

Viable Error-Mitigation Techniques for Noisy Intermediate-Scale Quantum Devices



Zohim Chandani

Department of Physics, The Cavendish Laboratory
University of Cambridge

This dissertation is submitted for the degree of
Master Of Advanced Study

May 2020

This thesis is dedicated to my parents whose sole purpose since my birth has been to provide me with the opportunities that they never had.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements.

COVID 19: Impact due to the pandemic has been minimal on my project due to its computational nature. However it has slowed down the progress of tasks that would have otherwise taken minimal effort to sort out. For example, learning how to use the high-performance computing cluster took longer than expected due to limited support available by the team. Moreover, getting in touch with members of the community at IBM and Rigetti proved to be a challenge as their efforts were redirected towards serving their main clients. This definitely hindered the acquisition of results from the quantum computer which otherwise would have been possible.

Zohim Chandani

May 2020

Acknowledgements

I would like to thank my project supervisor, Dr David Arvidsson-Shukur, for giving me the opportunity to delve deeper into this area of knowledge. I was initially interested in pursuing a project in particle theory but this opportunity has sparked my interest in quantum computation and information which I will be carrying forward. I have learnt a lot through various discussions and your overarching knowledge of the field has guided my thinking every step of the way.

Yordan Yordanov, a PhD student at the Cavendish, who initially helped me get started with the technicalities and was my first point of reference with any problems that I encountered throughout the duration of this project. Your knowledge of cloud quantum computing and recent advancements in the field was an invaluable asset and I highly appreciate the guidance you provided every step of the way.

A collective appreciation for the quantum technologies group led by Professor Simon Benjamin at The University of Oxford. The numerous discussions I had with Suguru Endo, Sam McArdle, and Armands Strikis allowed me to gain a clearer understanding of the techniques employed by other researchers and critique my own approaches. Our discussions also sparked an interest in exploring areas of the field out of curiosity that are not even mentioned in this thesis.

Alessandro Roussel, the man behind the YouTube channel ScienceClic. The invaluable support you provided me with to get started with Python will never be forgotten. You were my first point of contact with any issues relating to programming and I enjoyed watching you program combinatorial optimisation problems we first looked at.

And lastly, to the various members of the IBM team who I got to know via Slack over the course of this journey. The work you are doing in providing access to quantum computing to the layman is inspiring. I hope to contribute in a similar manner one day.

Abstract

We are currently rooted in the Noisy Intermediate-Scale Quantum (NISQ) technologies era and the community hopes to scale up considerably over the next 15-20 years. This requires further improvements in hardware which experimentalists have made huge strides in so far. At the same time, efforts are being poured into utilising NISQ devices to invent algorithms and generate results which will form a bedrock for future work when quantum computation is a feasible technique.

An area which shows great promise is quantum simulations of molecules which can be used to accelerate drug discovery. These algorithms rely on accurate estimates of the computational output which is currently plagued by decoherence.

I investigate two quantum error mitigation schemes: probabilistic error cancellation (PEC) and extrapolation, which aim to mitigate the effects of decoherence and require no additional qubit resources. Work by the original authors, Temme et al. [Phys. Rev. Lett. 119, 180509 (2017)] is extended to incorporate varying error rates for single and two-qubit gates and investigate the performance of the mitigation techniques under varying levels of algorithmic size as we incrementally move towards circuits with larger depths. Finally, the PEC method is experimentally implemented for the first time and its shortcomings in mitigating results is identified.

Table of contents

1	Global perspective	1
2	Fundamental concepts in quantum mechanics & quantum computing	3
2.1	Quantum bits	3
2.2	Time evolution	5
2.3	Measurement	5
2.4	Composite quantum systems	6
2.5	Density operator formalism	6
2.6	Quantum circuits	8
2.6.1	Single-qubit gate operations	8
2.6.2	Multi-qubit gates	9
2.6.3	Measurement	9
2.6.4	Combining qubits, gates & measurements	10
2.6.5	Circuit depth	11
3	Errors & mitigating errors	12
3.1	Framework to conceptualise noise	12
3.2	Quantum noise	13
3.2.1	Depolarising error channel	14
3.3	Error correction	15
3.4	Error mitigation	15
3.4.1	Extrapolation	15
3.4.2	Probabilistic error cancellation	17
4	Method & results	20
4.1	Technicalities	20
4.1.1	Random circuit generation	20
4.1.2	Implementing and inverting noise	20

Table of contents	vii
4.2 Numerical results	21
4.3 Experimental applications of probabilistic error cancellation	26
5 Discussion of results	29
6 Departing thoughts	31
References	33
Appendix A Quantum error correction	38
A.1 Three-qubit code	38
Appendix B Qiskit compiler	41
Appendix C Functionality of PEC and extrapolation as a function of circuit depth	42

Chapter 1

Global perspective

We are at the cusp of a new frontier. We might call this the *entanglement* frontier where for the first time we are able to precisely control entangled quantum states of many particles which cannot be simulated by our best classical computers [1].

The assumption that the entanglement frontier will yield fruitful results rests on two principles: 1) our preconception that quantum computing (QC) is powerful, and 2) quantum computers can be scaled to large devices.

We know QC is powerful because problems such as finding prime factors of large integers is exponentially costly in time for classical computers. Quantum algorithms such as one developed by Peter Shor, demonstrate results which are produced efficiently in polynomial time [2]. Other examples include Grover's [3] and Deutsch–Jozsa algorithms [4].

We must realise that the bounds of QC is not unlimited. Problems which are classically NP-hard are not expected to be solved by the advent of QC. Examples of problems which quantum computers could solve, and one which has generated huge interest in recent years, lie within the field of quantum chemistry [5, 6].

Feynman had envisioned this when he said: "Nature isn't classical dammit, and if you want to make a simulation of nature, you better make is quantum mechanical" [7]. 35 years after this proposal, we are beginning to reach the stage where QC can be used to provide solutions to equations describing many entangled particles.

Advancements in nanoscience has led to the advent of the current noisy intermediate-scale quantum devices (NISQ) era. NISQ summarises our current achievements in the field of QC. "Intermediate scale" refers to the size of quantum computers which are currently available, ranging from 50 to around 70 qubits [8]. "Noisy" emphasises that we have imperfect control over these qubits which imposes limitations on what quantum devices can achieve in the near future [9].

We still have a long way to go before we can make QC viable. Transitioning from the NISQ era is crucial if we are to successfully implement Shor's algorithm which is estimated to require several million qubits [10]. Progress is hindered by our experimental limitations: the error rates are relatively large compared to error occurrences in classical information processing [11]. Methods to suppress errors using encoding information have been researched for decades [12, 13] but require extra qubits which we do not have the freedom to employ in the current NISQ era. Techniques such as *error mitigation* have been developed, which is the focus of this report, to allow us to make practical use of NISQ era technology and not wholly fixate on large scale quantum devices [14].

Chapter 2

Fundamental concepts in quantum mechanics & quantum computing

We will embark upon a short voyage exploring the fundamental concepts of quantum theory and discuss its applications to quantum computing. This will serve as a bedrock for discussing material presented in Chapter 3 and beyond. It is fair to say that this summarises our current understanding of quantum mechanics, briefly, in under 5 pages, but is actually a culmination of an arduous journey spanning over 5 decades [15]. The word 'culmination' does not sit right as physics is a constant pursuit of validating theories with experiments and experiments with theories.

2.1 Quantum bits

Classical computing has information storage and manipulation rooted in the fundamental concept of a *bit* which can take values of 0 or 1. Quantum computing is built upon an analogous concept where the fundamental unit of information storage is the *quantum bit* or the *qubit*.

The arena in which the mathematics for qubits is performed is described by linear algebra and we say that the state of a qubit is a unit vector in a two-dimensional complex vector space with a well defined inner product which is also known as a Hilbert space, \mathcal{H} (note that Hilbert spaces are generally d-dimensional but for the case of qubits, d=2). The general state of a qubit is a linear combination of the possible basis states

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (2.1)$$

where $\alpha, \beta \in \mathbb{C}$. Measurement of a qubit yields the state $|0\rangle$ and $|1\rangle$ with probability $|\alpha|^2$ and $|\beta|^2$ respectively. This naturally extends to $|\alpha|^2 + |\beta|^2 = 1$ as a statement of conservation of probabilities requiring $\langle\psi|\psi\rangle = 1$ which is known as the *normalisation* condition for statevector's.

Physical realisations of two-level quantum systems take on several experimental forms. A popular choice employed by IBM, Rigetti and Google is the superconducting quantum circuit with an ingrained Josephson junction [16, 17]. Other popular choices are the optical photon quantum computer [18], trapped ions [19] and quantum dots [20]. We shall not delve into the experimental implementations of quantum computers but the interested reader is encouraged to follow up with the references provided. For our purposes, it will suffice to think of qubits in abstract terms as mathematical entities. For a geometric representation of $|\psi\rangle$, we can rewrite Equation 2.1 as

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle \quad (2.2)$$

where $0 \leq \phi < 2\pi$, and $0 \leq \theta \leq \pi$. From this, it is clear that there is a one to one mapping between qubit states encoded in \mathbb{C}^2 and the points on a unit sphere which lie on \mathbb{R}^3 . Figure 2.1 is a geometric visual of this mapping which is known as the Bloch sphere representation.

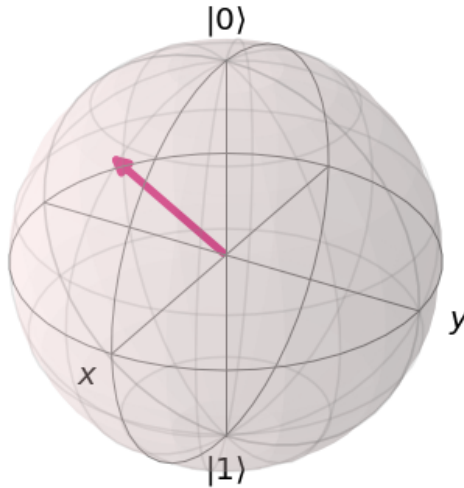


Fig. 2.1 The state of a qubit is denoted by a unit Bloch vector which lies on the surface of the Bloch sphere. $|0\rangle$ and $|1\rangle$ are defined to be along the z axis and are conventionally chosen to be the computational basis states. θ is the polar angle measured from the positive z axis. ϕ is the azimuthal angle in the x - y plane from the positive x axis.

2.2 Time evolution

Time evolution of the statevector $|\psi\rangle$, is governed by the *Schrödinger equation*,

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = H |\psi\rangle \quad (2.3)$$

where H is the Hamiltonian governing the system's dynamics. The Hamiltonian need not be time independent but we will skip discussion of perturbation theory which is the solution to time-dependent Hamiltonian's [21]. The solution to equation 2.3 is given by

$$|\psi(t_2)\rangle = U(t_1, t_2) |\psi(t_1)\rangle \quad (2.4)$$

for times t_1 and t_2 where U is a unitary operator defined as

$$U(t_1, t_2) = \exp\left(\frac{-iH(t_2 - t_1)}{\hbar}\right) \quad (2.5)$$

We can thus re-frame the evolution of the state $|\psi\rangle$ at time t_1 to $|\psi'\rangle$ at time t_2 , for $t_2 > t_1$, as a unitary transformation expressed as

$$|\psi'\rangle = U |\psi\rangle. \quad (2.6)$$

Examples of unitary operations, represented as matrices, which act on qubit states is given in section 2.6. Time evolution governed by the *Schrödinger equation* requires closed quantum systems. In principle, this means that the evolution of the qubit state must solely be governed by the interacting Hamiltonian and no other external factors. In reality, all experimental systems interact, to some degree, with the environment. This unwanted environmental influence will be explored in Chapter 3.

2.3 Measurement

Quantum-state preparation, information encoding and manipulation is an incomplete story. We need a way to measure the state of the quantum system after its evolution.

The result is determined by a set of measurement operators, $\{M_m\}$, where m refers to the possible outcomes that a measurement may yield. A measurement operator acts on the Hilbert space of the system to yield

$$|\psi\rangle \xrightarrow[\text{result}]{\text{measurement}} \frac{M_m |\psi\rangle}{\sqrt{p(m)}} \quad (2.7)$$

where, $p(m) = \langle \psi | M_m^\dagger M_m | \psi \rangle$, is the probability that the result m occurs.

For our qubit system given by Equation 2.1, measurement in the computational basis is defined by the two rank-1 projection operators: $M_0 = |0\rangle\langle 0|$ and $M_1 = |1\rangle\langle 1|$ where the probability that a measurement yields the state $|0\rangle$ is given by

$$p(0) = \langle \psi | M_0^\dagger M_0 | \psi \rangle = |\alpha|^2 \quad (2.8)$$

and similarly, $p(1) = \langle \psi | M_1^\dagger M_1 | \psi \rangle = |\beta|^2$.

2.4 Composite quantum systems

Extension of single qubit systems to multiple quantum bits is essential if we are to make quantum-information processing a feasible technique. Suppose quantum systems A and B have individual quantum states $|\psi_A\rangle$ and $|\psi_B\rangle$ encoded in Hilbert spaces \mathcal{H}_A and \mathcal{H}_B . We denote the state of the composite system AB as

$$|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle \quad (2.9)$$

where $|\psi_{AB}\rangle \in \mathcal{H}$ and $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. For brevity, we omit the \otimes and write the state as $|\psi_{AB}\rangle = |\psi_A\rangle |\psi_B\rangle = |\psi_A \psi_B\rangle$. This mathematical structure can be naturally extended and generalised to systems comprising of n qubits.

Entanglement is property fundamental to quantum mechanics which is unveiled to us from our definition of combining quantum systems [22]. Consider the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \quad (2.10)$$

where $|00\rangle$ represents the state of systems A and B . It is clear that Equation 2.10 cannot be expressed as a tensor product of states of individual subsystems. Quantum systems with this property are said to be *entangled* and entanglement persists even for spacelike separated events. This phenomenon plays a crucial role in quantum teleportation [23], superdense coding [24] and tests for Bells' inequality [25].

2.5 Density operator formalism

We have described the representation of our quantum system, time evolution and measurement using the statevector approach. An alternate formulation of the theory is the density operator formalism, which is mathematically equivalent, but provides more freedom to discuss systems

interacting with the environment. Below we state some key results of the density operator approach to highlight its usefulness.

A quantum system in state $|\psi\rangle$ is a *pure state* and its density operator, ρ , is defined as $\rho \equiv |\psi\rangle\langle\psi|$. Suppose we do not know the state of the system and can only assign classical probabilities to what it might be. More precisely, the quantum system is in the state $|\psi_i\rangle$ with probability p_i . The density operator in this case is defined as

$$\rho \equiv \sum_i p_i |\psi_i\rangle\langle\psi_i| \quad (2.11)$$

where ρ is a *mixed state*, a mixture of different pure states. The evolution of the statevector, as expressed in Equation 2.6, is given by $|\psi'\rangle = U|\psi\rangle$ where U is a unitary operator controlling the dynamics. Similarly, if the system is initially in the state $|\psi_i\rangle$ with probability p_i then the evolved system will be in the state given by $U|\psi_i\rangle$ with probability p_i . This is straightforwardly incorporated in Equation 2.11 as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \longrightarrow \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^\dagger = U\rho U^\dagger \quad (2.12)$$

which can be simplified to

$$\rho' = U\rho U^\dagger \quad (2.13)$$

where ρ' is the evolved density operator.

We can also formulate the density operator for a composite quantum system. Taking inspiration from our previous mathematical structure developed to combine statevector's, as expressed in Equation 2.9, we arrive at

$$\rho^{AB} = \rho^A \otimes \rho^B \quad (2.14)$$

where ρ^A and ρ^B are the density operators for individual subsystems A and B and ρ^{AB} is the density operator for the composite system. This yields the same results as the statevector approach and is thus correct (if we take the statevector approach as a definition).

Perhaps the most useful tool in the density operator approach is the partial trace which allows us to construct the *reduced density operator*. Suppose that our quantum system, A , inevitably gets entangled with the environment E via some potentially weak unwanted interaction. In experimental situations like this, we only have information about the evolution of subsystem A and have no knowledge of how the states of the environment evolve. In this case, one can trace over the environment to find the *reduced density operator* defined as

$$\rho^A = \text{tr}_E(\rho^{AE}) \quad (2.15)$$

where ρ^{AE} is the density operator for the composite system and tr_E is an operation known as the *partial trace* over subsystem E to recover the density operator, ρ^A , for our subsystem A . This allows theoreticians to mathematically manipulate subsystem A but provides no benefit for experimentalists who deal with the bane of experimental quantum mechanics: *decoherence*.

We omit further discussion of density operators as the main idea of this section is to highlight the usefulness of the formalism in tackling problems related to subsystems of a larger quantum entity. For the mathematically minded reader, further information can be found in a clear and concise approach to the subject by Vlatko Vedral [26]. Alternatively, see any standard quantum mechanics textbook.

2.6 Quantum circuits

We have discussed, in abstract mathematical terms, the concept of a statevector, time evolution and measurement. We shall now combine our efforts to construct quantum circuits.

2.6.1 Single-qubit gate operations

In section 2.2, we described how we can evolve statevector's through time using unitary operations. For purposes of quantum computation, these operations come in the form of quantum gates, application of which allows for manipulation of information encoded within qubits. This evolution of the statevector is represented by

$$|\psi'\rangle = U |\psi\rangle \quad (2.16)$$

where U is a unitary matrix representing the gate operation. The most general form of the unitary is

$$U(\theta, \phi, \lambda) = \begin{pmatrix} \cos \frac{\theta}{2} & -e^{i\lambda} \sin \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} & e^{i(\phi+\lambda)} \cos \frac{\theta}{2} \end{pmatrix} \quad (2.17)$$

for all $0 \leq \phi < 2\pi$, $0 \leq \theta \leq \pi$ and $0 \leq \lambda < 2\pi$. Different values of ϕ, θ and λ encompass all possible rotations of the Bloch vector around the Bloch sphere whilst ensuring that the norm of the vector is preserved. We shall later discuss *decoherence* which can be visualised as a contraction of the Bloch sphere representing information loss encoded within the qubit.





Gates	Circuit Symbol	Matrix Representation	Effect
Pauli-X		$U(\pi, 0, \pi) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	<ul style="list-style-type: none"> • $X 0\rangle \longrightarrow 1\rangle$ • $X 1\rangle \longrightarrow 0\rangle$
Pauli-Y		$U(\pi, \frac{\pi}{2}, \frac{\pi}{2}) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$	<ul style="list-style-type: none"> • $Y 0\rangle \longrightarrow i 1\rangle$ • $Y 1\rangle \longrightarrow -i 0\rangle$,
Pauli-Z		$U(0, 0, \pi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	<ul style="list-style-type: none"> • $Z 0\rangle \longrightarrow 0\rangle$ • $Z 1\rangle \longrightarrow - 1\rangle$
Hadamard		$U(\frac{\pi}{2}, 0, \pi) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$	<ul style="list-style-type: none"> • $H 0\rangle \longrightarrow \frac{1}{\sqrt{2}}(0\rangle + 1\rangle)$ • $H 1\rangle \longrightarrow \frac{1}{\sqrt{2}}(0\rangle - 1\rangle)$

Fig. 2.2 Names, circuit symbols and unitary matrix representation for common single qubit gates. Other examples of single qubit gates include the S , T and other combinations of θ , ϕ and λ to transform qubit states around the Bloch sphere.

Common single qubit gates derived from the general unitary U are the Pauli gates: X , Y & Z and the Hadamard denoted by H . Their circuit representation and effects on the computational qubit states $|0\rangle$ and $|1\rangle$ are shown in Figure 2.2.

2.6.2 Multi-qubit gates

We can extend ideas of single qubit gates to multiple qubits and a key example is the *controlled-not* or the CNOT gate. As shown in Figure 2.3, it has two inputs known as the control and target qubit. Its effect is to flip the state of the target if the control is in the $|1\rangle$ state.

2.6.3 Measurement

Finally, we introduce the measurement operation as shown in Figure 2.4. The operation converts a qubit state given by $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ into a probabilistic classical bit. This is distinguished by drawing the flow of quantum-information as a single line and classical

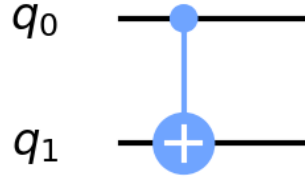


Fig. 2.3 Circuit representation of the CNOT gate operation. q_0 and q_1 are the control and target qubits respectively. The target is flipped if the control is in the $|1\rangle$ state.

information as a double line. Its effect can be thought of as projecting the quantum state onto the standard computational basis to yield $|0\rangle$ with probability $|\alpha|^2$ and $|1\rangle$ with probability $|\beta|^2$.

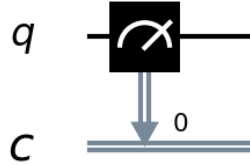


Fig. 2.4 Circuit representation for the projective measurement operation onto the Z basis. Note that q is the quantum bit and c is the classical bit.

2.6.4 Combining qubits, gates & measurements

A combination of quantum states and gates can be used to construct quantum circuits which are used to generate outputs to quantum algorithms. Many algorithms [27, 28] require the generation of a GHZ state which is defined as

$$|GHZ\rangle = \frac{|0\rangle^{\otimes n} + |1\rangle^{\otimes n}}{\sqrt{2}} \quad (2.18)$$

for $n \geq 1$ where n is the number of qubits. GHZ states are named after Greenberger, Horne, and Zeilinger, who were the first to study them in 1997 [29]. Figure 2.5 shows the circuit used to construct a 3-qubit GHZ state. With this circuit topology, one can construct n -qubit GHZ states by the application of $n-1$ CNOT's.

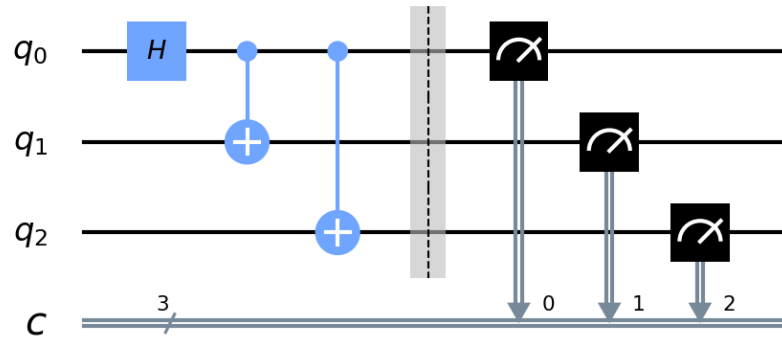


Fig. 2.5 The system is initialised in the $|000\rangle$ state. The Hadamard and the CNOT gate operations yield the entangled $(|000\rangle + |111\rangle)/\sqrt{2}$ GHZ state. The passage of time runs from left to right and the probability of measurement yielding $|000\rangle$ or $|111\rangle$ is a half.

2.6.5 Circuit depth

The complexity of a problem can be stated by a measure of the computational resources required to solve it. It is thus natural to use the number of gates or the circuit size to characterise the required resources. Alternatively, we might be interested in how much time is required to do the computation if several gates can be executed in parallel. Here we introduce the notion of *circuit depth* (cd) which is a measure of the number of time steps required assuming that gate operations on distinct qubits can operate simultaneously [30]. Figure 2.5 above is an example of a circuit with $cd = 4$. Note that the measurement operations, although shown in distinct layers, operate simultaneously.

Chapter 3

Errors & mitigating errors

We have mentioned the concept of *decoherence* in previous chapters without a firm footing for the idea to stand upon. The time has now come to explore this further.

We idealise quantum systems as closed entities which do not have interactions with the environment. In reality, ideal systems do not exist and any experimental setup has unwanted interactions which manifests itself as noise in information processing. Section 2.5 explored the usefulness of the partial trace operation which allowed theorists to reconstruct the *reduced density operator* in the quest for theoretical results. This is of no aid to the experimentalists. We simply cannot insert a tr_E whilst the qubit is interacting with the Josephson junction.

However, the community has made significant progress in its efforts to achieve current single and two-qubit gate error rates of 10^{-4} and 10^{-2} respectively [31]. At this point, it feels appropriate to acknowledge the immense experimental efforts that have been undertaken by countless researchers to achieve the so-called *quantum supremacy* (demonstrating that a quantum device can solve a problem no classical computer can feasibly solve) result by the team at Google [32], but more so by the whole QC community who have made incremental yet important advancements for the field to reach this milestone [33].

3.1 Framework to conceptualise noise

Consider a quantum circuit with n qubits initialised in the $|\bar{0}\rangle$ state where

$$|\bar{0}\rangle = |0\rangle^{\otimes n} = |000\dots 0\rangle. \quad (3.1)$$

In density operator formalism, the initial state is given by

$$\rho_0 = |\bar{0}\rangle\langle\bar{0}|. \quad (3.2)$$

N unitary gates applied to the initial state, ρ_0 , is represented as

$$\rho_0 = \mathcal{U}_N \dots \mathcal{U}_2 \mathcal{U}_1 (|\bar{0}\rangle \langle \bar{0}|) \quad (3.3)$$

where \mathcal{U} is the superoperator defined as $\mathcal{U}(\rho) = U\rho U^\dagger$. Information of interest is extracted from the circuit using a hermitian operator such that the ideal expectation value, $\langle O_0 \rangle$, of an observable is given by

$$\langle O_0 \rangle = \text{Tr}[\rho_0 O] \quad (3.4)$$

where O is the operator corresponding to the concerned observable. In reality, each gate operation, \mathcal{U}_i , is affected by a noise channel \mathcal{N}_i and hence the final state is

$$\rho = \prod_i \mathcal{N}_i \mathcal{U}_i (|\bar{0}\rangle \langle \bar{0}|). \quad (3.5)$$

From this we can extract our noisy observable

$$\langle O \rangle = \text{Tr}[\rho O]. \quad (3.6)$$

The noise channels, which are introduced in section 3.2.1, are induced due to interactions with the environment which is anything that has the possibility to interfere with the operations we perform on our system. These unwanted interactions reduce the magnitude of the off-diagonal terms in the density operator and can be referred to as *decoherence*.

Below, we shall mention techniques such as *error correction* which allows us to recover the noiseless state ρ_0 , from the noisy state ρ and discuss *error mitigation* which estimates noiseless results, $\langle O_0 \rangle$, from noisy ones, $\langle O \rangle$.

3.2 Quantum noise

Studying quantum noise starts from finding models of how interactions with the environment corrupt the information being processed. Once we have suitable noise models, we can try to mitigate their effects on observables either by improvements in hardware or post measurement during information processing.

There are many examples of noisy processes which researchers have tried to grasp the effects of [34]. Some of them are, bit and phase flip errors [35], amplitude damping [36] and thermal relaxation [37]. Below, we will focus on the depolarising error channel, its effect on quantum operations and how we can try to mitigate its unwelcomed intervention.

3.2.1 Depolarising error channel

The depolarising error channel, which is the focus of our investigation, is described by

$$\mathcal{D}(\rho) = \frac{p}{2}I + (1-p)\rho \quad (3.7)$$

which *depolarises* a single qubit with probability p . This is to say that the qubit is replaced by a completely mixed state, $I/2$, and left untouched with probability $1-p$.

Note that for arbitrary ρ , the mixed state is given by

$$\frac{I}{2} = \frac{\rho + X\rho X + Y\rho Y + Z\rho Z}{4} \quad (3.8)$$

and substituting this into Equation 3.7 yields

$$\mathcal{D}(\rho) = (1 - \frac{3p}{4})\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z). \quad (3.9)$$

We may choose to parametrise the error channel in such a way that the state ρ is left untouched with probability $1-p$ and the Pauli operations X, Y or Z are applied with probability $p/3$. In this case, one can represent Equation 3.9 as

$$\mathcal{D}(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z). \quad (3.10)$$

The effect of the depolarising channel on the Bloch sphere is illustrated in Figure 3.1.

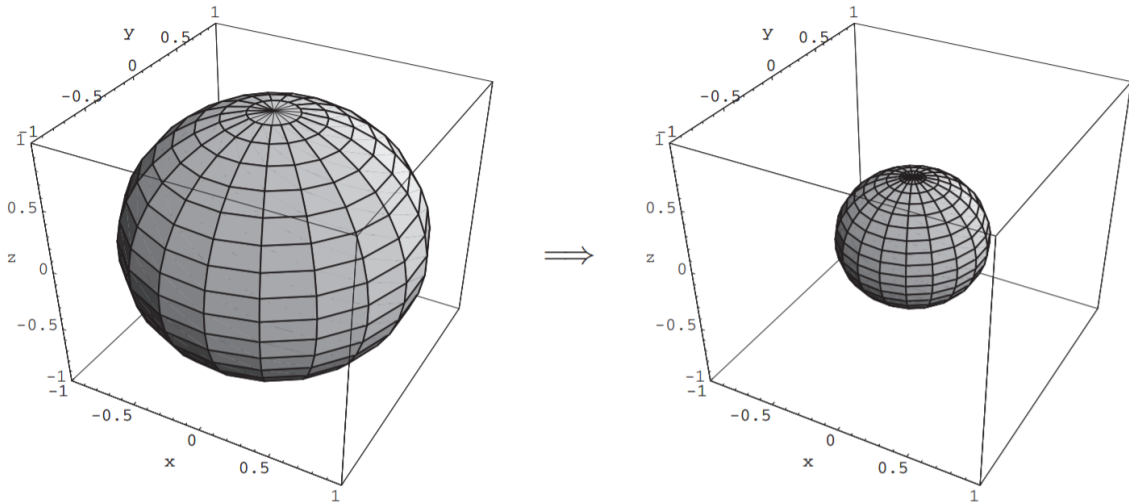


Fig. 3.1 Effect of the depolarising error channel on the Bloch sphere for $p = 0.5$. Note how the entire sphere contracts uniformly as a function of the error probability. This represents the loss of information encoded in the possible states the Bloch vector can occupy. Figure reproduced from [38].

3.3 Error correction

Error-correction codes guard quantum information against potential corruption due to noise by encoding information in extra qubits known as *ancilla* qubits. The state of ancilla qubits gives us information about the error and its location which we can then correct for. This assumes that the encoding and decoding process is error free [39]. An example of an error correction code is given in Appendix A.

The utilisation of extra ancilla qubits to store information is a freedom we do not have in the NISQ era of QC. Number of qubits is a restricted resource with a large cost. For this reason, the community has developed quantum error mitigation (QEM) techniques which we shall discuss now.

3.4 Error mitigation

Error mitigation methods require no extra qubits and can suppress errors with simple post-processing of different runs of the quantum circuits. Below we explore the extrapolation and probabilistic error cancellation techniques but note that there are several other methods such as subspace expansion [40] and stabiliser based checks [41].

3.4.1 Extrapolation

The linear extrapolation technique was independently introduced by the IBM group and Endo et al. [42, 43]. The basic premise in achieving a mitigated result is to boost the error rate in the quantum circuit in a systematic manner and extrapolate to the zero noise limit. This is familiar to anyone who has utilised a curve fitting toolbox but below we shall explore a more mathematical route to put our ideas on a firm footing.

Suppose the observable of interest has the expectation value, $\langle \hat{A} \rangle(\epsilon)$, which is a function of the minimum possible error rate, ϵ . This is inherently present due to the constraint provided by the current limitations in quantum-information processing. One can Taylor expand this as

$$\begin{aligned} \langle \hat{A} \rangle(\epsilon) &= \langle \hat{A} \rangle(0) + \langle \hat{A}' \rangle(0)\epsilon + \frac{1}{2} \langle \hat{A}'' \rangle(0)\epsilon^2 \\ &= \langle \hat{A} \rangle(0) + \sum_{j=1}^n A_j \epsilon^j + \mathcal{O}(\epsilon^{n+1}) \end{aligned} \tag{3.11}$$

where A_j is independent of ϵ and $\langle \hat{A} \rangle(0)$ is the noiseless expectation value. For $\epsilon \ll 1$ we have $\langle \hat{A} \rangle(\epsilon) \approx \langle \hat{A} \rangle(0)$. If $\langle \hat{A} \rangle(\epsilon)$ is measured for several different noise rates $a_j \epsilon$ where $a_0 = 1 < a_1 < a_2 \dots < a_n$ then the ideal value $\langle \hat{A} \rangle(0)$ can be estimated by the Richardson extrapolation method [44]

$$\begin{aligned} \langle \hat{A} \rangle(0) &= \sum_{i=0}^n \gamma_i \langle \hat{A} \rangle(a_i \epsilon) \\ &= \sum_{i=0}^n \gamma_i \langle \hat{A} \rangle(0) + \sum_{j=1}^n A_j \epsilon^j \sum_{i=0}^n \gamma_i a_i^j + \mathcal{O}(\epsilon^{n+1}) \\ &= \langle \hat{A} \rangle(0) + \mathcal{O}(\epsilon^{n+1}) \end{aligned} \quad (3.12)$$

where $\sum_{i=0}^n \gamma_i = 1$, $\sum_{i=0}^n \gamma_i a_i^j = 0$ for $j = 1, 2, \dots, n$ such that the higher order terms in ϵ cancel out.

Results shown in Figure 3.2 from the IBM team show a clear advantage of using a larger value of n to increase the accuracy of the expectation value. However the cost increases exponentially with n as confirmed in [45].

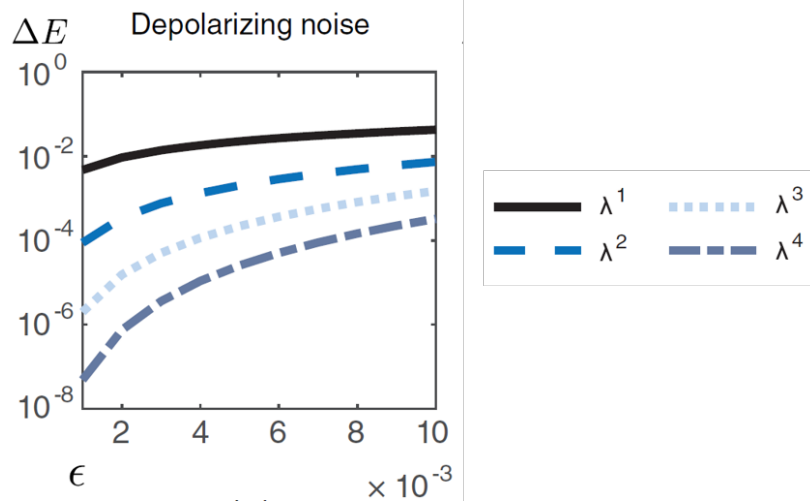


Fig. 3.2 Performance of the extrapolation technique for depolarising noise. The vertical axis is the deviation from the ideal value and the horizontal axis is the noise rate ϵ . n of λ^n is the number of points used in the extrapolation. The larger the value of n , the smaller the deviation from the ideal value. This figure is reproduced from [42].

By implementing the extrapolation technique, the IBM group was able to calculate the ground state energies of H_2 and LiH on a quantum computer [46]. Their results are shown in Figure 3.3 below. This is done by using a Variational-Quantum-Eigensolver (VQE) implemented on a hardware efficient ansatz. The race is now on for various groups around

the world to simulate molecules with a larger number of electronic orbitals and the reader is referred to the recent quantum chemistry review which summarises current ideas in this relatively young field [47].

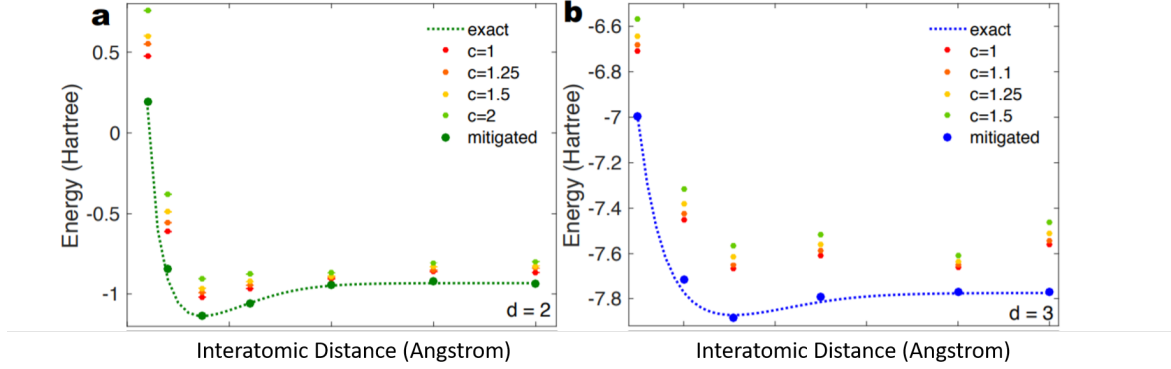


Fig. 3.3 Simulations of **a)** H_2 and **b)** LiH . The stretch factors, c , are proportional to the error rate in the circuit. The extrapolation technique was used to calculate the mitigated result in each case. The figure is reproduced from [46]. Note that some information from the original figure is omitted for brevity and the reader is referred to the original paper for further details.

3.4.2 Probabilistic error cancellation

The probabilistic error cancellation (PEC) method was independently introduced by Temme et al. [42] from the IBM group. Suppose that the ideal gate operation, \mathcal{U} , is affected by the depolarising error channel, \mathcal{D} , such that the actual implemented map is $\mathcal{D}\mathcal{U}$ where \mathcal{D} and \mathcal{U} are superoperators. The depolarising error channel is expressed as

$$\mathcal{D}(\rho) = (1 - p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z) \quad (3.13)$$

which is interpreted as the application of a random Pauli operator with probability $p/3$ inducing noise to the ideal gate operation. The main idea of mitigating the effect of the error channel \mathcal{D} , is to conjure up its inverse

$$\mathcal{D}^{-1}(\rho) = \gamma[p_1\rho - p_2(X\rho X + Y\rho Y + Z\rho Z)] \quad (3.14)$$

with

$$\gamma = \frac{4p+6}{6-8p}, \quad p_1 = \frac{12-4p}{8p+12}, \quad p_2 = \frac{4p}{8p+12} \quad (3.15)$$

such that $\mathcal{D}^{-1}(\mathcal{D}\mathcal{U}) = \mathcal{U}$. A schematic overview to realise the inverse channel is shown in Figure 3.4 below.

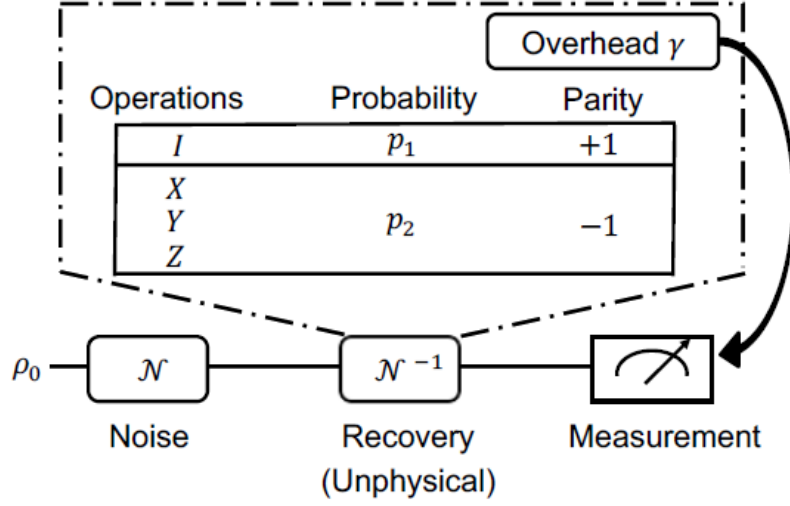


Fig. 3.4 After each gate operation, we apply I with probability p_1 and X, Y or Z each with probability p_2 . We multiply the measurement result by -1 if the Pauli operator was added. We repeat the experiment several times and multiply the average by the overhead coefficient γ to estimate the mitigated result. Figure reproduced from [47].

If the PEC technique is applied to a quantum circuit containing many gates, the error and the inverse channel, \mathcal{D} and \mathcal{D}^{-1} , is realised after each gate operation. Thus the overhead factor, γ , scales according to the number of times the inverse channel is called upon.

For two-qubit gates, the depolarising error channel is given by $\mathcal{D}_2(\rho_1, \rho_2) = \mathcal{D}(\rho_1) \otimes \mathcal{D}(\rho_2)$ where \mathcal{D}_2 contains terms such as

$$\mathcal{D}_2(\rho_1, \rho_2) = \frac{p - p^2}{3} \rho_1 X \rho_2 X + \frac{p^2}{9} X \rho_1 X X \rho_2 X + \dots \quad (3.16)$$

The first term in Equation 3.16 is interpreted as leaving qubit one idle and applying the X gate operation to qubit 2 with probability $p - (p^2/3)$. The inverse noise channel for two-qubit depolarising errors, \mathcal{D}_2^{-1} , can be constructed in a similar manner.

A major assumption implicitly applied in PEC is that the error suppression operations are themselves perfect. For example, applying I, X, Y, Z to realise \mathcal{D}^{-1} induces no further errors. One can think of these as inducing errors to the second order and thus consider them small.

Although the PEC technique is mathematically elegant and numerically mitigating, its practical usage is limited. Implementation of PEC requires full knowledge of the noise associated with each gate operation. This can be calculated via process tomography but is

heavily reliant upon the choice of hardware to realise qubits. There is a further restriction upon its usage as it is limited to mitigating Markovian noise and will not work for correlated errors [42].

Chapter 4

Method & results

Throughout the investigation, I employed the usage of IBM Quantum Experience (IBMQ) which allows free access to quantum devices via the cloud and also high-performance simulators [48]. The platform is accessed via a quantum-information kit (*Qiskit*), which is an open-source framework, allowing various levels of interactive freedom with the hardware. Moreover, the IBM package has an extensive collection of notebooks, guides and online forums to aid the layman in their pursuit of enhancing knowledge of this field [48].

There are a multitude of other quantum simulators available to use, namely pyQuil by Rigetti and QuEST by the group at Oxford University but their support is limited [49, 50]. Moreover, Rigetti also offers cloud based access to quantum hardware but this is on a pay-per-use basis.

4.1 Technicalities

4.1.1 Random circuit generation

To test our error mitigation techniques, it was crucial to develop a method to generate random circuits with a specified number of qubits and cd . The mean of the standard deviation of the average number of gates applied for a range of cd 's is 5. Theoretically, this could be lowered but at the time of writing this, I noticed a bug in the Qiskit compiler which affected this result. More details are given in Appendix B. Figure 4.1 is an example of a circuit with $cd = 10$.

4.1.2 Implementing and inverting noise

A *noise* function, with a single and two-qubit depolarising error probability, p_s and p_d was called upon after each gate operation in the quantum circuit which enacted the noise channel

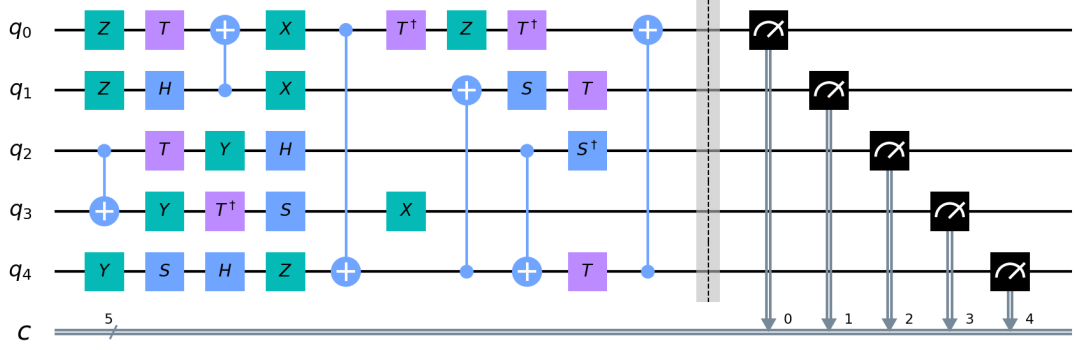


Fig. 4.1 Example of a randomly generated circuit. Each single qubit gate was picked randomly from the set $\{I, X, Y, Z, H, S, S^\dagger, T, T^\dagger\}$. Control and target for each CNOT were also picked at random.

given by Equation 3.13 and 3.16 respectively. Qiskit allows the implementation of defined noise models but I found the functionality of its usage to be restrictive for the purposes of this investigation. This is because I average over 1000 runs of each circuit where each run consists of 100 shots. For the purposes of my implementation technique, I require the dynamical freedom of noise varying with each run but the defined Qiskit noise model varies this with each shot.

A *correction* function enacting the single and two-qubit inverse channels, \mathcal{D}^{-1} and \mathcal{D}_2^{-1} , as explained in section 3.4.2, was called upon after each gate operation in the circuit. This is relevant for the purposes of the PEC technique.

4.2 Numerical results

Extrapolating to the zero noise limit

Verifying the successful implementation of the linear extrapolation technique was key in order to progress to the more complicated PEC technique. This was done on randomly generated circuits, as shown in Figure 4.1 with $n = 5$ qubits and $cd = 10$. Each gate operation was subject to the appropriate depolarising noise channel as given by Equations 3.16 and 3.13 with single and two-qubit error probabilities, $p_s = 10^{-4}$ and $p_d = 10^{-2}$ respectively. This signifies error rates achievable by current quantum hardware [51].

The error probabilities were increased in a systematic manner by a factor m where $m \in \{1, 2, 3, 4\}$. Four values were used to extrapolate to the zero noise limit as to realistically enact an experimentalist's approach in the lab. As mentioned in section 3.4.1, the estimated

error free value would be more accurate if more points were chosen to extrapolate from, however, this would increase the cost of the mitigation technique exponentially.

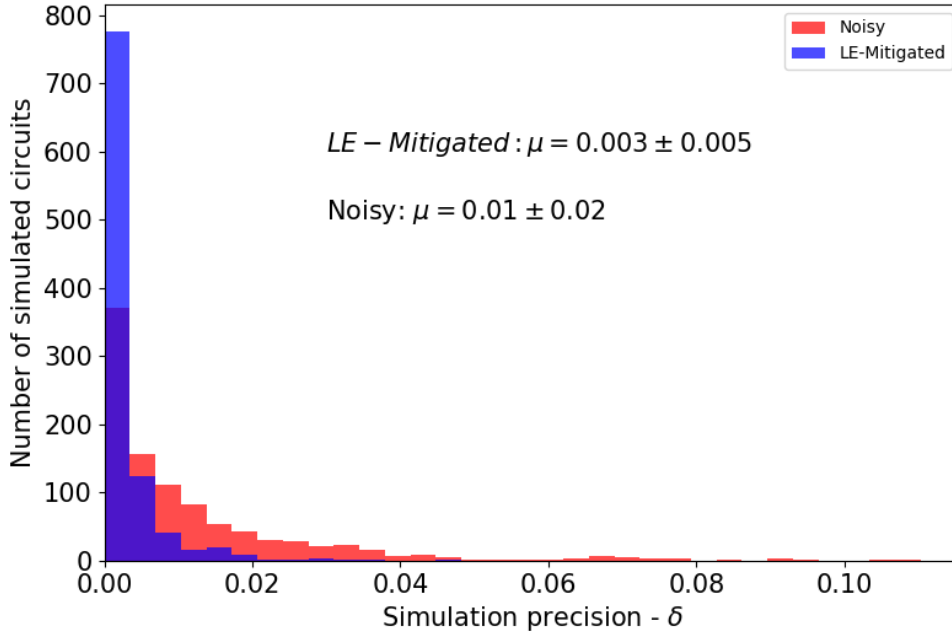
Quantum algorithms such as Trotterisation require outputs from which probabilities of achieving a given quantum state is calculated. [52, 53]. With this in mind, I define the measured quantity of a given circuit as the probability of achieving the state with the largest weighting out of the 2^n possible states in the ideal case. This is an arbitrary definition which I adopt throughout the use of this report, unless otherwise stated, but one has the freedom to pick any manipulated quantity of the final distribution of values. Results for simulations of the extrapolation technique is shown in Figure 4.2a below.

Note that in experimental situations, we do not have the freedom of employing noise functions to enact error channels. It is thus not possible to tune the error probability by multiplying the unwanted environmental influence by a factor of m . To circumvent this, the general method adopted by the community is to add additional gates into the original circuit which act as dummy operations inducing error but not changing the essence of the algorithm being simulated. For example, if a quantum circuit is described as $|\psi\rangle = X|0\rangle$, one could induce further errors by applying 2,4,6.. additional X gates where a pair of gate operations theoretically do not affect the output. This method has been recently used to calculate nuclear energy levels using imaginary time evolution [54].

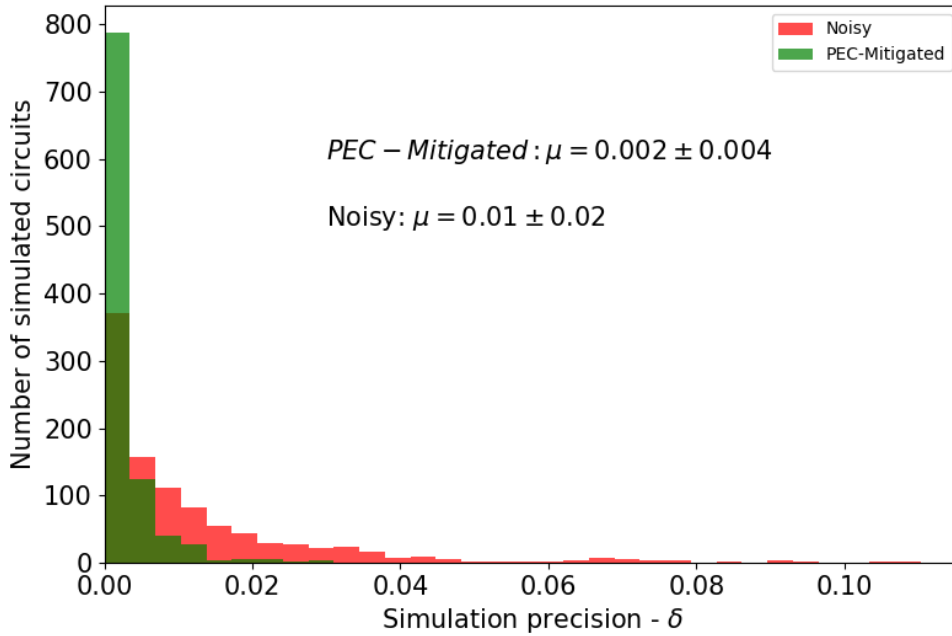
Probabilistic error cancellation

To date, the probabilistic error cancellation method has only been numerically implemented by the original authors, Temme et al. [42] and the group based at Oxford University [43].

PEC was also simulated on random circuits with the same parameters as mentioned in the implementation of the extrapolation technique. However, in this case, the noise function was followed by a correction function to enact the inverse channels as described in Section 3.4.2. Results comparing its operational viability is presented in Figure 4.2b below.



(a) Results for simulations with and without the extrapolation technique



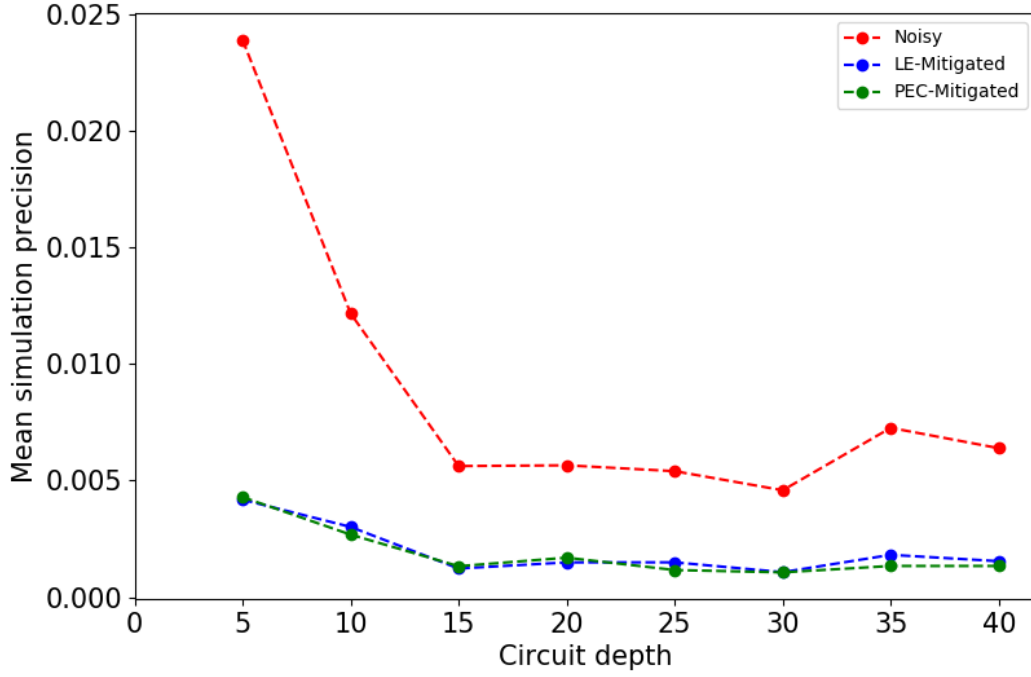
(b) Results for simulations with and without the PEC technique

Fig. 4.2 Histograms showing simulated results for the deviation of the defined measured quantity from the ideal value for 1000 randomly generated circuits with and without QEM. Each circuit has $n = 5$ qubits and $cd = 10$. An average from 1000 runs was calculated which is necessary to induce a range of noise and correction parameters. The range of runs also imposes a fixed constraint on our simulated time resource mirroring the actions of an experimentalist. The mean, μ , of each distribution is labelled on the figure. Notice that the mean shifts closer to 0 in the mitigated case compared to the noisy case. Results simulated on the Cambridge service for data-driven discovery, high-performance computing cluster [55]. Code available at [56].

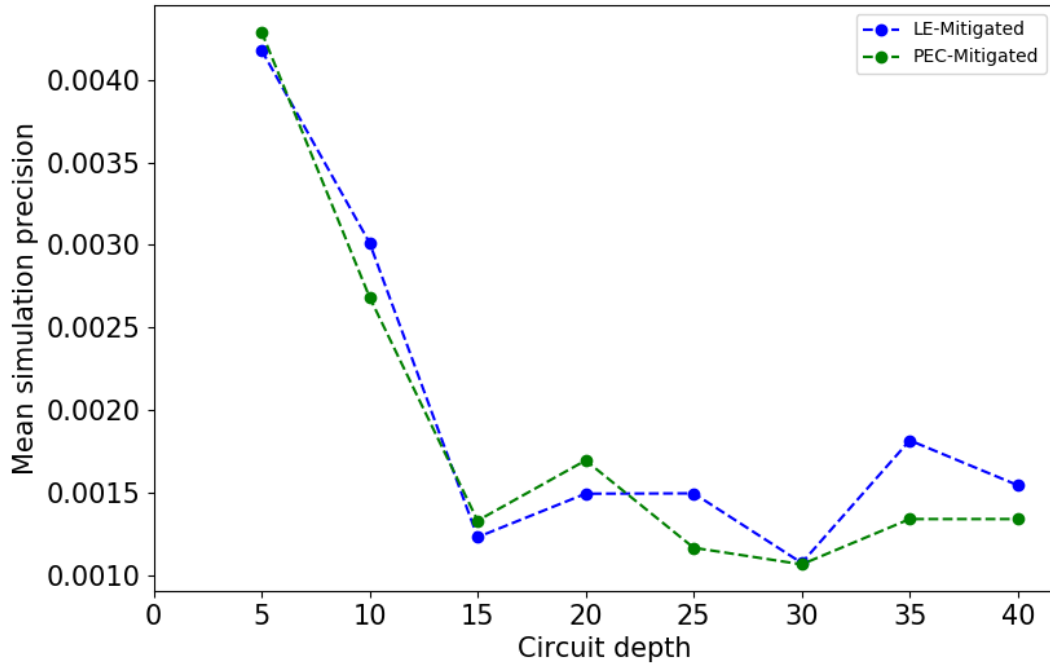
Functionality of PEC and extrapolation as a function of circuit depth

Quantum algorithms which simulate molecules on quantum computers scale according to the number of electronic orbitals involved [57]. It is thus crucial to test the performance of the error mitigation schemes against cd . One could also grow quantum circuits vertically by incorporating more qubits but this is not a readily available choice in the NISQ era.

I investigated the performance of PEC and extrapolation for cd 's in the range of 5 to 40 with the same simulation parameters as mentioned for Figure 4.2. For each cd , 100 circuits were simulated and their results averaged which is shown in Figure 4.3 below.



(a) Performance of error mitigation techniques against noisy simulations as a function of cd



(b) Performance of PEC and the extrapolation technique as a function of cd

Fig. 4.3 Performance of QEM schemes as a function of cd compared to the noisy simulations. The observed quantity is the probability of achieving the state with the largest weighting in the ideal case. Each data point is an average of 100 random circuits with $n = 5$ qubits. Results with error bars are shown in Appendix C. Results simulated on the Cambridge service for data-driven discovery, high-performance computing cluster [55]. Code available at [56].

4.3 Experimental applications of probabilistic error cancellation

I shall refrain from experimentally validating the applicability of the extrapolation technique as this has been conducted extensively in the literature for algorithms much more complex than the ones I have mentioned here [58]. I will however examine the experimental application of the PEC technique in light of the promising results shown in Figure 4.2b.

IBMQ 5-qubit devices were used to implement PEC where the quantum circuit is the maximally entangled n -qubit GHZ state as defined in section 2.6.4. In this setting, the measured quantity is defined to be

$$\mathcal{P} = \sum_{n=1}^5 P(0^{\otimes n}) + P(1^{\otimes n}) \quad (4.1)$$

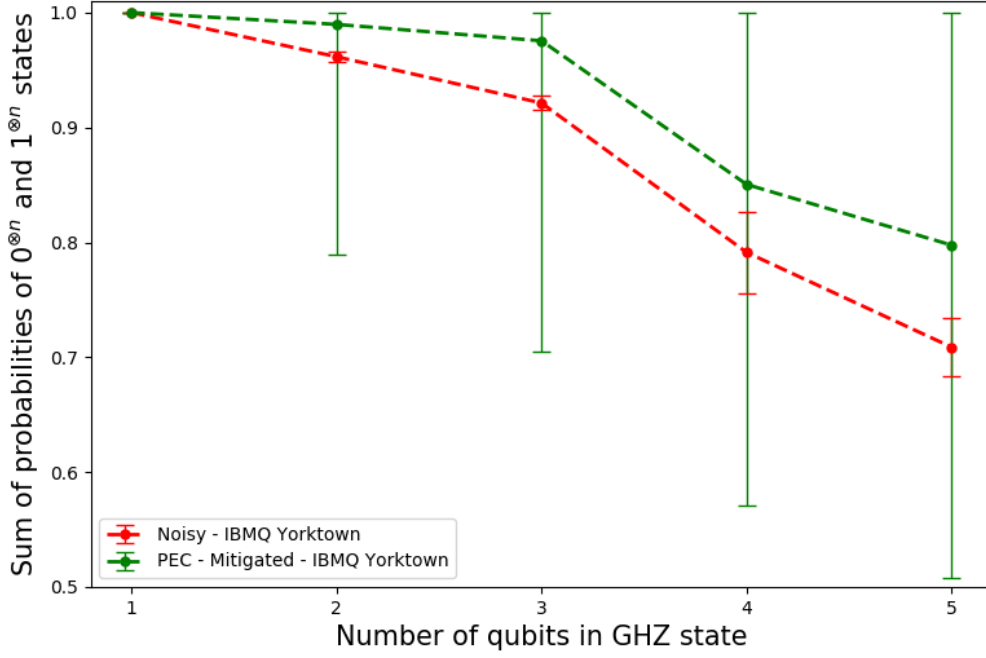
where $P(0^{\otimes n})$ and $P(1^{\otimes n})$ are the probabilities of achieving the $|0\rangle^{\otimes n}$ and $|1\rangle^{\otimes n}$ states and n is the number of qubits. A maximum value of n is restricted to 5 as that is the limitation imposed by the publicly available hardware. Ideally, for a given backend, the value of $\mathcal{P} = 5$.

Measuring \mathcal{P} for various devices allowed me to extract the device most prone to errors out of the 8 that are available. This can be thought of as a benchmarking procedure which yielded the IBMQ Yorktown device with $\mathcal{P} = 4.38 \pm 0.04$. It is to be noted that this procedure yields no information about the noise model which dominates a particular device.

To determine the full extent of the noise model that accompanies a device, one would have to employ methods of process tomography. For gate errors, this entails conducting a procedure termed randomised benchmarking in which sequences of gates is applied with the goal of taking the qubit on a random walk among certain points on the Bloch sphere and returning it to the $|0\rangle$ state it started in. As the number of gates in the sequence is increased, the chance of returning to $|0\rangle$ drops exponentially and eventually saturates near 50%. The gate error rate is extracted from the fit to this exponential decay [51].

For the purposes of this investigation, it will be interesting to to apply the PEC technique based on a depolarising error channel ansatz to investigate the extent of its role in hindering device performance. Details of single and two-qubit gate error rates, p_s and p_d , accompanying a device are available on the IBMQ website [48]. These values vary with each qubit in the device as the fabrication of each junction is not exactly identical. However, variations of error rates are on the same order of magnitude for each qubit and hence for the purposes of this investigation, it will suffice to take an average value of $p_s = 7.64 \times 10^{-4}$ and $p_d = 1.14 \times 10^{-2}$. Ideally, the error probabilities should vary with the qubit under consideration and this is currently possible with the compiler functionality provided by Qiskit. However, delving into

gate decomposition and qubit routing is something I briefly investigated but is not the focus of this thesis [59]. I will assume that p_s and p_d are the error rates for the depolarising error channel but in reality, these values represent a metric based on all possible errors that can be induced. Experimental and simulated results are shown in Figure 4.4a and 4.4b .



(a) Experimental implementation of PEC



(b) Simulated results for PEC

Fig. 4.4 (a) Experimental and (b) simulated results for the implementation of PEC. Each data point is averaged over 100 and 1300 runs for the noisy and PEC-Mitigated cases respectively. 1300 runs for the PEC-Mitigated case is a constraint imposed by the availability of run time on the IBMQ Yorktown device. Code available at [56].

Chapter 5

Discussion of results

A careful examination of Figure 4.2 proves that QEM techniques improve the results, i.e. the distribution of measured values corresponding to each circuit, out of the 1000 random circuits, is closer to zero compared to the noisy case. The mean of the distribution of values for the mitigated results is $\mu = 0.003 \pm 0.005$ and $\mu = 0.002 \pm 0.004$ for the extrapolation and PEC cases respectively whereas for the noisy simulations, $\mu = 0.01 \pm 0.02$. With these results in mind and their associated errors, it is unclear as to which technique is better to employ under the strain of the depolarising error channel.

Figure 4.3 explores the performance of QEM techniques under the varying parameter of cd . It was expected that their performance declines significantly as the cd increases due to the strain of the noise channel taking a toll on the desired outputs. However, my results show that the absolute difference in values between cd 40 and 5 is 0.003 ± 0.004 and 0.003 ± 0.005 for PEC and extrapolation techniques respectively. These values show that their individual performance is comparatively robust under the range of cd 's I have tested for. Moreover, a direct comparison yields no clear advantage as to which method should be employed considering the parameters under which they are simulated.

Near-term experiments in the NISQ era are exploring the capabilities of around 100 qubits and $cd < 100$ [30]. With this in mind, it is sufficient to say that both QEM techniques will be of use if the trend shown in Figure 4.3b continues. However, as discussed previously, the PEC method requires full knowledge of the noise model associated with a particular device. This information is not readily available and one has to undertake process tomography to gain this knowledge. Whereas the extrapolation method requires no knowledge of the noise model and is mitigating under the influence of all noise channels [57].

Figure 4.3 shows an initial spike in the simulation precision between cd 's 5 to 15. This behaviour suggests that there is more noise induced in the circuit for smaller cd values between 5-15. The simulations were carried out under error probabilities of $p_s = 10^{-4}$ and

$p_d = 10^{-2}$. Hence a larger number of CNOT gate operations for smaller cd values relative to $cd > 15$ can be a source of this anomaly. A detailed investigation into the method chosen to generate random circuits suggests this to be the case.

I define a measure denoted by ζ where ζ is the number of two-qubit gates divided by the cd . $\zeta = 0.5$ and $\zeta = 0.1$ for cd 's of 5 and 40 respectively. This shows that it is five times more likely for there to be a two-qubit gate compared to a single qubit gate and hence translates to an increase in the error probability of the depolarising error channel five fold for $cd = 5$ compared to $cd = 40$ respectively. This explains the five fold increase in the mean simulation precision shown in Figure 4.3a. This is a flaw in the random circuit generation method developed.

Finally, we investigate the results shown in Figure 4.4. Figure 4.4b shows data from the simulator which is affected purely by the depolarising error channel. The defined measure, $\mathcal{P} = 4.8 \pm 0.4$ and $\mathcal{P} = 4.9 \pm 0.7$ for the simulated noisy and mitigated cases respectively. The mitigated value is closer to the ideal result of $\mathcal{P} = 5$ but the deviation is broader due to the cost of the method as explained in [60]. It is expected that more runs will reduce the deviation.

The methodology I developed to simulate the numerical implementation of PEC induces a dynamic range of noise and correction parameters at the level of runs but is static for all shots within a given run. This is not the case for the Yorktown device as the noise is variant with each shot of each run. Hence adding correction parameters to mitigate a depolarising noise channel ansatz during each run when the noise is also varying at the level of shots is not the required methodology. Hence, a conclusion that the defined measured value of $\mathcal{P} = 4.38 \pm 0.04$ and $\mathcal{P} = 4.6 \pm 0.5$ for the experimental noisy and mitigated cases respectively, demonstrates that PEC increases accuracy on IBMQ Yorktown is drawn from the implementation of an imperfect method.

Moreover, error prone results shown in Figure 4.4a and 4.4b are an average of 100 runs with each run consisting of 1000 shots. This means that there are 10^5 independent noisy events occurring on the Yorktown device but only 10^2 on the simulator. Hence the probability of the depolarising error channel being induced is reduced in the simulated case and it cannot be concluded, with certainty, that various types of error channels are induced on the Yorktown device.

To achieve the wanted methodology, I would need research level access to the IBMQ devices which can be obtained through joining the IBMQ network. This would provide me with unrestricted access for a fixed time where I could conduct experiments and average over the required number of parameters.

Chapter 6

Departing thoughts

We started by introducing the basic principles behind quantum mechanics and applied it to computing to birth quantum-information science. Some examples of the practical uses of quantum computing are scattered throughout the thesis but it is fair to say that we have not done this justice. Decoherence was discussed and QEC was mentioned to get an idea for why we need require QEM. Extrapolation and PEC was explored, numerically implemented and an attempt within our means was made to gather utility of PEC in experimental situations.

Previous work in numerically simulating PEC was conducted for a fixed depolarising error probability for single and two-qubit gates. This does not represent current hardware. I have extended the numerical simulations to vary the error probability depending on the gate size. The utility of QEM techniques was investigated for varying ϵ 's with current algorithmic complexity in mind. It was found that the simulation precision is identical for extrapolation and PEC but the utility of PEC is limited to mitigating depolarising noise. It has been shown by [42] that PEC can also be applied to amplitude damping channels but more work needs to be carried out to generalise it to a range of error models.

This work could be directly extended to investigate a plethora of ideas. Firstly a successful experimental application of PEC is a natural progression. At the time of writing this, I have initiated the request to gain research level access to IBMQ systems. It is my hope that this will be ready for the next interested student. Secondly, it will be interesting to find a regime in circuit complexity where the simulation precision in QEM techniques deviate. This will bring about information which would help researchers decide the techniques they should utilise. Lastly, and one in which the community has great interest in, is the calculation of the ground state energies of Hamiltonian's. I predict a huge influx of research in quantum computational chemistry and it will be mandatory to learn techniques such as VQE's [61] if one would like to get involved.

I leave my personal investigations in this field unfinished as it has sparked more questions than it has answered.

References

- [1] John Preskill. Quantum computing in the nisq era and beyond, 2018.
- [2] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, 1995.
- [3] Richard Jozsa. Searching in grover’s algorithm. *arXiv preprint quant-ph/9901021*, 1999.
- [4] David Collins, KW Kim, and WC Holton. Deutsch-jozsa algorithm as a test of quantum computation. *Physical Review A*, 58(3):R1633, 1998.
- [5] Seth Lloyd. Universal quantum simulators. *Science*, pages 1073–1078, 1996.
- [6] Benjamin P Lanyon, James D Whitfield, Geoff G Gillett, Michael E Goggin, Marcelo P Almeida, Ivan Kassal, Jacob D Biamonte, Masoud Mohseni, Ben J Powell, Marco Barbieri, et al. Towards quantum chemistry on a quantum computer. *Nature chemistry*, 2(2):106–111, 2010.
- [7] Richard P Feynman. Simulating physics with computers. *Int. J. Theor. Phys*, 21(6/7), 1999.
- [8] David McKay, Lev Bishop, Antonio Corcoles, Petar Jurcevic, Abhinav Kandala, Jin-Sung Kim, Isaac Lauer, Seth Merkel, Zlatko Minev, Neereja Sundaresan, et al. Benchmarking tools for nisq systems. *Bulletin of the American Physical Society*, 2020.
- [9] Yuichiro Matsuzaki, Victor M Bastidas, Yuki Takeuchi, William J Munro, and Shiro Saito. One-way transfer of quantum states via decoherence. *Journal of the Physical Society of Japan*, 89(4):044003, 2020.
- [10] Joe O’Gorman and Earl T Campbell. Quantum computation with realistic magic-state factories. *Physical Review A*, 95(3):032338, 2017.
- [11] Abhilash Ponnath. Difficulties in the implementation of quantum computers. *arXiv preprint cs/0602096*, 2006.
- [12] Peter W Shor. Scheme for reducing decoherence in quantum computer memory. *Physical review A*, 52(4):R2493, 1995.
- [13] Andrew M Steane. Error correcting codes in quantum theory. *Physical Review Letters*, 77(5):793, 1996.
- [14] Garnet Kin Chan. Quantum algorithms of interest on nisq machines. 2020.

- [15] Mara Beller. The conceptual and the anecdotal history of quantum mechanics. *Foundations of Physics*, 26(4):545–557, 1996.
- [16] BD Joesphson. Possible new effects in superconductive tunneling. *Phys. Lett*, 1(7):251, 1962.
- [17] 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, Jun 2019.
- [18] C. Adami and N. J. Cerf. Quantum computation with linear optics. *Lecture Notes in Computer Science*, page 391–401, 1999.
- [19] Colin D. Bruzewicz, John Chiaverini, Robert McConnell, and Jeremy M. Sage. Trapped-ion quantum computing: Progress and challenges. *Applied Physics Reviews*, 6(2):021314, Jun 2019.
- [20] Daniel Loss and David P. DiVincenzo. Quantum computation with quantum dots. *Physical Review A*, 57(1):120–126, Jan 1998.
- [21] Francisco M Fernández. *Introduction to perturbation theory in quantum mechanics*. CRC press, 2000.
- [22] Martin B Plenio and Vlatko Vedral. Teleportation, entanglement and thermodynamics in the quantum world. *Contemporary physics*, 39(6):431–446, 1998.
- [23] Anders Karlsson and Mohamed Bourenane. Quantum teleportation using three-particle entanglement. *Physical Review A*, 58(6):4394, 1998.
- [24] Chuan Wang, Fu-Guo Deng, Yan-Song Li, Xiao-Shu Liu, and Gui Lu Long. Quantum secure direct communication with high-dimension quantum superdense coding. *Physical Review A*, 71(4):044305, 2005.
- [25] Alain Aspect. Bell’s inequality test: more ideal than ever. *Nature*, 398(6724):189–190, 1999.
- [26] Vlatko Vedral et al. *Introduction to quantum information science*. Oxford University Press on Demand, 2006.
- [27] Qian Dong, Guo-Hua Sun, Mohamad Toutounji, and Shi-Hai Dong. Tetrapartite entanglement measures of ghz state with nonuniform acceleration. *Optik*, 201:163487, 2020.
- [28] Bülent Demirel, Weikai Weng, Christopher Thalacker, Matty Hoban, and Stefanie Barz. Correlations for computation and computation for correlations. *arXiv preprint arXiv:2005.01780*, 2020.
- [29] Dik Bouwmeester, Jian-Wei Pan, Matthew Daniell, Harald Weinfurter, and Anton Zeilinger. Observation of three-photon greenberger-horne-zeilinger entanglement. *Physical Review Letters*, 82(7):1345, 1999.
- [30] John Preskill. Lecture notes for quantum information and computation. http://www.theory.caltech.edu/~preskill/ph219/chap5_13.pdf. (Accessed on 05/10/2020).

- [31] W. Huang, C. H. Yang, K. W. Chan, T. Tanttu, B. Hensen, R. C. C. Leon, M. A. Fogarty, J. C. C. Hwang, F. E. Hudson, K. M. Itoh, and et al. Fidelity benchmarks for two-qubit gates in silicon. *Nature*, 569(7757):532–536, May 2019.
- [32] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.
- [33] Sergio Boixo, Sergei V Isakov, Vadim N Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J Bremner, John M Martinis, and Hartmut Neven. Characterizing quantum supremacy in near-term devices. *Nature Physics*, 14(6):595–600, 2018.
- [34] Robin Harper, Steven T. Flammia, and Joel J. Wallman. Efficient learning of quantum noise, 2019.
- [35] Arman Babakhani, Herman Chan, Jeffrey Epstein, Song Zhang, Juan Atalaya, and Birgitta Whaley. Continuous quantum error correction with two-qubit annealing hamiltonian. In *APS Meeting Abstracts*, 2019.
- [36] Qihao Guo, Yuan-Yuan Zhao, Markus Grassl, Xinfang Nie, Guo-Yong Xiang, Tao Xin, Zhang-Qi Yin, and Bei Zeng. Testing a quantum error-correcting code on various platforms. *arXiv preprint arXiv:2001.07998*, 2020.
- [37] Eugen Dumitrescu and Pavel Lougovski. Assigning hamiltonians to open quantum systems. *Bulletin of the American Physical Society*, 2020.
- [38] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.
- [39] Andrew M Steane. A tutorial on quantum error correction. In *PROCEEDINGS-INTERNATIONAL SCHOOL OF PHYSICS ENRICO FERMI*, volume 162, page 1. IOS Press; Ohmsha; 1999, 2007.
- [40] Jarrod R McClean, Mollie E Kimchi-Schwartz, Jonathan Carter, and Wibe A de Jong. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Physical Review A*, 95(4):042308, 2017.
- [41] Xavi Bonet-Monroig, Ramiro Sagastizabal, M Singh, and TE O’Brien. Low-cost error mitigation by symmetry verification. *Physical Review A*, 98(6):062339, 2018.
- [42] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. Error mitigation for short-depth quantum circuits. *Physical Review Letters*, 119(18), Nov 2017.
- [43] Suguru Endo, Simon C. Benjamin, and Ying Li. Practical quantum error mitigation for near-future applications, 2017.
- [44] Lewis Fry Richardson and J Arthur Gaunt. Viii. the deferred approach to the limit. *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, 226(636-646):299–361, 1927.

- [45] Xiao Yuan, Zhen Zhang, Norbert Lütkenhaus, and Xiongfeng Ma. Simulating single photons with realistic photon sources. *Phys. Rev. A*, 94:062305, Dec 2016.
- [46] Abhinav Kandala, Kristan Temme, Antonio D. Corcoles, Antonio Mezzacapo, Jerry M. Chow, and Jay M. Gambetta. Extending the computational reach of a noisy superconducting quantum processor, 2018.
- [47] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. Quantum computational chemistry. *Rev. Mod. Phys.*, 92:015003, Mar 2020.
- [48] IBMQ Quantum Experience. <https://quantum-computing.ibm.com/>. Accessed: 05-05-2020.
- [49] pyQuil. <http://docs.rigetti.com/en/stable//>. Accessed: 05-05-2020.
- [50] QuEST. <https://quest.qtechtheory.org/about//>. Accessed: 05-05-2020.
- [51] Easwar Magesan, Jay M. Gambetta, and Joseph Emerson. Characterizing quantum gates via randomized benchmarking, 2011.
- [52] Bela Bauer, Dave Wecker, Andrew J. Millis, Matthew B. Hastings, and M. Troyer. Hybrid quantum-classical approach to correlated materials, 2015.
- [53] Ying Li and Simon C. Benjamin. Efficient variational quantum simulator incorporating active error minimization. *Physical Review X*, 7(2), Jun 2017.
- [54] Kübra Yeter-Aydeniz, Raphael C. Pooser, and George Siopsis. Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and lanczos algorithms, 2019.
- [55] University of Cambridge, Research Computing Service. <https://www.hpc.cam.ac.uk/>. Accessed: 18-05-2020.
- [56] Code for the project. <https://github.com/zohimchandani/Quantum-Computing-QEM>. Accessed: 16-05-2020.
- [57] 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. *Nature*, 549(7671):242–246, 2017.
- [58] Kübra Yeter-Aydeniz, Raphael C. Pooser, and George Siopsis. Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and lanczos algorithms, 2019.
- [59] Alexander Cowtan, Silas Dilkes, Ross Duncan, Alexandre Krajenbrink, Will Simmons, and Seyon Sivarajah. On the qubit routing problem. *arXiv preprint arXiv:1902.08091*, 2019.
- [60] Suguru Endo, Simon C. Benjamin, and Ying Li. Practical quantum error mitigation for near-future applications. *Physical Review X*, 8(3), Jul 2018.

-
- [61] Harper R Grimsley, Sophia E Economou, Edwin Barnes, and Nicholas J Mayhall. Adapt-vqe: An exact variational algorithm for fermionic simulations on a quantum computer. *arXiv preprint arXiv:1812.11173*, 2018.
 - [62] J. L. O’Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A. G. White. Quantum process tomography of a controlled-not gate, 2004.
 - [63] Vladimir Bužek and Mark Hillery. Quantum copying: Beyond the no-cloning theorem. *Physical Review A*, 54(3):1844, 1996.

Appendix A

Quantum error correction

As always, we take inspiration from the classical world to determine a starting point for our error correction technique. The simplest example is by *copying* the information stored in a *bit* multiple times, passing the information through a noisy information processing channel, *measuring* the resulting outputs and decoding through *majority voting*.

The astute reader may notice a major flaw with adopting a direct 1-1 mapping of ideas from the classical to the quantum world. The *no-cloning* theorem forbids making copies of quantum states and the *measurement* process destroys superposition which is a key resource in our quest for viable quantum computation. Nevertheless, the three-qubit code described below overcomes these problems and provides an insight into the underlying mechanisms behind the theory.

A.1 Three-qubit code

Suppose that we would like to protect our quantum state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ which is under attack by bit flip errors. This is to say that the error channel induces an X gate with a certain probability during information processing. Measuring this error probability is done using *process tomography* which is a standalone investigation in its own right [62]. We can encode the state as $|\psi\rangle \longrightarrow \alpha|000\rangle + \beta|111\rangle$ using the first two CNOT gate operations as shown in Figure A.1.

The no-cloning theorem forbids multiple copies of the same quantum state [63]. One cannot produce, $(\alpha|0\rangle + \beta|1\rangle)^{\otimes 3}$, but copying within the same basis is not restricted by quantum theory. We then pass our encoded state for information processing which is prone to errors. Ideally we would like to measure the output to determine which of the bits have suffered a flip error however this would destroy superposition. To overcome this, we introduce

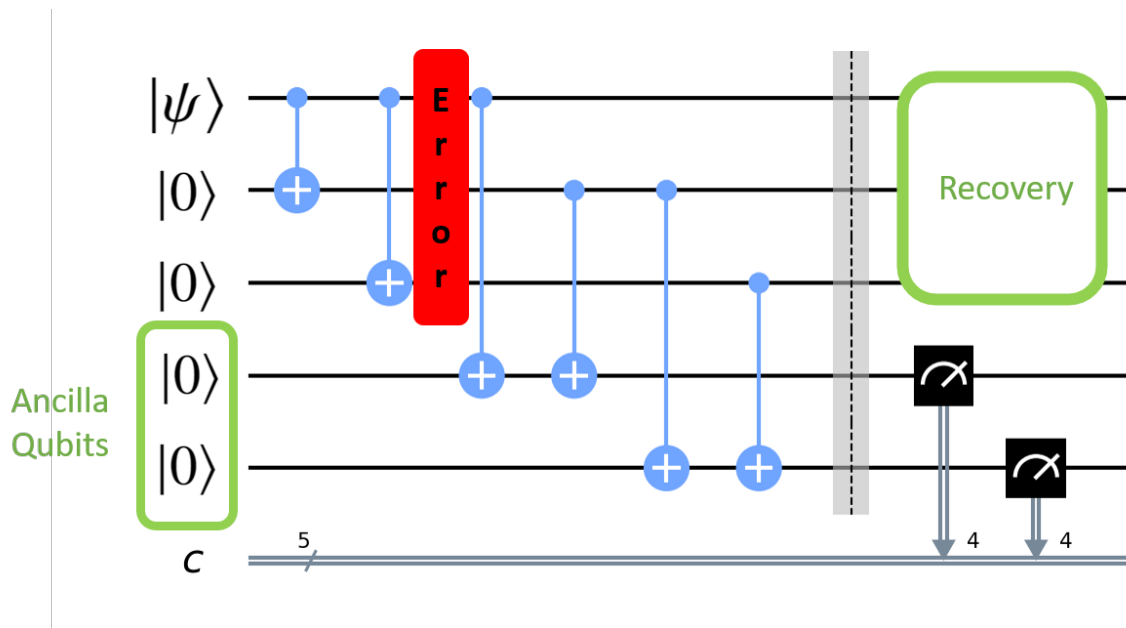


Fig. A.1 Circuit required to encode and correct for a bit-flip error. We assume that after encoding a single bit-flip occurs on one of the three qubits (or no error occurs). Two initialised ancilla are then coupled to the data block which only checks the parity between qubits. These ancilla are then measured, with the measurement result indicating where (or if) an error has occurred, without directly measuring any of the data qubits. Using this syndrome information, the error can be corrected with an additional X gate in the recovery process.

ancilla qubits which allow us to do a *syndrome* measurement by encoding the parity of states within them.

Our problem of figuring out which one of the 3 qubits has suffered a bit flip attack can be broken down into two sub-questions. Are qubits 0 and 1 the same? Are qubits 1 and 2 the same? By addressing these questions, we can figure out which qubit suffered damage and this can be corrected for in the recovery phase by adding an additional operation. This is summarised in Table A.1 below.

Error Location	Final State, $ data\rangle ancilla\rangle$
No Error	$\alpha 000\rangle 00\rangle + \beta 111\rangle 00\rangle$
Qubit 0	$\alpha 100\rangle 11\rangle + \beta 011\rangle 11\rangle$
Qubit 1	$\alpha 010\rangle 10\rangle + \beta 101\rangle 10\rangle$
Qubit 2	$\alpha 001\rangle 01\rangle + \beta 110\rangle 01\rangle$

Table A.1 Possible states of $|\psi\rangle$ and the ancilla qubits for a single X error on one of the qubits. Each possible error will result in a unique syndrome measurement result of the ancilla.

Error mitigation techniques such as this allow us to recover the state that we ideally desire but with a cost: we need extra qubits to store information which we cannot measure. This is not ideal as the current problem in scaling up quantum hardware is the number of logical qubits we can control. Hence, we cannot waste these precious resources on correcting errors.

Note that there are several other methods such as the Shor and the Steane code, concatenated code and topological error correction codes. The interested reader is referred to a set of excellent notes by Andrew Steane titled: '*A tutorial on quantum error correction*' [39].

Appendix B

Qiskit compiler

My method of random circuit generation focuses on the construction of a matrix in which values corresponding to each gate is inputted. This filled matrix is then realised into a circuit which can then be executed. However, the Qiskit compiler readjusts the order of some gates in the process. This is not ideal as it increases the variance in the number of gates for a specified cd. An example of this is shown in Figure B.1 below.

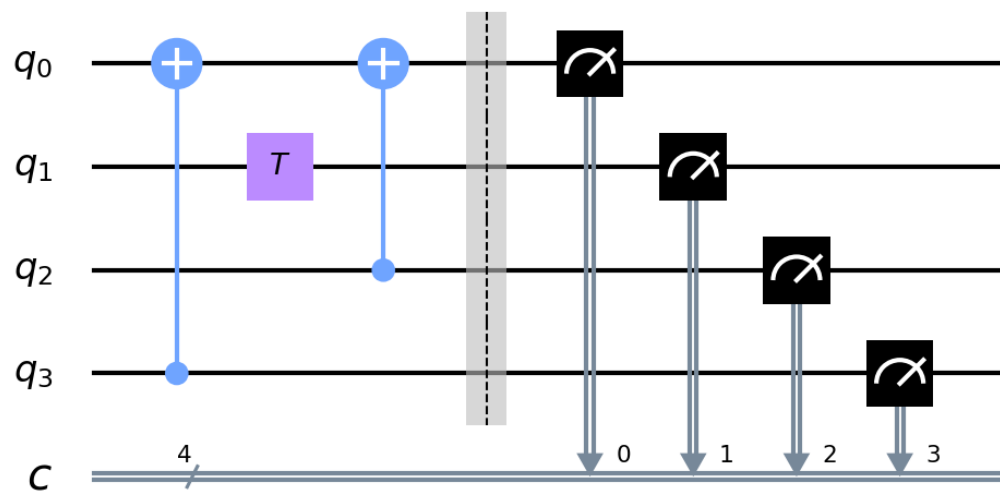


Fig. B.1 The circuit is initialised in the $|0000\rangle$ and two CNOT gates and T gate is added in the order stated. However the outcome is the reordered gate configuration shown in the figure.

Appendix C

Functionality of PEC and extrapolation as a function of circuit depth

Results with associated errors exploring the functionality of the QEM techniques is shown in Figure C.1.

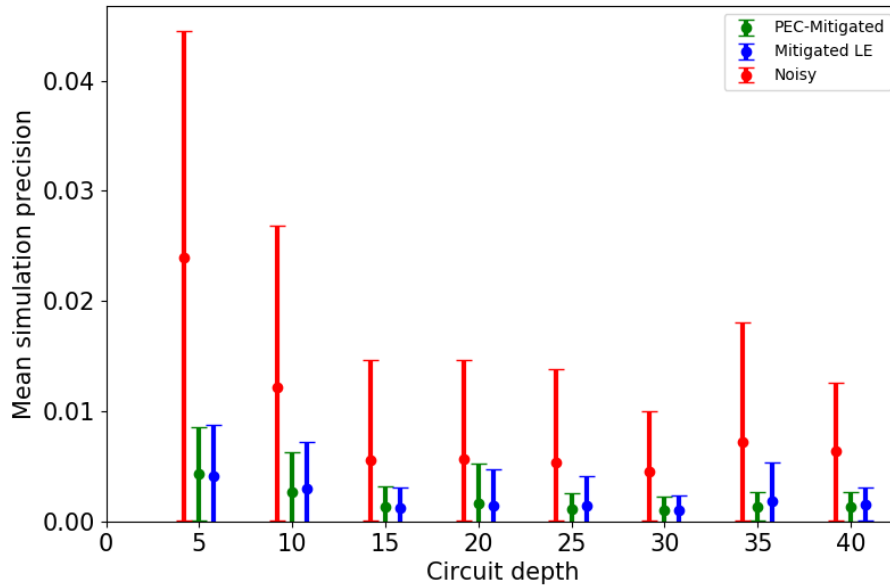


Fig. C.1 Performance of the error mitigation schemes as a function of cd compared to the noisy simulations. The observed quantity is the probability of achieving the state with the largest weighting in the ideal case. Each data point is an average of 100 random circuits with $n = 5$ qubits. Each circuit is simulated for 1000 runs and each run consists of 100 shots.