matrix and $|\psi\rangle$ is the pure state we are comparing.

For pure states, the definition reduces to the squared overlap between the states:

$$F_s(|\psi_1\rangle, |\psi_2\rangle) = |\langle\psi_1|\psi_2\rangle|^2 \tag{5.36}$$

where $|\psi_1\rangle$ and $|\psi_2\rangle$ are the two pure states we are comparing.

Entanglement Fidelity The entanglement fidelity was first proposed to characterise the performance of a quantum channel for quantum communication. Between physical process A and ideal process B, the entanglement fidelity F_e is defined as [236]

$$F_e(A,B) = F_s(\rho_{CJ}(A), \rho_{CJ}(B)) = \mathrm{Tr}\left(\sqrt{\sqrt{\rho_{CJ}(A)}\rho_{CJ}(B)\sqrt{\rho_{CJ}(A)}}\right)^2 \tag{5.37}$$

The entanglement fidelity can be described as the state fidelity of the Choi-Jamiolkowski isomorphism density matrix. We use $\epsilon_e = 1 - F_e$ to denote the entanglement infidelity.

Average Gate Fidelity The average gate fidelity $\bar{F}$ is defined by [194, 193]

$$\bar{F}(\Phi, U) = \int d\psi \, \langle\psi| U^\dagger \Phi(\psi) U |\psi\rangle \tag{5.38}$$

where the integral is over the Haar measure (uniform distribution) $d\psi$ on the state Hilbert space, Φ is the noisy quantum operation being characterised, and U is the ideal unitary operation. $\bar{F}(\Phi, U) = 1$ if and only if Φ implements U perfectly. It is sometimes useful to consider that the state $|\psi\rangle$ is an arbitrary fixed state, but introducing a unitary before and after the gate to change the input state. The average gate fidelity can also be written as an integral over the Haar measure of preparation operation V as [193]

$$\bar{F}(\Phi, U) = \int dV \, \langle\psi| V^\dagger \Phi(U\psi U^\dagger) V |\psi\rangle \tag{5.39}$$

The average gate fidelity is different from the entanglement fidelity. Their relation is given by [115]

$$\bar{F} = \frac{d \times F_e + 1}{d + 1} \tag{5.40}$$

where d is the dimension of the state space.

Error Generator A quantum operation can have a small error such that the reconstructed superoperator is too similar to interpret its difference from the ideal gate visually. *Error generators* [28] have been developed to quantify these small errors. Consider the ideal target map $\hat{G}_i$ for a physically reconstructed map G_i, the *post-gate error generator* L is defined as

$$G_i = e^{L_i}\hat{G}_i \tag{5.41}$$

It approximates the error as an error operation applied after the ideal operation. We could also apply the error before the ideal operation, which is called the *pre-gate error generator*. In this work, the error generators refer to the post-gate error generators.

Reference [28] also discusses the *projected error generator*, which structures the error generator into a different type of error. Here we discuss two simpler metrics proposed in [28] to quantify the coherent and non-coherent error.

Jamiołkowski probability *Jamiołkowski probability* is a metric to quantify the non-coherent error. It is defined by [28]

$$\epsilon_J(L) = -\mathrm{Tr}(\rho_J(L)\left|\psi\right\rangle\left\langle\psi\right|) \tag{5.42}$$

where $\left|\psi\right\rangle$ represents a fully entangled state.

Looking at the Choi-Jamiołkowski state of the error generator $\rho_{CJ}(L)$ is beneficial. Since the error generator is usually small, we can approximate $\rho_{CJ}(L) \approx \rho_{CJ}(G) - \rho_{CJ}(\hat{G}_i)$. Now $\rho_{CJ}(L)$ can be represented in a basis $\{\left|\psi\right\rangle, \left|\psi\right\rangle_1', \left|\psi\right\rangle_2', ..., \left|\psi\right\rangle_n'\}$ where $\left|\psi\right\rangle$ is the maximum entanglement state, and $\left|\psi\right\rangle_i'$ are basis orthogonal to $\left|\psi\right\rangle$.

From Figure 5.1, ϵ_J represents the probability of the state "jumping" from the fully entangled state to other states. Here the "jump" means the spectrum (distribution of the eigenvalues) of the density matrix changes and introduces mixed states. Therefore,

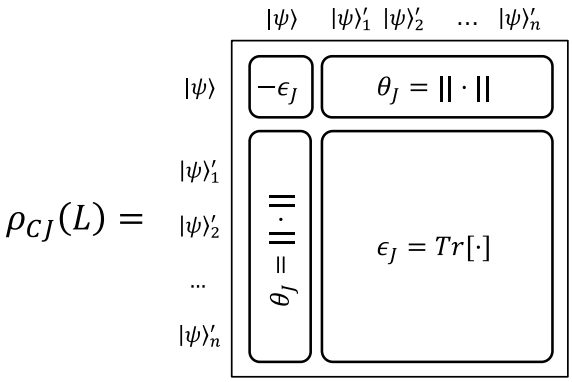


Figure 5.1: Choi-Jamiołkowski state of the error generator $\rho_{CJ}(L)$.

ϵ_J is related to the non-coherent part of the error generator.

Jamiołkowski amplitude The *Jamiołkowski amplitude* [28] is a metric for quantifying coherent error. It is defined by

$$\begin{aligned}\theta_J(L) &= |(1 - |\psi\rangle\langle\psi|)\rho_J(L)|\psi\rangle| \\ &= \sqrt{\langle\Psi|\rho_{CJ}(L)^2|\Psi\rangle - \langle\Psi|\rho_{CJ}(L)|\Psi\rangle^2}\end{aligned} \tag{5.43}$$

From $\rho_{CJ}(L)$ representation, $\mathcal{E}_J$ represents the probability of the state "rotating" from the fully entangled state to other states. Here "rotates" means the spectrum of the density matrix remains the same. Therefore, it is related to the coherent part of the error generator.

Generator infidelity The *generator infidelity* is defined by [28, 161]

$$\epsilon_g = \epsilon_J + \theta_J^2 \tag{5.44}$$

Under the common cases where coherent errors are much larger than stochastic errors, the entanglement fidelity can be approximated by generator infidelity $\epsilon_e \approx \epsilon_g$ [28, 161].

Quantum state tomography

The direct information we can get from measuring a superconducting qubit by conventional dispersive readout is whether the qubit is in its $|0\rangle$ state or the $|1\rangle$ state. From this information, we can construct the diagonal elements of the density matrix only. In order to find the off-diagonal elements, we need to change the "aspect" of looking at the quantum state: rotate the state before measurement to effectively change the measurement basis. Just as the word *tomography* describes reconstructing 3D images from sensors in different directions, it is borrowed to describe the protocol that reconstructs the complete information of a quantum state [194, 6, 258].

The goal of quantum state tomography is to construct the full density matrix of an unknown quantum state. A convenient method is to utilise Eq.5.22. The density operator in the matrix can be constructed by

$$\rho = \sum_i \langle\langle P_i|\rho\rangle\rangle \, P_i = \sum_i \hat{\rho}_i P_i \tag{5.45}$$

where $\hat{\rho} = \{\langle\langle P_0|\rho\rangle\rangle, \langle\langle P_1|\rho\rangle\rangle, ..., \langle\langle P_n|\rho\rangle\rangle\}$ is the *Pauli vector*. In order to construct the density matrix, we need to find all $\langle\langle i|\rho\rangle\rangle$. Some of the $\langle\langle i|\rho\rangle\rangle$ can be easily found when the $|i\rangle\rangle$ is directly measurable. For example, the Pauli Z operator can be measured directly on the quantum computing system that physically measures the population of $|0\rangle$, $|1\rangle$. For some other cases, we could choose a set of physically implementable measurement operators $\langle\langle M_i| \, F_j^{(m)}$ that spans the whole Hilbert space, where $\langle\langle M_i|$ are measurement operators implementable on the quantum system, and $F_j^{(m)}$ is called *measurement fiducials*, a set of operations that rotate the state before the measurement, to effectively measure with another operator.

In practice, the measurement operator is usually chosen to be $M_l = \mathrm{diag}(\{\delta_{i,l}\})$, which produces the expectation value equal to the probability of the state measured in $|l\rangle$. We could directly estimate the probability $p_{ij} = \langle\langle M_i|F_j^{(m)}|\rho\rangle\rangle$ from the experiment. Since $\langle\langle M_i| \, F_j^{(m)}$ can be represented as a linear combination of the Pauli

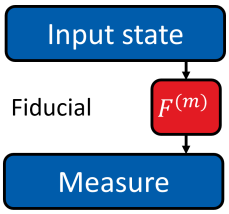


Figure 5.2: Experiment implementation of state tomography. The circuit repeatedly starts with an unknown state, and the state tomography applies a set of measurement fiducials and then measures the system and generates experimental data p_{ij}. Then the density matrix of the unknown state can be reconstructed with equation 5.47.

basis $\langle\langle P_k|$, we can rewrite the expression of p_{ij} to

$$p_{ij} = \sum_k \langle\langle E_l|F_j^{(m)}|P_k\rangle\rangle \langle\langle P_k|\rho\rangle\rangle \tag{5.46}$$

where we denote $A = \langle\langle E_l|F_j^{(m)}|P_k\rangle\rangle$. From the above relation, the value of the Pauli vector can be recovered as

$$\hat{\rho} = A^{-1}p_{ij} \tag{5.47}$$

Note that when an overcomplete measurement basis is used for measurement, A is no longer full-rank, and the inverse of A becomes a pseudo-inverse. The value is effectively estimated with the least square error linear estimation to the measured result.

Quantum process tomography

Quantum process tomography (QPT) is a method to estimate the complete quantum process of a given system. The process matrix or the Pauli transformation matrix describes a quantum process, including the noise. Since we can already extract Pauli vectors from state tomography, it is intuitive to move on and construct the PTM. By iterating the initial state on Pauli bases, the PTM can be constructed by applying the quantum gate and doing state tomography for each input state. More formally speaking, the PTM R_Λ for gate Λ is given by

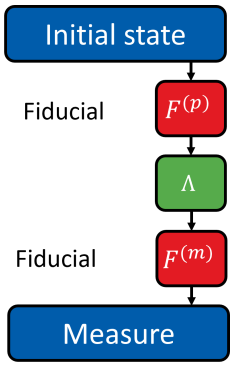


Figure 5.3: Experiment implementation of process tomography. The circuit repeatedly starts with a known initial state, usually the system's ground state. Then preparation fiducials $F_i^{(p)}$ are applied to the initial state and transfer the state into a set of states $|\rho_i\rangle$. Then the process we are interested in is applied to the state $|\rho_i\rangle$, followed by a state tomography gate sequence with measurement fiducial $F_j^{(p)}$. By iterating i and j for preparation and measurement fiducials, the experimental data p_{ijl} can be collected and the process can then be reconstructed by equation 5.51.

$$(R_\Lambda)_{ij} = \langle\langle \bar{P}_j | \, \Lambda \, | \bar{P}_i \rangle\rangle \tag{5.48}$$

So far we have very abstractly described how to evaluate each element of R_Λ. However, practically we cannot prepare $\bar{P}_i$ as a quantum state. Consider preparing a set of states $\{|\rho_i\rangle\} = \{F_i^{(p)} |\rho_0\rangle\rangle\}$ that spans the whole Hilbert space. Let the quantum channel Λ apply to all these states, and run state tomography for each state. The experiment implementing process tomography is shown in figure 5.3. We can obtain a set of probabilities given by

$$p_{ijl} = \langle\langle E_l | F_j^{(m)} | \Lambda | F_i^{(p)} | \rho_0 \rangle\rangle \tag{5.49}$$

$$= \sum_{kl} \langle\langle E_l | F_j^{(m)} | P_k \rangle\rangle \, \langle\langle P_k | \Lambda | P_l \rangle\rangle \, \langle\langle P_l | F_i^{(p)} | \rho_0 \rangle\rangle \tag{5.50}$$

Now if we define $B_{li} = \langle\langle P_l | F_i^{(p)} | \rho_0 \rangle\rangle$, the PTM can be recovered by

$$R_\Lambda = A^{-1} p_{ijl} B^{-1} \tag{5.51}$$

Just as with state tomography, when we include an overcomplete basis B^{-1} becomes a pseudo-inverse and gives the least-square estimate.

Gate Set tomography

While process tomography could produce the complete PTM of a quantum operation, it assumes the measurement and preparation of the states are perfect. Ignoring the state preparation and measurement (SPAM) errors could be detrimental when trying to characterise the small error of a quantum process and report its fidelity. Gate set tomography (GST) is developed to solve this issue [104, 191]. Instead of characterising a quantum process, GST characterises all the involved operators, including the initial state operator, measurement operators, all the fiducials and the quantum process we are interested in.

Suppose a gate set $\mathcal{G} = \{G_0, G_1, \ldots, G_n\}$ is defined so that it can compose all the fiducial F gates to implement state preparation and measurement. That is, for any F_i, we could always write it as

$$F_i = \prod_{j \in S} G_j \tag{5.52}$$

Then we could have a measurement result

$$m_{ijkl} = \langle\langle E_l | F_i^{(m)} G_k F_j^{(p)} | \rho_0 \rangle\rangle \tag{5.53}$$

The estimated PTM can be found with the maximum likelihood method by minimising the objective function $\mathcal{L}_m$

$$\mathcal{L}_m = \sum_{ijkl} (\langle\langle \tilde{E}_l | \tilde{F}_i \tilde{G}_k \tilde{F}_j | \rho \rangle\rangle - m_{ijkl})^2 \tag{5.54}$$

One variant of GST is the long-sequence GST, which repeats multiple different sequences of intrinsic gates defined in G to amplify the error of each gate. The experimental implementation of the long-sequence GST is shown in figure 5.4. Given σ_{ijk}^2 is the variance of the m_{ijk}, the optimisation problem is defined by the objective

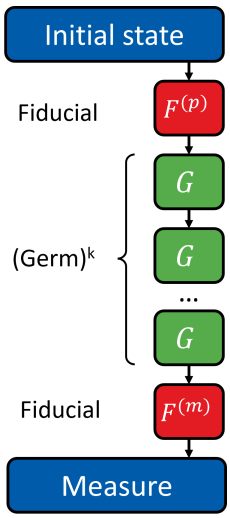


Figure 5.4: Experiment implementation of long sequence gate-set tomography. The circuit repeatedly starts with a known initial state, usually the system's ground state. Then preparation fiducials $F_i^{(p)}$ are applied to the initial state and transfer the state into a set of states $|\rho_i\rangle$. Then germs comprised of one or more gates in the gate set are applied to the state $|\rho_i\rangle$ k times. Then a state tomography gate sequence follows with measurement fiducial $F_j^{(p)}$. By iterating i and j for preparation and measurement fiducials, the experimental data p_{ijl} can be collected, and all the PTMs, initial state density matrix and measurement operators can be estimated by optimisation discussed in equation 5.55.

function

$$\mathcal{L}_m = \sum_{ijkl} (\langle\langle \tilde{E}_l | \tilde{F}_i \tilde{G}_k^{(1)} \tilde{G}_k^{(2)} \dots \tilde{G}_k^{(n)} \tilde{F}_j | \rho \rangle\rangle - m_{ijkl})^2 / \sigma_{ijk}^2 \tag{5.55}$$

Subject to

$$\rho_{CJ}(G) \succcurlyeq 0, \forall G \in \mathcal{G} \tag{5.56}$$

$$(R_G)_{0i} = \delta_{0i}, \forall G \in \mathcal{G} \tag{5.57}$$

$$(R_G)_{ij} \in [-1, 1], \forall G, i, j \tag{5.58}$$

$$\mathrm{Tr}(\rho) = 1 \tag{5.59}$$

$$I - E \succcurlyeq 0 \tag{5.60}$$

Constraint 5.56 ensures the PTM is positive by subjecting its CJ matrix to positive semidefinite. Constraint 5.57 sets the first row of the PTM to be one and all zeros, which is required by the TP constraints. Constraint 5.58 forces the value of PTM to be between -1 and 1. Constraints 5.59 and 5.60 enforce the initial state operator, and the measurement operator is physical.

The above maximum likelihood estimation still has gauge freedom missing and can produce non-unique results with the same dataset. Consider a quantum process B, and consider the following measurement result m'_{ijkl}

$$\begin{aligned} m'_{ijkl} &= \langle\langle E_l | B(B^{-1} F_i^{(m)} B)(B^{-1} G_k B)(B^{-1} F_j^{(p)} B) B^{-1} | \rho_0 \rangle\rangle \\ &= \langle\langle E'_l | F_i^{(m)\prime} G'_k F_j^{(p)\prime} | \rho'_0 \rangle\rangle \end{aligned} \tag{5.61}$$

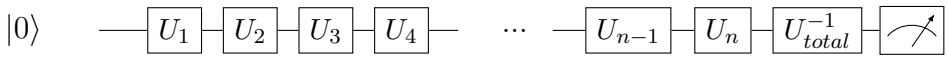
which would be identical to m_{ijkl}. For the estimated operators representing the physical process, we need to optimise the gauge operation B. This is done by minimising the sum of trace norms between the estimated operators and their ideal counterparts.

Randomised benchmarking

Another SPAM-error-insensitive method to characterise gate fidelity is Randomised Benchmarking (RB) [142, 54]. RB has been widely demonstrated on different quan-

tum systems [93, 91, 135, 197, 179], and has been developed for multilevel systems beyond qubits [137, 186, 123]. Compared to GST, RB scales efficiently and is more practical for larger quantum systems. However, the standard RB protocol only reports the average gate fidelity. It cannot determine the full error model of a quantum system and is insufficient for applications that require detailed error descriptions, such as quasi-probability error mitigation [82].

Standard randomised benchmarking An intuitive illustration of randomised benchmarking is to consider doing an identity operation that brings the quantum system to a random state and back to where it started, usually the ground state. If the operation is perfect, we should always observe that the measured system is in the ground state. However, when the operation contains errors, there are chances that the system will not land in the ground state. When the number of gates involved in implementing the operation grows, the probability that we measure the system in the ground state drops exponentially. We could extract the average gate fidelity by fitting the state probability decay curve, sometimes referred to as the *Success Probability*.



Unitary t-design While the idea of RB is straightforward, there are some problems to consider before implementing practical RB on quantum hardware. The original proposal of RB [80] defines the average gate fidelity to be the averaged fidelity over all possible unitary operations, which requires choosing the gate randomly from the Haar measure of the $SU(d)$ group, where d is the dimension of each quantum building block.

Now we set the arbitrary state of the average gate fidelity definition in Eq.5.39 to the ground state and get

$$\bar{F} = \int_{U(d)} dU \mathrm{Tr}[U |0\rangle \langle 0| U^\dagger \Lambda(U |0\rangle \langle 0| U^\dagger)] \tag{5.62}$$

Where $U(d)$ is the set of all $d \times d$ unitary matrices in the $SU(d)$ group. Such a configuration is difficult in practice, mainly because finding the inverse gate of a long sequence to put the system back to the initial state is difficult. It has in fact been shown that there is no need to strictly sample the gate from the Haar measure, but simply from the Clifford group is enough for average gate fidelity estimation [62]. This is because the Clifford group is an exact unitary 2-design [62]. Formally speaking, X is a unitary-t design if the following equation holds.

$$\frac{1}{|X|} \sum_{U \in X} U^{\otimes t} \otimes (U^*)^{\otimes t} = \int_{U(d)} U^{\otimes t} \otimes (U^*)^{\otimes t} dU \tag{5.63}$$

where $|X|$ denotes the number of gates in set X. From the definition, any homogeneous degree 2 polynomial function, sampling from a unitary-2 design is an adequate substitution of sampling from the Haar measure. The definition of average gate fidelity is a degree 2 polynomial.

Extracting average gate fidelity We have now observed a success probability of the RB experiment, now we need to find what exactly is the decay rate of the measured success probability, and how it relates to the average gate fidelity. More formally speaking, denote P_m as the success probability at sequence length m. P_m is exponentially decaying versus m, given as [163, 162]

$$P_m = A + Bp^m \tag{5.64}$$

where m is the length of the gate sequence, A and B are fitting parameters. The average infidelity per Clifford Gate can be estimated as

$$1 - \bar{F} \approx r = \frac{d-1}{d}(1-p) \tag{5.65}$$

where d is the dimension of the system. It has been proven that the RB protocols exactly measure the average gate-fidelity even under strong gate-dependent errors, but not non-Markovian (time-dependent) errors [218, 269, 44].

Characterising leakage *Leakage errors* are an important factor for operating superconducting circuits. Each gate defined on the superconducting circuits operates within a pre-defined subspace, usually the lowest d energy levels. The leakage errors result from the state of a quantum system "leaking" out of a pre-defined subspace to occupy an unwanted energy level.

Leakage randomised benchmarking (LRB) has been developed to characterise the leakage error for quantum logic gates with RB [270, 287]. While the gate set used in this work only operates between neighbouring energy levels, the transmon state can be transferred to unused energy levels, creating unwanted leakage states. Understanding the leakage error model for the device is beneficial for further applications such as quantum error corrections and error mitigation [178].

The LRB protocol is similar to the standard RB protocol, except it measures the leakage state population $p_m^{(l)}$ and keeps track of the overall sum of the population in the computation space $p_m^{(\mathcal{I})} = 1 - p_m^{(l)}$. Then the $p_m^{(\mathcal{I})}$ can be fit into a decay model

$$p_m^{(\mathcal{I})} = A + B\lambda_1^m \tag{5.66}$$

where A and B are fitting parameters with $0 \leq A, B$. The average leakage rate L_1 is given by [287]

$$L_1 = (1 - A)(1 - \lambda_1) \tag{5.67}$$

.

Interleaved randomised benchmarking While standard RB reports the average gate fidelity of a set of gates, finding the fidelity of a specific gate is sometimes beneficial. Interleaved randomised benchmarking is one method that estimates the fidelity of a specific gate [164]. The protocol first performs a standard randomised benchmarking and extracts the parameter p_s. The specific gate of interest is then inserted between two gates in the gate sequence. The last gate is updated so that the overall sequence still gives the identity operation. From the interleaved decay curve, extract parameter p_i.

And the average infidelity of the interleaved gate is estimated as

$$1 - \bar{F} \approx r_i = \frac{d}{d-1}(1 - \frac{p_i}{p_s}) \tag{5.68}$$

.

5.3 Benchmarking Results on Real Devices

We now move on to report the results of carrying out particular benchmarking techniques on two specific superconducting devices: first, the 2QB device, treating the transmons as qubits, and second, the 1QD device, treated as a 3 and 4-level system (which for simplicity we denote here as 'qutrit' and 'qudit' respectively). We first discuss carrying out RB and GST on the 2QB device, then report novel techniques to implement qutrit and qudit versions of GST on the 1QD device.

Benchmarking Results for the 2QB Device

The 2QB device is tuned with a 50-nanosecond Blackman DRAG pulse for a single-qubit gate. The two-qubit gate is implemented with a 300-nanosecond cross resonance drive.

Gate Set Tomography The gate set tomography result for this experiment is produced with the "std2Q_XYCNOT" model from pyGSTi [192]. The qubit-based system can use a simple choice of fiducials, which are $\frac{1}{2}(\hat{I}+\hat{Z})$, $\frac{1}{2}(\hat{I}-\hat{Z})$, $\frac{1}{2}(\hat{I}+\hat{X})$, $\frac{1}{2}(\hat{I}+\hat{Y})$. They correspond to the axes $|0\rangle$, $|1\rangle$, $|+x\rangle$, $|+y\rangle$ on the Bloch sphere. Applying $\hat{I}$, $\hat{Y}_{\frac{\pi}{2}}$, $\hat{X}_{-\frac{\pi}{2}}$, $\hat{X}_{\pi}$ and then measuring could yield the expected results. In the following table, we use $X_A(\theta)$ to denote the arbitrary rotation around the X-axis on qubit A with angle θ. The detailed fiducial configuration is listed in Appendix B.

The "std2Q_XYCNOT" model also includes six gates in the gate set G and 12 germs for error amplification. The detailed configuration is listed in Table C.1. The experiment is set as follows. The maximum germ length is set to 64, which gives 2900 gate sequences to measure from the given fiducials. The interval between shots is 500 microseconds. Each gate sequence is measured with 500 shots with a single-shot readout.

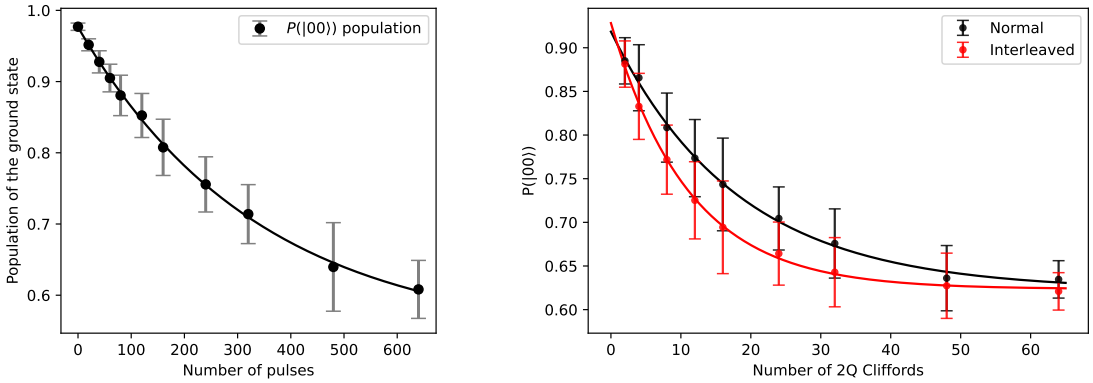
From the gate set tomography, we report the infidelities in Table 5.1. The reconstructed SPAM operators and the PTMs are detailed in Appendix C. From the result, we have very low readout fidelity, mainly due to the short readout time of the quantum amplifier installed in the setup. The error generator for each gate is shown in Appendix C. We found the ϵ_J characterising the amount of incoherent error for qubit B is much lower than qubit A, which agrees with the fact that qubit B has a longer coherence time.

Randomised Benchmarking First, we perform simultaneous randomised benchmarking on both qubits with only a single-qubit Clifford gate. The maximum sequence length is set to 640 Cliffords. It is averaged over 50 different random gate sequences, and each point is sampled 1000 times. The interval between two shots is 1000 microseconds. Then we perform the standard two-qubit Clifford randomised benchmarking and interleaved randomised benchmarking interleaving the $ZX(-\pi/2)$

Gate	$(1-\mathcal{F})\times 10^2$	$\epsilon_J \times 10^2$	$\theta_J \times 10^2$	Operator	$(1-\mathcal{F})\times 10^2$
CNOT	2.0 ± 0.3	2.4 ± 1.4	3.5 ± 0.4	M_0	6.4 ± 0.5
$X_A(\frac{\pi}{2})$	0.26 ± 0.09	0.3 ± 0.8	2.1 ± 0.2	M_1	9.7 ± 0.2
$Y_A(\frac{\pi}{2})$	0.02 ± 0.06	1.9 ± 1.0	2.5 ± 0.5	M_2	14.9 ± 0.2
$X_B(\frac{\pi}{2})$	0.1 ± 0.1	0.1 ± 0.8	1.4 ± 0.2	M_3	17.6 ± 0.2
$Y_B(\frac{\pi}{2})$	1.57 ± 0.06	0.1 ± 0.5	1.7 ± 0.2	ρ_0	7.6 ± 0.2

Table 5.1: Infidelity estimation of the gate set tomography. For each gate in the gate set, we report the average infidelity $(1-\mathcal{F})$, the Jamiołkowski probability (ϵ_J), and the Jamiołkowski amplitude (θ_J). We also report the average infidelity of measurement operators $(1-\mathcal{F})$ and the initial state density matrix. The error bar indicates a 95% confidence interval.

gate. The fidelities are reported in Table 5.2.



(a) (b)

Figure 5.5: Randomised benchmarking results. (a) The single-qubit simultaneous randomised benchmarking on the 2QB device. Single-qubit gates are applied to each qubit in parallel to extract the average single-qubit infidelity. (b) The two-qubit randomised benchmarking and interleaved randomised benchmarking. The $ZX(-\pi/2)$ gate is interleaved to extract its average infidelity.

Benchmarking Results for the 1QD Device

For the single transmon device, we treat it as a 3-level (qutrit) system and a 4-level system (qudit). For the qutrit configuration, we implemented the qutrit Hadamard gate for qutrit Clifford gate synthesis and implemented fiducials and tomography with only qutrit Clifford gates. Then we moved on to the qudit system and used

Parameter	Value
Single-qubit simultaneous Clifford average infidelity	$0.68 \pm 0.08 \times 10^{-3}$
Single-qubit simultaneous physical gate infidelity	$0.37 \pm 0.05 \times 10^{-3}$
Two-qubit Clifford average infidelity	$4.0 \pm 0.3 \times 10^{-2}$
Two-qubit gate $ZX(-\pi/2)$ infidelity	$2.6 \pm 0.4 \times 10^{-2}$

Table 5.2: Infidelity estimation from randomised benchmarking. Here we report the single-qubit randomised benchmarking average single-qubit Clifford gate infidelity and physical single-qubit gate infidelity. We also report the average two-qubit Clifford gate infidelity extracted from two-qubit randomised benchmarking and the $ZX(-\pi/2)$ infidelity extracted from interleaved two-qubit randomised benchmarking.

it to emulate a two-qubit system. We also performed qubit-like randomized benchmarking on all neighbouring computational levels and conducted Clifford randomized benchmarking on the qutrit and emulated two-qubit system.

The 4-level system is treated as a two-qubit emulator for benchmarking. The emulator maps $|0\rangle$, $|1\rangle$, $|2\rangle$, $|3\rangle$ to $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$ states of a two-qubit device. The single qubit gate for the first emulated qubit is achieved by driving transitions in the $\{|0\rangle, |1\rangle\}$ subspace and the $\{|2\rangle, |3\rangle\}$ subspace in sequence. The single qubit gate for the second qubit first drives $\{|1\rangle, |2\rangle\}$ transitions to implement a SWAP gate and then drives transitions the same way as operating the first virtual qubit. Then, transitions in the $\{|1\rangle, |2\rangle\}$ subspace are driven again to swap the state back. See Fig. 5.6.

The 1QD device is tuned up with a 50-nanosecond Blackman DRAG pulse to drive the single-photon transition. For all following experiments, each point is sampled with 500 shots with 1000-microsecond intervals between shots.

Gate Set Tomography Qutrit For qutrit GST, we chose a measurement basis that can be prepared with gates in the qutrit Clifford group shown in Table C.3. While previous physical proposals for tomography use an over-complete basis with 9 measurements, our proposal requires only 4. In the experiment, we used a standard dispersive measurement that can distinguish 3 different states with a single shot

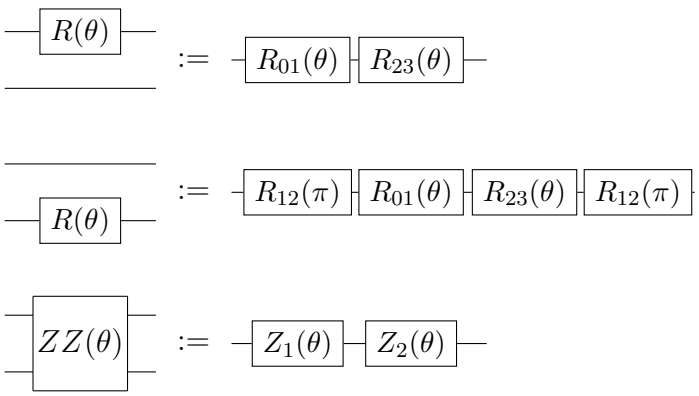


Figure 5.6: Gate sequence for emulating a two-qubit device with a 4-level transmon. The single qubit gate for the first emulated qubit is achieved by driving transitions in the $\{|0\rangle, |1\rangle\}$ subspace and the $\{|2\rangle, |3\rangle\}$ subspace in series. The single qubit gate for the second qubit first drives $\{|1\rangle, |2\rangle\}$ transitions to implement a SWAP gate and then drives transitions the same way as operating the first virtual qubit. Then, transitions in the $\{|1\rangle, |2\rangle\}$ subspace are driven again to swap the state back. The $ZZ(\theta)$ gate is implemented by two qudit virtual Z gates.

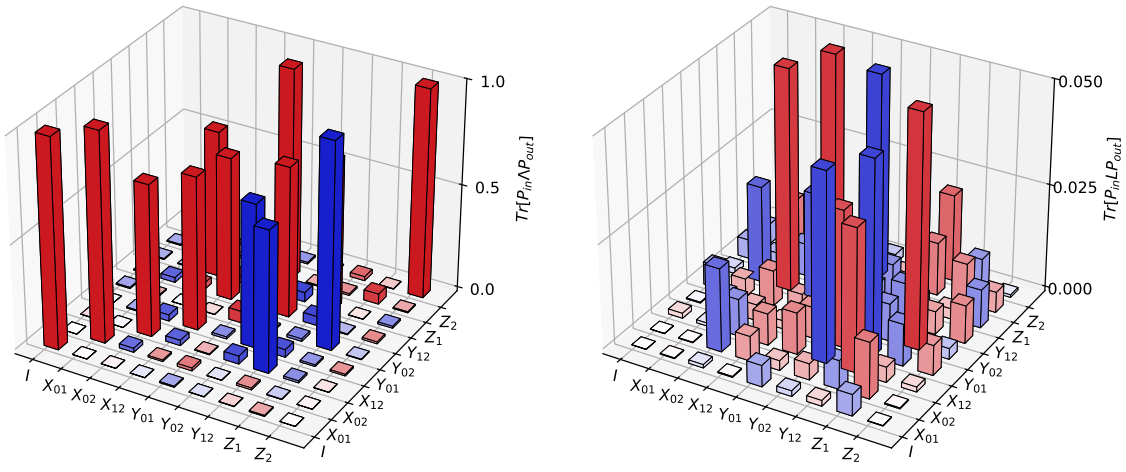
and determine 2 degrees of freedom of the qutrit density matrix. Since a qutrit density matrix has a total of 8 degrees of freedom, we require a minimum of 4 such measurements to implement qutrit state tomography.

The qutrit Hadamard gate H is directly parametrized in the GST gate set, instead of the physical gates from which we synthesise it. The brute-force approach would parameterise $Y_{01}(m)$ and replace H with a sequence of 5 gates. H frequently occurs during the fiducial synthesis; directly parametrising H shortens the gate sequence and reduces the required computational cost.

Since all fiducials are in the qutrit Clifford group, only virtual-Z gates and Hadamard gates are required to implement all the fiducials. The computational cost can be further reduced if we make the reasonable assumption that all virtual-Z gates are perfect. Since we still want to directly characterise the physical gates $X_{01}(\pi/2)$ and $X_{12}(\pi/2)$, we add them to the gate set but do not include them for fiducial synthesis. This implies that H is independent of $X_{01}(\pi/2)$ and $X_{12}(\pi/2)$, and ignores some information to more precisely characterise the $X_{01}(\pi/2)$ and $X_{12}(\pi/2)$. Our experimental results

show that under such configuration, our qutrit GST produces an accurate estimation of the quantum process.

We now proceed and implement the GST on the 1QD device and treat it as a qutrit. The maximum length of the germs sequence is 512. The collected data is then processed by the PyGSTi software with a customised model with the gate set and fiducials configuration described. The estimated parameters of the gates and SPAM operators are reported in Table 5.3. The reconstructed initial state density matrix and the measurement operators are reported in Appendix C. An example of $X_{01}(\pi/2)$ gate is shown in Figure 5.7. The infidelities of the virtual Z gates and Identity gates are at least an order of magnitude lower than the physical gates.



(a) (b)

Figure 5.7: (a) and (b) are the reconstructed PTM and error generator of $X_{01}(\pi/2)$ gate from qutrit gate-set tomography reconstructed. The blue colour denotes negative values and the red colour denotes positive values. For more details, please refer to Appendix C.

Gate Set Tomography Qudit For treating 1QD as a 4-level system,

we implemented a custom pyGSTi model that uses the single-photon transition between neighbouring levels and the virtual Z gate as the intrinsic gates. The gate

Gate	$(1-\mathcal{F})\times 10^3$	$\epsilon_J \times 10^3$	$\theta_J \times 10^3$	Operator	$(1-\mathcal{F})\times 10^2$
I	0.37 ± 0.01	1.5 ± 1.0	0.5 ± 0.2	ρ_0	11.37 ± 0.26
$Z_1(\frac{2\pi}{3})$	0.30 ± 0.08	0.8 ± 5.9	11.1 ± 0.9	M_0	0.4 ± 1.4
$Z_2(\frac{2\pi}{3})$	0.37 ± 0.1	1.0 ± 5.8	16.9 ± 1.4	M_1	9.2 ± 0.4
$X_{01}(\pi/2)$	2.8 ± 0.2	5.5 ± 9.3	56.0 ± 0.5	M_2	8.2 ± 1.4
$X_{12}(\pi/2)$	1.3 ± 0.1	1.8 ± 7.6	44.9 ± 0.9		
H	3.22 ± 0.11	12.4 ± 6.3	27.3 ± 1.5		

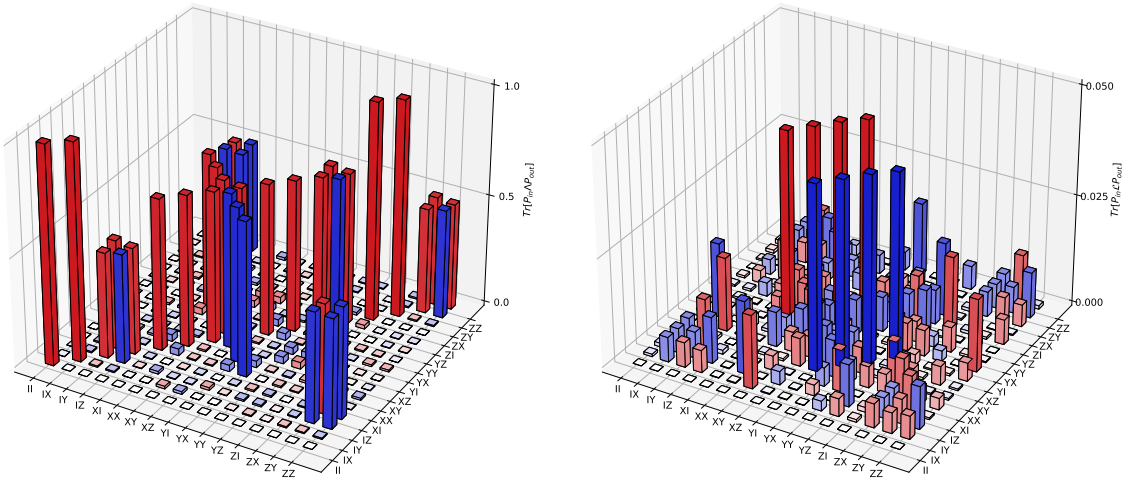
Table 5.3: Infidelity estimation from the qutrit GST. For each gate in the gate set, we report the average infidelity $(1-\mathcal{F})$, the Jamiołkowski probability (ϵ_J), and the Jamiołkowski amplitude (θ_J). We also report the average infidelity of measurement operators $(1-\mathcal{F})$ and the initial state density matrix. The error bar indicates a 95% confidence interval.

Gate	$(1-\mathcal{F})\times 10^3$	$\epsilon_J \times 10^3$	$\theta_J \times 10^3$	Operator	$(1-\mathcal{F})\times 10^3$
$Z_1(\frac{2\pi}{3})$	0.085 ± 0.015	0.04 ± 1.58	8.18 ± 0.18	ρ_0	0.102 ± 0.001
$Z_2(\frac{2\pi}{3})$	0.218 ± 0.007	0.05 ± 1.71	15.05 ± 0.33	M_0	1.2 ± 0.2
$Z_3(\frac{2\pi}{3})$	0.113 ± 0.008	0.07 ± 2.19	8.32 ± 0.19	M_1	0.9 ± 0.2
$X_{01}(\frac{\pi}{2})$	4.98 ± 0.11	5.62 ± 2.62	25.35 ± 0.35	M_2	0.20 ± 0.01
$X_{12}(\frac{\pi}{2})$	2.97 ± 0.05	2.88 ± 1.53	28.88 ± 0.10	M_3	0.1 ± 0.1
$X_{23}(\frac{\pi}{2})$	2.91 ± 0.07	3.43 ± 1.74	14.43 ± 0.10		

Table 5.4: Infidelity estimation from the qudit GST. For each gate in the gate set, we report the average infidelity $(1-\mathcal{F})$, the Jamiołkowski probability (ϵ_J), and the Jamiołkowski amplitude (θ_J). We also report the average infidelity of measurement operators $(1-\mathcal{F})$ and the initial state density matrix. The error bar indicates a 95% confidence interval.

set and germs are detailed in Appendix C. An example of $X_{01}(\pi/2)$ gate is shown in Figure 5.8. The fiducials for state preparation and measurement are the same as the two-qubit configuration but decomposed into intrinsic gates of a 4-level system as presented in Appendix C. The maximum germ length is set to 1024, which gives 19275 sequences for measurement. The estimated parameters of the gates and SPAM operators are reported in Table 5.4.

The results from GST show a lower fidelity for the $\{|0\rangle, |1\rangle\}$ subspace compared to the $\{|1\rangle, |2\rangle\}$ subspace. This is likely due to the choice of LO signal frequency closer to the transition frequency between $\{|1\rangle, |2\rangle\}$. The control pulses' RF signal is mixed with an IF signal generated by a 2 Gsps DAC and a fixed LO signal. To cover



(a) (b)

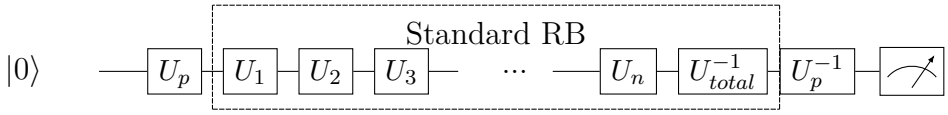
Figure 5.8: (a) and (b) are the reconstructed PTM and error generator of $X_{01}(\pi/2)$ gate from qudit gate-set tomography reconstructed. The blue colour denotes negative values and the red colour denotes positive values. For more details, please refer to Appendix C.

the transition frequency of the $\{|0\rangle, |1\rangle\}$, $\{|1\rangle, |2\rangle\}$ and $\{|2\rangle, |3\rangle\}$, the LO frequency is chosen to be 3.904 GHz, which causes the IF frequency of $\{|0\rangle, |1\rangle\}$ transition to be 230 MHz. 230 MHz is at the very edge of our DAC's operable frequency range. We argue that this can cause the gate in the $\{|0\rangle, |1\rangle\}$ subspace to have lower fidelity than in the $\{|1\rangle, |2\rangle\}$ subspace.

Randomised Benchmarking The qutrit Clifford gates are synthesised with only the qutrit Hadamard gate and virtual Z rotation. Then, both standard qutrit Clifford RB and interleaved qutrit RB are implemented, with the qutrit Hadamard gate interleaved.

To characterise the physical gate fidelity on different subspaces, the subspaces of $\{|0\rangle, |1\rangle\}$ and $\{|1\rangle, |2\rangle\}$ are considered as qubits respectively. Then single-qubit Clifford group randomised benchmarking is implemented on these two-level subspaces. The success probability P_m is renormalised as P_0/P_1 where P_1 is the sum probability

of the used subspaces.



The performance of the emulator is also benchmarked with RB. The single qubit randomised benchmarking is implemented on the emulator for each emulated qubit; see Figure 5.11 (a) and 5.11 (b). The single qubit simultaneous randomised benchmarking result is shown in Figure 5.11 (c). To compare the 2QB device and the emulator, the two-qubit Clifford randomised benchmarking is also demonstrated on the emulator, shown in Figure 5.11 (d). The decomposition of the Clifford group is strictly equivalent to the decomposition of the 2QB device. We report the result from randomised benchmarking in Table 5.2.

Parameter	Value $\times 10^2$
$\{\|0\rangle, \|1\rangle\}$ subspace Clifford gate infidelity	0.19 ± 0.10
$\{\|0\rangle, \|1\rangle\}$ subspace physical gate infidelity	0.11 ± 0.05
$\{\|1\rangle, \|2\rangle\}$ subspace Clifford gate infidelity	0.24 ± 0.11
$\{\|1\rangle, \|2\rangle\}$ subspace physical gate infidelity	0.13 ± 0.06
$\{\|2\rangle, \|3\rangle\}$ subspace Clifford gate infidelity	0.13 ± 0.03
$\{\|2\rangle, \|3\rangle\}$ subspace physical gate infidelity	0.07 ± 0.02
Average Qutrit Clifford infidelity	0.59 ± 0.08
Hadamard gate infidelity	0.6 ± 0.2
Infidelity per Single-qubit Clifford Emulated Q_A	1.9 ± 0.2
Infidelity per Single-qubit Clifford Emulated Q_B	2.9 ± 0.3
Infidelity per Single-qubit Clifford Emulated Q_AQ_B	6.2 ± 0.4
Infidelity per Two-qubit Clifford Emulated Q_AQ_B	9.5 ± 0.7

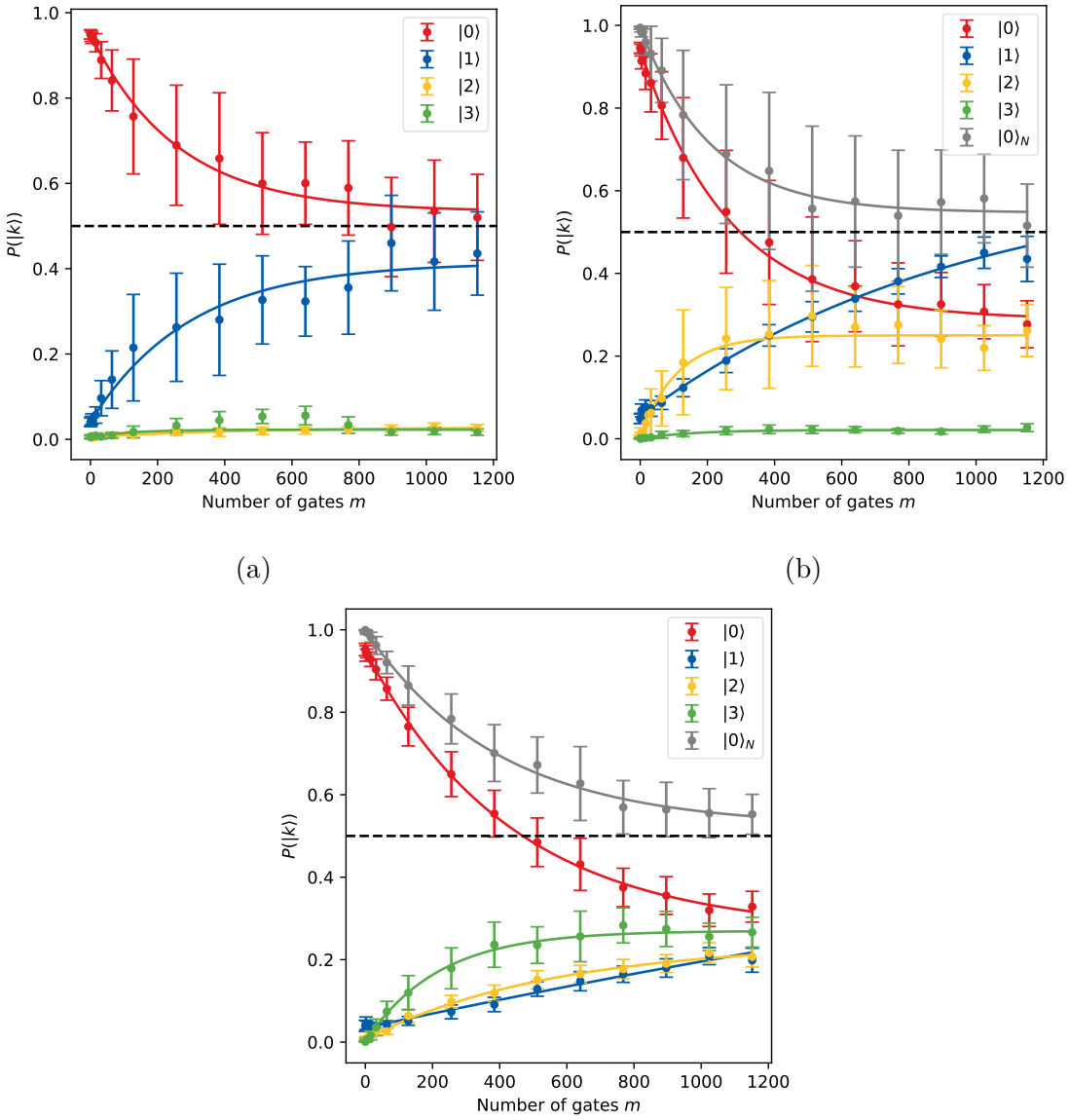
Table 5.5: Infidelity estimation from randomised benchmarking. Here we report the randomised benchmarking result on the neighbouring states of the 1QD device. We then report the average qutrit Clifford gate infidelity and qutrit Hadamard gate infidelity when the 1QD device is treated as a qutrit. We also report the average infidelity of the gates in the two-qubit emulator.

5.4 Discussion

The 1QD device has advantages in implementing specific gates. For instance, the CZ gate can be implemented on the qudit device using only virtual Z gates with high

fidelity. However, to emulate a two-qubit device, the single-qubit operation on the emulator needs to be implemented with multiple physical drives. The benchmarking results show that the single qubit gate operation on the emulator has about 1 magnitude higher infidelity than the physical device. The full two-qubit randomized benchmarking shows about 90% fidelity per Clifford, which is also significantly lower than the 2QB device. We can conclude that to get the most benefit from the multi-level system, the algorithm needs to be redesigned to fit the physical implementation of the device. The emulator cannot perform as well as the 2QB device.

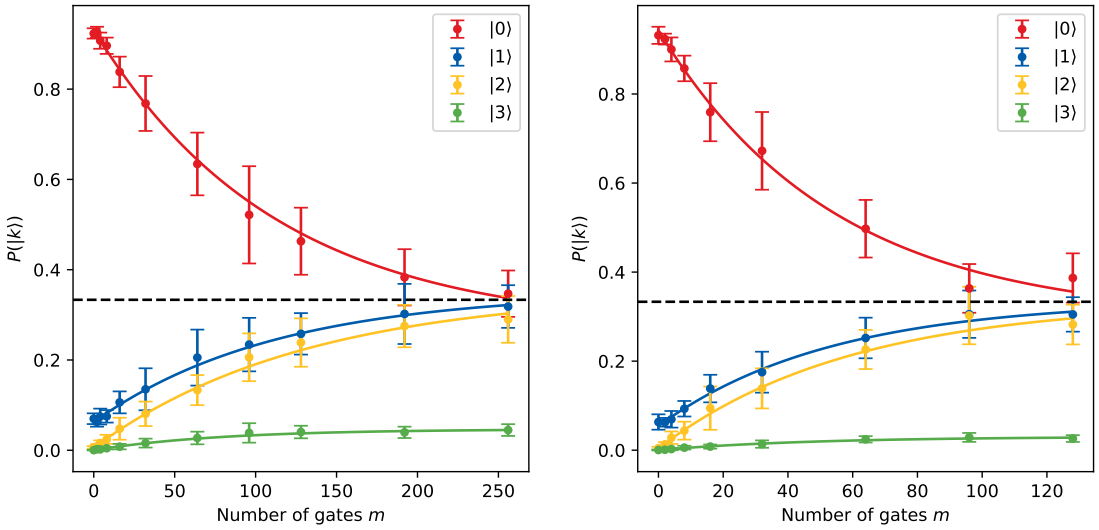
In addition, we have demonstrated using gate set tomography to characterise a single superconducting qutrit. The fiducials are synthesised with only the qutrit Hadamard gate and qutrit virtual Z gate to speed up the optimisation process of gate-set tomography. We also choose an efficient measurement basis to use the minimum measurement number to implement tomography. The above configuration significantly reduces the computation costs compared to previous proposals. The GST results show that coherent error is dominant for the physical gates, and the virtual Z gates have infidelity two orders of magnitude lower than the physical gates. We compare the GST results with the randomized benchmarking results and find a good agreement between characterised infidelity from the GST and RB.



(a) (b)

(c)

Figure 5.9: Population of each state versus the number of gates in neighbouring two-level randomised benchmarking. (a) to (c) are two-level randomised benchmarking results in $\{|0\rangle, |1\rangle\}$, $\{|1\rangle, |2\rangle\}$, $\{|2\rangle, |3\rangle\}$ subspaces, respectively. (b) (c) includes the renormalised population of the ground state, which makes the probability of the used subspaces sum up to 1.



(a) (b)

Figure 5.10: Population of each state versus the number of gates in neighbouring qutrit randomised benchmarking. (a) Standard Qutrit randomised benchmarking results. (b) Interleaved Qutrit randomised benchmarking results. The interleaved randomised benchmarking interleaves the qutrit Hadamard gate to extract its infidelity.

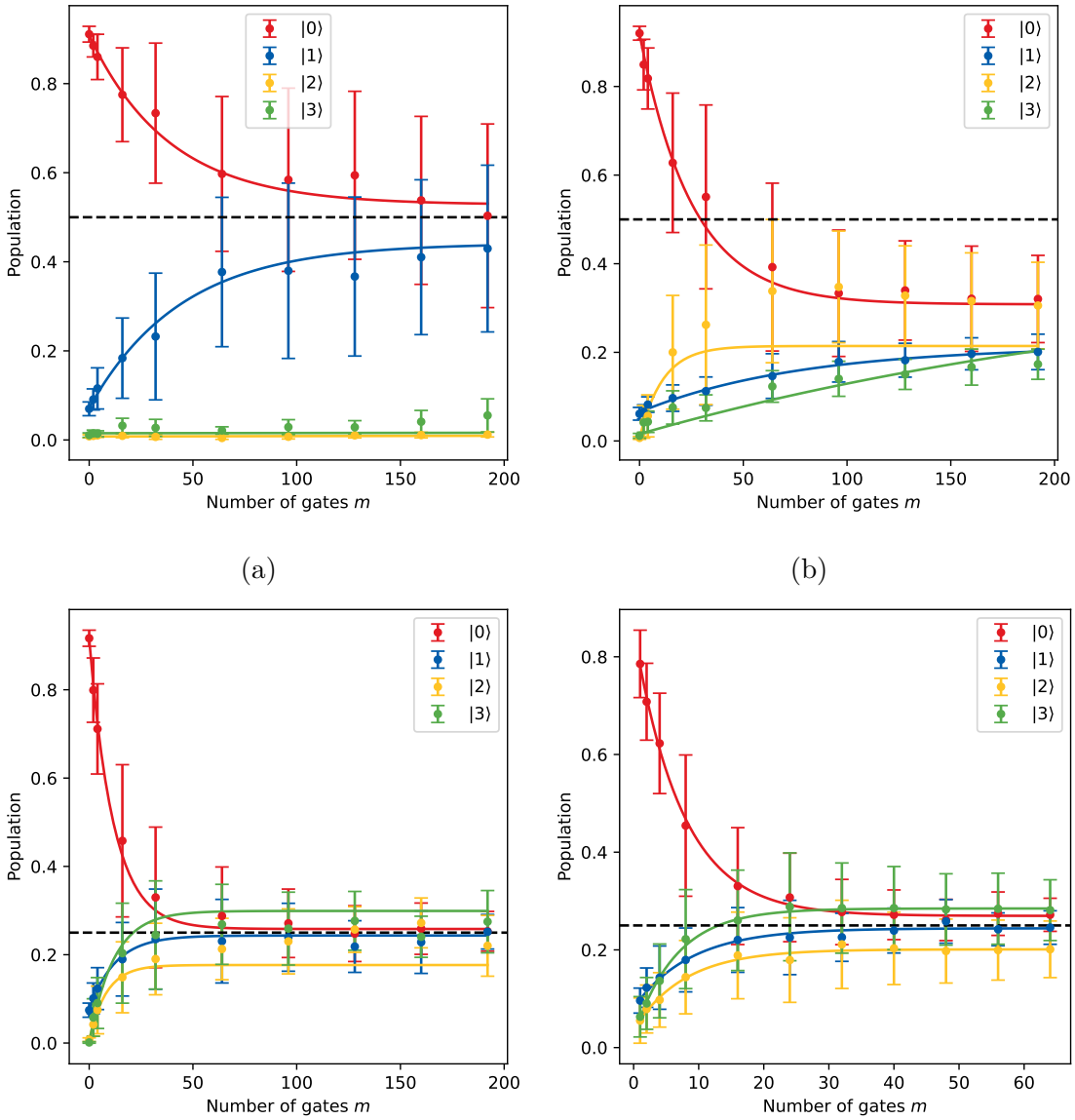


Figure 5.11: Population of each state versus the number of gates in the two-qubit randomised benchmarking on the emulator. (a) and (b) are randomised benchmarking results on the first and the second emulated qubit. (c) Single qubit simultaneous randomised benchmarking of both emulated qubits. (d) Two-qubit Clifford gate randomised benchmarking on the emulator.

Chapter 6

High-fidelity qudit ground state preparation by active feedback

This chapter explains the implementation of qudit active reset and presents experimental data to enhance transmon initial state fidelity. The content is entirely my original work. While active reset has been demonstrated, extending it to a 4-level qudit and implementing multilevel active reset is novel. I proposed the 4-state discrimination mechanism using the nearest neighbour method and implemented it on the FPGA, based on the existing firmware developed by Andrew Patterson. I conducted the experiments and performed data analysis.

Most quantum algorithms start with a pure state $|0\rangle^{\otimes N}$. A quantum computer must be able to prepare such an initial state with high fidelity. In superconducting circuits, the initial state is typically prepared by passively waiting for qubits to decay to their ground state and equilibrate with their thermal environment. There are two problems with this method. First, the temperature of the superconducting qubits cannot reach absolute zero; therefore, according to the Boltzmann distribution, a small population always occupies the excited states. Second, when the qubit coherence time has been significantly improved, it takes much longer to wait for the qubit to decay, considerably reducing the experiment's repetition rate and slowing the computation speed. Therefore, other mechanisms have been developed to prepare the initial state of the qubit.

This section discusses the implementation of quantum feedback control: active reset. Active reset prepares the qubit state into the ground state by repeatedly measuring the device and sending a pulse to move the state to the ground state, depending on the measurement results. Active reset has first been demonstrated on transmon qubits [96] and fluxonium qubits [165]. Active reset has also been demonstrated on transmon qudits [79]. In this chapter, we describe the method and gate sequences of implementing active reset and discuss our design of the FPGA firmware to implement active reset. We report the prepared initial state characterisation results with and without active reset.

6.1 Methods

The implementation of the active reset contains two parts. First, the readout signal must be processed on the FPGA, and state discrimination must be implemented. We use the lowest 4 levels of a transmon as a qudit; therefore, the transmon state could be any of the lowest 4 levels, and the single-shot readout must distinguish all 4 different states. Second, depending on the state of the transmon, a different reset pulse needs

to be sent to the transmon. Here we implement the conditioned pulse by playing a waveform conditioned on the measurement result of the transmon.

The framework for implementing active reset is shown in Figure 6.1. To implement the reset, first, we need to do the resonator spectroscopy, measurement calibration, and single qudit gate calibration. These calibrations are referred to in Appendix A. In this section, we have a detailed discussion about the readout procedure and then focus on discussing the implementation of active reset on the FPGA.

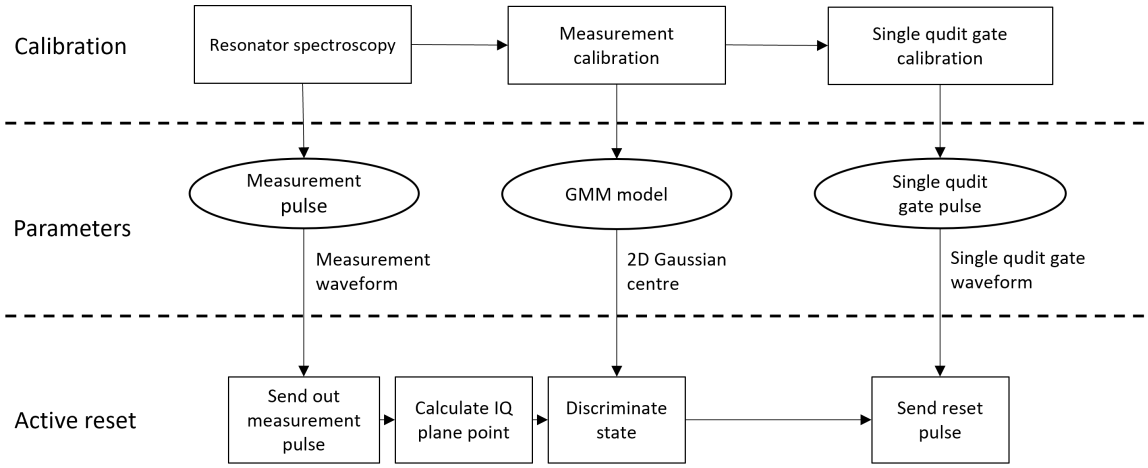
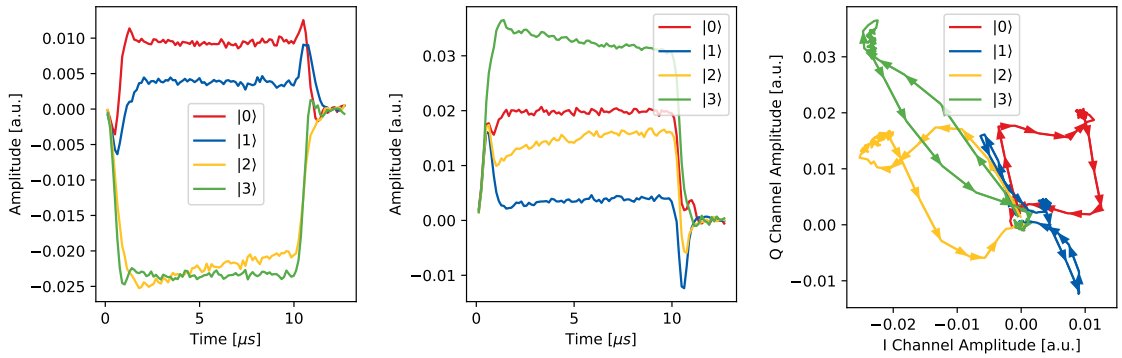


Figure 6.1: The framework of implementing active reset. The first row denotes the calibration experiments that need to be implemented before the active reset. The arrow between each experiment denotes the order of executing the calibration. The elliptic circle in the second row denotes the results generated from the calibration. The text next to the arrow denotes the information passed to the FPGA board for active reset. The third row denotes the procedure of implementing the active reset on FPGA. The arrow between the parameters and steps in the procedure denotes that these parameters are required. The arrow between the steps denotes the dependence of each step.

Device readout

A detailed investigation of the readout signal is required to have a high-fidelity readout for active reset. In this experiment, we set the readout frequency of the 1QD device to 8782.41 MHz to distinguish all four states with a single shot. The frequency comes from the calibration experiment, which sweeps the readout frequency around



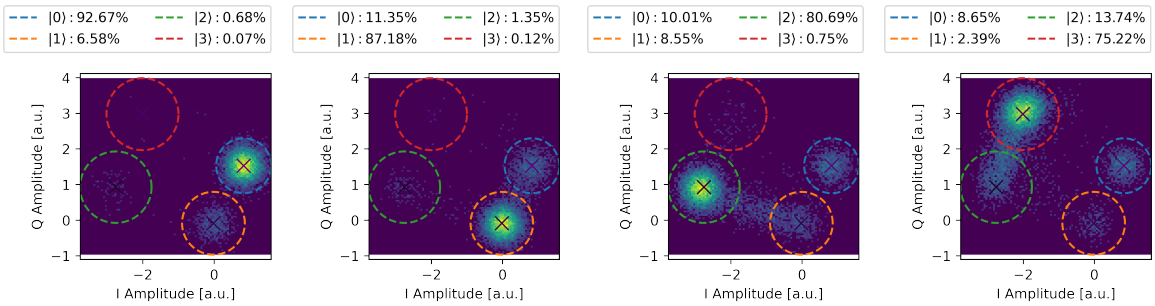
(a) (b) (c)

Figure 6.2: Demodulated readout signal versus transmon prepared in a different state. A 10μs soft square readout pulse with 0.5μs rise and drop sent to the qubit and the reflected signal from the resonator is demodulated and recorded. The signal is averaged over 10,000 shots. (a) I channel and Q channel trajectories of different transmon states in response to the readout pulse, respectively. (b) IQ-plane trajectories of different transmon states in response to the readout pulse, respectively.

the resonator resonance frequency and finds the frequency with the best SNR distinguishing all 4 states. The IF frequency is 8657.41 MHz, and the IF frequency is 125 MHz. The received signal is averaged among 10,000 shots, and the signal trace after demodulation is shown in Figure 6.2.

From the plot, we can observe that the readout signal moves from the origin, which denotes that both I and Q channels have no signal, to four locations at a steady state. Then it returns to the origin point after the readout signal is cut off. To distinguish the four different states, the sum of the trace signal of a single shot is calculated on the FPGA, which yields a single point on the IQ plane. The distribution of the points with different transmon states is shown in Figure 6.3.

The Gaussian Mixture Model (GMM) identifies the centre of each distribution and uploads it to the FPGA board. With the approximation of the standard deviation equal for each distribution, the point is assigned to the class with the closest distance to the centre of the GMM model. The FPGA is programmed to discriminate the qubit's state by comparing the distance between the measured signal on the IQ plane



(a) (b) (c) (d)

Figure 6.3: Single-shot measurement signal plotted on IQ plane. (a) to (d) are the heat maps of the single-shot measurement signal plotted on an IQ plane for state prepared to $|0\rangle$ to $|3\rangle$, respectively. The cross in the centre of the dashed line circle is the centre of the 2D Gaussian distribution fitted from the GMM model. The circle's radius denotes 3σ of the distribution, where σ is the standard deviation of the Gaussian distribution. The covariance matrix of the GMM model is chosen to be diagonal. Therefore, each Gaussian distribution is homogeneous on all dimensions, and the contour of the 3σ lines is strictly a circle. The percentage value in the legends for each colour denotes the population of transmon discriminated into each state by the GMM model.

and the centre of the four-state distribution.

Reset pulse

Once the state is discriminated, a reset pulse is sent to the device. In this implementation, we used a masking technique to play the waveform. A sequence of π pulse, $X_{\pi 23}$, $X_{\pi 12}$ and $X_{\pi 01}$ is pre-loaded to the DAC memory. A "mask" recording the start and the end of every single qubit π pulse is uploaded to the FPGA. When the measurement result is in state $|0\rangle$, all waveforms within masks will be zeroed, resulting in playing an empty waveform. If the measurement result is in state $|1\rangle$, waveform lies in the mask of $X_{\pi 23}$, $X_{\pi 12}$ will be zeroed to play empty waveform, and $X_{\pi 01}$ would be played as is. The same applies to $|2\rangle$ and $|3\rangle$. See Figure 6.4.

It is worth noticing that the resonator needs to return to its ground state before the reset pulse is sent to the qubit. This is because the readout pulse populates a large number of photons in the resonator and induces a significant shift in the transmon

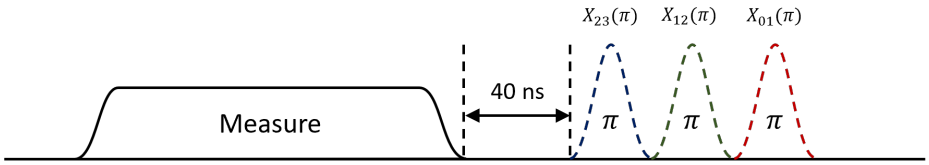


Figure 6.4: The reset pulse sequence of the active reset procedure. A waveform includes the measurement pulse and three π pulses $X_{\pi 23}$, $X_{\pi 12}$ and $X_{\pi 01}$ are preloaded to the FPGA. These three pulses will be played conditioned on the discriminated state of the transmon. The interval between reset and measurement pulse needs to be a minimum of 40 ns to allow the FPGA to finish the state discrimination.

transition frequency. The calibrated single qubit pulse won't be able to exactly drive a π rotation on the transmon if there are remaining photons in the resonator. The resonator population can be cleared with specifically designed pulse shapes [177]; in our setup, we utilise a soft-square pulse with 1 μs rise and drop edges to drive the resonator, plus a 0.5 μs gap before the reset pulse. This time

interval is much longer than the lifetime of the resonator[1] This configuration leaves no residual photon in the resonator that significantly shifts transmon transition frequency after the measurement pulse.

Implementation

The active reset firmware is integrated into the control firmware of our electronics setup. Our electronic setup for IF frequency generation consists of 5 Abaco PC821 boards with Xilinx Kintex UltraScale FPGA. One PC821 board provides the clock and trigger synchronisation between all boards, and all other four have an FMC230 board with 2.4 Gsps sampling rate DAC for qubit signal generation and an FMC110 board with 1.2 Gsps sampling rate ADC and DAC for readout pulse generation.

The FPGA firmware mainly comprises 4 modules; see Figure 6.5. The demodulator module accepts the sampled voltage data from the ADC and demodulates the signal into a trace of the I channel and Q channel. The trace is then sent to the

[1]The resonator lifetime has a similar definition as a qubit, which is defined as the time that the energy stored in the resonator decays to $1/e$ of the initial value. In our case, the lifetime of this resonator is approximately 300 ns.

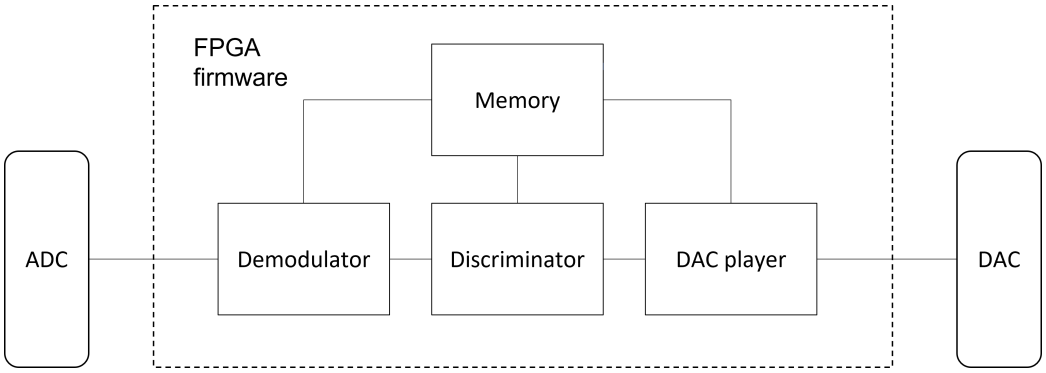


Figure 6.5: Simplified FPGA firmware block design. The firmware comprises 4 major components. The demodulator demodulates the signal sampled from ADC and returns signal traces of the I and Q channel. The discriminator accumulates the trace, generates the point on the IQ plane, and compares it with predefined centres of Gaussian distribution. The discriminator outputs the closest state label to the DAC player. The DAC player plays the reset waveform conditioned on the measurement result from the discriminator.

discriminator module. This module calculates the sum of the IQ trace and evaluates the distance between the sum point to the centre of four different Gaussian distributions. Then it compares the distance and outputs the discriminated state to the DAC player module. The DAC player module plays the uploaded waveform in the on-chip memory. When the readout discrimination is finished, it takes the result and conditionally plays the waveform for the reset pulses. The whole state discrimination takes 40 ns to finish; therefore, it has to be a minimum 40 ns interval between the readout pulse and the reset pulse.

6.2 Results

The active reset experiment is tested on the 1QD device. The single qubit randomized benchmarking shows the gate fidelity for the qubit-like gate in the $\{0, 1\}$ subspace is $\mathcal{F}_{01} = 99.91 \pm 0.09\%$, for the $\{1, 2\}$ subspace is $\mathcal{F}_{01} = 99.86 \pm 0.08\%$, and for the $\{2, 3\}$ subspace is $\mathcal{F}_{23} = 99.38 \pm 0.003\%$. The readout pulse shape is a soft-square pulse with a 0.2 microsecond rising and dropping time and a 10 microsecond flat top. First, the reset is applied multiple times; see Figure 6.6. There is an improvement after each

reset up to 3 resets, with which we obtain a preparation fidelity of 99.11%. Four or more resets were found not to improve this any further; see Figure 6.7. From this result, we chose to apply the reset three times as the standard procedure to prepare a high-fidelity initial state.

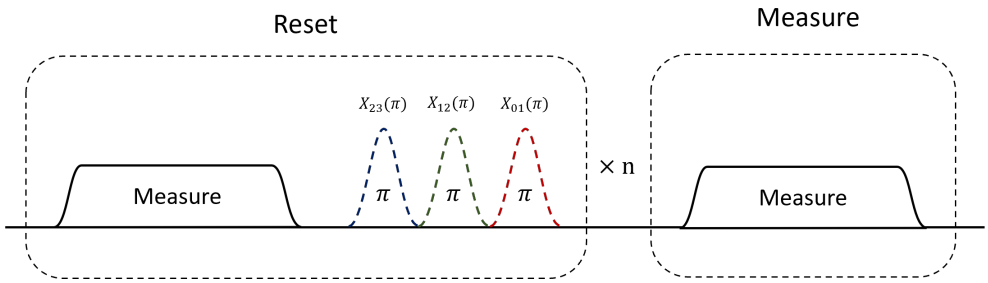
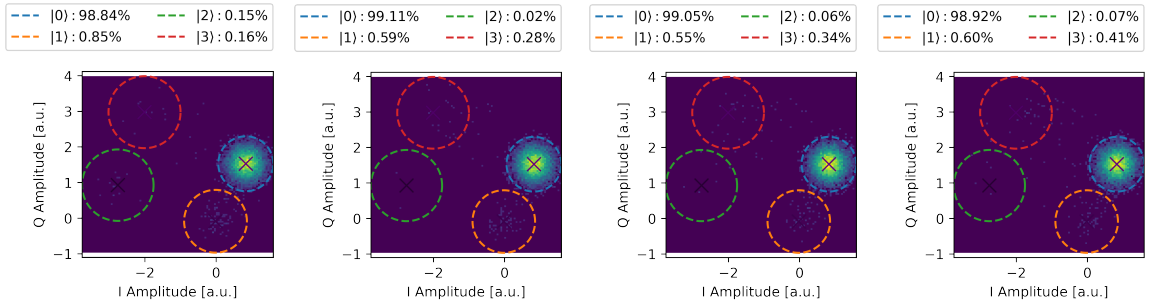


Figure 6.6: Pulse sequences for testing multiple resets. The reset sequences are the same as Figure 6.4, which is repeated multiple times to reset the transmon to its ground state. Then another measurement pulse is applied to measure the transmon state.



(a) (b) (c) (d)

Figure 6.7: Single-shot readout signal scatter plot, with respect to different times of active reset, is performed. The GMM model is the same as the calibrated readout in Figure 6.3, and the pulse sequence is referred to in Figure 6.6. (a) to (d) is the scatter plot of reset 1 time to 4 times.

Three readout-and-reset instructions are applied to the transmon at the beginning of each gate sequence for the following benchmarking of the active reset; see Figure 6.8. The readout assignment matrix, which describes the probability of the transmon being prepared and measured in a different state, is shown in Figure 6.9. It is shown that the correct assignment probability described in the diagonal element is significantly increased after applying the active reset.

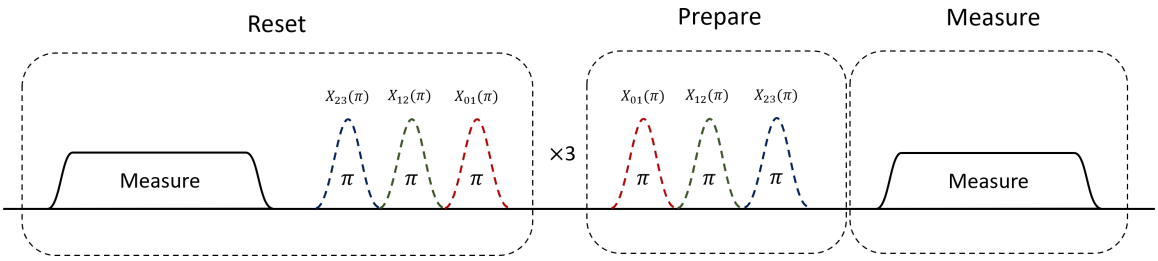


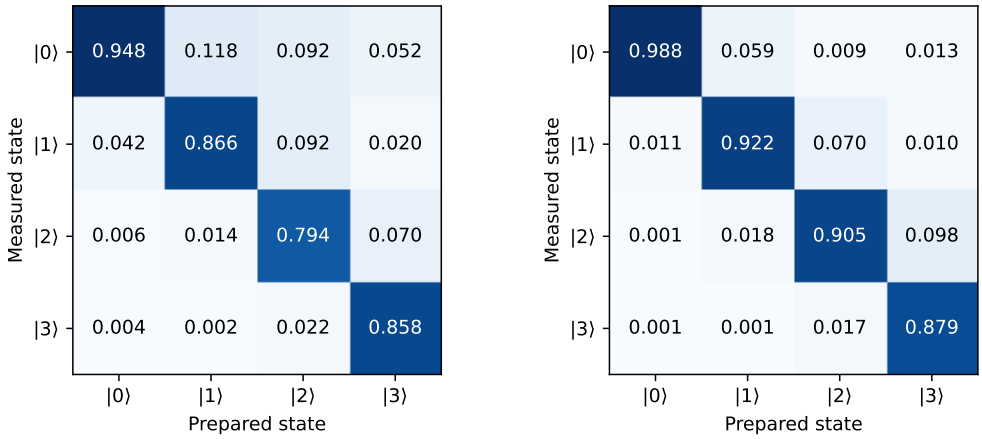
Figure 6.8: Pulse sequences for measuring the assignment matrix. When the active reset is turned on, the reset sequence is repeated three times, and then the transmon is prepared in the $|0\rangle$ to $|3\rangle$ state, respectively. Then, in the end, another measurement pulse was used to find the qubit population. The assignment matrix without reset is extracted with the sequence without the reset sequences.

The readout assignment matrix could not distinguish between the readout errors and initial state preparation errors. Therefore, we applied the GST to characterise the initial state, which results in Figure 6.10(b). From the plot, we found the initial state fidelity is improved from 0.900 ± 0.011 to 0.9932 ± 0.0013, and the effective temperature of the transmon is cooled from 88.82 mK to 37.21 mK.

6.3 Discussion

The active reset technique has been successfully demonstrated on the coaxmon architecture. As a simple application of quantum feedback control, the active reset includes the discrimination of multiple energy levels, which prepares the high-fidelity initial state for multilevel transmon and paves the path toward the applications of quantum algorithms on multilevel transmon in the next chapter. Issues and potential improvements could still be made to the active reset procedure.

First, the readout takes a significant amount of time (10 μs) to implement, making energy relaxation a significant contributor to errors. This is considered a major drawback of the active reset, which causes the readout fidelity to be relatively low. To alleviate this problem, quantum parametric amplifiers could be installed in our setup to improve the SNR [11]. However, our readout frequency band of 8-11 GHz is higher than the typical working frequencies of parametric amplifiers demonstrated



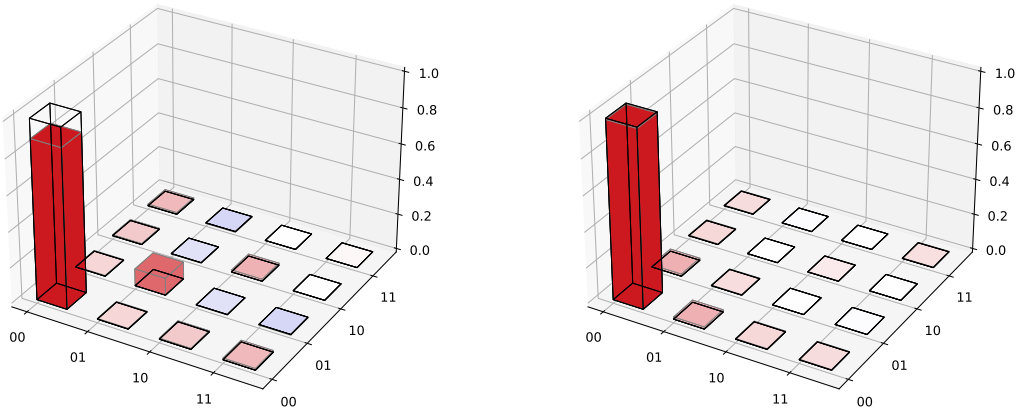
(a) (b)

Figure 6.9: The extracted assignment matrix with and without active reset for the 1QD device. The assignment matrix element A_{ij} denotes the probability that we tried to prepare the state in $|i\rangle$ but measured in $|j\rangle$. When the state preparation and readout assignment are ideal, it should be an identity matrix. (a) The assignment matrix was extracted without active reset. (b) The assignment matrix was extracted after repeating the active reset 3 times before preparing the state and measurement. The pulse sequence for extracting the assignment matrix is explained in Figure 6.8.

to date. On the other hand, the multi-tone readout method [47] could be applied to implement a single-shot readout with a shorter readout pulse. This method requires the FPGA to demodulate multiple frequencies.

Secondly, the initial state might be better prepared by continuously resetting it until the readout indicates it has reached the ground state rather than repeating reset a specific number of times. This method is not chosen because it requires multiple FPGAs to communicate and agree on the time to finish the reset, and our system requires extra effort to implement it. This strategy is certainly an improvement on the current approach.

Thirdly, a straightforward strategy that measures the initial state at the beginning and post-selects the ground state after the experiment can also be used to effectively prepare a high-fidelity initial state. If we are not demonstrating quantum feedback control, the post-selection method would be a straightforward approach to obtaining a



(a) $1 - F_{\rho_0} = 0.099 \pm 0.011$ (b) $1 - F_{\rho_0} = 0.0068 \pm 0.0013$

Figure 6.10: (a) and (b) are reconstructed ρ_0 without and with active reset, respectively. As the error in the assignment matrix comprises both the error of preparation and readout, here we use GST to reconstruct the initial state ρ_0, which is independent of the readout error.

high-fidelity initial state; however, it requires more shots to achieve the same sampling accuracy.

Chapter 7

Quantum Algorithms on a Transmon Qutrit and Qudit

This chapter demonstrates two variational quantum algorithms on the qudit device. Section 7.1 details an experiment implementing a qutrit-based quantum classifier. It is a collaborative effort with the National Physics Laboratory (NPL). Ivan Rungger and Weixi Zhang proposed the original idea, while I introduced innovations in the trainable hybrid encoding method. I wrote the code, and collected and analysed the data for this experiment. The results will be published in a journal article with NPL. Section 7.2 employs a two-qubit emulator on a 4-level transmon to execute a variational quantum eigensolver for solving the hydrogen molecule energy. The idea is originally from me, and I wrote the experimental code, conducted the experiment, and analysed the data. The results have been published in Ref. [40].

Having shown that we can implement high-fidelity preparation, measurement and universal control of transmon qudits, we can now proceed to implement quantum algorithms on them. In this chapter, we demonstrate two quantum variational algorithms utilising more than two energy levels of a single transmon. In the first experiment, we use the 1QD device as a qutrit and implement a variational quantum classifier. The second experiment utilises four transmon levels and emulates the behaviour of a two-qubit system. This emulated two-qubit device is used to demonstrate the variational quantum eigensolver of a hydrogen molecule.

7.1 Qutrit Iris Classifier

Machine learning is seen as a promising application of quantum computing. A detailed review of quantum machine learning can be found in chapter 2. The Iris flower dataset [7, 89] is a widely used dataset for benchmarking machine learning algorithms. It contains "setosa", "versicolor", "virginica", the three different species of the Iris flower. Each species has 50 samples. Each sample is a four-dimensional vector containing the sepal length, sepal width, petal length and petal width. Previous works have implemented the quantum Iris classifier [103, 230] on qubits. However, these methods either use more than one qubit or compress the data in advance. This is because a single qubit state has only two degrees of freedom which cannot encode all four dimensions of the data point without dimension reduction. A qutrit provides exactly four degrees of freedom to encode the whole Iris dataset.

Qutrit Quantum Circuit as a Classifier

A quantum classifier usually contains three components: First, an encoding component that translates the classical data into a quantum state. Second, the processing component manipulates the input of the quantum state with a parameterised circuit. The third is the discrimination component, where the qutrit is measured, and the

decision is made based on the obtained measurement outcome.

The gate set on the quantum device needs to be chosen before implementing these three components. First, we chose the qutrit virtual Z gate as part of the gate set. As an extension of the qubit virtual Z gate [180], the qutrit virtual Z gate is not a physical pulse but a shift of the phases of all following pulses. For more details of the implementation of qutrit virtual Z gates, see section 3.2.

$$Z_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad Z_2(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\theta} \end{pmatrix} \tag{7.1}$$

We would like to use a rotosolve-like optimisation method, so every gate needs to have exactly two different eigenvalues [190, 201]. However, directly driving a transition between two levels may introduce a three eigenvalue unitary[1]. Therefore, we define the following gates as the basic building blocks, where the parameterised gate has a maximum of two distinct eigenvalues:

$$U_{Y01}(\theta) = e^{i\frac{\theta}{2}\lambda_2}, \tag{7.2}$$

$$U_{Y12}(\theta) = e^{i\frac{\theta}{2}\lambda_7}, \tag{7.3}$$

$$H_1 = Z_1(\pi)U_{Y01}(-\pi/2), \tag{7.4}$$

$$H_2 = Z_2(\pi)U_{Y12}(-\pi/2), \tag{7.5}$$

$$X_{01}(\theta) = H_1 Z_1(\theta) H_1, \tag{7.6}$$

$$X_{12}(\theta) = H_2 Z_2(\theta) H_2, \tag{7.7}$$

where λ_2 and λ_7 are Gell-Mann matrices (see equation 5.7). H_1 and H_2 are two-level Hadamard gates on the $\{|0\rangle, |1\rangle\}$ subspace and the $\{|1\rangle, |2\rangle\}$ subspace respectively.

The above physical single qutrit gates are implemented with 100 ns microwave pulses with DRAG correction and Blackman window, slightly different from the standard configuration in the benchmarking and tomography. The qutrit readout is a

[1]The eigenvalues for $e^{i\frac{\theta}{2}\lambda_1}$ are $0, \frac{\theta}{2}, \text{and} -\frac{\theta}{2}$

standard dispersive readout for transmon devices and was recalibrated for this specific experiment. The readout pulse is a soft square pulse with a total width of 7 microseconds, including 1 microsecond rise and drop. The readout frequency is calibrated to give the best signal-to-noise ratio of three different states.

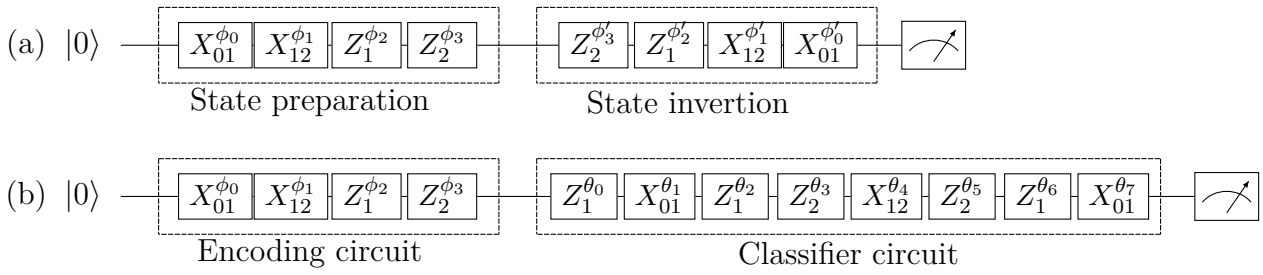


Figure 7.1: (a) Quantum circuit measuring overlap between two mixed states. The ground state population gives the estimation of quantity $\langle 0| U^\dagger(\phi')U(\phi) |0\rangle$. The mixed state overlap can be estimated by randomly sampling the ϕ and ϕ' from the corresponding dataset. (b) Quantum circuit implementing classification. After the data features are encoded into the quantum state, it is manipulated with a parameterised circuit so the state overlap with its label $|y\rangle, y \in \{0, 1, 2\}$ is maximised.

To implement a qutrit quantum classifier, we start by constructing the encoding component. Here is a parameterised circuit (See Figure. 7.1 (a)) to encode the data point into a quantum state.

$$|\phi\rangle = U(\phi) |0\rangle \tag{7.8}$$

And the ϕ value is given by

$$\phi = W\mathbf{x} + \mathbf{b} \tag{7.9}$$

where W is a four by four matrix, and $\mathbf{b}$ is a four-dimensional vector. After the encoding, the density matrix representing the whole class of the dataset is

$$\rho_i = \frac{1}{m_i} \sum_k |\phi_k\rangle \langle\phi_k|, x_k \in X_i \tag{7.10}$$

where m_i is the total data point number in class X_i. The value of W and $\mathbf{b}$ is trained to minimise the loss function $\mathcal{L}_e$.

$$\mathcal{L}_e = \sum_i \sum_{j,i\neq j} Tr(\rho_i \rho_j)^2 - \sum_i Tr(\rho_i \rho_i)^2 \tag{7.11}$$

The first term of $\mathcal{L}_e$ characterises the state overlap between two different categories, and the second term characterises the purity of the state in the same category. The value of $Tr(\rho_i \rho_j)$ can be estimated with quantum circuit 7.1 (a).

$$Tr(\rho_i \rho_j) = \frac{1}{m_i m_j} \sum_i \sum_j \langle \phi_i | \phi_j \rangle \tag{7.12}$$

$$= \frac{1}{m_i m_j} \sum_i \sum_j \langle 0 | U^\dagger(\phi') U(\phi) | 0 \rangle \tag{7.13}$$

The value of ϕ and ϕ' are evaluated from randomly sampled $x \in \mathcal{X}_i$ and $x' \in \mathcal{X}_j$, the estimated overlap is the probability measured in the ground state.

The processing component processes the encoded state with another parameterised quantum circuit. The circuit diagram is shown in 7.1 (b). And the discrimination assigns the label with the measured state with maximum probability. The classified output would be the state label with maximum probabilities. The parameter of the circuit is optimised to give the minimal value of loss function $\mathcal{L}_c$, given by

$$\mathcal{L}_c = 1 - P(\text{state with correct label}) \tag{7.14}$$

We intentionally chose a linear cost function and chose a gate set in which the generator has only two eigenvalues to apply a rotosolve-like optimisation method [190, 201]. See details from section 2.3.

Experiment Result

The experiment was implemented as follows. The Iris dataset is first split into training and testing sets with 100 and 50 samples. All the initial parameters for the circuit

parameterisation angle are randomly initialised, and the classical encoding circuits are initialised with a standard Gaussian distribution where standard deviation $\sigma = 1$.

The SPSA optimiser [246] is used to optimise the encoding training in this experiment[2]. 500 samples are randomly sampled with repetition from each class and paired up as the parameter for the overlap estimation circuit to estimate the distribution overlap. Then ten shots were executed for each pair of samples, and the overall ground state probabilities were used to estimate the state overlap. Here we show the estimated overlap and purity in figure 7.2 and figure 7.3.

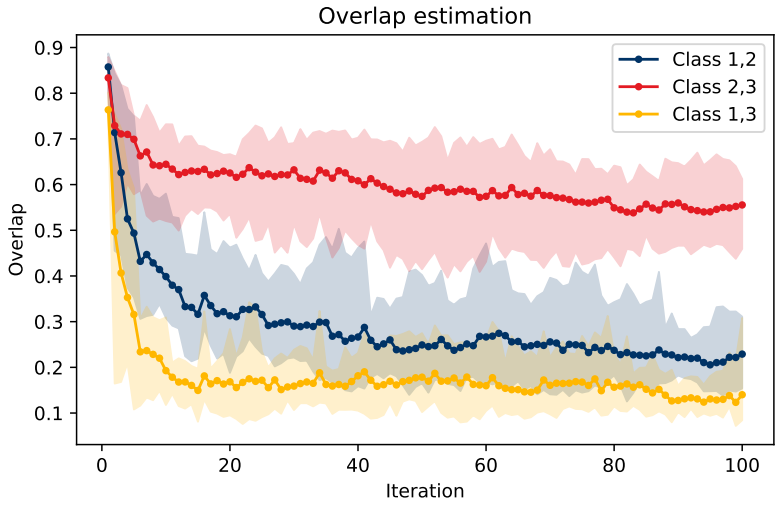


Figure 7.2: Estimated overlap $Tr(\rho_i \rho_j)$ of class i, j during the training of the encoder parameter $\mathbf{W}$. The encoding performance during the encoding training process. Here the upper error bar and lower error bar are the estimated value of two SPSA perturbations, and the estimated overlap is the average of two perturbed values.

For the classification training step, we modify the parameters for the classifier circuit one at a time. The loss function value $\mathcal{L}_e$ is estimated by evaluating the loss function for each data point in the training dataset with ten shots and taking the average. Here we iteratively sweep each parameter of the ansatz while fixing the others. Each sweep samples 16 points from 0 to 2π and fits the cost value with a sinusoidal function. Once the function is fitted, we could observe the function's

[2]We use *noisyopt* package, which provides an implementation of SPSA. The choice of parameters is $a = 0.2$, $c = 0.1$, $\alpha = 0.15$, $\gamma = 0.101$

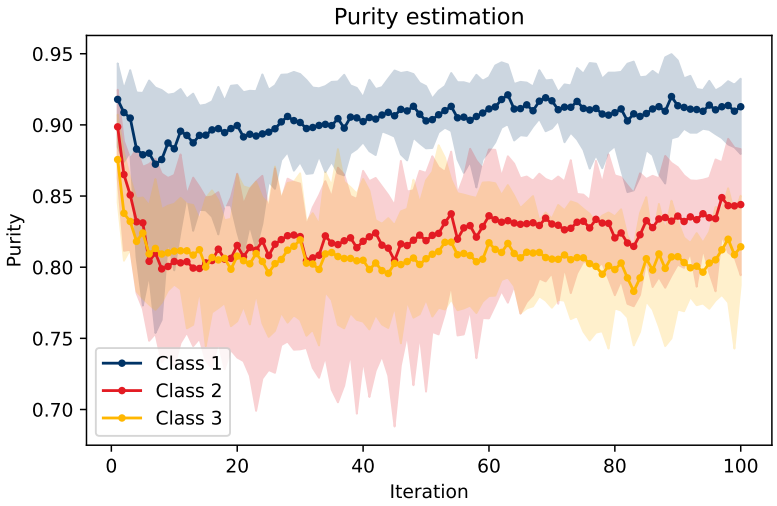


Figure 7.3: Estimated class purity of $Tr(\rho_i \rho_i)$ of class i during the training of the encoder parameter **W**. The error bar and colour scheme are the same as figure 7.2.

minimal value and choose that point as the selected value for the current parameter of the current iteration. See figure 7.4. After each fitting and updating a parameter to the current optimum value, we rigorously evaluate the classifier's performance with 500 shots. The classifier's performance versus the rotosolve steps is shown in figure 7.5.

After implementing the qutrit classifier, 8 hardware experiments were executed with different random seeds for training testing dataset partitioning and parameter initialisation. Here we show the accuracy result of training the classification circuit in Figure 7.5 and Table 7.1.

The experiment resulted in a test accuracy of 0.94 ± 0.04, while the baseline using Logistic Regression [146] to classify the dataset achieved an accuracy of 0.9487. Although the single qutrit quantum classifier did not significantly outperform the classical classifier, it still had comparable accuracy.

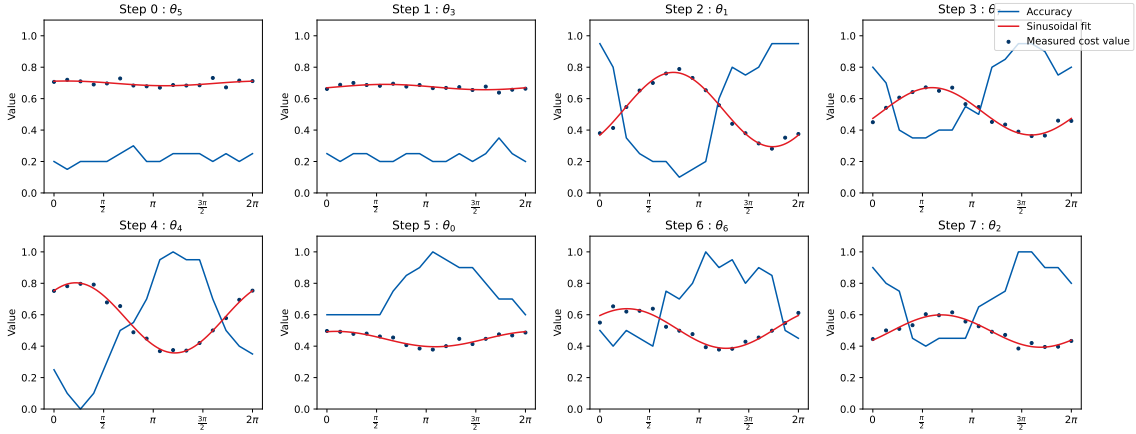


Figure 7.4: A typical rotosolve-like optimisation iteration during the training process. The optimisation algorithm sweeps each parameter from 0 to 2π and fits the result with a sinusoidal function. The parameter is modified to give the minimum of the fitted function. The blue line plots the estimated classification accuracy on the training dataset.

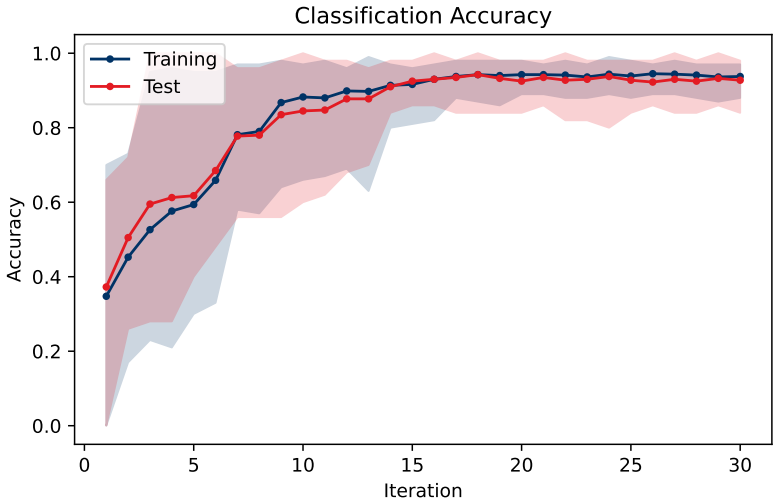


Figure 7.5: Classification accuracy versus rotosolve steps. The training process is implemented 8 times with different random seeds. The solid line shows the averaged accuracy among all experiments, and the shadowed area shows the maximum and minimum accuracy of all experiments.

Name	Value
Class 1 purity	0.91 ± 0.02
Class 2 purity	0.84 ± 0.03
Class 3 purity	0.81 ± 0.02
Class 1 and 2 overlap	0.23 ± 0.06
Class 2 and 3 overlap	0.56 ± 0.07
Class 1 and 3 overlap	0.14 ± 0.08
Training accuracy	0.96 ± 0.03
Test accuracy	0.94 ± 0.04

Table 7.1: Average training results from 8 hardware experiments with different random seeds. The purity and overlap are the average values after 100 SPSA steps. The accuracies are reported after 32 rotosolve iterations.

7.2 Qudit Hydrogen VQE

The Variational Quantum Eigensolver (VQE) is a promising application for near-term quantum devices. For more background on VQE, see Chapter 2. In this section, we implement the Variational Quantum Eigensolver to solve Hydrogen molecules with a single 4-level transmon. The original algorithm requires a minimum of two qubits and has been demonstrated on different superconducting circuit systems [198, 133]. Here, we use a 4-level system to emulate the gate set of the two-qubit system and implement the same algorithm to demonstrate the performance of our 4-level qudit.

Variational Quantum Eigensolver on a 4-Level Quantum System

VQE solves electronic structure problems by optimizing a guessed wave function, or "ansatz" $|\Psi\rangle$. The variational principle gives

$$\langle\Psi(\vec{\theta})|\, H\, |\Psi(\vec{\theta})\rangle \geq E_0 \tag{7.15}$$

where $|\Psi(\vec{\theta})\rangle$ is an ansatz parametrized by $\vec{\theta}$, and H is the Hamiltonian of the electronic system, given by

$$H = g_0 II + g_1 IZ + g_2 ZI + g_3 ZZ + g_4 YY + g_5 XX \tag{7.16}$$

where g_i is the coefficient that can be efficiently computed from the Hydrogen bond distance R.

By optimizing the parameter $\vec{\theta}$, E_0 can be approximated if the ansatz covers the lowest energy eigenstate of the electronic system. The wave function is usually prepared by a parametrized quantum circuit. The unitary coupled cluster theory shows the parameter ansatz for Hydrogen molecules can be chosen as

$$|\Psi(\vec{\theta})\rangle = e^{-i\theta X_0 Y_1}\, |01\rangle \tag{7.17}$$

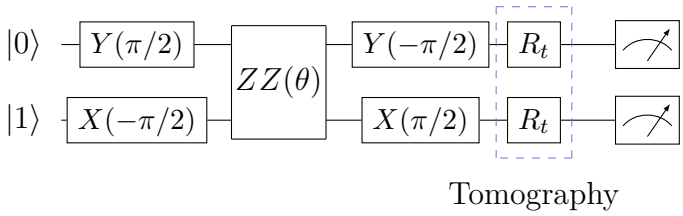


Figure 7.6: Circuit for implementing Hydrogen VQE on the two-qubit emulator.

We implement this ansatz on the two-qubit emulator with the following gate sequence shown in Fig. 7.6. The initial state $|01\rangle$ is the state of molecular hydrogen with the Hamiltonian as given in Equation 7.16. The energy of the Hydrogen molecule can be evaluated as

$$E(\vec{\theta}) = \sum_i g_i \langle \Psi(\vec{\theta}) | P_i | \Psi(\vec{\theta}) \rangle \tag{7.18}$$

Since the expectation value $\langle \Psi(\vec{\theta}) | P_i | \Psi(\vec{\theta}) \rangle$ can be directly measured on a quantum computer, the energy can be constructed after obtaining these expectation values.

Experimental Result

The experiment is implemented on the two-qubit emulator running on a transmon with the lowest 4 levels as the computation space. The physical single qudit gate is implemented with 30 ns microwave pulses with DRAG correction and a Blackman window. Apart from the physical gate, we also implemented the virtual Z gates (See Section 3.2). The ansatz is implemented with an optimized gate sequence, given as

$$|\psi(\theta)\rangle = U_2 U_{ZZ}(\theta) U_1 |0\rangle \tag{7.19}$$

where

$$\begin{aligned} U_1 &= X_{12}(\pi) X_{23}(\frac{\pi}{2}) X_{01}(\frac{\pi}{2}) X_{12}(\pi) Y_{23}(\frac{\pi}{2}) Y_{01}(\frac{\pi}{2}) \\ U_2 &= X_{12}(\pi) X_{23}(\frac{\pi}{2}) X_{01}(\frac{\pi}{2}) X_{12}(\pi) Y_{23}(-\frac{\pi}{2}) Y_{01}(-\frac{\pi}{2}) \\ U_{ZZ}(\theta) &= Z_1(\theta) Z_2(\theta) \end{aligned} \tag{7.20}$$

This ansatz starts with the ground state $|0\rangle$, which corresponds to the $|00\rangle$ for the emulated two-qubit system. An additional $X(\pi)$ gate for the emulated qubit is included

in U_1 to prepare the $|01\rangle$ state.

The measurement of the quantum state gives the probability distribution of the state in $|0\rangle$, $|1\rangle$, $|2\rangle$, and $|3\rangle$, which we denote as a vector $R = \{p_{|0\rangle}, p_{|1\rangle}, p_{|2\rangle}, p_{|3\rangle}\}$. From Equation 2.12, the expectation value for operator $M \in \{IZ, ZI, ZZ\}$ is evaluated as

$$\langle M \rangle = \text{diag}(M)^T R \tag{7.21}$$

For other Pauli operators, one of the operators in $\{IZ, ZI, ZZ\}$ is chosen after applying a rotation before measuring, which effectively measures the expectation value of the chosen operator.

The system uses a 10 μs readout pulse for single-shot readout, which has low readout fidelity (reported in Chapter 5). To compensate for this issue, active reset (described in Chapter 6) is applied, and an assignment matrix A is characterized before each experiment (described in Section 6.2).

The assignment matrix A describes the probability as follows. For an experiment that prepares the state to $|i\rangle$, A_{ij} gives the probability that the experiment measures the state in $|j\rangle$. By inverting the assignment matrix, a mitigated probability distribution $\tilde{R} = A^{-1}R$ can be obtained, and the mitigated expectation value can be evaluated as

$$\langle \tilde{M} \rangle = \text{diag}(M)^T A^{-1} R \tag{7.22}$$

There is only one parameter in the ansatz; in the experiment, the parameter is swept across 100 points from $-\pi$ to π. Each point is sampled 100 times to generate the heatmap plot of the expectation value shown in Fig. 7.7. The mitigated expectation value is also evaluated. The solved energy versus the Hydrogen bond distance is given in Fig. 7.8.

The ground truth for the expectation value is plotted as a dashed red line in the figure. We find the expectation value is slightly shifted for IZ and ZI terms. The

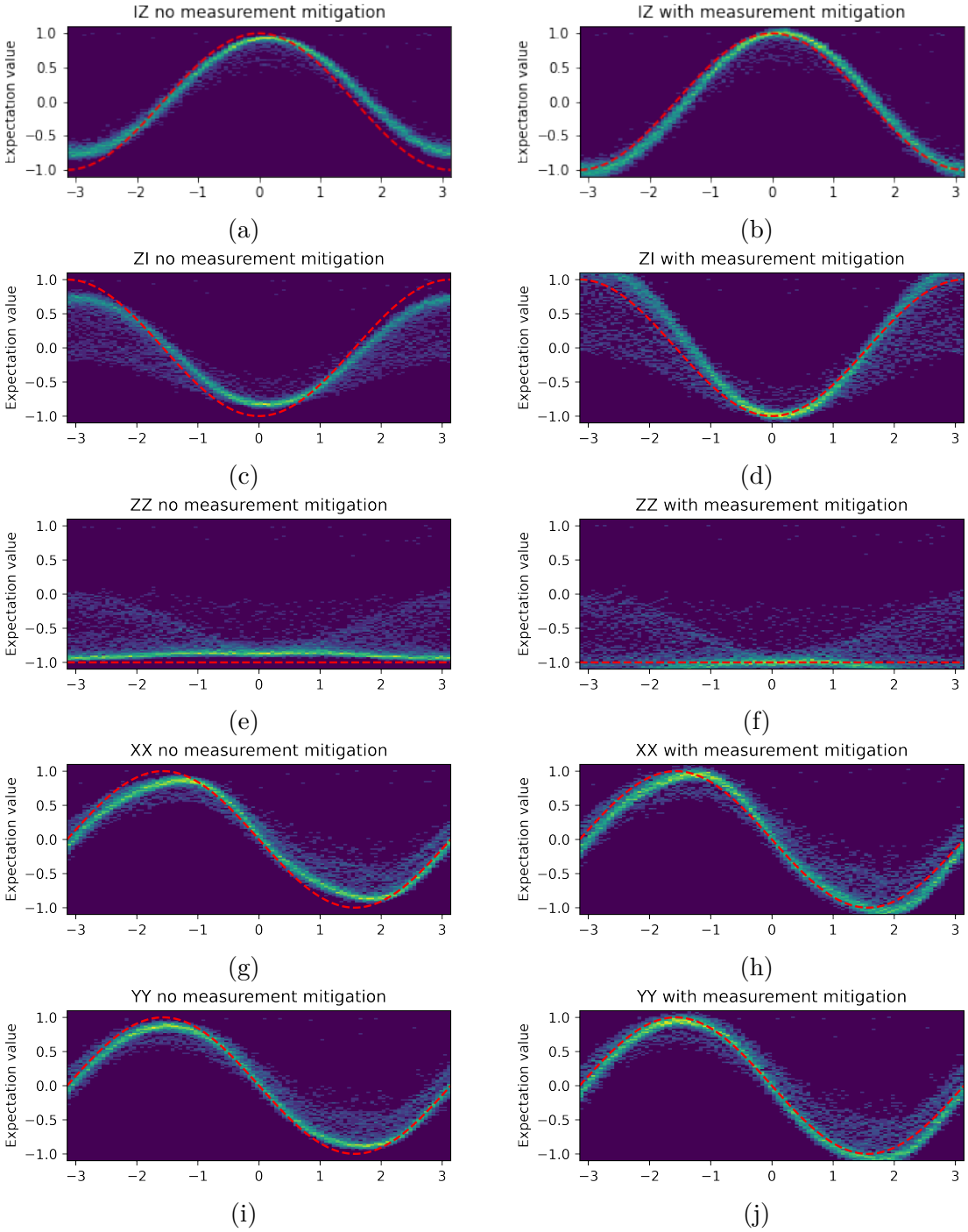
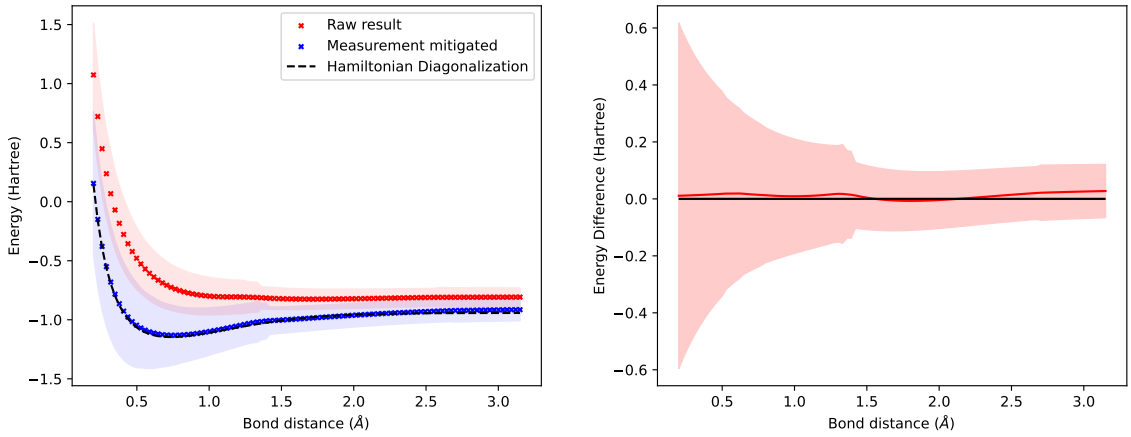


Figure 7.7: The measured expectation value versus the ansatz parameter. The experiment is repeated 100 times, and the frequency of the measured expectation value is shown as the heat plot. The left column shows the expectation value evaluated, and the right column shows the expectation value with measurement error mitigation. The red dashed line is the simulated ground truth for the ansatz.



(a) (b)

Figure 7.8: (a) The solved energy versus bond distance. The crosses denote the averaged energy value from multiple experiments, and the coloured shadow denotes the error bar equal to one standard deviation. The black line denotes the Hydrogen energy solved by the Hamiltonian diagonalisation. (b) The difference between the solved energy value with measurement error mitigation and the energy value solved by Hamiltonian diagonalisation. The solid red line denotes the averaged value between multiple experiments, and the coloured shadow denotes the error bar equal to one standard deviation.

ZZ term should not change with the parameter; however, it moved slightly. The XX and YY terms are distorted sinusoidal functions. These results show the system is exhibiting a complex noise model. Also, the expectation value has a "shadow," which could be caused by the decay of the high-level system.

From the solved energy results, we found that it has a rather large error bar due to the shadow; however, the averaged value predicts the result within 1.5×10^{-2} Hartree, reaching the chemical accuracy threshold [198].

In this experiment, only the measurement error was mitigated by linear inversion techniques, and the results are already close to the ground truth. This is because the total time to implement the gate sequence is short. It demonstrates an advantage in implementing a quantum algorithm on a multilevel transmon.

7.3 Conclusion and Outlook

This chapter demonstrates two variational quantum applications implemented using more than two energy levels of a transmon. The 1QD device is configured to behave as a qutrit for the Iris classifier or behave as a qudit four-level system to emulate a two-qubit device. Both experiments successfully demonstrated the algorithm, leading to the conclusion that utilising higher levels as computational space is both theoretically and practically possible.

Some benefits of using a higher-level system are clear. For the variational quantum eigensolver, the emulated two-qubit gate is implemented on the hardware with pure virtual gates. Therefore, the total execution time is much faster. The noise is relatively low, so only measurement assignment error mitigation is applied to the result, and a good result can be obtained.

The disadvantage of using higher levels in a transmon system is also encountered during the experiment. First, the noise from the higher levels is greater than from the lower levels, which leads to the variational quantum eigensolver having a much larger error bar compared to the other demonstrated experiment. Additionally, the dispersive readout is more difficult to tune since we need to distinguish more than two states for the transmon.

For future work, a more detailed error model for variational quantum algorithms can be investigated to find the origin of the distorted sinusoidal function and the "shadow" that appeared in the hydrogen VQE experiment. We can also explore utilising higher excited states of multiple coupled transmons. Such a configuration can significantly reduce the connectivity requirements of the transmons [264].

Chapter 8

Conclusions and Outlook

This work aims to explore the ability of higher levels in a transmon as computational space. Introducing the higher levels increases the Hilbert space and can run quantum algorithms. Utilising higher levels of a transmon has been demonstrated for analogue quantum simulation [252, 27, 253], and implementing qudit quantum logical gates [74, 288, 292]. This work extends further and demonstrates running quantum algorithms on transmon qudits. Two coaxmon prototype quantum devices were fabricated, characterised and benchmarked. The implementation of characterisation and benchmarking was explained in detail. The gate-set tomography of a three-level transmon and a four-level transmon in the coaxmon architecture was first demonstrated. Then variational quantum algorithms were demonstrated on the multilevel transmon device. This work demonstrates that the multilevel transmon system in the coaxmon architecture can be utilised as a qudit for near-term quantum applications.

Firstly, the coherence characterisation results are detailed. The two-qubit device has $T_{1A} = 127 \pm 19$ μs, $T_{2A}^{(e)} = 103 \pm 13$ μs and $T_{1B} = 26 \pm 14$ μs, $T_{2B}^{(e)} = 38 \pm 17$ μs, respectively. The characterised parameters for the multilevel transmon are $T_1^{(01)} = 189 \pm 41$ μs, $T_1^{(12)} = 119 \pm 21$ μs, $T_1^{(23)} = 87 \pm 23$ μs and $T_2^{(01)} = 118 \pm 21$ μs, $T_2^{(12)} = 76 \pm 27$ μs, $T_2^{(23)} = 35 \pm 14$ μs respectively. We found the charge-induced frequency change of the transition frequency between $|2\rangle$ and $|3\rangle$ is about 20 kHz. The coherence times of Q1 of the 2QB device and the 1QD device compare very well

to state-of-the-art, while the Q2 of the 2QB device also has good coherence for near-term algorithms. The lower coherence times in subspaces of higher excited states are expected from the theory of charge dispersion in the transmon, highlighting a limitation of transmon qudits.

Secondly, gate set tomography and randomised benchmarking are implemented to benchmark the performance of both devices. The simultaneous single qubit randomised benchmarking gives single qubit gate averaged infidelity $0.37 \pm 0.05 \times 10^{-3}$, and two-qubit gate $ZX(-\pi/2)$ averaged infidelity $2.6 \pm 0.4 \times 10^{-2}$. This result agrees with the GST result, which reports the two-qubit gate $CNOT$ averaged infidelity is $2.0 \pm 0.3 \times 10^{-2}$. For the multilevel transmon device, we successfully implemented qutrit GST with an efficient method that only requires 4 measurements to reconstruct the qutrit state. GST reported the qutrit Hadamard gate infidelity $3.22 \pm 0.11 \times 10^{-3}$ and the result of qutrit GST agrees with the qutrit randomised benchmarking. Comparing the benchmarking result between the two-qubit device and the 4-level emulated two-qubit device, the gate fidelities are close; however, the emulator uses multiple hardware pulses to emulate the two-qubit device, the two-qubit Clifford gate randomised benchmarking result (infidelity $9.5 \pm 0.7 \times 10^{-2}$) is worse than the two-qubit (infidelity $4.0 \pm 0.3 \times 10^{-2}$). The two-qubit Clifford group randomised benchmarking doesn't affect implementing near-term algorithms on the emulator, as the ansatz can be optimised for the 4-level system.

Thirdly, active reset is implemented for the 1QD device as a qudit to prepare a high-fidelity initial state. An FPGA firmware is implemented for fast state discrimination, requiring 40ns to distinguish 4 states from the board. A waveform includes the measurement pulse, and three π pulses for different transition levels are preloaded to the FPGA. These three pulses are played conditionally on the discriminated state of the transmon. We read and reset the multilevel transmon 3 times to get the best initial state preparation. Then the initial state fidelity is characterised by GST. We

found the initial state fidelity is improved from 0.900 ± 0.011 to 0.9932 ± 0.0013, which indicates a successful implementation of active reset on transmon qudit.

In the last chapter, we demonstrated two quantum variational algorithms on the qudit device. We first utilised the 1QD device as a qutrit and implemented a single qutrit classifier for the iris dataset. The single qutrit classifier consists of a hybrid encoding component and a classification circuit; two components are trained separately with SPSA and rotosolve optimiser. We report the hardware implementation of the qutrit quantum classifier is a success, with a training accuracy of 0.96 ± 0.03 and a testing accuracy of 0.94 ± 0.04. We also successfully demonstrated a variational quantum eigensolver for Hydrogen molecules on the two-qubit emulator built on the 1QD device. The solved energy versus the Hydrogen bond distance is within 1.5×10^{-2} Hartree, below the chemical accuracy threshold.

The characterisation, benchmarking results, and successful demonstration of two quantum algorithms show that higher levels of a transmon can be used as computational space and increase the size of Hilbert space with minimum extra costs.

Outlook

The initial demonstration of utilising the lowest 4 levels for computation is promising and encourages further studies aimed at scaling up quantum computers with transmons at higher levels. There are a few research directions that can be followed immediately after the work presented in the thesis: First, as the analysis from the qudit Hydrogen VQE experiment revealed that the readout error still dominates, proper quantum amplifiers can be utilised to improve the readout fidelity. Second, entanglement generation between multiplex multilevel transmons is necessary for scaling the system, and the physics behind generating entanglement with a multilevel system needs to be studied. Third, an optimised implementation for an arbitrary two-qubit Clifford group can also be explored to build a more efficient emulator.

For near-term applications, using qudit-based quantum processors might be a more practical choice over qubit-based quantum processors. By using higher levels of the transmon, the number of electronics required for each transmon can be reduced, making the qudit-based quantum processor more cost-effective and easier to build physically. Building intermediate scale qudit-based processors is a promising research direction to explore in the near term.

In the long term, fault-tolerant quantum computing needs to be achieved for more advanced quantum algorithms. Qudits can be used as the unit in error correction codes. Qudit stabilisers and error correction codes have already been developed, and the qudit toric code can tolerate larger physical noise than qubit codes [281], making it potentially easier to implement on the hardware. Using qudits to construct error correction codes provides an alternative solution for achieving fault-tolerant quantum computing.

Appendix A

Device Characterisation and Gate Tune-Up Procedures

This appendix discusses the procedures for finding the frequency of the resonators and transmon qubits, and then calibrating the parameters for implementing quantum gates. The process is shown in Figure A.1. For the qudit device, single qubit gate calibration procedures are repeated for each neighbouring transition.

A.1 Discover Qubit and Resonators

Resonator Spectroscopy We use the dispersive readout mechanism to infer the qubit state from the response of a resonator that couples to the transmon, as discussed in Chapter 3. Therefore, the first step to tuning the control system is finding the resonance frequency of the resonator. The coaxmon architecture processes the reflected measurement signal from the resonator. The resonance can be found by applying continuous wave frequency spectroscopy on the readout port. The experiment sends a continuous microwave signal to the resonator and analyses data of both amplitude and the phase of the reflected microwave signal. The first step in looking for the resonance is by sweeping through a large frequency range, usually indicated by the design parameter of the device. See Figure A.2.

Then we zoom in on the frequency region where the resonance exists. A typical example is shown in Figure A.3.

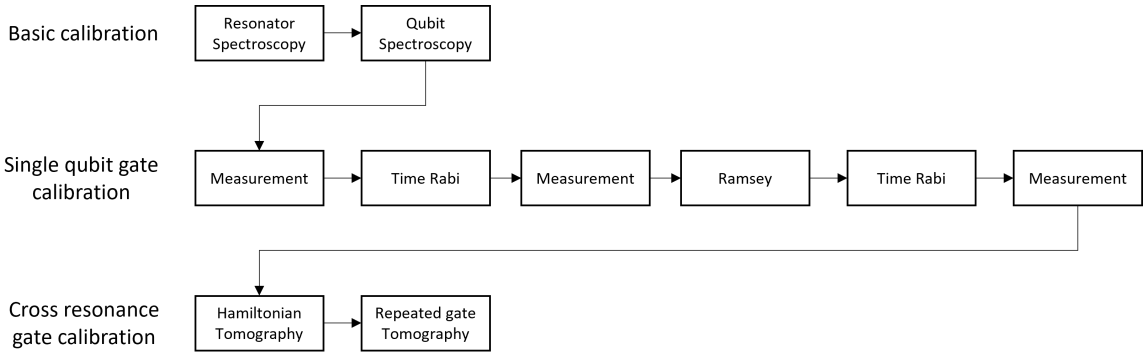
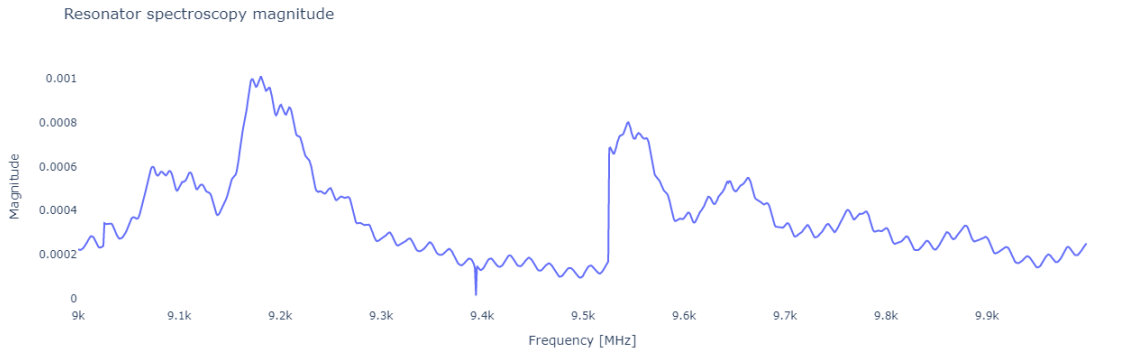
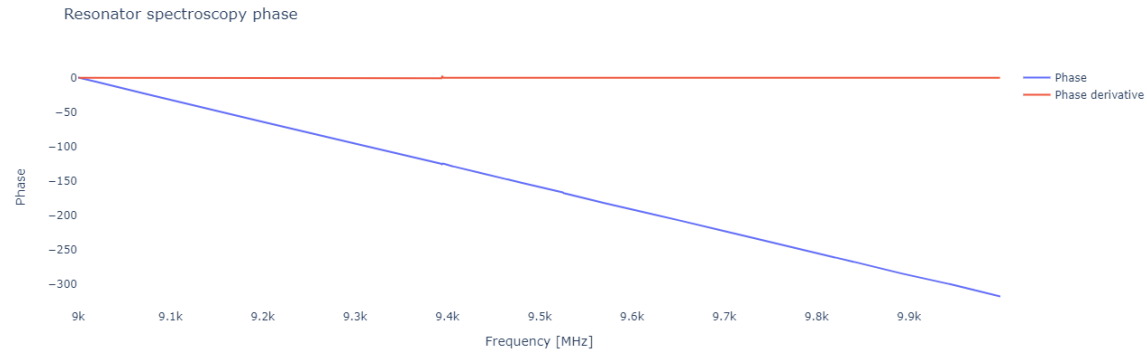


Figure A.1: Overview of tune-up procedures for calibrating parameters for quantum gates implementation. The calibration procedure starts with basic calibrations that look for resonator frequency and qubit frequency. The single-qubit gates are first calibrated with a sequence of experiments, and then two-qubit gates are also calibrated with multiple experiments.

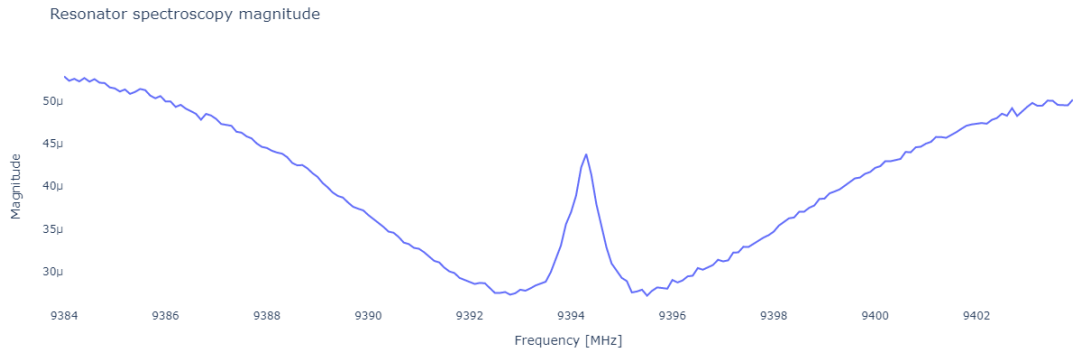


(a) Magnitude

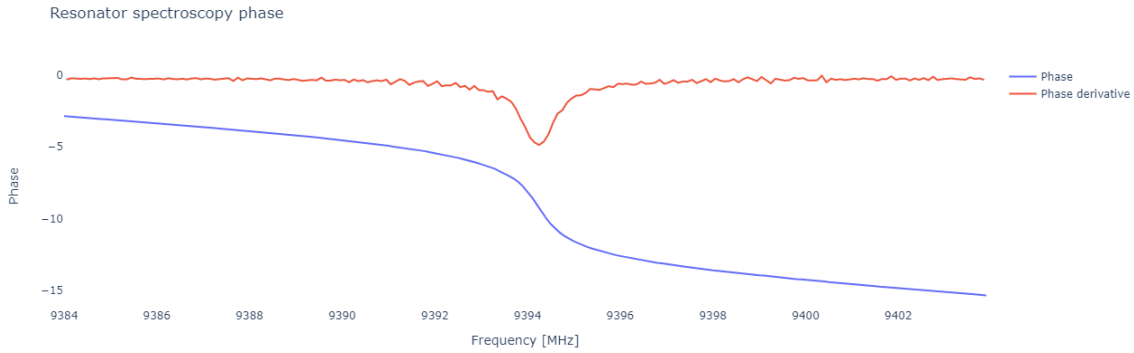


(b) Phase

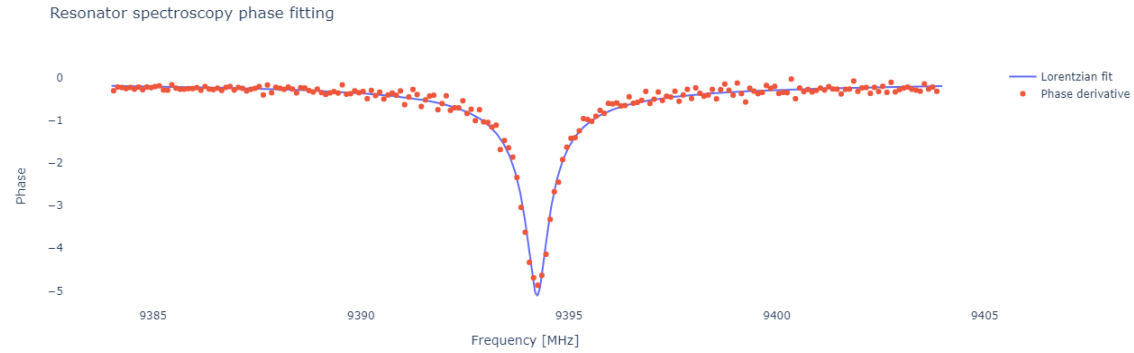
Figure A.2: A typical resonator spectroscopy experiment that sweeps through a large range of frequencies. (a) is the magnitude response, and (b) is the phase response.



(a) Magnitude



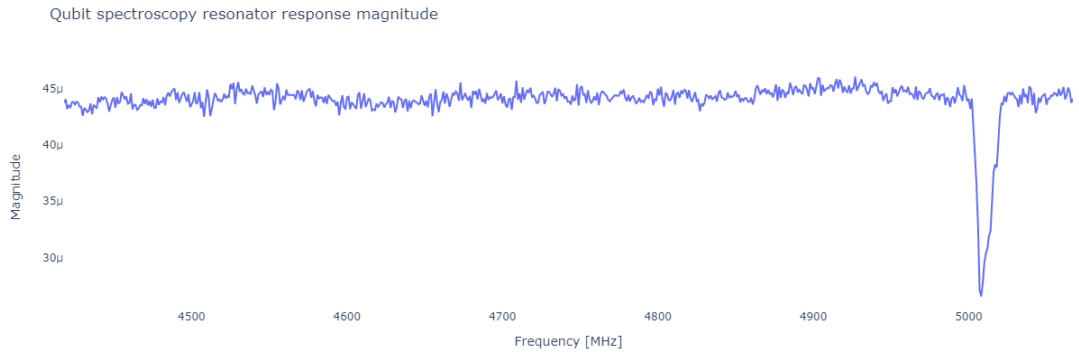
(b) Phase



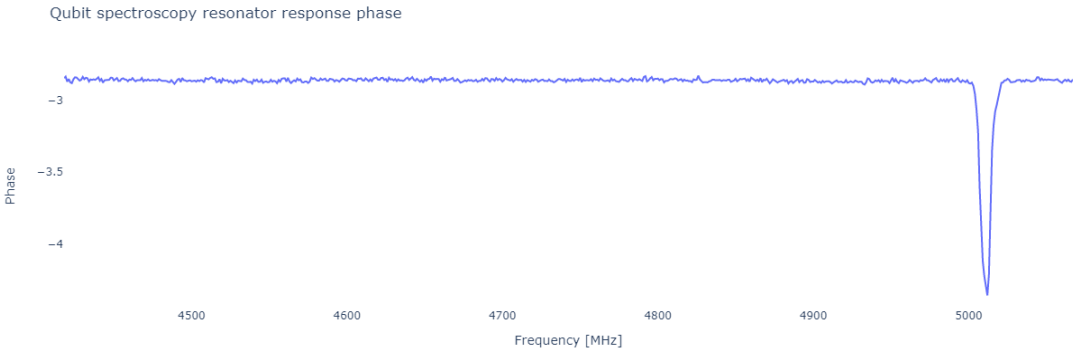
(c) Fitting

Figure A.3: A typical resonator spectroscopy experiment. (a) is the magnitude response, (b) the phase response, and (c) the fitting of the derivative of the phase, or the group delay of the qubit. From the Lorentzian fitting of the group delay, we can extract essential information about the resonator such as its resonance frequency and linewidth.

Qubit Spectroscopy The resonator exhibits a frequency shift that depends on its coupled qubit, and therefore when the qubit is perturbed by an on-resonance signal, the resonator shows a different response. Therefore, the qubit frequency can be discovered by applying a microwave signal perturbation to the qubit and sweeping its frequency while monitoring the response of the resonator. A typical experiment is shown in Figure A.4. The point with a significant change in the response indicates the frequency of the qubit.



(a) Magnitude

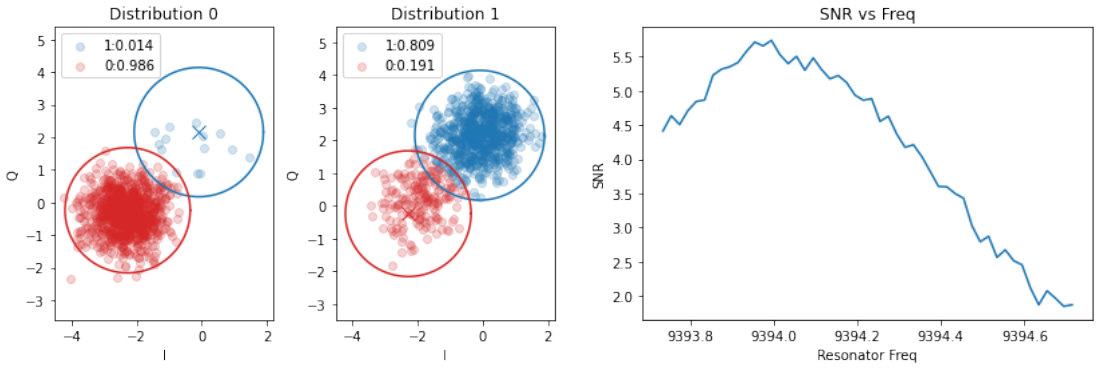


(b) Phase

Figure A.4: A typical qubit spectroscopy experiment that sweeps through a large range of frequencies. (a) is the magnitude response, and (b) the phase response of the resonator versus the frequency of the perturbation pulse sent to the qubit.

A.2 Calibrate Single Qubit Gate Parameters

Measurement Calibration While driving the qubit at its 01 transition frequency, we demodulate the response signal from the resonator and plot the IQ channel on a plane and generate a scattering plot. A Gaussian Mixture Model is trained on the dataset to discriminate different qubit states based on resonator response, as described in Chapter 4. To find the optimal readout frequency, we sweep the readout frequency around the resonator resonance frequency and measure the SNR (discussed in Chapter 4) predicted by the GMM model. Then the optimal readout frequency is chosen by selecting the frequency that maximises the SNR. See Figure A.5.



(a) (b)

Figure A.5: (a) Scatter plot of the demodulated signal and the GMM discriminator trained on this dataset. (b) The SNR versus readout frequency. The optimal readout frequency is chosen at the peak of the pattern.

Time Rabi Experiment From the theory discussed in Chapter 3, the Rabi rate of a qubit is strictly linear and proportional to the external field amplitude. The Time Rabi experiment aims to find the exact constant between the qubit pulse amplitude and the Rabi rate. A qubit control pulse at the qubit transition frequency is sent to the qubit, and the expectation value of the qubit is monitored while varying the lengths of the qubit pulse. See Figure A.6.

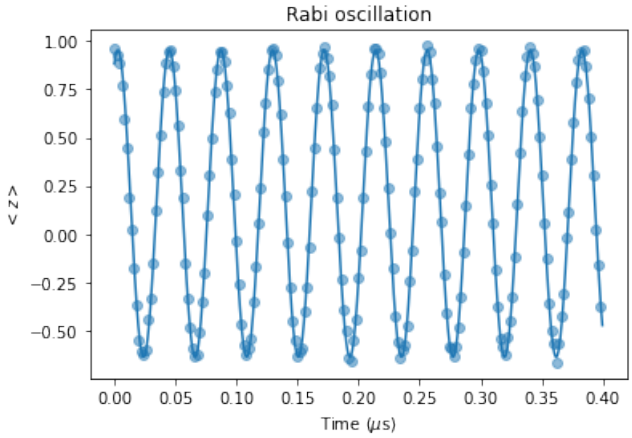
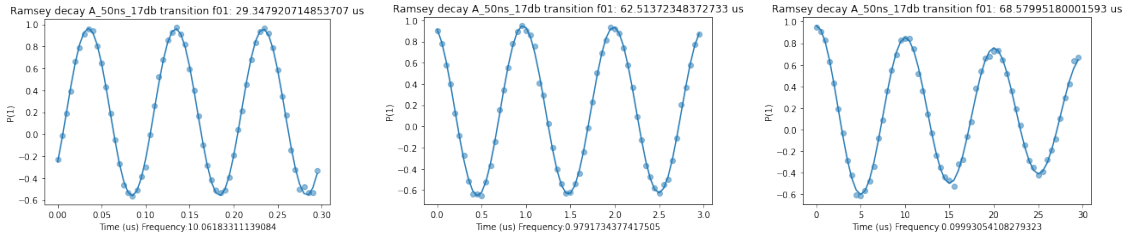


Figure A.6: Experimental data of a Rabi oscillation experiment. The Y axis is the $\langle Z \rangle$ expectation value of the qubit, evaluated by the probability distribution of the measured state. Due to the readout error, the expectation value usually cannot cover the full range of $[-1, 1]$.

Ramsey Experiment

The Ramsey interferometry experiment explained in Chapter 4 is also used to calibrate the qubit transition frequency. In our practice, three Ramsey experiments are performed three times with different offset frequencies and maximum duration. Each experiment gives a more accurate estimation of the qubit transition frequency.



(a) 10 MHz detuned (b) 1 MHz detuned (c) 0.1 MHz detuned

Figure A.7: Ramsey experiment to calibrate qubit driving frequency. The qubit drive frequency is offset by (a) 10 MHz, (b) 1 MHz, and (c) 0.1 MHz.

A.3 Two-Qubit Gate Calibration

After the single qubit gates are calibrated, we can move on and calibrate parameters for two-qubit interaction and generate entanglement. In this thesis, we use cross-

resonance interaction described in Chapter 3. The tune-up of the cross-resonance gate comprises two steps: Hamiltonian tomography to roughly calibrate the gate parameters and repeated gate tomography to fine

-tune the parameters and remove transient errors. The procedures are from [238, 208].

Hamiltonian Tomography Hamiltonian tomography measures $\langle X \rangle$, $\langle Y \rangle$, and $\langle Z \rangle$ axes of the target qubit under the cross-resonance drive. A cancellation pulse is applied to the target qubit to correct the unwanted IX and IY interactions. The pulse width t is swept to find the rate of the ZX interaction. A typical experiment is shown in Figure A.8.

Repeated Gate Tomography Similar to the Hamiltonian tomography, repeated gate tomography sweeps the number of repetitions of the gate sequence. A typical example is shown in Figure A.9.

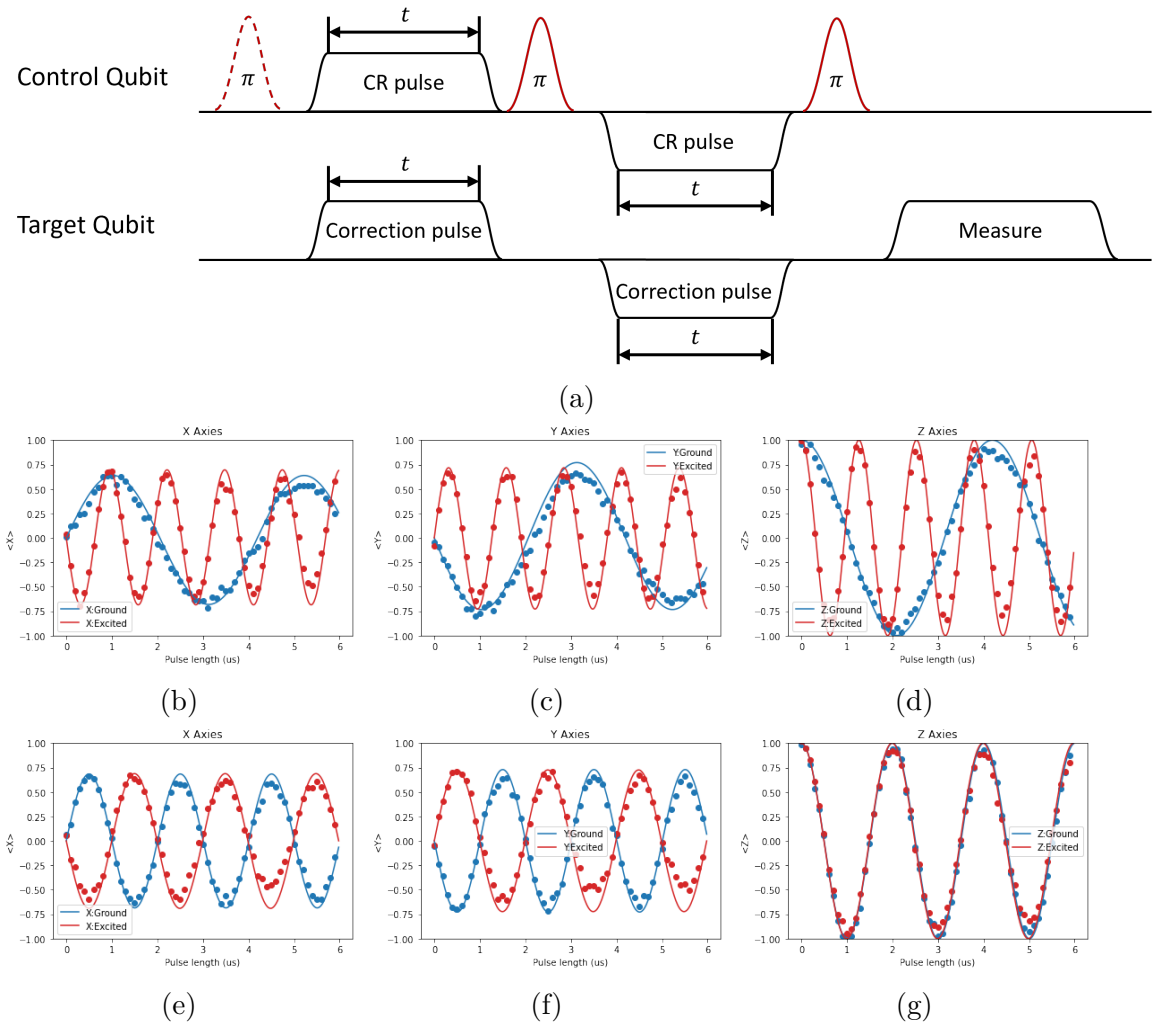


Figure A.8: Hamiltonian tomography. (a) The gate sequence of the Hamiltonian tomography experiment. t denotes the width of the cross-resonance pulse and is swept in this experiment. The experiment is repeatedly executed and updates the parameter until the fidelity of the gate cannot be further improved. (b) to (d) A typical experiment result of the $\langle X \rangle$, $\langle Y \rangle$ and $\langle Z \rangle$ expectation values from the initial experiment, respectively. (e) to (g) A typical experiment result of the $\langle X \rangle$, $\langle Y \rangle$ and $\langle Z \rangle$ expectation values from the last experiment, respectively.

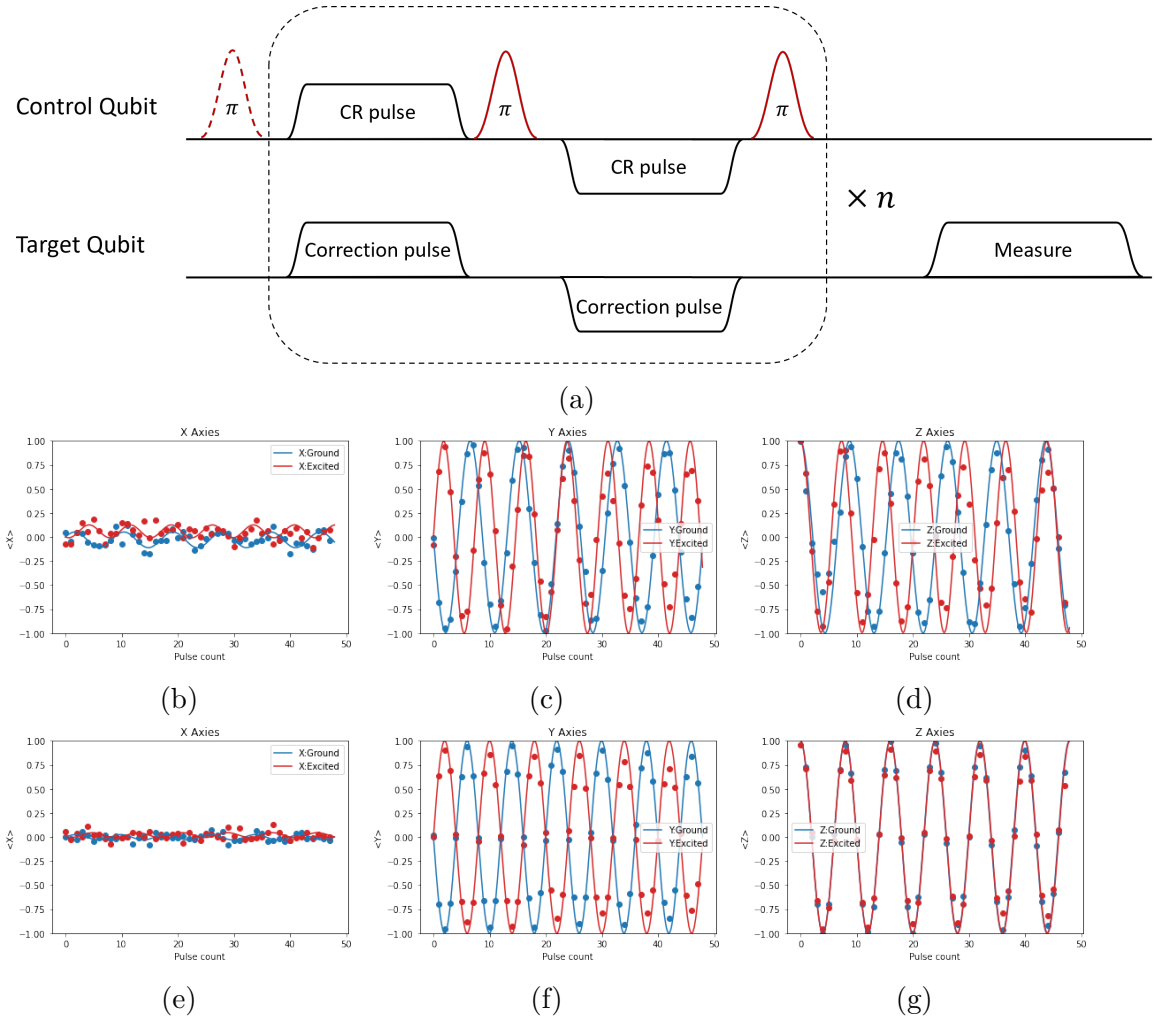


Figure A.9: Repeated gate tomography. (a) The gate sequence of the repeated gate tomography experiment. A full $ZX(-\pi/8)$ gate enclosed in the dashed lines is executed repeatedly n times. The experiment is repeatedly executed and updates the parameter until the fidelity of the gate cannot be further improved. (b) to (d) A typical experiment result of the $\langle X \rangle$, $\langle Y \rangle$ and $\langle Z \rangle$ expectation values from the initial experiment, respectively. (e) to (g) A typical experiment result of the $\langle X \rangle$, $\langle Y \rangle$ and $\langle Z \rangle$ expectation values from the last experiment, respectively.

Appendix B

Fabrication Process of Coaxmon Qubits

In this chapter, we summarise the method used to fabricate the sample devices with given parameters.

B.1 Pattern Generation

To fabricate our device, the first step is to generate the device pattern that we would like to print onto the substrate using the lithography method.

Automatic Pattern Generation Framework

The traditional method to generate the fabrication pattern is to edit manually with interactive CAD software (for example, AutoCAD). However, completing a pattern for a whole wafer by hand is extremely inefficient. It is very easy to make mistakes with more than a hundred chips of different parameters on a wafer. To solve this problem, a software package *qpat* was developed to generate the pattern with a hierarchical code structure. The structure of the *qpat* software is shown in Fig. B.1.

The core advantage of *qpat* is the flexibility of the rendering back-end model. The back-end model is fully decoupled from the pattern description model. Here we developed a GDS backend which renders the pattern into a GDS file, and a Raith backend, which generates a full Raith position list used in the Raith e-beam lithography tool.

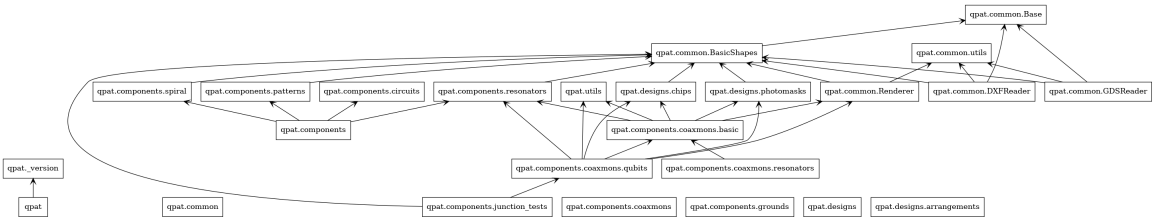


Figure B.1: The qpat package structure

Josephson Junction Pattern

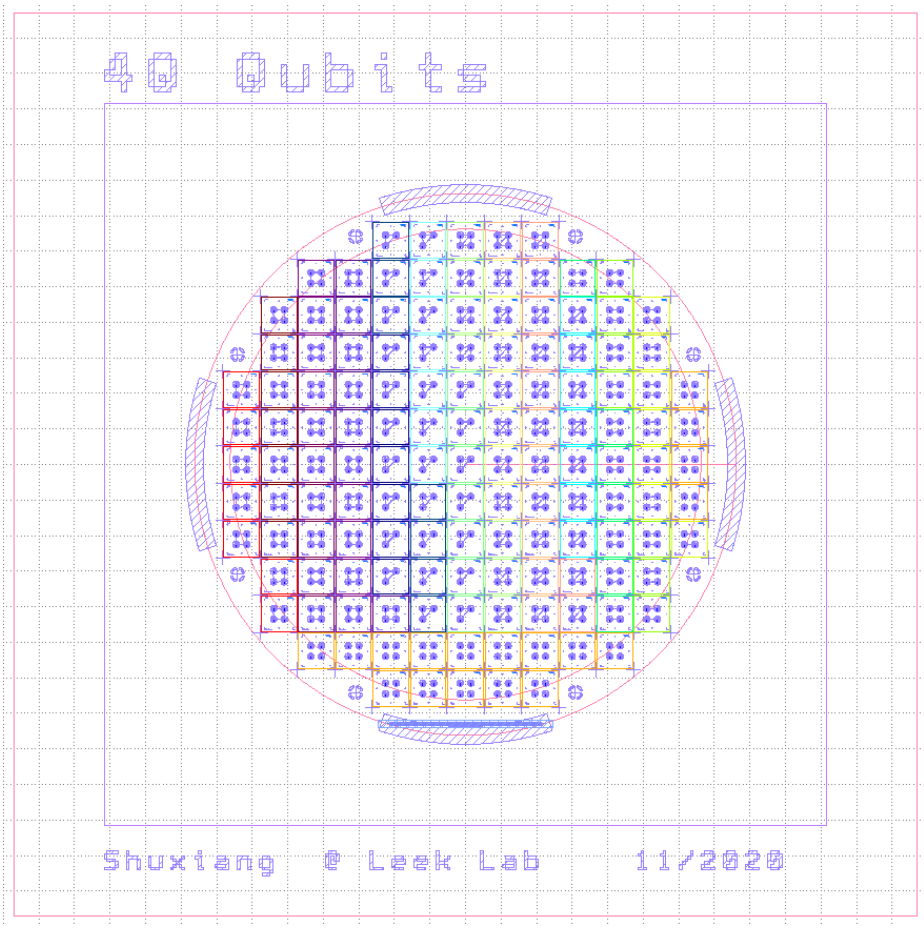
To build the Josephson junction, we used the shadowed evaporation method or the Dolan bridge method [73] to build the Josephson junction. The pattern is shown below.

A T-shaped arm and an L-shaped arm are designed into the pattern and the bridge is formed in between. The width of the T-arm and the L-arm defines the width and height of the junction. In our pattern, we added an extra undercut layer with exposure using a lower dose to create space to let the metal from shadowed evaporation stick onto the substrate. Also, we added extra exposure on the T-arm for the same reason. We found this helpful to get a cleaner liftoff process which reduces the chance of floating metal shorting the junction.

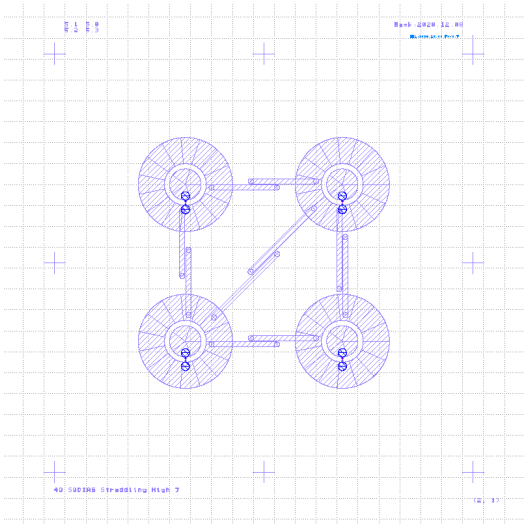
B.2 Fabrication Process

Wafer Cleaning

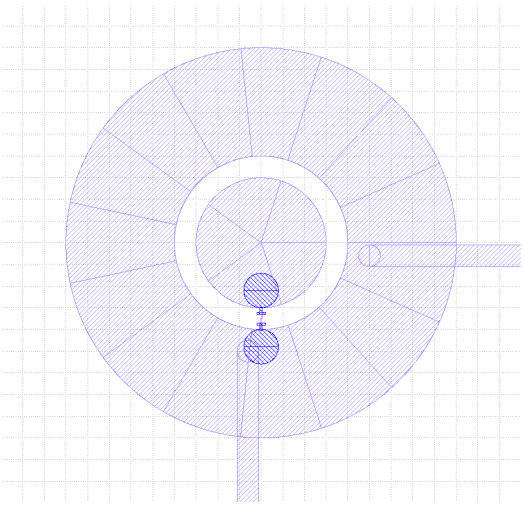
While the silicon wafer stays in the air, the surface of the silicon wafer will form a thin layer of silicon dioxide. This layer can increase the dielectric loss of our qubits, which is one of the most important contributions to the energy relaxation error [276]. The first step of our fabrication is to remove this oxidation layer. This is done with the standard RCA clean procedure which was developed in the integrated circuits industry [136]. The process is rather simple: we dip the wafer into a 1:10 BOE hydrofluoric acid for 150 seconds. The hydrofluoric acid is extremely toxic; the operator needs to



(a) Full mask pattern



(b) A chip pattern



(c) A qubit pattern

Figure B.2: Pattern generated from the qpat software.

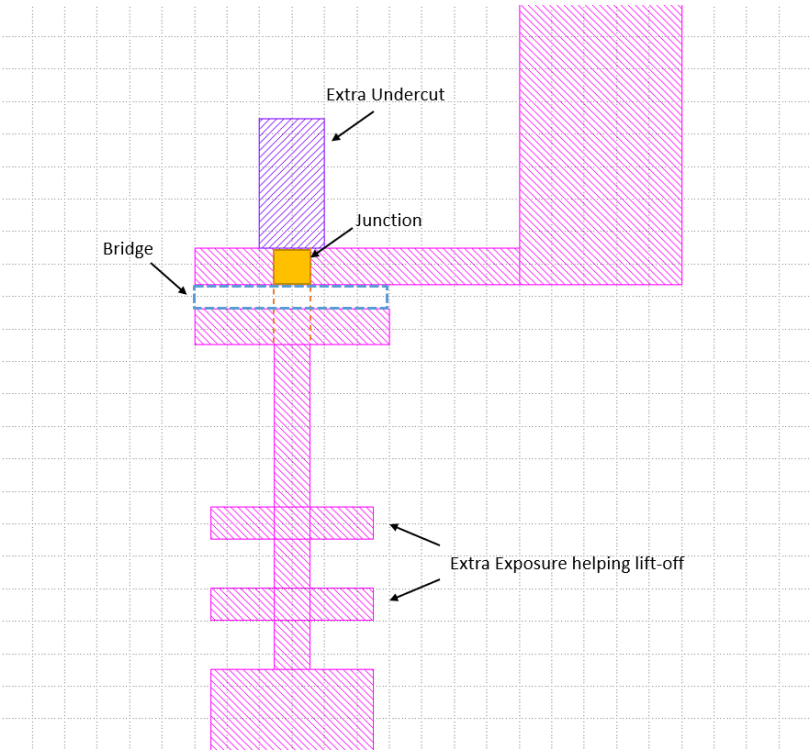


Figure B.3: Josephson junction pattern explained

be properly geared up and follow a very strict safety protocol to conduct this step.

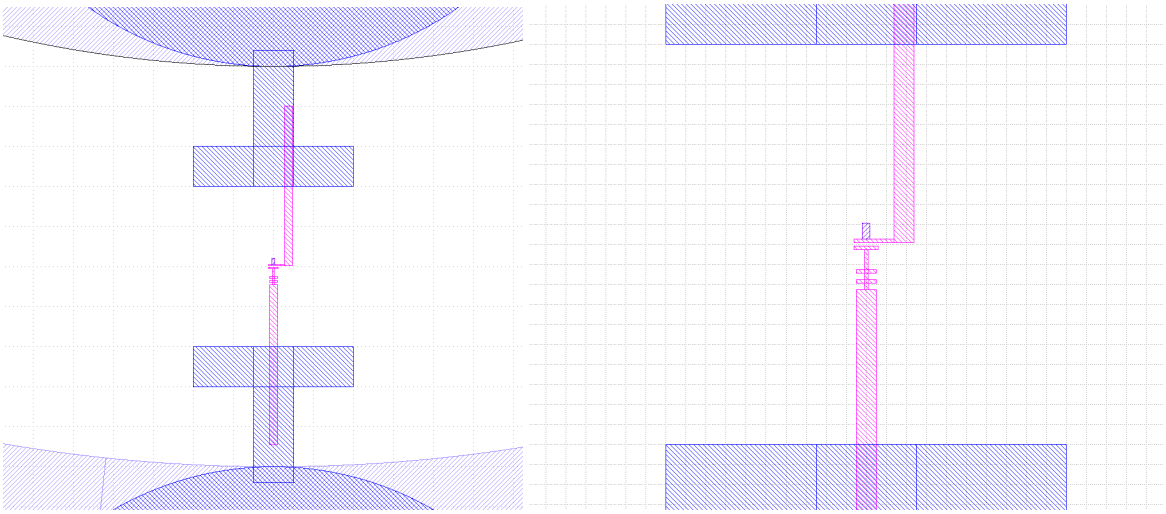
Then we clean the wafer with two DI water dips and immediately pack the wafer into the customised evaporator cassette and put it into a sealing bag full of nitrogen.

First Metal Deposition

After the cleaning process, we bake the wafer inside the evaporation chamber to remove potential water residues. Then we deposit 100 nm of aluminium onto both sides of our substrate.

Photolithography

Now we start our photolithography process. We use a positive process, which means where there is metal on the mask, there will be metal on the wafer. We choose AZ1514H as the photoresist. First, we bake the substrate above 100 Celsius (120 Celsius in practice) on a metal ring to remove potential water residues. The metal



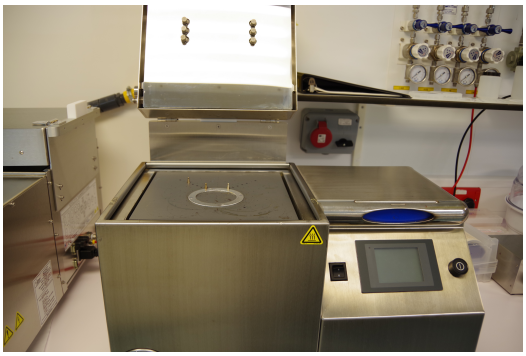
(a) Junction alignment (b) Junction pattern

Figure B.4: Junction pattern on the qubit.

Figure B.5: Doing HF cleaning.

ring can give good thermal conductance while protecting the bottom side of the wafer from touching the dirty hot plate surface. This step is often referred to as *prebake*. Then we take the wafer out and wait for 2 minutes until it cools down to room temperature.

Next, we coat the backside protection layer. We put the substrate onto the specialised vacuum-free spin-coater cassette. This is to protect the backside of the substrate, making sure it won't touch anything and damage the deposited metal layer. We take AZ1514H out of the fridge (4 Celsius) and immediately coat it onto the substrate to reduce the viscosity variation because of the temperature difference. The

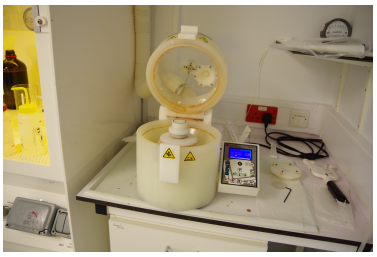


(a) Hotplate

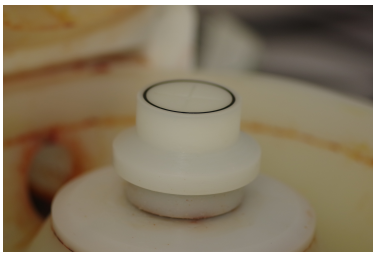
(b) Metal ring

Figure B.6: Hotplate used for fabrication. We used a metal ring to protect the centre area of the wafer from touching the plate.

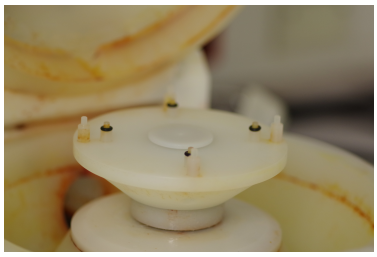
acceleration speed is 500 rpm/s, we enter a spread coating speed of 500 rpm for 10 seconds, then reach the coating speed of 4000 rpm for 50 seconds. Then we bake the substrate at 90 Celsius to remove the solvent and harden the resist.



(a) Spincoater



(b) Contact wafer chuck



(c) Contactless wafer chuck

Figure B.7: Spincoater used for fabrication.

Note that since the chuck is vacuum-free, the substrate just sits on the chuck. During the spinning, the substrate will oscillate and form a standing wave. This causes the non-uniformity of the resist. See Fig. B.8. After this step, we change the chuck into the standard vacuum contact chuck. Then we repeat the previous steps to coat the other side, and let the coated layer directly contact the vacuum chuck.

After both sides are prepared, we move to the mask aligner. We are using an EVG620 mask aligner; the exposure dose is 17-20 mJ/cm^2, and needs to be calibrated and tested before the lithography process.

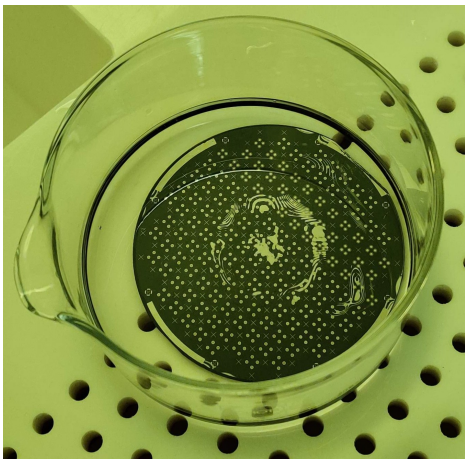
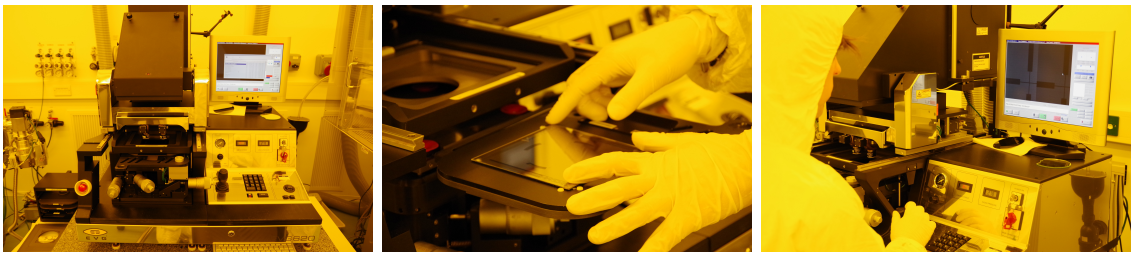


Figure B.8: Non-uniform resist when using a contactless chuck for fabrication.



(a) EVG620 (b) Load mask (c) Backside alignment

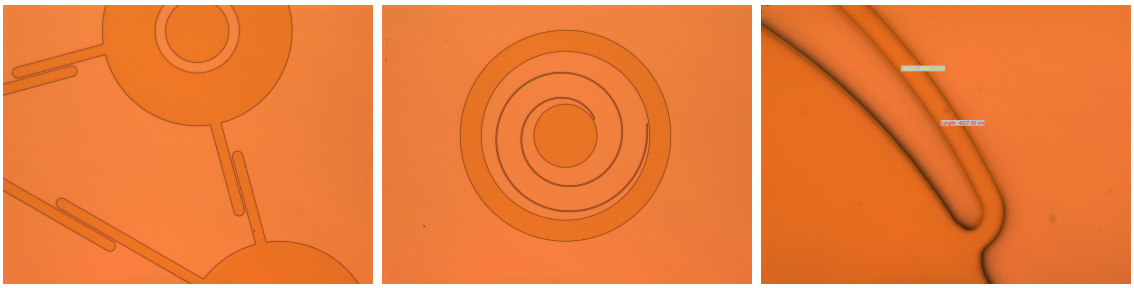
Figure B.9: EVG620 mask aligner used for fabrication.

Development

After the exposure, we dip the wafer into AZ726 MIF developer to remove the exposed resist. The developer's temperature was maintained at 25 Celsius, and the development time was 80 seconds. During the development, a proper water flow should be given. Then we dip the wafer into DI water for 30 seconds twice into different beakers to remove developer residues. Note that we don't dry the wafer at this step, because the pattern we have is rather small, and most of the area has no resist left. Blowing it with nitrogen can damage the resist profile.

Wet Etching

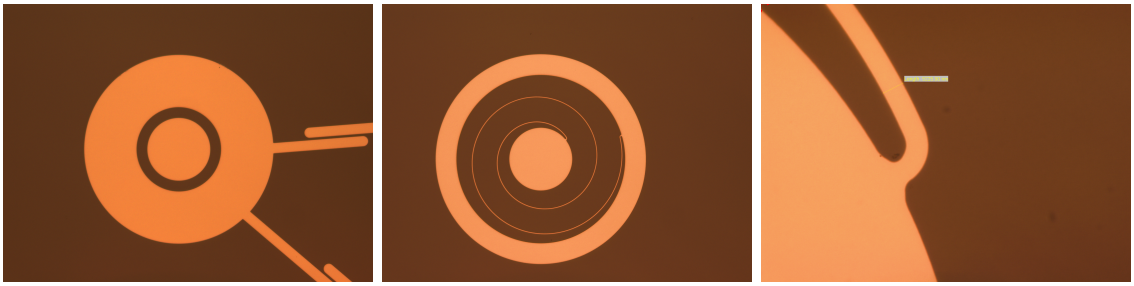
After the development, we dip our wafer into the aluminium etchant (What etchant). The temperature is maintained at 25 Celsius. Note we fabricate the pattern on both



(a) Qubit side 5x (b) Resornator side 5x (c) Resonator 100x

Figure B.10: Microscopic pictures of resist after development

sides of the substrate. For the first side, we expect to see the pattern after around 90 seconds. For the second side, we expect to see the patterns in about 150 seconds. This is because the resist residues couldn't be fully cleaned. We etch another 30 seconds after the pattern is clearly shown. Then we dip twice into different DI water beakers to stop the acid.



(a) Qubit side 5x (b) Resonator side 5x (c) Resonator 100x

Figure B.11: Microscopic pictures of resist after etching

Post-Etching Cleaning

After the etching process, we leave the wafer in the DMSO solvent at 80 Celsius for 15 minutes to remove the resist. Then we dip twice into different DI water beakers to remove DMSO.

E-Beam Lithography

Now we start the E-beam lithography process. First, we bake the wafer to 180 Celsius to remove the water, then just like photolithography, we spin coat the protection layer

with PMMA on the first side of the wafer, with the non-vacuum chuck. Then we bake at 180 Celsius to harden the resist. After it has cooled down (5 minutes), we spin the written layer.

Doses: 1500 junctions, 1300 overlap pads, 200 undercuts.

Second Metal Deposition

Ensure the wafer is oriented correctly when loading it. During the ion milling process, verify the parameters in the evaporator. For the first evaporation, set the angle to 60°, the thickness to 60 nm, and the rate to 0.5 nm/s. Proceed with oxidation for 10 minutes at 10 mbar. Finally, for the second evaporation, adjust the angle to 0°, the thickness to 70 nm, and maintain the rate at 0.5 nm/s.

Lift-Off

To perform the DMSO lift-off, use a plastic star to ensure the back side of the wafer does not touch the bottom of the beaker. First, leave the wafer in DMSO at 80°C for 30 minutes, using a pipette to remove metal from the wafer. If some metal remains, extend the duration and repeat the process until all metal is removed. Next, transfer the wafer to fresh DMSO at 30°C for 30 minutes and use the pipette again to remove any metal residues. Finally, rinse the wafer in three steps with warm spectroscopic grade DI water, allowing 30 seconds for each step.

Post-Fabrication Cleaning

After the etching process, we leave the wafer in the DMSO solvent at 80 Celsius for 15 minutes to remove the resist. Then we dip twice into different DI water beakers to remove DMSO. Then we do UV Ozone to clean the DMSO residues and resist residues.

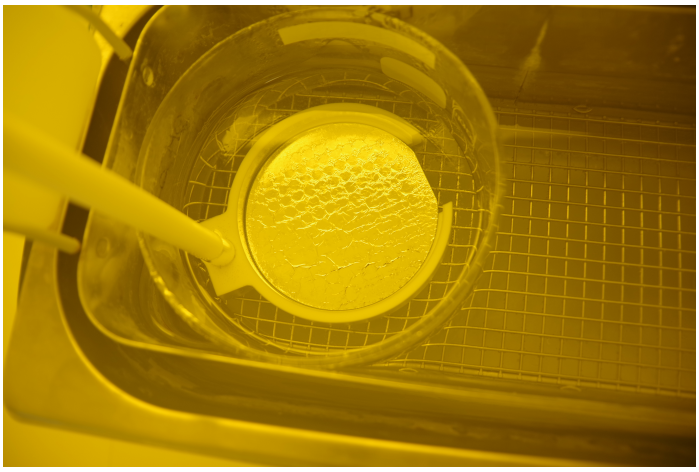
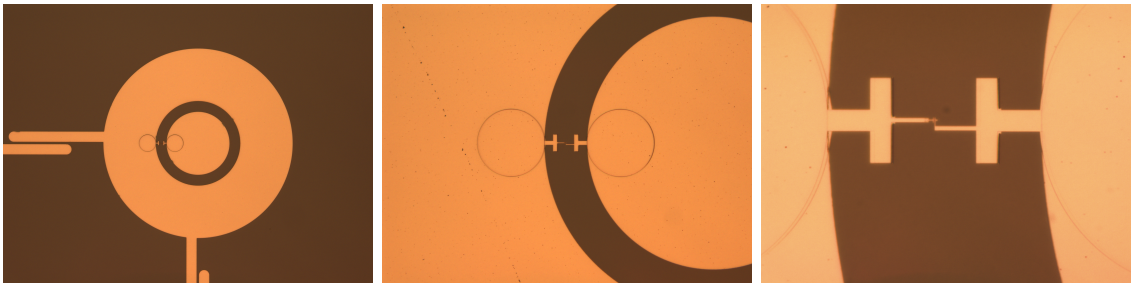


Figure B.12: Lift off in DMSO.



(a) 5x (b) 20x (c) 100x

Figure B.13: Junction microscopic picture after lift-off

Wafer Dicing

To prepare the wafer for dicing, first apply AZ1514H resist as a protection layer. Mount the non-vacuum chuck and spin the wafer with a ramp to 500 rpm for 10 seconds, followed by spinning at 4000 rpm for 50 seconds, and then a ramp down for 10 seconds. Bake the wafer at 100°C for 50 seconds on a metal ring, followed by a 5-minute cooldown. Then spin the other side of the wafer. It is then ready to be sent for dicing.

Before Mounting Cleaning

Before mounting the chips, we perform a final cleaning process. This involves three steps in acetone at 50°C for 5, 10, and 5 minutes respectively, followed by a rinse in

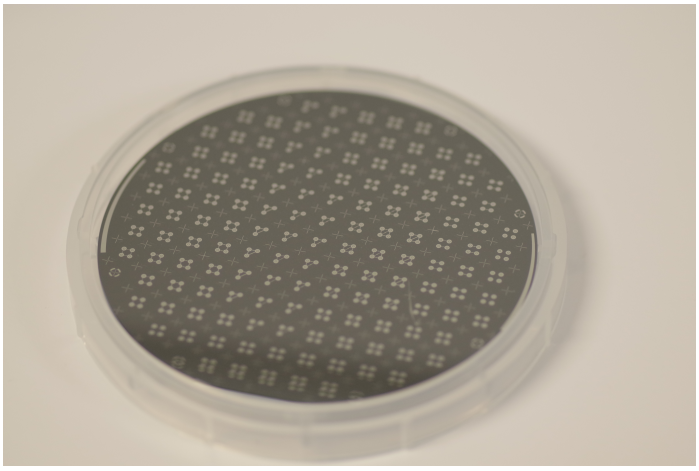


Figure B.14: Wafer finished cleaning.

IPA. After this, the chips are placed in DMSO at 80°C and subsequently rinsed three times in DI water: first at 80°C, then at 50°C, and finally at room temperature.

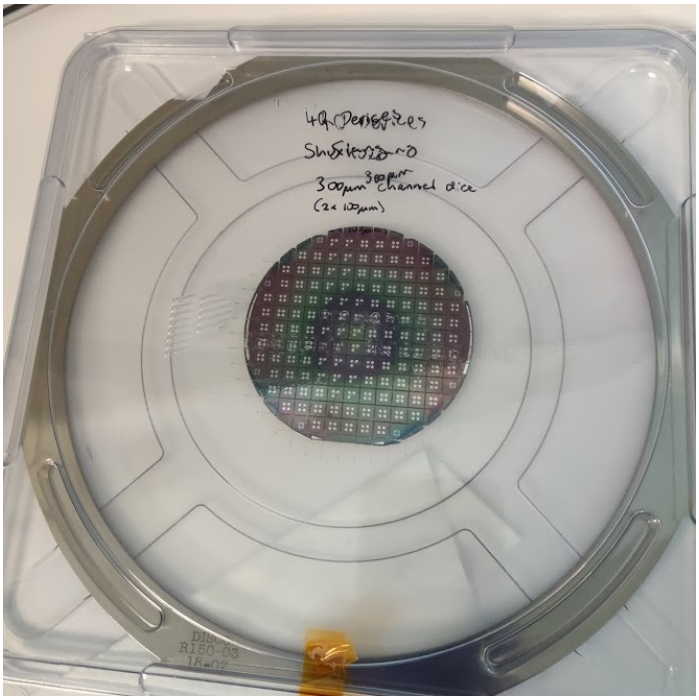


Figure B.15: Wafer after dicing.

Appendix C

Supplementary information for gate set tomography

C.1 Germs and Fiducials

Table C.1: The gate set and the Germs selected in 'std2Q_XYCNOT' model.

Gate set		
I	$X_A(\frac{\pi}{2})$	$X_B(\frac{\pi}{2})$
$Y_A(\frac{\pi}{2})$	$Y_B(\frac{\pi}{2})$	CNOT
List of germs		
I	$X_A(\frac{\pi}{2})$	$X_B(\frac{\pi}{2})$
$X_A(\frac{\pi}{2}).Y_A(\frac{\pi}{2})$	$X_B(\frac{\pi}{2}).Y_B(\frac{\pi}{2})$	$X_A(\frac{\pi}{2}).X_A(\frac{\pi}{2}).Y_A(\frac{\pi}{2})$
$X_B(\frac{\pi}{2}).X_B(\frac{\pi}{2}).Y_B(\frac{\pi}{2})$	CNOT.$(X_A(\frac{\pi}{2}).X_A(\frac{\pi}{2})) \otimes X_B(\frac{\pi}{2})$	
$X_A(\frac{\pi}{2}).X_B(\frac{\pi}{2}).X_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2}).Y_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2})$		
$X_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2}).X_B(\frac{\pi}{2}).Y_A(\frac{\pi}{2}).X_B(\frac{\pi}{2}).X_B(\frac{\pi}{2})$		
$Y_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2}).X_A(\frac{\pi}{2}).Y_A(\frac{\pi}{2}).X_A(\frac{\pi}{2})$.CNOT		
$Y_A(\frac{\pi}{2}).X_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2}).X_A(\frac{\pi}{2}).X_B(\frac{\pi}{2}).X_A(\frac{\pi}{2}).Y_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2})$		

Table C.2: Selected fiducials for two qubits gate set tomography.

Preparation Fiducials ($F^{(p)}$)			
1	I	9	$Y_B(\frac{\pi}{2})$
2	$X_B(\frac{\pi}{2})$	10	$Y_A(\frac{\pi}{2}).X_B(\frac{\pi}{2})$
3	$Y_B(\frac{\pi}{2})$	11	$Y_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2})$
4	$X_B(\frac{\pi}{2}).X_B(\frac{\pi}{2})$	12	$Y_A(\frac{\pi}{2}).X_B(\frac{\pi}{2}).X_B(\frac{\pi}{2})$
5	$X_A(\frac{\pi}{2})$	13	$X_A(\frac{\pi}{2}).X_A(\frac{\pi}{2})$
6	$X_A(\frac{\pi}{2}).X_B(\frac{\pi}{2})$	14	$X_A(\frac{\pi}{2}).X_A(\frac{\pi}{2}).X_B(\frac{\pi}{2})$
7	$X_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2})$	15	$X_A(\frac{\pi}{2}).X_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2})$
8	$X_A(\frac{\pi}{2}).X_B(\frac{\pi}{2}).X_B(\frac{\pi}{2})$	16	$X_A(\frac{\pi}{2}).X_A(\frac{\pi}{2}).X_B(\frac{\pi}{2}).X_B(\frac{\pi}{2})$
Measurement Fiducials($F^{(m)}$)			
1	I	7	$X_A(\frac{\pi}{2}).X_A(\frac{\pi}{2})$
2	$X_B(\frac{\pi}{2})$	8	$X_A(\frac{\pi}{2}).X_B(\frac{\pi}{2})$
3	$Y_B(\frac{\pi}{2})$	9	$X_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2})$
4	$X_B(\frac{\pi}{2}).X_B(\frac{\pi}{2})$	10	$Y_A(\frac{\pi}{2}).X_B(\frac{\pi}{2})$
5	$X_A(\frac{\pi}{2})$	11	$Y_A(\frac{\pi}{2}).Y_B(\frac{\pi}{2})$
6	$Y_A(\frac{\pi}{2})$		

Table C.3: The gate set and the Germs are chosen for qutrit GST.

Gate set		
I	H	$X_{01}(\frac{\pi}{2})$
$X_{12}(\frac{\pi}{2})$	$Z_1(\frac{2\pi}{3})$	$Z_2(\frac{2\pi}{3})$
List of germs		
I	H	$X_{01}(\frac{\pi}{2})$
$X_{12}(\frac{\pi}{2})$	$Z_1(\frac{2\pi}{3})$	$Z_2(\frac{2\pi}{3})$
$X_{01}(\frac{\pi}{2})X_{12}(\frac{\pi}{2})$	$H.Z_1(\frac{2\pi}{3})$	$H.Z_2(\frac{2\pi}{3})$
$H.X_{01}(\frac{\pi}{2})$	$X_{12}(\frac{\pi}{2}).Z_2(\frac{2\pi}{3})$	$X_{12}(\frac{\pi}{2}).Z_1(\frac{2\pi}{3})$
$H.X_{12}(\frac{\pi}{2})$	$X_{01}(\frac{\pi}{2}).Z_1(\frac{2\pi}{3})$	

Table C.4: Fiducials for qutrit GST.

	Preparation Fiducials ($F^{(p)}$)
1	I
2	$H.H.H.Z_2(\frac{2\pi}{3}).H$
3	$H.H.H.Z_1(\frac{2\pi}{3}).Z_2(\frac{2\pi}{3}).Z_2(\frac{2\pi}{3}).H$
4	$H.H.H.Z_1(\frac{2\pi}{3}).Z_1(\frac{2\pi}{3}).Z_2(\frac{2\pi}{3}).H$
5	$H.H.H.Z_1(\frac{2\pi}{3}).Z_2(\frac{2\pi}{3}).H$
6	$H.H.H.Z_1(\frac{2\pi}{3}).Z_1(\frac{2\pi}{3}).H$
7	$H.H.H.Z_1(\frac{2\pi}{3}).Z_1(\frac{2\pi}{3}).Z_2(\frac{2\pi}{3}).Z_2(\frac{2\pi}{3}).H$
8	$H.H.H.Z_2(\frac{2\pi}{3}).H.Z_2(\frac{2\pi}{3})$
9	$H.H.H.Z_1(\frac{2\pi}{3}).Z_1(\frac{2\pi}{3}).Z_2(\frac{2\pi}{3}).Z_2(\frac{2\pi}{3}).H.Z_2(\frac{2\pi}{3})$
	Measurement Fiducials($F^{(m)}$)
1	I
2	$Z_2(\frac{2\pi}{3}).H.H.H.Z_2(\frac{2\pi}{3}).H$
3	$H.H.H.Z_1(\frac{2\pi}{3}).Z_2(\frac{2\pi}{3}).H$
4	$H.H.H.Z_2(\frac{2\pi}{3}).H$

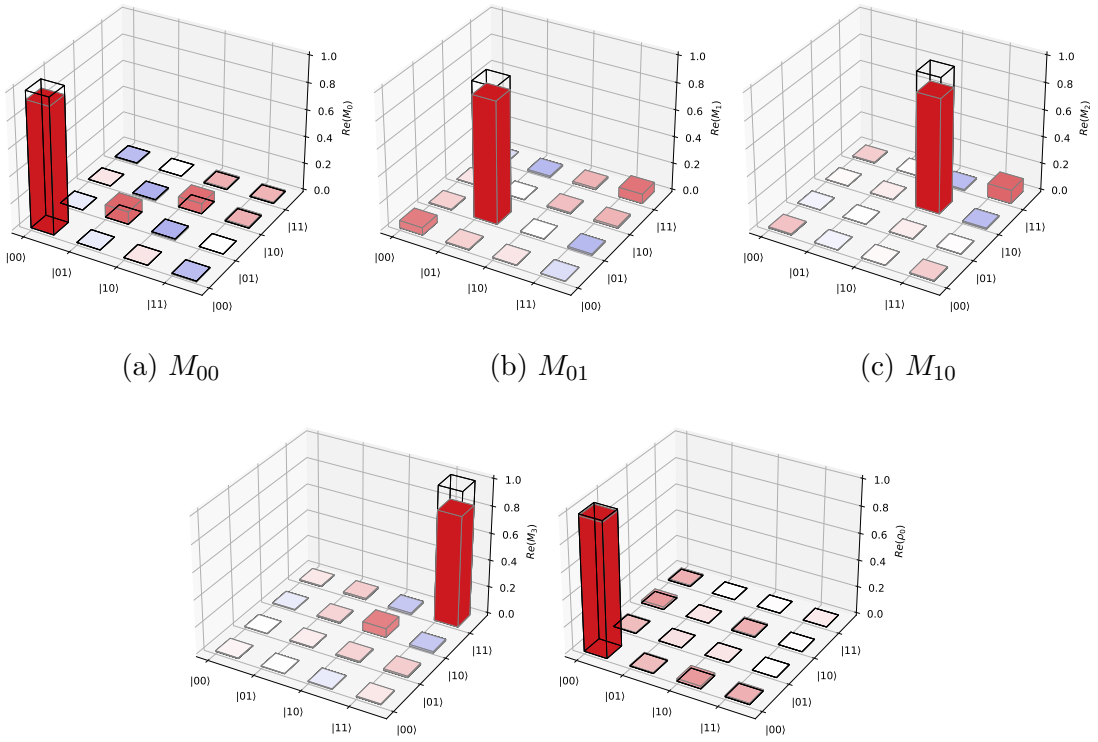
Table C.5: The gate set and the Germs are chosen for gate set tomography of this experiment.

Gate set		
$X_{01}(\frac{\pi}{2})$	$Z_1(\frac{\pi}{2})$	$X_{12}(\frac{\pi}{2})$
$Z_2(\frac{\pi}{2})$	$X_{23}(\frac{\pi}{2})$	$Z_3(\frac{\pi}{2})$
List of germs		
$X_{01}(\frac{\pi}{2})$	$Z_1(\frac{\pi}{2})$	$X_{12}(\frac{\pi}{2})$
$Z_2(\frac{\pi}{2})$	$X_{23}(\frac{\pi}{2})$	$Z_3(\frac{\pi}{2})$
$X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2})$	$X_{12}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2})$
$Z_1(\frac{\pi}{2}).Z_1(\frac{\pi}{2})$	$X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2})$	$X_{23}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
$Z_2(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$		

Table C.6: Fiducials for qutrit gate set tomography.

	Preparation Fiducials ($F^{(p)}$)
1	I
2	$X_{01}(\frac{\pi}{2})$
3	$X_{01}(\frac{\pi}{2}).Z_1(\frac{\pi}{2})$
4	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2})$
5	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2})$
6	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_2(\frac{\pi}{2})$
7	$X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2})$
8	$X_{01}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2})$
9	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2})$
10	$X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
11	$X_{01}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
12	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
13	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
14	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
15	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).Z_3(\frac{\pi}{2})$
16	$X_{01}(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
	Measurement Fiducials($F^{(m)}$)
1	I
2	$Z_1(\frac{\pi}{2}).Z_3(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
3	$X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
4	$X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_3(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).$ $X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2})$
5	$X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_3(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).$ $X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_3(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
6	$X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_3(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).$ $X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
7	$X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).$ $Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2})$
8	$X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).$ $Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_3(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$
9	$X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).X_{12}(\frac{\pi}{2}).$ $Z_1(\frac{\pi}{2}).Z_2(\frac{\pi}{2}).X_{01}(\frac{\pi}{2}).X_{23}(\frac{\pi}{2})$

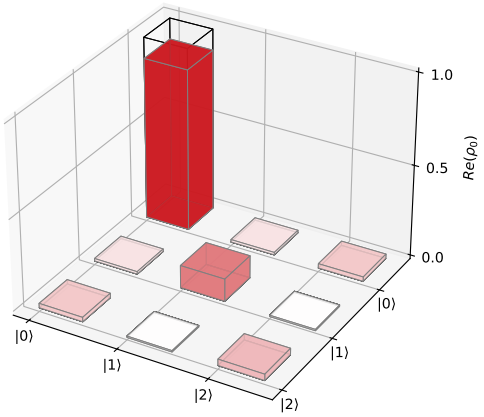
C.2 SPAM operators



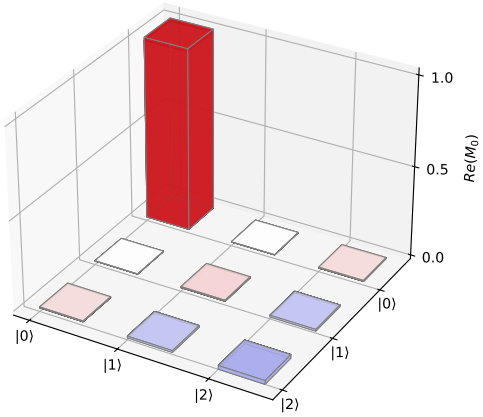
(a) M_{00} (b) M_{01} (c) M_{10}

(d) M_{11} (e) ρ_0

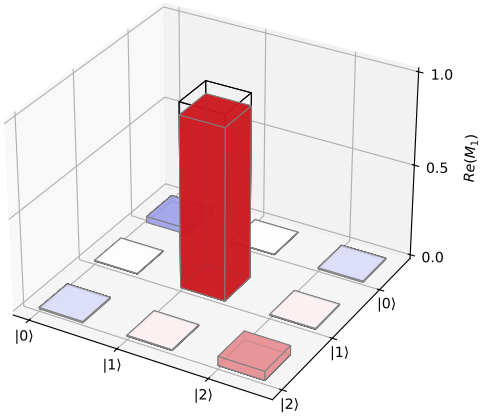
Figure C.1: Reconstructed SPAM operators for two-qubit device.



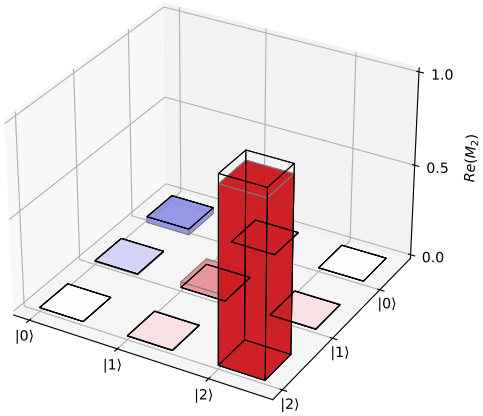
(a) ρ_0



(b) M_{00}

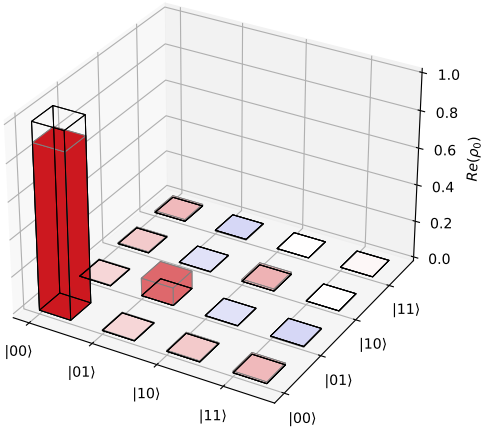


(c) M_{10}

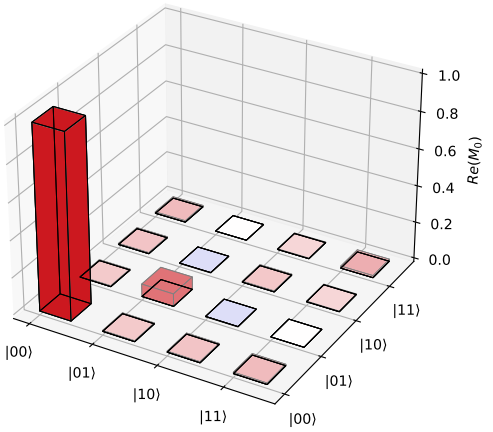


(d) M_{11}

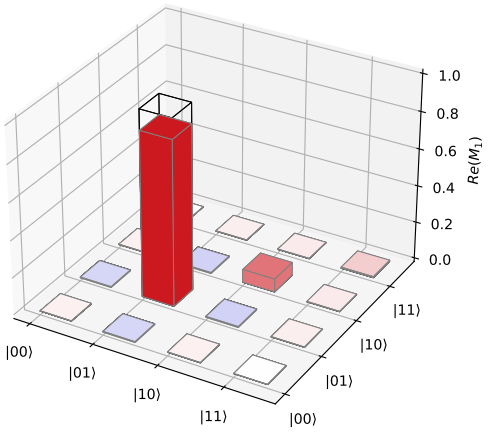
Figure C.2: Estimated SPAM operators for the qutrit device.



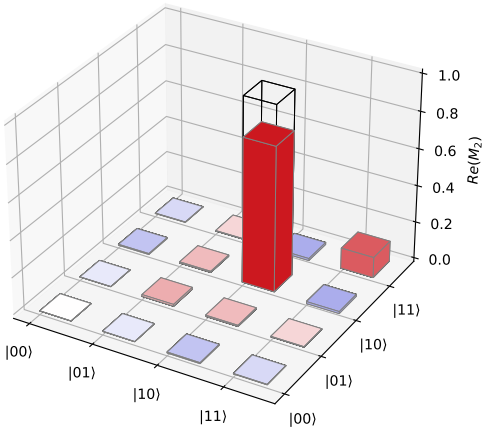
(a) ρ_0



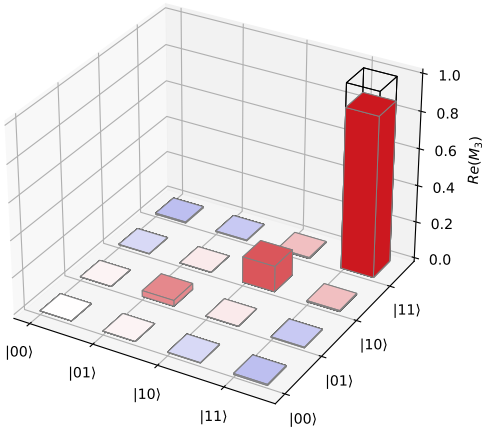
(b) M_0



(c) M_1



(d) M_2

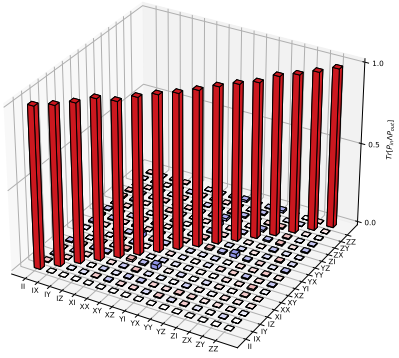
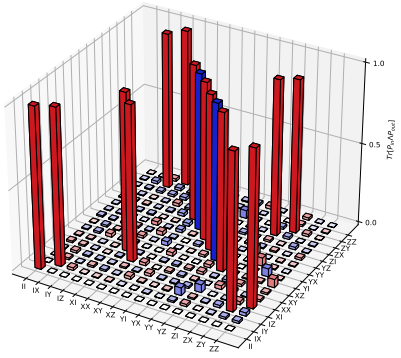
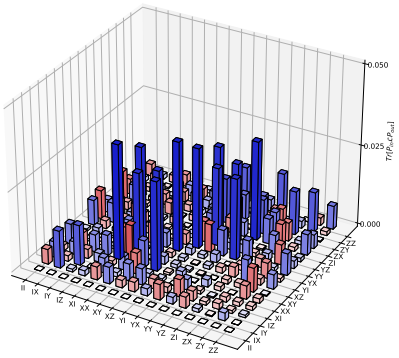
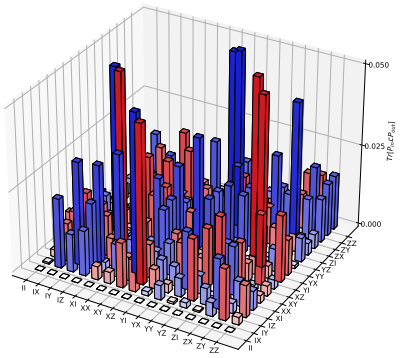
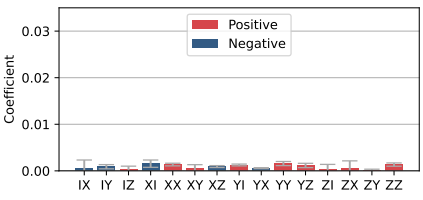
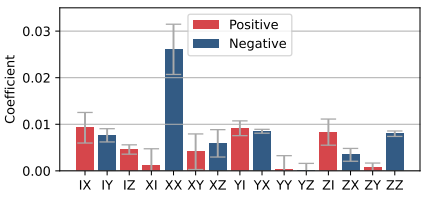


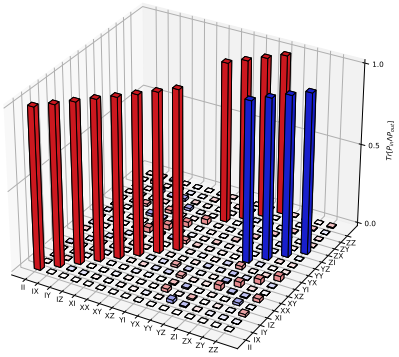
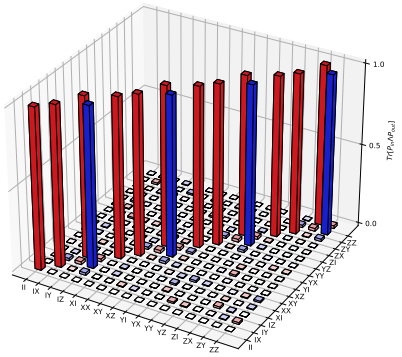
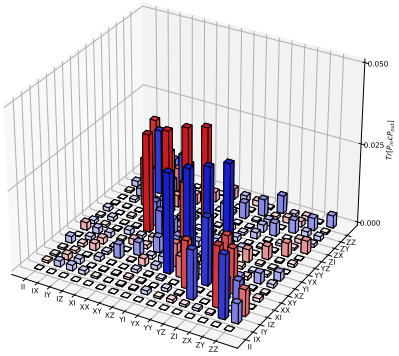
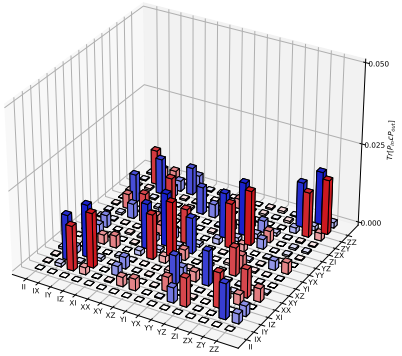
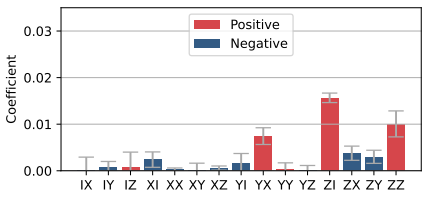
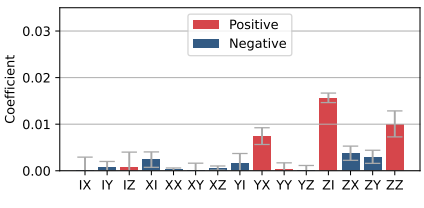
(e) M_3

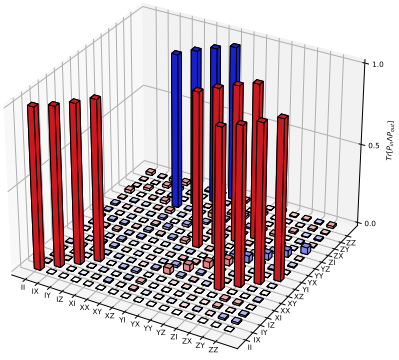
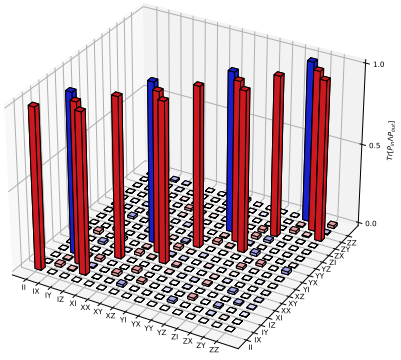
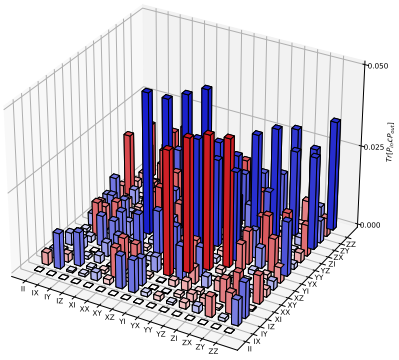
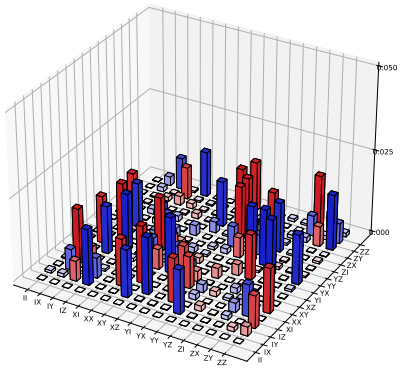
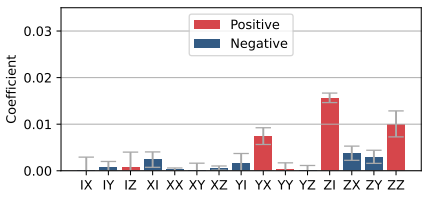
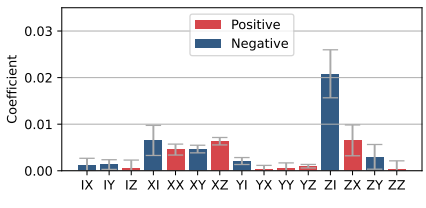
Figure C.3: Estimated SPAM operators for the qudit device.

C.3 Quantum processes

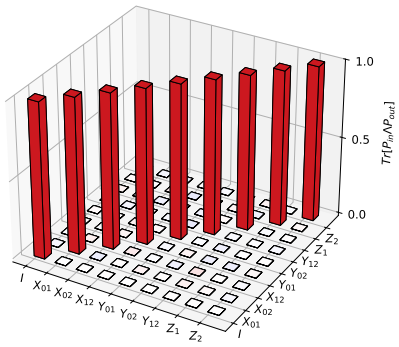
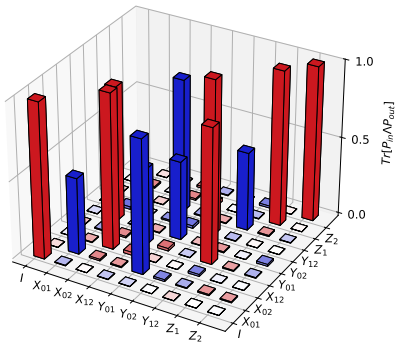
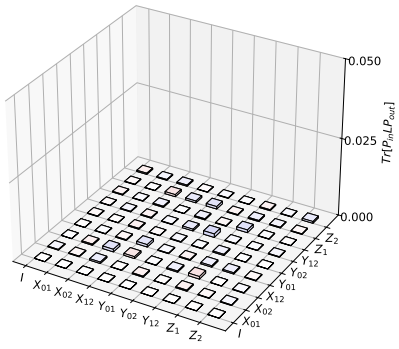
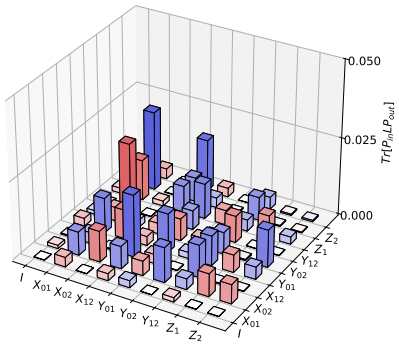
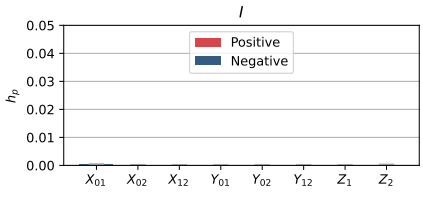
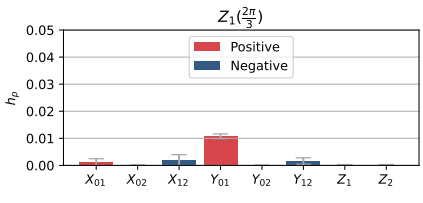
2QB device

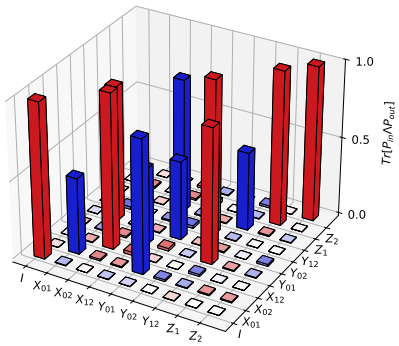
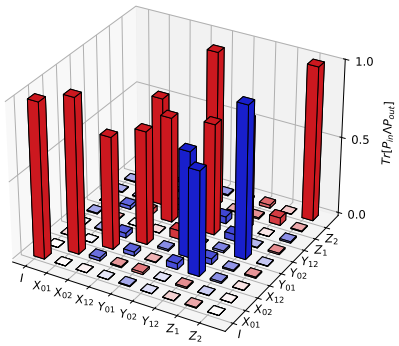
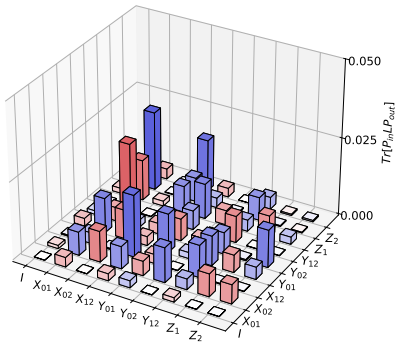
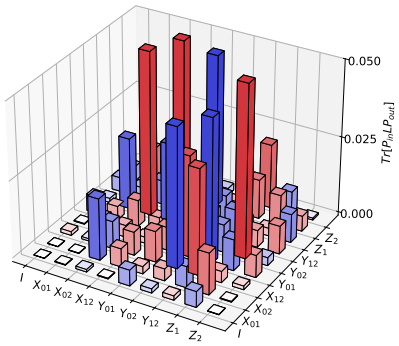
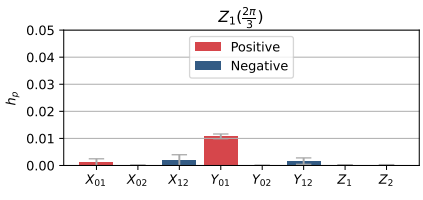
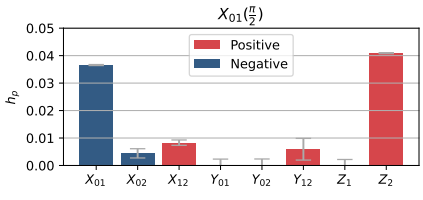
Gate name	I	CNOT
Average infidelity	0.0148(16)	0.0201(16)
Hamiltonian Error	0.059(27)	0.883(23)
Stochastic error	0.518(80)	0.063(18)
Active error	0.422(62)	0.054(25)
PTM		
Error generator		
Hamiltonian projected error generator		

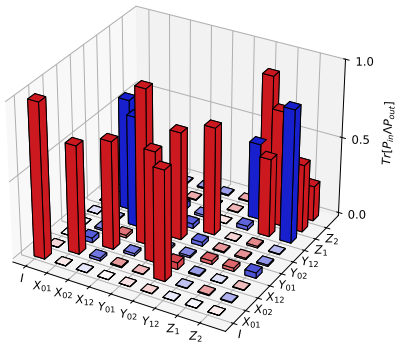
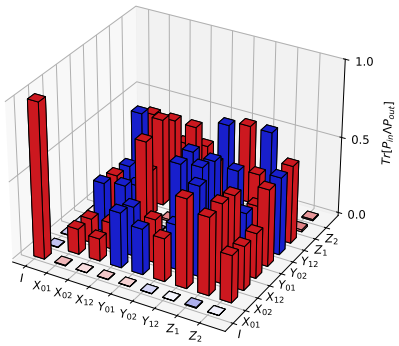
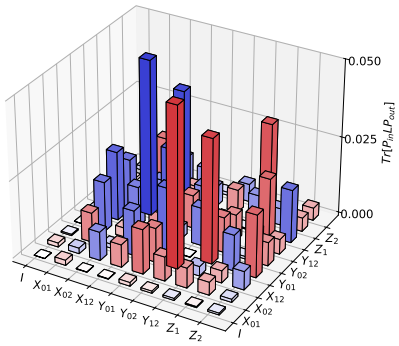
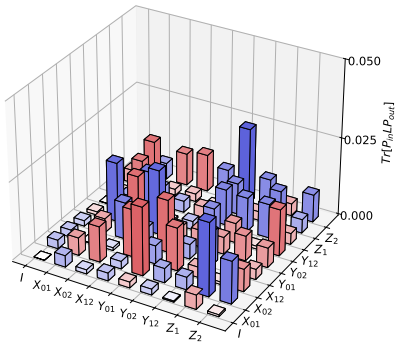
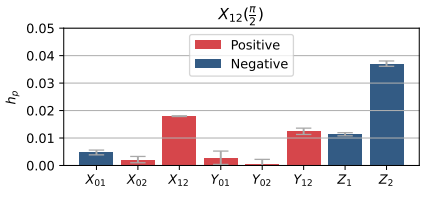
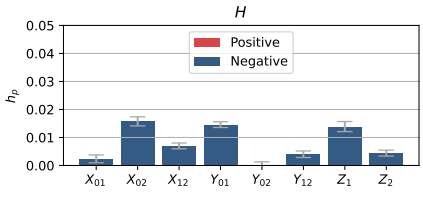
Gate name	$X_A(\frac{\pi}{2})$	$X_B(\frac{\pi}{2})$
Average infidelity	0.00261(95)	0.0010(12)
Hamiltonian Error	0.990(45)	0.9974(73)
Stochastic error	0.005(14)	0.0011(63)
Active error	0.005(32)	0.0015(63)
PTM		
Error generator		
Hamiltonian projected error generator		

Gate name	$Y_A(\frac{\pi}{2})$	$Y_B(\frac{\pi}{2})$
Average infidelity	0.01569(63)	0.0010(19)
Hamiltonian Error	0.710(48)	0.9987(78)
Stochastic error	0.161(80)	0.0005(41)
Active error	0.129(41)	0.0008(29)
PTM		
Error generator		
Hamiltonian projected error generator		

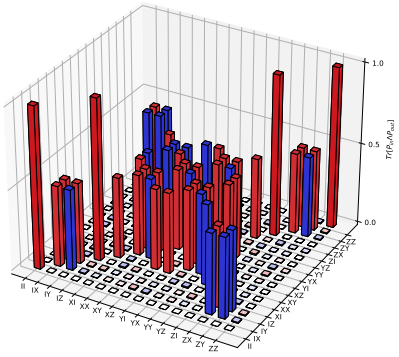
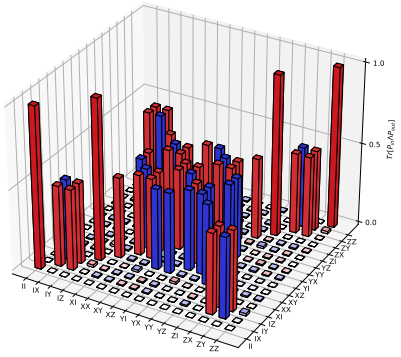
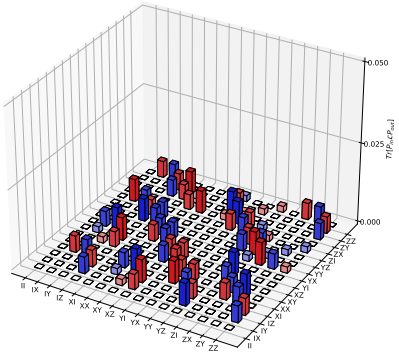
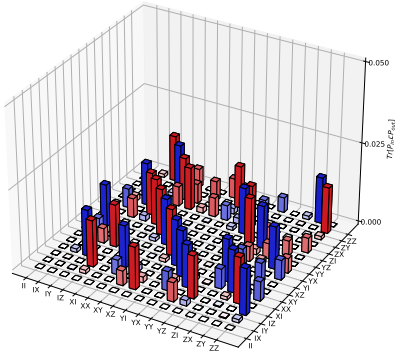
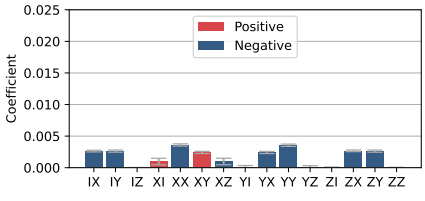
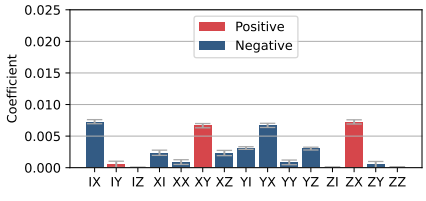
1QD device as a qutrit

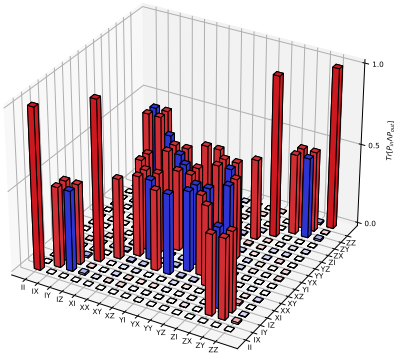
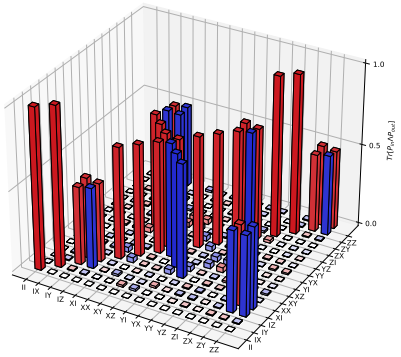
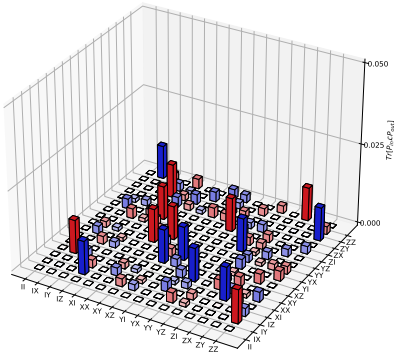
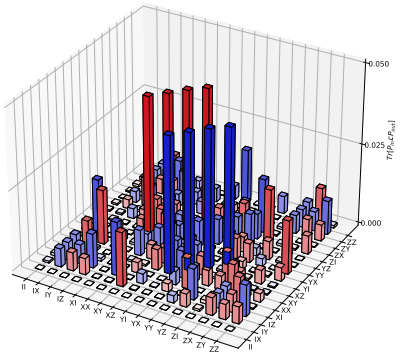
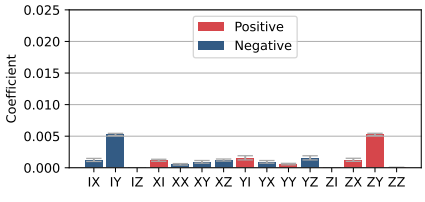
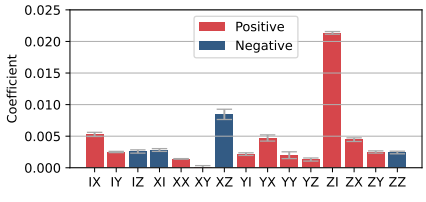
Gate name	I	$Z_1(\frac{2\pi}{3})$
Average infidelity	0.000365(97)	0.000221(57)
Hamiltonian Error	0.066(57)	0.070(15)
Stochastic error	0.28(16)	0.074(45)
Active error	0.66(16)	0.855(46)
PTM		
Error generator		
Hamiltonian projected error generator		

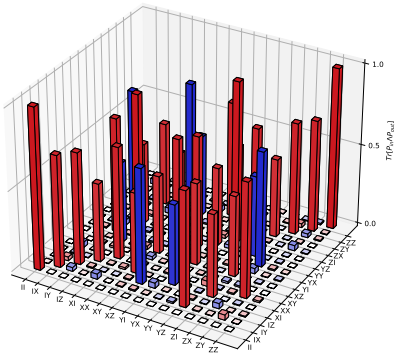
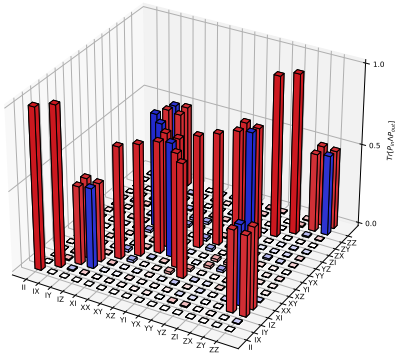
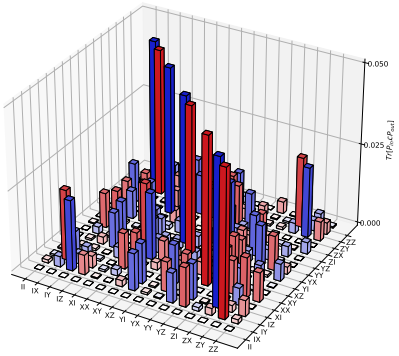
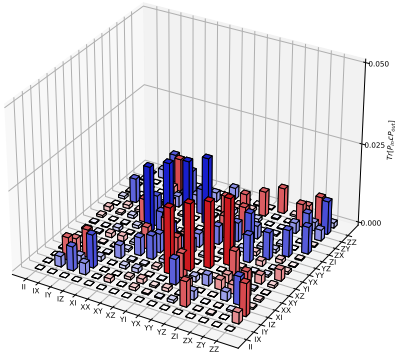
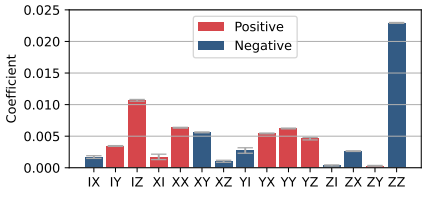
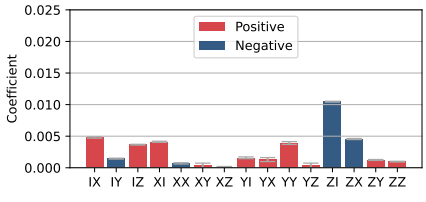
Gate name	$Z_2(\frac{2\pi}{3})$	$X_{01}(\pi/2)$
Average infidelity	0.000274(70)	0.002112(96)
Hamiltonian Error	0.198(40)	0.658(62)
Stochastic error	0.064(42)	0.066(30)
Active error	0.737(50)	0.275(63)
PTM		
Error generator		
Hamiltonian projected error generator		

Gate name	$X_{12}(\pi/2)$	H
Average infidelity	0.00099(7)	0.0032(11)
Hamiltonian Error	0.597(75)	0.399(47)
Stochastic error	0.013(11)	0.152(50)
Active error	0.388(76)	0.449(41)
PTM		
Error generator		
Hamiltonian projected error generator		

1QD device as a qudit

Gate name	$Z_1(\frac{2\pi}{3})$	$Z_2(\frac{2\pi}{3})$
Average infidelity	0.000085(15)	0.000218(70)
Hamiltonian Error	0.999991(78)	0.9999972(94)
Stochastic error	0.000003(44)	0.0000010(97)
Active error	0.000007(73)	0.000002(20)
PTM		
Error generator		
Hamiltonian projected error generator		

Gate name	$Z_3(\frac{\pi}{2})$	$X_{01}(\pi/2)$
Average infidelity	0.000113(80)	0.00498(11)
Hamiltonian Error	0.99998(25)	0.9744(15)
Stochastic error	0.000008(50)	0.01435(58)
Active error	0.00001(21)	0.0112(11)
PTM		
Error generator		
Hamiltonian projected error generator		

Gate name	$X_{12}(\pi/2)$	$X_{23}(\pi/2)$
Average infidelity	0.002966(52)	0.002906(68)
Hamiltonian Error	0.99452(71)	0.9724(49)
Stochastic error	0.00281(70)	0.0161(43)
Active error	0.00268(38)	0.0115(31)
PTM		
Error generator		
Hamiltonian projected error generator		

Bibliography

[1] Eagle's quantum performance progress | IBM Research Blog.

[2] Hands-On with Google's Quantum Computer - Scientific American.

[3] Rajeev Acharya, Igor Aleiner, Richard Allen, Trond I. Andersen, Markus Ansmann, Frank Arute, Kunal Arya, Abraham Asfaw, Juan Atalaya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Joao Basso, Andreas Bengtsson, Sergio Boixo, Gina Bortoli, Alexandre Bourassa, Jenna Bovaird, Leon Brill, Michael Broughton, Bob B. Buckley, David A. Buell, Tim Burger, Brian Burkett, Nicholas Bushnell, Yu Chen, Zijun Chen, Ben Chiaro, Josh Cogan, Roberto Collins, Paul Conner, William Courtney, Alexander L. Crook, Ben Curtin, Dripto M. Debroy, Alexander Del Toro Barba, Sean Demura, Andrew Dunsworth, Daniel Eppens, Catherine Erickson, Lara Faoro, Edward Farhi, Reza Fatemi, Leslie Flores Burgos, Ebrahim Forati, Austin G. Fowler, Brooks Foxen, William Giang, Craig Gidney, Dar Gilboa, Marissa Giustina, Alejandro Grajales Dau, Jonathan A. Gross, Steve Habegger, Michael C. Hamilton, Matthew P. Harrigan, Sean D. Harrington, Oscar Higgott, Jeremy Hilton, Markus Hoffmann, Sabrina Hong, Trent Huang, Ashley Huff, William J. Huggins, Lev B. Ioffe, Sergei V. Isakov, Justin Iveland, Evan Jeffrey, Zhang Jiang, Cody Jones, Pavol Juhas, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Tanuj Khattar, Mostafa Khezri, Mária Kieferová, Seon Kim, Alexei Kitaev, Paul V. Klimov, Andrey R. Klots, Alexander N. Korotkov, Fedor Kostritsa,

John Mark Kreikebaum, David Landhuis, Pavel Laptev, Kim-Ming Lau, Lily Laws, Joonho Lee, Kenny Lee, Brian J. Lester, Alexander Lill, Wayne Liu, Aditya Locharla, Erik Lucero, Fionn D. Malone, Jeffrey Marshall, Orion Martin, Jarrod R. McClean, Trevor Mccourt, Matt McEwen, Anthony Megrant, Bernardo Meurer Costa, Xiao Mi, Kevin C. Miao, Masoud Mohseni, Shirin Montazeri, Alexis Morvan, Emily Mount, Wojciech Mruczkiewicz, Ofer Naaman, Matthew Neeley, Charles Neill, Ani Nersisyan, Hartmut Neven, Michael Newman, Jiun How Ng, Anthony Nguyen, Murray Nguyen, Murphy Yuezhen Niu, Thomas E. O'Brien, Alex Opremcak, John Platt, Andre Petukhov, Rebecca Potter, Leonid P. Pryadko, Chris Quintana, Pedram Roushan, Nicholas C. Rubin, Negar Saei, Daniel Sank, Kannan Sankaragomathi, Kevin J. Satzinger, Henry F. Schurkus, Christopher Schuster, Michael J. Shearn, Aaron Shorter, Vladimir Shvarts, Jindra Skruzny, Vadim Smelyanskiy, W. Clarke Smith, George Sterling, Doug Strain, Marco Szalay, Alfredo Torres, Guifre Vidal, Benjamin Villalonga, Catherine Vollgraff Heidweiller, Theodore White, Cheng Xing, Z. Jamie Yao, Ping Yeh, Juhwan Yoo, Grayson Young, Adam Zalcman, Yaxing Zhang, and Ningfeng Zhu. Suppressing quantum errors by scaling a surface code logical qubit. 7 2022.

[4] Soumik Adhikary, Siddharth Dangwal, and Debanjan Bhowmik. Supervised learning with a quantum classifier using multi-level systems. *Quantum Information Processing*, 19(3):89, 3 2020.

[5] Esma Aïmeur, Gilles Brassard, and Sébastien Gambs. Machine Learning in a Quantum World. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 4013 LNAI, pages 431–442. Springer Verlag, 2006.

[6] Joseph B. Altepeter, Daniel F.V. James, and Paul G. Kwiat. 4 Qubit Quantum

State Tomography. In Jaroslav Řeháček Matteo Paris, editor, *Quantum State Estimation*, pages 113–145. Springer Berlin, Heidelberg, 8 2004.

[7] Edgar Anderson. The Species Problem in Iris. *Annals of the Missouri Botanical Garden*, 23(3):457, 9 1936.

[8] Eric Anschuetz, Jonathan Olson, Alán Aspuru-Guzik, and Yudong Cao. Variational Quantum Factoring. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11413 LNCS:74–85, 8 2018.

[9] B. Apolloni, C. Carvalho, and D. de Falco. Quantum stochastic optimization. *Stochastic Processes and their Applications*, 33(2):233–244, 12 1989.

[10] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G.S.L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman,

Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature 2019 574:7779*, 574(7779):505–510, 10 2019.

[11] Jose Aumentado. Superconducting Parametric Amplifiers: The State of the Art in Josephson Parametric Amplifiers. *IEEE Microwave Magazine*, 21(8):45–59, 8 2020.

[12] B. Jack Copeland. *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life: Plus The Secrets of Enigma*. OXFORD UNIVERSITY PRESS, 2004.

[13] Jonathan M. Baker, Casey Duckering, Pranav Gokhale, Natalie C. Brown, Kenneth R. Brown, and Frederic T. Chong. Improved Quantum Circuits via Intermediate Qutrits. *ACM Transactions on Quantum Computing*, 1(1):1–25, 12 2020.

[14] Adriano Barenco, Charles H. Bennett, Richard Cleve, David P. DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A. Smolin, and Harald Weinfurter. Elementary gates for quantum computation. *Physical Review A*, 52(5):3457–3467, 11 1995.

[15] Rodney J. Bartlett, Stanislaw A. Kucharski, and Jozef Noga. Alternative coupled-cluster ansätze II. The unitary coupled-cluster method. *Chemical Physics Letters*, 155(1):133–140, 2 1989.

[16] Rodney J. Bartlett and Monika Musiał. Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics*, 79(1):291–352, 2 2007.

[17] Atılım Günes Baydin, Barak A. Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *The Journal of Machine Learning Research*, 18(1):5595–5637, 2017.

[18] Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 4(4):043001, 11 2019.

[19] A. J. Berkley, H. Xu, R. C. Ramos, M. A. Gubrud, F. W. Strauch, P. R. Johnson, J. R. Anderson, A. J. Dragt, C. J. Lobb, and F. C. Wellstood. Entangled Macroscopic Quantum States in Two Superconducting Qubits. *Science*, 300(5625):1548–1550, 6 2003.

[20] Kishor Bharti, Tobias Haug, Vlatko Vedral, and Leong-Chuan Kwek. NISQ Algorithm for Semidefinite Programming. 6 2021.

[21] Jacob Biamonte. Lectures on Quantum Tensor Networks. 12 2019.

[22] R. Bianchetti, S. Filipp, M. Baur, J. M. Fink, M. Göppl, P. J. Leek, L. Steffen, A. Blais, and A. Wallraff. Dynamics of dispersive single-qubit readout in circuit quantum electrodynamics. *Physical Review A - Atomic, Molecular, and Optical Physics*, 80(4):043840, 10 2009.

[23] R. Bianchetti, S. Filipp, M. Baur, J. M. Fink, C. Lang, L. Steffen, M. Boissonneault, A. Blais, and A. Wallraff. Control and Tomography of a Three Level Superconducting Artificial Atom. *Physical Review Letters*, 105(22):223601, 11 2010.

[24] Lennart Bittel and Martin Kliesch. Training variational quantum algorithms is NP-hard. *Physical Review Letters*, 127(12), 1 2021.

[25] Alexandre Blais, Arne L. Grimsmo, S. M. Girvin, and Andreas Wallraff. Circuit quantum electrodynamics. *Reviews of Modern Physics*, 93(2):025005, 5 2021.

[26] Alexandre Blais, Ren-Shou Huang, Andreas Wallraff, S. M. Girvin, and R. J. Schoelkopf. Cavity quantum electrodynamics for superconducting electrical

circuits: An architecture for quantum computation. *Physical Review A*, 69(6):062320, 6 2004.

[27] M. S. Blok, V. V. Ramasesh, T. Schuster, K. O'Brien, J. M. Kreikebaum, D. Dahlen, A. Morvan, B. Yoshida, N. Y. Yao, and I. Siddiqi. Quantum Information Scrambling on a Superconducting Qutrit Processor. *Physical Review X*, 11(2):021010, 4 2021.

[28] Robin Blume-Kohout, Marcus P. da Silva, Erik Nielsen, Timothy Proctor, Kenneth Rudinger, Mohan Sarovar, and Kevin Young. A Taxonomy of Small Markovian Errors. *PRX Quantum*, 3(2):020335, 5 2022.

[29] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Kenneth Rudinger, Jonathan Mizrahi, Kevin Fortier, and Peter Maunz. Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nature Communications 2017 8:1*, 8(1):1–13, 2 2017.

[30] Alex Bocharov, Martin Roetteler, and Krysta M. Svore. Factoring with qutrits: Shor's algorithm on ternary and metaplectic quantum architectures. *Physical Review A*, 96(1):012306, 7 2017.

[31] V. Bouchiat, D. Vion, P. Joyez, D. Esteve, and M. H. Devoret. Quantum Coherence with a Single Cooper Pair. *Physica Scripta*, T76(1):165, 1998.

[32] Adam Bouland, Wim van Dam, Hamed Joorati, Iordanis Kerenidis, and Anupam Prakash. Prospects and challenges of quantum finance. 11 2020.

[33] Fernando G. S. L. Brandao and Krysta Svore. Quantum Speed-ups for Semidefinite Programming. *arXiv:1609.05537 [quant-ph]*, pages 1–18, 9 2016.

[34] Colin D. Bruzewicz, John Chiaverini, Robert McConnell, and Jeremy M. Sage. Trapped-ion quantum computing: Progress and challenges. *Applied Physics Reviews*, 6(2):021314, 5 2019.

[35] Jean-Luc Brylinski and Ranee Brylinski. Universal quantum gates. *Mathematics of Quantum Computation*, pages 101–116, 8 2001.

[36] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum Fingerprinting. *Physical Review Letters*, 87(16):167902, 9 2001.

[37] Mark Bydder, Andres Rahal, Gary D. Fullerton, and Graeme M. Bydder. The magic angle effect: A source of artifact, determinant of image contrast, and technique for imaging. *Journal of Magnetic Resonance Imaging*, 25(2):290–300, 2 2007.

[38] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret. Quantum error correction of a qubit encoded in grid states of an oscillator. *Nature*, 584(7821):368–372, 8 2020.

[39] Earl T. Campbell, Barbara M. Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, 12 2016.

[40] Shuxiang Cao, Mustafa Bakr, Giulio Campanaro, Simone D Fasciati, James Wills, Deep Lall, Boris Shteynas, Vivek Chidambaram, Ivan Rungger, and Peter Leek. Emulating two qubits with a four-level transmon qudit for variational quantum algorithms. *Quantum Science and Technology*, 9(3):035003, apr 2024.

[41] Shuxiang Cao, Deep Lall, Mustafa Bakr, Giulio Campanaro, Simone Fasciati, James Wills, Vivek Chidambaram, Boris Shteynas, Ivan Rungger, and Peter

Leek. Efficient characterization of qudit logical gates with gate set tomography using an error-free Virtual-Z-gate model. 10 2022.

[42] Shuxiang Cao, Leonard Wossnig, Brian Vlastakis, Peter Leek, and Edward Grant. Cost-function embedding and dataset encoding for machine learning with parametrized quantum circuits. *Physical Review A*, 101(5):052309, 5 2020.

[43] Ye Cao, Shi-Guo Peng, Chao Zheng, and Gui-Lu Long. Quantum Fourier Transform and Phase Estimation in Qudit System. *Communications in Theoretical Physics*, 55(5):790–794, 5 2011.

[44] Arnaud Carignan-Dugas, Kristine Boone, Joel J Wallman, and Joseph Emerson. From randomized benchmarking experiments to gate-set circuit fidelity: how to interpret randomized benchmarking decay parameters. *New Journal of Physics*, 20(9):092001, 9 2018.

[45] G. Catelani, R. J. Schoelkopf, M. H. Devoret, and L. I. Glazman. Relaxation and frequency shifts induced by quasiparticles in superconducting qubits. *Physical Review B - Condensed Matter and Materials Physics*, 84(6):064517, 8 2011.

[46] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics 2021 3:9*, 3(9):625–644, 8 2021.

[47] Liangyu Chen, Hang-Xi Li, Yong Lu, Christopher W. Warren, Christian J. Križan, Sandoko Kosen, Marcus Rommel, Shahnawaz Ahmed, Amr Osman, Janka Biznárová, Anita Fadavi Roudsari, Benjamin Lienhard, Marco Caputo, Kestutis Grigoras, Leif Grönberg, Joonas Govenius, Anton Frisk Kockum, Per Delsing, Jonas Bylander, and Giovanna Tancredi. Transmon qubit readout

fidelity at the threshold for quantum error correction without a quantum-limited amplifier. 8 2022.

[48] Samuel Yen-Chi Chen, Chao-Han Huck Yang, Jun Qi, Pin-Yu Chen, Xiaoli Ma, and Hsi-Sheng Goan. Variational Quantum Circuits for Deep Reinforcement Learning. 6 2019.

[49] Andrew M. Childs. Universal computation by quantum walk. *Physical Review Letters*, 102(18):180501, 5 2009.

[50] Andrew M. Childs, Richard Cleve, Enrico Deotto, Edward Farhi, Sam Gutmann, and Daniel A. Spielman. Exponential algorithmic speedup by a quantum walk. page 59, 2003.

[51] Andrew M. Childs, Leonard J. Schulman, and Umesh V. Vazirani. Quantum Algorithms for Hidden Nonlinear Structures. pages 395–404, 4 2008.

[52] I. Chiorescu, Y. Nakamura, C. J. P. M. Harmans, and J. E. Mooij. Coherent Quantum Dynamics of a Superconducting Flux Qubit. *Science*, 299(5614):1869–1871, 3 2003.

[53] Man Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 6 1975.

[54] J. M. Chow, J. M. Gambetta, L. Tornberg, Jens Koch, Lev S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Randomized benchmarking and process tomography for gate errors in a solid-state qubit. *Physical Review Letters*, 102(9):090502, 3 2009.

[55] Jerry M. Chow, A. D. Córcoles, Jay M. Gambetta, Chad Rigetti, B. R. Johnson, John A. Smolin, J. R. Rozen, George A. Keefe, Mary B. Rothwell, Mark B.

Ketchen, and M. Steffen. Simple All-Microwave Entangling Gate for Fixed-Frequency Superconducting Qubits. *Physical Review Letters*, 107(8):080502, 8 2011.

[56] P. T. Cochrane, G. J. Milburn, and W. J. Munro. Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping. *Physical Review A*, 59(4):2631, 4 1999.

[57] F. Coester and H. Kümmel. Short-range correlations in nuclear wave functions. *Nuclear Physics*, 17(C):477–485, 6 1960.

[58] John B Conway. *A Course in Functional Analysis*. Springer, 2 edition, 1 1994.

[59] Daniele Cozzolino, Beatrice Da Lio, Davide Bacco, and Leif Katsuo Oxenløwe. High-Dimensional Quantum Communication: Benefits, Progress, and Future Challenges. *Advanced Quantum Technologies*, 2(12):1900038, 12 2019.

[60] Gavin E. Crooks. Gradients of parameterized quantum gates using the parameter-shift rule and gate decomposition. 5 2019.

[61] Pierre-Luc Dallaire-Demers, Jonathan Romero, Libor Veis, Sukin Sim, and Alán Aspuru-Guzik. Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer. *Quantum Science and Technology*, 4(4):045005, 9 2019.

[62] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Physical Review A - Atomic, Molecular, and Optical Physics*, 80(1):012304, 8 2009.

[63] Vincent Danos, Elham Kashefi, and Prakash Panangaden. The measurement calculus. *Journal of the ACM (JACM)*, 54(2), 4 2007.

[64] Etienne de Klerk. Aspects of Semidefinite Programming. 65, 2002.

[65] Nachum Dershowitz and Evgenia Falkovich. A Formalization and Proof of the Extended Church-Turing Thesis. *Electronic Proceedings in Theoretical Computer Science*, 88:72–78, 7 2012.

[66] D. Deutsch. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117, 7 1985.

[67] Yao-Min Di and Hai-Rui Wei. Elementary gates for ternary quantum logic circuit. 5 2011.

[68] Alec Dinerstein, Caroline S Gorham, and Eugene F Dumitrescu. The hybrid topological longitudinal transmon qubit. *Materials for Quantum Technology*, 1(2):021001, 6 2021.

[69] David P. DiVincenzo. Two-bit gates are universal for quantum computation. *Physical Review A*, 51(2):1015, 2 1995.

[70] Shruti Dogra, Arvind, and Kavita Dorai. Determining the parity of a permutation using an experimental NMR qutrit. *Physics Letters A*, 378(46):3452–3456, 10 2014.

[71] Shruti Dogra, Arvind, and Kavita Dorai. Implementation of the quantum Fourier transform on a hybrid qubit-qutrit NMR quantum emulator. *International Journal of Quantum Information*, 13(7), 3 2015.

[72] Shruti Dogra, Kavita Dorai, and Arvind. Experimental demonstration of quantum contextuality on an NMR qutrit. *Physics Letters A*, 380(22-23):1941–1946, 5 2016.

[73] G. J. Dolan. Offset masks for lift-off photoprocessing. *Applied Physics Letters*, 31(5):337–339, 9 1977.

[74] Yuqian Dong, Qiang Liu, Jianhua Wang, Qingshi Li, Xiaoyan Yu, Wen Zheng, Yong Li, Dong Lan, Xinsheng Tan, and Yang Yu. Simulation of Two-Qubit Gates with a Superconducting Qudit. *physica status solidi (b)*, 259(3):2100500, 3 2022.

[75] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, and Dacheng Tao. The Expressive Power of Parameterized Quantum Circuits. 10 2018.

[76] Kai-Bo Duan and S. Sathiya Keerthi. Which Is the Best Multiclass SVM Method? An Empirical Study. pages 278–285. Springer, Berlin, Heidelberg, 2005.

[77] Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90(2):1007–1023, 1 1989.

[78] J. Ebel, T. Joas, M. Schalk, P. Weinbrenner, A. Angerer, J. Majer, and F. Reinhard. Dispersive readout of room-temperature ensemble spin sensors. *Quantum Science and Technology*, 6(3):03LT01, 6 2021.

[79] Salvatore S Elder, Christopher S Wang, Philip Reinhold, Connor T Hann, Kevin S Chou, Brian J Lester, Serge Rosenblum, Luigi Frunzio, Liang Jiang, and Robert J Schoelkopf. High-Fidelity Measurement of Qubits Encoded in Multilevel Superconducting Circuits. *Physical Review X*, 10, 2020.

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

[81] Dimitrios Emmanoulopoulos and Sofija Dimoska. Quantum Machine Learning in Finance: Time Series Forecasting. 2 2022.

[82] Suguru Endo, Simon C. Benjamin, and Ying Li. Practical Quantum Error Mitigation for Near-Future Applications. *Physical Review X*, 8(3):031027, 7 2018.

[83] Yale Fan. A Generalization of the Deutsch-Jozsa Algorithm to Multi-Valued Quantum Logic. In *37th International Symposium on Multiple-Valued Logic (ISMVL'07)*, pages 12–12. IEEE, 5 2007.

[84] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm. 11 2014.

[85] Edward Farhi and Hartmut Neven. Classification with Quantum Neural Networks on Near Term Processors. 2 2018.

[86] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics 1982 21:6*, 21(6):467–488, 6 1982.

[87] Robert Fickler, Radek Lapkiewicz, William N. Plick, Mario Krenn, Christoph Schaeff, Sven Ramelow, and Anton Zeilinger. Quantum entanglement of high angular momenta. *Science*, 338(6107):640–643, 11 2012.

[88] Maria Andreea Filip and Alex J.W. Thom. A stochastic approach to unitary coupled cluster. *The Journal of Chemical Physics*, 153(21):214106, 12 2020.

[89] R. A. FISHER. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2):179–188, 9 1936.

[90] C. Flühmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. P. Home. Encoding a qubit in a trapped-ion mechanical oscillator. *Nature*, 566(7745):513–517, 2 2019.

[91] M. A. Fogarty, M. Veldhorst, R. Harper, C. H. Yang, S. D. Bartlett, S. T. Flammia, and A. S. Dzurak. Non-exponential Fidelity Decay in Randomized Benchmarking with Low-Frequency Noise. *Physical Review A - Atomic, Molecular, and Optical Physics*, 92(2), 2 2015.

[92] Austin G. Fowler, Ashley M. Stephens, and Peter Groszkowski. High threshold universal quantum computation on the surface code. *Physical Review A - Atomic, Molecular, and Optical Physics*, 80(5), 3 2008.

[93] D S França and A K Hashagen. Approximate randomized benchmarking for finite groups. *Journal of Physics A: Mathematical and Theoretical*, 51(39):395302, 9 2018.

[94] Jay Gambetta, Alexandre Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting. *Physical Review A*, 74(4):042318, 10 2006.

[95] Z. Gedik, I. A. Silva, B. Çakmak, G. Karpat, E. L. G. Vidoto, D. O. Soares-Pinto, E. R. deAzevedo, and F. F. Fanchini. Computational speed-up with a single qudit. *Scientific Reports*, 5(1):14671, 12 2015.

[96] K. Geerlings, Z. Leghtas, I. M. Pop, S. Shankar, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret. Demonstrating a Driven Reset Protocol for a Superconducting Qubit. *Physical Review Letters*, 110(12):120501, 3 2013.

[97] Vlad Gheorghiu. Standard form of qudit stabilizer groups. *Physics Letters A*, 378(5-6):505–509, 1 2014.

[98] Vittorio Giovannetti. Quantum-Enhanced Measurements: Beating the Standard Quantum Limit. *Science*, 306(5700):1330–1336, 11 2004.

[99] Andrew N. Glaudell, Neil J. Ross, and Jacob M. Taylor. Canonical forms for single-qutrit Clifford+T operators. *Annals of Physics*, 406:54–70, 7 2019.

[100] Noah Goss, Alexis Morvan, Brian Marinelli, Bradley K. Mitchell, Long B. Nguyen, Ravi K. Naik, Larry Chen, Christian Jünger, John Mark Kreikebaum, David I. Santiago, Joel J. Wallman, and Irfan Siddiqi. High-fidelity qutrit entangling gates for superconducting circuits. *Nature Communications 2022 13:1*, 13(1):1–6, 12 2022.

[101] Daniel Gottesman. A Theory of Fault-Tolerant Quantum Computation. *Physical Review A - Atomic, Molecular, and Optical Physics*, 57(1):127–137, 2 1997.

[102] Daniel Gottesman. Fault-Tolerant Quantum Computation with Higher-Dimensional Systems. *Chaos, Solitons & Fractals*, 10(10):1749–1758, 2 1998.

[103] Edward Grant, Marcello Benedetti, Shuxiang Cao, Andrew Hallam, Joshua Lockhart, Vid Stojevic, Andrew G. Green, and Simone Severini. Hierarchical quantum classifiers. *npj Quantum Information*, 4(1):65, 4 2018.

[104] Daniel Greenbaum. Introduction to Quantum Gate Set Tomography. 9 2015.

[105] Lov K. Grover. Synthesis of quantum superpositions by quantum computation. *Physical Review Letters*, 85(6):1334–1337, 8 2000.

[106] Yanwu Gu, Rajesh Mishra, Berthold Georg Englert, and Hui Khoon Ng. Randomized linear gate-set tomography. *PRX Quantum*, 2(3):030328, 9 2021.

[107] E. L. Hahn. Spin Echoes. *Physical Review*, 80(4):580, 11 1950.

[108] Aram Harrow and John Napp. Low-depth gradient measurements can improve convergence in variational hybrid quantum-classical algorithms. 1 2019.

[109] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for solving linear systems of equations. *Physical Review Letters*, 103(15), 11 2008.

[110] Akel Hashim, Ravi K. Naik, Alexis Morvan, Jean Loup Ville, Bradley Mitchell, John Mark Kreikebaum, Marc Davis, Ethan Smith, Costin Iancu, Kevin P. O'Brien, Ian Hincks, Joel J. Wallman, Joseph Emerson, and Irfan Siddiqi. Randomized Compiling for Scalable Quantum Computing on a Noisy Superconducting Quantum Processor. *Physical Review X*, 11(4):041039, 12 2021.

[111] Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, 3 2019.

[112] Christoph Helmberg, Franz Rendl, Robert J. Vanderbei, and Henry Wolkowicz. An Interior-Point Method for Semidefinite Programming. *http://dx.doi.org/10.1137/0806020*, 6(2):342–361, 7 2006.

[113] Dylan A Herman, Cody Googin, Xiaoyuan Liu, Alexey Galda, Ilya Safro, Yue Sun, Marco Pistoia, Yuri Alexeev, JPMorgan Chase Bank, New York, and Usa dylanaherman. A Survey of Quantum Computing for Finance. 1 2022.

[114] Mark Hodson, Brendan Ruck, Hugh Ong, David Garvin, Stefan Dulman, and Rigetti Computing. Portfolio rebalancing experiments using the Quantum Alternating Operator Ansatz. 11 2019.

[115] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Physical Review A*, 60(3):1888–1898, 9 1999.

[116] A. A. Houck, J. A. Schreier, B. R. Johnson, J. M. Chow, Jens Koch, J. M. Gambetta, D. I. Schuster, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Controlling the spontaneous emission of a superconducting transmon qubit. *Physical Review Letters*, 101(8):080502, 8 2008.

[117] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. P. Song, C.-L. Zou, S. M. Girvin, L-M. Duan, and L. Sun. Quantum error correction and universal gate set operation on a binomial bosonic logical qubit. *Nature Physics*, 15(5):503–508, 5 2019.

[118] Wenhui Huang, Yuxuan Zhou, Ziyu Tao, Libo Zhang, Song Liu, Yuanzhen Chen, Tongxing Yan, and Dapeng Yu. A superconducting coplanar waveguide ring resonator as quantum bus for circuit quantum electrodynamics. *Applied Physics Letters*, 118(18):184001, 5 2021.

[119] Thomas Hubregtsen, Frederik Wilde, Shozab Qasim, and Jens Eisert. Single-component gradient rules for variational quantum algorithms. *Quantum Science and Technology*, 7(3):035008, 7 2022.

[120] William Huggins, Piyush Patel, K. Birgitta Whaley, E. Miles Stoudenmire, Piyush Patil, Bradley Mitchell, K. Birgitta Whaley, and E. Miles Stoudenmire. Towards Quantum Machine Learning with Tensor Networks. *Quantum Science and Technology*, 4(2):024001, 1 2018.

[121] S. S. Ivanov, H. S. Tonchev, and N. V. Vitanov. Time-efficient implementation of quantum search with qudits. *Physical Review A*, 85(6):062321, 6 2012.

[122] Artur F. Izmaylov, Robert A. Lang, and Tzu-Ching Yen. Analytic gradients in variational quantum algorithms: Algebraic extensions of the parameter-shift rule to general unitary transformations. *Physical Review A*, 104(6), 7 2021.

[123] Mahnaz Jafarzadeh, Mahnaz Jafarzadeh, Ya Dong Wu, Ya Dong Wu, Ya Dong Wu, Yuval R. Sanders, Barry C. Sanders, and Barry C. Sanders. Randomized benchmarking for qudit Clifford gates. *New Journal of Physics*, 22(6):063014, 6 2020.

[124] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 12 1972.

[125] Jay Gambetta. *IFF Spring School 2013 Quantum Information Processing Lecture Notes*, volume 339. 2013.

[126] Min Jiang, Shunlong Luo, and Shuangshuang Fu. Channel-state duality. *Physical Review A*, 87(2):022310, 2 2013.

[127] I. Jiř and E. K. Ĉíž. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *The Journal of Chemical Physics*, 45(11):4256, 5 2004.

[128] Don Johnson, Alfred Menezes, and Scott Vanstone. The Elliptic Curve Digital Signature Algorithm (ECDSA). *International Journal of Information Security 2001 1:1*, 1(1):36–63, 1 2014.

[129] Jonathan A. Jones. Quantum computing with NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 59(2):91–120, 8 2011.

[130] B. D. Josephson. Possible new effects in superconductive tunnelling. *Physics Letters*, 1(7):251–253, 7 1962.

[131] Richard Jozsa. Fidelity for Mixed Quantum States. *Journal of Modern Optics*, 41(12):2315–2323, 12 1994.

[132] Tadashi Kadowaki and Hidetoshi Nishimori. Quantum annealing in the transverse Ising model. *Physical Review E*, 58(5):5355, 11 1998.

[133] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow, and Jay M. Gambetta. Hardware-efficient variational

quantum eigensolver for small molecules and quantum magnets. Technical Report 7671, 2017.

[134] Amir H. Karamlou, William A. Simon, Amara Katabarwa, Travis L. Scholten, Borja Peropadre, and Yudong Cao. Analyzing the performance of variational quantum factoring on a superconducting quantum processor. *npj Quantum Information 2021 7:1*, 7(1):1–6, 10 2021.

[135] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I. C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J J O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and John M. Martinis. Optimal quantum control using randomized benchmarking. *Physical Review Letters*, 112(24):1–5, 2014.

[136] Werner Kern. The Evolution of Silicon Wafer Cleaning Technology. *Journal of The Electrochemical Society*, 137(6):1887–1892, 6 1990.

[137] E. O. Kiktenko, A. K. Fedorov, O. V. Man'ko, and V. I. Man'ko. Multilevel superconducting circuits as two-qubit systems: Operations, state preparation, and entropic inequalities. *Physical Review A - Atomic, Molecular, and Optical Physics*, 91(4):042312, 4 2015.

[138] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. 12 2014.

[139] A. Kitaev, A. Shen, and M. Vyalyi. *Classical and Quantum Computation*, volume 47 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 5 2002.

[140] Morten Kjaergaard, Mollie E. Schwartz, Jochen Braumüller, Philip Krantz, Joel I.-J. Wang, Simon Gustavsson, and William D. Oliver. Superconducting

Qubits: Current State of Play. *Annual Review of Condensed Matter Physics*, 11(1):369–395, 3 2020.

[141] P. V. Klimov, J. Kelly, Z. Chen, M. Neeley, A. Megrant, B. Burkett, R. Barends, K. Arya, B. Chiaro, Yu Chen, A. Dunsworth, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, T. Huang, E. Jeffrey, Erik Lucero, J. Y. Mutus, O. Naaman, C. Neill, C. Quintana, P. Roushan, Daniel Sank, A. Vainsencher, J. Wenner, T. C. White, S. Boixo, R. Babbush, V. N. Smelyanskiy, H. Neven, and John M. Martinis. Fluctuations of Energy-Relaxation Times in Superconducting Qubits. *Physical Review Letters*, 121(9):090502, 8 2018.

[142] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized Benchmarking of Quantum Gates. *Physical Review A - Atomic, Molecular, and Optical Physics*, 77(1), 7 2007.

[143] Jens Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Charge-insensitive qubit design derived from the Cooper pair box. *Physical Review A - Atomic, Molecular, and Optical Physics*, 76(4):1–21, 2007.

[144] Anton Frisk Kockum and Franco Nori. *Fundamentals and Frontiers of the Josephson Effect*, volume 286 of *Springer Series in Materials Science*. Springer International Publishing, Cham, 2019.

[145] Bálint Koczor and Simon C. Benjamin. Quantum analytic descent. *Physical Review Research*, 4(2):023017, 6 2022.

[146] S B Kotsiantis and P E Pintelas. Logitboost of Simple Bayesian Classifier. *Informatica*, 29:53–59, 2005.

[147] Jakob S. Kottmann, Abhinav Anand, and Alán Aspuru-Guzik. A feasible approach for automatically differentiable unitary coupled-cluster on quantum computers. *Chemical Science*, 12(10):3497–3508, 3 2021.

[148] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver. A quantum engineer's guide to superconducting qubits. *Applied Physics Reviews*, 6(2):021318, 6 2019.

[149] Sebastian Krinner, Nathan Lacroix, Ants Remm, Agustin Di Paolo, Elie Genois, Catherine Leroux, Christoph Hellings, Stefania Lazar, Francois Swiadek, Johannes Herrmann, Graham J. Norris, Christian Kraglund Andersen, Markus Müller, Alexandre Blais, Christopher Eichler, and Andreas Wallraff. Realizing repeated quantum error correction in a distance-three surface code. *Nature 2022 605:7911*, 605(7911):669–674, 5 2022.

[150] Rajath Krishna, Vishesh Makwana, and Ananda Padhmanabhan Suresh. A Generalization of Bernstein-Vazirani Algorithm to Qudit Systems. 9 2016.

[151] Michael Kues, Christian Reimer, Piotr Roztocki, Luis Romero Cortés, Stefania Sciara, Benjamin Wetzel, Yanbing Zhang, Alfonso Cino, Sai T. Chu, Brent E. Little, David J. Moss, Lucia Caspani, José Azaña, and Roberto Morandotti. On-chip generation of high-dimensional entangled quantum states and their coherent control. *Nature 2017 546:7660*, 546(7660):622–626, 6 2017.

[152] Joonho Lee, William J. Huggins, Martin Head-Gordon, and K. Birgitta Whaley. Generalized Unitary Coupled Cluster Wave functions for Quantum Computation. *Journal of Chemical Theory and Computation*, 15(1):311–324, 1 2019.

[153] Joonho Lee, David W. Small, Evgeny Epifanovsky, and Martin Head-Gordon. Coupled-Cluster Valence-Bond Singles and Doubles for Strongly Correlated

Systems: Block-Tensor Based Implementation and Application to Oligoacenes. *Journal of Chemical Theory and Computation*, 13(2):602–615, 2 2017.

[154] Joonho Lee, David W. Small, and Martin Head-Gordon. Open-shell coupled-cluster valence-bond theory augmented with an independent amplitude approximation for three-pair correlations: Application to a model oxygen-evolving complex and single molecular magnet. *The Journal of Chemical Physics*, 149(24):244121, 12 2018.

[155] A.W. Leissa. The historical bases of the Rayleigh and Ritz methods. *Journal of Sound and Vibration*, 287(4-5):961–978, 11 2005.

[156] Norbert M. Linke, Dmitri Maslov, Martin Roetteler, Shantanu Debnath, Caroline Figgatt, Kevin A. Landsman, Kenneth Wright, and Christopher Monroe. Experimental comparison of two quantum computing architectures. *Proceedings of the National Academy of Sciences*, 114(13):3305–3310, 3 2017.

[157] Ding Liu, Shi-Ju Ran, Peter Wittek, Cheng Peng, Raul Blázquez García, Gang Su, and Maciej Lewenstein. Machine learning by unitary tensor network of hierarchical tree structure. *New Journal of Physics*, 21(7):073059, 7 2019.

[158] Tong Liu, Bao-Qing Guo, Chang-Shui Yu, and Wei-Ning Zhang. One-step implementation of a hybrid Fredkin gate with quantum memories and single superconducting qubit in circuit QED and its applications. *Optics Express*, 26(4):4498, 2 2018.

[159] Hsuan-Hao Lu, Zixuan Hu, Mohammed Saleh Alshaykh, Alexandria Jeanine Moore, Yuchen Wang, Poolad Imany, Andrew Marc Weiner, and Sabre Kais. Quantum Phase Estimation with Time-Frequency Qudits in a Single Photon. *Advanced Quantum Technologies*, 3(2):1900074, 2 2020.

[160] MingXing Luo and XiaoJun Wang. Universal quantum computation with qudits. *Science China Physics, Mechanics & Astronomy*, 57(9):1712–1717, 9 2014.

[161] Mateusz T. Madzik, Serwan Asaad, Akram Youssry, Benjamin Joecker, Kenneth M. Rudinger, Erik Nielsen, Kevin C. Young, Timothy J. Proctor, Andrew D. Baczewski, Arne Laucht, Vivien Schmitt, Fay E. Hudson, Kohei M. Itoh, Alexander M. Jakob, Brett C. Johnson, David N. Jamieson, Andrew S. Dzurak, Christopher Ferrie, Robin Blume-Kohout, and Andrea Morello. Precision tomography of a three-qubit donor quantum processor in silicon. *Nature*, 601(7893):348–353, 1 2022.

[162] Easwar Magesan, J. M. Gambetta, and Joseph Emerson. Scalable and Robust Randomized Benchmarking of Quantum Processes. *Physical Review Letters*, 106(18):180504, 5 2011.

[163] Easwar Magesan, Jay M. Gambetta, and Joseph Emerson. Characterizing quantum gates via randomized benchmarking. *Physical Review A*, 85(4):042311, 4 2012.

[164] Easwar Magesan, Jay M. Gambetta, B. R. Johnson, Colm A. Ryan, Jerry M. Chow, Seth T. Merkel, Marcus P. da Silva, George A. Keefe, Mary B. Rothwell, Thomas A. Ohki, Mark B. Ketchen, and M. Steffen. Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking. *Physical Review Letters*, 109(8):080505, 8 2012.

[165] P. Magnard, P. Kurpiers, B. Royer, T. Walter, J.-C. Besse, S. Gasparinetti, M. Pechal, J. Heinsoo, S. Storz, A. Blais, and A. Wallraff. Fast and Unconditional All-Microwave Reset of a Superconducting Qubit. *Physical Review Letters*, 121(6):060502, 8 2018.

[166] J. Majer, J. M. Chow, J. M. Gambetta, Jens Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Coupling superconducting qubits via a cavity bus. *Nature*, 449(7161):443–447, 9 2007.

[167] Yuriy Makhlin, Gerd Schoen, and Alexander Shnirman. Quantum state engineering with Josephson-junction devices. *Rev. Mod. Phys.*, 73(2):357–400, 11 2000.

[168] Moein Malekakhlagh, Easwar Magesan, and David C. McKay. First-principles analysis of cross-resonance gate operation. *Physical Review A*, 102(4):042605, 10 2020.

[169] Mehul Malik, Manuel Erhard, Marcus Huber, Mario Krenn, Robert Fickler, and Anton Zeilinger. Multi-photon entanglement in high dimensions. *Nature Photonics 2016 10:4*, 10(4):248–252, 2 2016.

[170] Richard M. Martin. *Electronic Structure*. Cambridge University Press, 4 2004.

[171] L. A. Martinez, Z. Peng, D. Appelö, D. M. Tennant, N. Anders Petersson, J. L DuBois, and Y. J. Rosen. Noise-specific beats in the higher-level Ramsey curves of a transmon qubit. 11 2022.

[172] John M. Martinis, S. Nam, J. Aumentado, and C. Urbina. Rabi Oscillations in a Large Josephson-Junction Qubit. *Physical Review Letters*, 89(11):117901, 8 2002.

[173] John M. Martinis and Kevin Osborne. Superconducting Qubits and the Physics of Josephson Junctions. 2 2004.

[174] Sam McArdle, Suguru Endo, Alan Aspuru-Guzik, Simon Benjamin, and Xiao Yuan. Quantum computational chemistry. 8 2018.

[175] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1):4812, 12 2018.

[176] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, 2 2016.

[177] D T Mcclure, Hanhee Paik, L S Bishop, M Steffen, Jerry M Chow, and Jay M Gambetta. Rapid Driven Reset of a Qubit Readout Resonator. *PHYSICAL REVIEW APPLIED*, 5:11001, 2016.

[178] M. McEwen, D. Kafri, Z. Chen, J. Atalaya, K. J. Satzinger, C. Quintana, P. V. Klimov, D. Sank, C. Gidney, A. G. Fowler, F. Arute, K. Arya, B. Buckley, B. Burkett, N. Bushnell, B. Chiaro, R. Collins, S. Demura, A. Dunsworth, C. Erickson, B. Foxen, M. Giustina, T. Huang, S. Hong, E. Jeffrey, S. Kim, K. Kechedzhi, F. Kostritsa, P. Laptev, A. Megrant, X. Mi, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Niu, A. Paler, N. Redd, P. Roushan, T. C. White, J. Yao, P. Yeh, A. Zalcman, Yu Chen, V. N. Smelyanskiy, John M. Martinis, H. Neven, J. Kelly, A. N. Korotkov, A. G. Petukhov, and R. Barends. Removing leakage-induced correlated errors in superconducting quantum error correction. *Nature Communications 2021 12:1*, 12(1):1–7, 3 2021.

[179] David C. McKay, Sarah Sheldon, John A. Smolin, Jerry M. Chow, and Jay M. Gambetta. Three-Qubit Randomized Benchmarking. *Physical Review Letters*, 122(20):200502, 5 2019.

[180] David C. McKay, Christopher J. Wood, Sarah Sheldon, Jerry M. Chow, and Jay M. Gambetta. Efficient Z gates for quantum computing. *Physical Review A*, 96(2):022330, 8 2017.

[181] Marios H. Michael, Matti Silveri, R. T. Brierley, Victor V. Albert, Juha Salmilehto, Liang Jiang, and S. M. Girvin. New class of quantum error-correcting codes for a bosonic mode. *Physical Review X*, 6(3):031006, 7 2016.

[182] Mazyar Mirrahimi, Zaki Leghtas, Victor V. Albert, Steven Touzard, Robert J. Schoelkopf, Liang Jiang, and Michel H. Devoret. Dynamically protected cat-qubits: a new paradigm for universal quantum computation. *New Journal of Physics*, 16(4):045014, 4 2014.

[183] Kosuke Mitarai and Keisuke Fujii. Methodology for replacing indirect measurements with direct measurements. *Physical Review Research*, 1(1):013006, 8 2019.

[184] Kosuke Mitarai, Makoto Negoro, Masahiro Kitagawa, and Keisuke Fujii. Quantum Circuit Learning. *Physical Review A*, 98(3):032309, 3 2018.

[185] Gabriela Mogos. The Deutsch-Josza Algorithm for n-Qudits. In *Proceedings of the 7th Conference on 7th WSEAS International Conference on Applied Computer Science - Volume 7*, pages 49–51, 2007.

[186] A. Morvan, V. V. Ramasesh, M. S. Blok, J. M. Kreikebaum, K. O'Brien, L. Chen, B. K. Mitchell, R. K. Naik, D. I. Santiago, and I. Siddiqi. Qutrit Randomized Benchmarking. *Physical Review Letters*, 126(21):210504, 5 2021.

[187] Ahmed Abid Moueddene, Nader Khammassi, Sebastian Feld, and Said Hamdioui. A context-aware gate set tomography characterization of superconducting qubits. 3 2021.

[188] Clemens Müller, Jared H Cole, and Jürgen Lisenfeld. Towards understanding two-level-systems in amorphous solids: insights from quantum circuits. *Reports on Progress in Physics*, 82(12):124501, 12 2019.

[189] Y. Nakamura, Yu A. Pashkin, and J. S. Tsai. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature*, 398(6730):786–788, 4 1999.

[190] Ken M. Nakanishi, Keisuke Fujii, and Synge Todo. Sequential minimal optimization for quantum-classical hybrid algorithms. 3 2019.

[191] Erik Nielsen, John King Gamble, Kenneth Rudinger, Travis Scholten, Kevin Young, and Robin Blume-Kohout. Gate Set Tomography. *Quantum*, 5:557, 10 2021.

[192] Erik Nielsen, Kenneth Rudinger, Timothy Proctor, Antonio Russo, Kevin Young, and Robin Blume-Kohout. Probing quantum processor performance with pyGSTi. *Quantum Science and Technology*, 5(4), 2 2020.

[193] Michael A. Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, 303(4):249–252, 10 2002.

[194] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000.

[195] A. O. Niskanen, K. Harrabi, F. Yoshihara, Y. Nakamura, S. Lloyd, and J. S. Tsai. Quantum coherent tunable coupling of superconducting qubits. *Science*, 316(5825):723–726, 5 2007.

[196] Nissim Ofek, Andrei Petrenko, Reinier Heeres, Philip Reinhold, Zaki Leghtas, Brian Vlastakis, Yehan Liu, Luigi Frunzio, S. M. Girvin, L. Jiang, Mazyar Mirrahimi, M. H. Devoret, and R. J. Schoelkopf. Extending the lifetime of a quantum bit with error correction in superconducting circuits. *Nature*, 536(7617):441–445, 8 2016.

[197] P. J J O'Malley, J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I. C. Hoi, E. Jeffrey, A. Megrant,

J. Mutus, C. Neill, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Korotkov, A. N. Cleland, and John M. Martinis. Qubit Metrology of Ultralow Phase Noise Using Randomized Benchmarking. *Physical Review Applied*, 3(4):1–11, 2015.

[198] P. J.J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis. Scalable quantum simulation of molecular energies. *Physical Review X*, 6(3):031007, 7 2016.

[199] Takafumi Ono, Ryo Okamoto, Masato Tanida, Holger F. Hofmann, and Shigeki Takeuchi. Implementation of a quantum controlled-SWAP gate with photonic circuits. *Scientific Reports*, 7(1):45353, 6 2017.

[200] Roman Orus. A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States. *Annals of Physics*, 349:117–158, 6 2013.

[201] Mateusz Ostaszewski, Edward Grant, and Marcello Benedetti. Quantum circuit structure learning. Technical report, 2019.

[202] J. Paldus, J. Cízek, and M. Takahashi. Approximate account of the connected quadruply excited clusters in the coupled-pair many-electron theory. *Physical Review A*, 30(5):2193, 11 1984.

[203] J. Paldus, J. Íek, and I. Shavitt. Correlation Problems in Atomic and Molecular Systems. IV. Extended Coupled-Pair Many-Electron Theory and Its Application to the BH3 Molecule. *Physical Review A*, 5(1):50, 1 1972.

[204] Josef Paldus and Josef E Paldus R No. Correlation problems in atomic and molecular systems. V. Spin-adapted coupled cluster many-electron theory. *The Journal of Chemical Physics*, 67(1):303, 8 2008.

[205] Robert M. Parrish, Joseph T. Iosue, Asier Ozaeta, and Peter L. McMahon. A Jacobi Diagonalization and Anderson Acceleration Algorithm For Variational Quantum Algorithm Parameter Optimization. 4 2019.

[206] Dhrumil Patel, Patrick J. Coles, and Mark M. Wilde. Variational Quantum Algorithms for Semidefinite Programming. 12 2021.

[207] Raj B. Patel, Joseph Ho, Franck Ferreyrol, Timothy C. Ralph, and Geoff J. Pryde. A quantum Fredkin gate. *Science Advances*, 2(3):e1501531, 3 2016.

[208] A.D. Patterson, J. Rahamim, T. Tsunoda, P.A. Spring, S. Jebari, K. Ratter, M. Mergenthaler, G. Tancredi, B. Vlastakis, M. Esposito, and P.J. Leek. Calibration of a Cross-Resonance Two-Qubit Gate Between Directly Coupled Transmons. *Physical Review Applied*, 12(6):064013, 12 2019.

[209] Archimedes Pavlidis and Emmanuel Floratos. Quantum-Fourier-transform-based quantum arithmetic with qudits. *Physical Review A*, 103(3):032417, 3 2021.

[210] Alejandro Perdomo-Ortiz, Marcello Benedetti, John Realpe-Gómez, and Rupak Biswas. Opportunities and challenges for quantum-assisted machine learning in near-term quantum computers. *Quantum Science and Technology*, 3(3):030502, 6 2018.

[211] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man Hong Yung, Xiao Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications 2014 5:1*, 5(1):1–7, 7 2014.

[212] Michael J. Peterer, Samuel J. Bader, Xiaoyue Jin, Fei Yan, Archana Kamal, Theodore J. Gudmundsen, Peter J. Leek, Terry P. Orlando, William D. Oliver, and Simon Gustavsson. Coherence and decay of higher energy levels of a superconducting transmon qubit. *Physical Review Letters*, 114(1):010501, 6 2015.

[213] Piotr Piecuch, Robert Toboła, and Josef Paldus. Coupled-Cluster approaches with an approximate account of triply and quadruply excited clusters: Implementation of the orthogonally spin-adapted CCD + ST(CCD), CCSD + T(CCSD), and ACPQ + ST(ACPQ) formalisms. *International Journal of Quantum Chemistry*, 55(2):133–146, 7 1995.

[214] Ekaterina Pomerantseva, Si Hyoung Oh, Gianluigi Botton, al , Hadis Zarrin, Wenmu Li, Michael Fowler, Jason Wu, and Bernard Llanos. Coupled Cluster Approach. *Physica Scripta*, 21(3-4):251, 1 1980.

[215] D. Pozar. *Microwave Engineering Fourth Edition*. 2005.

[216] Shiroman Prakash, Akalank Jain, Bhakti Kapur, and Shubangi Seth. Normal form for single-qutrit Clifford+T operators and synthesis of single-qutrit gates. *Physical Review A*, 98(3):032304, 9 2018.

[217] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2, 1 2018.

[218] Timothy Proctor, Kenneth Rudinger, Kevin Young, Mohan Sarovar, and Robin Blume-Kohout. What Randomized Benchmarking Actually Measures. *Physical Review Letters*, 119(13):130502, 9 2017.

[219] J. Rahamim, T. Behrle, M. J. Peterer, A. Patterson, P. Spring, T. Tsunoda, R. Manenti, G. Tancredi, and P. J. Leek. Double-sided coaxial circuit QED with out-of-plane wiring. *Applied Physics Letters*, 110(22):1–4, 3 2017.

[220] T. C. Ralph, K. J. Resch, and A. Gilchrist. Efficient Toffoli gates using qudits. *Physical Review A*, 75(2):022313, 2 2007.

[221] Chandrasekhar Ramanathan, Nicolas Boulant, Zhiying Chen, David G. Cory, Isaac Chuang, and Matthias Steffen. NMR quantum information processing. *Experimental Aspects of Quantum Computing*, pages 15–44, 1 2015.

[222] R. Raussendorf and H. J. Briegel. A One-Way Quantum Computer. *Physical Review Letters*, 86(22):5188, 5 2001.

[223] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel. Measurement-based quantum computation on cluster states. *Physical Review A*, 68(2):022312, 8 2003.

[224] Douglas Reynolds. Gaussian Mixture Models. *Encyclopedia of Biometrics*, pages 659–663, 2009.

[225] Susanne Richer, Nataliya Maleeva, Sebastian T. Skacel, Ioan M. Pop, and David Divincenzo. Inductively shunted transmon qubit with tunable transverse and longitudinal coupling. *Physical Review B*, 96(17):174520, 11 2017.

[226] Chad Rigetti and Michel Devoret. Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies. *Physical Review B*, 81(13):134507, 4 2010.

[227] Martin Ringbauer, Michael Meth, Lukas Postler, Roman Stricker, Rainer Blatt, Philipp Schindler, and Thomas Monz. A universal qudit quantum processor with trapped ions. *Nature Physics*, 18(9):1053–1057, 9 2022.

[228] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 2 1978.

[229] Jonathan Romero, Ryan Babbush, Jarrod R McClean, Cornelius Hempel, Peter J Love, and Alán Aspuru-Guzik. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Science and Technology*, 4(1):014008, 10 2018.

[230] Vishal Sahni and C. Patvardhan. Iris Data Classification Using Quantum Neural Networks. *AIP Conference Proceedings*, 864(1):219, 11 2006.

[231] Yuval R. Sanders, Guang Hao Low, Artur Scherer, and Dominic W. Berry. Black-Box Quantum State Preparation without Arithmetic. *Physical Review Letters*, 122(2):020502, 1 2019.

[232] Rahul Sarkar and Theodore J. Yoder. The qudit Pauli group: non-commuting pairs, non-commuting sets, and structure theorems. 2 2023.

[233] Maria Schuld, Ville Bergholm, Christian Gogolin, Josh Izaac, and Nathan Killoran. Evaluating analytic gradients on quantum hardware. *Physical Review A*, 99(3):032331, 3 2019.

[234] Maria Schuld, Alex Bocharov, Krysta M. Svore, and Nathan Wiebe. Circuit-centric quantum classifiers. *Physical Review A*, 101(3):032308, 3 2020.

[235] Maria Schuld and Nathan Killoran. Quantum Machine Learning in Feature Hilbert Spaces. *Physical Review Letters*, 122(4):040504, 2 2019.

[236] Benjamin Schumacher. Sending entanglement through noisy quantum channels. *Physical Review A*, 54(4):2614–2628, 10 1996.

[237] R. Schutjens, F. Abu Dagga, D. J. Egger, and F. K. Wilhelm. Single-qubit gates in frequency-crowded transmon systems. *Physical Review A*, 88(5):052330, 11 2013.

[238] Sarah Sheldon, Easwar Magesan, Jerry M. Chow, and Jay M. Gambetta. Procedure for systematically tuning up cross-talk in the cross-resonance gate. *Physical Review A*, 93(6):060302, 6 2016.

[239] Peter W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Journal on Computing*, 26(5):1484–1509, 8 1995.

[240] Tycho Sleator and Harald Weinfurter. Realizable Universal Quantum Logic Gates. *Physical Review Letters*, 74(20):4087–4090, 5 1995.

[241] Sergei Slussarenko and Geoff J. Pryde. Photonic quantum information processing: A concise review. *Applied Physics Reviews*, 6(4):041303, 10 2019.

[242] David W. Small and Martin Head-Gordon. A fusion of the closed-shell coupled cluster singles and doubles method and valence-bond theory for bond breaking. *The Journal of Chemical Physics*, 137(11):114103, 9 2012.

[243] David W. Small, Keith V. Lawler, and Martin Head-Gordon. Coupled cluster valence bond method: Efficient computer implementation and application to multiple bond dissociations and strong correlations in the acenes. *Journal of Chemical Theory and Computation*, 10(5):2027–2040, 5 2014.

[244] Robert I. Soare. Computability and Recursion. *Bulletin of Symbolic Logic*, 2(3):284–321, 9 1996.

[245] Robert I. Soare. Turing oracle machines, online computing, and three displacements in computability theory. *Annals of Pure and Applied Logic*, 160(3):368–399, 9 2009.

[246] J. Spall. An overview of the simultaneous perturbation method for efficient optimization. *Johns Hopkins apl technical digest*, 19(4):482–492, 1998.

[247] Peter A. Spring, Shuxiang Cao, Takahiro Tsunoda, Giulio Campanaro, Simone Fasciati, James Wills, Mustafa Bakr, Vivek Chidambaram, Boris Shteynas, Lewis Carpenter, Paul Gow, James Gates, Brian Vlastakis, and Peter J. Leek. High coherence and low cross-talk in a tileable 3D integrated superconducting circuit architecture. *Science Advances*, 8(16):6698, 4 2022.

[248] Leandro Stefanazzi, Kenneth Treptow, Neal Wilcer, Chris Stoughton, Collin Bradford, Sho Uemura, Silvia Zorzetti, Salvatore Montella, Gustavo Cancelo, Sara Sussman, Andrew Houck, Shefali Saxena, Horacio Arnaldi, Ankur Agrawal, Helin Zhang, Chunyang Ding, and David I. Schuster. The QICK (Quantum Instrumentation Control Kit): Readout and control for qubits and detectors. *Review of Scientific Instruments*, 93(4):044709, 4 2022.

[249] E. Miles Stoudenmire and David J. Schwab. Supervised Learning with Quantum-Inspired Tensor Networks. *Advances in Neural Information Processing Systems*, 29:4799, 5 2016.

[250] A. Szabó and N. Ostlund. *Modern quantum chemistry : introduction to advanced electronic structure theory*. Dover Publications, illustrated, reprint edition, 1982.

[251] Shuhei Tamate, Yutaka Tabuchi, and Yasunobu Nakamura. Toward Realization of Scalable Packaging and Wiring for Large-Scale Superconducting Quantum Computers. *IEICE Transactions on Electronics*, E105.C(6):2021SEP0007, 6 2022.

[252] Xinsheng Tan, Dan Wei Zhang, Qiang Liu, Guangming Xue, Hai Feng Yu, Yan Qing Zhu, Hui Yan, Shi Liang Zhu, and Yang Yu. Topological Maxwell Metal Bands in a Superconducting Qutrit. *Physical Review Letters*, 120(13):130503, 3 2018.

[253] Xinsheng Tan, Dan Wei Zhang, Wen Zheng, Xiaopei Yang, Shuqing Song, Zhikun Han, Yuqian Dong, Zhimin Wang, Dong Lan, Hui Yan, Shi Liang Zhu, and Yang Yu. Experimental Observation of Tensor Monopoles with a Superconducting Qudit. *Physical Review Letters*, 126(1):017702, 1 2021.

[254] Andrew G. Taube and Rodney J. Bartlett. New perspectives on unitary coupled-cluster theory. *International Journal of Quantum Chemistry*, 106(15):3393–3401, 1 2006.

[255] Daniel M. Tennant, Luis A. Martinez, Kristin M. Beck, Sean R. O'Kelley, Christopher D. Wilen, R. McDermott, Jonathan L. Dubois, and Yaniv J. Rosen. Low-Frequency Correlated Charge-Noise Measurements Across Multiple Energy Transitions in a Tantalum Transmon. *PRX Quantum*, 3(3):030307, 7 2022.

[256] Dirk Oliver Theis. Optimality of Finite-Support Parameter Shift Rules for Derivatives of Variational Quantum Circuits. 12 2021.

[257] L. S. Theis, F. Motzoi, and F. K. Wilhelm. Simultaneous gates in frequency-crowded multilevel systems using fast, robust, analytic control shapes. *Physical Review A*, 93(1):012324, 1 2016.

[258] R. T. Thew, K. Nemoto, A. G. White, and W. J. Munro. Qudit quantum-state tomography. *Physical Review A*, 66(1):012303, 7 2002.

[259] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth, and Jonathan Tennyson. The Variational Quantum Eigensolver: a review of methods and best practices. 11 2021.

[260] A. M. Turing. *Systems of logic based on ordinals*. PhD thesis, 1939.

[261] John van de Wetering. ZX-calculus for the working quantum computer scientist. 12 2020.

[262] Lieven Vandenberghe and Stephen Boyd. Semidefinite Programming. *http://dx.doi.org/10.1137/1038003*, 38(1):49–95, 7 2006.

[263] L. M.K. Vandersypen and I. L. Chuang. NMR techniques for quantum control and computation. *Reviews of Modern Physics*, 76(4):1037–1069, 10 2004.

[264] Vladimir Vargas-Calderón, Nicolas Parra-A., Herbert Vinck-Posada, and Fabio A. González. Many-Qudit Representation for the Travelling Salesman Problem Optimisation. *https://doi.org/10.7566/JPSJ.90.114002*, 90(11), 10 2021.

[265] Farrokh Vatan and Colin Williams. Optimal quantum circuits for general two-qubit gates. *Physical Review A*, 69(3):032315, 3 2004.

[266] Javier Gil Vidal and Dirk Oliver Theis. Calculus on parameterized quantum circuits. 12 2018.

[267] Lorenza Viola and Seth Lloyd. Dynamical suppression of decoherence in two-state quantum systems. *Physical Review A*, 58(4):2733, 10 1998.

[268] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret. Manipulating the Quantum State of an Electrical Circuit. *Science*, 296(5569):886–889, 5 2002.

[269] Joel J. Wallman. Randomized benchmarking with gate-dependent noise. *Quantum*, 2:47, 1 2018.

[270] Joel J Wallman, Marie Barnhill, and Joseph Emerson. Robust characterization of leakage errors. *New Journal of Physics*, 18(4):043021, 4 2016.

[271] M Wallquist, K Hammerer, P Rabl, S M Girvin, M H Devoret, and R J Schoelkopf. Circuit QED and engineering charge-based superconducting qubits. *Physica Scripta*, 2009(T137):014012, 12 2009.

[272] M. Wallquist, V. S. Shumeiko, and G. Wendin. Selective coupling of superconducting charge qubits mediated by a tunable stripline cavity. *Physical Review B - Condensed Matter and Materials Physics*, 74(22):224506, 12 2006.

[273] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R.-S. Huang, J. Majer, S. Kumar, S. M. Girvin, and R. J. Schoelkopf. Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature*, 431(7005):162–167, 9 2004.

[274] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Approaching unit visibility for control of a superconducting qubit with dispersive readout. *Physical Review Letters*, 95(6):060501, 8 2005.

[275] P. Walther, K. J. Resch, T. Rudolph, E. Schenck, H. Weinfurter, V. Vedral, M. Aspelmeyer, and A. Zeilinger. Experimental one-way quantum computing. *Nature 2005 434:7030*, 434(7030):169–176, 3 2005.

[276] C. Wang, C. Axline, Y. Y. Gao, T. Brecht, Y. Chu, L. Frunzio, M. H. Devoret, and R. J. Schoelkopf. Surface participation and dielectric loss in superconducting qubits. *Applied Physics Letters*, 107(16):162601, 10 2015.

[277] Daochen Wang, Oscar Higgott, and Stephen Brierley. Accelerated Variational Quantum Eigensolver. *Physical Review Letters*, 122(14):140504, 4 2019.

[278] Jianwei Wang, Stefano Paesani, Yunhong Ding, Raffaele Santagati, Paul Skrzypczyk, Alexia Salavrakos, Jordi Tura, Remigiusz Augusiak, Laura Mančinska, Davide Bacco, Damien Bonneau, Joshua W. Silverstone, Qihuang Gong,

Antonio Acín, Karsten Rottwitt, Leif K. Oxenløwe, Jeremy L. O'Brien, Anthony Laing, and Mark G. Thompson. Multidimensional quantum entanglement with large-scale integrated optics. *Science*, 360(6386):285–291, 4 2018.

[279] Quanlong Wang. Qutrit ZX-calculus is Complete for Stabilizer Quantum Mechanics. *Electronic Proceedings in Theoretical Computer Science, EPTCS*, 266:58–70, 3 2018.

[280] Yuchen Wang, Zixuan Hu, Barry C. Sanders, and Sabre Kais. Qudits and High-Dimensional Quantum Computing. *Frontiers in Physics*, 8:479, 11 2020.

[281] Fern H.E. Watson, Hussain Anwar, and Dan E. Browne. Fast fault-tolerant decoder for qubit and qudit surface codes. *Physical Review A - Atomic, Molecular, and Optical Physics*, 92(3):032309, 9 2015.

[282] G. Wendin. Quantum information processing with superconducting circuits: a review. *Reports on Progress in Physics*, 80(10):106001, 9 2017.

[283] David Wierichs, Josh Izaac, Cody Wang, and Cedric Yen Yu Lin. General parameter-shift rules for quantum gradients. *Quantum*, 6:677, 3 2022.

[284] Colin P. Williams. *Explorations in Quantum Computing*. Texts in Computer Science. Springer London, London, 2011.

[285] J. Wills, G. Campanaro, S. Cao, S. D. Fasciati, P. J. Leek, and B. Vlastakis. Spatial Charge Sensitivity in a Multimode Superconducting Qubit. *Physical Review Applied*, 17(2):024058, 2 2022.

[286] Christopher J. Wood, Jacob D. Biamonte, and David G. Cory. Tensor networks and graphical calculus for open quantum systems. 11 2011.

[287] Christopher J. Wood and Jay M. Gambetta. Quantification and characterization of leakage errors. *Physical Review A*, 97(3):032306, 3 2018.

[288] Xian Wu, S L Tomarken, N Anders Petersson, L A Martinez, Yaniv J Rosen, and Jonathan L Dubois. High-Fidelity Software-Defined Quantum Logic on a Superconducting Qudit. *Physical Review Letters*, 125, 2020.

[289] Xiao Xue, Maximilian Russ, Nodar Samkharadze, Brennan Undseth, Amir Sammak, Giordano Scappucci, and Lieven M.K. Vandersypen. Quantum logic with spin qubits crossing the surface code threshold. *Nature 2022 601:7893*, 601(7893):343–347, 1 2022.

[290] T. Yamamoto, Yu. A. Pashkin, O. Astafiev, Y. Nakamura, and J. S. Tsai. Demonstration of conditional gate operation using superconducting charge qubits. *Nature*, 425(6961):941–944, 10 2003.

[291] Harry Yserentant. A Short Theory of the Rayleigh–Ritz Method. *Computational Methods in Applied Mathematics*, 13(4):495–502, 10 2013.

[292] Wen Zheng, Yu Zhang, Yuqian Dong, Jianwen Xu, Zhimin Wang, Xiaohan Wang, Yong Li, Dong Lan, Jie Zhao, Shaoxiong Li, Xinsheng Tan, and Yang Yu. Optimal control of stimulated Raman adiabatic passage in a superconducting qudit. *npj Quantum Information 2022 8:1*, 8(1):1–5, 1 2022.