

Solving Optimisation Problems on Near-Term Quantum Computers

Benjamin Jaderberg

Wolfson College



University of Oxford

A thesis presented for the degree of

Doctor of Philosophy

Trinity 2022

For a successful technology, reality must take precedence over public relations, for nature cannot be fooled.

— Richard Feynman

Acknowledgements

This thesis marks the culmination of the past 4 years of my life, as a PhD student in the Clarendon Laboratory and as a resident in Wolfson college. Both of these I have enjoyed profoundly, in no small part because of the support of many important individuals who deserve a great deal of thanks.

I am extremely grateful to my supervisor Professor Dieter Jaksch, whose guidance and tutelage sits at the core of all the work I have done. His close attention and rigour helped placed me on firm ground, yet he afforded me the freedom to pursue my own ideas and develop into my own researcher. Most of all he has taught me that nothing worth doing is easy.

A special thanks is due to several other members of the Jaksch group. To Martin Kiffner, your mentorship and supervision was instrumental in developing the skills and confidence that allowed me to thrive. To Karsten Leonhardt, who's patience and knowledge helped me find my footing, and to Lewis Anderson, who's hard work and creativity helped achieve some of the work I am most proud of. To the other members of the group, thank you for your friendship, which has made my time so memorable and enjoyable.

Needless to say, this thesis would not exist without the unwavering support of my family and friends, with whom I am indebted. To my parents, who have encouraged and motivated me every step of the way. To my siblings, each one of you an inspiration to me and a reminder to push to achieve your goals. To my friends, who filled my time outside of work with equal parts laughter and joy.

Finally, to my partner Ravina, who's kindness and uninterrupted belief in me has so often made the difference. You are the foundation on which the best of my achievements sit.

Abstract

Identifying the types of algorithms and applications that can be solved on today's quantum computers is one of the fundamental goals of modern quantum algorithms research. Underpinning this effort is the potential leap in human technology should it be shown that quantum computers can perform useful classically-intractable calculations in the absence of quantum error correction.

In this thesis we study the new paradigm of variational quantum algorithms (VQAs), designed specifically to solve optimisation problems on near-term hardware. We develop several novel algorithms spanning a range of application areas and provide estimates of the physical resources required to match classical methods. Thereafter, we attempt to run these algorithms on current quantum computers, probing their capabilities in the process. In this endeavour, we also develop our own algorithmic solution to device noise in the form of a new approach to approximate quantum circuit recompilation.

We begin with studying the feasibility of simulating the time evolution of quantum systems on current quantum computers, an important problem due to its classical hardness. We build an algorithm specifically focused on the electron-phonon Hamiltonian and subsequently realise the first demonstration of obtaining its dynamics on real quantum hardware. Subsequently, we test whether a similar result is possible for time-evolution based VQAs. For this we construct a near-term quantum computing approach to dynamical mean-field theory and find that whilst not possible currently, such an algorithm could be realistically evaluated on the next generation of quantum computers in the immediate future.

We then show how quantum circuits can be used as machine learning models and apply them to self-supervised learning, one of the most demanding tasks in deep learning. Through our experiments we observe a numerical advantage for the learning of visual representations using small-scale quantum neural networks over equivalently structured classical networks, a first step on the ladder towards general quantum advantage. Through this, we also highlight the potential of near-term quantum computing in problems with only empirically established classical complexity.

Table of Contents

1 Introduction **1**

1.1 Thesis Structure 10

1.2 Publications 13

1.3 Computational Resources 14

1.4 Funding 15

2 Incremental Structural Learning **16**

2.1 Introduction 16

2.2 ISL routine 20

2.2.1 Adding a layer 20

2.2.2 Optimising the layer 23

2.2.3 Finishing the layer 24

2.3 Conclusion 24

3 Recompilation-enhanced simulation of electron-phonon dynamics on IBM Quantum computers **28**

3.1 Introduction 28

3.2 Method 31

3.2.1 Model 31

3.2.2 Circuit implementation 32

3.2.3 Calculation of population 36

3.3 Results on simulated quantum hardware 38

3.3.1 With comparison to ED 38

3.3.1.1 Population with no phonon coupling 38

3.3.1.2 Weak electronic coupling 39

3.3.1.3 Strong electronic coupling 41

3.3.2 Simulator results of larger real world system 43
3.4 Real device results 44
3.5 Conclusion 49

4 Minimum Hardware Requirements for Hybrid Quantum-Classical DMFT 51
4.1 Introduction 51
4.2 Model 52
4.2.1 Hubbard model 52
4.2.2 DMFT 53
4.2.3 Two-site Hamiltonian-based DMFT 54
4.2.4 Hybrid quantum-classical DMFT routine 57
4.2.5 Noise model 61
4.3 Results 63
4.3.1 Fidelity requirement of original scheme 64
4.3.2 Circuit reduction using incremental structural learning . . 67
4.4 Conclusion 72

5 Quantum Self-Supervised Learning 75
5.1 Introduction 75
5.2 Method 78
5.2.1 Contrastive learning architecture 78
5.2.2 Quantum representation network 81
5.3 Results 84
5.3.1 Training 84
5.3.2 Linear probing 88
5.4 Classical width ablation 88
5.4.1 Quantum and classical results on the simulator 90
5.5 Quantum and classical results at different widths 91

5.6 Performance of alternative ansatz 92
5.6.1 Real device experiments 94
5.7 Conclusion 98
5.8 Code availability 101

6 Conclusions and Outlook **102**
6.1 Summary of Work 102
6.2 Outlook and Open Questions 107

Appendix A Appendix for Chapter 2 **148**
A.1 Figure 3.7 parameters 148

Appendix B Appendix for Chapter 3 **149**
B.1 Calculating the impurity Green's function with a reduced number of terms 149
B.2 Quantum circuit to obtain the impurity Green's function 150
B.3 Definition of single-qubit gates 152

List of Figures

1.1 Overview of a variational quantum algorithm. A quantum computer (blue) encodes an ansatz in the form of a circuit, which is evaluated for a specific trial solution given by parameters $\vec{\gamma}$. The result of this is measured and subsequently fed into a classical computer (green) running an optimiser, which suggests new parameters. The routine terminates when the convergence criteria is reached as defined by the optimiser. 5

2.1 In approximate quantum circuit recompilation, the problem of finding a set of gates $\hat{V}$ that approximates the action of another circuit $\hat{U}$ on an initial state can be considered as finding the inverse $\hat{V}^\dagger$ that recovers the initial state. As the form of the inverse is unknown, a parameterised ansatz $\hat{V}^\dagger(\boldsymbol{\theta})$ is optimised to try and achieve this. For an initial state $|0\rangle^{\otimes n}$, ISL avoids guessing a fixed structure of the ansatz and instead incrementally builds two-qubit layers with the aim of disentangling the circuit back to a product state. 19

2.2 A thinly-dressed CNOT gate is a CNOT gate surrounded by 4 single-qubit rotation gates $\hat{R}_i(\theta)$, where $i \in \{ x, y, z \}$ is the axis of rotation and θ is the angle. 22

3.1 Diagram of the model considered in this chapter. A lattice site i can be occupied by an electronic excitation with transition energy $\tilde{\epsilon}_i$, which can hop to connected sites with transition amplitude $\mathcal{V}_{ij}$. Each site is also connected to l phonon modes, approximated by quantum harmonic oscillators with frequency ω_{il}, which interact with the excitation via the electron-phonon coupling constant χ_{il}. . 33

3.2 Circuit diagram implementing the operator $\hat{XY} = \exp\left\{\frac{\mathcal{V}_{ij}\tau}{2\eta}(\hat{\sigma}_i^x\hat{\sigma}_j^x + \hat{\sigma}_i^y\hat{\sigma}_j^y)\right\}$. 34

3.3 Circuit diagram showing state initialisation and a single time evolution step. First, one of the site qubits are flipped to initialise the system with a single excitation, in this case site 0. The dashed box shows the application of H_{el}, with phase rotation gates to encode the electronic transition energies and $\hat{XY}$ gates encoding the electronic excitation hopping. Note that in this case, the sites are not fully connected and instead represent a model with a chain topology. The dotted boxes shows the application of H_{ph} to each oscillator, each of which consist of a $n_x = 2$ qubit register for a total of $d = 4$ energy levels. Afterwards, the interaction H_{ep} between each site and it's corresponding phonon mode is applied, the decomposition of which is too large to show. Finally, the site qubits are measured to obtain observables such as population. . . . 37

3.4 Population over time at sites 0 (blue) and 1 (orange), as calculated by exact diagonalisation (ED, dashed line) or quantum circuits run on the Qiskit Aer statevector simulator (SV, markers) with 6 (cross), 12 (square), 24 (circle) or 48 (triangle) Trotter steps. The Hamiltonian is simulated with parameters $V = 0.05\hbar\omega$, $\chi = 0.0\hbar\omega$. In this case there is no coupling to the phonon modes, so the system evolves under just $H = H_{\mathrm{el}}$. Since there are no non-commuting terms, the quantum circuit method incurs zero Trotter error and thus achieves equal accuracy to ED for any number of Trotter steps. Note that consequentially the dashed line is partially obscured by the symbols. 39

3.5 Left: population over time at sites 0 (blue) and 1 (orange), as calculated by exact diagonalisation (ED, dashed line) or quantum circuits run on the Qiskit Aer statevector simulator (SV, crossed marker). The Hamiltonian is simulated with parameters $V = 0.05\hbar\omega$, $\chi = 0.3\hbar\omega$ (upper panels), $\chi = 1.0\hbar\omega$ (lower panels) and $n_x = 1$. Moving across the panels left to right, for the quantum circuit approach, time evolution is discretised by an increasing number of Trotter steps η. Right: average energy level occupation of the oscillator coupled to site 0 during time evolution. The blue bar corresponds to the experiment shown on the left, where each phonon mode is represented by $n_x = 1$ qubit. The other bars represent subsequent experiments with increasing numbers of qubits per phonon mode, which allows simulation of up to $d = 2^{n_x}$ oscillator energy levels. These experiments were run with the number of Trotter steps required for convergence (e.g., $\eta = 48$ for $\chi = 0.3\hbar\omega$ and $\eta = 144$ for $\chi = 1.0\hbar\omega$). 41

3.6 Left: population over time at sites 0 (blue) and 1 (orange), as calculated by exact diagonalisation (ED, dashed line) or quantum circuits run on the statevector simulator (SV, crossed marker). The Hamiltonian is simulated with $V = 1.0\hbar\omega$, $\chi = 0.3\hbar\omega$ (upper panels), $\chi = 1.0\hbar\omega$ (lower panels) and $n_x = 1$ with increasing Trotter steps from left to right. Right: average energy level occupation of the oscillator coupled to site 0 during time evolution. . . 43

3.7 Population over time with non-uniform couplings and phonon frequencies, $\eta = 96$ and $n_x = 3$. All results were calculated using the quantum circuit implementation described in the main body of this work, run on the Qiskit Qasm simulator with 10,000 shots. The couplings $\vec{V}$, $\vec{\chi}$ and frequencies were randomly generated from the uniform interval [0.8, 1.2] and can be found in Appendix A.1. 45

3.8 Population over time as calculated by ED (dashed line) or quantum circuits executed on an IBM quantum computer directly (triangle marker) or first recompiled with ISL (crossed marker). All experiments are in the strong electronic coupling regime $V = 1.0\hbar\omega$. Left: system parameters $N = 2, \chi = 0.3\hbar\omega, \eta = 12$ and $n_x = 1$ run on *ibm_lagos*. Middle: same as the left panel, but with $N = 3$ sites, run on *ibmq_mumbai*. Right: $N = 2, \chi = 1.0\hbar\omega, \eta = 48$ and $n_x = 2$ run on *ibmq_mumbai*. The bottom panel shows the absolute difference Δ between experiments and the analytic solution for each site. 47

4.1 (a) The Hubbard model on an infinite dimensional Bethe lattice. Each site can be occupied by no electrons, a single electron or a pair of electrons of opposite spins with interaction energy U. Electrons can hop between adjacent lattice sites with energy t. (b) In dynamical mean-field theory (DMFT), we consider one site to be the impurity and approximate its interactions with the rest of the lattice as a mean-field. Two electrons occupying the impurity site have an interaction energy of U and electrons can be exchanged with the mean-field $\Delta(\omega)$. 54

4.2 (a) Hamiltonian-based DMFT replaces the mean field by a set of non-interacting bath sites. Electrons can occupy any lattice site, but can only move between the impurity (blue) and a bath site (grey). (b) In two-site DMFT, we use only a single bath site. In the half-filled case, the system dynamics are now described by two parameters, U and V, the on-site interaction and hybridization parameter respectively. 55

4.3 Diagram of hybrid quantum-classical DMFT. For given on-site interaction energy U, we iteratively discover the hybridization parameter V such that Eq. (4.5) is satisfied. For the first iteration, we start with a guess V_{guess}. We use a quantum computer to compute the impurity Green's function $G^R_{\text{imp}}(\tau)$, followed by a classical optimiser to suggest an improved hybridization V_{new}. The full loop is iterated until self-consistency is reached, such that $V_{\text{new}} = V$. . 58

4.4 Quantum circuit used to calculate the expectation value $\langle\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}_1^x\hat{\mathcal{U}}(\tau)\rangle$. The work qubits are first prepared into the ground state of the SIAM Hamiltonian using the sub-circuit $\hat{GS}$. They are then acted on by entangling gates with the ancilla qubit and the time evolution operator $\hat{\mathcal{U}}(\tau)$. The ancilla qubit itself undergoes single-qubit Hadamard $\hat{H}$ and bit-flip $\hat{X}$ gates. Repeated measurements of the ancilla in the $\hat{\sigma}_z$, $\hat{\sigma}_y$ bases build up the real and imaginary parts of the expectation value respectively. 61

4.5 CNOT gate infidelity for different noise model parameters. The qubit lifetime T is used to calculate the probability of thermal relaxation occurring. Additionally, a two-qubit depolarizing channel is applied with probability parameter λ. For this particular noise model, we set the CNOT gate time to be 300ns. Note that the color coding of the infidelity utilises a discretized logarithmic scale. 63

4.6 Quasiparticle weight Z as a function of interaction strength U and the Hubbard hopping amplitude in infinite dimensions t^*. For a given interaction strength, we iteratively obtain a self-consistent Z for 24, 36 and 48 Trotter steps and compare against the analytic solution. The inset focuses on the region near the critical value $U_c = 6.0t^*$. 65

4.7 Relative quasiparticle weight as a function of on-site interaction strength U for different two-qubit gate fidelities. Here Z_0 is the 24 Trotter step statevector simulator result shown in Fig. 4.6. For definition of error bars see text. 67

4.8 Laddering the recompiled solutions of the Green's function circuit at each time step allows the calculation of $\mathrm{i}G^R_{\mathrm{imp}}(\tau)$ for large τ without evaluating deep quantum circuits. A recompilation target $\hat{U}$ for the $\eta + 1$ Trotter steps Green's function circuit is created by adding one exact Trotter step to $\hat{RC}_\eta$, the recompiled circuit for η Trotter steps. A solution is constructed by trying to find a circuit $\hat{V}^\dagger = \hat{RC}^\dagger_{\eta+1}$ that approximately acts as the inverse of $\hat{U}$ but with fewer gates. The ansatz for $\hat{V}^\dagger$ is built up iteratively until the overlap between the output state and the input state $|\psi_0\rangle$ is sufficiently large. 70

4.9 Relative quasiparticle weight as a function of on-site interaction strength U for different two-qubit gate fidelities. Using incremental structural learning (ISL), a quantum computer with two-qubit gate fidelity of 99.9% can produce convergent DMFT results with perfect agreement to the noiseless solution. 72

5.1 Schematic of the overall neural network architecture and contrastive training method. For each input image $\vec{x}_i$, a pair of random augmentations are generated and applied to form a positive pair $\vec{x}^1_i$, $\vec{x}^2_i$. These are transformed by the encoder network, consisting of classical convolutional layers and a quantum or classical representation network, into representation vectors $\vec{y}^1_i$, $\vec{y}^2_i$. The projection head subsequently maps the representations to the vectors $\vec{z}^1_i$, $\vec{z}^2_i$, such that contrastive loss can be applied without inducing loss of information on the encoder. 80

5.2 General structure of a QNN. An input vector $\vec{v}$ is encoded into the qubits by a data loading unitary $\hat{D}(\vec{v})$. The variational ansatz consists of layers $\{\hat{U}_1(\vec{\theta}_1), \hat{U}_2(\vec{\theta}_2), \dots\}$ and is parameterised by trainable parameters $\{\vec{\theta}_1, \vec{\theta}_2, \dots\}$. The output of the QNN is taken as the average of repeated measurements in the $\hat{\sigma}_z$ basis. 82

5.3 Variational ansatz used in this chapter. Each layer consists of a single qubit $\hat{R}_y$ rotation on each qubit, followed by controlled $\hat{R}_x$ rotations, connecting the qubits in a ring topology. Every rotation gate is parameterised by a different variational parameter. 83

5.4 Contrastive learning with a quantum representation network. After each batch of 256 images the loss function (main graph) is recorded, alongside the average Hilbert Schmidt distance between positive and negative pairs $\bar{D}_{\text{HS}}$, the average positive pair clustering $\text{tr}\overline{(\rho^2)}$, the average clustering of all negative pairs $\text{tr}\overline{(\sigma^2)}$ and the ensemble inter-cluster overlap $\text{tr}\overline{(\rho\sigma)}$ (insets). 85

5.5 Classification accuracy achieved in linear probing experiments by classical representation networks with varying network widths at checkpoints across self-supervised training. The markers show the average of three independently trained self-supervised models and linear probe experiments, whilst the error bars show one standard deviation. 89

5.6 Classification accuracy achieved in linear probing experiments using the encoder at checkpoints across self-supervised training. Comparison between models trained with a classical representation network (blue), quantum representation network evaluated on a statevector simulator (orange) and quantum representation network evaluated on a sampling-based simulator with 100 shots (purple). The markers show the average of six independently trained models, whilst the error bars show one standard deviation. 90

5.7 Classification accuracy achieved in linear probing experiments for quantum and classical models with representation network widths 2 (upper left), 4 (upper right), 6 (lower left) and 8 (lower right). Within each graph, the circle (crossed) markers represent models trained with quantum (classical) representation networks. The markers show the average accuracy of three independently trained models, with error bars of one standard deviation, except for the $W = 8$ case which displays the same data as Fig. 5.6. 93

5.8 Alternative variational ansatz. Each layer consists of a single qubit $\hat{R}_y$ rotation on each qubit, followed by CNOT gates connecting all qubits to each other. After all layers have been applied, a final set of $\hat{R}_y$ rotations are applied. Every rotation gate is parameterised by a different variational parameter. 94

5.9 Classification accuracy achieved in linear probing experiments using the encoder at checkpoints across self-supervised training. Comparison between models trained with a classical representation network (blue), quantum representation network with the ring ansatz (orange) and quantum representation network with the all-to-all ansatz (green). All quantum circuits were evaluated on a statevector simulator. The markers show the average of six independently trained models, whilst the error bars show one standard deviation. 95

5.10 Confusion matrix from classifying 900 images using the best performing (a) classical model evaluated on a classical computer (b) quantum model evaluated on a real quantum computer with 100 shots per circuit. For a given true label (rows) and predicted label (columns), the number in each box shows the total number of times that prediction was made. 97

List of Tables

4.1 Parameters used to emulate different NISQ devices with our noise model. For a given qubit lifetime T and depolarizing channel probability λ, the corresponding gate fidelities $\mathcal{F}$ can be obtained. The single-qubit gates, $\hat{U}_1$, $\hat{U}_2$ and $\hat{U}_3$ are defined in Appendix B.3. 72

List of Abbreviations

AQCR . . . . . Approximate quantum circuit recompilation

DMFT . . . . . Dynamical mean-field theory

ISL . . . . . Incremental structural learning

PQC . . . . . Parameterised quantum circuit

QML . . . . . Quantum machine learning

QNN . . . . . Quantum neural network

VQA . . . . . Variational quantum algorithm

VQE . . . . . Variational quantum eigensolver

1 | Introduction

The past 40 years has seen the remarkable succession of quantum computing from a theoretical concept [1, 2] to a practical endeavour. The fruits of this labour are the handful of devices that exist today, spanning many different types of qubit architectures. Collectively, these devices are of the size and maturity that they are considered a new generation, distinct from the preliminary frontiers set by NMR quantum computing at the turn of the millennium. Ion trap quantum computers [3, 4] form one of the two pillars of current hardware, the size of which has increased from 2 qubits in 1995 [5] to 32 today [6], furthered yet by 53-qubit simulators [7]. More dramatic has been the rapid scaling of superconducting quantum computers [8, 9], from 9 qubits in 2016 [10] to 53 qubits in 2019 [11] to a staggering current record of 127 qubits [12]- albeit with diminishing qubit connectivity. During this growth, a transition occurred whereby the number of complex numbers required to fully describe the state space of the qubits exceeded the memory capacity of the largest classical supercomputers. This led to the landmark demonstration of quantum advantage for low fidelity sampling of random quantum circuits [11], indicating how far quantum hardware has come. Since then, many counter-claims have been presented with classical simulations of the quantum advantage experiment [13, 14], followed by their own rebuttals with larger quantum experiments [15, 16]. Thus, whilst it remains to be seen if the line of demonstrable quantum advantage has truly been crossed, it is certainly an exciting moment for quantum computing.

The swift increase in the size of current-generation quantum computers has captured significant scientific and media attention, leading to an arms race for the largest number of qubits. However, for almost all applications of quantum computing, the lifetimes of the qubits and error rates of the logic gates are just as

important. In this area progress has also been strong, with ion trap quantum computers currently able to execute two-qubit gates with a fidelity of 0.9994 [17] on a 2-qubit device or 0.997 on a 20-qubit device [18]. Superconducting qubit hardware has also seen strong progress in recent years in terms of error rates. In this case, rather than demonstrations of extremely high fidelity, superconducting quantum computers have seen error rates remain relatively steady whilst the device sizes have scaled up. This includes two-qubit fidelities of 0.9978 on a 10-qubit processor [19], 0.989 two-qubit fidelity on a 27-qubit processor [9, 20] and 0.994 on a 53-qubit processor [11].

Despite this progress, the field of quantum computing hardware still faces many of the same original challenges that have plagued practical usage for decades. Most notably, whilst fidelities are improving, the ratio of qubit lifetime to gate times are still small enough that decoherence of the qubits is a major source of noise during computation. Furthermore, imperfections in the control processes that act as quantum logic gates on the qubits create an additional layer of complex noise, resulting in effects such as crosstalk [21]. The overall noise produced by these effects is today prohibitive to the accurate execution of even small circuits. For example, despite being at a size for which storing its state is classicaly intractable, the previously referenced 127 qubit *ibmq_washington* device has a quantum volume [22, 23] of 32. This means that on average, only a circuit of width and depth 5 could be run without a two-qubit gate error occurring. It is for this reason that the current generation of quantum computers are often referred to as *noisy* intermediate-scale quantum (NISQ) computers [24]. An important consequence of this noise is that the foundational algorithms with proven quantum speedups, such as Shor's [25] and Grover's [26], cannot be executed. This calls into question whether current quantum computers can fulfil the potential of the technology to make a serious impact at a societal level.

Whilst the hardware community continue to work towards lower and lower error rates, learning to try and diminish the effect of this noise has become a major focus across the field of quantum algorithms. At first glance it may seem odd that quantum computing with noise should be a problem, given the many well established methods of quantum error correction (QEC) [27, 28]. Yet even though current quantum hardware lies within the threshold value in which error correcting schemes such as the surface code could be applied [29], the number of noisy qubits required to make a single logical qubit is expected to be between 10^3 and 10^4 [30]. Extended to Shor's algorithm, it is thought that breaking current encryption standards would require 20 million of today's noisy qubits [31]. Thus, large-scale fault tolerant quantum computing is not an achievable goal on current hardware. However, many recent algorithmic developments in QEC [32, 33], alongside small-scale experimental demonstrations of error correction [34, 35, 36], hold great promise for the future.

With QEC not a viable option right now, another ongoing effort in the quantum algorithm community to tackle the noise problem is quantum error *mitigation*. Rather than correct errors entirely, these methods work towards the less strict and rather broad goal of reducing noise as much as possible [37]. The complexity of such schemes vary greatly. At the simplest end, measurement calibration has proved a remarkably useful tool to account for coherent readout errors [38]. Here a map is created between preparation of all computational basis states and the observed noisy outcome, and subsequently the inverse of the map is applied to noisy experiments to correct for readout errors. More recently, zero-noise extrapolation [39] is another technique that has seen widespread success, whereby additional operations are added to a circuit to purposefully amplify the effect of noise. Crucially, operations are chosen which apply the identity (e.g., the insertion of pairs of CNOT gates [40] or replacing a unitary gate $\hat{U}$ by $\hat{U}\hat{U}^\dagger\hat{U}$ [41]), such that the

theoretical output of the circuit is not changed. After recording the value of an observable with several rounds of additional gates, an extrapolation is then applied backwards towards the "zero-noise" solution, which subsequently cancels powers of the noise perturbations. This and other error mitigation techniques remain a continuously advancing area of research, including hardware specific methods such as noise-aware transpilation [42] and hardware agnostic methods such as learning quantum noise with machine learning models [43, 44].

Unfortunately, error mitigation methods do not circumvent the noise problem entirely. Despite their success, such methods are not yet sufficiently powerful to allow the execution of traditional quantum algorithms on current hardware. Furthermore, it has been recently shown that for a wide class of error mitigation protocols, including zero-noise extrapolation, the number of samples of the circuit required to achieve a desired accuracy scales exponentially with the circuit depth [45, 46, 47]. This has led to a third and perhaps the most significant community effort to solve the noise problem: developing noise-tolerant algorithms. This idea has sparked one of the largest upheavals in quantum algorithms since its inception, with the invention and popularisation of variational quantum algorithms (VQAs) [48].

VQAs are a class of algorithms which seek to solve the optimisation problem of minimising or maximising a cost function $C(\vec{\gamma})$ by evaluating trial solutions of the cost function using a quantum computer. Such an idea was first proposed in 2014 [49], whereby the ground state of a Hamiltonian $\hat{H}$ is found by variationally solving to minimise the Rayleigh-Ritz quotient $\frac{\langle\psi(\vec{\gamma})|\hat{H}|\psi(\vec{\gamma})\rangle}{\langle\psi(\vec{\gamma})|\psi(\vec{\gamma})\rangle}$, in a routine that has become known as the variational quantum eigensolver (VQE). Since then, VQAs have been applied to a vast range of applications and domains. Within quantum chemistry, VQAs have also been used to calculate excited states [50] and the dynamics of various physically important Hamiltonians [51]. In the context of ma-

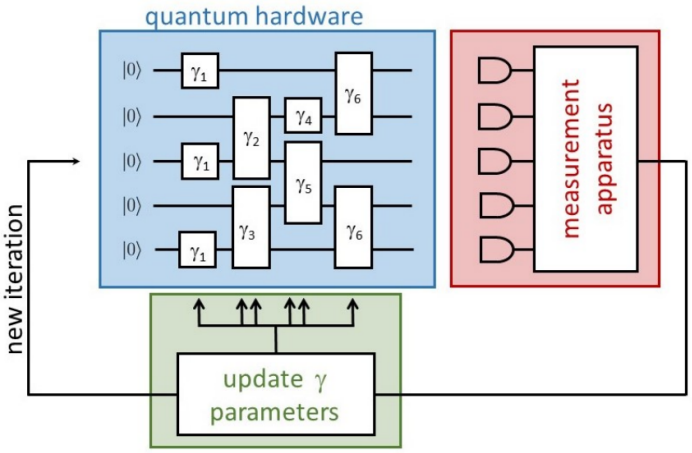


Figure 1.1: Overview of a variational quantum algorithm. A quantum computer (blue) encodes an ansatz in the form of a circuit, which is evaluated for a specific trial solution given by parameters $\vec{\gamma}$. The result of this is measured and subsequently fed into a classical computer (green) running an optimiser, which suggests new parameters. The routine terminates when the convergence criteria is reached as defined by the optimiser. Figure originally from [63].

chine learning, VQAs are often referred to as quantum kernels [52] or quantum neural networks (QNNs) [53], and have been trained via gradient methods to solve many problems in classification [54] and clustering [55]. VQAs have also been used to solve optimisation problems in quantum circuit compilation [56], combinatorial optimisation [57] and areas of mathematics including finding the solution to linear [58] and non-linear equations [59]. Progress in any of these problems could have an enormous impact on applications including artificial intelligence, simulating turbulence [60], protein folding [61] and computer chip design [62].

Figure 1.1 shows the overall structure of a VQA. The first important component is the quantum circuit shown in the blue box, which is the encoding of the parameters $\vec{\gamma}$ into the rotation angles of quantum logic gates. This is often referred to as the quantum ansatz, since its evaluation represents the trialing of specific values as the solution to the optimisation problem. For example, in VQE the ansatz pro-

duces the trial state vector $|\psi(\vec{\gamma})\rangle$. The structure of the ansatz and choice of logic gates varies greatly depending on the algorithm. In some cases, such as the algorithms put forward in chapters 3 and 4 of this thesis, the ansatz is entirely dictated by the cost function. In other cases, such as the quantum machine learning algorithm in chapter 5, there is total flexibility in the choice of ansatz and performance can depend greatly on the chosen ansatz structure. For such algorithms, methods to determine the optimal ansatz remains an important and open area of research. Notable recent developments include choosing ansatz structures inspired by the physical properties and symmetries of the problem, leading to significant performance improvements as seen in the ADAPT-VQE algorithm [64].

After the ansatz, the next component of a VQA is the measurements of the circuit. As with all quantum algorithms, this represents the interface between the quantum and classical realms and importantly, only a classical amount of information can be extracted from the circuit as per Holevo's bound. Thus, it is important that the cost function does not rely on an explicit representation of the statevector, which would require tomography schemes to obtain. For some VQAs, the order, number and basis of measurements to take are prescribed. For example, in VQE, obtaining the numerator of the cost function $\langle\psi(\vec{\gamma})|\,\hat{H}\,|\psi(\vec{\gamma})\rangle$ requires a decomposition of the Hamiltonian into Pauli operators, the expectation values of which are obtained through the repeated averaging of specific measurements. In other VQAs the measurement routine is flexible, and understanding how the choice of measurements can affect both the scalability and performance of VQAs is an interesting and emerging area of study [65, 66].

The final component of a VQA is the optimiser. Based on the evaluation of the cost function, the optimiser determines what the value of $\vec{\gamma}$ should be for the next iteration. There is no one best optimiser for every variational algorithm, but some examples include the well known gradient descent method [67], or the Nelder-

Mead algorithm [68] as used in VQE.

In the past decade, VQAs have dominated the landscape of quantum algorithms due to their increased noise tolerance over traditional algorithms, making them especially compatible with the noisy quantum computers that are available today. The noise tolerance of VQAs comes from three different effects. Firstly, when a quantum circuit is evaluated as part of a VQA, it is to obtain a single value of the cost function $C(\vec{\gamma})$. This is just one iteration of the wider algorithm, which could take many more iterations to converge to a minima or maxima. As such, the success of the overall algorithm is not reliant on the perfect execution of any given circuit, but rather that on average the circuits are executed accurately enough to allow for convergence. Whilst this relation between noise and convergence is highly nontrivial, it still allows the solution to be obtained in the presence of large but occasional errors. By contrast, the amount of noise that can be tolerated in Grover's algorithm vanishes exponentially with the number of qubits [69].

The second beneficial feature for mitigating noise is that for some applications, obtaining a solution with VQAs requires shorter quantum circuits than non-variational alternatives [70, 71]. Here the length of the circuit is more precisely defined as its depth, that is, the total number of non-simultaneous operations from start to finish. Importantly, a shallower circuit in turn can accommodate qubits with shorter coherence times. This is in fact the original motivation of the VQE, which executes much shallower circuits than the alternative quantum phase estimation algorithm, thus generating less noise when run on near-term noisy quantum computers. This advantage is not general to all VQAs, since not all have a known alternative quantum approach, however a a similar benefit can be realised more generally in that VQAs with flexible ansatzes can choose a structure with depth depending on the capabilities of the available hardware [72].

Finally, VQAs are well suited to running on noisy hardware due to their ability to factor out certain types of noise during the optimisation process [73, 74, 10]. For coherent (i.e., structured) noise, the mechanism behind this is simple to understand, for example the offsetting of systematic error with a constant valued bias added to the updated $\vec{\gamma}$ values. VQAs have also been shown to be resilient to incoherent random noise, with preliminary studies suggesting that this may be achieved through the inclusion of redundant parameters [75]. Overall however, understanding mechanisms by which VQAs can learn noise remains an open problem.

With a technical description of VQAs laid out, we turn our attention to the goals of this thesis. Whilst VQAs have the potential to be a powerful tool, many challenges remain if they are going to be used to solve practically relevant problems on current quantum computers. Firstly, analytically determining the properties of VQAs, such as the extent of their noise tolerance and the potential for quantum advantage, remains a difficult task. This is due to the complex interplay between random errors generated during evaluation of the cost function and its knock-on effect on the updated parameters provided by the classical optimiser, resulting in a stochastic optimisation problem. Subsequently, repeatedly running the same VQA on a noisy quantum computer may produce drastically different paths to convergence and even lead to divergent solutions depending on the cost landscape. Secondly, despite following the same blueprint, insights gained on one VQA may not apply to others. This is mainly due to the vastly different cost functions seen across different application areas. For example, the potential efficiency of generating trial wave functions in quantum chemistry based VQAs does not inform the suitability of using quantum circuits to represent the price distribution of financial assets [76].

For these reasons, the first goal of this thesis is to develop new variational quantum

algorithms for important optimisation problems that have not yet been studied. This goal is not for the sake of novelty itself, but rather follows a strong consensus within the field that a quantum speedup for optimisation problems cannot be ruled out yet and needs be explored [77]. The research contained within this thesis makes a important contribution towards this goal, developing novel VQAs across a range of domains and analysing their performance when using cost functions of known or unknown classical hardness. Furthermore, the algorithms that we study are of significant relevance to the larger scientific community, with applications in chapters 3 and 4 that could aid the development of next-generation materials and findings in chapter 5 that could enhance the already ubiquitous use of machine learning in scientific problem solving. More details on the structure of this thesis can be found in section 1.1.

The second goal of this thesis is to investigate how successfully these VQAs can be run on real current-generation quantum computers. This is a theme that runs through all work contained in this thesis, guiding both the size and complexity of the models we investigate. By exploring whether our algorithms can be run on current quantum hardware, we are able to quantify the hardware's capability to solve valuable problems. Our work is particularly timely, since alongside the recent growth of quantum computers has been a significant commercial effort to integrate them into established cloud services. This has left us uniquely placed in time to be the the first researchers for whom experiments can be conducted on real hardware without the barriers of building or obtaining our own. Through this effort, we are able to help answer the biggest questions in quantum computing, including determining how useful current devices really are. Where our algorithms are not able to run on current devices, we use noise modelling to determine the gap between available and required hardware. We hope that these results will in turn move the field of quantum hardware forward, providing clear feedback and

tangible goals as to where the next steps should be.

Earlier we laid out the three methods by which the quantum computing community is tackling noise: QEC, quantum error mitigation and noise tolerant algorithms. By reading this thesis, readers will gain a detailed understanding of the latter of these, including how to develop variational quantum algorithms and construct representations of the cost function and system Hamiltonians as a variational ansatz. Furthermore, this thesis will provide readers an insight into the quirks and intricacies of running these algorithms on current quantum hardware, including techniques that can be used to try and obtain the best results possible.

1.1 Thesis Structure

The structure of this thesis is as follows. In chapter 2, we start with a preliminary methods chapter introducing approximate quantum circuit recompilation (AQCR) and its use as a tool for error mitigation. This is an important part of achieving our second goal, running VQAs on real quantum computers, and the tools developed here are applied across many of the subsequent research chapters. Specifically we introduce incremental structural learning (ISL), our novel approach to AQCR where the trial solution is built and optimised incrementally, guided by the entanglement properties of the target circuit. The ISL algorithm was created out of a necessity during this DPhil to maximise the performance of the quantum computers used, and as such it constitutes an emergent contribution of this thesis to the field of quantum error mitigation. Its placement at the front of the thesis will allow the reader to gain an understanding of how we approach quantum error mitigation across the subsequent chapters of this thesis.

In chapter 3, we introduce a novel quantum algorithm to obtain the dynamics of electronic excitation transfer across systems evolving under the electron-phonon

Hamiltonian. Solving this problem has important applications in understanding natural light harvesting systems and by extension, designing better artificial equivalents. Whilst our algorithm is not directly a VQA, simulating the time evolution of quantum systems is an important application of quantum computing, since it exhibits one of the few known exponential quantum speedups. In this chapter, the primary goal is to establish the viability of time evolution simulation on near-term devices and by extension, lay the foundation for time-evolution based VQAs. To achieve this, we first consider the absolute resource cost for gate-based quantum simulation of small electron-phonon systems as dictated by the number of Trotter steps and bosonic energy levels necessary for the convergence of dynamics. We then apply these findings to perform experiments on IBM quantum hardware for both the weak and strong electron-phonon coupling regimes. We find that the noise generated by the device is prohibitive to obtaining accurate results, even for small systems with simple dynamics. To remedy this we apply the previously introduced ISL recompilation algorithm to find shallow approximate equivalents of the time evolution circuits, which we find dramatically reduces the number of gates required and thus mitigates a significant source of noise. Executing the recompiled circuits on a real quantum computer, we subsequently achieve results that match exact diagonalisation in both the strong and weak coupling regimes. Excitingly, our results mark the first ever accurate simulation of electron-phonon dynamics on real quantum hardware, highlighting the novelty of AQCR as an error mitigation tool.

In chapter 4 we extend upon our previous results to investigate whether a time-evolution based VQA could successfully run on current quantum computers. For this we propose hybrid quantum-classical dynamical mean-field theory (DMFT) - a variational approach to mapping the Hubbard model to an impurity problem. DMFT is itself an important problem, since a scalable approach to simulating

the Hubbard model may allow for the design of high-temperature superconductors and next generation ultra-fast electronics. First we analyse our algorithm's tolerance for fundamental error, a consequence of both the Suzuki-Trotter decomposition and the statistics of measuring qubits. After, we test its performance on emulated noisy quantum hardware, trialing different qubit lifetimes and logic gate error rates to establish the minimum hardware requirements in which a self-consistent solution can still be reached. We find this threshold to be two-qubit gate fidelities of 99.999%, several orders of magnitude higher than the current state of the art. Subsequently, we consider whether AQCR, this time applied repeatedly for each iteration of the VQA, could help improve improve upon this and increase the noise tolerance of the algorithm. Applying ladder-ISL, a more scalable extension of our previous recompilation algorithm, we find a factor of 100 reduction in the required gate fidelities to accurately run two-site DMFT, although the requirements remain just beyond the capabilities of current quantum computers at the time of the experiment. Promisingly, devices with sufficiently low errors have since been built and running the first ever quantum DMFT routine on a real quantum computer is now an ongoing collaboration with experimental groups.

In chapter 5 we turn our attention to VQAs as applied to machine learning. Here we build upon recent works demonstrating the possibility of a quantum advantage using QNNs for supervised learning of artificial data and challenge whether such a result is possible with real world visual data instead. Importantly, we reason that the extremely high classical performance of *supervised* learning on image datasets makes it an unlikely candidate for quantum advantage, which would require the underlying distribution to be classically hard to learn. Instead, through collaboration with computer vision and machine learning experts, we consider how a QNN might be integrated into *self-supervised* learning - a cutting edge deep learning technique. Self-supervised learning is particularly important as a scalable solu-

tion to the dramatically increasing size of real-world data sets without relying on human annotation. First, we build a code library that allows the popular machine learning software package PyTorch to train quantum and classical neural networks together as one architecture. Once completed, we then test how the inclusion of a QNN into the encoder affects the quality of the learnt image representations. Interestingly, when these representations are subsequently used for downstream classification, we observe a numerical advantage for an architecture using QNNs compared to equivalently structured classical networks, which holds even when the quantum circuits are sampled with only 100 shots. Furthermore, we then apply our best quantum model to classify unseen images on real quantum hardware and find that with the help of ISL, current noisy devices can already achieve equal accuracy to the equivalent classical model on downstream tasks. Whilst our results are very much a proof of concept, we hope they galvanise the community to consider classically hard problems in machine learning as an area of potential quantum advantage for VQAs.

Finally, in the conclusion we summarise the results of this thesis and explore both the ongoing and exciting future research directions spawned from it.

1.2 Publications

The research undertaken during this DPhil has led to the publication of several articles, which contain some of the results shown in this thesis. They are:

Jaderberg, Ben, Abhishek Agarwal, Karsten Leonhardt, Martin Kiffner, and Dieter Jaksch. "Minimum hardware requirements for hybrid quantum–classical DMFT." Quantum Science and Technology 5, no. 3 (2020): 034015.

Jaderberg, Ben, Lewis W. Anderson, Weidi Xie, Samuel Albanie, Martin Kiffner,

and Dieter Jaksch. "Quantum self-supervised learning." Quantum Science and Technology 7, no. 3 (2022): 035005.

Jaderberg, Ben, Alexander Eisfeld, Dieter Jaksch, and Sarah Mostame. "Recompilation-enhanced simulation of electron-phonon dynamics on IBM Quantum computers." arXiv preprint arXiv:2202.08270 (2022). Under review: New Journal of Physics.

1.3 Computational Resources

Across the experiments in this thesis, quantum hardware is accessed via the IBM quantum experience [20]. This is a scheme which since 2016 has allowed the general public open access their range of superconducting quantum computers, as well as access for institutional partners to larger and higher quality devices [78]. During the course of this DPhil, the University of Oxford was one such institutional partner of IBM, for which I am extremely grateful for.

For experiments in this thesis not on real quantum hardware, an emulated quantum computer was used instead. Specifically, the Qiskit `Statevector` and `Qasm` simulators were used [79]. The former simulates quantum circuits through the explicit vector and matrix representations of the quantum states and logic gates respectively, allowing for entirely noiseless evaluation. The latter computes the distribution generated by the circuit, replicating the more realistic process of sampling shots of the final state and allowing the introduction of emulated gate noise. In both cases, large experiments were run using the University of Oxford Advanced Research Computing (ARC) facility [80].

1.4 Funding

This DPhil was funded by the EPSRC National Quantum Technology Hub in Networked Quantum Information Technology (EP/M013243/1) and the EPSRC Hub in Quantum Computing and Simulation (EP/T001062/1).

2 | Incremental Structural Learning

2.1 Introduction

As introduced in chapter 1, today's quantum computers make up what is known as the NISQ era, predominantly characterised by high levels of noise that occurs randomly and destructively at all stages of computation. There are two key processes that can be considered as generating this noise. Firstly, the qubits themselves have finite lifetimes for which they can be held in coherent quantum states, eventually decohering to classical states through a process of exponential decay. This is not a fundamental property of the qubits, but rather a very practical reflection of the underlying nature of quantum mechanics, in which any interaction with the outside world can lead to a measurement of the quantum state that irreversibly changes it. Isolating qubits from their environment is a magnificently hard challenge, which includes shielding against an external bombardment of thermal radiation [81], sound vibrations [82] and even cosmic rays [83]. For the most popular qubit architectures, ongoing solutions to this include operating at cryogenic temperatures on the scale of mK [84] or within ultra-high vacuum conditions [85].

Secondly, the scientific instruments used to manipulate the qubits have a finite precision, such that the application of each logic gate results in a probability that the physical computation will diverge from the algorithmic expectation. For example, implementing a bit-flip gate on a superconducting qubit requires a pulse from the microwave drive line at the Rabi oscillation frequency. Errors in the precise amplitude required will lead to a different rotation on the Bloch sphere than desired, whilst limitations in the speed of the pulse may result in transitions of the qubit to excited states beyond the defined computational space [86]. Furthermore, errors in the frequency of the pulse caused by operational tolerance of the mi-

crowave source can lead to unintentional coupling between qubits [87]. Overall, a combination of qubit decay and imprecise qubit control can be approximated by a noise model in which there is a fixed probability of an error occurring for each gate executed, given equal gate times. Despite its simplicity, this model has been well substantiated at large sizes [11], including the implied scaling that the total computational error grows exponentially with the number of gates in the circuit.

Whilst this scaling cannot be fully removed until the physical realisation of QEC, the low qubit lifetimes and large gate infidelities of current quantum computers make the accurate execution of deep circuits particularly problematic. Therefore, given a computation which we wish to run on a quantum computer, it is vital that we implement it in as few quantum logic gates as possible. Let us consider the problem more formally. Suppose that our quantum algorithm requires the action of a unitary operation $\hat{U}$ on the qubits. Quantum circuit compilation describes the process of finding an implementation of $\hat{U}$ on a quantum computer using a series of m gate operations. On the other hand, quantum circuit *re*compilation is the process of trying to improve an existing circuit by replacing it with a circuit that generates the same unitary but takes less time to run. Note that in theory the distinction between the two can be difficult, since ideally the circuit compilation routine produces an optimally minimal implementation - yet practically speaking the composition of quantum algorithms often produces redundancies that are not immediately visible. Generally, without knowledge of specific gate times, quantum circuit recompilation can be reformulated as trying to find a circuit which implements $\hat{U}$ in $o < m$ gate operations. Solutions to this range from duplicate gate cancellation [88] to two-qubit block re-synthesis [9] involving the KAK decomposition [89]. Significantly, since the fidelity of the circuit evaluated on noisy quantum hardware scales exponentially with the number of gates, the recompiled implementation with fewer gates generates less noise. As such the process of

quantum circuit recompilation can be viewed as an error mitigation technique.

Despite the success of quantum circuit recompilation methods, there remains key limitations in its applicability to quantum computing in the NISQ era. Notably, as is demonstrated in the later chapters of this thesis, the unitaries that we would like to implement in VQAs are often too deep to be evaluated on current hardware, even in their theoretically minimal form. This motivates another consideration. Suppose that the evaluation of the minimal implementation of $\hat{U}$ on a noisy quantum computer produces an errored final state $|\tilde{\psi}_f\rangle$ with an overlap to the true final state $|\psi_f\rangle$ of $|\langle\tilde{\psi}_f|\psi_f\rangle| = 0.6$. In this case, is it better to instead evaluate an alternative circuit, which in the noiseless case produces the state $|\phi_f\rangle$ which is only approximately correct such that $|\langle\phi_f|\psi_f\rangle| < 1$, yet is implemented in so few gates that it's evaluation on noisy hardware produces a state $|\tilde{\phi}_f\rangle$ with higher overlap to the true solution $|\langle\tilde{\phi}_f|\psi_f\rangle| > 0.6$? In other words, when our computation is in the presence of heavy noise, do we really need to evaluate the exact decomposition of $\hat{U}$?

This is the question posed by approximate quantum circuit recompilation (AQCR). Here, rather than find an exact alternative implementation of $\hat{U}$, the goal is to find a shallower quantum circuit $\hat{V}$ which has approximately the same action on some initial state $\hat{U}|\psi\rangle \approx \hat{V}|\psi\rangle$. Reformulating this as $\langle\psi|\hat{V}^\dagger\hat{U}|\psi\rangle \approx 1$, we notice that the problem can be viewed as finding the set of gates $\hat{V}^\dagger$ that inverses the action of $\hat{U}$ as measured by the overlap between the initial and final state, to within some desired accuracy. Once the inverse is found, each gate that makes up $\hat{V}^\dagger$ can be inverted individually to produce our desired approximately equivalent circuit $\hat{V}$. This process of trying to find the inverse can itself be recognised as an energy minimisation problem, leading to the popular approach to solve it as a variational quantum algorithm [90, 56, 91]. However, like many variational algorithms, the best solution often requires finding the optimal ansatz through

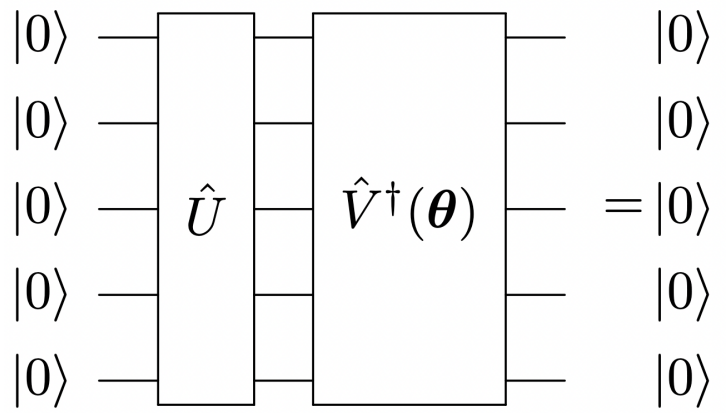


Figure 2.1: In approximate quantum circuit recompilation, the problem of finding a set of gates $\hat{V}$ that approximates the action of another circuit $\hat{U}$ on an initial state can be considered as finding the inverse $\hat{V}^\dagger$ that recovers the initial state. As the form of the inverse is unknown, a parameterised ansatz $\hat{V}^\dagger(\boldsymbol{\theta})$ is optimised to try and achieve this. For an initial state $|0\rangle^{\otimes n}$, ISL avoids guessing a fixed structure of the ansatz and instead incrementally builds two-qubit layers with the aim of disentangling the circuit back to a product state.

a potentially lengthy trial-and-error process. Furthermore, what constitutes the best ansatz may vary dramatically for different inputs, such that training multiple ansatzes for every circuit to be recompiled dramatically increases the complexity of general AQCR. Attempts to use a single problem-agnostic ansatz for multiple inputs would require the ansatz to have a random structure, leading to trainability issues imposed by barren plateaus [92]. These limitations combined are what we consider to be *the ansatz problem* for AQCR.

Our novel contribution to AQCR begins with the consideration that for all quantum algorithms, the convention is for the qubits to begin in the initial state $|\psi_0\rangle = |0\rangle^{\otimes n}$. Here the problem can then be expressed as finding the set of gates that maps the original circuit back to $|\psi_0\rangle$, a particular formulation shown in Fig. 2.1. We reiterate here that the goal of AQCR is to find a recompiled solution for only one initial quantum state, satisfying $\hat{U}|\psi_0\rangle \approx \hat{V}|\psi_0\rangle$ and not the more general condi-

tion $\hat{U} \approx \hat{V}$. One notable feature of our choice of initial state, and thus target final state of Fig. 2.1, is that it is a product state, with no pairwise entanglement. Thus instead of using a random ansatz for $\hat{V}^\dagger$, our solution should be physically inspired, specifically, one that works to disentangle the state $\hat{U}\left|\psi_0\right\rangle$. This represents the structural component of our incremental structural learning (ISL) algorithm.

The second novel component of ISL is the incremental construction of the ansatz. Instead of using an ansatz with a fixed layout, we build its structure layer-by-layer, evaluating the cost function

$$C = 1 - \left|\left\langle\psi_0\right| \hat{V}^\dagger \hat{U} \left|\psi_0\right\rangle\right|^2 \tag{2.1}$$

for each layer added. This approach not only offers a greater range of potential solutions, but also reduces the likelihood of a redundantly deep ansatz wasting resources or an inadequately shallow ansatz failing to to find the optimal solution. Importantly, our application of incrementally built ansatzes to AQCR follows notable successes using similar schemes to solve other VQAs [64], whilst also avoiding the issue of barren plateaus.

2.2 ISL routine

2.2.1 Adding a layer

We describe the process of ISL through the routine of adding a single layer. Suppose that after the application of $n-1$ layers, we have our current best guess of the inverse $\hat{V}^\dagger = \hat{V}_{n-1}^\dagger ... \hat{V}_1^\dagger$, but evaluating Eq. (2.1) results in a cost greater than the required threshold $C > C_t$. Subsequently, we continue the recompilation and add

the nth layer. In ISL, the structure of every layer is a thinly-dressed CNOT gate as shown in Fig. 2.2. This choice of two-qubit block is based on the regular dressed CNOT gate, used successfully in previous works as an ansatz layer specifically for the problem of variational quantum compiling [75]. Here we describe our layer as *thinly* dressed because the single-qubit gates are restricted to a parameterised rotation along one axis, as opposed to the combination of three parameterised rotations along all three axes in the original scheme. This was chosen in order to make our layers amenable to the specific optimiser described in section 2.2.2. Interestingly, one layer of the (thinly) dressed CNOT gate does not contain the 15 independent parameters required to span the special unitary group SU(4) which contains all two-qubit interactions [93]. However, this is a feature shared amongst the overwhelming majority of ansatzes used in VQAs and has not been shown to hamper practical performance. Moreover, the non-universality of our ansatz layer has the pragmatic benefit of requiring fewer gates to implement, making it more efficient to implement on real quantum computers.

Since each ansatz layer acts only on two qubits, if our recompilation target $\hat{U}$ spans more than two qubits we must first decide which qubits the thinly dressed CNOT gate should apply across. To solve this, first we evaluate an entanglement measure E between each pair of qubits, such as entanglement of formation [94], concurrence [95] or negativity [96]. This can be calculated from the state vector when using emulated quantum hardware or directly on the quantum circuit. Since our goal is to reduce the entanglement of the system to zero, we choose our layer to apply to the qubit pair with the largest E.

It is also possible that all qubit pairs have $E = 0$. For example, the maximally entangled state $|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ or the non-superposition state $|\psi\rangle = |011\rangle$ does not have any pairwise local entanglement and will result in $E = 0$ for all qubit pairs. This typically happens towards the end of recompilation rather than

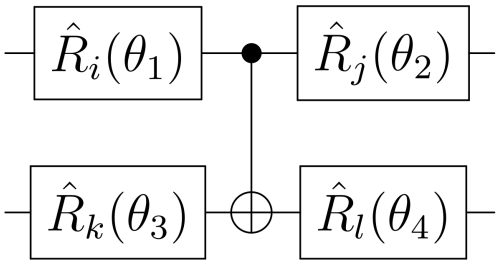


Figure 2.2: A thinly-dressed CNOT gate is a CNOT gate surrounded by 4 single-qubit rotation gates $\hat{R}_i(\theta)$, where $i \in \{ x, y, z \}$ is the axis of rotation and θ is the angle.

as the initial output of the target circuit, after several disentangling ansatz layers have already been applied. In this case, pairwise entanglement is not a good metric to base which qubits to apply the next layer on. Instead, recall that ISL converges when a set of gates are found which inverts the target circuit to produce the $|0\rangle^{\otimes n}$ state, as in Fig. 2.1. Thus, we can impose that convergence requires both $E = 0$ and that the single-qubit expectation value of each qubit in the $\hat{\sigma}_z$ basis equals 1, since $\langle 0| \hat{\sigma}_z |0\rangle = 1$. Practically then, when faced with a scenario where all qubit pairs have $E = 0$, the ISL algorithm measures each qubit and then applies a thinly-dressed CNOT layer to the two qubits with the lowest and second lowest expectation values, since these are the furthest from the solution. In the example of $|\psi\rangle = |q_3\, q_2\, q_1\rangle = |011\rangle$, a single layer would be added to the first and second qubits, which would be optimised to produce the net effect of a bit flip on both.

One constraint that we impose on the choice of the control and target is that it must not be the same as the control and target for the previous layer. This is because in general, adding layers to different choices of control and target qubits allows us to explore a greater region of the available Hilbert space. This also avoids creating circuits with large depth but small numbers of gates. Hence, if the qubit pair with the highest E is the same as in the previous layer, we choose different qubits with next largest value of E instead, working our way through the remaining possible

pairs.

Once we have chosen the control and target qubits, we add the thinly-dressed CNOT layer $\hat{V}_n^\dagger$ with initial rotations $\theta = 0$ about the Z axis.

2.2.2 Optimising the layer

After the layer $\hat{V}_n^\dagger$ is added, the axes and angles of rotation of the single-qubit gates are optimised using the Rotoselect algorithm [91], with respect to minimising Eq. (2.1). This procedure works by fixing three of the gates and varying the rotation axes and angle for the remaining one. Importantly, adjusting the angle of rotation for the single unfixed gate will cause a sinusoidal change in the cost function, displaying a clear minima. This is then repeated over the remaining 3 gates, concluding the first cycle of Rotoselect. As the changing of later gates in the layer will affect the loss landscape of the first gates, we repeatedly cycle over all 4 rotation gates until a termination criteria is reached. Here we define the termination criteria to be when the reductions in the cost function between cycles is less than 1%, or when an upper-limit of 1000 cycles is reached. We note that whilst the number of cycles to convergence depends on the particular loss landscape, across the recompilation results presented in this thesis the 1000 cycle limit is never reached. Instead, Rotoselect convergence is typically achieved within 20 cycles, with a worst case observed of up to 200 cycles.

Once the single-qubit gates of this particular layer have been optimised, we then optimise the whole ansatz $\hat{V}^\dagger$ using the Rotosolve algorithm [91]. This is a similar procedure to Rotoselect, but optimises only the rotation gate angles rather than the bases and this cheaper execution allows us to run it over the entire ansatz.

2.2.3 Finishing the layer

Once the Rotosolve procedure is terminated, we perform simple non-approximate recompilation of the new best guess $\hat{V}^\dagger$. This includes the merging of adjascent rotation gates in the same basis, removing two-qubit gate blocks that form an identity (e.g., consecutive CNOT gates) and removing of any gates for which the rotation angle is below 0.001.

After this we take one final evaluation of the cost function. If the cost is now below the threshold $C < C_t$, the ISL algorithm is terminated and we recursively invert all of the gates in the ansatz to return the recompiled solution $\hat{V}$. If not, then the next layer is added and the steps outlined above are repeated.

2.3 Conclusion

Overall, ISL is an AQCR algorithm developed over the course of this DPhil. Like other approximate recompilers, the end goal of the algorithm is to find an approximate equivalent $\hat{V}$ of a quantum circuit $\hat{U}$ that for a given initial state $|\psi\rangle$ produces the same output state to within a threshold, with fewer quantum logic gates. The novelty of ISL lies in its incremental and structural approach, recognising that if the initial state is $|0\rangle^{\otimes n}$ then the candidate solution can be guided through the measured entanglement properties of the qubits. Promisingly, this unique approach has already been shown by others in the community to outperform fixed ansatz alternatives in recompilation of circuits corresponding to the simulation of atom-cavity coupled systems [97].

As motivated prior, the benefit of AQCR is realised when running $\hat{V}$ in place of $\hat{U}$ on near-term quantum computers. Despite the accuracy of the state $\hat{V}\,|0\rangle^{\otimes n}$ being theoretically upper-bounded due to the approximations made, the shorter circuit

generates less noise, which leads to an overall higher accuracy result in the presence of significant noise. This potential benefit is what drove the development of ISL in the first place and the method was derived heuristically out of the need to reduce the errors encountered when solving optimisation problems on real hardware. As such, we do not present a theoretical analysis of the ISL method and its convergences under different conditions, which we consider explicitly beyond the scope of this thesis. We do however numerically test the effectiveness of ISL in its application to the research projects in chapters 3, 4 and 5. In each of these cases ISL performs excellently, producing significantly shallower circuits that allow us to obtain high quality results on real quantum hardware that would not be possible without it.

This detail of which type of hardware ISL is run on requires particular attention. In the use of ISL across this thesis, the circuit that generates the cost function as shown in Fig. 2.1 is evaluated on noiseless emulated quantum hardware that directly computes the statevector. Thus, despite being a quantum algorithm (i.e., computation expressed in terms of quantum circuits), our application of ISL to these research projects is practically speaking a classical recompilation method, the cost of which scales exponentially with the number of qubits in the target circuit. However, the promising results obtained in our work certainly warrants further study of ISL, including how it might scale to classically intractable systems by performing the recompilation itself on quantum hardware. Further discussion of this is given in the thesis outlook in section 6.2.

There is also a second important caveat to the impressive results obtained when using ISL in this thesis. Looking again at Fig. 2.1, it should be noted that obtaining the cost requires an accurate evaluation of the recompilation target circuit $\hat{U}$. This is not specific to ISL, but rather is part of the definition of AQCR itself and a feature shared by other well established variational recompilers [98]. Consequen-

tially, the requirement to evaluate $\hat{U}$ in the recompilation process is as to say, that one must already know the answer in order to get a solution. Nevertheless, this does not diminish the utility of ISL for a number of reasons. Firstly, when run on a classically emulated quantum computer, despite being able to access the target state vector $|\psi_f\rangle = \hat{U}|\psi_0\rangle$, finding a new circuit that approximately prepares that state in fewer operations than $\hat{U}$ is still not a trivial task. This is perhaps best exemplified by the results of chapters 3 and 5, as well as other uses of ISL in the literature [97], in which this classical mode of ISL is able to produce sufficiently accurate equivalents of the original circuit, with shallower depth than achieved by other classical recompilation algorithms or analytic methods. In turn, this allows us to run algorithms on noisy quantum computers that would otherwise not be possible. Thus, in the near term, ISL can be used to maximise the performance of current quantum computers.

Secondly, we can consider the value of running ISL on real quantum hardware. Once again, this seems odd to suggest, since obtaining the recompiled solution to the circuit relies on us being able to accurately run $\hat{U}|\psi_0\rangle$ on the hardware to begin with. However, the quantum computers in existence today span a large range of error rates. This provides an interesting future application of ISL for *bootstrapping* quantum computers, in which the resources of a low error device are used to improve the performance of higher error devices. For example, in chapter 5 we train quantum circuits as machine learning models on visual data. Supposing that the number of gates required to encode a certain high-resolution image exceeds the capability of the high error device. Instead, one could recompile the circuit that encodes the image using ISL on the low error device and use the solution in lieu of the exact encoding on the high error device. Importantly, this would then allow the user of the high error device to continue with the task at hand (e.g., classifying if the image is a cat or dog).

Finally, in chapter 4 we present an adapted version of ISL, in which the time evolution of a quantum system is sequentially recompiled in discrete time steps. This method, which we call ladder-ISL, is able to produce recompiled circuits that approximate the dynamics at later times without ever evaluating the full time evolution operator. Excitingly, when modified this way, we ensure that the quantum computational resources to run ISL does not exceed the complexity of evaluating the original un-recompiled circuit. This extension is just one example demonstrating the flexibility of the ISL approach. Overall, the development of ISL is an exciting byproduct of our research into solving optimisation problems on near-term quantum computers, and we look forward to investigating its wider potential and impact in the future.

3 | Recompilation-enhanced simulation of electron-phonon dynamics on IBM Quantum computers

With the ISL method explained, we begin this thesis' journey towards developing novel VQAs and using them to solve optimisation problems on current quantum computers. In this chapter, we first introduce an algorithm to simulate the time evolution of the electron-phonon Hamiltonian. Whilst not a VQA itself, the importance of our algorithm is clear when considering that simulating quantum dynamics is thought to be one of the first applications for which quantum computers may demonstrate a useful advantage due to its classical hardness. By extension, this means that simulating quantum dynamics would form an excellent basis for a VQA with the potential to exhibit quantum advantage. Thus, the goal of this chapter is to establish the feasibility of simulating time evolution on near-term quantum computers. In doing so, this will help us evaluate how realistic a goal running a time-evolution based VQA on a near-term quantum computer is. Our choice of the electron-phonon model is also ambitious, as despite efficient schemes to simulate d-level systems with a digital quantum computing approach, no accurate simulation of this system has been produced on real quantum hardware to date.

3.1 Introduction

Since its inception, the number of potential applications for quantum computers has grown tremendously, including at present, computational chemistry and energy minimisation [99, 49, 100, 101], pattern finding [102, 103] and optimisation [57, 22]. Yet the number of algorithms with provable speedup over their clas-

sical counterparts is small [25, 11, 26] and further still when discounting those which aren't useful [104]. Of the applications remaining, simulating the evolution of quantum systems is perhaps the most promising, with its exponential advantage [105] in memory and time over classical simulation and broad range of use cases. For quantum computers built on two-level qubits, the simulation of purely fermionic systems has been explored widely due to their naturally equivalent degrees of freedom. However, it is the inclusion of vibrational modes that gives rise to specific important phenomena such as phonon-mediated superconductivity [106]. Thus, the ability to accurately simulate systems with both electronic and vibrational modes on a quantum computer could result in significant technological advances.

One approach to simulating electron-phonon systems is through the use of analogue quantum devices [107, 108, 109, 110, 111, 112, 113, 114, 115, 116]. This can be achieved by representing the electronic components of the system with qubits, whilst the phonon modes are represented by a physical system with the desired vibrational interactions, such as inductor–resistor–capacitor oscillators [117], standing light waves [118] or Rydberg atoms [119], depending on the type of qubit. Whilst analog quantum simulators continue to produce impressive results and address interesting problems in physics and chemistry, their application is nevertheless limited by their inflexibility and non-universality. By contrast, a digital quantum computing approach could modify the coupling strength, topology, system size and even introduce additional interactions with no cost.

Implementing d-level systems (e.g., phonons and other bosons) on a quantum computer can be achieved in a number of ways, including a hybrid quantum-classical approach discussed in [120] or both first [121, 122, 123] and second quantisation mappings [124] of the wavefunction. In this chapter we focus on the latter, for which schemes were recently demonstrated requiring an efficient

number of qubits and circuit operations [125]. Such considerations are important if simulations are going to be run on near-term quantum computers, which are limited in both their size and number of operations that can applied before the qubits decohere. Nevertheless, despite efficient scaling, accurate evolution of the electron-phonon Hamiltonian has not been demonstrated on real quantum hardware to date, owing to the large number of gates required to evolve the Hamiltonian for even one time step.

In this chapter we study the dynamics of electron-phonon systems through experiments on the Qiskit [79] statevector simulator (SV) and real quantum hardware and consider the limits required to obtain accurate results in each case. In section 3.2 we introduce the electron-phonon Hamiltonian and detail the circuit implementation of its time propagator, including initialisation and measurement of the electronic excitation population. In section 3.3 we present results achieved on emulated quantum hardware, starting in section 3.3.1 with establishing the resource requirements to obtain accurate dynamics of small systems. Specifically, we consider not only the number of qubits required, but also the necessary number of time slices the evolution is discretised into, referred to as Trotter steps. Following this we demonstrate the scalability of our approach in section 3.3.2, presenting a larger 28 qubit experiment for a more complex system with non-uniform couplings.

Subsequently, we turn our attention to simulating the evolution of the electron-phonon Hamiltonian on real quantum computers. A key component of our strategy is the use of AQCR to reduce the depth of the implementation of each Trotter step, specifically using the ISL algorithm introduced in chapter 2. In section 3.4, we apply this method to simulate systems with up to 3 lattice sites (6 qubits) on real quantum computers. Excitingly, through our time-evolution algorithm and maximising the performance of current quantum hardware with AQCR, we

achieve results that match exact diagonalisation in the strong and weak coupling regimes. Finally, in section 3.5 we summarise our findings.

3.2 Method

3.2.1 Model

The electron-phonon model is used to describe many systems with electronic and vibrational degrees of freedom, such as molecular aggregates [126] including natural and artificial light harvesting systems [127]. In this model we consider N sites, where each site consists of an electronic two-level system. The electronic part of the Hamiltonian describing this model is given by

$$H_{\mathrm{el}} = \sum_{i=1}^{N} \tilde{\epsilon}_i |i\rangle\langle i| + \sum_{j=1}^{N} \sum_{i<j}^{N} \mathcal{V}_{ij} (|i\rangle\langle j| + |j\rangle\langle i|) \tag{3.1}$$

where $\tilde{\epsilon}_i$ is the electronic transition energy for site i and $\mathcal{V}_{ij}$ is the dipole-dipole coupling between sites i and j. Here we focus on the case of a single electronic excitation, such that $|i\rangle = |g_1 \cdots e_i \cdots g_N\rangle$ denotes the state where all sites are in the electronic ground state, except site i which is populated by an excitation. In this framework, we can also consider $\mathcal{V}_{ij}$ as the hopping amplitude of the excitation between sites. The vibrational modes of the system can be described under the harmonic approximation by a finite set of quantum harmonic oscillators (QHOs) [128, 122]. In this case, the phonon Hamiltonian can be written as

$$H_{\mathrm{ph}} = \sum_{i=1} \sum_{l} \hbar\omega_{il} (\hat{a}_{il}^{\dagger} \hat{a}_{il} + 1/2) \tag{3.2}$$

where $\hat{a}_{il}^{\dagger}$ ($\hat{a}_{il}$) is the bosonic creation (annihilation) operator for the vibrational excitation of the lth phonon mode of site i and ω_{il} is the frequency of oscillation.

The interactions between the electronic excitations and phonons is given by

$$H_{\text{ep}} = \sum_{i=1}^{N} |i\rangle\langle i| \left[\sum_{l} \chi_{il} (\hat{a}_{il}^{\dagger} + \hat{a}_{il}) \right] \tag{3.3}$$

where χ_{il} is the interaction strength between site i and it's l^{th} phonon mode and has dimension of energy. Finally, the total electron-phonon Hamiltonian can be expressed as the sum of these three components

$$H = H_{\text{el}} + H_{\text{ph}} + H_{\text{ep}}. \tag{3.4}$$

The interactions described in the electron-phonon Hamiltonian and their corresponding parameters can be seen in Figure 3.1. We note that for this Hamiltonian each QHO is local to a single site, which models the internal vibrational modes of molecular aggregates, and as such extensions would be required to simulate systems with shared modes.

3.2.2 Circuit implementation

To study the dynamics of electron-phonon systems we need to simulate the action of the time evolution operator $\hat{\mathcal{U}}(\tau) = \exp\Big(-i\hat{H}\tau/\hbar\Big)$ on some initial state $|\psi_{\text{init}}\rangle$. This can be implemented on a quantum computer by considering the logic gates that implement the exponential of each term in Eq. (3.4), the products of which approximate the time evolution operator via a first order Suzuki-Trotter decomposition [129]

$$\hat{\mathcal{U}}(\tau) = \left(e^{\frac{-iH_{el}\tau}{\hbar\eta}} e^{\frac{-iH_{ph}\tau}{\hbar\eta}} e^{\frac{-iH_{ep}\tau}{\hbar\eta}} \right)^{\eta} \tag{3.5}$$

to within error $O(\delta^2)$ where $\delta = \frac{\tau}{\eta}$ and η is the number of steps the evolved time is discretised into, often called the Trotter steps.

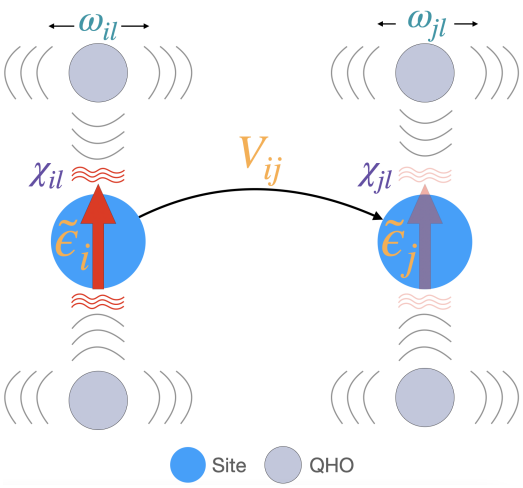


Figure 3.1: Diagram of the model considered in this chapter. A lattice site i can be occupied by an electronic excitation with transition energy $\tilde{\epsilon}_i$, which can hop to connected sites with transition amplitude $\mathcal{V}_{ij}$. Each site is also connected to l phonon modes, approximated by quantum harmonic oscillators with frequency ω_{il}, which interact with the excitation via the electron-phonon coupling constant χ_{il}.

The electronic part of the Hamiltonian H_{el} can be implemented on a quantum computer by rewriting it in terms of Pauli matrices. First, we recognise that the single-qubit operators spanning the computational basis set can be substituted for Pauli operators via the relations

$$|0\rangle\langle 0| = \frac{1}{2}\left(\mathbb{1} + \hat{\sigma}^z\right) \tag{3.6a}$$

$$|0\rangle\langle 1| = \frac{1}{2}\left(\hat{\sigma}^x + \mathrm{i}\hat{\sigma}^y\right) \tag{3.6b}$$

$$|1\rangle\langle 0| = \frac{1}{2}\left(\hat{\sigma}^x - \mathrm{i}\hat{\sigma}^y\right) \tag{3.6c}$$

$$|1\rangle\langle 1| = \frac{1}{2}\left(\mathbb{1} - \hat{\sigma}^z\right). \tag{3.6d}$$

Subsequently, site operators (e.g., $|i\rangle = |g_1 \cdots e_i \cdots g_N\rangle$) can be represented as

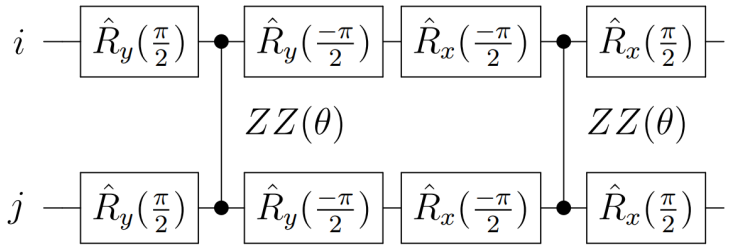


Figure 3.2: Circuit diagram implementing the operator $\hat{XY} = \exp\Big\{\frac{\mathcal{V}_{ij}\tau}{2\eta}(\hat{\sigma}_i^x\hat{\sigma}_j^x + \hat{\sigma}_i^y\hat{\sigma}_j^y)\Big\}$.

the tensor product of the above relations. Applying this to Eq. (3.3) and simplifying we obtain

$$H_{el} = \frac{1}{2}\sum_{i=1}^{N} \tilde{\epsilon}_i \, \hat{\sigma}_i^z + \frac{1}{2}\mathcal{V}_{ij}(\hat{\sigma}_i^x\hat{\sigma}_j^x + \hat{\sigma}_i^y\hat{\sigma}_j^y), \tag{3.7}$$

where $\hat{\sigma}_i^\alpha$ is the Pauli operator $\alpha \in \{ x, y, z \}$ acting on qubit i. Exponentiating Eq. (3.7), each term of the first summation can be implemented by a single qubit rotation $\hat{R}_i^z(\theta) = \exp\{-\theta\hat{\sigma}_i^z/2\}$, where substituting the general rotation parameter $\theta = \tilde{\epsilon}_i\tau/\eta$ gives us the desired evolution for one Trotter step. For the second summation in Eq. (3.7), the exponential of each term can be implemented by the circuit shown in Figure. 3.2 [130]. Here simulating one Trotter step of evolution requires the ZZ rotation $\theta = \mathcal{V}_{ij}\tau/\eta$.

To implement H_{ph} and H_{ep}, we first require a representation of the QHO in the qubit basis. This is achieved using the standard binary mapping, a method which although established in previous works [131, 128, 132, 133], we briefly describe here for completeness. Considering a QHO in the second quantisation formalism, we represent each of its d energy levels by binary strings (e.g., $|0\rangle, |1\rangle, |2\rangle, |3\rangle \longrightarrow |00\rangle, |01\rangle, |10\rangle, |11\rangle$), which can subsequently be encoded into the computational basis states of $n_x = \log_2(d)$ qubits. Since the QHO has infinite bosonic degrees of freedom, any value of d will introduce truncation errors. Therefore, it is important to find the number of levels required for the dynamics to converge, as we do later.

In the second quantisation, any operator acting across the QHO energy levels can be expressed in terms of the ladder operators $\hat{a}^\dagger$, $\hat{a}$. In turn, these can be written as binary operators, for example the term $\hat{a}^\dagger + \hat{a}$ in Eq. (3.3) can be expanded as

$$\hat{a}^\dagger + \hat{a} = \sum_{d=0}^{2^{n_x}-2} \sqrt{(d+1)}|d+1\rangle\langle d| + \sum_{d=1}^{2^{n_x}-1} \sqrt{d}|d-1\rangle\langle d|, \tag{3.8}$$

where $|d\rangle$ is the eigenstate of the dth excited state. Considering the specific example of $n_x = 2$ $(d = 4)$, we obtain

$$\begin{aligned} \hat{a}^\dagger + \hat{a} = &|01\rangle\langle 00| + |00\rangle\langle 01| \\ &+ \sqrt{2}(|10\rangle\langle 01| + |01\rangle\langle 10|) \\ &+ \sqrt{3}(|11\rangle\langle 10| + |10\rangle\langle 11|). \end{aligned} \tag{3.9}$$

Subsequently, the action on each qubit can be mapped to Pauli operators as shown in Eq. (3.6). As an example in the context of ladder operators, the first two terms of Eq. (3.9) would then become

$$|01\rangle\langle 00| + |00\rangle\langle 01| = \frac{1}{2}\left(\hat{\sigma}_1^x + \hat{\sigma}_2^z\hat{\sigma}_1^x\right) \tag{3.10}$$

after cancellations.

Overall, using this method allows the transformation of either ladder operator to $O(d^2)$ Pauli strings [128], which we define as $\hat{P} = \bigotimes_{i=1}^{n_x} \hat{\sigma}_i^\alpha$. For our time evolution algorithm, we implement the exponential $\exp\left(-\mathrm{i}\theta\hat{P}\right)$ in a quantum circuit using the well-known CNOT staircase method [134, 125, 135].

Using the qubit representations of the QHO described above, we can decompose the time evolution of the remaining Hamiltonians H_{ph} and H_{ep} into quantum cir-

cuits. For H_{ph}, the bosonic number operator $\hat{a}_{il}^{\dagger}\hat{a}_{il}$ is a diagonal evenly-spaced operator which has the known decomposition [125]

$$\hat{a}_{il}^{\dagger}\hat{a}_{il} = \sum_{k=0}^{n_x} 2^k \hat{\sigma}_{k+1}^z = 2^0\hat{\sigma}_1^z + 2^1\hat{\sigma}_2^z + ... \tag{3.11}$$

where the index k spans specifically the qubit register representing the lth phonon mode of site i. Thus

$$\exp\Big(-\mathrm{i}\theta\hat{a}_{il}^{\dagger}\hat{a}_{il}\Big) = \prod_{k=0}^{n_x} \exp\left(-\mathrm{i}\frac{\theta}{2}2^{k+1}\hat{\sigma}_{k+1}^z\right) \tag{3.12a}$$

$$= \prod_{k=1}^{n_x} \hat{R}_z^k(-\mathrm{i}\theta 2^k), \tag{3.12b}$$

implements the desired Trotterised phonon evolution for each term in Eq. (3.2) given $\theta = \hbar\omega_{il}\tau/\eta$.

Finally, the time evolution circuit for H_{ep} is generated first by following the decomposition of $\hat{a}_{il}^{\dagger} + \hat{a}_{il}$ in Eq. (3.8) through (3.10), continuing to generate Pauli strings $\hat{P}_{il}$ for all the binary operators. Combining these with the prepended site operator, the terms $\hat{\sigma}_j^z \bigotimes \hat{P}_{il}$ build up the Trotter step circuit $\exp(-\mathrm{i}H_{ep}) = \prod_i \prod_l \exp\Big(-\mathrm{i}\theta\hat{\sigma}_i^z \bigotimes \hat{P}_{il}\Big)$, where each term in the product is implemented using the previously cited CNOT staircase method. This implements the Trotterised evolution of H_{ep} given $\theta = \chi_{il}\tau/\eta$.

3.2.3 Calculation of population

In section 3.2.2 we outlined the circuits required to implement one Trotter step of the electron-phonon Hamiltonian. Simulating the time dynamics of a system which evolves under this Hamiltonian requires the application of multiple Trotter steps on an initial state $|\psi_{\text{init}}\rangle$, which here we consider to be the electronic and

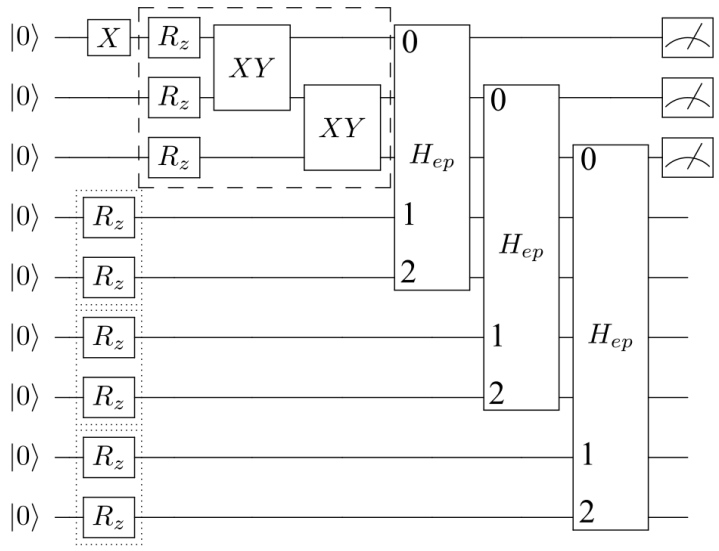


Figure 3.3: Circuit diagram showing state initialisation and a single time evolution step. First, one of the site qubits are flipped to initialise the system with a single excitation, in this case site 0. The dashed box shows the application of H_{el}, with phase rotation gates to encode the electronic transition energies and $\hat{XY}$ gates encoding the electronic excitation hopping. Note that in this case, the sites are not fully connected and instead represent a model with a chain topology. The dotted boxes shows the application of H_{ph} to each oscillator, each of which consist of a $n_x = 2$ qubit register for a total of $d = 4$ energy levels. Afterwards, the interaction H_{ep} between each site and it's corresponding phonon mode is applied, the decomposition of which is too large to show. Finally, the site qubits are measured to obtain observables such as population.

vibrational ground state apart from a single electronic excitation. This is achieved by initialising all of the qubits in the $|0\rangle$ state and then performing a bit flip on the one site where the excitation is localised to, requiring only 1 gate in total. We note that this is an advantage compared to first quantisation schemes [122], in which the vibrational ground state requires the preparation of a Gaussian across the amplitudes of the phonon qubits. The overall circuit for state initialisation followed by a single Trotter step can be seen in Figure. 3.3

Once the desired time evolution has been reached by repeated Trotter steps, the qubits are measured to retrieve the desired time-dependent observable under study.

Here we use a single measurement of each site in the $\hat{\sigma}^z$ basis, which results in the qubits collapsing to the binary $|0\rangle$ or $|1\rangle$ state depending on the absence or presence of the excitation respectively. Through repeated sampling of the circuit we obtain the average electronic excitation population on each site at the given time. From here we refer to this simply as the population.

3.3 Results on simulated quantum hardware

3.3.1 With comparison to ED

We first apply our algorithm to small systems of only two sites and one phonon mode per site. This serves two purposes. Foremost, this allows us to verify our gate-based quantum computing implementation of the time evolution propagator and compare the results obtained to exact diagonalisation (ED) methods. Secondly, this allows us to check the convergence of dynamics with respect to errors caused by our mapping of the Hamiltonian into a qubit basis, namely Trotterisation error and truncation from representing infinite bosonic degrees of freedom with a finite number of qubits. In this section, for simplicity we only consider the case whereby site dependent parameters are equal (e.g., $\mathcal{V}_{ij} = \mathcal{V}_{kl} = \mathcal{V}$), and subsequently set $\tilde{\epsilon} = 0$ since it only contributes a global phase.

3.3.1.1 Population with no phonon coupling

Figure 3.4 shows the site populations over time for the case of weak electronic coupling $V = 0.05\hbar\omega$ and no phonon coupling. Here, as expected for the purely electronic case, the exact diagonalisation results (dashed line) demonstrate a simple pattern of the excitation transferring from its initial placement on site 0 to site 1. The different markers represent obtaining the population using the quantum circuit method described in the main text, repeated with increasing numbers of

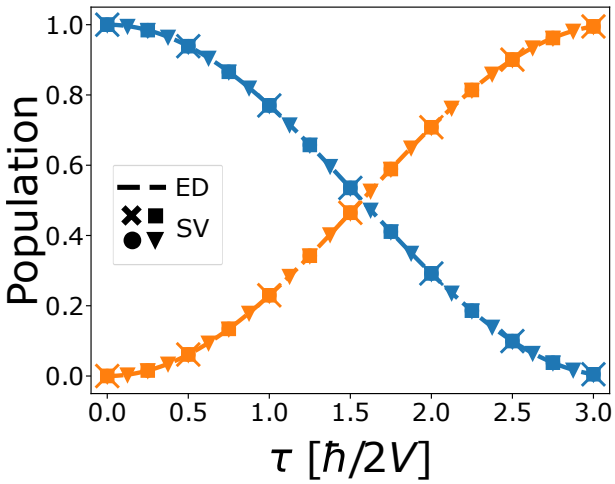


Figure 3.4: Population over time at sites 0 (blue) and 1 (orange), as calculated by exact diagonalisation (ED, dashed line) or quantum circuits run on the Qiskit Aer statevector simulator (SV, markers) with 6 (cross), 12 (square), 24 (circle) or 48 (triangle) Trotter steps. The Hamiltonian is simulated with parameters $V = 0.05\hbar\omega$, $\chi = 0.0\hbar\omega$. In this case there is no coupling to the phonon modes, so the system evolves under just $H = H_{\mathrm{el}}$. Since there are no non-commuting terms, the quantum circuit method incurs zero Trotter error and thus achieves equal accuracy to ED for any number of Trotter steps. Note that consequentially the dashed line is partially obscured by the symbols.

Trotter steps. Here the lack of electron-phonon coupling simplifies things, since the system effectively evolves under the Hamiltonian $H = H_{el}$. Therefore, without any non-commuting terms in the Hamiltonian, the Trotter error is zero and adding more Trotter steps does not improve the accuracy of the solution with respect to the ED results.

3.3.1.2 Weak electronic coupling

Figure 3.5 shows the population for the case of weak electronic coupling $V = 0.05\hbar\omega$ with non-zero electron-phonon coupling. Starting with $\chi = 0.3\hbar\omega$ on the upper panels, the ED results show that overall, the dynamics look very similar to the purely electronic case in section 3.3.1.1 except that the crossing where the populations are equal occurs slightly later. This is caused by excitation of the vi-

brational modes of the system, leading to an effective re-scaling of the electronic coupling [136]. However, the results as calculated using quantum circuits on the SV simulator are different to the purely electronic case in terms of convergence with number of Trotter steps. With the electron-phonon interactions now switched on, the non-commuting terms of the Hamiltonian lead to an error in the observed population when using a small number of Trotter steps. As we increase the number of Trotter steps, we see that this error reduces, reaching convergence with the ED solution at $\eta = 48$.

Next we consider the strong electron-phonon coupling case $\chi = 1.0\hbar\omega$, shown on the lower panels of Figure 3.5. Here the dynamics evolve differently to $\chi = 0.3\hbar\omega$. Firstly, the time taken for the excitation to transfer between sites is further extended, to the point where the crossing doesn't occur within the simulated time. Furthermore, we also observe a second mode of population transfer, visible as rapid oscillations. Looking across at the Trotter step convergence, we find that $\eta = 144$ Trotter steps are required to produce these more complicated dynamics with accuracy equal to ED.

After finding the number of Trotter steps required to accurately simulate the dynamics for $V = 0.05\hbar\omega$, we next consider convergence with respect to the truncation of the QHO. In the furthest right graph of Figure 3.5, the blue bar shows the average occupation of each energy level of the oscillator coupled to site 0 during the converged experiments on the left. This is then repeated with increasing numbers of QHO energy levels $d = 2^{n_x}$ up to $n_x = 4$. Looking at the results of all the bars, for $\chi = 0.3\hbar\omega$ we find that only the ground state $d = 0$ and first excited state $d = 1$ are occupied to any observable level. This implies that simulation of this system is indeed converged with only $n_x = 1$ qubit per phonon mode. For strong electron-phonon coupling however, the truncation when choosing $n_x = 1$ is clear, since for larger values of n_x there is a non-zero occupation of the $d = 2$

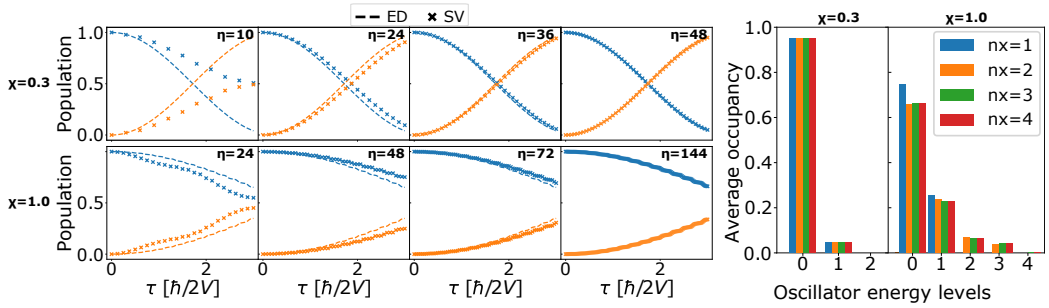


Figure 3.5: Left: population over time at sites 0 (blue) and 1 (orange), as calculated by exact diagonalisation (ED, dashed line) or quantum circuits run on the Qiskit Aer statevector simulator (SV, crossed marker). The Hamiltonian is simulated with parameters $V = 0.05\hbar\omega$, $\chi = 0.3\hbar\omega$ (upper panels), $\chi = 1.0\hbar\omega$ (lower panels) and $n_x = 1$. Moving across the panels left to right, for the quantum circuit approach, time evolution is discretised by an increasing number of Trotter steps η. Right: average energy level occupation of the oscillator coupled to site 0 during time evolution. The blue bar corresponds to the experiment shown on the left, where each phonon mode is represented by $n_x = 1$ qubit. The other bars represent subsequent experiments with increasing numbers of qubits per phonon mode, which allows simulation of up to $d = 2^{n_x}$ oscillator energy levels. These experiments were run with the number of Trotter steps required for convergence (e.g., $\eta = 48$ for $\chi = 0.3\hbar\omega$ and $\eta = 144$ for $\chi = 1.0\hbar\omega$).

and $d = 3$ excited states. Furthermore, whilst the $d = 4$ level has a very small average occupation of 0.003 and is barely visible, we can conclude that its inclusion is necessary for convergence of the dynamics. This can be seen in the $n_x = 2$ case, whereby not being able to access the $d = 4$ excited state leads to an overestimation of the occupancy at $d = 1$ and underestimation at $d = 3$. Thus, for a two-site $V = 0.05\hbar\omega, \chi = 1.0\hbar\omega$ system, each phonon mode needs to be represented by at least $n_x = 3$ qubits for full convergence.

3.3.1.3 Strong electronic coupling

Next we consider our two-site system in the strong electronic coupling regime $V = 1.0\hbar\omega$, repeating the experiments to determine convergence with respect to η and n_x. These results can be seen in Figure 3.6. Notably, although the larger

value of V leads to the excitation hopping between sites at a much faster rate, this is factored out by plotting time in units of $\hbar/2V$. With this accounted for, looking at the results for $\chi = 0.3\hbar\omega$ on the upper panels, we see that the population looks quite similar to its equivalent in Figure 3.5, albeit with a slight damping of the oscillation. However, there is a large difference when comparing convergence with respect to Trotter steps, which occurs here for as few as $\eta = 6$ as shown in the second panel. Unlike the other regimes studied in this chapter, such few Trotter steps are required for an accurate solution because it is the closest to the purely electronic case based on the ratio V/χ. Thus, this regime may be the first that can be simulated accurately on near term quantum computers, due to its low quantum circuit depth requirements. This is further demonstrated by the average oscillator energy level occupancy, as shown in the rightmost graph. Here we see the state of the oscillator can be represented without truncation by only $n_x = 1$ qubit, with only the ground and first excited state being occupied. However, we note that this regime may not be a desirable goal in the context of quantum advantage, as the domination of the electronic interactions makes this system particularly amenable to classical approximate solvers such as the density-matrix renormalization group (DMRG) method [137]. Such methods can efficiently simulate a truncated local Hilbert space with high accuracy in this regime, only becoming unyieldingly costly for large electron-phonon interaction strengths [138].

Finally we examine the case of both strong electronic coupling $V = 1.0\hbar\omega$ and strong electron-phonon coupling $\chi = 1.0\hbar\omega$, shown in the lower panels of Figure 3.6. Here, unlike the weak regime, the populations do not cross and we instead see an inflection at $\tau = 2$. Simulation over a longer time would be required to see whether this is in fact a secondary higher frequency mode of oscillation, or is a reflection of the primary excitation wave. Looking across the panels, we see that $\eta = 48$ Trotter steps are required for convergence. Considering the average QHO

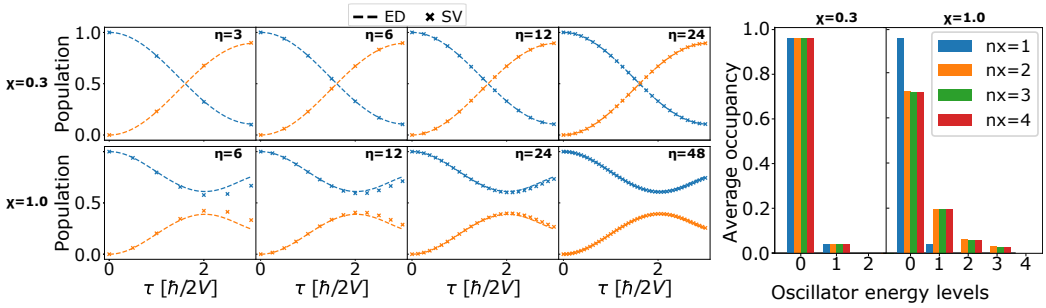


Figure 3.6: Left: population over time at sites 0 (blue) and 1 (orange), as calculated by exact diagonalisation (ED, dashed line) or quantum circuits run on the statevector simulator (SV, crossed marker). The Hamiltonian is simulated with $V = 1.0\hbar\omega$, $\chi = 0.3\hbar\omega$ (upper panels), $\chi = 1.0\hbar\omega$ (lower panels) and $n_x = 1$ with increasing Trotter steps from left to right. Right: average energy level occupation of the oscillator coupled to site 0 during time evolution.

energy level occupancy, the rightmost graph demonstrates again that in the strong electron-phonon regime the oscillators are excited to higher energy levels than the weak regime. Similar to the $V = 0.05\hbar\omega$ case, whilst the scarcely visible $d = 4$ occupation may suggest that $n_x = 2$ is sufficient, the average occupancy across all levels only converges for $n_x = 3$ qubits.

3.3.2 Simulator results of larger real world system

In the previous section, we presented results obtained using a statevector simulator for systems with two sites, one phonon mode per site and up to $n_x = 4$ for a maximum of ten qubits. This allowed us to study different coupling regimes and by keeping the Hilbert space small, compare them easily to exact diagonalisation results. In this section however, we demonstrate the flexibility of the quantum circuit approach to scale to much larger systems. Unlike analogue simulation of the electron-phonon Hamiltonian, our method requires no redesigning of the algorithm or hardware to scale up, since the circuits are created using the general rules laid out in section 3.2.2. Furthermore, whilst our previous simulations had

uniform couplings, here we can consider systems with non-uniform couplings with no additional cost.

Figure 3.7 shows the result of our larger scale simulation, run on the Qasm simulator with $k = 10,000$ shots. Here we have $N = 7$ sites, $l = 1$ phonon mode per site and $n_x = 3$ qubits per oscillator for a total of 28 qubits. The couplings $\vec{V}$ and $\vec{\chi}$ are randomly generated from a uniform distribution in the range $[0.8, 1.2]\hbar\omega$ and can be found in Appendix A.1. We choose this range such that the mean value still represents the strong coupling regime, which we found in section 3.3.1 to have interesting and non-trivial dynamics.

Looking at Figure 3.7 up to $\tau = 3$, we see a pattern whereby the excitation originally localised to site 0 distributes down the chain. This is the primary excitation wave, causing visible peaks on sites 1, 2 and 3 on this short time scale. However, beyond this point the dynamics become much more complex, with many oscillations and peaks of different scales. This is caused first by the reflection of the primary wave off each site, and subsequently by the complex interference between counterpropogating wavepackets. By the end of our simulation, the net transfer across the sites begins to slow, such that the dynamics appear to begin reaching an equilibrium.

3.4 Real device results

Up until this point, all results in this chapter have been obtained by executing the quantum circuits that make up the time evolution operator on a classically emulated quantum computer. This was done to factor out the noise generated by real quantum computers, which for current generation devices is prohibitively large to accurately simulate the time evolution circuits required. However, we posture that with AQCR, it is in fact possible to obtain accurate results which capture the dy-

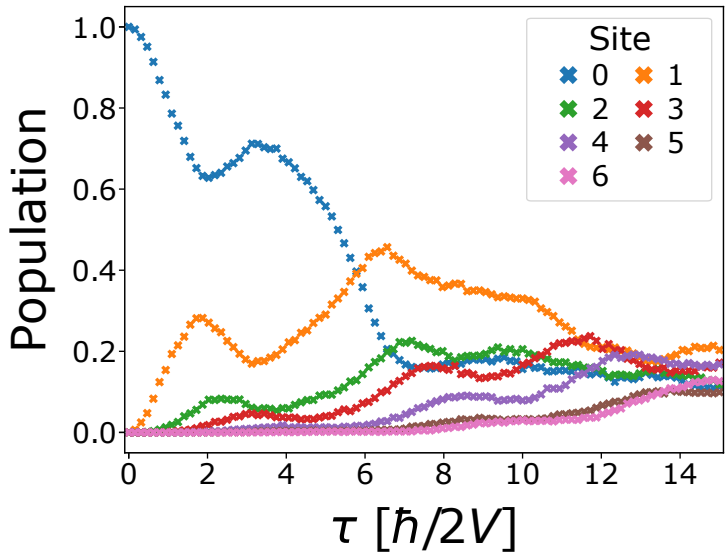


Figure 3.7: Population over time with non-uniform couplings and phonon frequencies, $\eta = 96$ and $n_x = 3$. All results were calculated using the quantum circuit implementation described in the main body of this work, run on the Qiskit Qasm simulator with 10,000 shots. The couplings $\vec{V}$, $\vec{\chi}$ and frequencies were randomly generated from the uniform interval [0.8, 1.2] and can be found in Appendix A.1.

namics of the electron-phonon Hamiltonian on a real quantum computer. Specifically, we apply our incremental structural learning (ISL) algorithm detailed in chapter 2, which novelly uses the entanglement properties of the circuit to avoid trialing different ansatzes. This allows us to learn approximate equivalents of all the Trotterised electron-phonon evolution circuits that contain the Hamiltonian dynamics.

To start, we first try the regime from section 3.3.1 which achieves convergence with the fewest quantum resources. For this we consider a two-site system with couplings $V = 1.0\hbar\omega$, $\chi = 0.3\hbar\omega$, which produces accurate results with as few as $n_x = 1$ qubit per oscillator (4 qubits total) and $\eta = 12$ Trotter steps. The left panel of Figure 3.8 shows the population, obtained through experiments on the 7 qubit, 32 QV (quantum volume) *ibm_lagos* quantum computer with $k = 8192$ shots

per circuit. The triangle markers indicate the results obtained when implementing the Trotterised time evolution operator without AQCR. Due to the small size of the system, only 6 CNOT and 12 single-qubit gates are required per Trotter step, producing accurate results for up to 6 Trotter steps of evolution. For this simple problem, that corresponds to half the evolved time due to its low total Trotter step requirements. Yet the number of gate operations still grows linearly with the number of Trotter steps, such that simulating 7 through 12 Trotter steps of evolution, requiring 42 to 72 CNOT gates respectively, leads to a sharp decrease in accuracy. Here we focus on the number of two-qubit gates, as opposed to single-qubit gates, because their error rate is an order of magnitude higher on the devices used.

By contrast, the crossed markers show the results obtained running on the same device but first recompiling the circuits with ISL. Here, for *any* number of Trotter steps, ISL finds an equivalent circuit producing the same state to 99% overlap with on average 3.4 CNOT gates and 8.8 single-qubit gates. Excitingly, this enables the results obtained on the real device to accurately reproduce those calculated via exact diagonalisation, even in the region $\tau > 1.5$. It is important to remember that in this case ISL is more akin to a classical recompilation method, running on emulated quantum hardware with direct access to the target state vector. Nevertheless, even in this paradigm, we are able to use ISL to simulate the dynamics of our Hamiltonian with shallower circuits than previously found through other classical techniques. This in turn allows us to maximise the accuracy achieved on real quantum hardware, as shown in this result.

We note that whilst for the non-recompiled experiments we recorded the number of gate operations *per* Trotter step, here we express the number of gates as an average over all different evaluated time points. This is because in the non-recompiled case the overall number of gates scales linearly with the number of Trotter steps

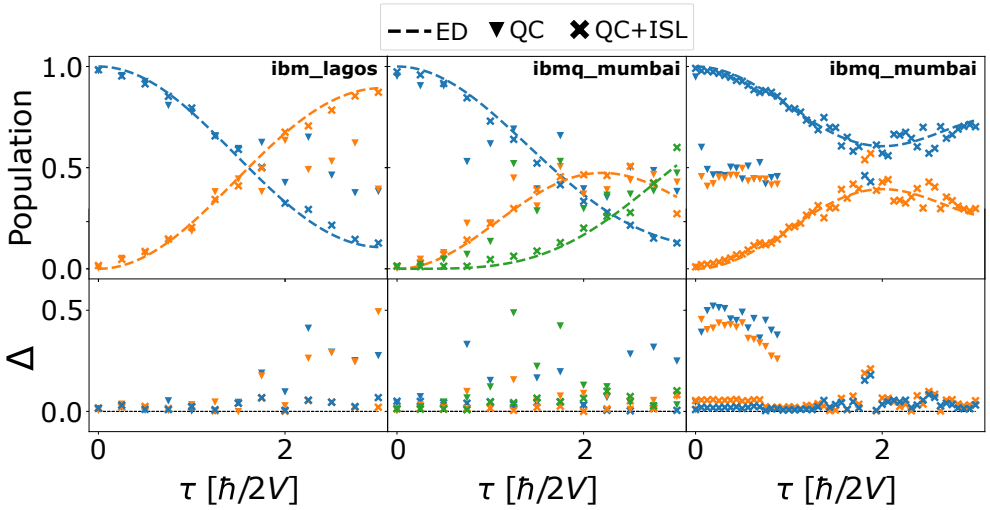


Figure 3.8: Population over time as calculated by ED (dashed line) or quantum circuits executed on an IBM quantum computer directly (triangle marker) or first recompiled with ISL (crossed marker). All experiments are in the strong electronic coupling regime $V = 1.0\hbar\omega$. Left: system parameters $N = 2, \chi = 0.3\hbar\omega, \eta = 12$ and $n_x = 1$ run on *ibm_lagos*. Middle: same as the left panel, but with $N = 3$ sites, run on *ibmq_mumbai*. Right: $N = 2, \chi = 1.0\hbar\omega, \eta = 48$ and $n_x = 2$ run on *ibmq_mumbai*. The bottom panel shows the absolute difference Δ between experiments and the analytic solution for each site.

of evolution, so the total cost is easy to infer (e.g., obtaining the population at $\tau = 3.0$ requires evaluating a circuit consisting of 12 sequential Trotter steps). By contrast, the depth and number of operations of the ISL solution depends only on the entanglement complexity of the state of the system, which is not guaranteed to grow with time. Indeed in the specific experiment here, the recompiled solution to the $\eta = 5$ and $\eta = 12$ Trotter step circuits both contain 4 CNOT gates.

Following this, we simulate a larger system to further test what is possible on a real quantum computer. We repeat the previous simulation but with $N = 3$ sites, giving a total of 6 qubits required. Whilst the inclusion of extra qubits itself incurs no error, the additional gates required to simulate a larger chain and another oscillator mode leads to more noise. The middle panel of Figure 3.8 shows results

obtained on the 27 qubit, 128 QV *ibmq_mumbai* device with $k = 8192$ shots per circuit. Here, a direct implementation of the time evolution requires 10 CNOT gates and 20 single-qubit gates per Trotter step, the results of which are shown by the triangle marker. Despite running on a higher fidelity quantum computer, the increased number of gates per Trotter step are such that the population obtained at each site is only accurate for the first few Trotter steps. Beyond this the accuracy decreases, such that for the 12 Trotter step circuit used to calculate the population at $\tau = 3.0$, requiring 120 CNOT gates, the qubit measurement statistics are random. We note that a random qubit state translates to a measured population of 0.5, which in this graph is deceivingly close to the real dynamics at later time steps.

When recompiling this larger system, ISL finds for any number of Trotter steps an equivalent circuit producing the same state to 95% overlap with on average 10.0 CNOT gates and 18.7 single-qubit gates. Although this is more operations than the two site case, it is sufficiently shallow so that the noisy device can produce dynamics similar to that of ED for all but two of the observed points. The increased error at $\tau = 2.5$ and 3.0 can be explained by the observation that the recompiled solutions at these times contained more operations than average, with 21 and 16 CNOT gates respectively.

As mentioned in section 3.3.1, although the low Trotter step and n_x requirements of the small χ regime make it appealing for quantum simulation, it is unlikely one would require non-classical methods for such simple dynamics. Therefore, in this final experiment we test what can be simulated on a real quantum computer for the nontrivial strongly coupled case $V = 1.0\hbar\omega, \chi = 1.0\hbar\omega$. Here we use $\eta = 48$ Trotter steps as required for convergence and $n_x = 2$, which suffers from very minimal truncation at the benefit of less qubits and gates. In this case the limitation of the direct implementation becomes apparent, with each Trotter step requiring 42 CNOT and 66 single-qubit gates. Furthermore, because this regime requires

$n_x > 1$ qubits per oscillator for convergence, the graph of connected qubits in the circuit no longer matches the honeycomb structure of the hardware used in these experiments. Mapping between the topologies requires SWAP operations which dramatically increases the number of CNOT gates per Trotter step to 133. The result of executing these circuits on *ibmq_mumbai* is shown by the triangle markers in the right panel of Figure 3.8. Here, it is not possible to obtain an accurate population after even 1 Trotter step of evolution, and the signal rapidly decays to random noise. Conversely, the populations obtained using ISL match the ED results well up to $\tau = 1.5$ and follow the correct trend after. This is possible because the recompiled circuits allow us to obtain the dynamics for any point in time up to $\tau = 3.0$ with on average 15.6 CNOT gates and 26.9 single-qubit gates. Furthermore, the use of SWAP gates can be avoided by restricting the qubit connections allowed during ISL to only those which are connected on the physical device.

3.5 Conclusion

In this chapter we demonstrate the first known simulation of electron-phonon dynamics on a real quantum computer. To achieve this, we first establish the previously unknown Trotter step and qubit requirements for accurate simulation of systems described by the electron-phonon Hamiltonian, which we find grows with the ratio of the electron-phonon to dipole-dipole couplings χ/V. We then run the converged time evolution circuits on real quantum computers and demonstrate that, despite the resource efficiency of the scheme, accurate results cannot be obtained even for small systems. To remedy this, we use incremental structural learning to find approximately equivalent circuits for each time step in the simulation, leading to a significant reduction in the number of gates required. Executing the recom-

piled circuits on a real quantum computer, we obtain highly accurate dynamics for small systems with weak electron-phonon coupling and partially accurate results for the more complicated case of strong coupling.

In the long term, the study of electron-phonon dynamics will benefit greatly from simulation on error-corrected quantum computers and its associated exponential speedup. However, the ongoing development of quantum hardware demonstrates that the ability to execute deep circuits will remain a longstanding problem to obtain accurate results for interesting physical systems. In this way, our work provides a significant first step to address this, and our novel approach illustrates that electron-phonon Hamiltonians can be simulated on near-term quantum computers when using AQCR. Looking forward, evolving electron-phonon systems on a quantum computer is not only beneficial for the direct study of dynamics, but also for optimisation routines whose cost functions contain time dependent observables of such systems. Important example of this include studying the conditions required for optimal energy transfer [139] and fitting values to unknown system parameters based on experimental data [140], both of which would make interesting extensions to this chapter. Notably, looking beyond just the electron-phonon model, the results presented here pave the way more generally for the goals of thesis and provide a promising sign that it may be possible to evaluate evolution-based VQAs on real quantum hardware.

4 | Minimum Hardware Requirements for Hybrid Quantum-Classical DMFT

In the previous chapter, we demonstrated that simulating the time evolution of small quantum systems is possible on current quantum computers. In this chapter, we extend this concept to investigate whether a VQA in which the cost function requires evaluating the time evolution operator could be amenable to NISQ hardware. Specifically, we introduce a variational algorithm for dynamical mean-field theory, an optimisation procedure used to simulate strongly correlated electronic systems. By emulating noisy quantum devices, we gauge the minimal hardware requirements for a digital quantum computer to run such a scheme. When the required fidelity is discovered to be significantly beyond currently available quantum computers, we consider strategies to reduce the quantum overhead, including the previously referenced ISL circuit recompilation algorithm. With these improvements, we subsequently demonstrate through noise emulation that hybrid quantum-classical DMFT could be realised on real quantum computers in the immediate future, and is perhaps already possible at the time of writing.

4.1 Introduction

Ongoing efforts to build a quantum computer are currently characterised by hardware with less than 100 qubits, large gate errors and no error correction [141]. Nevertheless, these NISQ devices are believed to be well suited to solving optimisation problems using hybrid quantum-classical algorithms [49, 74, 63] and a number of variational quantum algorithms have been successfully demonstrated on them. For example, quantum chemistry calculations were carried out on superconducting [10, 99, 73] and ionic [142] NISQ devices, and nuclear structure calcu-

lations were performed on quantum processors accessed via cloud servers [143]. Furthermore, nuclear magnetic resonance systems were used to demonstrate a hybrid quantum-classical approach to quantum optimal control [144].

The success of NISQ devices in solving small-scale electronic structure problems is substantiated by theoretical results showing that quantum computers can solve correlated electronic structure problems in polynomial time [145, 146], e.g., via phase estimation algorithms [147]. It is therefore natural to consider if other electronic structure methods could benefit from a quantum computational approach. For example, dynamical mean-field theory (DMFT) [148] is a standard approach for simulating materials with strong electronic correlations, and proposals for hybrid quantum-classical DMFT algorithms have been put forward recently [149, 150, 151]. Experimental realisations have been achieved for the insulating phase [152] and in the case of an alternative approach to DMFT [153], which uses the variational quantum eigensolver method to calculate ground and excited states of the system. However, to the best of our knowledge the precise hardware requirements for obtaining high-quality DMFT results on a quantum computer are not known.

4.2 Model

4.2.1 Hubbard model

Strongly correlated materials in thermodynamic equilibrium are often described by the Fermi-Hubbard model:

$$\hat{H} = -t \sum_{\langle j,k \rangle \sigma} (\hat{c}^{\dagger}_{j,\sigma} \hat{c}_{k,\sigma} + H.c.) + U \sum_{j} \hat{n}_{j,\downarrow} \hat{n}_{j,\uparrow} \tag{4.1}$$

where $\hat{c}^{\dagger}_{j,\sigma}$ ($\hat{c}_{j,\sigma}$) is the fermionic creation (annihilation) operator and $\hat{n}_{j,\sigma} = \hat{c}^{\dagger}_{j,\sigma}\hat{c}_{j,\sigma}$ is the number operator acting on site j with spin component $\sigma \in \{\uparrow, \downarrow\}$. This Hamiltonian describes a system of fermions, such as electrons, that have two possible spin eigenstates. Assuming localised orbitals, electrons can hop between adjacent lattice sites with amplitude t, and lattice sites occupied by a pair of electrons experience an energy penalty U.

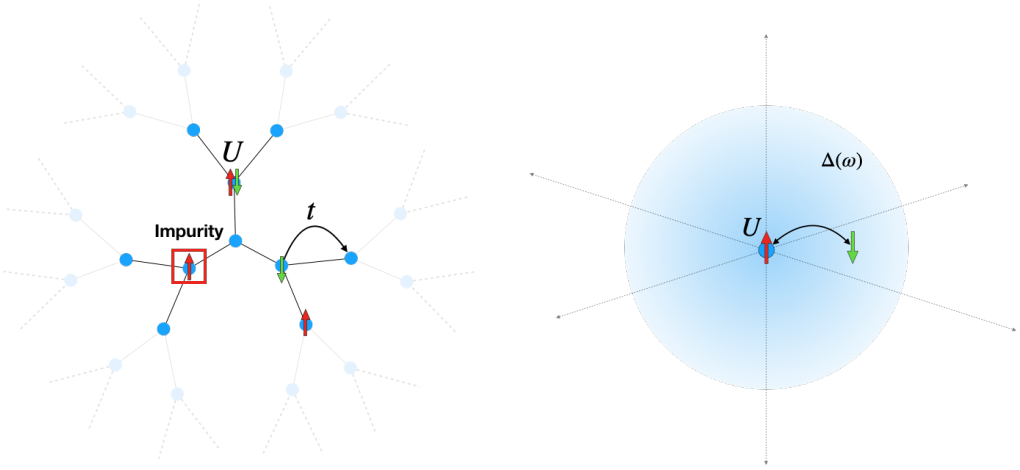
There are several popular methods used to solve the Fermi-Hubbard model that have proven success, however each are faced with their own problems. Exact diagonalisation has exponentially scaling computational complexity, quantum Monte Carlo methods suffer from the numerical sign problem and tensor networks struggle to solve systems beyond one dimension. Thus, DMFT is a popular alternative for determining the dynamics of strong electronic correlations.

4.2.2 DMFT

DMFT translates the many-body problem of the Hubbard model to a single-site impurity model. To do this, we arbitrarily choose one lattice site and label it as the impurity, as shown in figure 4.1a. We can then represent the interactions between the impurity and the surrounding fermions as a time-varying mean-field, which the impurity site can exchange electrons with as shown in figure 4.1b. This mapping from a lattice model to an impurity model is exact in the limit where the number of spatial dimensions goes to infinity [154].

In order to make this mapping accurate, the DMFT routine must self-consistently determine a mean field such that the retarded impurity Green's function is equal to the local retarded lattice Green's function,

$$G^{R}_{\mathrm{imp}}(\omega) = G^{R}_{\mathrm{latt},jj}(\omega). \tag{4.2}$$



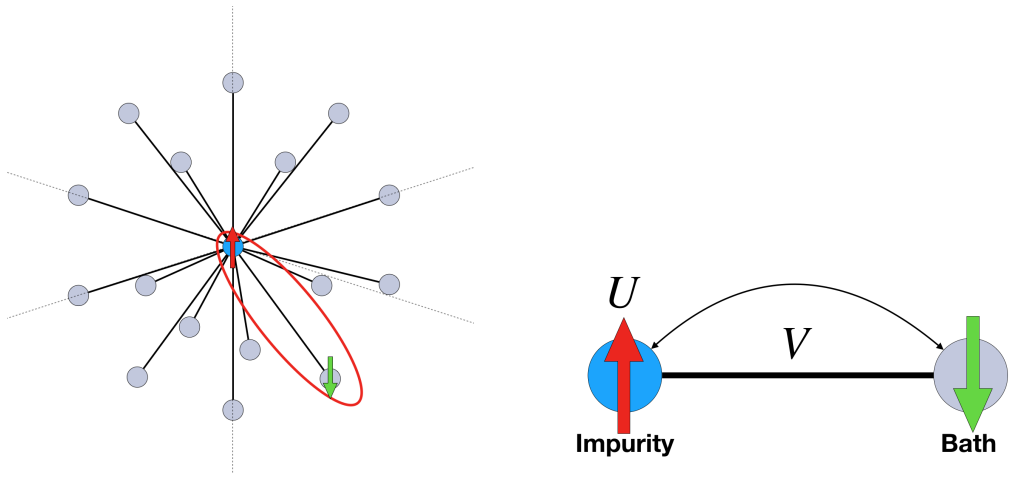
(a) (b)

Figure 4.1: (a) The Hubbard model on an infinite dimensional Bethe lattice. Each site can be occupied by no electrons, a single electron or a pair of electrons of opposite spins with interaction energy U. Electrons can hop between adjacent lattice sites with energy t. (b) In dynamical mean-field theory (DMFT), we consider one site to be the impurity and approximate its interactions with the rest of the lattice as a mean-field. Two electrons occupying the impurity site have an interaction energy of U and electrons can be exchanged with the mean-field $\Delta(\omega)$.

Here we consider the Fermi-Hubbard model embedded in an infinite dimensional Bethe lattice [148], as has been done previously for hybrid quantum DMFT [150, 153]. To account for a lattice model with infinite coordination number $z \to \infty$, the Hubbard hopping amplitude t needs to scale as $t \sim t^*/\sqrt{z}$ to avoid a diverging kinetic energy per lattice site. This defines a new constant, t^*, which is the Hubbard hopping amplitude in infinite dimensions.

4.2.3 Two-site Hamiltonian-based DMFT

In Hamiltonian-based DMFT, the mean-field is parametrised by a set of non-interacting bath sites, as shown in Fig. 4.2a. This formulation of the impurity model is particularly conducive to being solved using a quantum computer, as for a given Hamiltonian $\hat{H}$, it requires evaluating the time evolution operator



(a) (b)

Figure 4.2: (a) Hamiltonian-based DMFT replaces the mean field by a set of non-interacting bath sites. Electrons can occupy any lattice site, but can only move between the impurity (blue) and a bath site (grey). (b) In two-site DMFT, we use only a single bath site. In the half-filled case, the system dynamics are now described by two parameters, U and V, the on-site interaction and hybridization parameter respectively.

$\hat{\mathcal{U}}(\tau) = \exp\left(-i\hat{H}\tau/\hbar\right)$. This is known to be exponentially faster on a quantum computer [145].

The self-consistency condition in Eq. (4.2) can only be satisfied exactly for an infinite number of bath sites. Here we consider the minimal implementation of Hamiltonian-based DMFT which involves just two sites - one for the impurity and another to approximate the mean field, see Fig. 4.2b. This model is known as two-site DMFT [155], and provides an approximate yet qualitatively correct description of strongly correlated phenomena in the Hubbard model. The system

is now described by the SIAM Hamiltonian

$$\begin{aligned}\hat{H}_{\mathrm{SIAM}} =& U\hat{n}_{1\downarrow}\hat{n}_{1\uparrow} - \mu\sum_{\sigma}\hat{n}_{1\sigma} + \sum_{\sigma}\epsilon_c\hat{c}^{\dagger}_{2\sigma}\hat{c}_{2\sigma} \\ &+ \sum_{\sigma}V(\hat{c}^{\dagger}_{1\sigma}\hat{c}_{2\sigma} + \mathrm{H.c.}),\end{aligned} \tag{4.3}$$

where U is the same on-site interaction as in our original lattice model and μ is the impurity chemical potential.

In general, the bath site energy ϵ_c and the hybridization between the two sites V need to be determined such that the self-consistency condition in Eq. (4.2) is approximately satisfied. In the following we focus on the half-filled case, which exhibits interesting effects such as the metal-insulator transition [156] and maximal antiferromagnetic spin correlations [157]. In this case, $\mu = U/2$ and $\epsilon_c = 0$, such that Eq. (4.3) reduces to

$$\hat{H}_{\mathrm{SIAM}} = U\hat{n}_{1\downarrow}\hat{n}_{1\uparrow} - \frac{U}{2}\sum_{\sigma}\hat{n}_{1\sigma} + \sum_{\sigma}V(\hat{c}^{\dagger}_{1\sigma}\hat{c}_{2\sigma} + \mathrm{H.c.}). \tag{4.4}$$

The hybridization parameter V in Eq. (4.4) is now the only free parameter that needs to be determined for a given U such that Eq. (4.2) is approximately fulfilled. This self-consistency condition is shown to be equivalent to satisfying [155]

$$V^2 = Zt^{*2}, \tag{4.5}$$

where Z is the quasiparticle weight, which physically represents both the sign and magnitude of interactions in a Fermi liquid [158].

Determining V can be achieved via an iterative procedure incorporating a quantum processor and classical feedback loop, which we describe in the next Sec. 4.2.4.

4.2.4 Hybrid quantum-classical DMFT routine

The iterative process of hybrid quantum-classical DMFT is illustrated in Fig. 4.3 and consists of the following steps [150]:

1. Set the value of the impurity on-site interaction energy U.
2. Make an initial guess for the value of the hybridization parameter V.
3. Obtain the impurity Green's function $\mathrm{i}G^R_{\mathrm{imp}}(\tau)$ from the quantum computer as a function of time τ (i is the imaginary unit).
4. At half-filling the impurity Green's function has the form

$$\mathrm{i}G^R_{\mathrm{imp}}(\tau) = \alpha\cos(\omega_1\tau) + (1-\alpha)\cos(\omega_2\tau). \tag{4.6}$$

   Using the result for $\mathrm{i}G^R_{\mathrm{imp}}(\tau)$ obtained from the quantum computer, find the best fit for the parameters α, ω_1 and ω_2, which make up the residues and poles of $G^R_{\mathrm{imp}}(\omega)$ respectively. In contrast to other works [150, 152], we use the normalization iG(0)=1 to reduce the number of fitting parameters from four to three.
5. Calculate the quasiparticle weight according to

$$Z = \left[V^4\left(\frac{\alpha}{\omega_1^4} + \frac{1-\alpha}{\omega_2^4}\right)\right]^{-1}. \tag{4.7}$$

   If the values for Z and V satisfy Eq. (4.5), then self-consistency has been reached.
6. Otherwise, update the hybridization parameter V to one that would be self-consistent with the current system (i.e., $V_{\mathrm{new}} = \sqrt{Z}t^*$) and repeat from step 3.

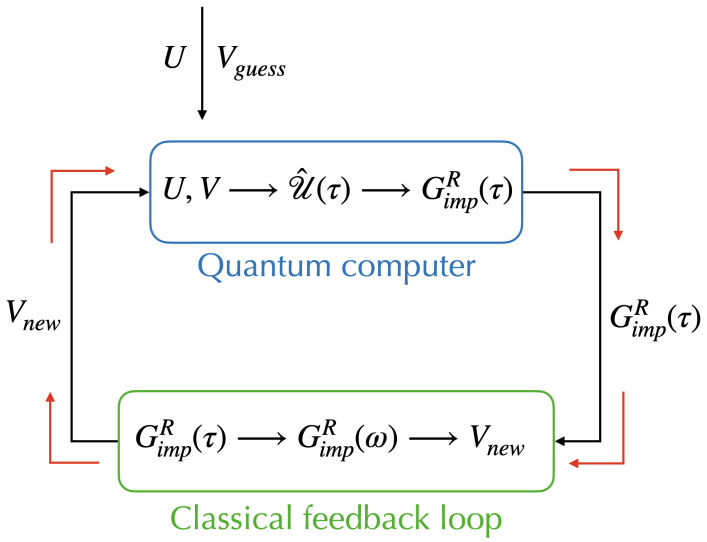


Figure 4.3: Diagram of hybrid quantum-classical DMFT. For given on-site interaction energy U, we iteratively discover the hybridization parameter V such that Eq. (4.5) is satisfied. For the first iteration, we start with a guess V_{guess}. We use a quantum computer to compute the impurity Green's function $G^R_{\text{imp}}(\tau)$, followed by a classical optimiser to suggest an improved hybridization V_{new}. The full loop is iterated until self-consistency is reached, such that $V_{\text{new}} = V$.

Next we show how the impurity Green's function can be measured using a quantum computer as required in step 3. To do this, we first map the impurity model onto a qubit system. Applying a Jordan-Wigner transformation [159] to Eq. (4.4), we obtain

$$\hat{H}_{\text{SIAM}} = \frac{U}{4}(\hat{\sigma}_1^z\hat{\sigma}_3^z) + \frac{V}{2}(\hat{\sigma}_1^x\hat{\sigma}_2^x + \hat{\sigma}_1^y\hat{\sigma}_2^y + \hat{\sigma}_3^x\hat{\sigma}_4^x + \hat{\sigma}_3^y\hat{\sigma}_4^y), \tag{4.8}$$

where $\hat{\sigma}_n^\alpha$ is the Pauli operator $\alpha \in \{x, y, z\}$ acting on qubit n. As part of this process, we assign two qubits to represent each site, due to its occupation and spin degrees of freedom.

Next, we note that the impurity Green's function can be written as [130]

$$G^R_{\text{imp}}(\tau) = \theta(\tau)[G^>_{\text{imp}}(\tau) - G^<_{\text{imp}}(\tau)], \tag{4.9}$$

where θ is the heavyside step function and the greater and lesser Green's functions are defined as

$$G^>_{\text{imp}}(\tau) = -\mathrm{i}\langle \hat{c}_{1\sigma}(\tau)\hat{c}^\dagger_{1\sigma}(0)\rangle, \tag{4.10}$$

$$G^<_{\text{imp}}(\tau) = \mathrm{i}\langle \hat{c}^\dagger_{1\sigma}(0)\hat{c}_{1\sigma}(\tau)\rangle, \tag{4.11}$$

respectively, where the average is computed in the ground-state of Eq. (4.4). We express the impurity Green's function in terms of Pauli operators via the Jordan-Wigner transformation

$$\hat{c}^\dagger_{1\downarrow} = \frac{1}{2}(\hat{\sigma}^x_1 - i\hat{\sigma}^y_1) \tag{4.12}$$

$$\hat{c}_{1\downarrow} = \frac{1}{2}(\hat{\sigma}^x_1 + i\hat{\sigma}^y_1). \tag{4.13}$$

Substituting these into Eq. (4.10) and Eq. (4.11) for the specific spin eigenvalue $\hat{c}_{1\sigma} = \hat{c}_{1\downarrow}$ we obtain

$$\begin{aligned} G^>_{imp}(\tau) &= -\frac{i}{4}\langle \hat{\mathcal{U}}^\dagger(\tau)(\hat{\sigma}^x_1 + i\hat{\sigma}^y_1)\hat{\mathcal{U}}(\tau)(\hat{\sigma}^x_1 - i\hat{\sigma}^y_1)\rangle \\ &= -\frac{i}{4}(\langle \hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^x_1\hat{\mathcal{U}}(\tau)\hat{\sigma}^x_1\rangle - i\langle \hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^x_1\hat{\mathcal{U}}(\tau)\hat{\sigma}^y_1\rangle \\ &\quad + i\langle \hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^y_1\hat{\mathcal{U}}(\tau)\hat{\sigma}^x_1\rangle + \langle \hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^y_1\hat{\mathcal{U}}(\tau)\hat{\sigma}^y_1\rangle) \end{aligned} \tag{4.14}$$

and

$$\begin{aligned} G^{<}_{imp}(\tau) &= \frac{i}{4}\langle(\hat{\sigma}^x_1 - i\hat{\sigma}^y_1)\hat{\mathcal{U}}^\dagger(\tau)(\hat{\sigma}^x_1 + i\hat{\sigma}^y_1)\hat{\mathcal{U}}(\tau)\rangle \\ &= \frac{i}{4}(\langle\hat{\sigma}^x_1\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^x_1\hat{\mathcal{U}}(\tau)\rangle + i\langle\hat{\sigma}^x_1\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^y_1\hat{\mathcal{U}}(\tau)\rangle \\ &\quad - i\langle\hat{\sigma}^y_1\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^x_1\hat{\mathcal{U}}(\tau)\rangle + \langle\hat{\sigma}^y_1\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^y_1\hat{\mathcal{U}}(\tau)\rangle) \end{aligned} \tag{4.15}$$

where

$$\hat{\mathcal{U}}(\tau) = \exp\Big(-\mathrm{i}\hat{H}_{\mathrm{SIAM}}\tau/\hbar\Big) \tag{4.16}$$

is the time evolution operator. Each term of the form $\langle\hat{\sigma}^\alpha_1\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^\beta_1\hat{\mathcal{U}}(\tau)\rangle$ can be obtained through evaluating two single-qubit Ramsey interferometry circuits [130], allowing the calculation of Eq. (4.9) with the result of 16 different quantum circuits. However, through considering the conditions $G^{>}_{imp}(\tau) = G^{<}_{imp}(\tau)^*$ and $\mathrm{Im}[iG^R_{imp}(\tau)] = 0$ that occur in the half-filling case (see Appendix B.1), in this thesis we are able to reduce Eq. (4.9) to simply

$$\mathrm{i}G^R_{\mathrm{imp}}(\tau) = \mathrm{Re}[\langle\hat{\sigma}^x_1\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^x_1\hat{\mathcal{U}}(\tau)\rangle], \tag{4.17}$$

which can be evaluated by a single quantum circuit, shown in Fig. 4.4. This circuit, based on the findings of [160], constructs the expectation value $\mathrm{Re}[\langle\hat{\sigma}^x_1\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}^x_1\hat{\mathcal{U}}(\tau)\rangle]$ through repeated measurements of the ancilla qubit in the $\hat{\sigma}_z$ basis, as detailed in Appendix B.2. Thus, measurements of this circuit directly give us $\mathrm{i}G^R_{\mathrm{imp}}(\tau)$ and as such we call it the Green's function circuit. Notably, obtaining the impurity Green's function with this circuit requires measuring only one qubit, which is true

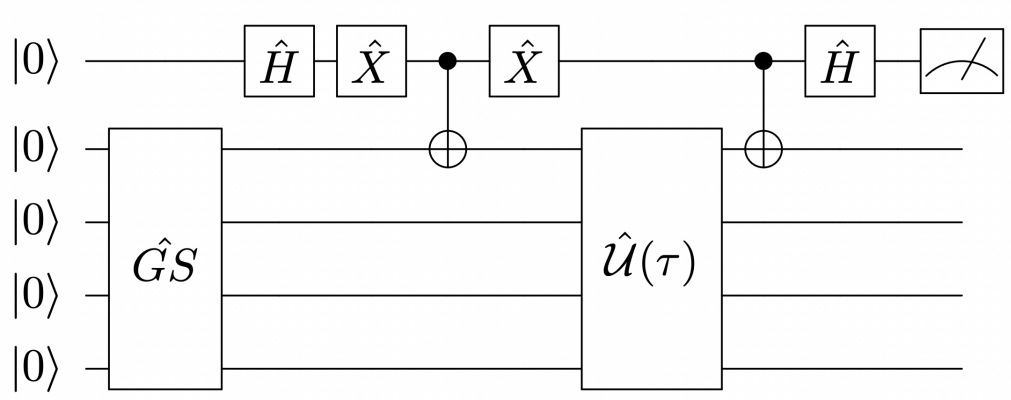


Figure 4.4: Quantum circuit used to calculate the expectation value $\langle\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}_1^x\hat{\mathcal{U}}(\tau)\rangle$. The work qubits are first prepared into the ground state of the SIAM Hamiltonian using the sub-circuit $\hat{GS}$. They are then acted on by entangling gates with the ancilla qubit and the time evolution operator $\hat{\mathcal{U}}(\tau)$. The ancilla qubit itself undergoes single-qubit Hadamard $\hat{H}$ and bit-flip $\hat{X}$ gates. Repeated measurements of the ancilla in the $\hat{\sigma}_z$, $\hat{\sigma}_y$ bases build up the real and imaginary parts of the expectation value respectively.

even as we increase the number of bath sites in the impurity model.

In order to represent the time-evolution operator $\hat{\mathcal{U}}(\tau)$ in Eq. (4.17) in terms of quantum logic gates, we approximate it by a first order Suzuki-Trotter decomposition [161] as shown in [150]. By executing the Green's function circuit several times with different numbers of Trotter steps, we numerically reconstruct $iG^R_{\rm imp}$ as a function of τ. The circuit $\hat{GS}$ in Fig. 4.4 which prepares the ground state of the SIAM Hamiltonian can be obtained via arbitary state preparation techniques [162].

4.2.5 Noise model

Next we describe the noise model that we use in our simulations of NISQ devices presented in Sec. 5.3. Our model, implemented using Qiskit, accounts for both imperfections in qubits and gates. It is applied to all operations allowed in our

emulator, made up of the $\hat{U}_1$, $\hat{U}_2$ and $\hat{U}_3$ single-qubit gates (see Appendix B.3), the CNOT two-qubit gate and measurement.

Firstly, when an operation is applied to a qubit, we model the qubit to undergo thermal relaxation based on its lifetimes T_1, T_2 and the gate time, where T_1 and T_2 are the relaxation and dephasing time constants respectively. For simplicity, we set $T_1 = T_2 = T$ here and estimate operation times using guidance from both the literature [163, 164, 165] and example noise models given in Qiskit. To calculate the probability of thermal relaxation during a two-qubit gate, we tensor product the single-qubit error channels of each of the two qubits involved.

Secondly, we model the imperfections of quantum gates using a depolarizing quantum error channel [166]. When applied to a single qubit, this has the form

$$\epsilon(\rho) = (1-\lambda)\rho + \frac{\lambda}{3}\left(\sigma^x\rho\sigma^x + \sigma^y\rho\sigma^y + \sigma^z\rho\sigma^z\right), \tag{4.18}$$

where ρ is the density matrix of the qubit. The physical interpretation of this error channel is that when a gate is applied, an additional Pauli operation occurs with probability λ. The depolarizing channel is often used to characterise quantum noise [167, 168], particularly as a worst case scenario where we have little information about the true noise channels, which makes it an apt description of NISQ devices. We subsequently implement the depolarizing channel for both single and two-qubit gates.

We combine the thermal relaxation and depolarizing error channels to produce a realistic emulation of noisy quantum computers [169, 79]. From this, individual fidelities can be extracted for any operation - including single-qubit gates, two-qubit gates and measurements. For example, Figure 4.5 shows the infidelity of the CNOT gate, as a function of the noise model parameters. We see that if the depolarizing error is negligible, (i.e., small values of λ), the gate infidelity only

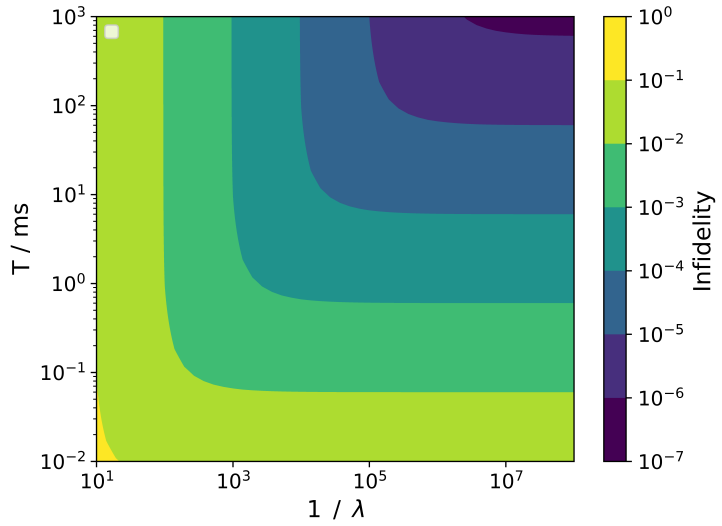


Figure 4.5: CNOT gate infidelity for different noise model parameters. The qubit lifetime T is used to calculate the probability of thermal relaxation occurring. Additionally, a two-qubit depolarizing channel is applied with probability parameter λ. For this particular noise model, we set the CNOT gate time to be 300ns. Note that the color coding of the infidelity utilises a discretized logarithmic scale.

depends on qubit lifetime. Conversely, in the limit of very long qubit lifetimes T, the depolarizing error becomes the dominant source of error. Moreover, we find that in this case the gate infidelity is equal to the value of the depolarizing parameter λ. It follows that achieving high fidelity requires a combination of both long qubit lifetime and low depolarizing error probability.

4.3 Results

We now implement the DMFT routine described in section 4.2 in Python, constructing the relevant quantum circuits in Qiskit. In section 4.3.1, we find the minimum number of Trotter steps required to reproduce the analytic two-site DMFT solution and consider the number of shots of the Green's function circuit required to mitigate statistical errors. We use these results to subsequently find the lowest

gate fidelities that can produce accurate results compared to the noiseless solution. In section 4.3.2 we apply incremental structural learning (ISL), our circuit recompilation algorithm, and compare by how much the minimum hardware requirements change.

4.3.1 Fidelity requirement of original scheme

We run the full DMFT scheme described in Sec. 4.2 using a noiseless statevector simulator for different numbers of Trotter steps. The results, seen in Fig. 4.6, show excellent agreement of the converged quasiparticle weight Z to the analytic solution [155], particularly in the conducting phase at $U < 3.0t^*$. As we approach the metal-insulator phase transition at $U = 6.0t^*$,our hybrid algorithm underestimates the quasiparticle weight and in the cases of $\eta = 24$ and $\eta = 36$ Trotter steps, incorrectly identifies where the transition occurs. This is an expected consequence of the approximations made during a Trotter decomposition and can be rectified by increasing the number of Trotter steps. Indeed, for $\eta = 48$ Trotter steps we see excellent agreement to the analytic solution, even at the phase transition.

To minimise circuit depth, we now focus on the $\eta = 24$ Trotter steps case, which still provides accurate results in the range $2.0t^* < U < 3.5t^*$. In order to apply our noise model to the simulated hardware, we must first switch from using a statevector simulator to a measurement-based one. In doing so, we add a source of error to our simulation in the form of shot noise (i.e., the number of measurements required to build up an accurate expectation value). Through experimentation, we find 75,000 shots to be sufficient for the statistical error to be less than the error generated by our noise models. This is well within the capabilities of NISQ devices.

We then look to apply our noise model to test the performance of DMFT. Using

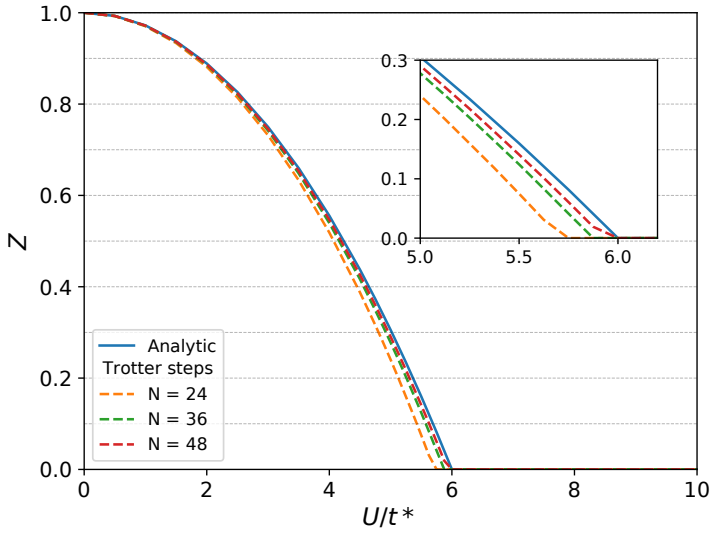


Figure 4.6: Quasiparticle weight Z as a function of interaction strength U and the Hubbard hopping amplitude in infinite dimensions t^*. For a given interaction strength, we iteratively obtain a self-consistent Z for 24, 36 and 48 Trotter steps and compare against the analytic solution. The inset focuses on the region near the critical value $U_c = 6.0t^*$.

Fig. 4.5, we find noise model parameters on the boundary of each infidelity contour such that they correspond to CNOT gate fidelities separated by an order of magnitude each, e.g., 99%, 99.9%, 99.99% and so forth. These parameters, τ, λ, subsequently determine the single-qubit gate and measurement fidelities, which are always larger than the CNOT fidelity. We implement these parameters in our simulations and run DMFT, transpiling all quantum circuits in Qiskit with the "heavy" optimisation option.

As shown in Fig. 4.7, simulations with higher gate fidelities produce quasiparticle weights closer to the noiseless solution, as obtained on the statevector simulator. We observe that a two-qubit fidelity of 99.9% is not sufficient for DMFT to converge consistently, as shown by the absence of a result in the $U = 2.0t^*$ case. Increasing the fidelity to 99.99% (noise parameters A) allows DMFT to converge to a quasiparticle weight within 4% of the exact solution. The full details of these

noise parameters and all others referenced in this section are shown in Table 4.1.

The error bars shown in Fig. 4.7 are determined as follows. We find that in the presence of noise, DMFT oscillates around the self-consistent solution without settling to a finite value. To account for this, after the self-consistent threshold is met [170] we run 50 additional iterations and take the average quasiparticle weight as our solution. We then use the standard deviation σ of these iterations to produce the error bars in Fig. 4.7 of size 2σ.

The low-fidelity results in Fig. 4.7, shown by the blue and orange data points, demonstrate that the magnitude of the error bars does not fully account for the deviation from the exact solution. We find that the applied noise channels dampen the oscillations of the impurity Green's function, restricting its ability to represent the analytically correct solution. If the gate error is too large, and the impurity Green's function too misshapen, the quasiparticle weight at each DMFT iteration will jump too much to converge. However, subsequent steps may fall within the self-consistency tolerance [170] at a larger incorrect solution $Z > Z_0$. This is because of the non-linear relationship between V_{new} and V, evident when substituting the self-consistency condition Eq. (4.5) into Eq. (4.7). In particular, we find $V_{\mathrm{new}} \propto V^{-2}$ and thus large values of V result in smaller step sizes, which explains the spurious convergence observed for the low-fidelity results in Fig. 4.7.

The high-fidelity result, with two-qubit gate fidelity of 99.999% (noise parameters B) are shown by the green data points in Fig. 4.7. We find that this represents the maximum noise that can be tolerated whilst reproducing the statevector simulator within the noise-induced error bars.

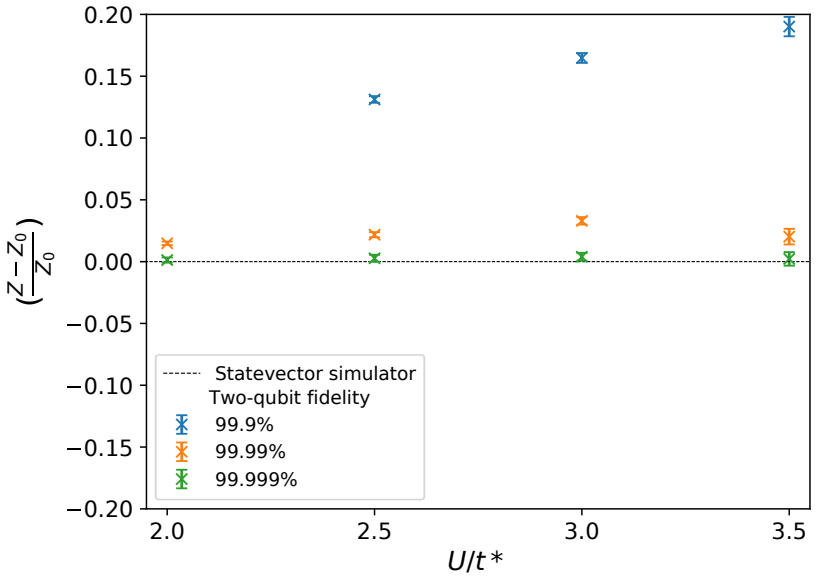


Figure 4.7: Relative quasiparticle weight as a function of on-site interaction strength U for different two-qubit gate fidelities. Here Z_0 is the 24 Trotter step statevector simulator result shown in Fig. 4.6. For definition of error bars see text.

4.3.2 Circuit reduction using incremental structural learning

The total noise incurred in the execution of a quantum circuit scales exponentially with the number of logic gates. Therefore, we focus our efforts on lowering the fidelity requirements of quantum DMFT by reducing the number of gates that make up the Green's function circuit, shown in Fig. 4.4. This is achieved in two different ways. First, we use the variational quantum eigensolver [49] (VQE) to find an approximate representation of the circuit $\hat{GS}$ which prepares the ground state of the SIAM Hamiltonian. In this way, we reduce the depth of $\hat{GS}$ from 72 using exact initialisation techniques to 4. Furthermore, another advantage of this approach is that the ground state of the SIAM Hamiltonian does not need to be calculated on a classical computer and its initialisation decomposed into logic gates, both of which are classically intractable.

Secondly we use incremental structural learning (ISL), the AQCR algorithm described in chapter 2, to reduce the depth of the full Green's function circuit. Here, the complexity of the recompilation problem is far greater than the previous chapter for two reasons. Firstly, since ISL requires the input state to be $|\psi_0\rangle = |0\rangle^{\otimes n}$, the recompilation target includes both the VQE ground state solution and the initial gates applied to the ancilla qubit for the Hadamard test, in addition to the unitary evolution operator $\hat{\mathcal{U}}(\tau)$. Once an approximate equivalent is found for this part of the circuit, for hybrid quantum-classical DMFT we then need to append the final ancilla operations to create the full Green's function circuit. Secondly, since different Hamiltonian parameters are used for each DMFT iteration, and subsequently a different set of circuits are run to produce the impurity Green's function, the recompilation procedure needs to be repeated for each DMFT iteration.

Furthermore, in this chapter we introduce an extension to the ISL routine that address the issue of its scalability to intractable systems. Specifically, ISL requires the repeated execution of a circuit containing the recompilation target followed by the current best guess of the inverse $\hat{V}^\dagger \hat{U}|\psi_0\rangle$, where $\hat{V}^\dagger$ contains between 0 and L layers of thinly-dressed CNOT gates. As discussed in chapter 2, in this thesis we evaluate this circuit on emulated quantum computers that store the statevector, allowing optimal convergence to a shallow recompiled solution in the absence of noise. Yet working with the explicit form of the quantum state vector is a luxury only afforded when studying classically tractable systems, which by definition could be solved without quantum computers. The natural remedy to this would be to run the circuit $\hat{V}^\dagger \hat{U}|\psi_0\rangle$ on real quantum hardware, yet it is clear that we wouldn't be able to evaluate this accurately, otherwise we would simply run the original $\hat{U}|\psi_0\rangle$ and do away with recompilation.

To solve this, we introduce a different way of using ISL. Rather than recompile the

entire time evolution operator $\hat{\mathcal{U}}(\tau)$ at once, we break it into its individual Trotter steps and recompile each one sequentially to obtain the impurity Green's function at each increment. To further clarify, this process begins by first recompiling the action of a single Trotter step $\hat{TS}$ on the initial state $|\psi_0\rangle = |0\rangle^{\otimes n}$. This produces a recompiled circuit $\hat{RC}_1$, which prepares the correct state after one time step to within the specified ISL overlap such that $\hat{RC}_1 |\psi_0\rangle \approx \hat{TS} |\psi_0\rangle$. Subsequently, we wish to find a recompiled solution for the evolution of two Trotter steps. Rather than recompile the circuit $\hat{TS}\hat{TS} |\psi_0\rangle$, we use our approximate representation of the action of the first Trotter step, instead recompiling the circuit $\hat{TS}\hat{RC}_1 |\psi_0\rangle$. This produces the recompiled circuit that approximately prepares the evolution after two time steps $\hat{RC}_2 |\psi_0\rangle$. Generally, for the $\eta + 1$ Green's function circuit, where η is the number of Trotter steps, we use the ISL solution of the previous η Green's function circuit and add one exact Trotter step to create the recompilation target $\hat{U}$. This allows us to "ladder" recompiled circuits together to obtain late-time dynamics, as illustrated in Fig. 4.8, in a process we name ladder-ISL. This strategy was also developed independently in another work, where it is given the name restarted quantum dynamics [171].

Although more complicated than running one recompilation of the entire time evolution, ladder-ISL has considerable benefits when considering the depth of the circuits that need to be run during the ISL procedure. Whereas regular ISL would require evaluation of $\hat{V}^\dagger \hat{\mathcal{U}}(\tau) |\psi_0\rangle$, the full 24 Trotter step circuit plus the current best guess of the recompilation solution, ladder-ISL needs to only evaluate $\hat{V}^\dagger \hat{TS}\, \hat{RC}_\eta |\psi_0\rangle$. Since by definition ISL aborts if the solution is deeper than the target circuit, we can define that the depth $\mathcal{D}$ follows the relations

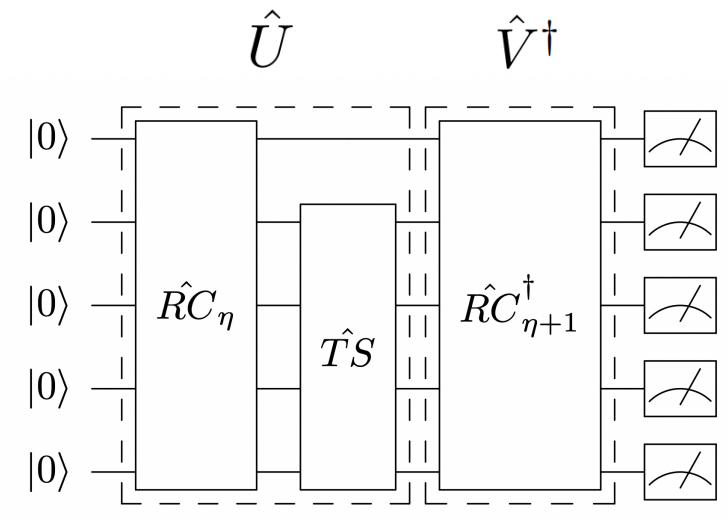


Figure 4.8: Laddering the recompiled solutions of the Green's function circuit at each time step allows the calculation of $\mathrm{i}G^R_{\mathrm{imp}}(\tau)$ for large τ without evaluating deep quantum circuits. A recompilation target $\hat{U}$ for the $\eta+1$ Trotter steps Green's function circuit is created by adding one exact Trotter step to $\hat{RC}_\eta$, the recompiled circuit for η Trotter steps. A solution is constructed by trying to find a circuit $\hat{V}^\dagger = \hat{RC}^\dagger_{\eta+1}$ that approximately acts as the inverse of $\hat{U}$ but with fewer gates. The ansatz for $\hat{V}^\dagger$ is built up iteratively until the overlap between the output state and the input state $|\psi_0\rangle$ is sufficiently large.

$$\mathcal{D}(\hat{RC}_\eta) < 2\mathcal{D}(\hat{TS})$$
$$\mathcal{D}(V^\dagger) < \mathcal{D}(\hat{TS})$$
$$\text{thus } \mathcal{D}(\hat{V}^\dagger \hat{TS} \hat{RC}_\eta) < 4\mathcal{D}(\hat{TS})$$

By using this iterative approach, the depth of circuits evaluated in the ladder-ISL circuit is never larger than four times the depth of a single Trotter step. Thus, our method provides a way to recompile the target dynamics $\hat{\mathcal{U}}(\tau)\,|\psi_0\rangle$ without the requirement of running the original circuit to begin with. Promisingly, this pro-

vides a scalable path for recompilation to be executed on real quantum computers which are too noisy to evaluate the target circuit. We note that this benefit does not come for free, as it introduces an additional error scaling from the chaining of the approximate solutions together. Furthermore, recompiling each Trotter step individually means that ISL needs to be run to convergence $O(\eta)$ times, as opposed to $O(1)$ previously.

We apply the ladder-ISL method to recompile the Green's function circuits at each of the 24 Trotter steps of evolution used in our DMFT algorithm. Excitingly, the recompilation procedure still leads to significant gate reductions in the Green's function circuit. Once completed, ladder-ISL produces a Green's function circuit containing on average 6 two-qubit gates and 11 single-qubit gates for any number of Trotter steps. This is in contrast to the 24 Trotter step Green's function circuit in the original scheme, which contains 510 two-qubit gates and 752 single-qubit gates.

We rerun hybrid quantum DMFT, using the same noise parameters as in section 4.3.1, but this time applying ladder-ISL to each Green's function circuit. In Fig. 4.9 we show that in this case, a two-qubit fidelity of 98% (noise parameters C) or 99% (noise parameters D) is enough for DMFT to converge within 35% or 10% of the exact solution respectively. Furthermore, we find that a quantum computer with 99.9% two-qubit gate fidelity (noise parameters E) is sufficient to produce results that perfectly agree with the statevector simulator within noise-induced error margins. Therefore, by applying ISL, we see a factor of 100 improvement in the noise tolerance of two-site hybrid DMFT compared to using non-AQCR techniques.

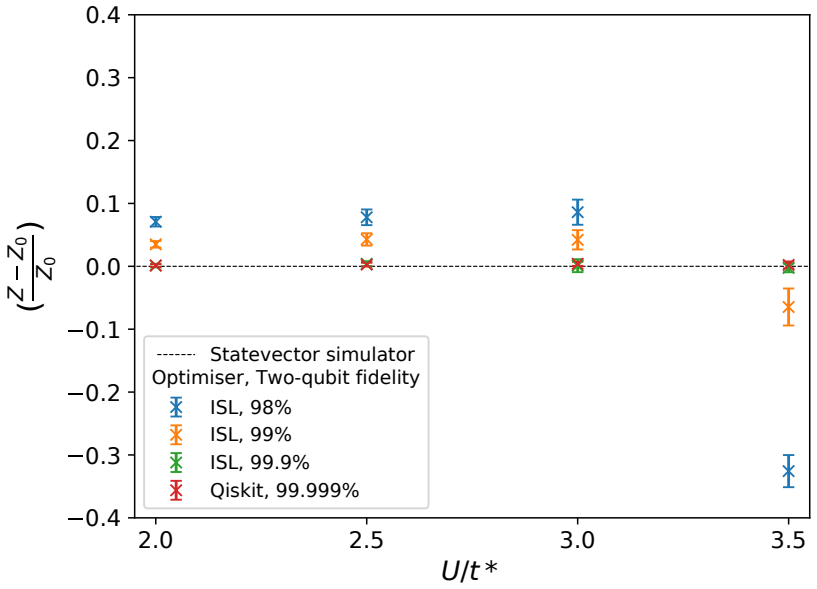


Figure 4.9: Relative quasiparticle weight as a function of on-site interaction strength U for different two-qubit gate fidelities. Using incremental structural learning (ISL), a quantum computer with two-qubit gate fidelity of 99.9% can produce convergent DMFT results with perfect agreement to the noiseless solution.

4.4 Conclusion

In this chapter, we find that hybrid quantum-classical DMFT can be accurately solved within noise-induced error margins, provided quantum hardware capable of executing 75,000 shots, two-qubit gate fidelity of 99.999% and average single-

Noise parameters	T/ms	λ	$\mathcal{F}(U_1)$	$\mathcal{F}(U_2)$	$\mathcal{F}(U_3)$	$\mathcal{F}$(CNOT)	$\mathcal{F}$(Measurement)
A	10	4e-5	0.99997	0.99997	0.99996	0.9999	0.9999
B	100	4e-6	0.999997	0.999997	0.999996	0.99999	0.99999
C	0.04	5e-3	0.996	0.995	0.994	0.980	0.975
D	100	4e-6	0.997	0.997	0.996	0.990	0.991
E	1.1	4e-4	0.9997	0.99974	0.9996	0.999	0.999

Table 4.1: Parameters used to emulate different NISQ devices with our noise model. For a given qubit lifetime T and depolarizing channel probability λ, the corresponding gate fidelities $\mathcal{F}$ can be obtained. The single-qubit gates, $\hat{U}_1$, $\hat{U}_2$ and $\hat{U}_3$ are defined in Appendix B.3.

qubit gate fidelity of 99.9997% (noise parameters B). However, by finding shallow approximations of the Green's function circuits using our ISL recompiler, we show that DMFT can be self-consistently solved by quantum hardware with two-qubit and average single-qubit fidelities of 99% and 99.7% respectively (noise parameters D), within 10% of the exact solution. These results are consistent with those for the implementation of the alternative DMFT scheme in [153], where a solution at U=4 was found within 2.6% of the exact solution on an IBM device using circuit recompilation and SPAM error mitigation methods. Note that the algorithm in [153] requires one to calculate all excited states via the variational quantum eigensolver method, and thus it scales exponentially with the number of sites. Furthermore, we find that increasing the two-qubit and single-qubit fidelities to 99.9% and 99.97% respectively (noise parameters E) allows one to produce results in perfect agreement with the exact solution, within the noise-induced error bounds.

Excitingly, these findings show that our scheme could produce accurate results on noisy quantum computers in the near future. For superconducting qubit architectures, Google's *Sycamore* 53 qubit device has two-qubit and single-qubit gate fidelities of 99.64% and 99.85% respectively [172]. Given that the total noise scales with the number of qubits, these figures suggest that our fidelity requirements could already be met by a smaller, high fidelity, 5 qubit device.

A different perspective can be gained on the capabilities of NISQ computers if we consider quantum volume (QV) instead, as calculated via the protocol defined in [173]. Using randomised circuit benchmarking, we calculate the QV corresponding to noise parameters E to be 32. By comparison, at the time of this experiment the best available device on the IBM quantum experience had a QV 16, with an impending release of the 28 qubit *ibmq_raleigh* device with the highest ever measured QV 32. Given that at the time of writing this thesis there exists three IBM

quantum computers with QV 128, as well as hardware from other providers with claimed QV 4096 [174], it is highly likely that our scheme is runnable on real quantum computers at present.

Looking forward, an open problem remains to determine the fidelity requirements for hybrid quantum-classical DMFT with more than just two sites. This is particularly true for achieving a quantum advantage, which would require more than 25 bath sites (50 qubits). Since the additional bath sites would increase the depth of the circuit implementation of the time evolution operator, our results for two-site DMFT establish a lower bound for this ultimate goal. Furthermore, whilst the scalability of variational algorithms such as VQE and ISL remains an open question, the number of gates in our scheme grows sub-exponentially with the number of DMFT sites. In this way, hybrid quantum-classical DMFT may prove to be another candidate for displaying quantum advantage before the era of fault-tolerant qubits.

5 | Quantum Self-Supervised Learning

In the previous two chapters of this thesis, we have investigated the application of near-term quantum computing to problems in quantum chemistry. In this final research chapter we consider something different, namely the use of VQAs as machine learning models. Here not only is the problem domain different, but the ansatzes are no longer prescriptive, instead built as a general trainable structure that has led to the nomenclature quantum neural networks (QNNs). The flexibility of this approach has led to the use of QNNs to solve an enormous range of machine learning problems, yet in most of these cases an analytic form of the solution distribution is not known. Thus, QNNs have a notable lack of guarantees for the classical hardness of the problem they are solving. In this chapter, we approach this problem from the other direction, considering one of the computationally hardest problems in classical deep learning and investigating how QNNs can be used to help. In proof-of-concept experiments, we then achieve the minimum threshold for a potential future quantum advantage, numerically comparing the performance of QNNs to equivalently sized classical networks.

5.1 Introduction

In the past decade, machine learning has revolutionised scientific analysis, yielding breakthrough results in protein folding [175], black hole imaging [176] and heart disease treatment [177]. At the forefront of this progress is deep learning [178], characterised by the successive application of artificial neural network layers [179, 180]. Notably, its use in computer vision has seen the top-1 accuracy on benchmark datasets such as ImageNet soar from 52% [181] to over 90% [182], fuelled by shifts in the underlying techniques used [183, 184]. However, what

has remained consistent in these top performing models is the use of labelled data to supervise the representation learning process. Whilst effective, the reliance on large quantities of human-provided annotations presents a significant challenge as to whether such approaches will scale into the future. Crucially, modern datasets such as the billions of images uploaded to social media are both vast and unbounded in their subject, quickly making the task of labelling unfeasible.

This has reignited interest in an alternative approach, termed *self-supervised learning* [185], which seeks instead to exploit structure in the data itself as a learning signal. Rather than predict human annotations, a model is trained to perform a *proxy task*, that makes use of attributes of the data that can be inferred without labelling. Furthermore, the proxy task should encourage the model to learn representations that capture useful factors of variation in the visual input, such that solving it ultimately correlates with solving tasks of interest after training. Recent progress in the self-supervised learning of visual data has been driven by the success of contrastive learning [186, 187, 188, 189, 190], in which the proxy task is differentiating augmented instances of the same image from all other images. Provided the correct choice of augmentations, this produces a model which is invariant to transformations that do not change the semantic meaning of the image, allowing the learning of recognisable features and patterns in unlabelled datasets.

With these techniques, contrastive learning is able to learn visual representations with comparable quality to supervised learning [191, 190], without the bottleneck of labelling. However, it is a fundamentally more difficult task than its supervised counterpart [192], and capturing complex correlations between augmented views requires more training data, more training time and larger network capacity [188, 189]. Therefore, it is important to consider whether emerging technologies can contribute to the growing requirement for more powerful neural networks [193].

In this chapter we consider the potential of VQAs as one such new paradigm. While VQAs have been used to solve many types of optimisation problems [194, 195, 196], it is their application to supervised learning [54, 197, 198], unsupervised learning [55], generative models [199, 200] and reinforcement learning [201, 202, 203] which has led to them being referred to as quantum neural networks (QNNs) [204, 205, 206]. In theory, the power of these models comes from their access to an exponentially large feature space [197] and ability to represent complex high-dimensional distributions, as formalised by the effective dimension [207]. Importantly, early evidence suggests that quantum models can achieve an advantage over their classical counterparts, yet these works focus only on the supervised learning of either artificial data [205, 208] or simple historical datasets [207]. For example, whilst widely used to study QNNs [209, 210], classical supervised learning of MNIST can already achieve 99.3% top-1 accuracy with a two-layer 784-800 width multi-layer perceptron (MLP) [211]. Thus, it is highly unlikely that this problem would practically benefit from a quantum model with access to a $> 2^{50}$ dimensional feature space and careful consideration should be made about whether supervised learning is the best setting to try to achieve quantum advantage. By comparison, self-supervised learning of ImageNet with the widely-used ResNet 50 architecture [212] (with maximum channel width 2048) achieves only 76.5% top-1 accuracy [191]. The necessity for large capacity models means that self-supervised learning may be a better setting in which to seek useful quantum advantage through quantum neural networks.

In this chapter, we construct a contrastive learning architecture in which classical and quantum neural networks are trained together. By randomly augmenting each image in the dataset, our hybrid network learns visual representations which groups different views of the same image together in both classical and Hilbert space. Afterwards, we test the quality of the representations by using them to

train a linear classifier, which then makes predictions on an unseen test set. We find that our hybrid encoder, constrained in both size and training time by quantum simulation overheads, achieves an average test accuracy of $(46.51 \pm 1.37)\%$. In contrast, replacing the QNN with a classical neural network of equivalent width and depth results in a model which obtains $(43.49 \pm 1.31)\%$ accuracy. Thus, our results provide the first indication that a quantum model may capture the complex correlations required for self-supervised learning better than classical neural networks, albeit in the uncharted territory of how to equitably compare such models.

We then apply the best performing quantum model to classify test images on a real quantum computer. Notably, the accuracy achieved using the *ibmq_paris* [20] device equals the best performing classical model, despite significant device noise. This illustrates the capability of our algorithm for real-world applications using current devices, with flexibility to assign more of the encoding to QNNs as quantum hardware improves. While further research is required to demonstrate scalability, our scheme provides a strong foundation for quantum self-supervised learning. Excitingly, given that contrastive learning has also been successfully applied to non-visual data [189, 213, 214, 215, 216], our work opens the possibility of using QNNs to learn large, unlabelled datasets across a range of disciplines.

5.2 Method

5.2.1 Contrastive learning architecture

Given an unlabelled dataset, the objective of self-supervised learning is to find low dimensional encodings of the images which retain important higher level features. In this chapter, we train a model to do this by adapting the widely used SimCLR algorithm [191], the steps of which can be seen in Fig. 5.1. Firstly for a given

image, the data of which is contained within $\vec{x}_i$, we generate two augmentation functions. Each one randomly crops, rotates, blurs and colour distorts the picture, such that two augmented views $\vec{x}_i^1$, $\vec{x}_i^2$ of the same base image are produced. Importantly, these augmentations still allow for the underlying object to remain visually distinguishable. This enables us to assert that these two views contain a recognisable description of the same class, which we call a positive pair.

Once this positive pair is generated, each view is passed through a set of neural networks. First, an encoder network is applied, which maps the high dimensional input data $\vec{x}_i^1$, $\vec{x}_i^2$ to low dimensional representations $\vec{y}_i^1$, $\vec{y}_i^2$. Then the output of the encoder network is passed to the projection head, a small multi-layer-perceptron (MLP) [217] consisting of two fully connected layers. This produces the final representations $\vec{z}_i^1$, $\vec{z}_i^2$.

Given a batch of $\mathcal{N}$ images, the above process is repeated such that we are left with $2\mathcal{N}$ representations corresponding to $2\mathcal{N}$ augmented views. Looking at all possible pairings of these representations, we have not only positive pairs (e.g., $\vec{z}_i^1$, $\vec{z}_i^2$) but also negative pairs (e.g., $\vec{z}_i^1$, $\vec{z}_k^1$ where $i \neq k$), which we cannot definitely say contain the same class. For each training step, all of these possible pairs are used to calculate the normalised temperature-scaled cross entropy loss (NT-Xent) [218]. For a given positive pair $\mathcal{P}_i = \{ \vec{x}_i^1, \vec{x}_i^2 \}$, this is defined as

$$\mathcal{L}_i = -\log \frac{\exp\left(\vec{z}_i^1 \cdot \vec{z}_i^2 / \tau\right)}{\sum_{\substack{k=1 \\ \alpha \in \{1,2\}}}^{\mathcal{N}} \mathbb{1}_{[k \neq i]} \exp\left(\vec{z}_i^1 \cdot \vec{z}_k^\alpha / \tau\right)}, \tag{5.1}$$

where $\vec{z}_i^1 \cdot \vec{z}_k^\alpha$ is the cosine similarity between the two vectors, $\mathbb{1}_{[k \neq i]}$ is an indicator function evaluating to 1 iff $k \neq i$ and τ is a temperature hyperparameter used to control how conservative the model is. Subsequently, the total loss is calculated as the sum over all positive pairs $\mathcal{L} = \sum_{i=1}^{\mathcal{N}} \mathcal{L}_i$. Intuitively, minimising

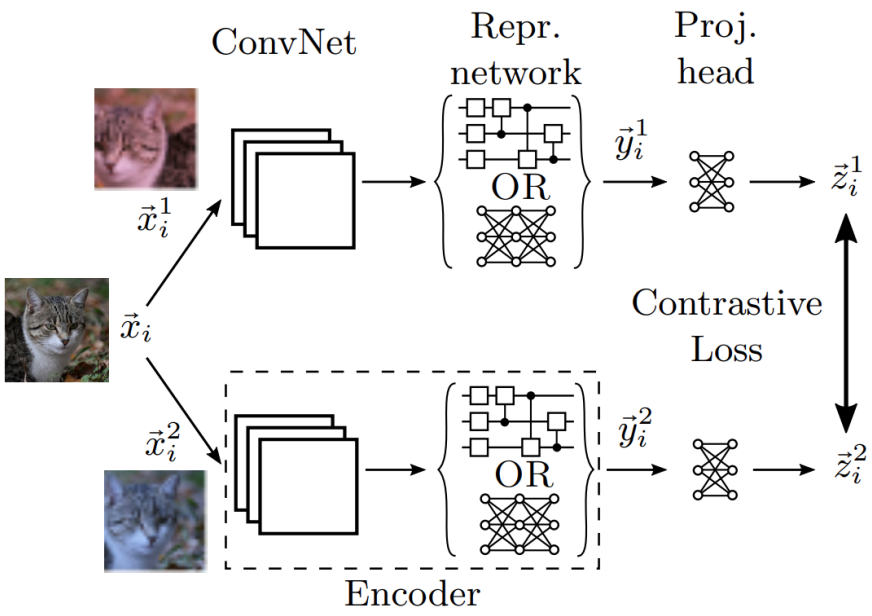


Figure 5.1: Schematic of the overall neural network architecture and contrastive training method. For each input image $\vec{x}_i$, a pair of random augmentations are generated and applied to form a positive pair $\vec{x}_i^1$, $\vec{x}_i^2$. These are transformed by the encoder network, consisting of classical convolutional layers and a quantum or classical representation network, into representation vectors $\vec{y}_i^1$, $\vec{y}_i^2$. The projection head subsequently maps the representations to the vectors $\vec{z}_i^1$, $\vec{z}_i^2$, such that contrastive loss can be applied without inducing loss of information on the encoder.

this loss function, which is done through stochastic gradient descent [219], can be understood as training the network to produce representations in which positive pairs are mapped close together and negative pairs far apart, as measured by their cosine similarity. This idea is a core concept in contrastive learning and many machine learning techniques [220]. Note that whilst it is possible to train the network by applying NT-Xent directly to the output of the encoder, the contrastive loss function is known to induce loss of information on the layer it is applied to [191]. Therefore, the addition of the projection head ensures that the encoder remains sensitive to image characteristics (e.g., colour, orientation) that improves performance on downstream tasks.

In order to incorporate QNNs, we modify the encoder to contain both classical and quantum layers working together. The first part of the encoder consists of a convolutional neural network, specifically the widely used ResNet-18. This produces a 512 length feature vector, which is already an initial encoding of the augmented image. However, we then extend the encoder with a second network, which we call the representation network as it acts directly on the representation space. This consists of either a multi-layer QNN of width W, or a classical fully connected MLP with equivalent width and depth. Ideally the representation network would have width $W = 512$, so as to minimise loss of information. However, we instead look to work in a regime which is realisable on current quantum computers, and as such here we use $W = 8$. This is achieved by following the convolutional network with a single classical layer that compresses the vector, a common technique used to link classical and quantum networks together [221, 222]. After the representation network is applied, the resultant encoding is passed onto the previously described projection head. To maintain the structure of the original SimCLR architecture, we limit the projection head to be no wider than the width of the QNN.

5.2.2 Quantum representation network

The quantum representation network follows the structure shown in Fig. 5.2, beginning with a data loading unitary $\hat{D}(\vec{v})$. Whilst schemes exist to encode data into quantum circuits with exponential compression [223, 224], these require a prohibitively large number of logic gates compared to current hardware capabilities. By compressing the output of the ConvNet as described in section 5.2.1, we need only to solve the simpler issue of loading a vector $\vec{v}$ of length W into equally as many qubits. This is achieved by applying a single qubit rotation $\hat{R}_x$ to each qubit in the register; $\hat{D}(\vec{v}) = \bigotimes_{i=1}^{W} \hat{R}_x(v^k)$. Here, v^k is the kth element of input

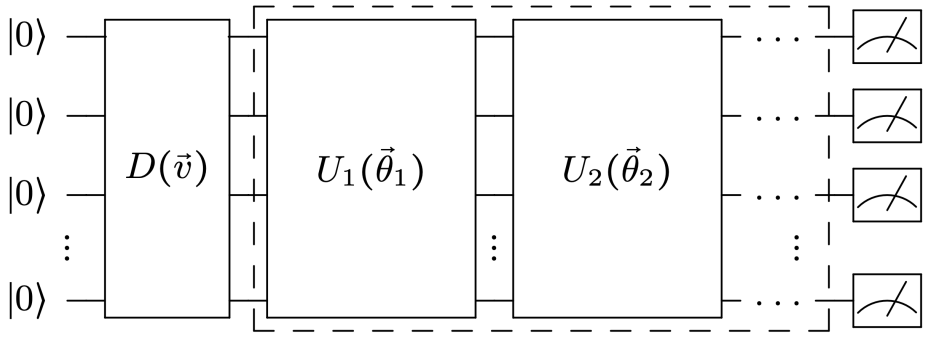


Figure 5.2: General structure of a QNN. An input vector $\vec{v}$ is encoded into the qubits by a data loading unitary $\hat{D}(\vec{v})$. The variational ansatz consists of layers $\{\hat{U}_1(\vec{\theta}_1), \hat{U}_2(\vec{\theta}_2), \dots\}$ and is parameterised by trainable parameters $\{\vec{\theta}_1, \vec{\theta}_2, \dots\}$. The output of the QNN is taken as the average of repeated measurements in the $\hat{\sigma}_z$ basis.

vector $\vec{v}$ and is mapped to the range $[0, \pi]$ to prevent large values wrapping back around the Bloch sphere.

Once the input data is loaded, we apply the learning component of our QNN, a parameterised quantum circuit ansatz. In applications where the ansatz is used to solve optimisation problems relating to a physical system (e.g., the simulation of molecules), the circuit structure and choice of logic gates can be inspired by the underlying Hamiltonian [64]. However, without such symmetries to guide our choice, we use a variational ansatz based on recent theoretical findings in expressibility and entangling capability [225]. The ansatz is shown in Fig. 5.3, the structure of which is derived from circuit 14 of Ref. [225] and was chosen due to its performance in both these metrics.

After the application of several ansatz layers, the network is finished by measuring each qubit to obtain an expectation value in the $\hat{\sigma}_z$ basis. When evaluated on a real quantum computer or sampling-based simulator, the expectation value is constructed by averaging the sampled eigenvalues over a finite number of shots. If evaluated on a statevector simulator, the expectation value is calculated exactly.

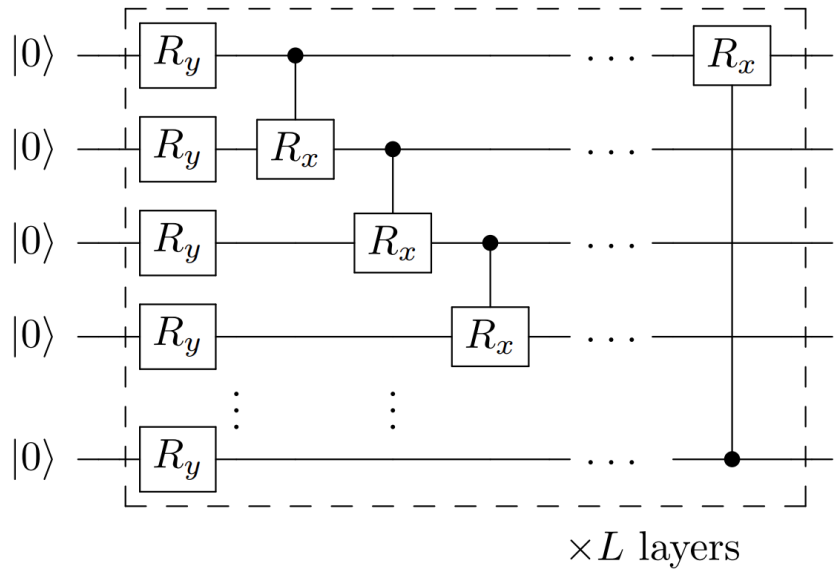


Figure 5.3: Variational ansatz used in this chapter. Each layer consists of a single qubit $\hat{R}_y$ rotation on each qubit, followed by controlled $\hat{R}_x$ rotations, connecting the qubits in a ring topology. Every rotation gate is parameterised by a different variational parameter.

The gradients of the QNN output with respect to the trainable parameters and the input parameters are calculated using the parameter shift rule [204, 226], which we describe here. Consider an observable $\hat{O}$ measured on the state

$$\left|\psi(\vec{\theta})\right\rangle = \prod_i \hat{U}_i(\theta_i)\hat{V}_i \left|0\right\rangle, \tag{5.2}$$

resulting from the application of M parameterised gates $\hat{U}_1, \hat{U}_2, \ldots, \hat{U}_M$ and M fixed gates $\hat{V}_1, \hat{V}_2, \ldots, \hat{V}_M$, where gates $\hat{U}_i = e^{i\theta_i \hat{P}_i/2}$ are generated by operators $\hat{P}_i \in \{\mathbb{1}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}^{\otimes n}$ that are tensor products of the Pauli operators. According to the parameter shift rule, the gradient of the expectation value $f = \left\langle \psi(\vec{\theta}) \right| \hat{O} \left| \psi(\vec{\theta}) \right\rangle$ with respect to parameter θ_i is given by

$$\frac{\partial f(\vec{\theta})}{\partial \theta_i} = \frac{1}{2}\left[f\left(\theta_i + \frac{\pi}{2}\right) - f\left(\theta_i - \frac{\pi}{2}\right)\right]. \tag{5.3}$$

For each parameterised gate within the circuit, including both the variational ansatz and data loading unitary, an unbiased estimator for the gradient is calculated by measuring the QNN with the two shifted parameter values given in Eq. (5.3).

Once the QNN gradients have been calculated, we combine them with gradients of the classical components to obtain gradients of the loss function with respect to all trainable quantum and classical parameters via backpropagation [227]. In this way, the QNN is trained simultaneously with the classical networks, and the quality of the gradients produced on quantum hardware play a crucial role in the training ability of the whole network.

5.3 Results

5.3.1 Training

To examine whether the proposed architecture can successfully train, we apply it to the CIFAR-10 dataset [228]. In this preliminary experiment we restrict the dataset to the first two classes, leaving 10,000 32×32 colour images containing either an aeroplane or automobile. We also train this initial model without a projection head, since it is not being used for classification later. The quantum representation network here is a simulated two-layer QNN and is trained together with the classical components from scratch. Practically speaking, training both paradigms together under one optimiser required the building of a custom code library which integrates Qiskit [79] quantum neural networks with the popular PyTorch machine learning framework [229]. The key step in this code is the creation of our custom class `QNet` as a subclass of `nn.Module`, with custom implementations of the `forward()` and `backward()` functions. The `forward()` function is

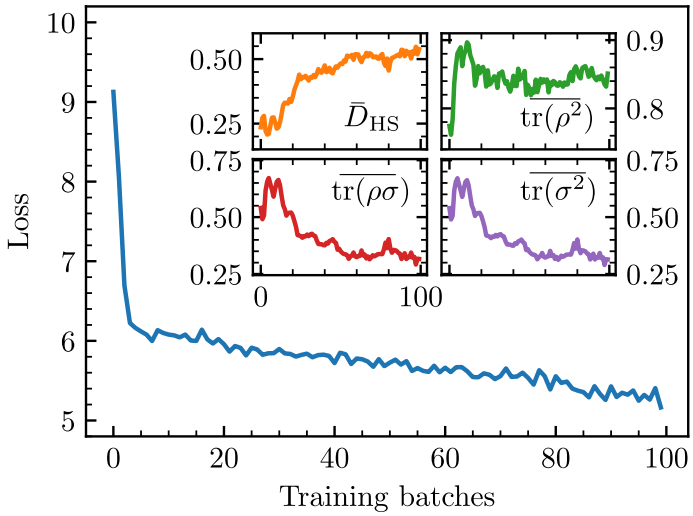


Figure 5.4: Contrastive learning with a quantum representation network. After each batch of 256 images the loss function (main graph) is recorded, alongside the average Hilbert Schmidt distance between positive and negative pairs $\bar{D}_{\mathrm{HS}}$, the average positive pair clustering $\mathrm{tr}(\overline{\rho^2})$, the average clustering of all negative pairs $\mathrm{tr}(\overline{\sigma^2})$ and the ensemble inter-cluster overlap $\mathrm{tr}(\overline{\rho\sigma})$ (insets).

the simplest, assigning the input vector and the PyTorch parameters to the data loading unitary and variational ansatz of the QNN respectively, then evaluating the circuit on the specified Qiskit backend. In the `backward()` pass, we use the aforementioned parameter shift rule to calculate the gradient of the loss with respect to all quantum circuit parameters. These gradient tensors are then batched and passed backwards to the classical layers of our hybrid architecture, continuing the backpropagation process. The library and examples of its usage can be found at `https://github.com/bjader/quantum-neural-network`. Furthermore, in this experiment our training hyperparameters are; batch size: 256, optimiser: ADAM [230] with $\beta_1 = 0.9, \beta_2 = 0.999$, learning rate: 10^{-3}, weight decay: 10^{-6} and softmax temperature: 0.07.

Fig. 5.4 shows the results of several key metrics after training for 100 batches. Firstly, we record the loss after each batch, the minimisation of which repre-

sents the ability to produce representations in the classical W dimensional space whereby positive pairs have high similarity. Our results show that the loss decreases from 9.13 to 5.16 over the course of training, indicating that our model is able to learn. Importantly, since the quantum and classical parameters are trained together, this shows that information is successfully passed both forwards and backwards between these different network paradigms.

Secondly, we log the Hilbert-Schmidt distance (D_{HS}), a metric that has been applied in quantum machine learning previously to study data embedding in Hilbert space [222]. Here, we use it to track the separation between our pseudo classes in the 2^W dimensional quantum state space while optimising the classical loss function. For a given positive pair $\vec{x}_i^1, \vec{x}_i^2$, we calculate the statistical ensembles

$$\rho_i = \frac{1}{2}\left(\left|\psi_i^1\right\rangle\left\langle\psi_i^1\right| + \left|\psi_i^2\right\rangle\left\langle\psi_i^2\right|\right), \tag{5.4a}$$

$$\sigma_i = \frac{1}{2N-2}\sum_{j\neq i}\left(\left|\psi_j^1\right\rangle\left\langle\psi_j^1\right| + \left|\psi_j^2\right\rangle\left\langle\psi_j^2\right|\right), \tag{5.4b}$$

where $|\psi_i^\alpha\rangle$ is the state vector produced by the hybrid encoder given augmented view $\vec{x}_i^\alpha$. The Hilbert-Schmidt distance is then given by

$$D_{\text{HS},i} = ((\rho_i - \sigma_i)^2). \tag{5.5}$$

We repeat this for each positive pair in the batch and record the mean, $\bar{D}_{\text{HS}} = \frac{1}{N}\sum_i D_{\text{HS},i}$. Focusing on the inset of Fig. 5.4, we see in the upper-left panel that $\bar{D}_{\text{HS}}$ increases consistently across the range of training, indicating that the QNN successfully learns to separate positive and negative pairs in Hilbert space. Expanding out the quadratic in Eq. (5.5), we can break down the metric into the

so-called purity terms $\mathrm{tr}(\rho^2)$ and $\mathrm{tr}(\sigma^2)$, which are measures of the intra-cluster overlaps, and the term $\mathrm{tr}(\rho\sigma)$, which is the inter-cluster overlap. Looking at the upper-right panel, we see that the average positive pair clustering $\overline{\mathrm{tr}(\rho^2)}$ increases rapidly at the start of training, before steadying at a value around 0.85. This demonstrates one mechanism by which $\bar{D}_{\mathrm{HS}}$ increases, through the QNN producing representations which group positive pairs close together in Hilbert space. The bottom panels of Fig. 5.4 show the average negative pair clustering $\overline{\mathrm{tr}(\sigma^2)}$ and average negative-positive pair overlaps $\overline{\mathrm{tr}(\rho\sigma)}$, which decrease consistently throughout training. This demonstrates a second behaviour, whereby the QNN produces representations in which negative pairs are well separated. We note that these two values are very similar, which occurs in our self-supervised learning algorithm because of both the need to average over all positive pairs and because of the fixed size of ρ_i. Thus, in the limit $N \to \infty$ the ensemble σ_i contains the entire batch and both metrics are effectively measuring the clustering of all data points.

Overall, Fig. 5.4 shows that the quantum component of the encoder contributes to the overall learning process, despite the network's parameters being optimised explicitly in a classical space. It is notable that the training time presented here is significantly less than classical benchmarks, which would typically be 100s of *epochs*. Due to its technological infancy, executing quantum circuits on real or simulated hardware is computationally expensive. Thus, the 1-2 epochs of training used in this chapter represents the limit of our current experiment. Practically speaking, this limit is set by the number of 8 qubit circuits that can be simulated by Qiskit's `Statevector` or `Qasm` simulator, running on a CPU node of the University of Oxford Advanced Research Computing (ARC) facility [80] for the maximum wall time of 5 days. The choice of Qiskit is significant here, which is chosen in our proof-of-principle experiments due to its feasible integration with PyTorch

and ease of transferring experiments to real hardware. However, benchmarks suggest that running our experiments on more efficient simulators may lead to the ability to evaluate significantly more training epochs [231], furthered still by the imminent public release of GPU-enhanced simulators in the coming years [232]. Nevertheless, our use of Qiskit is sufficient to evaluate at least one epoch of the CIFAR-10 data set, which is a still used as a benchmark for state of the art classical deep learning algorithms [233, 234, 235].

5.3.2 Linear probing

Once training is complete, we require a way to test the quality of the image representations learnt by the encoder. Specifically, a good encoding will produce representations whereby different classes are linearly separable in the representation space [236]. Therefore, we numerically test the encoder using the established linear evaluation protocol [236], in which a linear classifier is trained on the output of the encoder network, whilst the encoder is frozen to stop it training any further. Once this linear probe experiment has trained for 100 epochs, we apply the whole network to unseen test data and record the classification accuracy.

5.4 Classical width ablation

In order to incorporate QNNs that can be run on current quantum devices into contrastive learning, a compression of the feature vector is required after ConvNet. Since this would not be necessary in a purely classical setting, its impact on final performance is not well understood. To this end, we perform a study of the accuracy achieved by models with different representation network widths. We do this with classical representation networks to remove the quantum specific considerations of statistical noise and optimal circuit architecture, focusing purely

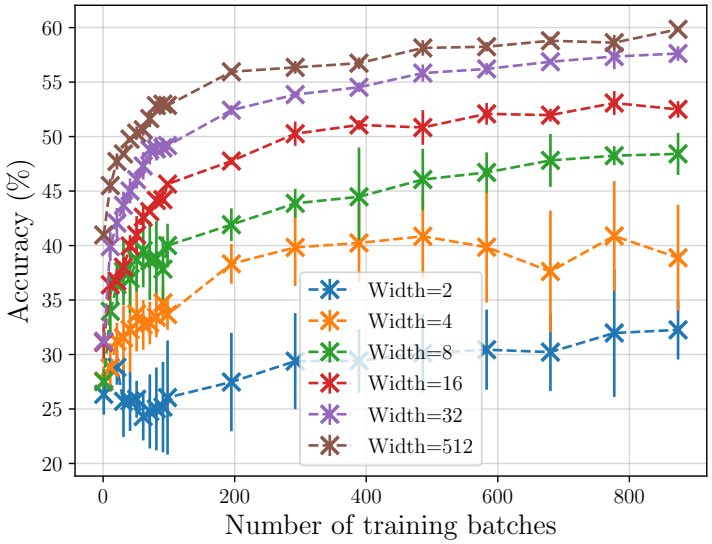


Figure 5.5: Classification accuracy achieved in linear probing experiments by classical representation networks with varying network widths at checkpoints across self-supervised training. The markers show the average of three independently trained self-supervised models and linear probe experiments, whilst the error bars show one standard deviation.

on width. The classical representation network is a two-layer, width W MLP, with Leaky ReLu activation functions after each layer and with bias.

Each model is trained on the first five classes of the CIFAR-10 dataset and a linear probe experiment evaluates the performance at regular checkpoints during training. Fig. 5.5 shows the result comparing models with different representation network widths, including the $W = 512$ case which corresponds to no compression. Starting from $W = 2$, we see that increasing the width of the representation network improves the test accuracy. Furthermore, we find that $W = 8$ is the lowest width network in which test accuracy retains the same qualitative behaviour as the uncompressed network. Therefore, in our proof-of-principle quantum experiments, we use an eight width representation network corresponding to eight qubits.

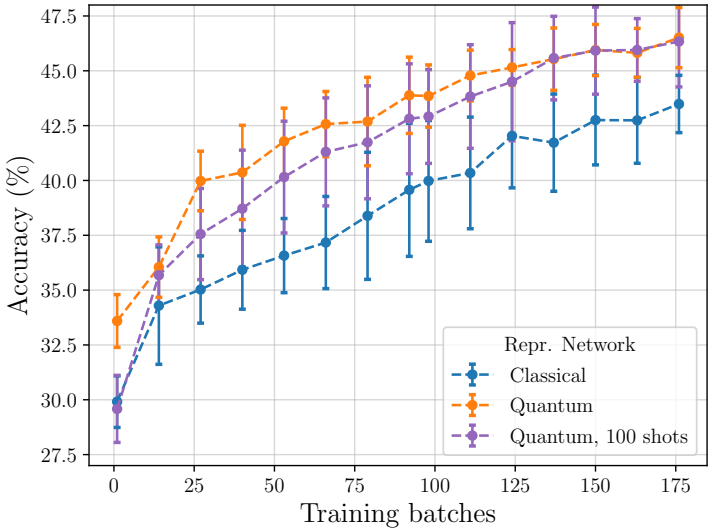


Figure 5.6: Classification accuracy achieved in linear probing experiments using the encoder at checkpoints across self-supervised training. Comparison between models trained with a classical representation network (blue), quantum representation network evaluated on a statevector simulator (orange) and quantum representation network evaluated on a sampling-based simulator with 100 shots (purple). The markers show the average of six independently trained models, whilst the error bars show one standard deviation.

5.4.1 Quantum and classical results on the simulator

We repeat training, this time with the first five classes of CIFAR-10 and a projection head. We train models with three different types of representation networks; classical MLP with bias and Leaky ReLU activation functions after each layer, quantum trained on a statevector simulator and quantum trained on a sampling-based simulator. We choose the representation networks to be width $W = 8$ in order to minimise the simulation overhead, whilst still being in a compression regime where training is stable (see section 5.4). Quantitatively, this means our two-layer classical and quantum representation networks have 144 and 32 learnable parameters respectively..

Fig. 5.6 shows the result of linear probe experiments at checkpoints across 176

batches of contrastive training. We find that when the quantum circuits are evaluated using a statevector simulator, the quantum representation network produces higher average accuracy on the test set than the equivalent classical network at all points probed throughout training, and is separated by more than one standard deviation for over half of these. In these results, the confidence interval corresponds to one standard deviation on the mean of six independently trained models.

Subsequently, we explore whether using a finite number of shots limits this advantage. We train another quantum model on a simulator where the expectation values of measured qubits are sampled from 100 shots, both in the forward pass (generating the representations) as well as the backwards pass (calculating gradients). We find that beyond the first batch, the average accuracy of this model is still above what is achieved by the classical representation network, reaching $(46.34 \pm 2.07)\%$ by the end of training. Significantly, this matches the performance of the statevector simulator, which represents the limit of infinite shots, demonstrating resilience of our scheme to shot noise. However, we note that the additional uncertainty introduced by the sampling does manifest as a larger standard deviation between repeated runs, compromising the consistency of the advantage.

5.5 Quantum and classical results at different widths

In section 5.4.1 we demonstrate that for an architecture with a $W = 8$ representation network, using a QNN to form a hybrid model leads to higher performance in linear probing experiments than the purely classical case. Here we supplement this with additional experiments for the $W = 2$, 4 and 6 cases alongside an equally sized classical comparison for each one. The same problem setup and training parameters are used as in Fig. 5.6.

Fig. 5.7 shows the accuracy achieved by these additional models in linear probing experiments at intervals across training, as well as the $W = 8$ results from the main text. Focusing on the circle markers representing the quantum models, we see that the accuracy improves consistently when increasing the QNN width. This matches the behaviour of the classical models, represented in this figure by the crossed markers, illustrating that our intuition for how compression of the network affects performance can be applied to both the quantum and classical regimes. Secondly, we compare between quantum and classical models of the same width, as shown by the lines of the same colour. Here we see that for the new cases of $W = 2, 4, 6$, there is a numerical improvement in the average accuracy achieved across all training checkpoints sampled, consistent with the $W = 8$ case. Whilst these are still small models, they provide further impetus to consider whether this improvement would remain for models with width $W > 8$, eventually competing directly with the uncompressed SimCLR algorithm at $W = 512$. Looking forward, testing this hypothesis towards the $W = 60$ qubit range may be possible with more efficient simulators [237, 238] as well as by employing training shortcuts such as calculating gradients directly with the quantum state rather than using the parameter shift rule [209].

5.6 Performance of alternative ansatz

In section 5.3 all QNNs are constructed using the variational ansatz seen in Fig. 5.3, which connects the qubits in a ring of parameterised controlled rotation gates. Here we introduce a second ansatz, as seen in Fig. 5.8, which is different in that it connects all of the qubits together and only has single qubit parameterised gates. Notably, this ansatz was recently shown to exhibit a larger effective dimension when applied to supervised learning than equivalent classical networks [207].

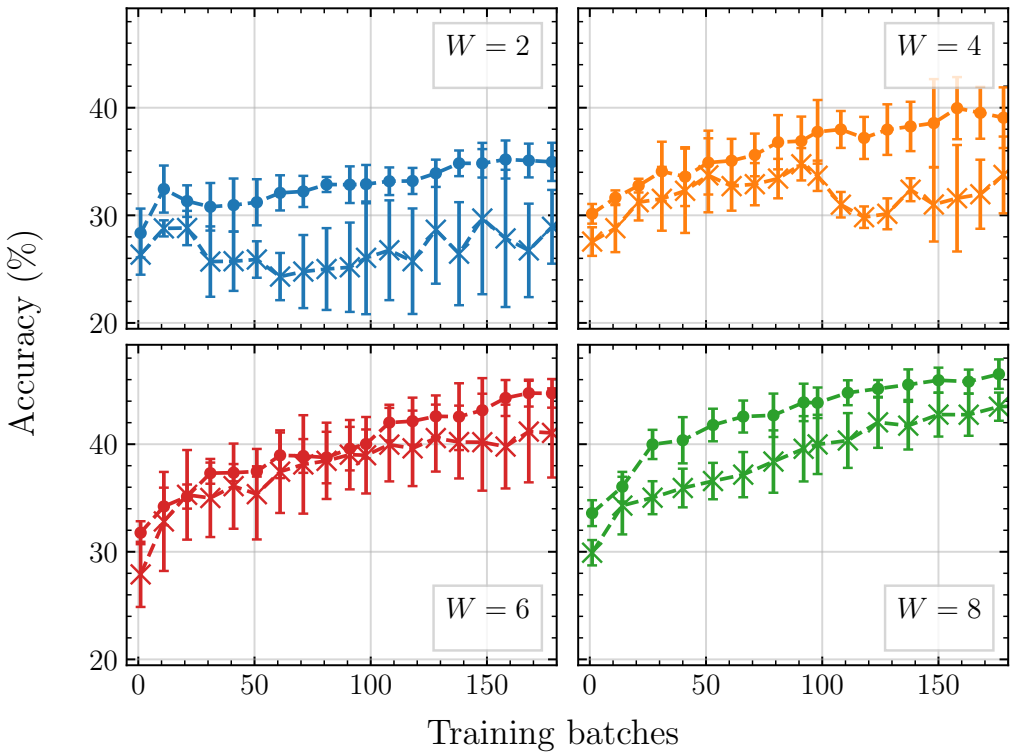


Figure 5.7: Classification accuracy achieved in linear probing experiments for quantum and classical models with representation network widths 2 (upper left), 4 (upper right), 6 (lower left) and 8 (lower right). Within each graph, the circle (crossed) markers represent models trained with quantum (classical) representation networks. The markers show the average accuracy of three independently trained models, with error bars of one standard deviation, except for the $W = 8$ case which displays the same data as Fig. 5.6.

Therefore, we test whether this circuit structure is also a good candidate for improved performance in a self-supervised setting.

We train a model with a quantum representation network structured as the new all-to-all ansatz, simulated on a statevector simulator. The dataset consists of the first five classes of CIFAR-10 and the model is trained with a projection head. Importantly, for a fair comparison, we apply three layers of the all-to-all ansatz, so that it has the same number of learnable parameters as two layers of the ring

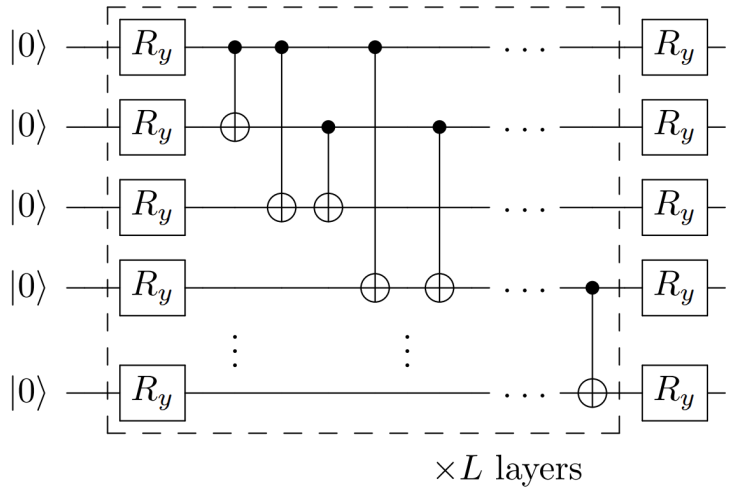


Figure 5.8: Alternative variational ansatz. Each layer consists of a single qubit $\hat{R}_y$ rotation on each qubit, followed by CNOT gates connecting all qubits to each other. After all layers have been applied, a final set of $\hat{R}_y$ rotations are applied. Every rotation gate is parameterised by a different variational parameter.

ansatz. The result of the linear probe experiments can be seen in Fig. 5.9, along with the previous models for comparison. We see that for the all-to-all ansatz, test accuracy is no higher than the classical model beyond the statistical variance of repeating training with different initial parameters, and below the ring ansatz. Indeed, by the end of training, the all-to-all ansatz achieves a final accuracy of $(43.46 \pm 1.68)\%$, which is similar to the classical model. Thus, we show that achieving an advantage using quantum neural networks in contrastive learning is highly dependent on the correct choice of quantum circuit structure.

5.6.1 Real device experiments

In section 5.4.1, we showed that a numerical advantage can be achieved for self-supervised learning with a quantum representation network, even when sampling the quantum circuits with only 100 shots. However, it does not follow that such an improvement can necessarily be realised on current quantum devices. The biggest

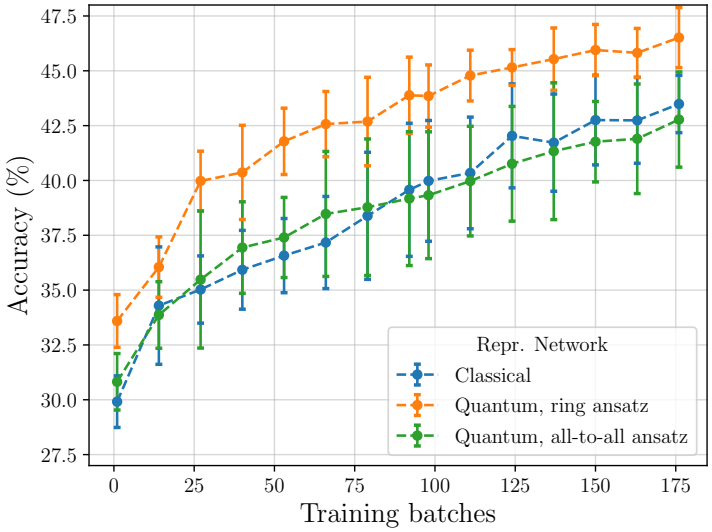


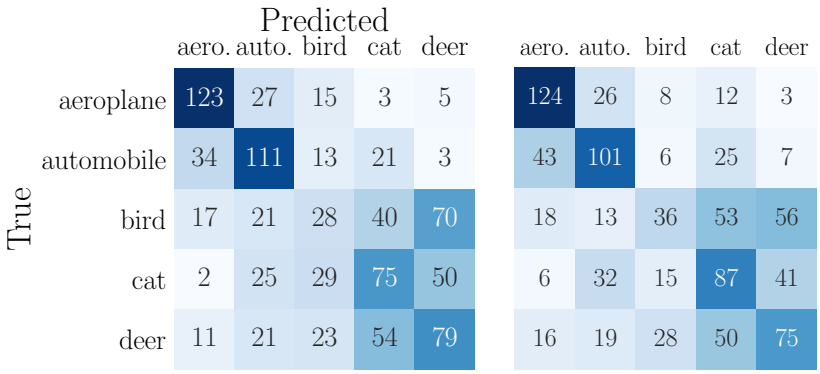
Figure 5.9: Classification accuracy achieved in linear probing experiments using the encoder at checkpoints across self-supervised training. Comparison between models trained with a classical representation network (blue), quantum representation network with the ring ansatz (orange) and quantum representation network with the all-to-all ansatz (green). All quantum circuits were evaluated on a statevector simulator. The markers show the average of six independently trained models, whilst the error bars show one standard deviation.

barrier to this is the complex noise present on quantum hardware, a product of both the finite lifetime that qubits can be held in coherent states for and imperfections in the application of logic gates. To this end, we test the ability of real devices to accurately prepare representations produced by a pretrained quantum model and how this changes downstream accuracy on the test set.

We construct a linear probe experiment with a quantum representation network and load in weights from the best performing pretrained model in which circuits were evaluated with 100 shots. Freezing all of the layers so that the entire network no longer trains, we repeat classification of images from the test set, however this time the circuits are executed on IBM's 27-qubit *ibmq_paris* quantum computer. When executing our QNNs on *ibmq_paris*, translating the ring topology of our

variational ansatz to the honeycomb structure that the qubits are physically connected by requires a significant number of SWAP operations. Quantitatively this increases the number of two-qubit gates in the circuit from 16 to 143, which poses a significant challenge to obtaining a predictive signal beyond random noise since the total circuit error scales exponentially with the number of gates. To mitigate this, the quantum representation network is first recompiled with ISL. Although the quantum ansatz parameters are fixed for this prediction task, the data loading unitary is different for each image. As such, this experiment requires the evaluation of 900 different circuits corresponding to 900 randomly sampled images from the test set, each of which we recompile. Importantly, we first attempt to recompile the circuits with an adapted version of ISL, in which the two-qubit layers added incrementally to the trial solution can only be applied to pairs of qubits that are physically connected on the real device. If successful, this has the benefit of producing approximately equivalent circuits that are both shallower and will not require SWAP operations to run on the real device. We find that in this first pass, for over half of the circuits a device-aware ISL solution is found which produces the same state vector with at least 99% overlap using on average 14 CNOT gates. For the remaining images, we apply ISL once again, but this time without any constraints on the connectivity of the circuit. This produces an approximately equivalent circuit with at least 99% overlap using on average 8 CNOT gates, although after accounting for SWAP gates when transpiling to the real device this grows to an average of 30 CNOT gates. Nevertheless, both cases represent a significant reduction in the depth of the circuit to be evaluated and total error incurred when running this experiment.

Fig. 5.10a shows the result of classifying 900 images randomly sampled from the test set, using the best performing classical model and evaluated on a classical computer. Fig. 5.10b shows the result when classifying the same images using



(a) (b)

Figure 5.10: Confusion matrix from classifying 900 images using the best performing (a) classical model evaluated on a classical computer (b) quantum model evaluated on a real quantum computer with 100 shots per circuit. For a given true label (rows) and predicted label (columns), the number in each box shows the total number of times that prediction was made.

the best performing 100-shot quantum model, evaluated on *ibmq_paris*. Overall, the classical and quantum models achieve an accuracy of 47.27% and 47.00% respectively. Excitingly, this demonstrates that in this experiment, error induced by noise on the quantum computer is able to be offset by the enhanced theoretical performance of quantum neural networks, provided the circuit depth is reduced with recompilation techniques. Furthermore, in both setups the most correctly predicted class was aeroplanes (71.1% and 71.7%) whilst the most incorrectly predicted class was birds (15.9% and 20.5%), both of which the quantum model performed better on. We propose that birds and deer were most likely to be mistaken with one another due to the images sharing a common background of the outdoor natural environment.

5.7 Conclusion

In this chapter, we propose a hybrid quantum-classical architecture for self-supervised learning and demonstrate a numerical advantage in the learning of visual representations using small-scale QNNs. We train quantum and classical neural networks together, such that encodings are learnt that maximise the similarity of augmented views of the same image in the representation space, as well as implicitly in Hilbert space. After training is complete, we determine the quality of the embedding by tasking a linear probe to classify images from different classes. We find that an encoder with a QNN acting in the representation space achieves higher average test set accuracy than one in which the QNN is replaced by a classical neural network with equivalent width and depth, even when evaluating quantum circuits with only 100 shots. We note that although making such a comparison has been established in previous works [207], how to fairly compare quantum and classical neural networks remains a significant open question. In theory, the comparison made in this chapter lends itself to improved quantum network performance, since it trains with access to an exponentially larger feature space than the counterpart classical network. Yet this is not a bygone conclusion, least of which because the quantum network has fewer parameters. Our result provides an important benchmark and the first step towards a useful comparison between the paradigms. Looking forward, a next step could be to compare quantum and classical networks of equivalent run time or energy consumption, although this is not relevant until the entire QNN training can be done on real quantum hardware.

After this first experiment, we then apply our best performing pretrained classical and quantum models to downstream classification, whereby the quantum circuits were evaluated on a real quantum computer. The observation of a quantum predictive signal with equivalent accuracy to that of the classical model, despite the

complex noise present on current quantum devices, is representative of the potential practical benefit of our setup. If recent progress in superconducting qubit hardware continues [11, 239, 9], it is possible that QNNs running on real devices will outperform equally sized classical neural networks in the near future in this experiment.

One advantage of the hybrid approach taken here is the resulting flexibility in how much of the encoder is quantum or classical. In fact, there now exist numerous software solutions for producing and testing such hybrid architectures [79, 240, 241]. As the quality and size of quantum hardware improves, our scheme allows classical capacity to be substituted for quantum, eventually replacing ResNet entirely. By optimising directly for the Hilbert-Schmidt distance, it is also possible with a fully quantum encoder to apply our setup to problems in which the data is itself quantum [242, 243, 244, 245, 246]. Promisingly, in this regime it may prove that the advantage observed in this chapter is further extended, given the ability of a quantum model to inherently exploit the dimensionality of the input [247]. Recently developed data sets consisting of entangled quantum states [248, 249] serve as an obvious target for such work. In this case contrastive augmentations could be quantum operations that change the state but conserve the properties of interest, for example LOCC operations that do not affect the amount of entanglement in the system. With classical contrastive learning having been applied to non-visual problems in biology [215] and chemistry [216], our work provides a strong foundation for applying quantum self-supervised learning to fundamentally quantum problems in the natural sciences [65].

An open question remains as to whether a general quantum advantage for self-supervised learning may prove possible [11, 250], in which no classical computer of any size can produce accuracies equal to that of a quantum model. In [207, 251], the authors define *effective-dimension*, a metric measuring the expressive

power of classical and quantum neural networks. In general, quantum models are able to achieve a higher effective dimension, and therefore capture a larger space of functions, than classical models with comparable width and number of parameters. Although it does not necessarily increase monotonically, the effective dimension of quantum models can remain larger than classical as the model and data set size are increased. Such behaviour indicates that the expressive power available to QNNs may allow for an advantage over classical neural networks, particularly for a problem such as self-supervised learning where highly expressive, large capacity models are believed to be particularly important for achieving highly accurate predictions [191].

Achieving experimental quantum advantage would require, as a minimum, a QNN with width greater than 60 qubits, such that the dimensionality of the accessible feature space becomes classically intractable. Furthermore, the QNNs would need to be trained on real devices, which remains a challenge due short qubit lifetimes and low gate fidelity. Therefore, considerable research still remains into the scalability of our scheme, which was only demonstrated at the small sizes feasible on current quantum hardware. Promisingly however, our method can be adapted to use different QNN structures and training schemes that avoid the scaling issue of barren plateaus [92]. Specifically, our choice of random ansatz could be replaced with a quantum convolutional neural network [252], or initialised to form the identity operation [253], both of which avoid the variance of the gradient of the parameters vanishing exponentially with system size. Alternatively, our sampling of global observables of the quantum circuit could be replaced with local observables [254], which is known to avoid such concentration of measure. Excitingly, scaling our scheme to larger networks could already be attempted with the recent development of more efficient quantum computing simulators [231]. Looking forward, the rate at which quantum hardware continues to progress pro-

vides the possibility of representing intractable distributions using QNNs. In this way, quantum computers may yet push self-supervised learning beyond the performance afforded by classical hardware.

[252, 253, 254]

5.8 Code availability

The code used to train the models described in this chapter can be found at `https://github.com/bjader/QSSL`. The code used to incorporate and train Qiskit quantum neural networks into PyTorch can be found at `https://github.com/bjader/quantum-neural-network` and is required to build quantum representation networks.

6 | Conclusions and Outlook

Across this thesis we have explored the theory of solving optimisation problems using VQAs, a new paradigm in quantum algorithms. By gathering domain expertise in different fields, our research has demonstrated both the development of novel VQAs and their potential impact in areas such as quantum chemistry and machine learning. In this section, we recap the main chapters of this thesis and consider the overarching conclusions that can help advance the field forward. Finally, we suggest significant next steps that can be taken in light of the findings of this thesis, both specific to the algorithms studied and for quantum computing as a whole.

6.1 Summary of Work

A brief recap of the work presented in the main chapters of this thesis is as follows. In chapter 2 we introduce approximate quantum circuit recompilation (AQCR) and its motivation as a form of quantum error mitigation which we use across our research. Our main contributions are:

- We highlight the limitations of existing approximate quantum circuit recompilation algorithms, which use fixed ansatz structures.
- We develop a new recompilation algorithm, ISL, which incrementally builds a candidate solution by connecting the most entangled pairs of qubits together.
- We detail the process of adding and optimising each incremental layer, as well as the termination conditions for the whole algorithm.

In chapter 3, we test whether the dynamics of electronic excitations evolving under the electron-phonon Hamiltonian can be obtained on current digital quantum computers. This is an important first step towards using VQAs to solve dynamics based optimisation problems, which are a promising avenue for near-term quantum advantage. Our main contributions are:

- We develop the first implementation simulating the evolution of the electron-phonon Hamiltonian in the second quantisation formalism on a quantum computer.
- We report the number of Trotter steps and number of qubits necessary for convergence of the dynamics compared to the exact diagonalisation solution.
- We find that the electronic excitation population cannot be accurately obtained on superconducting hardware even for a minimal number of sites.
- To overcome this barrier we test the effectiveness of the previously introduced ISL algorithm, applying it to the Trotterised evolution circuits. We observe that with ISL, a recompiled solution can consistently be found which approximates the system dynamics with less gates than the original circuit. This helps to reduce the noise generated on near-term quantum computers and maximise performance.
- Subsequently, we show that using the ISL recompiled circuits, dynamics can be achieved on current quantum computers comparable to the exact diagonalisation solution.

In chapter 4 we pursue a natural next step to the above research, namely testing whether a time-evolution based VQA can be run accurately on near-term quantum computers. For this we propose a scheme for hybrid quantum-classical DMFT, in

which the impurity Green's function is obtained using a quantum computer. Our main contributions are:

- We find the number of Trotter steps necessary for the self-consistently determined quasiparticle weight to match the analytic solution, including identifying the Mott metal-insulator transition. This is a stricter requirement than just convergence of the dynamics as in chapter 3.
- Through noise modelling, we determine the minimum two-qubit fidelity required to obtain results within one standard deviation of the true value is 99.999%, well beyond the capabilities of near-term quantum computers.
- We show that with the use of ISL, the minimum fidelity can be decreased by a factor of 100 to 99.9%. Whilst this was beyond the ability of quantum hardware at the time of the experiment, near-term devices look soon to fulfil these requirements.

In chapter 5 we investigate the use of VQAs to solve problems in machine learning, often described as quantum neural networks. Whilst the potential impact of QNNs is enormous, evaluation of the cost function (i.e., a forward pass of the network) is not classically hard in general, unlike time evolution based VQAs. We argue that the supposed benefit of QNNs having access to an exponentially large feature space would thus be most likely recognised when applied to deep learning problems for which performance is known to be limited by today's available classical computing power. One of the most significant examples of this is self-supervised learning, which requires orders of magnitudes more training resources than other methods. Our main contributions are:

- We build and open source the first ever library allowing for integration of Qiskit QNNs into PyTorch. This enables the construction of large hybrid neural network architectures where quantum and classical networks can be

trained concurrently under one optimiser.

- We demonstrate constrastive self-supervised learning of image representations and show that a QNN embedded into the encoder contributes to the learning process.
- We find that the inclusion of a QNN leads to a numerical advantage in classification accuracy over equivalently sized classical networks, even when including realistic qubit sampling noise. This provides preliminary evidence that self-supervised learning could benefit from an encoder with exponential dimensionality.
- We perform classification of 900 images using real quantum hardware and achieve tied performance with the fully classical model.

With this work completed, we are well placed to consider the insights and conclusions to the wider community that can be drawn from our results. One of the goals of this thesis was to test the capabilities of current quantum computers, exploring their ability to execute useful algorithms and in the process providing feedback on their limitations. Our findings suggest that firstly, both the size of available hardware and the number of shots that can be evaluated on them are sufficient for near-term algorithms, with neither attribute presenting a problem when building and proving out the different algorithms in this thesis. Secondly, despite significant progress in hardware error rates, the levels of noise on current quantum computers does not allow for the direct implementation of algorithms to solve optimisation problems on them, even at small scales. This is true both for the evaluation of the cost function itself, as in chapter 3, and also when we integrate the quantum circuit based cost function into wider optimisation routines as per chapters 4 and 5. This demonstrates that despite the theoretical noise-tolerant capabilities of VQAs, there is still a large gap to be bridged. As a consequence,

researchers developing VQAs should give significant consideration to strategies to deal with this noise if they are to be run on near-term devices. Our work shows that AQCR is one such strategy and can be integrated into many different types of VQAs. Therefore, we wholly recommend the uptake of AQCR as an error mitigation technique into the wider community, which at the moment is not widely adopted. We hope that this, in tandem with more well-known post-processing schemes, may provide a feasible algorithmic approach for quantum computing in the presence of noise.

Another finding of this thesis, based on our many experiments with real quantum hardware, is that the problem of noise is particularly exacerbated when the desired quantum circuit connectivity does not match the physical qubit connectivity. Mapping between the two subsequently requires SWAP gates that significantly increase the circuit depth and thus noise generated. Despite this, superconducting quantum computers continue to be demonstrated at large scales using only nearest-neighbour connections in a square or heavy-hexagonal lattice. We conclude that working towards superconducting quantum computers with high qubit connectivity, so that many different graph structures can be represented, would significantly improve their ability to accurately execute VQAs in the near term. Alternatively, from a quantum algorithms perspective, we conclude that the required connectivity of an algorithm should be equally prioritised when considering the suitability of different qubit architectures for an algorithm, alongside other properties such as qubit number and error rate.

The other aim of this thesis was to develop and experiment with a range of different algorithms, such that we could explore what a useful quantum advantage on NISQ devices might look like. In this regard, one of the main conclusions from this thesis is that not all VQAs are created equal, and the steps to achieve an advantage will require a specific set of conditions for each one. In chapters 3 and

4, we explore applications which require the computation of the time evolution of a quantum system, which is well understood to be classically intractable. Yet in the short term, it is still vital to establish how errors affect the observables of the cost function and understand the subsequent effect of this on convergence of the optimiser. Despite their classical hardness, if such algorithms are exponentially sensitive to noise then they will not result in a practical advantage in the NISQ era. By contrast, in chapter 5 we explore the use of VQAs for machine learning, whose equivalent in the form of classical learning models are not intractable. Therefore, the steps required to achieve an advantage for QNN based algorithms looks very different to quantum chemistry, and the larger burden of proof may require the development of theory beyond the currently understood heuristics. Conversely, because they encode a fundamentally different cost function to their classical alternative, an advantage for QNNs may be provable not in time or memory but in training iterations. This opens the possibility of discovering a quantum advantage in an entirely numerical setting, mirroring the way in which the progress of classical deep learning has historically led the theoretical understanding.

6.2 Outlook and Open Questions

Alongside the work presented in this thesis, we have also discussed some of the interesting questions and open problems arising from our results. Here we consider them in more detail and discuss future avenues of research, first as specific extensions to our research and subsequently the questions our research poses to the wider field.

In chapter 2 we describe ISL and explain how it can be used to find approximately equivalent shallow circuits that produce less noise when run on near-term quantum computers. Applying this method in our research significantly aids the

achievement of a number of notable goals in this thesis including: demonstrating the first electron-phonon Hamiltonian simulation in chapter 3, identifying hybrid quantum-classical DMFT as an imminent application in chapter 4 and using a quantum self-supervised learning model to classify images on real hardware in chapter 5. These examples demonstrate the great potential that ISL holds despite being developed here as a practical and heuristic solution. Therefore, one clear direction of research would be the further development, analysis and benchmarking of the ISL algorithm as its own publication. Importantly, where ISL was used in this thesis, the recompilation was done using classically simulated quantum circuits, the computational complexity of which scales exponentially. Thus, a significant next step would be to establish whether ISL evaluated on a quantum computer would scale efficiently. This is immediately important if we are to reproduce and extend the results of this thesis at scales of 30+ qubits, where the memory required to simulate the circuit quickly exceeds that even of computing clusters.

For many of the subroutines in ISL, it is clear that they could be efficiently run on real quantum hardware. For example, obtaining the cost function scales only linearly with the number of qubits of the target circuit, since the overlap in Eq. (2.1) can be obtained via the so-called swap test [255, 256]. The core of this scheme is evaluating a controlled SWAP operation, where the control is on an ancilla qubit and the SWAP occurs over two registers of qubits acted on by the original circuit and recompiled circuit respectively. Thus, recompiling a N qubit circuit requires only $2N+1$ qubits. Evaluation of the cost function is also efficient in time, since by definition the recompiled solution is shallower than the target circuit and if not, ISL aborts. This means that at worst, recompiling a target of depth D requires the evaluating circuits no deeper than $2D$. Furthermore, the number of shots required for the error in the measured overlap to be smaller than

the ISL threshold cost C_t is $O(\frac{1}{C_t^2})$ [257].

At a higher level, the process of adding and optimising each layer is also scalable, since there are only quadratically many pairwise entanglement measure to obtain and the Rotoselect algorithm is efficient [91]. Thus, the largest question for whether ISL is scalable or not centres on how the number of layers to reach convergence scales with the size of the target. This is a difficult question to analytically reason about since on a real quantum computer the cost is randomly sampled, and as such ISL is a stochastic optimisation problem whereby the scaling of convergence with respect to system size is unknown, like all VQAs. Numerically, one study has shown that the number of circuit evaluations required to reach convergence for ISL n_e is an order of magnitude larger than non-incremental recompilers for certain problems [97], a trade-off for the improved flexibility and trainability. Therefore, an important next step would be to determine numerically how n_e scales with the size of the target circuit for a wider range of systems. Importantly, if this future work could establish ISL as a mechanism by which deep circuits can be substituted for a non exponential number of shallow circuits, this would represent a significant step towards a practical advantage for near-term quantum computers for any application.

In chapter 4 we determine the minimum hardware requirements for hybrid quantum-classical DMFT. The conclusion of this was that even with error mitigation methods, accurately running two-site DMFT was out of reach for cloud accessible superconducting quantum computers. However, an important caveat is that there exists many smaller purpose built quantum computers with higher gate fidelity than those available over the cloud. Working with such devices, it may be possible to experimentally demonstrate the first quantum DMFT procedure, a significant step in VQAs for quantum chemistry. This future research avenue is the subject of ongoing collaboration with experimental research groups.

In chapter 5 we report on a proof-of-principle experimental advantage for quantum self-supervised learning over equivalently sized classical models. Following on from this, there are two significant avenues of research which can build off the foundations of this result, improving on some of its limitations in the process. Firstly, due to the inefficiency of Qiskit's simulator, and the almost one million circuits required to be evaluated for every training epoch, we were only able to present results for an eight qubit network trained for under two epochs. Despite this requiring significant computational resources, a large disparity still remains to the hundreds of epochs that self-supervised learning models are normally trained for to reach convergence. This is important, since until the model stops improving with additional training time, it is impossible to conclude that a tangible advantage exists. Instead, it may be possible that the inclusion of a quantum network leads to some immediate improvement, but ultimately provides worse performance relative to classical alternatives as training continues. To test this theory, it is vital that an extended study is done with significantly longer training times. Notably, this may be possible by leveraging the recent landmark release of cuQuantum [232, 258], a GPU enhanced quantum simulator built by NVIDIA which can achieve an order of magnitude speedup over CPU simulators. This will allow not only training to convergence, but also a more thorough analysis of whether the advantage scales to larger numbers of qubits.

Another worthy future research direction from our work in chapter 5 would be the study of quantum self-supervised learning with fully quantum networks, by which we mean without the classical components that make our architecture hybrid. These classical networks were included originally to play the important role of reducing the dimensionality of the input data, making it amenable to encoding into intermediate sized quantum computers. However, a hybrid architecture has the disadvantage of removing the clarity by which we can correlate changes

in the QNN to changes in performance, due to a significant proportion of the learning occurring classically. This is especially restrictive for producing analytic arguments to support a potential quantum advantage for self-supervised learning. Therefore, further experiments are warranted into whether an advantage can be more fundamentally proved using only quantum networks. To avoid the expensive schemes required to encode images into quantum circuits [224, 223], this future work could evaluate performance on one of the quantum benchmark datasets developed recently [248], such as classifying the entanglement of quantum states [249].

Despite the explicit goal of this thesis to develop and test different types of VQAs, we reflect here on whether such a strategy should be pursued in the future. Over the past few years of research, the landscape of near-term quantum algorithms has changed dramatically, shaped by Google's demonstration of quantum advantage [11]. Whilst this experiment was itself not a VQA, it has renormalised expectations within the community of what might be possible in the near future and if such an advantage could be demonstrated with VQAs. Our research demonstrates one approach to tackling this problem, widely pursued by the field at large, in which we try different applications in order to learn heuristics and probe the limit of current performance. However, an alternative strategy is to start from the ground up, considering analytically whether for a certain optimisation problem there exists a low-depth quantum circuit which is both useful to solve the problem and classically hard. For example, recent efforts have shown that for one of the most popular structures of QNNs, the quantum model is effectively represented as a Fourier series in the input data [259, 260] and can only become a universal function approximate in the limit of exponential depth. This raises an important consideration: even VQAs with intractable ansatzes may not be generally useful if they can only represent a limited class of functions. We conclude that as the

field of near-term quantum algorithms matures, such analyses for other models of VQAs will allow a constructive and informed approach as to which applications should then be developed further and trialled on the real devices.

In this thesis, we have shown that despite the noise tolerance of VQAs, there exists many optimisation problems for which current superconducting quantum computers are not able to solve accurately. As variational algorithms gain ever increasing popularity, we suggest that looking forward, there is a growing importance for the wider field to start establishing bounds on what sort of noise tolerance can be expected. Alongside analytic research into these bounds, our results could provide a strong foundation for future numerical comparisons between errors accumulated in the execution of quantum circuits versus errors in the solution of variational algorithms using that circuit as a cost function. This will help the near-term quantum algorithm community understand if the current focus on noise tolerance using VQAs is justified, and how specific or general this benefit is. Furthermore, it is equally important that these future works compare between different types of qubit architectures, furthering our understanding of whether certain cost functions and applications are better fitted for specific architecture due to differences in connectivities, error rate distribution or natural qubit interactions. Such comparisons are becoming more and more accessible even for researchers without access to in-house hardware, with recent expansions of commercial cloud offerings already including ion-trap architectures since our research was conducted [261]. Hopefully in the future, this will also include the next generation of quantum computers, such as neutral-atom [262] and photonic architectures [263].

Bibliography

[1] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982.

[2] IU I Manin. *Vychislimoe i nevychislimoe*. Sov. radio,, 1980.

[3] Hartmut Häffner, Christian F Roos, and Rainer Blatt. Quantum computing with trapped ions. *Physics reports*, 469(4):155–203, 2008.

[4] Caroline Figgatt, Aaron Ostrander, Norbert M Linke, Kevin A Landsman, Daiwei Zhu, Dmitri Maslov, and Christopher Monroe. Parallel entangling operations on a universal ion-trap quantum computer. *Nature*, 572(7769):368–372, 2019.

[5] Chris Monroe, David M Meekhof, Barry E King, Wayne M Itano, and David J Wineland. Demonstration of a fundamental quantum logic gate. *Physical review letters*, 75(25):4714, 1995.

[6] Elizabeth Gibney. Quantum computer race intensifies as alternative technology gains steam. *Nature*, 587(7834):342–344, 2020.

[7] Jiehang Zhang, Guido Pagano, Paul W Hess, Antonis Kyprianidis, Patrick Becker, Harvey Kaplan, Alexey V Gorshkov, Z-X Gong, and Christopher Monroe. Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator. *Nature*, 551(7682):601–604, 2017.

[8] Göran Wendin. Quantum information processing with superconducting circuits: a review. *Reports on Progress in Physics*, 80(10):106001, 2017.

[9] Petar Jurcevic, Ali Javadi-Abhari, Lev S Bishop, Isaac Lauer, Daniela Borgorin, Markus Brink, Lauren Capelluto, Oktay Gunluk, Toshinari Itoko,

Naoki Kanazawa, et al. Demonstration of quantum volume 64 on a superconducting quantum computing system. *Quantum Science and Technology*, 2021.

[10] Peter JJ O'Malley, Ryan Babbush, Ian D Kivlichan, Jonathan Romero, Jarrod R McClean, Rami Barends, Julian Kelly, Pedram Roushan, Andrew Tranter, Nan Ding, et al. Scalable quantum simulation of molecular energies. *Physical Review X*, 6(3):031007, 2016.

[11] 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.

[12] Jerry Chow, Oliver Dial, and Jay Gambetta. Ibm quantum breaks the 100-qubit processor barrier, 2021.

[13] Cupjin Huang, Fang Zhang, Michael Newman, Junjie Cai, Xun Gao, Zhengxiong Tian, Junyin Wu, Haihong Xu, Huanjun Yu, Bo Yuan, et al. Classical simulation of quantum supremacy circuits. *arXiv preprint arXiv:2005.06787*, 2020.

[14] Feng Pan and Pan Zhang. Simulating the sycamore quantum supremacy circuits. *arXiv preprint arXiv:2103.03074*, 2021.

[15] Yulin Wu, Wan-Su Bao, Sirui Cao, Fusheng Chen, Ming-Cheng Chen, Xiawei Chen, Tung-Hsun Chung, Hui Deng, Yajie Du, Daojin Fan, et al. Strong quantum computational advantage using a superconducting quantum processor. *Physical review letters*, 127(18):180501, 2021.

[16] Qingling Zhu, Sirui Cao, Fusheng Chen, Ming-Cheng Chen, Xiawei Chen, Tung-Hsun Chung, Hui Deng, Yajie Du, Daojin Fan, Ming Gong, et al.

Quantum computational advantage via 60-qubit 24-cycle random circuit sampling. *Science Bulletin*, 2021.

[17] Craig R Clark, Holly N Tinkey, Brian C Sawyer, Adam M Meier, Karl A Burkhardt, Christopher M Seck, Christopher M Shappert, Nicholas D Guise, Curtis E Volin, Spencer D Fallek, et al. High-fidelity bell-state preparation with ca+ 40 optical qubits. *Physical Review Letters*, 127(13):130505, 2021.

[18] Eli Chertkov, Zihan Cheng, Andrew C Potter, Sarang Gopalakrishnan, Thomas M Gatterman, Justin A Gerber, Kevin Gilmore, Dan Gresh, Alex Hall, Aaron Hankin, et al. Characterizing a non-equilibrium phase transition on a quantum computer. *arXiv preprint arXiv:2209.12889*, 2022.

[19] Demonstrating benefits of quantum upgradable design strategy: System model h1-2 first to prove 2,048 quantum volume. Dec 2021.

[20] IBM Quantum. https://quantum-computing.ibm.com/, 2023.

[21] Pranav Mundada, Gengyan Zhang, Thomas Hazard, and Andrew Houck. Suppression of qubit crosstalk in a tunable coupling superconducting circuit. *Physical Review Applied*, 12(5):054023, 2019.

[22] Nikolaj Moll, Panagiotis Barkoutsos, Lev S Bishop, Jerry M Chow, Andrew Cross, Daniel J Egger, Stefan Filipp, Andreas Fuhrer, Jay M Gambetta, Marc Ganzhorn, et al. Quantum optimization using variational algorithms on near-term quantum devices. *Quantum Science and Technology*, 3(3):030503, 2018.

[23] Andrew W Cross, Lev S Bishop, Sarah Sheldon, Paul D Nation, and Jay M Gambetta. Validating quantum computers using randomized model circuits. *Physical Review A*, 100(3):032328, 2019.

[24] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.

[25] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.

[26] Lov K Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 212–219, 1996.

[27] Peter W Shor. Fault-tolerant quantum computation. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 56–65. IEEE, 1996.

[28] Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error rate. *SIAM Journal on Computing*, 2008.

[29] Earl T Campbell, Barbara M Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, 2017.

[30] Austin G Fowler, Matteo Mariantoni, John M Martinis, and Andrew N Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86(3):032324, 2012.

[31] Craig Gidney and Martin Ekerå. How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits. *Quantum*, 5:433, 2021.

[32] Ben W Reichardt. Fault-tolerant quantum error correction for steane's seven-qubit color code with few or no extra qubits. *Quantum Science and Technology*, 6(1):015007, 2020.

[33] Rui Chao and Ben W Reichardt. Flag fault-tolerant error correction for any stabilizer code. *PRX Quantum*, 1(1):010302, 2020.

[34] Laird Egan, Dripto M Debroy, Crystal Noel, Andrew Risinger, Daiwei Zhu, Debopriyo Biswas, Michael Newman, Muyuan Li, Kenneth R Brown, Marko Cetina, et al. Fault-tolerant control of an error-corrected qubit. *Nature*, 598(7880):281–286, 2021.

[35] Zijun Chen, Kevin J Satzinger, Juan Atalaya, Alexander N Korotkov, Andrew Dunsworth, Daniel Sank, Chris Quintana, Matt McEwen, Rami Barends, Paul V Klimov, et al. Exponential suppression of bit or phase errors with cyclic error correction. *Nature*, 595(7867):383–387, 2021.

[36] KJ Satzinger, Y-J Liu, A Smith, C Knapp, M Newman, C Jones, Z Chen, C Quintana, X Mi, A Dunsworth, et al. Realizing topologically ordered states on a quantum processor. *Science*, 374(6572):1237–1241, 2021.

[37] Suguru Endo, Zhenyu Cai, Simon C Benjamin, and Xiao Yuan. Hybrid quantum-classical algorithms and quantum error mitigation. *Journal of the Physical Society of Japan*, 90(3):032001, 2021.

[38] Jerry M Chow, Jay M Gambetta, Easwar Magesan, David W Abraham, Andrew W Cross, Blake R Johnson, Nicholas A Masluk, Colm A Ryan, John A Smolin, Srikanth J Srinivasan, et al. Implementing a strand of a scalable fault-tolerant quantum computing fabric. *Nature communications*, 5(1):1–9, 2014.

[39] Kristan Temme, Sergey Bravyi, and Jay M Gambetta. Error mitigation for short-depth quantum circuits. *Physical review letters*, 119(18):180509, 2017.

[40] Andre He, Benjamin Nachman, Wibe A de Jong, and Christian W Bauer. Resource efficient zero noise extrapolation with identity insertions. *arXiv preprint arXiv:2003.04941*, 2020.

[41] Tudor Giurgica-Tiron, Yousef Hindy, Ryan LaRose, Andrea Mari, and William J Zeng. Digital zero noise extrapolation for quantum error mitigation. In *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 306–316. IEEE, 2020.

[42] Swamit S Tannu and Moinuddin K Qureshi. Not all qubits are created equal: a case for variability-aware policies for nisq-era quantum computers. In *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, pages 987–999, 2019.

[43] Armands Strikis, Dayue Qin, Yanzhu Chen, Simon C Benjamin, and Ying Li. Learning-based quantum error mitigation. *PRX Quantum*, 2(4):040330, 2021.

[44] Changjun Kim, Kyungdeock Daniel Park, and June-Koo Rhee. Quantum error mitigation with artificial neural network. *IEEE Access*, 8:188853–188860, 2020.

[45] Ryuji Takagi, Suguru Endo, Shintaro Minagawa, and Mile Gu. Fundamental limits of quantum error mitigation. *arXiv preprint arXiv:2109.04457*, 2021.

[46] Ryuji Takagi, Hiroyasu Tajima, and Mile Gu. Universal sample lower bounds for quantum error mitigation. *arXiv preprint arXiv:2208.09178*, 2022.

[47] Yihui Quek, Daniel Stilck França, Sumeet Khatri, Johannes Jakob Meyer, and Jens Eisert. Exponentially tighter bounds on limitations of quantum error mitigation. *arXiv preprint arXiv:2210.11505*, 2022.

[48] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.

[49] 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*, 5:4213, 2014.

[50] Oscar Higgott, Daochen Wang, and Stephen Brierley. Variational quantum computation of excited states. *Quantum*, 3:156, 2019.

[51] Ying Li and Simon C Benjamin. Efficient variational quantum simulator incorporating active error minimization. *Physical Review X*, 7(2):021050, 2017.

[52] Maria Schuld and Nathan Killoran. Quantum machine learning in feature hilbert spaces. *Physical review letters*, 122(4):040504, 2019.

[53] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. The quest for a quantum neural network. *Quantum Information Processing*, 13(11):2567–2586, 2014.

[54] 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, 2018.

[55] JS Otterbach, R Manenti, N Alidoust, A Bestwick, M Block, B Bloom, S Caldwell, N Didier, E Schuyler Fried, S Hong, et al. Unsupervised machine learning on a hybrid quantum computer. *arXiv preprint arXiv:1712.05771*, 2017.

[56] Sumeet Khatri, Ryan LaRose, Alexander Poremba, Lukasz Cincio, Andrew T Sornborger, and Patrick J Coles. Quantum-assisted quantum compiling. *Quantum*, 3:140, 2019.

[57] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.

[58] Carlos Bravo-Prieto, Ryan LaRose, Marco Cerezo, Yigit Subasi, Lukasz Cincio, and Patrick Coles. Variational quantum linear solver: A hybrid algorithm for linear systems. *Bulletin of the American Physical Society*, 65, 2020.

[59] Michael Lubasch, Jaewoo Joo, Pierre Moinier, Martin Kiffner, and Dieter Jaksch. Variational quantum algorithms for nonlinear problems. *Physical Review A*, 101(1):010301, 2020.

[60] Nikita Gourianov, Michael Lubasch, Sergey Dolgov, Quincy Y van den Berg, Hessam Babaee, Peyman Givi, Martin Kiffner, and Dieter Jaksch. A quantum-inspired approach to exploit turbulence structures. *Nature Computational Science*, pages 1–8, 2022.

[61] Anton Robert, Panagiotis Kl Barkoutsos, Stefan Woerner, and Ivano Tavernelli. Resource-efficient quantum algorithm for protein folding. *npj Quantum Information*, 7(1):1–5, 2021.

[62] Yong-Hyuk Kim, Yourim Yoon, and Zong Woo Geem. A comparison study of harmony search and genetic algorithm for the max-cut problem. *Swarm*

and evolutionary computation, 44:130–135, 2019.

[63] Gian Giacomo Guerreschi and Mikhail Smelyanskiy. Practical optimization for hybrid quantum-classical algorithms, 2017.

[64] Harper R. Grimsley, Sophia E. Economou, Edwin Barnes, and Nicholas J. Mayhall. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature Communications*, 10(1):3007, 2019.

[65] Iris Cong, Soonwon Choi, and Mikhail D Lukin. Quantum convolutional neural networks. *Nature Physics*, 15(12):1273–1278, 2019.

[66] M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J. Coles. Cost-function-dependent barren plateaus in shallow quantum neural networks, 2020.

[67] J. Kiefer and J. Wolfowitz. Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, 23(3):462–466, 1952.

[68] J. A. Nelder and R. Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, 1965.

[69] Pedro J Salas. Noise effect on grover algorithm. *The European Physical Journal D*, 46(2):365–373, 2008.

[70] Mauro ES Morales, Timur Tlyachev, and Jacob Biamonte. Variational learning of grover's quantum search algorithm. *Physical Review A*, 98(6):062333, 2018.

[71] Eric Anschuetz, Jonathan Olson, Alán Aspuru-Guzik, and Yudong Cao. Variational quantum factoring. In *International Workshop on Quantum Technology and Optimization Problems*, pages 74–85. Springer, 2019.

[72] Enrico Fontana, Nathan Fitzpatrick, David Muñoz Ramo, Ross Duncan, and Ivan Rungger. Evaluating the noise resilience of variational quantum algorithms. *Physical Review A*, 104(2):022403, 2021.

[73] James I Colless, Vinay V Ramasesh, Dar Dahlen, Machiel S Blok, Mollie E Kimchi-Schwartz, Jarrod R McClean, Jonathan Carter, Wibe A de Jong, and Irfan Siddiqi. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Physical Review X*, 8(1):011021, 2018.

[74] 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.

[75] Kunal Sharma, Sumeet Khatri, Marco Cerezo, and Patrick J Coles. Noise resilience of variational quantum compiling. *New Journal of Physics*, 22(4):043006, 2020.

[76] Shouvanik Chakrabarti, Rajiv Krishnakumar, Guglielmo Mazzola, Nikitas Stamatopoulos, Stefan Woerner, and William J Zeng. A threshold for quantum advantage in derivative pricing. *Quantum*, 5:463, 2021.

[77] John Preskill. Quantum computing 40 years later. *arXiv preprint arXiv:2106.10522*, 2021.

[78] Simon J Devitt. Performing quantum computing experiments in the cloud. *Physical Review A*, 94(3):032329, 2016.

[79] Gadi Aleksandrowicz, Thomas Alexander, Panagiotis Barkoutsos, Luciano Bello, Yael Ben-Haim, David Bucher, Francisco Jose Cabrera-Hernández, Jorge Carballo-Franquis, Adrian Chen, Chun-Fu Chen, Jerry M. Chow, An-

tonio D. Córcoles-Gonzales, Abigail J. Cross, Andrew Cross, Juan Cruz-Benito, Chris Culver, Salvador De La Puente González, Enrique De La Torre, Delton Ding, Eugene Dumitrescu, Ivan Duran, Pieter Eendebak, Mark Everitt, Ismael Faro Sertage, Albert Frisch, Andreas Fuhrer, Jay Gambetta, Borja Godoy Gago, Juan Gomez-Mosquera, Donny Greenberg, Ikko Hamamura, Vojtech Havlicek, Joe Hellmers, Łukasz Herok, Hiroshi Horii, Shaohan Hu, Takashi Imamichi, Toshinari Itoko, Ali Javadi-Abhari, Naoki Kanazawa, Anton Karazeev, Kevin Krsulich, Peng Liu, Yang Luh, Yunho Maeng, Manoel Marques, Francisco Jose Martín-Fernández, Douglas T. McClure, David McKay, Srujan Meesala, Antonio Mezzacapo, Nikolaj Moll, Diego Moreda Rodríguez, Giacomo Nannicini, Paul Nation, Pauline Ollitrault, Lee James O'Riordan, Hanhee Paik, Jesús Pérez, Anna Phan, Marco Pistoia, Viktor Prutyanov, Max Reuter, Julia Rice, Abdón Rodríguez Davila, Raymond Harry Putra Rudy, Mingi Ryu, Ninad Sathaye, Chris Schnabel, Eddie Schoute, Kanav Setia, Yunong Shi, Adenilton Silva, Yukio Siraichi, Seyon Sivarajah, John A. Smolin, Mathias Soeken, Hitomi Takahashi, Ivano Tavernelli, Charles Taylor, Pete Taylour, Kenso Trabing, Matthew Treinish, Wes Turner, Desiree Vogt-Lee, Christophe Vuillot, Jonathan A. Wildstrom, Jessica Wilson, Erick Winston, Christopher Wood, Stephen Wood, Stefan Wörner, Ismail Yunus Akhalwaya, and Christa Zoufal. Qiskit: An Open-source Framework for Quantum Computing, January 2019.

[80] Andrew Richards. *University of Oxford Advanced Research Computing*, August 2015.

[81] Antonio D Córcoles, Jerry M Chow, Jay M Gambetta, Chad Rigetti, James R Rozen, George A Keefe, Mary Beth Rothwell, Mark B Ketchen, and Matthias Steffen. Protecting superconducting qubits from radiation.

Applied Physics Letters, 99(18):181906, 2011.

[82] LB Ioffe, VB Geshkenbein, Ch Helm, and G Blatter. Decoherence in superconducting quantum bits by phonon radiation. *Physical review letters*, 93(5):057001, 2004.

[83] John M Martinis. Saving superconducting quantum processors from decay and correlated errors generated by gamma and cosmic rays. *npj Quantum Information*, 7(1):1–9, 2021.

[84] Philip Krantz, Morten Kjaergaard, Fei Yan, Terry P Orlando, Simon Gustavsson, and William D Oliver. A quantum engineer's guide to superconducting qubits. *Applied Physics Reviews*, 6(2):021318, 2019.

[85] CJ Ballance, TP Harty, NM Linke, MA Sepiol, and DM Lucas. High-fidelity quantum logic gates using trapped-ion hyperfine qubits. *Physical review letters*, 117(6):060504, 2016.

[86] Yvonne Y Gao, M Adriaan Rol, Steven Touzard, and Chen Wang. Practical guide for building superconducting quantum devices. *PRX Quantum*, 2(4):040202, 2021.

[87] Ron Schutjens, F Abu Dagga, Daniel J Egger, and Frank K Wilhelm. Single-qubit gates in frequency-crowded transmon systems. *Physical Review A*, 88(5):052330, 2013.

[88] D Michael Miller and Zahra Sasanian. Lowering the quantum gate cost of reversible circuits. In *2010 53rd IEEE International Midwest Symposium on Circuits and Systems*, pages 260–263. IEEE, 2010.

[89] Stephen S Bullock and Igor L Markov. Arbitrary two-qubit computation in 23 elementary gates. *Physical Review A*, 68(1):012318, 2003.

[90] Tyson Jones and Simon C Benjamin. Quantum compilation and circuit optimisation via energy dissipation. *arXiv preprint arXiv:1811.03147*, 2018.

[91] Mateusz Ostaszewski, Edward Grant, and Marcello Benedetti. Structure optimization for parameterized quantum circuits. *Quantum*, 5:391, 2021.

[92] 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, 2018.

[93] Todd Tilma, Mark Byrd, and ECG Sudarshan. A parametrization of bipartite systems based on su (4) euler angles. *Journal of Physics A: Mathematical and General*, 35(48):10445, 2002.

[94] William K Wootters. Entanglement of formation of an arbitrary state of two qubits. *Physical Review Letters*, 80(10):2245, 1998.

[95] Scott Hill and William K Wootters. Entanglement of a pair of quantum bits. *Physical review letters*, 78(26):5022, 1997.

[96] Guifré Vidal and Reinhard F Werner. Computable measure of entanglement. *Physical Review A*, 65(3):032314, 2002.

[97] Nathan Fitzpatrick, Harriet Apel, and David Muñoz Ramo. Evaluating low-depth quantum algorithms for time evolution on fermion-boson systems. *arXiv preprint arXiv:2106.03985*, 2021.

[98] Tyson Jones and Simon C Benjamin. Robust quantum compilation and circuit optimisation via energy minimisation. *Quantum*, 6:628, 2022.

[99] 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*, 549:242–246, 2017.

[100] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. Quantum chemistry in the age of quantum computing. *Chemical Reviews*, 119(19):10856–10915, 2019.

[101] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. Quantum computational chemistry. *Rev. Mod. Phys.*, 92:015003, Mar 2020.

[102] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[103] 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, 2019.

[104] David Deutsch and Richard Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 439(1907):553–558, 1992.

[105] Seth Lloyd. Universal quantum simulators. *Science*, pages 1073–1078, 1996.

[106] F Marsiglio and JP Carbotte. Electron-phonon superconductivity. In *Superconductivity*, pages 73–162. Springer, 2008.

[107] Sarah Mostame, Joonsuk Huh, Christoph Kreisbeck, Andrew J Kerman, Takatoshi Fujita, Alexander Eisfeld, and Alán Aspuru-Guzik. Emulation of complex open quantum systems using superconducting qubits. *Quantum Information Processing*, 16(2):1–16, 2017.

[108] D. W. Schönleber, A. Eisfeld, M. Genkin, S. Whitlock, and S. Wüster. Quantum simulation of energy transport with embedded rydberg aggregates. *Phys. Rev. Lett.*, 114:123005, Mar 2015.

[109] Feng Mei, Vladimir M. Stojanović, Irfan Siddiqi, and Lin Tian. Analog superconducting quantum simulator for holstein polarons. *Phys. Rev. B*, 88:224502, Dec 2013.

[110] Vladimir M. Stojanović, Tao Shi, C. Bruder, and J. Ignacio Cirac. Quantum simulation of small-polaron formation with trapped ions. *Phys. Rev. Lett.*, 109:250501, Dec 2012.

[111] Juha Leppäkangas, Jochen Braumüller, Melanie Hauck, Jan-Michael Reiner, Iris Schwenk, Sebastian Zanker, Lukas Fritz, Alexey V. Ustinov, Martin Weides, and Michael Marthaler. Quantum simulation of the spin-boson model with a microwave circuit. *Phys. Rev. A*, 97:052321, May 2018.

[112] Sabrina Patsch, Sabrina Maniscalco, and Christiane P. Koch. Simulation of open-quantum-system dynamics using the quantum zeno effect. *Phys. Rev. Research*, 2:023133, May 2020.

[113] Anton Potočnik, Arno Bargerbos, Florian A. Y. N. Schröder, Saeed A. Khan, Michele C. Collodo, Simone Gasparinetti, Yves Salathé, Celestino Creatore, Christopher Eichler, Hakan E. Türeci, Alex W. Chin, and Andreas Wallraff. Studying light-harvesting models with superconducting circuits. *Nature Communications*, 9:904, 2018.

[114] TH Johnson, Stephen R Clark, Martin Bruderer, and Dieter Jaksch. Impurity transport through a strongly interacting bosonic quantum gas. *Physical Review A*, 84(2):023617, 2011.

[115] Martin Bruderer, Alexander Klein, Stephen R Clark, and Dieter Jaksch. Polaron physics in optical lattices. *Physical Review A*, 76(1):011605, 2007.

[116] M Bruderer, Alexander Klein, Stephen R Clark, and Dieter Jaksch. Transport of strong-coupling polarons in optical lattices. *New Journal of Physics*, 10(3):033015, 2008.

[117] Sarah Mostame, Patrick Rebentrost, Alexander Eisfeld, Andrew J Kerman, Dimitris I Tsomokos, and Alán Aspuru-Guzik. Quantum simulator of an open quantum system using superconducting qubits: exciton transport in photosynthetic complexes. *New Journal of Physics*, 14(10):105013, 2012.

[118] Vladimir M Stojanović, Tao Shi, C Bruder, and J Ignacio Cirac. Quantum simulation of small-polaron formation with trapped ions. *Physical review letters*, 109(25):250501, 2012.

[119] JP Hague, S Downes, C MacCormick, and PE Kornilovitch. Cold rydberg atoms for quantum simulation of exotic condensed matter interactions. *Journal of Superconductivity and Novel Magnetism*, 27(4):937–940, 2014.

[120] Alicia B Magann, Matthew D Grace, Herschel A Rabitz, and Mohan Sarovar. Digital quantum simulation of molecular dynamics and control. *Physical Review Research*, 3(2):023165, 2021.

[121] Pauline J. Ollitrault, Guglielmo Mazzola, and Ivano Tavernelli. Nonadiabatic molecular quantum dynamics with quantum computers. *Phys. Rev. Lett.*, 125:260511, Dec 2020.

[122] Alexandru Macridin, Panagiotis Spentzouris, James Amundson, and Roni Harnik. Digital quantum computation of fermion-boson interacting systems. *Physical Review A*, 98(4):042312, 2018.

[123] Alexandru Macridin, Panagiotis Spentzouris, James Amundson, and Roni Harnik. Electron-phonon systems on a universal quantum computer. *Phys. Rev. Lett.*, 121:110504, Sep 2018.

[124] Rolando Somma, Gerardo Ortiz, Emanuel Knill, and James Gubernatis. Quantum simulations of physics problems. *International Journal of Quantum Information*, 1(02):189–206, 2003.

[125] Nicolas PD Sawaya, Tim Menke, Thi Ha Kyaw, Sonika Johri, Alán Aspuru-Guzik, and Gian Giacomo Guerreschi. Resource-efficient digital quantum simulation of d-level systems for photonic, vibrational, and spin-s hamiltonians. *npj Quantum Information*, 6(1):1–13, 2020.

[126] Nicholas J Hestand and Frank C Spano. Expanded theory of h-and j-molecular aggregates: the effects of vibronic coupling and intermolecular charge transfer. *Chemical reviews*, 118(15):7069–7163, 2018.

[127] Tobias Brixner, Richard Hildner, Jürgen Köhler, Christoph Lambert, and Frank Würthner. Exciton transport in molecular aggregates–from natural antennas to synthetic chromophore systems. *Advanced Energy Materials*, 7(16):1700236, 2017.

[128] Sam McArdle, Alexander Mayorov, Xiao Shan, Simon Benjamin, and Xiao Yuan. Digital quantum simulation of molecular vibrations. *Chemical science*, 10(22):5725–5735, 2019.

[129] Masuo Suzuki. Generalized trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-

body problems. *Communications in Mathematical Physics*, 51(2):183–190, 1976.

[130] Juha M Kreula, Laura García-Álvarez, Lucas Lamata, Stephen R Clark, Enrique Solano, and Dieter Jaksch. Few-qubit quantum-classical simulation of strongly correlated lattice fermions. *EPJ Quantum Technology*, 3(1):1–19, 2016.

[131] Libor Veis, Jakub Višňák, Hiroaki Nishizawa, Hiromi Nakai, and Jiří Pittner. Quantum chemistry beyond born–oppenheimer approximation on a quantum computer: A simulated phase estimation study. *International Journal of Quantum Chemistry*, 116(18):1328–1336, 2016.

[132] Nicolas PD Sawaya and Joonsuk Huh. Quantum algorithm for calculating molecular vibronic spectra. *The journal of physical chemistry letters*, 10(13):3586–3591, 2019.

[133] Nicolas PD Sawaya, Francesco Paesani, and Daniel P Tabor. Near-and long-term quantum algorithmic approaches for vibrational spectroscopy. *arXiv preprint arXiv:2009.05066*, 2020.

[134] James D Whitfield, Jacob Biamonte, and Alán Aspuru-Guzik. Simulation of electronic structure hamiltonians using quantum computers. *Molecular Physics*, 109(5):735–750, 2011.

[135] Yordan S Yordanov, David RM Arvidsson-Shukur, and Crispin HW Barnes. Efficient quantum circuits for quantum computational chemistry. *Physical Review A*, 102(6):062612, 2020.

[136] Jan Roden, Georg Schulz, Alexander Eisfeld, and John Briggs. Electronic energy transfer on a vibronically coupled quantum aggregate. *The Journal of chemical physics*, 131(4):044909, 2009.

[137] Ulrich Schollwöck. The density-matrix renormalization group. *Reviews of modern physics*, 77(1):259, 2005.

[138] David Jansen, Janez Bonča, and Fabian Heidrich-Meisner. Finite-temperature density-matrix renormalization group method for electron-phonon systems: Thermodynamics and holstein-polaron spectral functions. *Physical Review B*, 102(16):165155, 2020.

[139] Patrick Rebentrost, Masoud Mohseni, Ivan Kassal, Seth Lloyd, and Alán Aspuru-Guzik. Environment-assisted quantum transport. *New Journal of Physics*, 11(3):033003, 2009.

[140] Jasper Knoester. Nonlinear optical line shapes of disordered molecular aggregates: motional narrowing and the effect of intersite correlations. *The Journal of chemical physics*, 99(11):8466–8479, 1993.

[141] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.

[142] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos. Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator. *Phys. Rev. X*, 8:031022, 2018.

[143] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski. Cloud Quantum Computing of an Atomic Nucleus. *Physical Review Letters*, 120(21):210501, May 2018.

[144] Jun Li, Xiaodong Yang, Xinhua Peng, and Chang-Pu Sun. Hybrid quantum-classical approach to quantum optimal control. *Phys. Rev. Lett.*, 118:150503, Apr 2017.

[145] S. Lloyd. Universal Quantum Simulators. *Science*, 273(5278):1073–1078, 1996.

[146] Alan Aspuru-Guzik, Anthony D. Dutoi, Peter J. Love, and Martin Head-Gordon. Simulated Quantum Computation of Molecular Energies. *Science*, 309(5741):1704–1707, 2005.

[147] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White. Towards quantum chemistry on a quantum computer. *Nature Chemistry*, 2(2):106–111, 2010.

[148] Antoine Georges, Gabriel Kotliar, Werner Krauth, and Marcelo J. Rozenberg. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Reviews of Modern Physics*, 68(1):13–125, 1996.

[149] J. M. Kreula, S. R. Clark, and D. Jaksch. Non-linear quantum-classical scheme to simulate non-equilibrium strongly correlated fermionic many-body dynamics. *Scientific Reports*, 6(1):32940, 2016.

[150] Juha M. Kreula, Laura García-Álvarez, Lucas Lamata, Stephen R. Clark, Enrique Solano, and Dieter Jaksch. Few-qubit quantum-classical simulation of strongly correlated lattice fermions. *EPJ Quantum Technology*, 3(1):11, 2016.

[151] Bela Bauer, Dave Wecker, Andrew J. Millis, Matthew B. Hastings, and Matthias Troyer. Hybrid quantum-classical approach to correlated materials. *Phys. Rev. X*, 6:031045, Sep 2016.

[152] Trevor Keen, Thomas Maier, Steven Johnston, and Pavel Lougovski. Quantum-classical simulation of two-site dynamical mean-field theory on

noisy quantum hardware. *Quantum Science and Technology*, 5(3):035001, 2020.

[153] I. Rungger, N. Fitzpatrick, H. Chen, C. H. Alderete, H. Apel, A. Cowtan, A. Patterson, D. Munoz Ramo, Y. Zhu, N. H. Nguyen, E. Grant, S. Chretien, L. Wossnig, N. M. Linke, and R. Duncan. Dynamical mean field theory algorithm and experiment on quantum computers. *arXiv e-prints*, Oct 2019.

[154] Walter Metzner and Dieter Vollhardt. Correlated Lattice Fermions in d = 8 Dimensions. *Physical Review Letters*, 62(9):1066–1066, 1989.

[155] Michael Potthoff. Two-site dynamical mean-field theory. *Phys. Rev. B*, 64:165114, Oct 2001.

[156] Ekkehard Lange. Renormalized Versus Unrenormalized Perturbation-Theoretical Approaches to the Mott Transition. *Modern Physics Letters B*, 12(22):915–919, 1998.

[157] Lawrence W. Cheuk, Matthew A. Nichols, Katherine R. Lawrence, Melih Okan, Hao Zhang, Ehsan Khatami, Nandini Trivedi, Thereza Paiva, Marcos Rigol, and Martin W. Zwierlein. Observation of spatial charge and spin correlations in the 2D Fermi-Hubbard model. *Science*, 353(6305):1260–1264, Sep 2016.

[158] Elmer V. H. Doggen and Jami J. Kinnunen. Momentum-resolved spectroscopy of a fermi liquid. *Scientific Reports*, 5(1):9539, 2015.

[159] P. Jordan and E. Wigner. Uber das Paulische Aquivalenzverbot. *Zeitschrift fur Physik*, 47(9-10):631–651, 1928.

[160] Juan Pablo Paz and Augusto Roncaglia. Quantum gate arrays can be programmed to evaluate the expectation value of any operator. *Phys. Rev. A*,

68:052316, Nov 2003.

[161] Masuo Suzuki. Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Communications in Mathematical Physics*, 51(2):183–190, 1976.

[162] V.V. Shende, S.S. Bullock, and I.L. Markov. Synthesis of quantum-logic circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 25(6):1000–1010, 2006.

[163] T. Bækkegaard, L. B. Kristensen, N. J. S. Loft, C. K. Andersen, D. Petrosyan, and N. T. Zinner. Realization of efficient quantum gates with a superconducting qubit-qutrit circuit. *Scientific Reports*, 9(1):13389, 2019.

[164] 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, 2017.

[165] Morten Kjaergaard, Mollie E. Schwartz, Jochen Braumüller, Philip Krantz, Joel I-Jan Wang, Simon Gustavsson, and William D. Oliver. Superconducting Qubits: Current State of Play. *e-prints*, May 2019.

[166] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[167] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, Nov 1996.

[168] Carlo Cafaro and Stefano Mancini. Quantum stabilizer codes for correlated and asymmetric depolarizing errors. *Phys. Rev. A*, 82:012306, Jul 2010.

[169] E. Knill. Quantum computing with realistically noisy devices. *Nature*, 434(7029):39–44, 2005.

[170] Which we define in this case to be 3 subsequent iterations with V_{new} values within 0.4% of each other. However, this definition should be reconsidered on an individual basis depending on the computational resources available and desired accuracy of the final solution.

[171] Matthew Otten, Cristian L Cortes, and Stephen K Gray. Noise-resilient quantum dynamics using symmetry-preserving ansatzes. *arXiv preprint arXiv:1910.06284*, 2019.

[172] 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*, 574(7779):505–510, 2019.

[173] Andrew W. Cross, Lev S. Bishop, Sarah Sheldon, Paul D. Nation, and Jay M. Gambetta. Validating quantum computers using randomized model circuits. *Physical Review A*, 100(3):032328, Sep 2019.

[174] Quantinuum announces quantum volume 4096 achievement.

[175] Andrew W Senior, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Žídek, Alexander WR Nelson, Alex Bridgland, et al. Improved protein structure prediction using potentials from deep learning. *Nature*, 577(7792):706–710, 2020.

[176] K. Akiyama et al. First m87 event horizon telescope results. i. the shadow of the supermassive black hole. *The Astrophysical Journal*, 875:1–17, 2019.

[177] Christina V Theodoris, Ping Zhou, Lei Liu, Yu Zhang, Tomohiro Nishino, Yu Huang, Aleksandra Kostina, Sanjeev S Ranade, Casey A Gifford, Vladimir Uspenskiy, et al. Network-based screen in ipsc-derived cells reveals therapeutic candidate for heart valve disease. *Science*, 371(6530), 2021.

[178] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[179] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[180] Anders Krogh. What are artificial neural networks? *Nature biotechnology*, 26(2):195–197, 2008.

[181] Yuanqing Lin, Fengjun Lv, Shenghuo Zhu, Ming Yang, Timothee Cour, Kai Yu, Liangliang Cao, and Thomas Huang. Large-scale image classification: fast feature extraction and svm training. In *CVPR 2011*, pages 1689–1696. IEEE, 2011.

[182] Hieu Pham, Zihang Dai, Qizhe Xie, and Quoc V Le. Meta pseudo labels. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11557–11568, 2021.

[183] David G Lowe. Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision*, volume 2, pages 1150–1157. Ieee, 1999.

[184] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[185] Virginia R de Sa. Learning classification with unlabeled data. In *Advances in neural information processing systems*, pages 112–119, 1994.

[186] Zhirong Wu, Alexei A Efros, and Stella X Yu. Improving generalization via scalable neighborhood component analysis. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 685–701, 2018.

[187] Zhirong Wu, Yuanjun Xiong, Stella Yu, and Dahua Lin. Unsupervised feature learning via non-parametric instance-level discrimination. *arXiv preprint arXiv:1805.01978*, 2018.

[188] Olivier Henaff. Data-efficient image recognition with contrastive predictive coding. In *International Conference on Machine Learning*, pages 4182–

4192. PMLR, 2020.

[189] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

[190] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9729–9738, 2020.

[191] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations, 2020.

[192] Guangrun Wang, Keze Wang, Guangcong Wang, Phillip HS Torr, and Liang Lin. Solving inefficiency of self-supervised representation learning. *arXiv preprint arXiv:2104.08760*, 2021.

[193] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. Masked autoencoders are scalable vision learners. *arXiv preprint arXiv:2111.06377*, 2021.

[194] Google AI Quantum et al. Hartree-fock on a superconducting qubit quantum computer. *Science*, 369(6507):1084–1089, 2020.

[195] Leo Zhou, Sheng-Tao Wang, Soonwon Choi, Hannes Pichler, and Mikhail D Lukin. Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *Physical Review X*, 10(2):021067, 2020.

[196] He Ma, Marco Govoni, and Giulia Galli. Quantum simulations of materials on near-term quantum computers. *npj Computational Materials*, 6(1):1–8, 2020.

[197] 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, 2019.

[198] Maria Schuld, Alex Bocharov, Krysta M. Svore, and Nathan Wiebe. Circuit-centric quantum classifiers. *Phys. Rev. A*, 101:032308, Mar 2020.

[199] Marcello Benedetti, Delfina Garcia-Pintos, Oscar Perdomo, Vicente Leyton-Ortega, Yunseong Nam, and Alejandro Perdomo-Ortiz. A generative modeling approach for benchmarking and training shallow quantum circuits. *npj Quantum Information*, 5(1):1–9, 2019.

[200] Christa Zoufal, Aurélien Lucchi, and Stefan Woerner. Quantum generative adversarial networks for learning and loading random distributions. *npj Quantum Information*, 5(1):1–9, 2019.

[201] 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. *IEEE Access*, 8:141007–141024, 2020.

[202] Owen Lockwood and Mei Si. Reinforcement learning with quantum variational circuit. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 16, pages 245–251, 2020.

[203] V. Saggio, B. E. Asenbeck, A. Hamann, T. Strömberg, P. Schiansky, V. Dunjko, N. Friis, N. C. Harris, M. Hochberg, D. Englund, and et al. Experimental quantum speed-up in reinforcement learning agents. *Nature*, 591(7849):229–233, Mar 2021.

[204] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Physical Review A*, 98(3), Sep 2018.

[205] Kerstin Beer, Dmytro Bondarenko, Terry Farrelly, Tobias J. Osborne, Robert Salzmann, Daniel Scheiermann, and Ramona Wolf. Training deep quantum neural networks. *Nature Communications*, 11(1):808, 2020.

[206] Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 4(4):043001, 2019.

[207] Amira Abbas, David Sutter, Christa Zoufal, Aurélien Lucchi, Alessio Figalli, and Stefan Woerner. The power of quantum neural networks. *Nature Computational Science*, 1(6):403–409, 2021.

[208] Hsin-Yuan Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R McClean. Power of data in quantum machine learning. *Nature communications*, 12(1):1–9, 2021.

[209] Johannes Bausch. Recurrent quantum neural networks. *Advances in Neural Information Processing Systems*, 33, 2020.

[210] Andrea Skolik, Jarrod R McClean, Masoud Mohseni, Patrick van der Smagt, and Martin Leib. Layerwise learning for quantum neural networks. *Quantum Machine Intelligence*, 3(1):1–11, 2021.

[211] Patrice Y Simard, David Steinkraus, John C Platt, et al. Best practices for convolutional neural networks applied to visual document analysis. In *Icdar*, volume 3, 2003.

[212] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[213] Andriy Mnih and Koray Kavukcuoglu. Learning word embeddings efficiently with noise-contrastive estimation. In *Conference on Neural Information Processing Systems*, 2013.

[214] Aditya Grover and Jure Leskovec. Node2vec: Scalable feature learning for networks. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2016.

[215] Amy X Lu, Haoran Zhang, Marzyeh Ghassemi, and Alan M Moses. Self-supervised contrastive learning of protein representations by mutual information maximization. *bioRxiv*, 2020.

[216] Sabrina Jaeger, Simone Fulle, and Samo Turk. Mol2vec: unsupervised machine learning approach with chemical intuition. *Journal of chemical information and modeling*, 58(1):27–35, 2018.

[217] Ke-Lin Du and Madisetti NS Swamy. *Neural networks and statistical learning*. Springer Science & Business Media, 2013.

[218] Kihyuk Sohn. Improved deep metric learning with multi-class n-pair loss objective. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, pages 1857–1865, 2016.

[219] Herbert Robbins and Sutton Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951.

[220] Raia Hadsell, Sumit Chopra, and Yann LeCun. Dimensionality reduction by learning an invariant mapping. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 2, pages 1735–1742. IEEE, 2006.

[221] Andrea Mari, Thomas R. Bromley, Josh Izaac, Maria Schuld, and Nathan Killoran. Transfer learning in hybrid classical-quantum neural networks.

Quantum, 4:340, Oct 2020.

[222] Seth Lloyd, Maria Schuld, Aroosa Ijaz, Josh Izaac, and Nathan Killoran. Quantum embeddings for machine learning, 2020.

[223] Phuc Q Le, Fangyan Dong, and Kaoru Hirota. A flexible representation of quantum images for polynomial preparation, image compression, and processing operations. *Quantum Information Processing*, 10(1):63–84, 2011.

[224] Yi Zhang, Kai Lu, Yinghui Gao, and Mo Wang. NEQR: a novel enhanced quantum representation of digital images. *Quantum Information Processing*, 12(8):2833–2860, 2013.

[225] Sukin Sim, Peter D. Johnson, and Alán Aspuru-Guzik. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12):1900070, 2020/06/08 2019.

[226] Maria Schuld, Ville Bergholm, Christian Gogolin, Josh Izaac, and Nathan Killoran. Evaluating analytic gradients on quantum hardware. *Phys. Rev. A*, 99:032331, Mar 2019.

[227] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[228] Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009.

[229] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703*, 2019.

[230] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[231] 2021.

[232] Taylor L Patti, Jean Kossaifi, Anima Anandkumar, and Susanne F Yelin. Variational quantum optimization with multi-basis encodings. *arXiv preprint arXiv:2106.13304*, 2021.

[233] Yanping Huang, Youlong Cheng, Ankur Bapna, Orhan Firat, Dehao Chen, Mia Chen, HyoukJoong Lee, Jiquan Ngiam, Quoc V Le, Yonghui Wu, et al. Gpipe: Efficient training of giant neural networks using pipeline parallelism. *Advances in neural information processing systems*, 32:103–112, 2019.

[234] Ekin D Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, and Quoc V Le. Autoaugment: Learning augmentation strategies from data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 113–123, 2019.

[235] Nguyen Huu Phong and Bernardete Ribeiro. Rethinking recurrent neural networks and other improvements for image classification. *arXiv preprint arXiv:2007.15161*, 2020.

[236] Alexander Kolesnikov, Xiaohua Zhai, and Lucas Beyer. Revisiting self-supervised visual representation learning. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 1920–1929, 2019.

[237] Xiu-Zhe Luo, Jin-Guo Liu, Pan Zhang, and Lei Wang. Yao. jl: Extensible, efficient framework for quantum algorithm design. *Quantum*, 4:341, 2020.

[238] Yasunari Suzuki, Yoshiaki Kawase, Yuya Masumura, Yuria Hiraga, Masahiro Nakadai, Jiabao Chen, Ken M Nakanishi, Kosuke Mitarai, Ryosuke Imai, Shiro Tamiya, et al. Qulacs: a fast and versatile quantum circuit simulator for research purpose. *Quantum*, 5:559, 2021.

[239] 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*, 11:369–395, 2020.

[240] Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, M Sohaib Alam, Shahnawaz Ahmed, Juan Miguel Arrazola, Carsten Blank, Alain Delgado, Soran Jahangiri, et al. Pennylane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2018.

[241] Michael Broughton, Guillaume Verdon, Trevor McCourt, Antonio J Martinez, Jae Hyeon Yoo, Sergei V Isakov, Philip Massey, Murphy Yuezhen Niu, Ramin Halavati, Evan Peters, et al. Tensorflow quantum: A software framework for quantum machine learning. *arXiv preprint arXiv:2003.02989*, 2020.

[242] Gael Sentís, John Calsamiglia, Ramón Munoz-Tapia, and Emilio Bagan. Quantum learning without quantum memory. *Scientific reports*, 2(1):1–8, 2012.

[243] Unai Alvarez-Rodriguez, Lucas Lamata, Pablo Escandell-Montero, José D Martín-Guerrero, and Enrique Solano. Supervised quantum learning without measurements. *Scientific reports*, 7(1):1–9, 2017.

[244] Mohammad H Amin, Evgeny Andriyash, Jason Rolfe, Bohdan Kulchyt-

skyy, and Roger Melko. Quantum boltzmann machine. *Physical Review X*, 8(2):021050, 2018.

[245] Ming Gong, He-Liang Huang, Shiyu Wang, Chu Guo, Shaowei Li, Yulin Wu, Qingling Zhu, Youwei Zhao, Shaojun Guo, Haoran Qian, et al. Quantum neuronal sensing of quantum many-body states on a 61-qubit programmable superconducting processor. *arXiv preprint arXiv:2201.05957*, 2022.

[246] Tomasz Szołdra, Piotr Sierant, Maciej Lewenstein, and Jakub Zakrzewski. Unsupervised detection of decoupled subspaces: many-body scars and beyond. *arXiv preprint arXiv:2201.07151*, 2022.

[247] Gael Sentís, Alex Monras, Ramon Muñoz-Tapia, John Calsamiglia, and Emilio Bagan. Unsupervised classification of quantum data. *Physical Review X*, 9(4):041029, 2019.

[248] Elija Perrier, Akram Youssry, and Chris Ferrie. Qdataset: Quantum datasets for machine learning. *arXiv preprint arXiv:2108.06661*, 2021.

[249] Louis Schatzki, Andrew Arrasmith, Patrick J Coles, and M Cerezo. Entangled datasets for quantum machine learning. *arXiv preprint arXiv:2109.03400*, 2021.

[250] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu, et al. Quantum computational advantage using photons. *Science*, 370(6523):1460–1463, 2020.

[251] Amira Abbas, David Sutter, Alessio Figalli, and Stefan Woerner. Effective dimension of machine learning models. *arXiv preprint arXiv:2112.04807*, 2021.

[252] Arthur Pesah, M Cerezo, Samson Wang, Tyler Volkoff, Andrew T Sornborger, and Patrick J Coles. Absence of barren plateaus in quantum convolutional neural networks. *Physical Review X*, 11(4):041011, 2021.

[253] Edward Grant, Leonard Wossnig, Mateusz Ostaszewski, and Marcello Benedetti. An initialization strategy for addressing barren plateaus in parametrized quantum circuits. *Quantum*, 3:214, 2019.

[254] Marco Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J Coles. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature communications*, 12(1):1–12, 2021.

[255] Adriano Barenco, Andre Berthiaume, David Deutsch, Artur Ekert, Richard Jozsa, and Chiara Macchiavello. Stabilization of quantum computations by symmetrization. *SIAM Journal on Computing*, 26(5):1541–1557, 1997.

[256] Harry Buhrman, Richard Cleve, John Watrous, and Ronald De Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902, 2001.

[257] Ronald De Wolf. Quantum computing: Lecture notes. *arXiv preprint arXiv:1907.09415*, 2019.

[258] Sam Stanwyck, Harun Bayraktar, and Tim Costa. cuquantum: Accelerating quantum circuit simulation on gpus. *Bulletin of the American Physical Society*, 2022.

[259] Francisco Javier Gil Vidal and Dirk Oliver Theis. Input redundancy for parameterized quantum circuits. *Frontiers in Physics*, 8:297, 2020.

[260] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103(3):032430, 2021.

[261] 2022.

[262] Loïc Henriet, Lucas Beguin, Adrien Signoles, Thierry Lahaye, Antoine Browaeys, Georges-Olivier Reymond, and Christophe Jurczak. Quantum computing with neutral atoms. *Quantum*, 4:327, 2020.

[263] Fulvio Flamini, Nicolo Spagnolo, and Fabio Sciarrino. Photonic quantum information processing: a review. *Reports on Progress in Physics*, 82(1):016001, 2018.

A | Appendix for Chapter 2

A.1 Figure 3.7 parameters

Here we present the random couplings used to generate Fig. 3.7. Note that the structure of the $\vec{V}$ matrix corresponds to sites connected in a chain topology. One could translate this to a ring topology by adding non-zero values in the lower-left and upper-right corners of the matrix, or a fully connected topology by filling all elements with non-zero values.

$$\vec{V} = \begin{bmatrix} 0 & 0.981 & 0 & 0 & 0 & 0 & 0 \\ 0.981 & 0 & 1.024 & 0 & 0 & 0 & 0 \\ 0 & 1.024 & 0 & 1.170 & 0 & 0 & 0 \\ 0 & 0 & 1.170 & 0 & 0.986 & 0 & 0 \\ 0 & 0 & 0 & 0.986 & 0 & 1.003 & 0 \\ 0 & 0 & 0 & 0 & 1.003 & 0 & 1.035 \\ 0 & 0 & 0 & 0 & 0 & 1.035 & 0 \end{bmatrix} \hbar\omega \tag{A.1}$$

$$\vec{\chi} = \begin{bmatrix} 1.149 \\ 1.002 \\ 1.192 \\ 1.079 \\ 1.089 \\ 0.963 \\ 1.024 \end{bmatrix} \hbar\omega \tag{A.2}$$

B | Appendix for Chapter 3

B.1 Calculating the impurity Green's function with a reduced number of terms

To begin, we rewrite equations 4.14 and 4.15 in a simplified format

$$G^{<}_{imp}(\tau) = \frac{i}{4}(xx + i(xy) - i(yx) + yy), \tag{B.1}$$

where the notation $\alpha\beta = \langle\hat{\sigma}_1^{\alpha}\hat{\mathcal{U}}^{\dagger}(\tau)\hat{\sigma}_1^{\beta}\hat{\mathcal{U}}(\tau)\rangle$. Note that the terms of $G^{>}_{imp}(\tau)$ are in fact all complex conjugates of terms in $G^{<}_{imp}(\tau)$. Thus, we can write

$$G^{>}_{imp}(\tau) = -\frac{i}{4}(\bar{x}\bar{x} + i(\bar{x}\bar{y}) - i(\bar{y}\bar{x}) + \bar{y}\bar{y}). \tag{B.2}$$

Substituting these into equation 4.9 we get

$$G^{R}_{imp}(\tau) = -\frac{i}{4}(\bar{x}\bar{x} + i(\bar{x}\bar{y}) - i(\bar{y}\bar{x}) + \bar{y}\bar{y}) - \frac{i}{4}(xx + i(xy) - i(yx) + yy) \tag{B.3}$$

$$= -\frac{i}{4}(\bar{x}\bar{x} + xx + yy + \bar{y}\bar{y} + i(\bar{x}\bar{y} + xy - (\bar{y}\bar{x} + yx))) \tag{B.4}$$

$$iG^{R}_{imp}(\tau) = -\frac{i}{4}(2Re(xx) + 2Re(yy) + i(2Re(xy) - 2Re(yx))). \tag{B.5}$$

In the half-filled case we expect $iG^{R}_{imp}(\tau)$ to be real. This is due to particle-hole symmetry as explained by [150] equation 33. It is also clear in the re-expressed form of [150] equation 34 that there are no real parts. From this we can conclude

that equation B.5 can be simplified:

$$iG^R_{imp}(\tau) = \frac{1}{2}(Re(xx) + Re(yy)) \tag{B.6}$$

Indeed numerical calculations show that the terms $Re(xy)$, $Re(yx)$ in equation B.5 are equal.

One final observation in the half-filled case is that the terms $Re(xx)$, $Re(yy)$ are also equal. This is not fully understood yet, but has been verified by numerical results. This reduces equation B.6 even further to just:

$$iG^R_{imp}(\tau) = Re(xx) \tag{B.7}$$

We have shown that using only equation B.7 recreates the results in [150]. Remarkably, this means in the half-filled case we have now reduced the number of quantum circuits required for our calculation from 16 to 1. Combined with the removal of the C and D gates in [150] figure 3, it's clear to see why the half-filled case is the most amenable for near-term quantum computers

B.2 Quantum circuit to obtain the impurity Green's function

Here we show that measurement of the Green's function circuit seen in figure 4.4 obtains the term $Re[\langle\psi|\,\hat{\sigma}^x_1\hat{\mathcal{U}}^\dagger\hat{\sigma}^x_1\hat{\mathcal{U}}\,|\psi\rangle]$.

Let $|\Phi\rangle$ be the combined state of all 5 qubits; q_0, q_1, q_2, q_3, q_4. Let ψ be the input state of the four work qubits; q_1, q_2, q_3, q_4. Stepping through the circuit:

$$
\begin{aligned}
|\Phi\rangle &= |0\rangle_0 \otimes |\psi\rangle_{1,2,3,4} \\
H_0 &\rightarrow \frac{1}{\sqrt{2}}[(|0\rangle_0 + |1\rangle_0) \otimes |\psi\rangle_{1,2,3,4}] \\
\hat{X}_0\ \hat{\sigma}^x_{0,1}\ \hat{X}_0 &\rightarrow \frac{1}{\sqrt{2}}[|0\rangle_0 \otimes \hat{\sigma}^x_1 |\psi\rangle_{1,2,3,4} + |1\rangle_0 \otimes |\psi\rangle_{1,2,3,4}]
\end{aligned}
$$

Where $\hat{\sigma}^x_{0,1}$ is a controlled $\hat{\sigma}^x$ gate with qubit 0 as the control and qubit 1 as the target. Continuing, dropping the subscript of $|\psi\rangle_{1,2,3,4}$ for ease of reading:

$$
\begin{aligned}
\hat{\mathcal{U}}_{1,2,3,4} &\rightarrow \frac{1}{\sqrt{2}}[|0\rangle_0 \otimes \hat{\mathcal{U}}\hat{\sigma}^x_1 |\psi\rangle + |1\rangle_0 \otimes \hat{\mathcal{U}} |\psi\rangle] \\
\hat{\sigma}^x_{0,1} &\rightarrow \frac{1}{\sqrt{2}}[|0\rangle_0 \otimes \hat{\mathcal{U}}\hat{\sigma}^x_1 |\psi\rangle + |1\rangle_0 \otimes \hat{\sigma}^x_1\hat{\mathcal{U}} |\psi\rangle] \\
H_0 &\rightarrow \frac{1}{2}[(|0\rangle_1 + |1\rangle_1) \otimes \hat{\mathcal{U}}\hat{\sigma}^x_1 |\psi\rangle + (|0\rangle_1 - |1\rangle_1) \otimes \hat{\sigma}^x_1\hat{\mathcal{U}} |\psi\rangle]
\end{aligned}
$$

Multiplying out and grouping terms we get:

$$
|\Phi\rangle = \frac{1}{2}[|0\rangle_1 \otimes (\hat{\mathcal{U}}\hat{\sigma}^x_1 |\psi\rangle + \hat{\sigma}^x_1\hat{\mathcal{U}} |\psi\rangle) + |1\rangle_1 \otimes (\hat{\mathcal{U}}\hat{\sigma}^x_1 |\psi\rangle - \hat{\sigma}^x_1\hat{\mathcal{U}} |\psi\rangle)]
$$

The result of a measurement in the Z basis on the ancilla qubit corresponds to the expectation value $\langle\Phi|\hat{\sigma}^z_0|\Phi\rangle$. Keeping in mind $\hat{\sigma}^z |0\rangle = |0\rangle$ and $\hat{\sigma}^z |1\rangle = -|1\rangle$:

$$\langle\Phi|\hat{\sigma}_0^z|\Phi\rangle = \frac{1}{4}[\langle 0|_1 \otimes (\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger + \langle\psi|\,\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x) + \langle 1|_1 \otimes (\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger - \langle\psi|\,\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x)]$$
$$\cdot [|0\rangle_1 \otimes (\hat{\mathcal{U}}\hat{\sigma}_1^x\,|\psi\rangle + \hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle) - |1\rangle_1 \otimes (\hat{\mathcal{U}}\hat{\sigma}_1^x\,|\psi\rangle - \hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle)]$$

$$= \frac{1}{4}[(\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger + \langle\psi|\,\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x)\cdot(\hat{\mathcal{U}}\hat{\sigma}_1^x\,|\psi\rangle + \hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle)$$
$$- (\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger - \langle\psi|\,\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x)\cdot(\hat{\mathcal{U}}\hat{\sigma}_1^x\,|\psi\rangle - \hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle)]$$

$$= \frac{1}{4}[\cancel{\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger\hat{\mathcal{U}}\hat{\sigma}_1^x\,|\psi\rangle} + \langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle + \langle\psi|\,\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\hat{\sigma}_1^x\,|\psi\rangle$$
$$+ \cancel{\langle\psi|\,\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle} - \cancel{\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger\hat{\mathcal{U}}\hat{\sigma}_1^x\,|\psi\rangle} + \langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle$$
$$+ \langle\psi|\,\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\hat{\sigma}_1^x\,|\psi\rangle - \cancel{\langle\psi|\,\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle}]$$

$$= \frac{1}{4}[2\,\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle + 2\,\langle\psi|\,\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\hat{\sigma}_1^x\,|\psi\rangle]$$
$$= \frac{1}{2}[\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle + \langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle^*]$$

$$\langle\Phi|\hat{\sigma}_0^z|\Phi\rangle = Re[\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle]$$

Notably, it would be possible to also obtain $Im[\langle\psi|\,\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger\hat{\sigma}_1^x\hat{\mathcal{U}}\,|\psi\rangle]$ by repeating the above process measuring in the Y basis instead, allowing the construction of the full expectation $\langle\hat{\sigma}_1^x\hat{\mathcal{U}}^\dagger(\tau)\hat{\sigma}_1^x\hat{\mathcal{U}}(\tau)\rangle$. However, we showed in B.1 that this is no longer needed and as such the single circuit in 4.4 is sufficient.

B.3 Definition of single-qubit gates

The single-qubit unitary gates, U_1, U_2 and U_3, are defined as [79]

$$\hat{U}_3(\theta,\phi,\lambda) = \begin{pmatrix} \cos(\theta/2) & -e^{i\lambda}\sin(\theta/2) \\ e^{i\phi}\sin(\theta/2) & e^{i\lambda+i\phi}\cos(\theta/2) \end{pmatrix} , \tag{B.8}$$

$$\hat{U}_2(\phi,\lambda) = \hat{U}_3(\pi/2,\phi,\lambda) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -e^{i\lambda} \\ e^{i\phi} & e^{i(\phi+\lambda)} \end{pmatrix} , \tag{B.9}$$

$$\hat{U}_1(\lambda) = \hat{U}_3(0,0,\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\lambda} \end{pmatrix} . \tag{B.10}$$

Although the $\hat{U}_3$ gate is universal, it is useful to distinguish these three separate gates operations for noise modelling purposes. This is because the $\hat{U}_1$, $\hat{U}_2$ and $\hat{U}_3$ gates require 0, 1 and 2 X90 pulses respectively. This in turn affects how long it takes to run each gate.