

Quantum algorithms for near-term quantum computers



Lewis William Anderson

Keble College, University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas term 2024

Acknowledgements

Despite a number of months spent physically isolated from them, this thesis would not have been possible without the support of many people.

Firstly, I'd like to thank my supervisor, Dieter Jaksch. Dieter, throughout my DPhil, you have allowed me the freedom to pursue my own ideas while providing a critical eye and scientific guidance without which, I'd not be the physicist that I am today. I'd also like to thank my collaborators Samuel Albanie, Panagiotis Barkotsos, Jason Crain, Thomas O'Leary, Ivano Tavernelli, Weidi Xie and, in particular; Martin Kiffner and Ben Jaderberg. Martin, I thank you for your constant guidance and support, and Ben, I am grateful for our time spent working together and the important, if not timely, lessons we learnt about the value of unit testing. Beyond those that I have directly worked with, I'd like to thank the Jaksch Group past and present. In particular, to Cameron, Chris, Jess and Nik, I thank you for both stimulating discussions as well as putting up with my snoring in the office.

I would not have made it to this point in my DPhil without the friendship of many amazing people I met in Oxford. I am grateful to my housemates George, James, Liam and Phil. Together, we have experienced the trials and tribulations of graduate study (as opposed to the 'real world' of work) and, more importantly, been able to identify the best Indian takeaway in Oxford and a definitive ranking of all commercially available Christmas puddings. Additionally, my time in Oxford would not have been what it was without Keble College Boat Club and its members, past and present. Special thanks goes to Charlie, Jordie, Marko, and of course Raphi. Your friendship has been second to none and our trips to Hank's, without which this thesis may have been completed far sooner, have been incredibly memorable - or so I'm told.

Finally, and most importantly, I'd like to thank my parents Jane and Phil. Who'd have thought that after all these years after having to do so with my GCSE poetry assignments, you would still be spell-checking my work! You've been immensely supportive; always managed to get me out of whatever scrapes I find myself in; and experienced all (with a bit more on top) of the stress I felt during this DPhil. Without your love over the last four years, or preceding twenty-two, I would not be where I am today.

Abstract

One of the most promising applications of quantum computers is computing the ground state energy of quantum mechanical systems. Realisation of this potential requires a comprehensive understanding and refinement of algorithms tailored for such systems, along with pinpointing applications where their impact can be most significant. This thesis aims to contribute to this understanding by investigating two quantum algorithms and exploring their application in novel contexts where they have not been previously employed. The objective is twofold: firstly, to establish a framework for their effective utilisation, and secondly, to assess their feasibility on near-term quantum processors. We focus here on systems of interacting Bosons and Bosons interacting with Fermions which, in contrast to purely Fermionic systems, have seen comparatively little research.

In quantum chemistry, variational quantum algorithms (VQAs) have gained traction for simulating strongly-bound systems with covalent bonds. However, they have so far seen little application to non-covalent molecular systems. We propose the first application of quantum computers to the weakly-bound intermolecular and non-covalently bonded regime. To achieve this, we develop a coarse-grained representation of the electronic response, wherein molecular charge fields are represented by Bosonic oscillators. This is ideally suited for determining the ground state of weakly interacting molecules using a VQA. We require qubit numbers that grow linearly with the number of molecules and derive scaling behaviour for the number of circuits and measurements required, which compare favourably to traditional variational quantum eigensolver (VQE) methods. We perform experiments using IBM superconducting processors which demonstrates our algorithm’s efficacy in resolving van der Waals dispersion energy directly from zero-point quantum fluctuations for non-polar molecule pairs. We thus conclude that by using our coarse-grained framework, current-generation quantum hardware can effectively explore energies in the weakly bound, yet chemically significant, intermolecular regime. Additionally, we conduct experiments on simulated quantum computers, extending the analysis to systems with three, four, and five oscillators and those with anharmonic onsite binding potentials. The incorporation of the latter, often neglected in large systems using classical computational methods, introduces minimal computational overhead in our quantum approach.

We build on this work by studying the use of quantum computers to solve the Schwinger model, an archetypal lattice gauge theory representing quantum electrodynamics in one spatial dimension. In order to avoid some of the drawbacks of the VQAs, we investigate the use of the quantum subspace expansion (QSE) algorithm implemented using qubitisation. Numerical simulations, including the effect of measurement noise, are used to extrapolate the resources required to compute the ground state of the model to a desired accuracy. These simulations represent the first application of QSE to the ground state problem for a lattice gauge theory. We perform a full resource estimate for the gate cost of implementing the qubitisation procedure, including the development of a novel method of performing the qubitisation for a lattice gauge theory that scales linearly, up to a logarithmic factor, with lattice size. Combining this with our numerical experiments allows us to evaluate the feasibility of applying our algorithm to near-term or early fault-tolerant quantum processors.

Contents

List of Figures	xi
List of Tables	xiii
List of Abbreviations	xv
1 Introduction	1
1.1 Quantum algorithms for quantum systems	2
1.1.1 Quantum phase estimation	4
1.1.2 Variational quantum eigensolver	5
1.1.3 Quantum subspace expansion	9
1.2 Thesis structure	12
1.3 Publications	14
2 Variational quantum algorithm for coarse-grained quantum chemistry	15
2.1 Introduction	16
2.1.1 Computational chemistry on classical and quantum computers	16
2.1.2 Our contribution	19
2.1.3 Chapter layout	22
2.2 Coarse-grained modelling on digital quantum computers	22
2.2.1 Coarse-grained model	22
2.2.2 Algorithm description	26
2.2.3 Measurement cost	29

2.2.4	Variational ansatz	33
2.3	Experimental results using IBM quantum computers	34
2.4	Extending the model	38
2.4.1	Three-dimensional QDOs	38
2.4.2	Induction, anharmonic and non-linear terms	40
2.5	Further experimental results	42
2.5.1	Anharmonic systems	42
2.5.2	Many oscillators	45
2.6	Discussion and conclusions	46
3	Solving lattice gauge theories using the quantum Krylov algorithm and qubitisation	51
3.1	Introduction	52
3.1.1	Classical and quantum computing for gauge field theories . .	52
3.1.2	Our contribution	55
3.1.3	Chapter layout	56
3.2	Hamiltonian lattice formulation of the single flavour Schwinger model	57
3.2.1	Full Hamiltonian	57
3.2.2	Gauging out Bosonic fields	59
3.3	Quantum subspace expansion and quantum singular value transforms	60
3.3.1	Block encoding and QSVT	60
3.3.2	Generating Krylov basis using QSVT	65
3.4	Numerical experiments	67
3.4.1	Modelling measurement noise	70
3.4.2	Dealing with measurement noise	71
3.4.3	Results	72
3.5	Estimating resource requirements	76
3.5.1	Sketch of proofs for gate costs	76

3.5.2	Total resource cost of algorithm	78
3.6	Discussion and conclusions	81
4	Conclusions and outlook	85
Appendices		
A	In depth analysis for digital quantum simulation of Bosonic systems	91
A.1	London dispersion in linear molecules and one-dimensional quantum Drude oscillators	91
A.2	Truncated Fock space	92
A.3	Decomposition and measurement of Hamiltonian	93
A.3.1	Scaling in the total number of Pauli operators	93
A.3.2	Scaling of number of circuits required	94
A.3.3	Number of circuits for three-dimensional QDOs	99
A.4	Shot requirements	100
B	Experimental details on variational quantum algorithms for coarse grained systems	111
B.1	Noise mitigation for anharmonic experiments	111
B.2	IBM quantum hardware	113
B.3	Symplectic eigenvalues of harmonic oscillators with linear coupling .	114
C	Resource analysis for block encoding Schwinger model	115
C.1	Preparing $ G\rangle$ and U using linear combination of unitaries	115
C.1.1	Implementing U	115
C.1.2	Preparing $ G\rangle_a$	116
C.1.3	Applying signs to $\hat{P}_i$'s and preparing $ \tilde{G}_a\rangle$	118
C.2	Gate costs for block encoding Schwinger model	118
C.2.1	Some useful mathematics	118
C.2.2	Pauli decomposition of the Hamiltonian	120

C.2.3	Generating the block encoding state $ G\rangle$	124
C.2.4	T gate cost of rotation gates	134
C.2.5	Applying the block encoding unitary U	136
C.2.6	Total gate cost for Π_φ and $\tilde{\Pi}_\varphi$	138
C.2.7	Total gate cost of the full algorithm	139
D	Additional simulations of QSE with measurement noise	141
E	Hardware information for current generation quantum processors	145
	References	149

List of Figures

2.1	Schematic of the VQE optimisation procedure for coarse grained model.	28
2.2	Estimated number of shots required to measure $\langle \hat{H}(R_{\text{vdW}}) \rangle$ with error $\epsilon = 10\% \times \Delta E(R_{\text{vdW}})$ for a range of different molecules. . . .	32
2.3	Circuits used in VQE experiments for coarse-grained model.	34
2.4	Experimental measurements used for error subtraction in coarse-grained model.	37
2.5	London dispersion energy (ΔE) calculated using variational quantum algorithm for one-dimensional quantum Drude oscillators computing interacting I_2 dimers.	39
2.6	Interaction energy for system of two coupled anharmonic oscillators with optimisation performed on simulated <i>ibmq_montreal</i> and interaction energy for 1D anharmonic oscillator in external field. . .	43
2.7	Simulation of VQA for systems of three, four and five QDOs arranged in regular polygons using ansatze of varying depth.	46
3.1	Energy error as function of Krylov order and fitting parameters as function of lattice size in the case of zero measurement error	69
3.2	Estimated number Krylov order needed for QSE to achieve fractional energy error of 10^{-4} in the case of no measurement noise	69
3.3	Fractional energy error with PQSE procedure with finite number of measurements for range of different lattice sizes.	74
3.4	Energy error as function of calls to qubitisation procedure and fitting parameters as function of lattice size	75
3.5	Estimated number of call to qubitisation procedure using PQSE to achieve fractional energy error of 10^{-4}	75
3.6	Upper bound gate and qubit costs for block encoding procedure . .	79

3.7	Total number of T-gates required for the whole PQSE procedure for the single flavour Schwinger model	81
A.1	Ground state energy of one-dimensional QDO pair as a function of coupling constant for various values of d	93
A.2	Fully connected graph K_4^2 along with the three corresponding partitions of edges.	97
A.3	Number of circuits required to measure Hamiltonian as a function of N and d	98
B.1	Coupling map of the IBM Quantum devices.	113
D.1	Fractional energy error using TQSE procedure with finite number of measurements for range of different lattice sizes.	142
D.2	Estimated number of call to qubitisation procedure with TQSE to achieve fractional energy error of 10^{-4}	142

List of Tables

1.1	Estimated qubit and time requirements required for VQE algorithm from different works	10
2.1	Number of unique circuits required using our exact grouping as well as expected number of shots required to achieve absolute measurement error ϵ for a system of N d -level one-dimensional QDOs.	31
C.1	Number of multi-controlled gates for each Hamming weight b needed to generate state $(N-1) G^{\text{field}}(1)\rangle + (N-1) G^{\text{int}}(1)\rangle + \sum_{i=1}^N G^{\text{spin}}(n)\rangle$ for each term in the Schwinger model Hamiltonian.	130
C.2	Gate counts to partially swap gauge field and interaction terms for a gauge link site into the remaining $N - 2$ sites implementing the transformation $(N - 1) G^{\text{field}}(1)\rangle + (N - 1) G^{\text{int}}(1)\rangle + \sum_{i=1}^N G^{\text{spin}}(n)\rangle \longrightarrow G\rangle$ for Schwinger model.	130
E.1	Hardware information for range of quantum processors	147
E.2	Estimated block encoding run time for different quantum processors assuming that two-qubit gates are executed in series	147
E.3	Estimated block encoding run time for different quantum processors assuming two-qubit gates are maximally parallelised	148

List of Algorithms

3.1	Partitioned quantum subspace expansion	72
-----	--	----

List of Abbreviations

CPU	Central processing unit (a classical computer).
DMFT	Dynamical field theory, computation method for strongly correlated materials.
Full CI	Full configuration interaction, variational ansatz for electronic structure problem in which all Slater determinants are included.
LCU	Linear combination of unitaries, method for performing block encoding.
LFT	Lattice field theory.
LGT	Lattice gauge theory
UCC	Unitary coupled cluster, common ansatz used in the study of fermionic chemical systems.
UVCC	Unitary vibrational coupled cluster ansatz, ansatz analogous to UCC used for vibrational (bosonic) systems.
vdW	van der Waals (interaction, coefficient etc...)
VQA	Variational quantum algorithm, hybrid quantum-classical algorithm based on minimisation cost function using a parameterised quantum circuit.
VQE	Variational quantum eigensolver, example of VQA in which solution is groundstate of some (typically chemical) Hamiltonian.
PQSE	Partitioned quantum subspace expansion, method for mitigating affect of noise in within generalised eigenvalue problem of QSE.
QCD	Quantum chromodynamics.
QDO	Quantum Drude oscillator, a quantum harmonic oscillator used to represent the polarisation response of electronic distribution in atom or molecule to external field.
QED	Quantum electrodynamics.

QFT	Quantum field theory.
QPE	Quantum phase estimation (algorithm), quantum algorithm for finding ground states using a fault tolerant quantum computer.
QPU	Quantum processing unit (a quantum computer).
QSE	Quantum subspace expansion algorithm.
QSVT	Quantum singular value transform, method for implementing functions of block-encoded operators on quantum computer.
TQSE	Thresholded quantum subspace expansion, method for mitigating affect of noise in within generalised eigenvalue problem of QSE.

“Maintaining speed is a good thing.”

— Tim Foster MBE

1

Introduction

Since Richard Feynman’s 1981 proposal to use quantum devices to solve quantum systems that are hard to study using conventional (classical) computers [1], the goal of solving computational problems using quantum systems (namely quantum computers) has developed into a rapidly growing and potentially highly disruptive field. While considerable effort within academic and industrial groups goes into engineering such a quantum computer - reducing error rates and increasing qubit counts, the study of quantum algorithms, namely algorithms that can be used to solve problems on a quantum computer, is necessary to make use of quantum computers when they exist at sufficient scale and quality. In addition, understanding the feasibility of different algorithmic approaches and application areas, as well as the specific hardware requirements, can help inform the hardware design process to develop architectures that are tailored towards the most impactful applications.¹

There are a number of known quantum algorithms with proven asymptotic speed ups

¹Alongside these engineering and algorithmic improvements, methods for dealing with device errors, namely quantum error correction [2–4] and quantum error mitigation [5] are studied in order to exponentially suppress errors to give fault tolerant computation (in the case of error correction) or to reduce the effect of noise for systems of limited size to enable the extraction of data that is close enough to the true value to be of use (in the case of error mitigation).

over classical algorithms². These include the seminal proposals of Shor’s algorithm for prime factorisation [6] and Grover’s database search algorithm [7]. Taking the cracking of RSA 2048 as a benchmark for the utility of Shor’s algorithm, although there is an asymptotic exponential speed up compared to classical algorithms, state of the art resource estimates for reasonable decryption times (8 hours) show that Shor’s algorithm would need many millions of physical qubits to implement the level of fault tolerance required [8]. Similarly, when there is a “Grover-like” quadratic speedup for industrially relevant optimisation problems, the system sizes at which classical methods fail requires fault tolerant devices as well as time and physical resource requirements that make achieving quantum advantage with such approaches unlikely [9]. Despite rapid advances in experimental demonstrations of quantum computers [10, 11]³, the hardware requirements for useful applications of these algorithms are beyond those that we can reasonably expect to have within the next couple of decades. Therefore, there is considerable onus to reduce the cost of such algorithms as well as identify novel algorithmic approaches and application areas to make use of quantum computers sooner.

1.1 Quantum algorithms for quantum systems

Since Shor and Grover’s works, many new quantum algorithms have been proposed⁴ in application areas such as optimisation [14], differential equations [15–18], machine learning [19–23] and pure mathematics (including algebra, number theory and graph theory) [24–29]. However, in searching for such novel approaches and applications to achieve quantum utility in the near term, the most promising avenue is to return to Feynman’s original proposal to use quantum computers to study complex quantum systems, in particular with application to studying molecules and novel materials in

²These could be complexity theoretic separations between the limits of classical and quantum computations, as is the case for Grover’s algorithm, or proven speedups over the best known classical algorithms as is seen for Shor’s algorithm.

³Although such results often claim to outperform best known classical algorithms, novel approaches based on tensor networks quickly develop to refute such claims [12, 13].

⁴Stephen Jordan maintains a comprehensive list of quantum algorithms at www.quantumalgorithmzoo.org.

quantum chemistry and materials science [30–32]. While improving computational approaches for such problems plays a crucial role in developing the fundamental science as well as generating economic and societal impact through the development of new technologies, they also exhibit a number of properties that make them appealing targets for the earliest use of quantum computation to solve problems beyond those of purely academic or complexity theoretic interest. Some of these properties are as follows.

1. The exponentially large Hilbert space of highly entangled quantum systems means that classical computers cannot efficiently represent such states⁵. In contrast quantum computers, which are themselves quantum systems, have access to exponentially growing Hilbert spaces. This is particularly apparent when studying electronic or spin systems, for which the state space can be mapped directly onto that of qubits within a quantum computer.
2. When studying real-world chemicals and materials, in which there is always some level of production or measurement uncertainty, achieving finite precision when computing ground state or dynamical properties is often satisfactory for comparison to experimental quantities⁶. Therefore, quantum algorithms that compute physical quantities up to some finite error will still be of value, particularly if the size of the error is known and controllable, and opens the possibility of using approximate or heuristic quantum algorithms as well as possibly making only partial use of or completely forgoing error mitigation or correction schemes. This is in contrast to many other quantum algorithms. For example, there is little value in Shor’s algorithm if it outputs a result like “5.5 and 2.9 are approximate prime factors of 15”.
3. Algorithm primitives involved in studying physical systems often have natural and relatively low cost ways of being performed in quantum hardware. For

⁵In particular, while tensor network methods have seen considerable success in solving highly correlated systems in 1D, there are still considerable challenges to using them to solve systems in two or more dimensions demonstrating volume law entanglement.

⁶In the case of quantum chemistry the level of accuracy required to make realistic chemical prediction is typically given taken as 1 kcal/mol and given the name ‘chemical accuracy’.

example, given access to unitary evolution of operators making up the Hamiltonian, time evolution implemented through Trotterisation can be performed (relatively) easily [33]. Primitives based on numerical operations such as addition and multiplication by comparison, which are very cheap to do on classical processors, require significant resources when performed on quantum processors for algorithms based on integer or floating point arithmetic [34].

Quantum algorithms for complex quantum systems usually perform one of two tasks. The first is to calculate how dynamical properties, such as particle density or magnetisation of a material, change as a system is time evolved under a given Hamiltonian. The second is to calculate groundstate properties of a Hamiltonian, which will be the focus of this thesis. Accurate computation of groundstate energies is needed to determine stability of molecules and crystals, calculate reaction pathways and energies for chemical and biological processes and calculate the potential energy surfaces and time-evolution of molecules with the Born-Oppenheimer approximation. Such computations play essential roles in understanding and developing new semi-conducting, superconducting materials and energy storage solutions, as well as developing and determining the efficacy and safety of new drugs. Other groundstate properties, often related to entanglement or macroscopic ordering which have no classical analogue, are of significant importance for understanding novel phases of matter in high-energy and condensed matter physics.

1.1.1 Quantum phase estimation

Originally proposed by Kitaev [35] the quantum phase estimation (QPE) algorithm [36] has become one of the most studied quantum algorithms to calculate groundstates. Assuming we have efficient access to an input state that has inverse polynomial overlap with the true ground state⁷, the QPE can be used to calculate the groundstate energy with resources that scale polynomially in the desired precision and system size [37]. It is hoped that this may offer asymptotic speedups over

⁷This appears to be a very strong assumption

classical algorithms allowing for the calculation of energies for larger systems or with fewer approximations - and therefore more accurate outputs, than possible with classical processors. Due to the need for high accuracy controlled time-evolution under the problem Hamiltonian, the number of gates and qubits required means that QPE will need to be implemented within a fault tolerant architecture.

Since it's original proposal and early works studying QPE using Trotterisation, there has been much work to find more resource efficient methods for performing QPE. These include algorithmic improvements based on qubitisation [38–40], as well as more recent investigations using a first, rather than a second, quantised approach [41–43]. However, despite cost reductions by many order of magnitude, state-of-the-art QPE implementation for electronic Hamiltonians with applications to industrial catalysts or lithium ion battery technologies still require upwards of 10^{10} T-gates⁸ and many millions of physical qubits [41, 45]. Resource estimates for biologically relevant compounds are similarly discouraging [46].

1.1.2 Variational quantum eigensolver

Almost two decades after Kitaev's proposal of the phase estimation algorithm, Peruzzo et al. [47] developed a new approach for groundstate computation for electronic systems called the variational quantum eigensolver (VQE). In contrast to QPE, VQE is a heuristic algorithm specifically designed for use on small, non-error corrected quantum processors.

The idea of VQE algorithm is to use a short *ansatz* circuit $\hat{U}(\vec{\theta})$, which contains quantum gates parameterised by real values $\vec{\theta}$, acting on an input reference state $|\psi_0\rangle$ to produce parameterised states $|\psi(\vec{\theta})\rangle := \hat{U}(\vec{\theta})|\psi_0\rangle$. The algorithm computes an approximate value of the groundstate energy E_0 of a given Hamiltonian $\hat{H}$ by

⁸The primary metric for evaluating the cost of fault tolerant algorithms is usually considered to be number of T-gates or the closely related Toffoli gate. These non-Clifford gates cannot be implemented transversally in the surface code and so dominate the run time of the algorithm due to required magic-state creation and distillation methods used to implement them. See [44] for details on such methods.

finding parameters $\vec{\theta}_\star$ that solve the optimisation problem

$$E_{\vec{\theta}_\star} = \min_{\vec{\theta}} \langle \psi(\vec{\theta}) | \hat{H} | \psi(\vec{\theta}) \rangle. \quad (1.1)$$

This problem is solved by using a quantum computer to generate trial states $|\psi(\vec{\theta})\rangle$ and measure the expectation value of $\hat{H}$. After each evaluation of the cost function $E_{\vec{\theta}} := \langle \psi(\vec{\theta}) | \hat{H} | \psi(\vec{\theta}) \rangle$, a classical co-processor uses this to calculate a new set of trial parameters to reduce the cost function. This is repeated as a feedback loop between the quantum and classical processor until the optimisation procedure, which can use gradient or non-gradient based methods, converges within some precision.

By the variational principle, $E_{\vec{\theta}} \geq E_0$ for all $\vec{\theta}$ (including $\vec{\theta}_\star$). However, if the ansatz is chosen such that it can generate states close to the true groundstate, and the optimisation process converges to the global minimum of the cost function landscape, then the value of $E_{\vec{\theta}_\star}$ calculated will be approximately equal to the desired E_0 .

For VQE to provide some quantum advantage, it is necessary for the ansatz to be able to generate states close to the true groundstate of the problem and be hard to simulate using classical computers (otherwise we would not need to use a quantum computer). In the case of the original VQE proposal, this was achieved by using the unitary coupled cluster (UCC) ansatz [48] for the electronic part of a molecular Hamiltonian. A UCC ansatz, consisting of single and double excitation operators, for N electrons is given by

$$\hat{U}_{\text{UCC-SD}}(\vec{\theta}) := e^{\hat{T} - \hat{T}^\dagger}, \quad (1.2)$$

where $\hat{T} = \hat{T}_1 + \hat{T}_2$ is the cluster operator containing single and double excitations $\hat{T}_1 = \sum_{p,r} \theta_r^p \hat{a}_p^\dagger \hat{a}_r$ and $\hat{T}_2 = \sum_{p,q,r,s} \theta_{r,s}^{p,q} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s$, where $\hat{a}_p^\dagger$ and $\hat{a}_q$ are creation and annihilation operators for orbital p . This ansatz requires a polynomial number of variational parameters and is known to be hard to simulate classically. Compared to coupled cluster ansätze, given by $e^{\hat{T} + \hat{T}^\dagger}$, which are some of the most prevalent and accurate methods in classical algorithms for quantum chemistry, UCC obeys

the variational principle, is hard to simulate classically and can better deal with strong correlation [49]. Peruzzo et al. used a Hartree-Fock state as $|\psi_0\rangle$ and implemented the Hamiltonian and ansatz by mapping Fermions onto qubits using a Jordan-Wigner transformation. The UCC circuit operators were then implemented using Pauli rotations and Trotterisation.

The use of shallow parameterised quantum circuits in the VQE algorithm and optimisation within a hybrid quantum-classical feedback loop has since been used to develop various other variational quantum algorithms (VQAs). VQAs are used to solve optimisation problems where, rather than minimising a physical energy, more general cost functions can be mapped onto operators (taking the place of the molecular Hamiltonian in VQE) and solutions encoded in appropriately structured ansatz circuits. They have been applied to a range of different problems including combinatorial optimisation [14, 50], strongly correlated materials [51–55], non-linear partial differential equations [18, 56, 57] and machine learning [58, 59]. A longer discussion about specific applications of VQAs within quantum chemistry is presented in Chapter 2.

There have been a range of further developments to improve ansatz structure of VQAs, these aim to improve solution quality, increase noise robustness, target specific hardware architectures or to make optimisation landscapes more favourable. In general however, there is often a trade off made between ansätze that exploit structure or symmetries in the problem restricting them to ‘reasonable’ trial solutions [10, 60, 61]; leading to higher quality solutions and easier trainability, and ansätze that are designed to access as much of Hilbert space as possible while minimising the gate cost [62, 63]; making them more feasible for use in real, non-error corrected devices. The difficulty in choosing between these two approaches may be somewhat mitigated by using adaptive methods that construct the ansatz systematically as part of the optimisation routine [64–66].

Further algorithmic improvements include sophisticated methods for calculating gradients of $|\psi(\vec{\theta})\rangle$ with respect to each variational parameter [67, 68] as well as better

ways of measuring the expectation values. In most variational algorithms, $|\psi(\vec{\theta})\rangle$ are measured by decomposing the Hamiltonian into its constituent Pauli operators, measuring each operator using an appropriate basis rotation and summing results with the corresponding prefactors. This requires a number of measurements that is inversely proportional to the desired measurement precision squared and increases with the number of terms within the Hamiltonian. The number of measurements required can be reduced by grouping operators that commute⁹ then measuring them in their common eigenbases [71, 72] and truncating low weight Hamiltonian terms [73] as well as using inference methods based on partial state reconstruction and tomography [74, 75].

An appealing feature of VQAs is that they include some inherent noise robustness. Coherent hardware errors such as systematic over- or under-rotations in rotation gates can be counteracted by the optimiser learning to introduce the opposite bias within $\vec{\theta}_*$. Detrimental effects of incoherent statistical noise on measured values for the cost-function or gradients may also have reduced impact on solution quality as imperfect values. These become manifest as steps in the ‘wrong’ direction during the optimisation procedure, which can be counteracted by following optimisation steps, particularly in the case of gradient based methods which still perform well when given noisy gradient values¹⁰. A robust quantitative understanding of the noise resilience of VQAs remains an open problem however.

Since their initial proposal, a number of drawbacks have been discovered that may limit the utility and scalability of variational quantum algorithms. The largest of these is the occurrence of ‘barren plateaus’ in which the gradients within the cost function landscape vanish exponentially with system size and volumes of the parameter space with non-vanishing gradients become exponentially concentrated around the global minimum [76]. Performing successful optimisation in the presence of a barren plateau is expected to require exponentially many optimisation steps as

⁹The problem of calculating optimal groupings can be mapped onto a minimum clique colouring problem [69], which is in general NP hard [70] and best solved using heuristics.

¹⁰This is the principle behind stochastic gradient descent which is widely used in the field of (classical) machine learning.

well as exponential measurement precision in order to resolve vanishing gradients. Barren plateaus are induced by a number of causes including non-local cost functions [77, 78], incoherent measurement noise [79] and ansätze that exhibit volume law entanglement (which we typically require for hard classical simulation) [80]. All these sources appear to be related to the exponentially large algebra generated by parameterised circuits [81–83]. Finding strategies to avoid these is an active area of interest, with proposals based on developing ansatz structures [84–87], parameter initialisation techniques [88–90], and adaptive ansätze [91] amongst others [92–95]. Other works aim to improve optimisation algorithms [96, 97] and modify the optimisation problems [98] to improve the trainability of unfavourable cost function landscapes (with many local minima for example).

Without considering the number of optimisation steps needed to achieve convergence, a number of works have aimed to estimate the resources required to perform measurement of expectation values for just a single optimisation step for various molecules of industrial importance. Some of these estimates are summarised in Table 1.1. Although the types of molecules investigated, assumed gate times and algorithmic methods vary between the different estimates, each suggests that current state of the art VQE methods are too costly for useful application in molecular simulation, even ignoring the prefactor arising from the number of optimisation steps (which is difficult to estimate but expected to be very large).

1.1.3 Quantum subspace expansion

Originally proposed as an error mitigation technique [103], a quantum algorithm called quantum subspace expansion (QSE) has emerged as a method for groundstate computation on noisy intermediate scale quantum processors. Like the VQE algorithm, QSE aims to use a relatively shallow quantum circuit to efficiently estimate expectation values which are then used by an inexpensive classical computation. However, it does not rely on the hybrid quantum-classical feedback loop for optimisation with the hope of avoiding the scalability issues related to barren plateaus.

Ref.	Molecule	Qubits	Run time	Notes
[99]	Fe ₂ S ₂	112	10 ¹³ samples	Single energy evaluation
[100]	C ₁₀ H ₈	800 – 1500	10 ⁸ two-qubit gates	Single energy estimation
[101]	C ₂ H ₆ O	260	71 days	Single energy estimation
[102]	Cr ₂	52	24 years	Gradients for single step
[102]	Cr ₂	2.9 × 10 ¹¹	3 ms	As above, parallelised

Table 1.1: Estimated qubit and time requirements required for VQE algorithm from different works. With the exception of [102], which uses a hardware efficient ansatz, all works use some version of the UCC-SD ansatz. Each assumes different algorithmic options, notably measurement method, as well as gate times (for those given as times). Notes indicate the expectation values measured. For single energy estimates, this is the number required to achieve chemical accuracy or some similar ‘practically useful’ statistical error. For those labelled ‘gradients’, this includes expectations required to calculate gradients with respect to all parameters, and in the parallelised case performs different measurements using additional qubits and multiple copies of the ansatz. For precise details, the original references should be consulted.

For groundstate calculation, the QSE algorithm operates an approximate ground-state within a subspace spanned by a D dimensional non-orthogonal basis

$$\{|\psi_0\rangle, |\psi_1\rangle, \dots, |\psi_{D-1}\rangle\} \quad (1.3)$$

which can be generated by a quantum processor. Within this subspace a trial state has the form $|\psi(\vec{c})\rangle = \sum_{i=0}^{D-1} c_i |\psi_i\rangle$, minimising the functional $\langle\psi(\vec{c})|\hat{H}|\psi(\vec{c})\rangle$ to find optimal coefficients $\vec{c}$ is equivalent to solving the generalised eigenvalue problem

$$\mathbf{H}\vec{c} = E\mathbf{S}\vec{c}, \quad (1.4)$$

for minimum eigenvalue E , where matrices $\mathbf{H}$ and $\mathbf{S}$ are given by

$$\mathbf{H}_{i,j} = \langle\psi_i|\hat{H}|\psi_j\rangle, \quad \mathbf{S}_{i,j} = \langle\psi_i|\psi_j\rangle. \quad (1.5)$$

The power of the QSE algorithm relies on making a good choice of basis and being able to efficiently measure expectation values for $\mathbf{H}_{i,j}$ and $\mathbf{S}_{i,j}$. The basis $\{|\psi_k\rangle | k = 0, \dots, (D-1)\}$ is typically generated by applying (often repeatedly) operators to an input state $|\psi_0\rangle$ which is chosen to be close to the true groundstate. Many types of operators have been used to generate this basis including real and imaginary time evolution [104–109], Fermionic excitation operators [110], Pauli

operators [111] and using parameterised quantum circuits in variational manner [112]. Depending on the basis used, the expectation values can be calculated in different ways such as a SWAP test [113, 114] or by measuring constituent Pauli expectation values and recombining as done in VQE.

A particularly powerful basis, called a Krylov basis, is formed by repeated application of the Hamiltonian to the reference state: $\{|\psi_k\rangle = \hat{H}^k |\psi_0\rangle |k = 0, 1, \dots, D-1\}$. This forms the basis for the well studied classical Lanczos algorithm for groundstates [115, 116]. In the infinite D limit, the Krylov basis includes infinite imaginary time evolution and, assuming non-zero overlap between the initial and ground states and infinite precision arithmetic will lead to the exact groundstate. In the case of finite sized bases for a given problem instance, the Lanczos algorithm converges exponentially with basis size, and so very good solutions can often be found using reasonable basis sizes [117, 118]. For quantum systems in a Hilbert space, the cost of exactly evaluating and storing the Hamiltonian acting on the reference states grows exponentially with system size, making the use of a quantum processor to generate the basis within a QSE procedure very appealing. Moreover, when using a Krylov basis, the matrix elements required simplify to $\mathcal{O}(D)$ unique values given by $\mathbf{H}_{i,j} = \langle\psi_0|\hat{H}^{i+j+1}|\psi_0\rangle$, $\mathbf{S}_{i,j} = \langle\psi_0|\hat{H}^{i+j}|\psi_0\rangle$, rather than the $\mathcal{O}(D^2)$ unique values required when using operators to generate the basis that do not commute with the Hamiltonian. Of course the Hamiltonian is not a unitary operator in general, and so methods to compute expectation values must use some additional classical computation generate the Krylov basis. Strategies to compute the non-unitary expectation values in $\mathbf{H}, \mathbf{S}$ include classical post processing to combine expectation values from time evolutions [119] or methods based on qubitisation [120] which we will discuss further in Chapter 3.

As is the case for measurements within VQE by performing measurements as Bernoulli trials, the precision of measured expectation values is inversely proportional to the square root of the number of measurements. The stability of the generalised eigenvalue problem of (1.4) relies on these matrices being well conditioned, however

when using a Krylov basis, the condition number of the matrices $\mathbf{H}$ and $\mathbf{S}$ grows exponentially with D [121]. Unless this ill-conditioning is controlled, finite numerical precision as well as measurement noise can lead to failure of the QSE procedure. Recently, theoretical and numerical analysis has begun to elucidate the effect that such noise has on the generalised eigenvalue problem. Strategies to regularise the problem by discarding eigenvalues of the matrices less than a chosen threshold can be used to increase the noise robustness [120, 122]. Although such methods give upper bounds on the limiting behaviour of the performance of such methods, such bounds are typically very loose and require a number of assumptions that mean they often fail to accurately predict the performance of realistic algorithmic implementations [123]. Additional methods have also been developed to try to improve the conditioning including using a modified bases formed from Gaussian functions of the Hamiltonian [124] as well iterative [85] and partitioned [125] methods in which the effective Krylov basis is constructed systematically from a series of smaller bases.

So far, the QSE algorithm has been applied to ground [105, 126] and excited states [127] of electronic Hamiltonians in quantum chemistry as well as ground states of spin models [120, 128].

1.2 Thesis structure

The aim of this thesis is to apply and develop quantum algorithms to compute the ground state of Hamiltonians of quantum systems. We use two different algorithmic approaches, namely VQE and QSE, each to systems to which they have not been applied before. We study firstly a system of interacting oscillators used to model coarse grained intermolecular interactions and then a lattice gauge field theory (the Schwinger model) representing quantum electrodynamics in one spatial dimension. The former sees application in modelling dispersive interactions in systems such as noble gas liquids [129] and water [130], while the latter describes fundamental particle interactions analogous to full QCD and QED [131]. Although these application

areas differ significantly from one another; the theoretical tools used to understand the resource cost for each, in particular in mapping Bosonic degrees of freedom onto qubits, are similar in both. Through a combination of theoretical analysis, numerical simulation and real device experiments we begin to understand how scalable each approach is and how applicable they are to near term devices. The structure of this thesis is as follows.

In Chapter 2 we develop a variational quantum algorithm for studying intermolecular interactions within a coarse grained molecular model and in doing so present the first application of quantum computing to studying intermolecular dispersion interactions. We describe the coarse grained model and how groundstate and binding energies can be computed using the VQA. We compute bounds on the number of shots required to measure expectation values of the energy within the algorithm and find that this scales more favourably with system size than other VQE methods for quantum chemistry. We perform proof of principle experiments on an IBM quantum computer demonstrating that, when combined with error mitigation methods, our variational algorithm is able to calculate the interaction energy between two Iodine dimers using currently available quantum hardware. We further show how the model can be extended to include anharmonic and non-linear interactions to capture a richer set of physics with minimum additional cost for the quantum algorithm. We then perform numerical simulations demonstrating the algorithm within such a context, as well as for larger system sizes.

In Chapter 3 we apply quantum subspace expansion algorithm to the problem of computing the ground state of the lattice Schwinger model - a prototypical lattice gauge theory used to describe quantum electrodynamics in one spatial dimension. We use qubitisation and quantum signal processing to generate the expectation values required for the algorithm. We develop a novel method for performing qubitisation of the lattice gauge theory and perform a full resource analysis. Compared to previous methods used within the context of QSE, our method reduces the cost by multiple orders of magnitude for system sizes of interest. We perform numerical

simulation of the algorithm, including the effect of measurement noise, to estimate the number of basis sizes required and, combining with the theoretical calculation of gate costs, estimate the entire gate cost of the algorithm. This work presents the first application of QSE methods to finding the ground states of a lattice gauge theory and, excluding previous works performing tensor networks, performs numerical simulation of a quantum algorithm solving the Schwinger model for larger system sizes than other numerical or experimental works. Furthermore, compared to studies using VQE for LGTs in which it is difficult to predict how the optimisation procedure scales with system size, we perform a more comprehensive estimate and extrapolation of the total resource cost for the quantum algorithm.

Finally, in Chapter 4 we discuss our results and present further research directions.

1.3 Publications

- Chapter 2 is based on work presented in Lewis W. Anderson, Martin Kiffner, Panagiotis Kl. Barkoutsos, Ivano Tavernelli, Jason Crain, and Dieter Jaksch. “Coarse-grained intermolecular interactions on quantum processors.” *Physical Review A*, **105**, 062409, 2022.
- Chapter 3 is based on work presented in Lewis W. Anderson, Martin Kiffner, Thomas O’Leary, Jason Crain, and Dieter Jaksch. “Solving lattice gauge theories using the quantum Krylov algorithm and qubitization.” *arXiv:2403.08859*.
- While undertaking my DPhil, I worked on the following publications that are not included within this thesis.
Ben Jaderberg*, Lewis W. Anderson*, Weidi Xie, Samuel Albanie, Martin Kiffner, and Dieter Jaksch. “Quantum self-supervised learning.” *Quantum Science and Technology*, **7**(3), 035005, 2022. *joint first authors.
Thomas O’Leary, Lewis W. Anderson, Dieter Jaksch, and Martin Kiffner. “Partitioned Quantum Subspace Expansion.” *arXiv:2403.08868*.

2

Variational quantum algorithm for coarse-grained quantum chemistry

Contents

2.1	Introduction	16
2.1.1	Computational chemistry on classical and quantum computers	16
2.1.2	Our contribution	19
2.1.3	Chapter layout	22
2.2	Coarse-grained modelling on digital quantum computers	22
2.2.1	Coarse-grained model	22
2.2.2	Algorithm description	26
2.2.3	Measurement cost	29
2.2.4	Variational ansatz	33
2.3	Experimental results using IBM quantum computers .	34
2.4	Extending the model	38
2.4.1	Three-dimensional QDOs	38
2.4.2	Induction, anharmonic and non-linear terms	40
2.5	Further experimental results	42
2.5.1	Anharmonic systems	42
2.5.2	Many oscillators	45
2.6	Discussion and conclusions	46

2.1 Introduction

2.1.1 Computational chemistry on classical and quantum computers

The computer simulation of atomic and molecular matter plays a vital role in the discovery of new functional materials. For systems more complex than atomic hydrogen, numerical computation relies on various approximations of the system and methods must strike a balance between the physical realism of an underlying model and the practicality of its implementation.

This trade-off varies depending on the system being studied and is made considerably more difficult by the deep hierarchy of molecular forces that determine the structure and behaviour of condensed phases of matter. These interactions span energy scales that vary from greater than 1 eV for the *intra*-molecular regime down to $10^{-1} - 10^{-2}$ eV for the *inter*-molecular regime [132]. Within the inter-molecular regime, the weakest of these interactions are dispersion interactions (van der Waals forces) that arise from quantum fluctuations in the electron distribution within atoms and molecules. Such interactions give rise to correlated induced multipole moments which, at leading-order, present as dipolar coupling with energy contributions following a $1/R^6$ decay with distance. This is known as the London interaction [133]. Although dispersion interactions are typically weaker than electrostatic interactions or Hydrogen bonds, their ubiquity and long-ranged nature means that they influence all non-covalent associations.

The effects of these dispersion interactions extend beyond cohesive energy and play important roles in determining structural and mechanical properties [134, 135] as well as phase transitions and barrier heights [136]. Accurate calculation of such effects is therefore vital in understanding a wide range of condensed-phase phenomena including low-dimensional and layered system stacking, molecular crystal stability and liquid state physics as well as organic–inorganic interface properties and biophysical interactions such as drug-ligand binding. Since these interactions

scale with system size, they are expected to play even larger roles in systems of large molecules and supramolecular assemblies [137].

To capture these molecular interactions through computational models, two possible strategies can be used which vary significantly in their computational requirements. Firstly, empirical potential methods (also known as force field methods) in which the system energy is decomposed into bonded terms (stretches, bends and torsions) and non-bonded (electrostatic and dispersive) terms. Parameterising the strength of each of these interactions using experimental reference quantities, Newtonian mechanics can be used to calculate the motion of the system. Static properties of the system can be calculated by averaging over many of these trajectories. By limiting oneself to a classical description of the motion, empirical potential methods are extremely computationally efficient allowing them to be applied to relatively large systems. The necessary trade-off for such a description however is that they generally lack an explicit description of the electronic response and are therefore limited in their transferability to environmental or thermodynamic regimes outside the range used for the experimental parameterisation.

Secondly there are “first-principles” approaches where the Schrödinger equation with various levels of approximation is solved directly. State of the art approaches for this are based on Kohn-Sham density functional theory. Unlike empirical potential methods, there is no use of empirical force laws, rather forces are computed directly from the electronic structure according to the Hellmann-Feynman [138, 139] theorem and evolving along the Born-Oppenheimer surface [140]. In such a model, because of the mass difference between nuclei and molecular electrons, the nuclear coordinates are considered fixed and the electronic Hamiltonian within the corresponding nuclear potentials is solved. The mean force felt by the nuclei is calculated from the energy of these electronic configurations and is used to calculate the dynamics of the nuclei which follow the changing energy landscape as the electrons react to the modified nuclear coordinates. These methods are far more resource intensive than empirical force field methods and are thus limited

to smaller system sizes. Additionally, common density functional approximations struggle to capture non-local correlations leading to dispersive interactions (although work is being done to correct for this [141–144]). High-level quantum chemistry methods (such as configuration interaction (CI) [145–147] or canonical coupled cluster (CC) [148, 149] with perturbative - single, double, triple - excitations) are able to capture dispersive interactions to varying levels of approximation, albeit with particularly poor system size scaling [150–152].

The high computational cost of such ab-initio approaches has led to a significant amount of interest in using quantum computing to study strongly correlated electron systems at the first principles level. First developed in the 1990s, the quantum phase estimation (QPE) [36] algorithm may offer guaranteed speedups for solving electronic Hamiltonians. However, even with many years of theoretical improvement, current state of the art QPE approaches still require many hundreds or thousands of *logical* (error-corrected) qubits to solve chemical Hamiltonians of industrial interest [40, 153].

In the near future, in which we are limited to small, non-error corrected quantum computers, one possible route towards the practical application of such devices is to use variational quantum algorithms (VQAs) [154–156] to solve the Schrödinger equation for chemical systems [31]. Due to the small size of current quantum processors and limitations in classical methods for QPU simulation, current studies of molecular properties using VQAs have been limited.

As a result, chemical applications of VQAs have been limited to small systems within the strongly-bound (covalently bonded) intra-molecular regime. Examples of such results include studies of dissociation curves of small molecules [62, 64, 157], ground state energies for intramolecular bonds [158], molecular excitation energies [159], covalent binding energies of hydrogen chains, isomerisation barriers [160], nuclear effects (using post-Born Oppenheimer methods) in hydrogen and deuterium molecules [161], molecular vibrations [60, 162] and, two-site DMFT calculations [163, 164]. Excluding the latter two examples, these all use full electronic basis sets such

as Slater-type or Gaussian-type orbitals [165]. Within the rest of this work, we will refer to these methods as *orbital-based* VQE methods.

The quantum algorithms to solve the inter-molecular regime on the other hand have seen very little investigation. As mentioned previously however, such interactions play a vital role in many aspects of condensed phase and biological systems and must be investigated in order to fully understand the material properties of such systems using quantum technologies. As with classical algorithms, directly scaling up existing orbital-based VQE approaches will prove to be prohibitive due to the necessary inclusion of a large number of core electrons to describe dispersive forces in large molecules [166, 167]. It is necessary to find strategies for reducing the number of electrons and orbitals explicitly treated if we hope to find practical applications of quantum algorithms.

2.1.2 Our contribution

As a step towards the practical application of quantum algorithms to studying large non-covalently bonded systems, we develop a coarse-grained model of the electronic interaction that is well suited for VQAs. Our model has the further advantage of being easily expanded to a wide range of possible interactions not yet studied for intermolecular systems. Our approach extends the maximally coarse-grained model that has recently been developed for classical computations [168]. In this coarse grained approach the responsive part of the electronic distribution of a molecule (or part of a molecule) is modelled by a quantum harmonic oscillator. By constructing a system based on *quantum* mechanical oscillators, zero-point multi-polar fluctuations arise naturally in the model. The properties of the oscillator are tuned to match reference values for polarisability or dispersion coefficients for real molecular species which are found independently through experiments or ab-initio computations. As a result, although model parameters are empirically fitted to real molecular parameters, interaction potentials between atoms or molecules are not set a-priori and are instead determined through the multipolar fluctuations in the electronic distribution.

One can therefore think of this method as lying somewhere between empirical potential and first principle methods. Under certain model limitations¹, the coarse-grained model has already been solved using a path integral approach on classical processors. This has been used to study noble gas liquids and solids (that interact entirely through dispersion) as well as water (where both polarisation and dispersion are present) across its phase diagram where it has correctly predicted vaporisation enthalpy, temperature of maximum density, temperature dependence of the dielectric constant and surface tension [130, 168, 169] and revealed molecular signatures of Widom line crossing in the supercritical fluid [170]. These results have been restricted to only include harmonic trapping confining potentials. This means that the model is governed by Gaussian statistics in which interactions are related by simple algebraic relationships and some higher order interactions are missing completely [171]. Such relations only hold approximately for real systems and extension beyond Gaussian statistics is needed for a complete description of the chemistry.

Compared to classical approaches, by using quantum processors to capture a coarse-grained representations our method provides a more general framework in which we can extend the model to anharmonic oscillators and non-linear interactions. This in theory allows us to capture these missing interactions as well as introduce more complex non-linear physics.

In the following we summarise the main features of our algorithm. Firstly, the space requirement of our algorithm scales efficiently with system size whereas a classical representation of the Hilbert space would have exponentially growing memory requirements. Explicitly, the number of qubits required for our algorithm grows linearly with the number of molecules. Secondly, the number of unique circuits needed to evaluate the cost-function within our VQA scales as $\mathcal{O}(N)$ for a N oscillator Hamiltonian with all-to-all dipolar interactions. By comparison, this is far better than the $\mathcal{O}(N^3)$, where N is the number of orbitals, needed for full orbital-based VQE [172]. Furthermore, relaxation of the harmonic approximation

¹Namely, a restriction to harmonic trapping potentials with linear coupling. See Sec. 2.4 amongst other parts of this chapter for further discussion on this.

can be done with practically zero cost with our VQA. This is not the case for classical methods such as path integral and Monte Carlo techniques which typically rely on the efficient sampling [173] of two point correlations of many body Gaussian states. Ground states for anharmonic Hamiltonians are non-Gaussian and are therefore much more costly to sample from. This has meant that many body systems with anharmonic potentials have been largely unexplored using classical computational approaches.

As an experimental demonstration of our algorithms, we use IBM superconducting quantum processors to solve our model. We consider a system of two linear non-polar molecules in which dispersive contributions arise from dipolar quantum fluctuations (the London dispersion interactions). By using model parameters corresponding to an iodine (I_2) molecular dimer, we use this to show the first experimental demonstration of quantum computers being used to accurately model dispersive inter-molecular interactions.

We present proof of principle experiments to larger systems of three, four and five oscillators using simulated quantum devices on a classical computer. We investigate how ansatz size affects the performance of our algorithm for these larger systems and find that we are able to accurately calculate ground states with relatively low depth circuit even for these larger systems. We further extend our experimental results by using an IBM superconducting quantum processor to solve a system featuring strong anharmonic on-site potentials demonstrating how our algorithm can be used to find non-Gaussian ground states. We find that, although the algorithm is capable of finding good ground states for such a system, noise levels on the device must be improved in order to accurately measure the anharmonic contributions to the dispersion energy.

The coarse-grained quantum algorithm presented here allows us to model inter-molecular interactions between large molecules. Understanding such effects is an essential part of characterising a wide range of chemical and biological systems. This extends existing classical methods for such problems and opens up opportunities

for using quantum computers to study systems at time and length scale that were previously inaccessible. Comparing the scalings of the measurement cost and qubit requirements to full-orbital based VQE calculations, our algorithm shows evidence of being more readily scalable and may be a stronger candidate for demonstrating useful quantum advantage.

2.1.3 Chapter layout

For the remainder of the chapter the structure will be as follows. In Sec. 2.2 we outline the coarse-grained model and describe how it can be solved using a variational quantum eigensolver algorithm. This includes Sec. 2.2.1 defining the coarse-grained Hamiltonian; Sec. 2.2.2 describing how this is mapped onto a quantum circuit; Sec. 2.2.3 in which we give a scheme to group commuting operators in order to efficiently measure the Hamiltonian as well as giving bounds on the number of experimental shots required to achieve a given statistical error; and Sec. 2.2.4 which describes the parameterised ansatz circuit used in our experiments. In Sec. 2.3 we show proof of principle experiments using our model to accurately model the London dispersion interaction of an I_2 dimer on an IBM superconducting quantum processor. In Sec. 2.4 we show how the one-dimensional model considered thus far can be extended to three-spatial dimensions along with the inclusion of non-linear interactions and anharmonic binding potentials. In Sec. 2.5 we give further experimental results on real and simulated quantum processors. This includes Sec. 2.5.1 considering anharmonic oscillator systems; as well as Sec. 2.5.2 considering systems of more oscillators. Finally, in Sec. 2.6 we discuss our results and describe possible future research directions.

2.2 Coarse-grained modelling on digital quantum computers

2.2.1 Coarse-grained model

Within the harmonic coarse-grained model, the electronic distribution of an atom or molecular moiety is modelled by a charged quantum Drude oscillator (QDO). For

neutral species, this is a quantum harmonic oscillator consisting of a (quasi-)particle of charge $-q$ bound to charge $+q$ which is fixed in the atomic or molecular frame. Allowing these charges to interact through a Coulomb interaction (truncated to linear order) leads to correlated displacement of these QDOs which correspond to induced instantaneous charge density fluctuations in the molecules. This allows us to represent dispersive interactions between non-polar molecules. Although such a model is very simplistic, it nonetheless exhibits a full hierarchy of long range electronic responses.

Within this model, many body polarisation and dispersion interactions arise naturally at all orders and the resulting intermolecular forces emerge without the need for a-priori force laws or interaction potentials.

One feature (or drawback) of this model is that it obeys distinguishable particle statistics rather than Fermi statistics. In order to capture exchange interactions, the model must be augmented with short range repulsive terms. Such an extension is beyond the scope of what we present here and is often unnecessary when using the model in condensed phases systems dominated by van der Waals interactions and where exchange interactions have minimal contribution.

This basic model has been applied with great success on classical processors for noble gas solids and single-component liquids [174, 175]. By defining the model parameters using just properties of single, isolated molecules, these models have predicted emergent phenomena across large portions of their respective phase diagrams. For a recent review of such results see Ref. [168].

To present the coarse-grained model, the algorithm and the first set of experimental results within this section (Sec. 2.2) as well as Sec. 2.3, we will consider only harmonic one dimensional QDOs with dipole order coupling². Even within the simple model one-dimensional model, a rich set of physics emerges. With just two of these oscillators we generate the London dispersion interaction which is proportional to R^{-6} and α^2 as expected, where α is the dipole polarisability and R

²We will later describe how to extend this to more spatial dimensions and complex Hamiltonians.

the physical separation, with proportionality constant matching that for the linear non-polar molecules in the limit of infinitely anisotropic polarisability. We will briefly demonstrate this in the following.

We define a pair of identical QDOs with charge q , effective mass μ and harmonic trapping frequency ω ; these being three model parameters that can be fitted to experimental data to achieve the required dispersive interaction. We separate these QDOs by distance R and allow the negative charges to move perpendicular to the inter-oscillator axis and parallel to each other. The position and momentum operators (which can be defined in terms of ladder operators obeying the usual bosonic commutation relations) corresponding to the motion of the quasi-particles are given by $\hat{x}_1, \hat{p}_1$ and $\hat{x}_2, \hat{p}_2$ for the first and second QDO respectively. We arrive at the following Hamiltonian by considering kinetic, onsite potential and (leading order) Coulomb interactions between charges within each operator;

$$\hat{H} = \frac{\hat{p}_1^2}{2\mu} + \frac{\hat{p}_2^2}{2\mu} + \frac{1}{2}\mu\omega^2(\hat{x}_1^2 + \hat{x}_2^2 + \gamma\hat{x}_1\hat{x}_2), \quad (2.1)$$

where coupling constant $\gamma = -4\frac{q^2}{\mu\omega^2}\frac{1}{R^3}$. Alternatively $\gamma = -4\frac{\alpha}{R^3}$ when written in terms of the dipole polarisability $\alpha = \frac{q^2}{\mu\omega^2}$. Defining the Bogluibov transformation

$$\hat{p}_{\pm} = \frac{1}{\sqrt{2}}(\hat{p}_1 \pm \hat{p}_2), \quad \hat{x}_{\pm} = \frac{1}{\sqrt{2}}(\hat{x}_1 \pm \hat{x}_2), \quad (2.2)$$

the Hamiltonian can be rewritten in this new basis as a pair of uncoupled oscillators:

$$\hat{H} = \left(\frac{\hat{p}_+^2}{2\mu} + \frac{1}{2}\mu\omega^2 \left(1 + \frac{\gamma}{2} \right) \hat{x}_+^2 \right) + \left(\frac{\hat{p}_-^2}{2\mu} + \frac{1}{2}\mu\omega^2 \left(1 - \frac{\gamma}{2} \right) \hat{x}_-^2 \right). \quad (2.3)$$

By identifying the prefactors on the $\hat{x}_{\pm}^2$ terms as trapping frequencies of the uncoupled oscillators we can read off the ground state energy

$$\begin{aligned} E &= \frac{1}{2}\hbar\omega \left(\sqrt{1 + \frac{\gamma}{2}} + \sqrt{1 - \frac{\gamma}{2}} \right) \\ &= \frac{1}{2}\hbar\omega \left(2 - \frac{\gamma^2}{16} + \mathcal{O}(\gamma^4) \right). \end{aligned} \quad (2.4)$$

Using the definition $\gamma = -\frac{4\alpha}{R^3}$ gives a leading order dispersive energy correction of

$$\Delta E = -\frac{1}{32}\gamma^2\hbar\omega = -\frac{1}{2}\frac{\alpha^2}{R^6}\hbar\omega. \quad (2.5)$$

The scaling of this proportional to α^2/R^6 matches that of the London dispersion interaction and the proportionality constant exactly matches that of the completely anisotropic non-polar linear molecules (See Appendix A.1 for further details).

Moving on to systems of more than two oscillators, the non-dimensional Hamiltonian for N one-dimensional QDOs with linear coupling is given by

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (2.6)$$

where $\hat{x}_i$ and $\hat{p}_i$ are the bosonic position and momentum operators for oscillator i and $\gamma_{i,j} \in \mathbb{R}$ is the coupling between oscillators i and j which are separated by distance $R_{i,j}$. From now on we will assume that all QDOs have the same trapping frequency and we will implicitly work in energy units of $\hbar\omega/2$. As with the pair of one dimensional oscillators, the coupling constant is given by $\gamma_{i,j} = -2(q_i q_j / \sqrt{\mu_i \mu_j} \omega^2) R_{i,j}^{-3} = -4\sqrt{\alpha_i \alpha_j} R_{i,j}^{-3}$, where for i^{th} QDO the effective charge of the Drude particle q_i , the effective mass μ_i and the oscillator frequency ω are model parameters which can be chosen. The dipole polarisability is given by $\alpha_i = q_i^2 / \mu_i \omega^2$. If pairs of QDOs instead align parallel to each other and parallel to the inter-oscillator axis, the coupling constant is given by $\gamma_{i,j} = 2\sqrt{\alpha_i \alpha_j} R_{i,j}^{-3}$. We will not consider such a geometry in this work.

The Hamiltonian in (2.6) is quadratic, this means that its ground state and thermal states consist of Gaussian states and can be solved using efficient classical methods. This allows us to efficiently benchmark our quantum algorithm in the experimental results in the following sections. The price paid for this computational simplicity is that interactions at various orders are related by Gaussian statistics - simple algebraic relations that relate the different values to one another. This results in symmetries within the model that lead to the absence of certain terms, such as the second hyperpolarisability (dipole-dipole-dipole-dipole interaction) [175].

In the following we will present how (2.6) can be solved using a variational quantum algorithm along with experimental results of doing so. We will later show how this model can be extended to three dimensional oscillators and include anharmonic

trapping and non-linear interaction terms, taking us out of the regime of Gaussian statistics and into one where classical algorithms are no longer efficient.

2.2.2 Algorithm description

We find the ground state of the 1D QDO Hamiltonians using the common procedure of the Variational Quantum Eigensolver (VQE) algorithm [47]. The aim of VQE is to use a quantum circuit containing gates parameterised by real parameters $\boldsymbol{\theta} = \{\theta_0, \theta_1, \dots\}$ to generate a variational quantum state $|\psi(\boldsymbol{\theta})\rangle$. The parameters $\boldsymbol{\theta}$ are varied by using a quantum-classical feedback loop in order to find an optimal set of parameters corresponding to the ground state of the problem Hamiltonian. Guided by the variational principle, using this procedure to solve the minimisation problem

$$\min_{\boldsymbol{\theta}} \langle \psi(\boldsymbol{\theta}) | \hat{H} | \psi(\boldsymbol{\theta}) \rangle \quad (2.7)$$

will lead to us finding the ground state of the model as well as ground state energy, provided the structure of our parameterised circuit (or ansatz) contains within it the ability to represent the true ground state. As is common with optimisation algorithms (particularly in the case of noisy measurements, complex cost function landscape or an ansatz that is not fully expressible³), we rather find an approximate solution to the optimisation problem above, hoping that our approximate ground state calculated approximates the true ground state and energy within a required precision.

To represent the states of QDOs we work within the (Bosonic) Fock basis. Since the single oscillator Fock states form an infinite basis and we will be working with finite dimensional qubits, we will necessarily have to restrict ourselves to basis containing a finite subset of the Fock states. This will not allow us to produce exact solutions

³In general if we do not know what form the ground state should take, a variational ansatz is only guaranteed to find the true ground state if it is fully expressive (can generate all states within the full Hilbert space). As will be discussed in Sec. 2.6, we typically want to avoid this and instead settle for ansatz that cover a subset of the Hilbert space that we believe contains reasonably good approximate solutions.

to the true infinite dimensional Hamiltonian. However, since dispersive effects in real systems are typically much smaller than the bare energy scale of $\hbar\omega$, we expect true ground states to be dominated by contributions from lowest energy Fock states. Thus, we expect that truncating the Hilbert space to just the first few Fock states will not severely impact the accuracy of solutions.

For a single QDO, we use a truncated basis of the first d Fock states. We will use a compact encoding [176], which has previously been proposed for calculation of molecular vibrational molecular modes on a quantum computer [60, 177]. For such an encoding, we require $m = \lceil \log_2 d \rceil$ qubits and use the Fock state to qubit state mapping

$$\mathcal{H}_{d-\text{Fock}} \rightarrow \mathcal{H}_2^{\otimes m} : |\underline{n}\rangle \rightarrow |\text{bin}_m(n)\rangle \quad (2.8)$$

where $\text{bin}_m(n)$ denotes the m -length binary representation of integer $0 \leq n < d$. To make the most use of a fixed number of qubits per QDO, we will typically choose d to be a power of two, i.e., $d = 2^m$. Throughout this work, underlined integers within a ket will denote a Fock state and non-underlined integers within kets will correspond to the computational basis of the qubits unless otherwise stated. The basis for the system of N QDOs will then be formed of a tensor product of the individual QDO bases: $(\mathcal{H}_2^{\otimes m})^{\otimes N}$ which will require a total of $mN =: M$ qubits. Fig. 2.1 shows a schematic representation of the algorithm and correlated charge fluctuations captured by the model for two oscillators.

In order to measure the cost function for a given variational state a during the optimisation procedure, the Hamiltonian in (2.6) must be decomposed into Pauli operators that can be measured on the quantum computer. Formally, this requires us to write $\hat{H}$ in terms of the basis on M qubits Pauli operators: $\{\mathbb{1}, X, Y, Z\}^{\otimes M}$, calculating the coefficients with which each contributes to the Hamiltonian:

$$\hat{H} = \sum_{\hat{P}_i \in \{\mathbb{1}, X, Y, Z\}^{\otimes M}} \beta_i \hat{P}_i \quad (2.9)$$

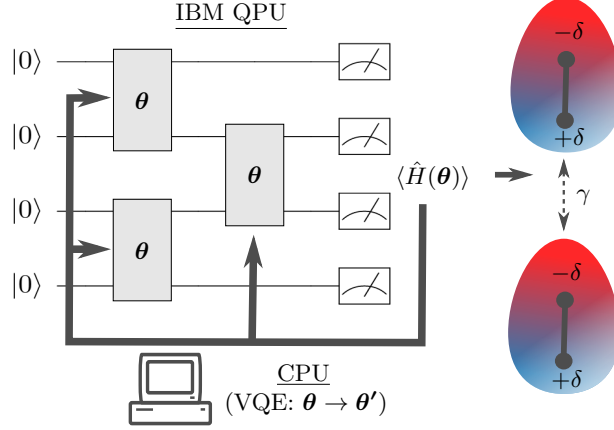


Figure 2.1: Schematic of the VQE optimisation procedure and output states showing instantaneous asymmetry of the molecular charge distribution. These correlated charge density fluctuations lead to dispersion interaction which, to the lowest order, is the familiar $1/R^6$ London attraction.

For each non-zero value of β_i , we measure $\langle \hat{P}_i \rangle$ on the QPU then sum them all on a CPU to give $\langle \hat{H} \rangle = \sum_{\hat{P}_i \in \{1, X, Y, Z\}^{\otimes M}} \beta_i \langle \hat{P}_i \rangle$. In the following, we derive the values of β_i for each $\hat{P}_i$ and calculate the number of non-zero values giving the number of unique circuits that must be measured for each evaluation of the cost-function. We will see that, although there are 4^M different operators in the Pauli basis only a small subset of them need to be measured to evaluate the cost-function.

For the non-interacting and interacting terms within the Hamiltonian, we will present the Pauli decomposition for systems of a single and two QDOs respectively. We note that the same decompositions apply to the N QDO Hamiltonian by applying these one and two oscillator decompositions to the correct subsets of m and $2m$ qubits respectively. Since we represent our states in the truncated Fock basis, the non-interacting contribution of a single QDO is simply a measurement in the computational basis consisting of m non-identity Z operators:

$$\hat{x}^2 + \hat{p}^2 = 2^m \mathbf{1}^{\otimes m} - \sum_{i=0}^{m-1} 2^i Z_i, \quad (2.10)$$

where Z_i is the Pauli Z operator acting on qubit i .

For a pair of m qubits QDOs, the coupled part of the Hamiltonian consists of $\left(\frac{1}{2}d \log_2 d\right)^2$ non-zero Pauli terms. The full structure along with a proof for the number of terms is given in Appendix. A.3. As examples, the Pauli decompositions for $m = 1$ and $m = 2$ are given by

$$\hat{x}_1 \hat{x}_2|_{m=1} = X_1 X_2, \quad (2.11)$$

$$\begin{aligned} \hat{x}_1 \hat{x}_2|_{m=2} = & \frac{\sqrt{3}+2}{4} X_1 X_3 + \frac{\sqrt{3}+1}{4\sqrt{2}} X_1 X_2 X_3 + \frac{\sqrt{3}+1}{4\sqrt{2}} X_1 X_3 X_4 \\ & + \frac{1}{4} X_1 X_2 X_3 X_4 + \frac{\sqrt{3}+1}{4\sqrt{2}} Y_1 Y_2 X_3 + \frac{1}{4} Y_1 Y_2 X_3 X_4 \\ & + \frac{\sqrt{3}+1}{4\sqrt{2}} X_1 Y_3 Y_4 + \frac{1}{4} X_1 X_2 Y_3 Y_4 - \frac{\sqrt{3}-1}{4\sqrt{2}} Y_1 Y_2 X_3 Z_4 \\ & - \frac{\sqrt{3}-1}{4\sqrt{2}} X_1 Z_2 Y_3 Y_4 - \frac{1}{4} X_1 Z_2 X_3 - \frac{\sqrt{3}-1}{4\sqrt{2}} X_1 Z_2 X_3 X_4 \\ & - \frac{1}{4} X_1 X_3 Z_4 - \frac{\sqrt{3}-1}{4\sqrt{2}} X_1 X_2 X_3 Z_4 + \frac{2-\sqrt{3}}{4} X_1 Z_2 X_3 Z_4 \\ & + \frac{1}{4} Y_1 Y_2 Y_3 Y_4. \end{aligned} \quad (2.12)$$

The Pauli terms corresponding to the interacting part of the Hamiltonian for N QDOs is constructed by combining these pair-wise interaction terms for all $N(N-1)/2$ pairs of QDOs, making sure to multiply by the corresponding coupling constants $\{\gamma_{i,j}\}$.

2.2.3 Measurement cost

If all coupling constants $\gamma_{i,j}$ are non-zero, the total number of non-identity Pauli operators that must be measured to evaluate the $\langle \hat{H} \rangle$ is $(\log_2 d)N + \frac{1}{8}d^2(\log_2 d)^2 N(N-1)$. Naively, $\hat{H}$ can be measured by measuring each of these terms as a separate circuit with some fixed number of experimental shots. The number of measurements required can be significantly reduced however, by grouping Pauli operators into collections that mutually commute, and since mutually commuting operators can be simultaneously diagonalised by a single unitary, measured at the same time. In general calculating the commuting collection optimally is expected to be NP-hard there exist many heuristic methods that efficiently find reasonably good

partitions, see for example Refs. [72, 172, 178–182]. Using the structure of our Hamiltonian, we derive an exact grouping of the terms into mutually-commuting subsets and therefore find an upper bound on the number of unique measurements that need to be performed to achieve a desired measurement accuracy. We restrict ourselves to finding subsets that *qubit-wise* commute since the rotation needed to perform the simultaneous measurements consists of just single qubit operators⁴. For our Hamiltonian (2.6), we find the number of collections of qubit-wise mutually commuting operators needed to measure $\langle \hat{H} \rangle$ to be $(N - 1) \left(\frac{d}{2} \log_2 d \right)^2 + 1$ or $N \left(\frac{d}{2} \log_2 d \right)^2 + 1$, for even and odd N respectively. In general, for Hamiltonian with all-to-all k -order couplings, the number of collections will be $\mathcal{O} \left((d \log_2 d)^k N^{k-1} \right)$. Heuristic sorting algorithms do reduce the number of collections relative to this, however an analytic prediction of how they will perform in general is not known. The scalings given here serve as an upper bound that can be used for theoretical analysis of algorithmic resource requirements. For a description of the grouping method, proof of these scalings as well as examples of the result of using heuristic grouping algorithms, see Appendix A.3.

Using this exact grouping of mutually-commuting operators, we estimate the expected number of experimental shots to measure $\langle \hat{H} \rangle$ within absolute error ϵ . Without knowing the exact form of the ground state of $\langle \hat{H} \rangle$ (if we did know this a priori, we would not need to use computational algorithms to find them!), we calculate the expected number of shots with respect to distributions of states within Hilbert space. Firstly, we consider a uniform spherical measure on the full Hilbert space [183, Chapter 7], effectively assuming no knowledge of the distribution of possible states generated during the algorithm. By assuming nothing about the nature of these states, this can be considered an upper bound on the number of shots required. Secondly, we consider a distribution consisting of just a single

⁴By qubit-wise commute, we mean that each single qubit operator within one of the multi-qubit Pauli operators commutes with the corresponding single qubit oscillator in the other. For example, the operators $\{X \otimes X, \mathbb{1} \otimes X\}$ qubit-wise commute (and so can be measured simultaneously by applying a Hadamard gate to each qubit and measuring in computational basis), whereas the operators $\{X \otimes X, Z \otimes Z\}$ commute, but do not qubit-wise commute (simultaneous measurement will involve performing a rotation into an entangled basis).

Measure	Ungrouped		Grouped	
	Circuits	Shots	Circuits	Shots
Spherical	$\mathcal{O}(N^2(\log_2 d)^2 d^2)$ ^a	$\mathcal{O}\left(\frac{\gamma^2 N^4 d^8}{\epsilon^2}\right)$ ^b	$\mathcal{O}(Nd^2)$ ^c	$\mathcal{O}\left(\frac{\gamma^2 N^3 (\log_2 d)^4 d^6}{\epsilon^2}\right)$ ^d
Pure		$\mathcal{O}\left(\frac{\gamma^2 N^4 d^4}{\epsilon^2}\right)$ ^e		$\mathcal{O}\left(\frac{\gamma^2 N^3 d^4}{\epsilon^2}\right)$ ^f

For full expressions and proofs see Appendix A.4: ^aTheorem A.3.1, ^bTheorem A.4.2, ^cTheorem A.3.6, ^dTheorem A.4.1, ^eTheorem A.4.4 and, ^fTheorem A.4.3.

Table 2.1: Number of unique circuits required using our exact grouping as well as expected number of shots required to achieve absolute measurement error ϵ for a system of N d -level one-dimensional QDOs. Here γ is a typical scale for the inter-oscillator couplings. The expectations for number of shots are calculated with respect to distribution over the uniform spherical measure of states (Spherical) as well as the pure state of the uncoupled QDOs (Pure). For ungrouped results all Pauli operators in the Hamiltonian are measured separately, for grouped results Pauli operators are measured in groups according to the method described in Appendix A.3 Theorem A.3.6.

pure state equal to the ground state of the uncoupled system i.e., $|\psi_0\rangle = |0\rangle^{\otimes mN}$, which we expect to be close to the ground state with small coupling. For this uncoupled state, there is no contribution to the error from coupling terms and so this can effectively be considered a lower bound in the number of shots required. The resulting requirements for each of these distributions are given in Table 2.1 for the case of measuring all Pauli operators separately as well as when grouping and measuring commuting operators as described previously. We see that for both distributions, grouping commuting operators reduces the required number of shots from $\mathcal{O}(N^4)$ to $\mathcal{O}(N^3)$ with respect to N . For the case of the uniform spherical measure, grouping also significantly reduces the polynomial scaling with respect to d . In all cases, the precise estimate of the shots required (with prefactors) as well as proofs for the number of shots required can be found in Appendix A.4. It is worth noting that the estimated number of shots quoted is an upper bound for the spherical measure and tight for the pure uncoupled state.

To learn about the number of shots required for modelling real molecules of interest, we calculated the number of shots given by the above bounds required to achieve a measurement error on $\langle \hat{H} \rangle$ of $\epsilon = 10\% \times \Delta E(R_{\text{vdW}})$. Here $\Delta E(R_{\text{vdW}})$ is the dispersion interaction energy between a pair of one-dimensional QDOs representing

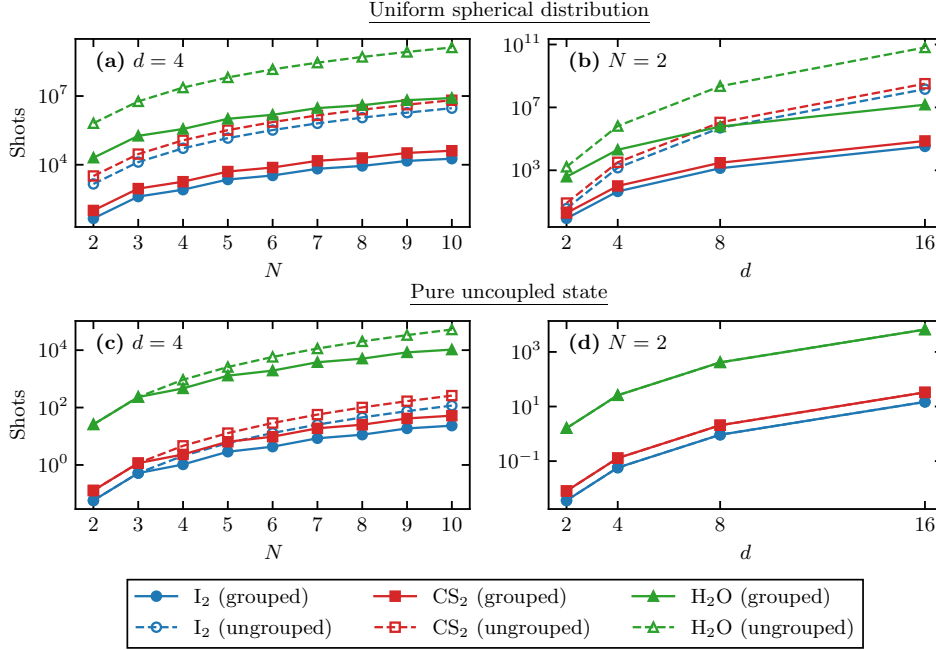


Figure 2.2: Estimated number of shots required to measure $\langle \hat{H}(R_{\text{vdW}}) \rangle$ with error $\epsilon = 10\% \times \Delta E(R_{\text{vdW}})$ for a range of different molecules. Here R_{vdW} is the van der Waals diameter of the respective molecules. The expected number of shots is calculated with respect to the spherically uniform measure [(a) and (b)] as well as with the uncoupled state [(c) and (d)] with and without grouping commuting operators. The number of shots is shown both as a function N for fixed $d = 4$ and as a function of d for fixed $N = 2$. Note that in (d), the grouped and ungrouped lines are indistinguishable from one another. Values for R_{vdW} and α are taken from Refs. [184–187].

molecules separated by their van der Waals radius (R_{vdW}). The results of these calculations are presented in Fig. 2.2. Assuming distributions of the uniform spherical measure as well as the pure uncoupled state, we see that the grouping can reduce the number of shots required by an order of magnitude, even for the relatively small values of N and d considered.

It is worth briefly comparing the measurement cost of this algorithm to that of orbital-based VQE. Although there are no good estimates for the number of shots required for orbital based VQE in the literature, there does exist scalings for the total number of collections of mutually commuting operators. The best scaling shown for the number of these collections is $\mathcal{O}(N^3)$ [172] where N is the number of orbitals. Comparing this to the $\mathcal{O}(N)$ scaling in the terms of the number of

QDOs N shown here⁵, we see that the measurement cost for our algorithm scales significantly better as a function of problem size.

2.2.4 Variational ansatz

In order to minimise circuit depth and to match device connectivity, for the experimental results in this paper we use a variational ansatz consisting of layers of two qubit operators. Because our problem Hamiltonian is positive we know that the ground state we desire will be real valued (up to a global phase). Therefore, we restrict our variational ansatz to consist of only real valued, two qubit operators. We use the optimal construction for $\text{SO}(4)$ given by Vatan and Williams [188]. In order to reduce the total number of variational parameters, we remove two parameterised R_z gates from the $\text{SO}(4)$ block, reducing the number of parameters from six to four and limiting the space accessible to a single block to a subset of $\text{SO}(4)$. This modified block is shown in Fig. 2.3b with the full ansatz (consisting of two layer) shown in Fig. 2.3a.

For future work aimed at achieving practical quantum advantage, it is likely that the ansatz used should be further inspired by the structure of the problem Hamiltonian. This would limit the space accessible to the ansatz (while hopefully containing the solution) in order to mitigate the effect of barren plateaus (we will discuss this further in Sec. 2.6). Although more costly than ansätze designed to make efficient use of the topology of existing quantum hardware [62], in the case of orbital based VQE, chemistry inspired ansätze such as unitary-coupled-cluster (UCC) [189] are believed to be better candidates for achieving quantum advantage. Although we do not consider them in this work, analogous ansätze for bosonic systems can be constructed [60, 162, 190] called the unitary-vibrational coupled-cluster (UVCC). This is inspired by the vibrational coupled cluster (VCC) used in classical computations [191, 192] but, unlike the VCC ansatz, the UVCC ansatz obeys the variational principle and has been suggested to better deal with strong static correlations [60].

⁵this is because our Hamiltonian only includes pairwise interaction terms, whereas the fermionic hamiltonian for orbital-based VQE includes four-body operators.

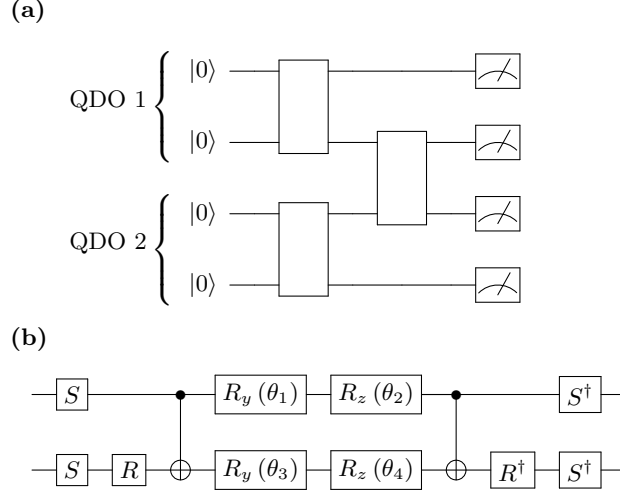


Figure 2.3: Circuits used in experiments. **(a)** Four qubit variational ansatz representing the state of two $d = 4$ QDOs. The upper (lower) qubit within the two-qubit register for each QDO corresponds to the most (least) significant bit in the binary encoding of the corresponding QDO Fock states. Each two-qubit block consists of a four-parameter real-valued operator. **(b)** Composition of the four-parameter real-valued operators. R_z and R_y gates in the centre are Pauli Z and Y rotation gates, each one contains a single variational parameter. The remaining gates are fixed with $S = R_z(\frac{\pi}{2})$ and $R = R_y(\frac{\pi}{2})$. This block is based on the general SO(4) operator from Ref. [188] with two parameterised R_z gates removed.

2.3 Experimental results using IBM quantum computers

As a demonstration of our algorithm on current generation quantum hardware we calculate the dispersion curve for a pair of interacting linear, non polar molecules. We choose model parameters to match those of an I_2 dimer, which forms a molecular liquid above its melting point of 114°C in which dispersion forces dominate the cohesion. In order to plot the dispersion energy, we calculate the ground state energy with our VQA for a range of intermolecular separations R .

Using a total of four qubits, we represent each molecule as a single $d = 4$ level QDO on two qubits. Since we model what are really three-dimensional molecular charge distributions with one-dimensional oscillators, we align the molecules with their molecular axis along the intermolecular axis such that the induced polarisation of the 1D QDO matches the direction in which the largest polarisation occurs in

the real linear molecule - along the molecular axis corresponding to the largest anisotropic polarisability coefficient.

We use a range of different IBM superconducting devices for the experiments presented here. For the optimisation part of the algorithm, we use the *ibmq_santiago* superconducting processor for the $R = \infty$ data point and we use the *ibmq_montreal* for all other values of R . See Appendix B.2 for further details of the hardware used in this work. In order to measure $\langle \hat{H} \rangle$ for the optimisation, the 21 Pauli operators making up the two oscillator Hamiltonian are partitioned into ten mutually-commuting groups as described in Sec. 2.2.3. Limited by hardware capabilities, rather than optimally distribute shot between these groups to minimise the variance, we run each of the ten measurement circuits with the maximum allowed 8192 shots. We use the gradient based ADAM optimiser (learning rate 0.25, decay rates 0.9, 0.99) [193] using the parameter shift rule [194] to evaluate gradients of the cost function with respect to the variational parameters. For the gates making up our ansatz, the parameter shift rule allows us to exactly⁶ calculate a single one of these gradients by making two measurements of cost function, one shifting the parameter of interest up by $\frac{\pi}{4}$ and the other by shifting it down by $\frac{\pi}{4}$. More precisely,

$$\frac{d}{d\theta_i} \langle \hat{H}(\boldsymbol{\theta}) \rangle = \left\langle \hat{H} \left(\boldsymbol{\theta} + \frac{\pi}{4} \mathbf{e}_i \right) \right\rangle - \left\langle \hat{H} \left(\boldsymbol{\theta} - \frac{\pi}{4} \mathbf{e}_i \right) \right\rangle, \quad (2.13)$$

where $\mathbf{e}_i$ is the unit vector corresponding to the θ_i direction.

When the optimisation procedure is complete and an optimal set of parameters found, we extracted the binding energy by using a simple scheme to subtract off device noise. The binding energy for a given finite value of R is given by

$$\Delta E(R) = E(R) - E(R = \infty), \quad (2.14)$$

where $E(R)$ is the ground state energy at R and with $E(R = \infty)$ corresponding to the ground state energy of the completely uncoupled state. If we assume the output of noisy quantum processors can be described by the unitary operation of

⁶by exactly, we mean that there is no error due to a finite difference scheme. The gradient will have errors associated with finite sampling and device noise.

the circuit governed by $\boldsymbol{\theta}$ along with a completely depolarising channel with error rate λ , the output state is given by

$$\rho(\boldsymbol{\theta}) = (1 - \lambda)|\psi(\boldsymbol{\theta})\rangle\langle\psi(\boldsymbol{\theta})| + \frac{\lambda}{2^{mN}}\mathbb{1}, \quad (2.15)$$

where $|\psi(\boldsymbol{\theta})\rangle$ is the state corresponding to a noiseless ansatz with parameters $\boldsymbol{\theta}$. Assuming optimisation was successful such that $|\psi(\boldsymbol{\theta})\rangle$ is the required ground state, then the energy evaluated on the noisy device is given by

$$\langle\hat{H}(R)\rangle_{\rho(\boldsymbol{\theta})} = (1 - \lambda)E(R) + \frac{\lambda}{2^{mN}} \text{Tr}[\hat{H}(R)]. \quad (2.16)$$

Therefore, if we assume λ is independent of the variational parameters and does not change between experiments, the dispersion energy is given by

$$\Delta E(R) = \frac{\langle\hat{H}(R)\rangle_{\rho(\boldsymbol{\theta})} - \langle\hat{H}(R = \infty)\rangle_{\rho(\boldsymbol{\theta}_\infty)}}{(1 - \lambda)}, \quad (2.17)$$

where $\boldsymbol{\theta}_\infty$ are variational parameters found for the ground state of $\hat{H}(R = \infty)$ and we have used that $\text{Tr}[\hat{H}(R)] = \text{Tr}[\hat{H}_{\text{non-int}}] = dN \forall R$.

We can then use (2.17) to estimate the dispersion energy using a noisy quantum device. We estimate λ by measuring the fidelity of the uncoupled ground state on the noisy computer $\mathcal{F} = \text{Tr}[\rho(\boldsymbol{\theta}_\infty)|0\rangle\langle 0|^{\otimes mN}] = (1 - \lambda) + \frac{\lambda}{2^{mN}}$. This can be done without additional cost by reusing the counts measured when calculating $\langle\hat{H}\rangle_{\rho(\boldsymbol{\theta}_\infty)}$. Although there are a number of large assumptions in getting to this result, we find that for the devices and short depth circuits used in these experiments, this method to subtract off the effect noise gives reasonably accurate estimates of the London dispersion energy, without the need for more complex or costly error mitigation techniques.

To perform this subtraction method to extract the binding energy for optimised parameter values we used the *ibm_lagos* quantum processor. Since we assume that the error rate λ is constant across all experiments, it was important to run these experiments under conditions that led to the most similar noise profiles across all

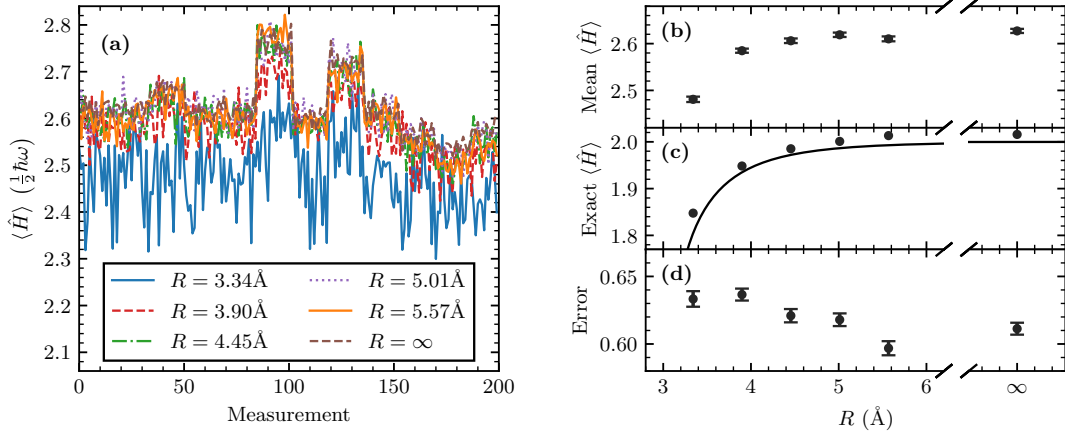


Figure 2.4: Experimental measurements used for error subtraction. **(a)** Measurements of $\langle \hat{H}(R) \rangle_{\rho(\theta_R)}$ where θ_R are the optimal parameters found for each separation R evaluated on *ibm_lagos* quantum computer. As described in the main text, each of the 200 experiments is performed for each value of R before moving onto the next one. **(b)** Mean values of the measurements for each R . **(c)** Exact calculation of $\hat{H}$ with no noise for the optimised parameters. Solid curve shows the true ground state. **(d)** Absolute difference between exact value of $\langle \hat{H} \rangle$ and mean of noisy measurements. This shows that, although the effect of noise varies with each experiment, by measuring each R in the order described, the effect of noise is approximately equal for each data point. Data points at $R = \infty$ correspond to the uncoupled Hamiltonian and parameters found using the optimisation procedure.

measurements. Firstly, the same qubits and recompiled circuits were used in all experiments as the noise on the device varies significantly for different qubits and gate operations. Secondly, due to the significant time dependence of the noise levels on the superconducting device, we had to evaluate expectation values for all values of R as close to simultaneously as possible. To measure the noisy Hamiltonian expectation values, we took the mean of 200 repetitions of each measurement (using 8192 shots per circuit and the same grouping scheme as during optimisation), performing the measurements for each R sequentially before moving onto the next of the 200 repetitions. We maximised the number of the circuit performed in a single job submitted via the cloud to the quantum processor and all jobs were submitted simultaneously to minimise the time between successive jobs. We see from Fig. 2.4 that the levels of noise vary significantly through the 200 evaluations performed, but by submitting the measurements in an inter-collated order, the drift seen in expectation values is approximately matched for each different value of R .

The final dispersion spectrum achieved by the optimisation and error subtraction method described is shown in Fig. 2.5. We see that when using our variational procedure and ansatz on real quantum compressors, a set of optimal parameters can be found that, when evaluated without device noise, corresponds to energy values close to the true ground state values of the system. Although significant level of device noise means that raw experimental energies are much larger than even the uncoupled energy of the system, our simple subtraction method allows us to partially account for these errors and to measure dispersion energies using the real device that closely match the exact London dispersion interaction. We find that for all values of R tested, the measured binding energy is within two standard deviations of the exact values.

It should be noted that, when evaluated the energy without device noise but with parameters found using a noisy quantum device (as shown in Fig. 2.4(b)), the values measured are all slightly above the true ground state of the system. There are a number of compounding factors that could contribute to this. Firstly, this could be the results of unconverged optimisation or the result of the algorithm getting stuck in local minima. Secondly, systematic errors in the rotation gates applied by the device can be accounted for as over/under rotations in the variational parameters learnt by the algorithm. These are obviously not reflected in expectation values measured using error-free simulation. Such issues can affect all VQAs and we expect their effects will be reduced as quantum hardware performance improves and could be somewhat mitigated by fine tuning the optimisation method used (to achieve better convergence to the global minimum).

2.4 Extending the model

2.4.1 Three-dimensional QDOs

As discussed in Sec. 2.2.1, full three-dimensional QDOs are required for a complete description of the dipole-dipole interaction and higher order terms. In the following

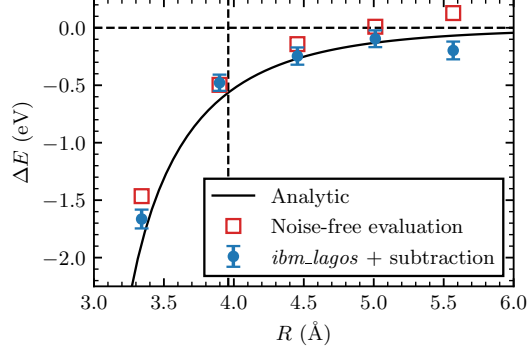


Figure 2.5: London dispersion energy (ΔE) calculated using variational quantum algorithm for one-dimensional quantum Drude oscillators computing interacting I_2 dimers. The black line gives the dispersion energy obtained analytically. Red squares correspond to noise-free evaluation (via state vector simulations) of the experimentally optimized coupled ground state. Blue circles correspond to energy directly obtained from experimentally optimized ground states (error bars represent one-standard deviation on the mean). The horizontal line marks zero dispersion interaction corresponding to uncoupled oscillators at $R = \infty$ and vertical dashed line corresponds to the I_2 van der Waals diameter.

we show how our variational algorithm can be simply extended to systems of three-dimensional QDOs.

The Hamiltonian for coupled three-dimensional QDOs is given by

$$\hat{H} = \sum_{i=1}^N \left(\frac{1}{2\mu} \hat{\mathbf{p}}_i^2 + \frac{1}{2} \mu \omega^2 \hat{\mathbf{x}}_i^2 \right) + q^2 \sum_{j>i} \frac{3(\hat{\mathbf{x}}_i \cdot \mathbf{R}_{i,j})(\hat{\mathbf{x}}_j \cdot \mathbf{R}_{i,j}) - \hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_j R_{i,j}^2}{R_{i,j}^5}, \quad (2.18)$$

where $\hat{\mathbf{x}}_i = (\hat{x}_i, \hat{y}_i, \hat{z}_i)$ and $\hat{\mathbf{p}}_i = (\hat{p}_{x_i}, \hat{p}_{y_i}, \hat{p}_{z_i})$ are vectors containing position and momentum operators for QDO i in each of the three Cartesian coordinates. $\mathbf{R}_{i,j}$ is the vector separating the centres of the pairs of QDOs i and j ; and μ , ω and q represent the effective mass, natural frequency and charge for the QDO particles which are assumed to be equal for all oscillators and isotropic in all three spatial dimensions. It is simple to relax these requirements if needed but we will work with them here for simplicity of notation.

Identifying the position and momentum operators $\{(\hat{x}, \hat{p}_x), (\hat{y}, \hat{p}_y), (\hat{z}, \hat{p}_z)\}$ for each of the oscillators into their own set of creation and annihilation operators we can define a (truncated) Fock basis for each one and encode them each into m qubits. Thus, the state of a single three dimensional oscillator state is encoded in a quantum

register of $3m$ qubits:

$$\begin{aligned} \mathcal{H}_{d-\text{Fock}}^{\otimes 3} &\rightarrow \mathcal{H}_2^{\otimes m} \otimes \mathcal{H}_2^{\otimes m} \otimes \mathcal{H}_2^{\otimes m} : \\ |\underline{n}_x, \underline{n}_y, \underline{n}_z\rangle &\rightarrow |\text{bin}_m(n_x), \text{bin}_m(n_y), \text{bin}_m(n_z)\rangle. \end{aligned} \quad (2.19)$$

As in Sec. 2.2.2, we can find an upper bound for the number circuits that are required to measure H is $(N-1)\left(\frac{d}{2}\log_2 d\right)^2 + 1$ or $N\left(\frac{d}{2}\log_2 d\right)^2 + 1$, for even and odd N respectively. For further details see Appendix A.3.

For $N = 2$ three dimensional oscillators, the Hamiltonian simplifies to

$$\hat{H} = \sum_{u \in \{x, y, z\}} (\hat{u}_1^2 + \hat{p}_{u_1}^2) + (\hat{u}_2^2 + \hat{p}_{u_2}^2) + \gamma_{uu} \hat{u}_1 \hat{u}_2, \quad (2.20)$$

where $\gamma_{xx} = -\alpha_{xx}R^{-3}$, $\gamma_{yy} = -\alpha_{yy}R^{-3}$ and $\gamma_{zz} = 2\alpha_{zz}R^{-3}$. Here α_{xx} , α_{yy} , α_{zz} represent the dipole polarisability along each of the three spatial axes (which we now allow to be anisotropic, corresponding to quasi-particle charge q being an anisotropic tensor) and the intermolecular axis is in the z direction; along which the physical separation is R . From this, we see that the Hamiltonian is equivalent to that of three independent interacting pairs of one dimensional oscillators (one corresponding to $u = x$, one to $u = y$ and one to $u = z$ in the above) with the appropriate choices of coupling constants. Thus, in Sec. 2.3 where we modelled the interaction of two linear molecules aligned along an axis with just one-dimensional QDOs, extension to a full three-dimensional system is trivial.

2.4.2 Induction, anharmonic and non-linear terms

Thus far we have only considered quadratic Hamiltonians in which onsite potentials are harmonic and inter-oscillator interactions are linear. As explained in Sec. 2.2.1 such a model can be solved efficiently using classical algorithms and is an incomplete description of the molecular system and a richer set of physics can be revealed by extending beyond quadratic Hamiltonians. Here we show how the Hamiltonian can be extended to include these higher-order interactions and anharmonic terms. We show how the decomposition and grouping strategies Sec. 2.2.3 can be applied when applying our VQA to such systems.

The simplest modification to the model is the inclusion of induction effects by applying a uniform external field E to the oscillators. For systems of two or more oscillators, such a modification can lead to non-additive effects [168]. For one-dimensional oscillators, the additional interaction term for a Hamiltonian of N oscillators is of the form $\hat{H}_{\text{E-field}} = \sum_{i=1}^N q_i E \hat{x}_i$, where q_i is the quasi-particle charge of QDO i . The expectation values of this can be measured using the same decompositions for $\hat{x}$ as were used previously. Since these are guaranteed to qubit-wise commute with terms of the form $\hat{x} \otimes \hat{x}$, these can be measured simultaneously and incur no additional circuits than the interacting harmonic Hamiltonian.

We can produce more realistic induction models by considering oscillators with a Coulomb potential resulting from one or more point charges. The additional Hamiltonian term will have the form $\hat{H}_{\text{coulomb}} = \sum_{i=0}^N \sum_{j=0}^{N_c} \phi_j(\hat{x}_i)$ for the case of N oscillators and in potentials ϕ_j resulting from N_c fixed charges. In order to measure a Hamiltonian of this form, each potential must be Taylor expanded to some finite order w :

$$\phi_j(\hat{x}_i) \sim \frac{1}{|r_j - \hat{x}_i|} \approx \frac{1}{r_j} + \frac{\hat{x}_i}{r_j^2} + \frac{\hat{x}_i^2}{r_j^3} + \cdots + \frac{\hat{x}_i^{(w-1)}}{r_j^w}, \quad (2.21)$$

where $r_j \gg |\hat{x}_i|$ is the position of the point charge relative to the centre of the QDO. Such systems could be used to model the effect of permanent charges doped into a system dominated by van der Waals interactions. Since each of the form $\hat{x}_i^w$ acts on distinct oscillators they can be measured simultaneously. The expected contribution to the required number of shots will also be $\mathcal{O}(N)$ with respect to N , leaving the asymptotical scaling with respect to number of oscillators unchanged.

Clearly the use of Taylor expansions to add anharmonic potentials is not limited to those described above. Limited only by the total Taylor order w , additional terms of the form $\hat{H}_{\text{anharmon}} = \omega_3 \hat{x}_i^3 + \cdots + \omega_w \hat{x}_i^w$ with appropriate model parameters defining coefficients $\{\omega_3, \omega_4, \dots, \omega_w\}$ can be used to produce arbitrary, possibly asymmetric, potentials. Using anharmonic trapping potentials of this form extends the coarse-grained model to regimes that have not been accessible for large systems using classical computation methods [175]. Such anharmonic trapping potentials

can lead to the inclusion of some higher order interactions (such as four-dipole hyperpolarisability [129, 175]) as well as non-linear optical effects such as the Kerr effect, second- and third- harmonic generation and, forms of optical mixing [195].

Finally, we can also introduce non-linear interaction terms. Consider as an example explicit three-body interaction terms of the form $\hat{H}_{\text{non-lin}} = \sum_{i,j,k=0}^N \beta_{\langle i,j,k \rangle} \hat{x}_i \hat{x}_j \hat{x}_k$. Terms similar to this have been considered for quantum algorithms of vibrational spectra where it has been noticed that they are particularly difficult to deal with using classical computational methods [60, 196]. If all $\mathcal{O}(N^3)$ possible terms in $\hat{H}_{\text{non-lin}}$ are non-zero, the exact grouping method described in Appendix A.4 will contribute an $\mathcal{O}(\frac{d^9 N^5}{\epsilon^2})$ cost to the number of required measurements. In practice, the scaling with respect to both N and d may be significantly lower as heuristic sorting methods can make use of qubit-wise commutation relations not identified in our exact grouping scheme. For higher order interactions, these commutation relations may become increasingly common. Additionally, for many physical systems we may not require all-to-all interactions or coupling values may decay strongly with distance such that many terms can be ignored reducing the cost significantly compared to that given. The amount of flexibility afforded by these terms as well as anharmonic potentials allows VQAs to be used to investigate a wide range of chemical and physical phenomena not yet studied using classical methods.

2.5 Further experimental results

2.5.1 Anharmonic systems

We test the performance of our algorithm on systems of one-dimensional anharmonic oscillators on both simulated and real quantum computers.

We first consider a system consisting of a single anharmonic oscillator with a quartic onsite potential term within an external field. For this, we use the Hamiltonian $\hat{H} = \hat{x}^2 + \hat{p}^2 + \frac{1}{5}\hat{x}^4 + E\hat{x}$ and vary the strength of the external field E . We use two qubits to store the state of the QDO with $d = 4$ Fock levels. As a variational ansatz we use a single four-parameter two qubit block given in Fig. 2.3(b).

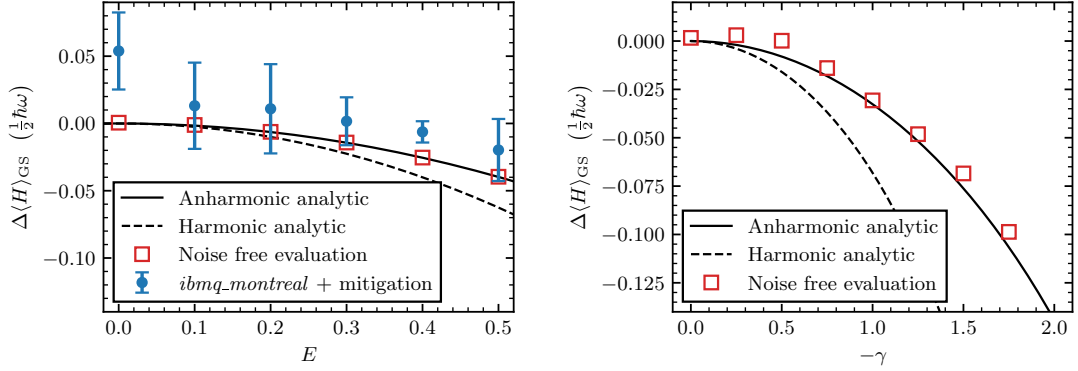


Figure 2.6: (a) Deviation from zero field ground state energy of 1D anharmonic oscillator in external field E . Blue circles correspond to evaluating the energy on the real *ibmq_montreal* quantum processor using error mitigation techniques described in the main text. Red squares correspond to evaluating the final energy with no device or shot noise. Both sets of data points use the same variational parameters which were found through optimisation performed on a simulated version of *ibmq_montreal*. Black solid line shows the exact spectrum found by numerically diagonalising the Hamiltonian. Black dashed line shows the exact spectrum for the analogous harmonic system (setting x^4 term to zero). (b) Deviation from uncoupled ground state energy for system of two coupled anharmonic oscillators with optimisation performed on simulated *ibmq_montreal* device. Data points show ground state energy relative to true ground state energy of uncoupled system, evaluated without device or shot noise. Solid line gives exact spectrum of the anharmonic system calculated by direct matrix diagonalisation. Dashed line gives exact spectrum of the corresponding harmonic system.

To optimise the variational parameters, we use a simulated version of the *ibmq_montreal* quantum processor and same ADAM optimiser and settings as in Sec. 2.3. For the first data point at $E = 0$, initial values of θ were taken randomly on the uniform interval $[-\pi/2, \pi/2]$ and optimised for 200 gradient descent steps. Following data points were run in order of increasing E . Initial values of θ were set by linearly extrapolating the optimised parameters from the previous data points with respect to the two corresponding values of E .

Using the values of θ calculated through the optimisation procedure, the final values of the ground state energy for each E were evaluated on *ibmq_montreal*. Since the system was no longer harmonic the ground state of the non-interacting ($E = 0$) system was no longer the zero state and so we were unable to use the method given in (2.17) to subtract off the effect of noise from the ground state energy. To account for the effect of device noise, we instead used two forms of error mitigation, namely

zero noise extrapolation [197, 198] and one based on the Lanczos method [199] which were used in combination. For a description of these two methods, as well as experimental settings used in this work see Appendix B.1.

The results of this experiment are shown in Fig. 2.6(a). The value of the energy found by the algorithm, when evaluated without device noise, agree well with the true value calculated through direct numerical diagonalisation of the Hamiltonian matrix. As with other experiments, despite high levels of (simulated) device noise in the experiment, the gradient descent procedure for this small system is able to find a good set of variational parameters. When these parameters are combined with error mitigation, some agreement is also found when evaluating the energy on the real quantum device however the large uncertainties in the measured values⁷ meant that the difference between interaction energies for harmonic and anharmonic systems cannot be resolved on the quantum device. We note that for the model parameters considered here, the energy resolution required to make this distinction is significantly smaller than that of the dispersion interaction in Sec. 2.3.

We also perform experiments for system of linearly coupled anharmonic oscillators. Adding a quartic onsite potential to the system in Sec. 2.3 we get the Hamiltonian $\hat{H} = \hat{x}_1^2 + \hat{p}_1^2 + \frac{1}{5}\hat{x}_1^4 + \hat{x}_2^2 + \hat{p}_2^2 + \frac{1}{5}\hat{x}_2^4 + \gamma\hat{x}_1\hat{x}_2$. Again, we use two qubits per oscillator as well as the ansatz given in Fig 2.3(a) and optimisation settings in Sec. 2.3. Initial values for θ for $\gamma = 0$ were again randomly chosen and for following data points chosen by extrapolating the optimal values found for previous values of γ . For the optimisation, we used a simulated version of *ibmq_montreal* device. Evaluating the energy corresponding to the parameters found using exact (noiseless) simulation we find that, without device noise, the states found by the algorithm give energies in close agreement to the true ground state. The results are shown in Fig. 2.6(b).

⁷A common drawback of many error mitigation methods is the significant increase in measurement variance. An increase in experimental shots (and therefore run time) as well as device noise will decrease this variance.

2.5.2 Many oscillators

Finally, we also test our algorithm for larger systems of $N = 3, 4$ and 5 oscillators. Using the quadratic Hamiltonian in (2.6) again, we define coupling constants $\gamma_{i,j}$ to represent one-dimensional QDOs placed on the edges of N sided polygons. The QDO were aligned such that they oscillate perpendicular to the inter-oscillator plane (i.e., normal to the polygon on which they are placed). For this system, $\gamma_{i,j} = 2\alpha/R_{i,j}^3 \propto 2\alpha/D^3$ where $R_{i,j}$ are the distances between oscillators i and j , D is the diameter of the circle that encloses the N vertex polygons. By varying $2\alpha/D^3$ (effectively shrinking or growing these polygons) we calculate the spectra shown in Fig. 2.7. The insets of this figure also show the geometry of each system.

The application of the VQA to these systems, both in the optimisation procedure and final energy evaluation, was simulated without device or shot noise. Again, we use qubits per oscillator to give local Hilbert space dimensions of $d = 4$. For $N = 3, 4, 5$ this corresponds to six, eight and ten qubits respectively. We use ansätze following the layer-wise structure of Fig. 2.3 for systems of more qubits, with the slight modification of including a cyclic geometry by applying additional $\text{SO}(4)$ blocks between the first and last qubits on odd-numbered layers. For each data point, we perform 500 optimisation steps using the same optimiser settings as before. For the $\alpha/D^3 = 0$ data points, $\boldsymbol{\theta}$ were initialised to values taken uniformly on the interval $[-\pi/2, \pi/2]$ with $\boldsymbol{\theta}$ for further data-points being initialised through extrapolation. For each geometry we test ansätze of varying depth to see how the increase in system size affect the circuit size and number of parameters required.

Fig. 2.7 shows that our VQA successfully recreates the energy spectrum for these systems of three, four and five oscillators. The three oscillator system is of particular interest as it is the smallest system in which effective many body effects, namely three body effects (known as the full Axilrod-Teller interaction energy [200] for a full three-dimensional oscillator), are induced from polarisable systems with only pairwise coupling. Across all three systems and coupling parameters considered, the truncated $d = 4$ spectrum matches the true analytic spectrum (corresponding to

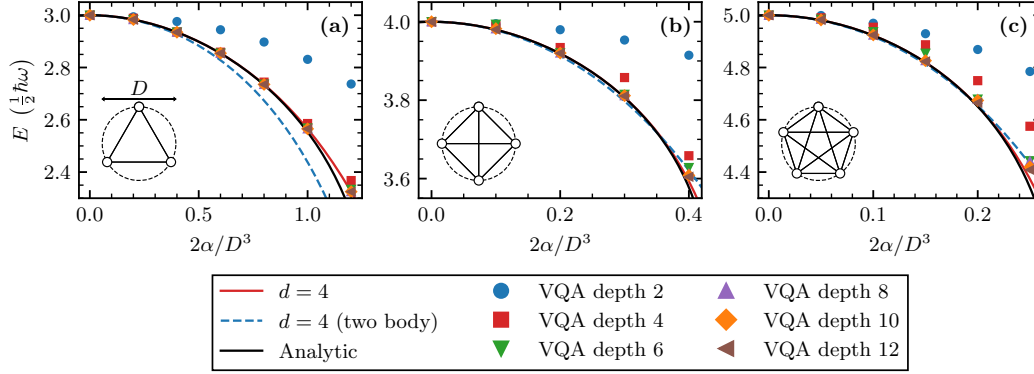


Figure 2.7: Simulation of VQA for systems of (a) three, (b) four and (c) five QDOs arranged in regular polygons using ansatz of varying depth. Insets show respective oscillator geometries, labelling the diameter D of the circle enclosing the polygons which was effectively varied to generate the spectra. Red solid line (labelled $d = 4$) corresponds to numerically diagonalising Hamiltonian of oscillators of $d = 4$ energy levels. Blue dashed line corresponds to the same system but only considering the two body pairwise interactions. Black solid line (labelled analytic) corresponds to using symplectic representation of many oscillator system (as described in Appendix B.3) which effectively corresponds to $d = \infty$. Data points correspond to outputs of VQA algorithm for varying depth ansatz.

$d = \infty$). Such a $d = \infty$ comparison is only possible with the restriction to quadratic Hamiltonians which allows us to calculate the analytic spectrum using symplectic methods. See Appendix B.3 for further details on this.

2.6 Discussion and conclusions

In this chapter we developed a VQA for modelling intermolecular interactions within a coarse-grained representation. We introduced a Fock state encoding scheme and variational ansatz circuit structure required for solving the model on digital quantum computers. We presented a method for grouping together mutually commuting operators and derived bounds on the number of circuits and shots that are required to measure the problem Hamiltonian. From this analysis, we demonstrated the significant reduction in cost achieved by applying this grouping method and have shown that the VQA we presented scales more favourably with system size than the most common (orbital based) VQE algorithms. We have shown how the

model can be simply extended to include anharmonic potential terms and higher order interactions with negligible measurement cost. Such modifications allow our algorithm to probe systems not easily accessible to quantum algorithms.

As with other VQE methods, the overall scaling of the algorithm we present - determined by the number of optimisation steps required to find a good ground state - is not known. It is however known that the classical optimisation required for VQAs is in general NP-hard[201]. In addition, the occurrence of barren plateaus [76, 77, 79, 80] in which optimisation gradients vanish with system size also suggest that the scalability of VQAs may be significantly limited. However, as seen in the field of classical machine learning, worst case complexity arguments [202] do not necessarily preclude the use of heuristic optimisation methods for practical applications, particularly when using gradient based methods to optimise systems of many parameters which seem to work remarkably well. Additionally, there are a number of algorithmic improvements, both in circuit design and optimisation procedure being investigated to mitigate these issues.

It is likely that many of the ideas being investigated to address barren plateaus in prototypical or general VQE algorithms can be applied to the Hamiltonian in this work. For the work presented here, a physically inspired ansatz such as the UVCC described in Sec. 2.2.4 may be necessary to avoid barren plateaus likely induced by the highly expressive ansätze [76] that we have used in this work. However, the local Hamiltonian⁸ presented here may be more resilient to barren plateaus than the global cost function used in orbital-based VQE.

We have also presented a number of experiments using our VQA on real and simulated quantum processes. We experimentally demonstrated that van der Waals dispersion energies, corresponding to some of the weakest but most ubiquitous intermolecular interactions, are resolvable on current generation quantum hardware when using this coarse grained approach. This is the first example of using a

⁸Local Hamiltonians contain Pauli terms which act non-trivially on a constant subset of the qubits. In comparison, a global cost function has terms acting on number of qubits proportional to the total system size.

quantum processor to study such non-covalent interactions which govern the behaviour of a large range of condensed states and molecular assemblies. While the quadratic model used in such experiments can be solved efficiently using classical methods, this work serves as a first step towards developing efficient and well characterised quantum algorithms to address systems outside the reach of classical methods.

Such quantum algorithms can serve to expand the palette of problems accessible to quantum computers by providing a level of coarse graining which still describes a complete set of long electronic responses, fluctuations and correlation phenomena. One future development could be to combine the coarse grained methodology with other electronic structure strategies in order to treat different parts of complex or multiscale systems to varying levels of physical accuracy. Such mixed system might consist of covalently bonded inorganic atoms treated to high levels of physical realism, placed within large baths of neutral but highly polarisable biological molecules described with a coarse grained representation.

Secondly, the coarse grained framework could be used to enhance existing bead-based models for branched polymer and protein folding. The combinatorial complexity of such models has already prompted investigations into solving such models using quantum computing [203, 204]. The simplistic interactions currently used to represent the interactions between different parts of the protein could be augmented or replaced by a coarse-grained representation in order to develop more realistic models for such folding problems.

In our work, we have shown how the quadratic Hamiltonian for the original harmonic coarse-grained system can be extended to include anharmonic and non-linear terms. We have shown that the experimental overhead arising from such terms is in many cases negligible and thus our VQA could be used to study coarse grained systems exhibiting a more rich set of physics than those accessible by classical methods. Furthermore, such methods are not limited to coarse grained chemical applications. VQAs have already been applied to vibrational chemical systems [60, 196] which

involve similar Hamiltonians. Lattice field theories (LFTs), such as ϕ^4 models, are also described by Hamiltonians similar to those investigated in this work. LFTs lead to a rich set of physical phenomena and play important roles in particle physics and theoretical condensed matter physics [205, 206]. Quantum algorithms have been proposed to study the time evolution of such models [207, 208]. Our VQA could be used to find vacuum states of LFTs, either within their own right for studying broken symmetry states or as inputs to scattering problems solved using time evolution on a quantum computer.

3

Solving lattice gauge theories using the quantum Krylov algorithm and qubitisation

Contents

3.1	Introduction	52
3.1.1	Classical and quantum computing for gauge field theories	52
3.1.2	Our contribution	55
3.1.3	Chapter layout	56
3.2	Hamiltonian lattice formulation of the single flavour Schwinger model	57
3.2.1	Full Hamiltonian	57
3.2.2	Gauging out Bosonic fields	59
3.3	Quantum subspace expansion and quantum singular value transforms	60
3.3.1	Block encoding and QSVT	60
3.3.2	Generating Krylov basis using QSVT	65
3.4	Numerical experiments	67
3.4.1	Modelling measurement noise	70
3.4.2	Dealing with measurement noise	71
3.4.3	Results	72
3.5	Estimating resource requirements	76
3.5.1	Sketch of proofs for gate costs	76
3.5.2	Total resource cost of algorithm	78
3.6	Discussion and conclusions	81

3.1 Introduction

3.1.1 Classical and quantum computing for gauge field theories

Gauge theories are quantum field theories (QFTs) in which a gauge field imposes the fundamental symmetry on the Lagrangian and mediates interactions. They play foundational roles within various fields of theoretical physics. Within particle physics, gauge fields describe fundamental particles and their interaction through the exchange of force carriers. The nature of the gauge symmetry imposed determines the nature of the theory with quantum electrodynamics (QED) arising from an abelian $U(1)$ symmetry and quantum chromodynamics (QCD) from a non-abelian $SU(3) \times SU(2)$ symmetry [209]. Outside of particle physics, gauge field theories play important roles in cosmology [210] and condensed matter physics [211].

Many gauge theories are too complex to study analytically and because many phenomena of interest arise non-perturbatively, efficient numerical methods have become increasingly important in studying gauge theories. Originally formulated in seminal work by Wilson [212], numerical study of such phenomena is now routinely performed using lattice gauge theories (LGTs) in which spacetime is discretized onto a lattice of points on which the values of matter and gauge fields are defined. Wilson's path integral formulation of LGTs saw extensive computation study using Monte-Carlo methods [213]. Applying Monte-Carlo algorithm to some Hamiltonians however, leads to the so called *sign problem* or *complex action problem* [214] in which the integrand contains negative terms or the action becomes complex (and so the Boltzmann weight cannot be directly interpreted as a probability density function). In dealing with these issues, for example by using a Taylor expansion [215], computing expectation values can involve summing many highly oscillatory terms of which many cancel due to alternating signs. This behaviour leads to statistical error that grows exponentially with system size [214]. Due to the minus sign in the Fermionic commutation relation the sign problem poses a significant challenge to

simulations involving interacting electrons, as well as lattice QCD with non-zero Baryon density.

One of the most widely studied gauge field theories is the Schwinger model [131] which describes charged particles and their interactions with electromagnetic fields within QED in $1+1D$ ¹. Being the simplest gauge theory, the Schwinger model plays an important role as a benchmark for numerical methods for solving gauge theories as well as a pedagogical tool that can be used to study a range of complex phenomena that arise in more complex quantum field theories. Such phenomena include chiral symmetry breaking in which an effective Fermionic mass arises from pairing of Fermion and anti-Fermions [216], confinement where Fermions only exist alongside their corresponding anti-particle and separating such pairs requires increasing energy as they are drawn further apart (analogous to quark confinement in $3+1D$ QCD) [217] and, phase transitions between massive and massless s [218].

Simulating the real time dynamics of the Schwinger model as well as the ground state in the presence of a chemical potential imbalance between different Fermion flavours both give rise to the sign problem. As such, the lattice Schwinger model has been used to benchmark novel computational approaches which which do not suffer from the sign problem². A promising classical approach are tensor network methods [219, 220] which use a Hamiltonian, rather than path-integral, representation and thus don't suffer from the sign problem. Matrix product states (MPS), a popular tensor network ansatz, are particularly well suited to gapped locally interacting systems in one spatial dimension. MPS have been used to study massless to massive phase transitions [218], CP-violating phase transitions [221], chiral condensation [222], real time dynamics [223, 224] and ground state entanglement structure [225] in the Schwinger model. Despite their successes in $1+1D$ however, extending tensor network methods to QFTs in two or more spatial dimensions pose significant challenges. MPS methods in more than one dimension often require bond-dimensions that grow

¹one space and one time dimension

²this is not to say that such computational approaches do not lead to their own computational bottlenecks

rapidly with system size and tensor networks embedded in higher dimensions, such as projected entangled pairs states (PEPS) [226], cannot be contracted efficiently.

Quantum computing may offer a pathway towards simulation of lattice gauge theories when the sign problem arises or tensor networks methods fail [227]. Multiple works have investigated algorithmic approaches to lattice gauge theories [228], and the Schwinger model in particular. These quantum computing methods typically aim to solve one of two problems. The first is to simulate time evolution of the model. Notably Shaw et al. [229] provide rigorous bounds for Trotterised time evolution of the model in both NISQ and fault tolerant settings which scale as $\tilde{\mathcal{O}}(N^{3/2}T^{3/2}\sqrt{x}\Lambda)$, where N is system size, T is simulation time, x is interaction strength and Λ is the local dimension of the gauge field (see Sec. 3.2 for further details of these model parameters)³. Other authors focused on performing experimental or simulated implementations and have demonstrated real time dynamics of the Schwinger model using Trotter methods [232], open quantum system dynamics [233] and variational approaches [234].

The second approach to using quantum computing for simulating the Schwinger model has been ground state computation. This has included methods based on variational quantum eigensolver (VQE) algorithms [234–241] and adiabatic algorithms [242, 243]. To date, the largest ground state computation using VQE has included 50 lattice sites performed on an IBM superconducting processor [240]. In the case of variational proposals, the ansatz can be engineered to preserve model symmetries such as approximate translation [235] and charge symmetry [239] which may help alleviate the impact of noise and barren plateaus, however it is still expected that such issues will affect the scalability of the optimisation procedure. Therefore, there is an onus to study alternate quantum algorithms to solve the Schwinger model and other LGTs that may be more scalable than VQE. Numerical investigations using tensor network methods have been able to simulate adiabatic quantum algorithms with up to 100 lattice sites by truncating the local Gauge

³They note that this scaling with Λ is more favourable than other methods based on qubitisation [230] or random compilation for Hamiltonian simulation (QDRIFT) [231].

field dimension to relatively small values ($\Lambda \leq 9$) and noting that the gap is approximately independent of system size and Hilbert space dimension for small interaction strength x . For larger values of x however, the gap decreases with system size which will increase the run time of the adiabatic algorithm.

3.1.2 Our contribution

Research studying the ground state computation for LGTs, and in particular the Schwinger model, on quantum computers has so far focused on the use of VQE methods [234–239, 244–247]. Given the similarity between methods for mapping high dimensional Fock spaces used for the coarse grained work in Chapter 2 and those used to represent gauge fields on a digital quantum computer, it would seem natural to apply the analysis and experimental methods we have previously used to the new application area of lattice gauge theories. It has become apparent however, both as part of our previous work as well as within the wider community that VQE methods present significant issues to scalability, particularly in regards to barren plateaus [76–83]. Moreover, the highly stochastic and time intensive optimisation procedure of VQE means that numerical studies of the cost and scaling of the optimisation routine is challenging using both simulated and real quantum processors.

Therefore, for the second half of this thesis we investigate the use of the more recent quantum subspace expansion (QSE) [103, 248] algorithm for solving the lattice Schwinger model. Since the QSE algorithm does not involve the stochastic optimisation procedure, not only does it avoid the issue of barren plateaus (although obstacles to the scaling of the algorithm may exist) but analysis of the performance and scaling of the algorithm with numerical methods may be more fruitful.

We present the first use of QSE for finding ground states of a lattice gauge theory. We use methods based on qubitisation and quantum singular value transforms to generate the necessary expectation values and apply this to the prototypical lattice Schwinger model. We provide a full description and resource estimation for the

qubitisation procedure required based on the linear combination of unitaries. As part of this, we propose a novel method for applying the LCU procedure to lattice models in which we exploit translational symmetry of the lattice. Compared to an existing proposal performing QSE with qubitisation [120], which can be applied to the Schwinger model with a gate cost of $\tilde{O}(N^2)$, the gate cost of our methods is $\tilde{O}(N)$ which reduces the resources by multiple orders of magnitude for system sizes of hundreds of lattice sites.

To estimate the total resource cost and run time of the QSE algorithm we perform extensive numerical simulation of the algorithm, including the effect of measurement noise for systems of up to $N = 26$ lattice sites. Such sizes are well beyond those observed in works studying the Schwinger model with VQE. We compare two algorithmic approaches for studying device noise based on thresholding [122] and partitioning [125] the subspace expansion method. Extrapolating from the numerical results we estimate the resource cost of using QSE to solve the lattice Schwinger model for lattice sizes comparable to state of the art classical results.

Although we consider the single flavour Schwinger model, for which the ground state problem does not lead to the sign problem, the analysis and numerics we present here can be extended to more complex models for which Monte Carlo methods fail. The work we present here serves as a first step towards understanding and using a new quantum algorithm to study lattice gauge theories.

3.1.3 Chapter layout

The remainder of the chapter is structured as follows. In Sec. 3.2 we introduce the Hamiltonian for the Schwinger model. This includes Sec. 3.2.1 detailing the full Hamiltonian explicitly including the gauge fields; and Sec. 3.2.2 showing how we can apply a gauge transformation to remove gauge fields at the expense of adding in non-local spin-spin interactions. In Sec. 3.3 we introduce the concept of quantum singular values transforms (QSVT) and qubitisation. This includes Sec. 3.3.1 defining a block encoding and QSVT and then Sec. 3.3.2 in which we show

how this can be used to generate a Krylov basis for QSE. In Sec. 3.4 we present our numerical simulations. This includes a description of how we model measurement noise in Sec. 3.4.1, how we use partitioned quantum subspace expansion to deal with this noise in 3.4.2, and a set of results in Sec. 3.4.3 showing the number of measurements and basis size needed to achieve a desired energy error. In Sec. 3.5 we calculate the resources required to run the algorithm. This includes Sec. 3.5.1 presenting a sketch of the proof method (full proofs are given in the Appendix), and Sec. 3.5.2 presenting the total gate costs of the algorithm. Finally in Sec 3.6 we discuss the results presented in this chapter.

3.2 Hamiltonian lattice formulation of the single flavour Schwinger model

3.2.1 Full Hamiltonian

The Schwinger model is usually represented on a lattice by using a “staggered Fermion” representation [249]. As we will show in the following, the staggered Fermion representation uses odd lattice sites to encode Fermions and even Fermions to encode anti-Fermions. The single flavour model is constructed as follows. We define a single component Fermion field at lattice site labelled by integer n with creation and annihilation operators $\phi^\dagger(n)$ and $\phi(n)$ respectively. These obey the usual commutation relations

$$\{\phi^\dagger(n), \phi(m)\} = \delta_{m,n}, \quad \{\phi(n), \phi(m)\} = 0. \quad (3.1)$$

The gauge field is defined on links connecting sites $(n, n+1)$ with operator

$$U(n, n+1) := e^{i\theta(n)}, \quad (3.2)$$

which mediates interactions between neighbouring Fermionic sites. Within the “compact” formalism, the onsite gauge-field becomes an angular variable on the lattice with conjugate spin variable $L(n)$ obeying

$$[\theta(n), L(m)] = i\delta_{m,n} \quad (3.3)$$

such that $L(n)$ has eigenvalues $L(n) = 0, \pm 1, \pm 2, \dots$ and $e^{i\theta(n)}$ and $e^{-i\theta(n)}$ act as raising and lowering operators respectively. That is $e^{i\theta(n)}$ ($e^{-i\theta(n)}$) has the effect of increasing (decreasing) the value of $L(n)$ linking lattice sites $(n, n+1)$ by one, while having no effect on $L(m \neq n)$ for other lattice sites.

We define model parameters m, g, a to be Fermionic mass, coupling constant and lattice spacing respectively. The full Hamiltonian in units of $g^2 a/2$ is

$$H = H_0 + xV \quad (3.4)$$

with

$$H_0 = \sum_{n=1}^{N-1} L^2(n) + \mu \sum_{n=1}^N (-1)^n \phi^\dagger(n) \phi(n) \quad (3.5)$$

$$V = -i \sum_{n=1}^{N-1} \left(\phi^\dagger(n) e^{i\theta(n)} \phi(n+1) - \text{h.c.} \right), \quad (3.6)$$

where $\mu = 2m/g^2 a$ and $x = 1/g^2 a^2$ are the re-scaled mass and coupling [250]. As we will see shortly, the alternating sign on the mass term allows us to incorporate both matter and anti-matter particles on a bipartite lattice.

Physical states must obey Gauss's law which ensures that the electric flux into and out of a lattice site is balanced by the onsite charge. On the lattice, Gauss's law is given by [250]

$$L(n) - L(n-1) = \phi^\dagger(n) \phi(n) - \frac{1}{2} (1 - (-1)^n) \quad \forall 0 \leq n < N. \quad (3.7)$$

This means that excitations $\psi^\dagger(n) \psi(n) = 1$ to an even n site generates -1 unit of flux between the gauge fields into and out of the lattice site. Thus, on an even lattice site $\psi^\dagger(n) \psi(n) = 1$ is interpreted as the existence of a quark, while $\psi^\dagger(n) \psi(n) = 0$ is an absence. Conversely, for an odd n lattice site, $\psi^\dagger(n) \psi(n) = 0$ generates $+1$ units of flux. Thus $\psi^\dagger(n) \psi(n) = 0$ can be interpreted as the existence of an anti-quark and $\psi^\dagger(n) \psi(n) = 1$ as the absence [243]. We can see from the mass term in (3.5) that this interpretation gives the expected mass addition μ for both quark and anti-quark.

It is common to impose zero field to the boundaries, namely $L(0) = 0$. In the $x \rightarrow 0$ (no coupling) limit, the ground state can be found by inspection to be

$$L(n) = 0 \quad \forall n, \quad \phi^\dagger(n)\phi(n) = \begin{cases} 0 & \text{even } n \\ 1 & \text{odd } n \end{cases}. \quad (3.8)$$

As one might expect for a non-interacting theory, this vacuum state corresponds to a complete absence of particles and anti-particles as well as zero gauge field.

Remark 3.2.1. *Since the occupation of neighbouring gauge fields can vary by at most ± 1 , the maximum value of $|L(n)|$ is $N - 1$ for states obeying Gauss's law.*

This is particularly useful when studying such models with quantum computers or tensor network methods in which it is necessary to truncate the dimensionality of the gauge fields to map them onto finite dimensional systems. In the case of a qubit or spin system, the number of qubits to encode the gauge field subspace required is $m = \lceil \log_2(N) \rceil + 1$, which gives maximum and minimum gauge field occupancies $\pm \Lambda$ of $\Lambda \geq (N - 1)$. It is likely that further truncation can be applied with minimal impact on accuracy of solution. It has been shown [251] that for time evolution and the computation of spectrally isolated eigenstates it suffices to approximate the gauge field with a truncated Fock space of dimension Λ that scales polylogarithmically with inverse error⁴. The ability to significantly truncate the dimensionality of the gauge fields has been noted in numerical tests [242], and is standard practice in tensor network methods for LGTs.

3.2.2 Gauging out Bosonic fields

The Fermionic operators within the Hamiltonian can be mapped to spin operators using the Jordan-Wigner transformation

$$\phi(n) = \prod_{j < n} [i\sigma_3(j)] \sigma^-(n), \quad \phi^\dagger(n) = \prod_{j < n} [-i\sigma_3(j)] \sigma^+(n), \quad (3.9)$$

⁴This theoretical bound on truncation scaling applies to many systems in which only a single Bosonic field is involved in each interaction term such as the Hubbard-Holstein [252] and Fröhlich [253] models, as well as other gauge field theories. The best known truncation bound for systems in which multiple Bosons interact, such as scalar ϕ^4 models [254] or the systems of interacting oscillators in Chapter 2, is less favourable; with permissible truncation scaling polynomially with inverse error desired [207, 251].

where $\sigma_1, \sigma_2, \sigma_3$ are the spin-half Pauli operators and $\sigma^\pm = (\sigma_1 \pm i\sigma_2)/2$. Under this transformation the Hamiltonian becomes

$$H_0 = \sum_{n=1}^{N-1} L^2(n) + \sum_{n=1}^N (-1)^n \left(\frac{\mu}{2} + \frac{\mu}{2} \sigma_3(n) \right), \quad (3.10)$$

$$V = \sum_{n=1}^{N-1} \left(\sigma^+(n) e^{i\theta(n)} \sigma^-(n+1) + \text{h.c.} \right) \quad (3.11)$$

and Gauss' law can be reformulated as

$$L(n) - L(n-1) = \frac{1}{2} (\sigma_3(n) + (-1)^n) \quad \forall 0 \leq n < N. \quad (3.12)$$

If we assume free boundaries $L(0) = 0$ we can eliminate the gauge field by using Gauss' law combined with the gauge transformation

$$\sigma^-(n) \rightarrow \prod_{j < n} \left[e^{-i\theta(j)} \right] \sigma^-(n), \quad (3.13)$$

in which case the Hamiltonian becomes

$$H_0 = \sum_{n=1}^N (-1)^n \left(\frac{\mu}{2} + \frac{\mu}{2} \sigma_3(n) \right) + \sum_{n=1}^{N-1} \left[\frac{1}{2} \sum_{m=1}^n (\sigma_3(m) + (-1)^m) \right]^2, \quad (3.14)$$

$$V = \sum_{n=1}^{N-1} \left(\sigma^+(n) \sigma^-(n+1) + \text{h.c.} \right). \quad (3.15)$$

Within this spin-only representation, the $x \rightarrow 0$ ground state corresponds to anti-Ferromagnetic state with $\sigma_3(n) = -(-1)^n \forall n$ [232].

3.3 Quantum subspace expansion and quantum singular value transforms

3.3.1 Block encoding and QSVT

Qubitisation and quantum singular value transformations (QSVT) have emerged as powerful methods for implementing arbitrary functions of non-unitary operators on quantum computers. It has been shown how many paradigmatic quantum algorithms can be implemented using such methods which in some cases offers significant improvements over resources requirements and scalings of the original proposals [255]. QSVT has become one of the most popular methods for implementing QPE and

time evolution algorithms. Although such methods are expected to require some level of fault tolerance, the latter application has already been demonstrated on a trapped-ion hardware platform [256].

In this work, we use QSVT via qubitisation to generate the polynomial expectation values required to generate the Krylov basis within QSE. In the following, we describe the key idea behind qubitisation and block encoding and then quantum singular value transforms.

Definition 3.3.1 (Block encoding, modified from [255]). *Let H be a Hamiltonian acting on Hilbert space $\mathcal{H}_s$ with spectral norm $\|H\| \leq 1$. A block-encoding of H is three unitary operators $(U, G, \tilde{G})$, where U acts on $\mathcal{H}_a \otimes \mathcal{H}_s$, for some auxiliary Hilbert space $\mathcal{H}_a$, and $|G\rangle_a := G|0\rangle_a \in \mathcal{H}_a$ and $|\tilde{G}\rangle_a := \tilde{G}|0\rangle_a \in \mathcal{H}_a$, are states such that*

$$(\langle \tilde{G}|_a \otimes \mathbb{1}_s) U (|G\rangle_a \otimes \mathbb{1}_s) = H. \quad (3.16)$$

From this definition, we see that block encoding allows us to encode the operation of a H within the subspace labelled by auxiliary states $|G\rangle, |\tilde{G}\rangle$ ⁵. It is often the case, such as in the original proposal [230], that the auxiliary states are taken to be equal, i.e., $|G\rangle = |\tilde{G}\rangle$. This is also the case in previous works using qubitisation for QSE. We will see later that we can reduce the algorithm cost by using an asymmetric encoding $|G\rangle \neq |\tilde{G}\rangle$.

Provided we can find unitary operators U, G and $\tilde{G}$ (if not equal to G), qubitisation allows us to measure operators and, as we will see later, functions of non-unitary operators using quantum circuits. In most situations the restriction on the spectral norm of H can be satisfied by rescaling the Hamiltonian.

The linear combination of unitaries (LCU) method [230] is a commonly used method for block encoding local Hamiltonians that we will use in this work. LCU works

⁵In the case where $G = \tilde{G}$ it is this state that gives rise to the name qubitisation. $\mathcal{H}_a$ acts as a qubit subspace spanned by $|G\rangle$ and its orthogonal complement. Functions of the block encoded operator are implemented by repeated rotations around the Bloch sphere of this effective qubit.

as follows. We express an n qubit Hamiltonian as a linear combination of T Pauli operators:

$$H = \sum_{i=0}^{T-1} \alpha_i \hat{P}_i, \quad (3.17)$$

where α_i are real coefficients and $\hat{P}_i \in \{\mathbb{1}, X, Y, Z\}^{\otimes n}$ is an n Pauli operator. In most situations we can assume that $\sum_{i=0}^{T-1} |\alpha_i| = 1$, such that $\|H\| < 1$, by renormalising the Hamiltonian. The block encoding unitary is

$$U = \sum_{i=1}^{T-1} |i\rangle\langle i|_a \otimes \hat{P}_i, \quad (3.18)$$

which corresponds to the application of each $\hat{P}_i$ controlled on the state of the auxiliary register being in $|i\rangle$.

As for the corresponding block encoding states, we divert from previous implementations of QSE with qubitisation [120] by using an assymetric block encoding $G \neq \tilde{G}$. For our work we will use

$$|G\rangle_a = \sum_{i=0}^{T-1} \sqrt{|\alpha_i|} |i\rangle_a \quad (3.19)$$

and

$$|\tilde{G}\rangle_a = \sum_{i=0}^{T-1} \text{sgn}(\alpha_i) \sqrt{|\alpha_i|} |i\rangle_a. \quad (3.20)$$

It is straight forward to confirm that these definitions of U , $|G\rangle$ and $|\tilde{G}\rangle$ satisfy Definition 3.3.1 for a block encoding. We will see later in this chapter that since G and $\tilde{G}$ differ only by the possible signs of α_i that they can be prepared using almost identical circuits.

The skill in performing a block encoding this way is to choose a labelling procedure $\{|i\rangle\}$ that makes efficient use of auxiliary qubits and leads to an efficient method of generating the block encoding state. We use a labelling procedure in which $|i\rangle$ are computational basis states corresponding to a pair of n length binary numbers $(\vec{x}, \vec{z})_i$ that give a binary encoding of n qubit Pauli operator $\hat{P}_i$. If $\vec{x}$ ($\vec{z}$) contains a one in bit j , then $\hat{P}_i$ includes an X (Z) operator on qubit j . Constituent Y operators arise when both vectors contain a one in the same position as $ZX = iY$. The operation of U in (3.18) can then be implemented with the following three sets

of gates. (1) A set of CX gates, each controlled on a qubit within the $\vec{x}$ register and targeting the corresponding qubit within the system register. (2) A set of CZ gates doing the same thing, except each controlled on one of the qubits within the $\vec{z}$ register. (3) A controlled S (i phase gate) acting between a qubit in the $\vec{x}$ register and the qubit in the same position within the $\vec{z}$ register.

To prepare $|G\rangle$ we must prepare a real superposition with amplitudes $\{\sqrt{|\alpha_i|}\}$ over computational basis states $|\vec{x}, \vec{z}\rangle$. To do so, a series of multi-controlled partial CNOT and SWAP gates must be applied with control and target qubits as well as rotation angles chosen in such a way as to rotate into different states with increasing Hamming weights with the correct amplitudes. The process and order by which this is done is quite complex, a full description is given in Appendix C.1.2. As described by Kirby [120] this method requires us to apply a different multi-controlled rotation gate for each Pauli operator within the Hamiltonian. We present later, a method for reducing the number of these gates required by exploiting translational symmetry within the Hamiltonian such that the most costly of these gates need only be applied to single gauge link within the lattice and then effectively copied over to the remaining lattice sites. This reduces the cost from $\tilde{\mathcal{O}}(N^2)$ (if applying the method given by Kirby [120] to our Hamiltonian) to $\tilde{\mathcal{O}}(N)$. The other block encoding state $|\tilde{G}\rangle$ can be generated using the same set of rotation gates except flipping the phase of some of the rotation angles in order to add an additional minus sign phase where needed.

In previous work [120] in which $\tilde{G} = G$, to include the negative signs on values of α_i a potentially large number of multicontrolled Z gates must be applied within U . In using our asymmetric encoding to include the required signs within $|\tilde{G}\rangle$ we remove this cost which, depending on the number of negative α_i , can dominate the cost of the algorithm.

Given access to the block encoding, functions of H can be implemented through the quantum singular value transforms (QSVT) [257]. For our purpose of computing a Krylov basis, the following theorem to generate polynomials of H will suffice.

Theorem 3.3.2 (Polynomials through QSVT, modified from [255, 257]). *Given Hamiltonian H with eigenvalues on the range $[-1, 1]$ and block encoding $(U, G, \tilde{G})$, we define the projectors*

$$\Pi := |G\rangle\langle G|, \quad \tilde{\Pi} := |\tilde{G}\rangle\langle \tilde{G}| \quad (3.21)$$

onto the subspaces spanned by $|G\rangle$ and $|\tilde{G}\rangle$ and corresponding rotations by angle $\varphi \in \mathbb{R}$,

$$\Pi_\varphi := e^{2i\varphi\Pi}, \quad \tilde{\Pi}_\varphi := e^{2i\varphi\tilde{\Pi}}. \quad (3.22)$$

Then, for any degree k polynomial $f : [-1, 1] \rightarrow [-1, 1]$,

- *if k is odd, there exists angles $\vec{\varphi} \in \mathbb{R}^k$ defining the operator*

$$U_{\vec{\varphi}} := \tilde{\Pi}_{\varphi_1} U \prod_{j=1}^{\frac{k-1}{2}} \left[\Pi_{\varphi_{2j}} U^\dagger \tilde{\Pi}_{\varphi_{2j+1}} U \right] \quad (3.23)$$

such that

$$\langle \tilde{G} | U_{\vec{\varphi}} | G \rangle = f(H), \quad (3.24)$$

- *if k is even, there exists angles $\vec{\varphi} \in \mathbb{R}^k$ defining the operator*

$$U_{\vec{\varphi}} := \prod_{j=1}^{\frac{k}{2}} \left[\Pi_{\varphi_{2j-1}} U^\dagger \tilde{\Pi}_{\varphi_{2j}} U \right], \quad (3.25)$$

such that

$$\langle G | U_{\vec{\varphi}} | G \rangle = f(H). \quad (3.26)$$

Moreover, there exists an efficient classical algorithm for computing $\vec{\varphi}$.

Corollary 3.3.3. *There exists $\vec{\varphi} \in \mathbb{R}^k$ such that if k is odd*

$$\left(\langle \tilde{G} | \otimes \langle \psi_0 | \right) U_{\vec{\varphi}} (|G\rangle \otimes |\psi_0\rangle) = \langle \psi_0 | H^k | \psi_0 \rangle \quad (3.27)$$

and if k is even

$$\left(\langle G | \otimes \langle \psi_0 | \right) U_{\vec{\varphi}} (|G\rangle \otimes |\psi_0\rangle) = \langle \psi_0 | H^k | \psi_0 \rangle. \quad (3.28)$$

We can use this to measure the expectation values required for $\mathbf{H}$ and $\mathbf{S}$ within a Krylov basis. The rotation gates required can be implemented as follows.

$$\Pi_{\varphi_j} = \begin{array}{c} |0\rangle \text{---} \oplus \text{---} \boxed{e^{2i\varphi_j Z}} \text{---} \oplus \text{---} |0\rangle \\ \text{---} \boxed{G} \text{---} \circ \text{---} \text{---} \circ \text{---} \boxed{G^\dagger} \text{---} \text{---} \end{array} \quad (3.29)$$

Proof. This can be confirmed by hand with the construction

and then cancelling the product $G^\dagger G$ which commute through the rotation on the ancilla qubit. Here $\overline{\text{CNOT}}$ denotes the multi-controlled NOT gate controlled on the qubits being in the zero, rather than one, state. $\square$

Recalling the definition of $\mathbf{H}$ and $\mathbf{S}$ for a Krylov basis given in (1.5), the matrix elements for a D dimensional Krylov space $\langle \psi_0 | H^k | \psi_0 \rangle$ for $k = 0, 1, \dots (2D - 1)$ must be estimated on a quantum computer. Given access to G , $\tilde{G}$ and U this can be done according to Corollary 3.3.3 as follows.

1. Generate reference state $|\psi_0\rangle$ in the system register and apply G to the auxiliary register to prepare $|G\rangle_a \otimes |\psi_0\rangle_s$.
2. For each φ_j in $\vec{\varphi}$, apply the block encoding operator U to both registers as well as the rotation operator Π_{φ_j} or $\tilde{\Pi}_{\varphi_j}$ (depending on if j is odd or even) to the auxiliary register.
3. Apply the reverse operation for initial state preparation to the system registers, as well as $G^\dagger$ or $\tilde{G}^\dagger$ depending on the parity of k to the auxiliary register.

4. Measure all system and auxiliary qubits in the computational basis and count result if the all zero state is measured.
5. Return to step 1. and repeat until enough measurements are performed to give desired precision. The expectation value of $\langle \psi_0 | H^k | \psi_0 \rangle$ will be the fraction of times that zero state is measured.

We apply the LCU procedure along with the method for generating $|G\rangle$ described by Kirby et al. (Appendix C in Ref. [120]) to the single flavour Schwinger model to give accurate gate and qubit counts for U and $|G\rangle$ required for the qubitisation and QSVT procedure. In doing so, we use and extend our analysis for Pauli decomposition of truncated Bosonic operators used in Chapter 2. Moreover, we exploit translational symmetry within our Hamiltonian to improve the gate scaling of G by a factor of system size N . Secondly, by combining this method with the use of an asymmetrical block encoding such that $\tilde{G} \neq G$, we are able to reduce the scaling of U by a factor of N . The cost of $\tilde{G}$ is the same as G . Since the procedures for generating $(U, G, \tilde{G})$ along with analysis resource requirements are quite involved, we defer detailed descriptions and calculation to Appendix C.

If we use m qubits to encode each Bosonic gauge field, then the total number of qubits required to represent the system and implement the block encoding procedure is $\mathcal{O}(mN)$. The number of gates required for U is $\mathcal{O}(N)$ and the number of gates required for G is $\mathcal{O}(mN)$. Both of these gate costs are reduced from $\mathcal{O}(mN^2)$ by using our method to exploit translational symmetry. Exact gate costs and prefactors are given in Theorems C.2.26 and C.2.21 in Appendix C. By using our novel method for exploiting translational symmetry, we reduce the gate cost of the algorithm by multiple orders of magnitude for relevant system sizes, as compared to the qubitisation based on the work of Kirby [120]. We will discuss the gate costs of the QSVT procedure, as well as the total algorithmic costs, in Sec. 3.5.

3.4 Numerical experiments

To analyse the performance of the QSE algorithm for the Schwinger Hamiltonian, we perform classical simulation based on sparse matrix multiplication. For the quantum computing algorithm, we propose working with the single flavour 1+1D Schwinger Hamiltonian of (3.10) in which we explicitly include the gauge fields. This is in contrast to using the gauge transformed model of (3.14) in which the transformed interaction term contains non-local terms which are costly to implement on quantum hardware with local qubit connectivity. When performing matrix sparse matrix multiplication with the Hamiltonian the cost of doing so is not dependant on the locality of interactions, but the rather memory usage required to store states. Thus, for our numerical experiments we work with the gauge transformed Hamiltonian of (3.14). This allows us to use dense vectors of length 2^N to represent the corresponding spin states, rather than length $2^{(N-1)m+N}$ which would be required if we included the Hilbert space of the gauge fields. Using this representation, we are able to generate Krylov bases for up to $N = 30$ lattice sites through sparse matrix multiplication and perform direct diagonalisation of the ground state for up to $N = 26$ using nodes with up to 3TB of RAM accessed via the Oxford Advanced Research Computing (ARC) Facility [258]. In all of our experiments we fix $\mu = 1.5$ and $x = 0.5$.

As input state we use the ground state of the corresponding $x = 0$ system given by (3.8). Since this state satisfies Gauss's law, and the Gauss law operators commute with the Hamiltonian used to generate the Krylov basis, the solutions generated by the QSE algorithm will automatically obey Gauss's law. This input state is simply a computational basis state generated by applying X operators to the odd lattice sites.

We begin by simulating the algorithm with no measurement noise, running the QSE algorithm for system sizes ranging from $N = 4$ to $N = 26$. For each system size, we perform QSE with increasing basis size D until the generalised eigenvalue problem becomes unstable. Although we do not introduce measurement noise at this

point, finite machine precision still means that the generalised eigenvalue problem becomes unstable for sufficiently large Krylov basis. These results are shown in Fig. 3.2(a). We see that for each value of N , the fractional energy error achieved by the algorithm decreases exponentially with Krylov basis size. As seen in Fig. 3.2(b), for the system sizes considered, our chosen initial state has reasonably large overlap with the true ground state. This overlap decreases approximately linearly with system size N over the range we consider.

To estimate the basis size required to achieve a given accuracy we perform a fit of the form $\log \Delta E/E = \chi D + \lambda$ for each value of N , where $\Delta E/E$ is the ground state fractional energy error, and D is the Krylov basis order. χ and λ are N dependent parameters that correspond to the gradient and intercept of the lines in Fig. 3.2 (a). For each value of N , we fit to data-points up to $D = 10$ ($D = 4$ in the case of $N = 4$ only) with the fitting parameters χ and λ shown in Figs. 3.2 (c) and (d) respectively. Using these fitting parameters, we are able to estimate the Krylov basis size required to compute the ground state energy within a desired precision, there is no noise due to measurement or numerical precision (and therefore no risk of instability in the generalised eigenvalue problem). Such estimates are given in Fig. 3.2 for a desired energy error of $\Delta E/E = 10^{-4}$. Considering the gradient of the curve drawn out by the data points, the rate at which D increases with N , decreases as we go from $N = 4$ to $N = 26$, with the value of D becoming approximately linear with N for the later data points. Assuming the required order continues to grow linearly, or the gradient with respect to N further decreases, by fitting a straight line to the final two data points and extrapolating this to larger N , we can estimate an upper bound on the required basis for a given N . Such a line is shown in Fig. 3.2 and corresponds to an upper bound of $D \approx 0.049N + 3.041$. This linear relationship, characterised by a shallow slope, suggests that the basis size cost does not exhibit a substantial increase as the system size grows assuming zero measurement error and infinite precision. However, to gain a full understanding of the resource cost of the algorithm we must analyse the problem in the presence of finite measurement

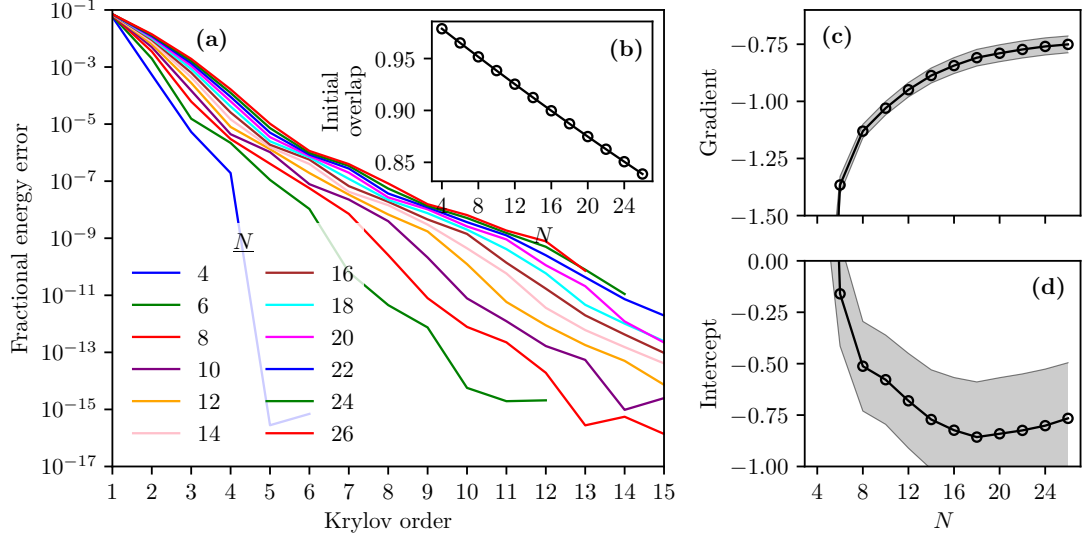


Figure 3.1: (a) Energy error as a function of maximum Krylov order for different value of Lattice size N . No measurement noise has been applied. (b) Overlap between initial state $|\psi_0\rangle$ and exact ground state. (c) & (d) Gradient and intercept for straight lines (on log-linear plot) fit to data in (a). Filled area indicates standard errors on estimates.

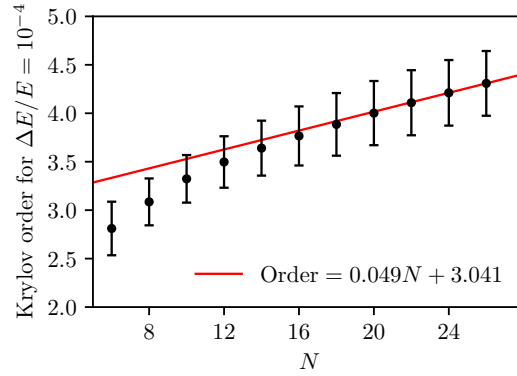


Figure 3.2: Estimated Krylov order required for QSE to achieve fractional energy error of 10^{-4} calculated using data from Fig. 3.1. Red line corresponds to linear fit to final two data points.

noise. As we will see from the following sections, the behaviour and cost of the QSE algorithm in such a scenario differs significantly from the zero-noise case.

3.4.1 Modelling measurement noise

Since we assume that the number of qubits and gates involved in the qubitisation procedure will require fault tolerant operation, we only consider the effect of measurement noise on the expectations values of $\mathbf{S}$ and $\mathbf{H}$. The effect of measurement noise is to perturb these matrices as $\mathbf{S} \rightarrow \mathbf{S} + \Delta_{\mathbf{S}}$ and $\mathbf{H} \rightarrow \mathbf{H} + \Delta_{\mathbf{H}}$. Like the original overlap matrices, the perturbation matrices $\Delta_{\mathbf{S}}, \Delta_{\mathbf{H}}$ are Hankel matrices⁶. Since expectation values are measured as Bernoulli trials, elements of the perturbation matrices can be modelled as random variables drawn from a normal distribution with width given by the expected variance of the expectation value.

Let $\overline{\langle \psi_0 | H^k | \psi_0 \rangle}$ denote the best estimate for the expectation value of $\langle \psi_0 | H^k | \psi_0 \rangle$ as measured according to the procedure in Sec. 3.3.2. If M_k measurements are used to calculate this estimate then the variance is given by

$$\text{Var } \overline{\langle \psi_0 | H^k | \psi_0 \rangle} = \frac{\text{Var } \langle \psi_0 | H^k | \psi_0 \rangle}{M_k} = \frac{\langle \psi_0 | H^{2k} | \psi_0 \rangle - \langle \psi_0 | H^k | \psi_0 \rangle^2}{M_k}. \quad (3.31)$$

From the above, we are able to calculate the variance of each matrix element used in the generalised eigenvalue problem by calculating values of $\langle \psi_0 | H^k | \psi_0 \rangle$ for k twice as large as the maximum value needed for the expectation values. We choose values of M_k as follows.

We use the total calls to the block encoding operator $\Pi_{\varphi}U$ as a measure of the cost of performing the measurements required (counting Π_{φ} and $\tilde{\Pi}_{\varphi}$ in the same way since they have the same cost). By controlling this, we effectively control the number of gates required. Using the QSVT procedure of Theorem 3.3.2, the circuit required to generate H^k requires k applications of $\Pi_{\varphi}U$. Thus, given a distribution of shots $\{M_k | k = 1, \dots, 2D + 1\}$, the total number of calls to $\Pi_{\varphi}U$ is

$$\# \Pi_{\varphi}U \text{ calls} = \sum_{k=1}^{2D+1} k \cdot M_k. \quad (3.32)$$

⁶in which ascending skew-diagonal from left to right is constant

We choose $\{M_k\}$ such that the measurement error as a fraction of expectation value is the same for all k , and as low as possible given a total number of U applications. Thus, using (3.31) we assign M_k as

$$M_k = \mathcal{M} \frac{\text{Var} \langle \psi_0 | H^k | \psi_0 \rangle}{\langle \psi_0 | H^k | \psi_0 \rangle^2}, \quad (3.33)$$

where proportionality constant $\mathcal{M}$ is calculated to normalise the total number of shots according to (3.32).

We can then sample elements of the perturbation matrices as $\Delta_{S_{i,j}} = \delta_{i+j}$, $\Delta_{H_{i,j}} = \delta_{i+j+1}$, where δ_k are drawn from a normal distribution of width

$$\sqrt{\text{Var} \langle \psi_0 | H^k | \psi_0 \rangle / M_k}. \quad (3.34)$$

3.4.2 Dealing with measurement noise

For simulations with measurement noise we use partitioned quantum subspace expansion (PQSE) which has been seen to perform better than thresholded approaches to avoid numerical instability in the presence of measurement noise [125]. Although the development of this algorithm is not the focus of this thesis, we describe the procedure in Algorithm 3.1.

To mitigate the effect of ill-conditioned matrices, PQSE breaks down the full D dimensional Krylov space into a series of smaller subspaces of dimension $\{d_i\}$ such that $\sum_i d_i = D$ in which the generalised eigenvalue problem contains smaller matrices which reduces the condition number. Each subspace of dimensions d_i is a Krylov subspace in which the input state is the optimal state from the previous subspace of dimension d_{i-1} . The size of each partitioned subspace is selected systematically throughout the PQSE algorithm by choosing d_i such that the solution to the generalised eigenvalue problem gives the smallest energy variance. Although the partitioned subspaces correspond to different input states, the overlap matrices $\mathbf{H}, \mathbf{S}$ required for each partition can be efficiently calculated from the same set of overlaps. As such, PQSE requires the same overlaps (and therefore the same quantum computation resources) as normal QSE, i.e. $\langle \psi_0 | H^k | \psi_0 \rangle$ for $k = 0, 1, \dots, (2D + 1)$.

For a full description of the PQSE algorithm, we refer the reader to the original publication [125].

Algorithm 3.1: Partitioned quantum subspace expansion

```

Data:  $\langle \psi_0 | H^k | \psi_0 \rangle$ ,  $k = 0, 1, \dots, (2D + 1)$  // Measured on QPU
 $d \leftarrow 1$ ;
 $\mathbf{S} \in \mathbb{R}^{D \times D} \leftarrow S_{i,j} = \langle \psi_0 | H^{i+j} | \psi_0 \rangle$ ;
 $\mathbf{H} \in \mathbb{R}^{D \times D} \leftarrow H_{i,j} = \langle \psi_0 | H^{i+j+1} | \psi_0 \rangle$ ;
while  $d < D$  do
     $\text{Var}_* \leftarrow \infty$ ;
    for  $d' \in \{1, 2, \dots, D - d\}$  do
        // Solve generalised eigenvalue problem for  $d'$  dimension
        // subspace (square brackets denote  $d' \times d'$  upper left
        // submatrices)
         $\vec{c}, E \leftarrow \min E \text{ s.t. } \mathbf{H}[:, d', : d'] \vec{c} = E \mathbf{S}[:, d', : d'] \vec{c}$ ;
         $\text{Var } E \leftarrow \sum_{i,j=0}^{d'-1} c_i^* H_{i+1,j} c_j - \left( \sum_{i,j=0}^{d'-1} c_i^* H_{i,j} c_j \right)^2$ ;
        // Choose subspace that gives lowest variance
        if  $\text{Var } E < \text{Var}_*$  then
             $E_* \leftarrow E$ ;
             $\text{Var}_* \leftarrow \text{Var } E$ ;
             $\vec{c}_* \leftarrow \vec{c}$ ;
             $d_* \leftarrow d'$ ;
        end
    end
    // Update overlap matrices with new subspace
     $d \leftarrow d + d_*$ ;
     $\mathbf{S} \in \mathbb{R}^{(D-d) \times (D-d)} \leftarrow S_{i,j} = \sum_{\alpha,\beta=0}^{(d_*-1)} c_{*,\alpha}^* S_{i+\alpha,j+\beta} c_{*,\beta}$ ;
     $\mathbf{H} \in \mathbb{R}^{(D-d) \times (D-d)} \leftarrow H_{i,j} = \sum_{\alpha,\beta=0}^{(d_*-1)} c_{*,\alpha}^* H_{i+\alpha,j+\beta} c_{*,\beta}$ ;
end
return  $E_*$ ;
  
```

3.4.3 Results

To test the performance of the PQSE algorithm in the presence of measurement noise we simulate noise as described in Section 3.1 for a range of different calls to block encoding operators $\Pi_\varphi U$. For a given value of ΠU calls, we split these calls among varying numbers of expectation values to apply the PQSE procedure with different total Krylov basis size D .

The results of this for lattice sizes varying from $N = 4$ to $N = 26$ are shown in

Fig. 3.3. For each choice of $\Pi_\varphi U$ calls and N , 100 noise instances are simulated with the median and upper and lower quartiles plotted. From this we notice three things. Firstly, for a given number of calls to $\Pi_\varphi U$, as we increase maximum Krylov order the fractional energy error starts off decreasing, until reaching a minimum at which point it begins to increase. We can interpret the initial decreasing behaviour as the error reduces due to the increased basis size - of course the error is worse than the noiseless equivalent due to error on the expectation values. Once the minimum is reached however, the matrices become sufficiently ill-conditioned such that any finite measurement noise is enough for the generalised eigenvalue problem to become unstable and so increasing order doesn't improve solution quality. In fact, when this happens increasing the maximum value of D increases the measurement error on expectation values as we have fixed the number of shots, and so error increases rapidly. Secondly, the value of D giving the lowest error increases with N . This is to be expected as noiseless results for this, as well as other Hamiltonians, typically require larger Krylov bases for larger systems to achieve the same error. Finally, in all cases increasing the number of calls to $\Pi_\varphi U$, and therefore reducing measurement error decreases overall solution error.

As in the noiseless case, we want to predict the number of resources required, this time total calls to $\Pi_\varphi U$ rather than basis size, to achieve a desired solution accuracy for a given system size. To estimate this, we begin by taking the smallest median energy error for each number of calls to U . These correspond to the lowest points on each curve in Fig. 3.3. This assumes that we have a reasonably good guess for the maximum order required to give the best results through PQSE. In practice, this could be inferred through some method of scaling up system size through real or simulated experiments. We see that for larger systems in Fig. 3.3, there are a number of choices of Maximum Krylov order that lead to near-optimal output error.

The best median error achieved by PQSE as a function of $\Pi_\varphi U$ calls is shown in Fig. 3.4 (a) for each value of N . Unlike for the noiseless case in which we saw an approximately log-linear relationship between the error and basis size, for a given N

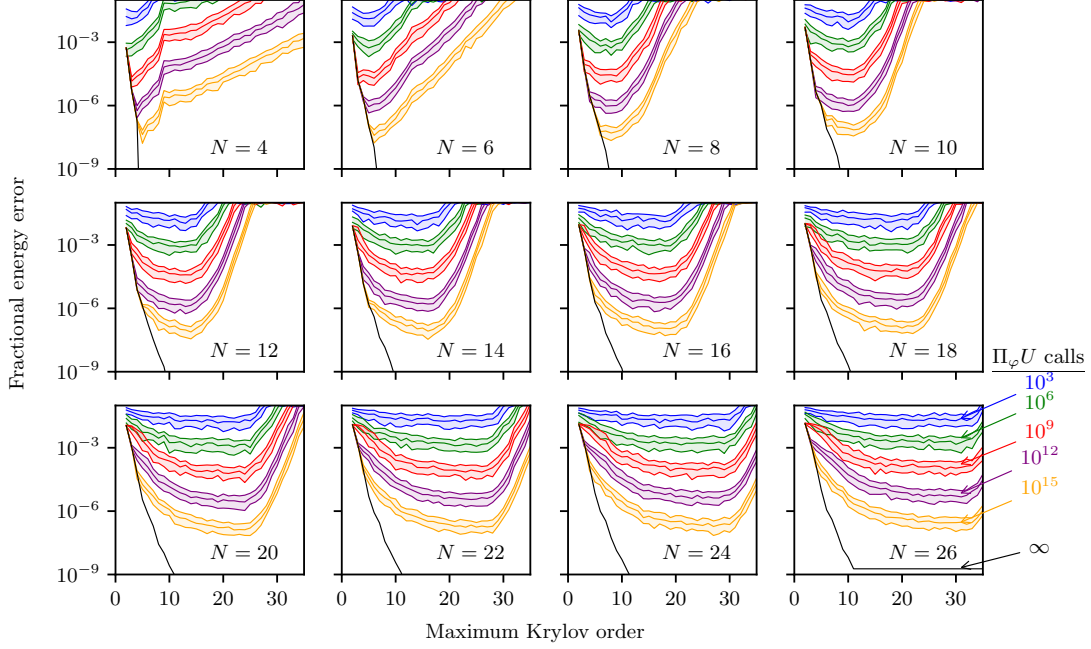


Figure 3.3: Fractional energy error with PQSE procedure with finite number of measurements for range of different lattice sizes N . Values are shown as a function of the maximum allowed Krylov basis for the PQSE procedure and as a function of number of calls to the block encoding operator U (which affect the number of measurements as defined in the main text). Lines and shaded area corresponds to the median as well as upper and lower quartiles of energy error. The black line (labelled as infinite calls to $\Pi_\varphi U$) corresponds to no measurement noise.

we observe a strong log-log dependence between the fractional energy error and the circuit cost (number of calls to $\Pi_\varphi U$). Fitting to this data (see Fig. 3.3 (b) and (c) for fitting parameters) allows us to estimate the number of calls required to achieve a specified accuracy. The result of this, to achieve a fractional energy error of 10^{-4} is shown in Fig. 3.5. For the values of N studied, the number of calls to U grows approximately exponentially with N as $\#\Pi_\varphi U \text{ calls} \approx 10^{7.888} \cdot 1.119^N$.

Applying thresholded QSE as described by Epperly et al. [122] for the same lattice sizes gives energy errors that are approximately an order of magnitude worse than PQSE with same number of resources. See Appendix D for further details.

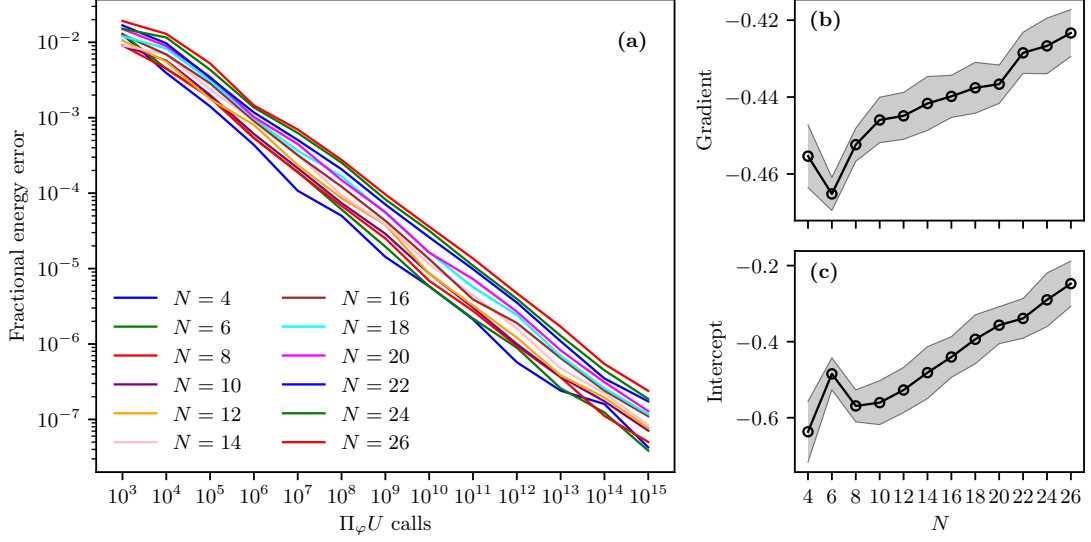


Figure 3.4: (a) Median energy achieved in Fig. 3.3 for optimal choices of maximum Krylov order as a function of calls to $\Pi_\varphi U$. (b) & (c) Gradient and intercept for straight lines (on log-log plot) fit to data in (a). Filled area indicates standard errors on estimates.

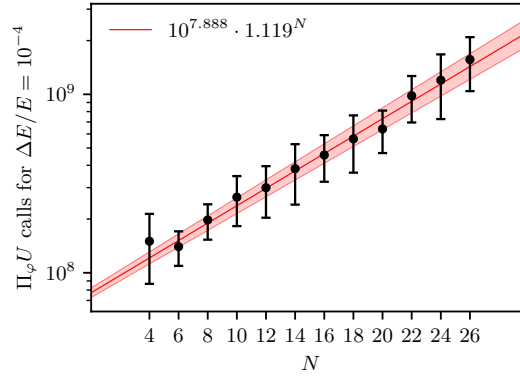


Figure 3.5: Estimated number of call to qubitisation procedure using PQSE to achieve fractional energy error of 10^{-4} calculated using data from Fig. 3.3. Red line corresponds to exponential fit (straight line on log-linear axes) with filled area indicating uncertainty in fitting and extrapolation.

3.5 Estimating resource requirements

After making an estimate of the basis size; and therefore number of calls to the qubitisation procedure required, to understand the total resource cost of solving the Schwinger model we must now compute the number of gates required to implement each part of the quantum circuit. A full description of the different parts of the circuits, along with the mathematics needed to estimate the gate cost is long-winded. Therefore, we defer a full description of the algorithmic implementation along with proofs for upper bounds of the required gate costs to Appendix. C. The final results of this are presented in Theorems C.2.21, C.2.26, C.2.29, C.2.32. We give a brief outline of the results and proofs in the following.

3.5.1 Sketch of proofs for gate costs

We start by calculating the Pauli decomposition for each term within the Schwinger model Hamiltonian. This is done in Appendix C.2.2. Since the gauge fields are Bosonic fields these Pauli decompositions take similar forms to the operators describing the oscillators used in Chapter 2. For each term in the Pauli decomposition we compute the Hamming weight of the bit string used to encode the operator within the LCU qubitisation procedure.

3.5.1.1 Calculating the gate cost of U

Calculating an upper bound for the cost of U consists of two parts. The first set of two qubit gates to apply the controlled X, controlled Z and controlled S (i phase) gates required to apply Pauli operators $\hat{P}_i$. For the CX and CZ gates, there is one of these for each qubit within the corresponding $\vec{x}$ and $\vec{z}$ auxiliary register. The controlled S gates are each controlled on a qubit within the $\vec{x}$ register, targeting the corresponding qubit within the z register. Thus the total cost of this step is just $\sim 3mN$ two qubit gates. A full proof of this cost is given in Theorem C.2.26

We also give, in Theorem C.2.28, the cost of U if we apply the phases of α_i rather than using an asymmetric block encoding and applying phases within $\tilde{G}$. Applying

phases within U gives a gate cost that is $\mathcal{O}(N^2 \log(N))$ so we achieve a significant saving by applying phases within $\tilde{G}$.

3.5.1.2 Calculating the gate cost of G and $\tilde{G}$

As described by Kirby et al. [120], $|G\rangle$ can be generated by applying partial rotations between different computational basis states of the $|G\rangle$ register where the different states rotated into encode the binary encoding of each Pauli decomposition with the Hamiltonian. This cost is calculated in Appendix C.2.3 by computing sums for the number of terms with a given Hamming weight and therefore the number of multicontrolled SWAP and partial CNOT gates required. Having computed a decomposition of the multi-qubit gates into T-gates, CNOTs and R_z gates (Lemma C.2.20), we can then calculate the total number of non-Clifford (T and rotation gates) along with two qubit CNOT gates.

We are able to reduce the cost of G compared to the method described by Kirby et al. by exploiting the translational symmetry within the Hamiltonian. Noticing that gauge field and interaction terms within the Hamiltonian are the same for each gauge-link site, rather than rotating between the binary states required for all lattice sites, we perform the costly rotations into basis states $|\vec{x}, \vec{z}\rangle$ for just a single link site and then perform a less costly partial swap of all terms into the other sites. This method is described in Appendix C.2.3.6 and the total gate cost for G , using this method is presented in Theorem C.2.21.

To convert the single qubit R_z gate counts into the more relevant T-gates, we make use of the approximate decomposition of R_z given by Ross et al. [259]. This decomposition requires a number of T-gates per rotation gate that scales logarithmically with the required rotation precision (see Proposition C.2.23). Assuming that required rotation error scales inversely with the average values of the amplitudes of each Pauli decomposition (and therefore the number of terms in the Hamiltonian) allows us to estimate the number of T-gates required for the rotations within G . For further details see Appendix C.2.4.

If we use m qubits for each gauge link field (such that they have maximum local occupancy of $\Lambda = 2^{m-1}$, the number of number of gates required scales as $\mathcal{O}(m2^m + Nm)$. Recalling from Remark 3.2.1 that the maximum gauge field value is equal to the number of lattice sites, the maximum value of Λ required is N and this gate cost is $\mathcal{O}(\log(N)N)$. Since the number of qubits required to represent the system is also $\mathcal{O}(\log(N)N)$, we expect that such a scaling for the gate cost, up to constant prefactors, is optimal. Without using our method to reduce the cost through translational symmetry, the gate cost of G would be $\mathcal{O}(\log(N)N^2)$.

We note in Corollary C.2.31 that $\tilde{G}$ can be implemented with the same rotation gates as G except flipping the sign of some of the angles. Thus, the cost for $\tilde{G}$ is the same as the cost for G .

3.5.2 Total resource cost of algorithm

Now that we have calculated the gate costs for G and $\tilde{G}$, it is easy to calculate the gate costs of corresponding rotations Π_φ and $\tilde{\Pi}_\varphi$. This is given in Theorem C.2.29 and is twice the cost of G with an additional number $\mathcal{O}(mN)$ number of CNOT and T-gates. Combining these with the cost of U allows us to calculate the gate cost to perform a single measurement of $\langle \psi_0 | H^k | \psi_0 \rangle$. The cost of this, as described in Theorem C.2.32, is $\mathcal{O}(k \log(N)N)$.

In Fig. 3.6 (a) we plot the number of gates as a function of system size N for a single step of the block encoding procedure. This is given by the cost of U plus Π_φ (which is the same as $\tilde{\Pi}_\varphi$). Fig. 3.6 (b) gives the qubit cost of the algorithm which, for systems of N in the hundreds or thousands, is not far off the thousands of physical qubits targeted by, for example IBM, in the next few years. Figs. 3.6 (c) and (d) show a direct comparison between our improved implementations of the block encoding procedure described in the previous section and the more general method of block encoding using a linear combination of unitaries. We see that for both T-gate count and CNOT count, our optimisations lead to significant reductions in the gate cost.

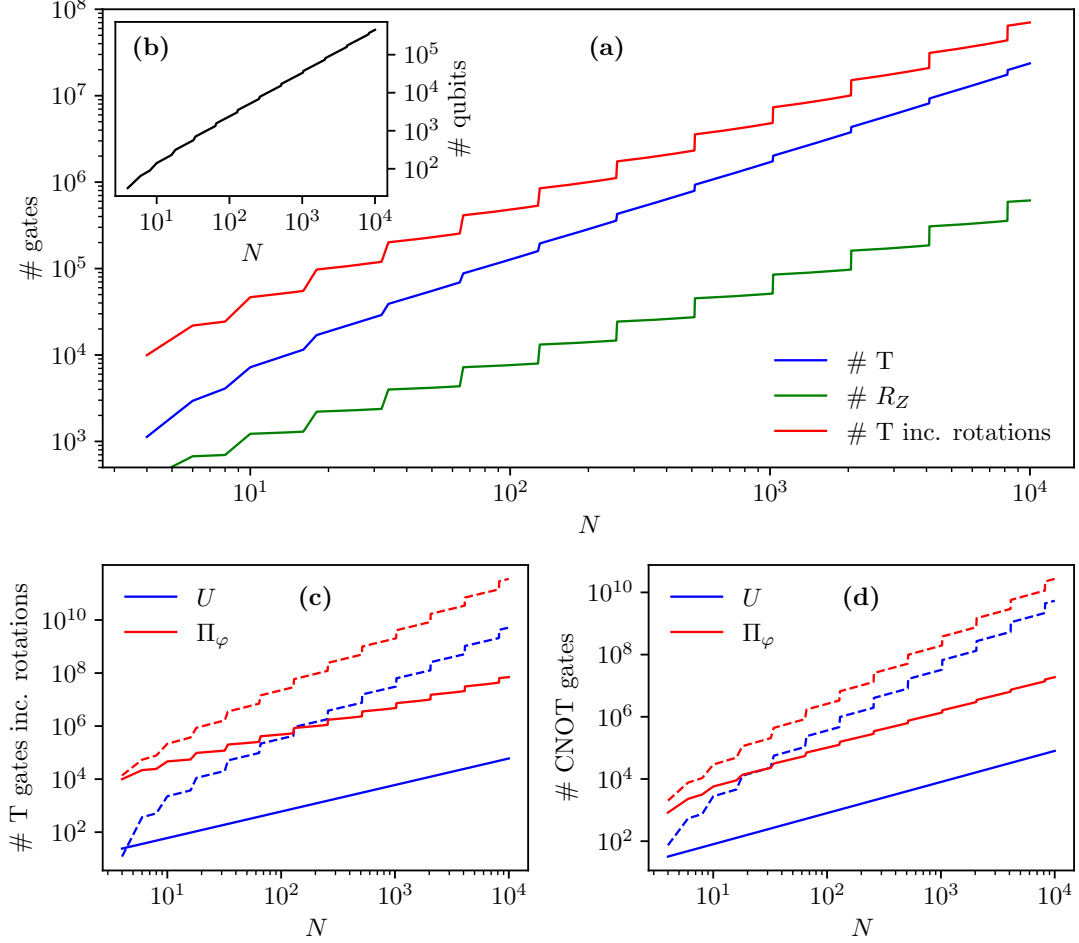


Figure 3.6: Upper bounds for resource costs in a single step of QSVT procedure (single application of U and Π_φ). **(a)** Total gate costs as a function of lattice size. The line labelled “# T-gates” corresponds to T-gates arising from controlled operations excluding rotation gates. The line labelled “# T-gates inc. rotations” includes the additional cost of converting single qubit rotation gates (line labelled “ R_z gate”) using Ref. [259]. Although not shown, the CNOT gate cost is indistinguishable from the line T-gate line on this scale. **(b)** Total logical qubit count. **(c)** T-gate cost for individual operators U and Π_φ . Dashed lines give costs for U and Π_φ using LCU procedure as described by Kirby et al. (see Appendix C.1.2). Solid lines give cost using our improved methods (see Appendix C.2.3.6 for details). **(d)** Same as (c) except for CNOT gates.

One way to assess the feasibility of our algorithm is to compare the expected run time of the algorithmic gates to the coherence times for existing, non fault-tolerant, processors. Assuming that two qubit gates are run in series and give the largest contribution to run time we find that the time needed to execute the CNOTs required for a single step of the QSVT procedure (the application of $\tilde{\Pi}_\varphi U$) is larger than the coherence times systems of $N = 100$ for a range of current generation quantum processors. For systems of $N = 1000$ and $N = 10000$, the run time of the gates is many orders of magnitude larger than the coherence times of the devices. See Appendix E for further details on what hardware specifications were used for these comparison.

Even assuming that two-qubit gates can be maximally parallelised (running $N_{\text{qubits}}/2$ two-qubit gates at the same time), the time to execute $\tilde{\Pi}_\varphi U$ can be of the order of $\sim 0.1\%$ to $\sim 1000\%$ times the coherence times. Such estimates indicate the circuit lengths required to solve the Schwinger model, at least using the implementation presented here, are beyond what is possible on currently available devices and those that we can expect in the foreseeable future. Therefore, for system sizes of hundreds of lattice site the algorithm we present must be run in an error-corrected setting.

In an error-corrected setting, rather than CNOT gates dominating the execution time of the algorithm, the number of non-Clifford gates determine how long a computation will take to run. We estimate the total number of T-gates required to calculate the ground state energy by combining the circuit costs for a single block encoding step with the total number of calls to the block encoding estimated in extrapolating the numerical results of Section 3.4 as in Fig. 3.5. This cost is shown for a range of desired errors in Fig. 3.7. We see that, for systems of hundreds of lattice sites, the exponential number of measurements we require in order to control the measurement noise leads to T-gate costs prohibitive within even the most optimistic predictions of fault tolerant computations. However, given that we have only been able to simulate systems of up to $N = 26$ and extrapolated to systems orders of magnitude larger, it is possible that the exponential scaling

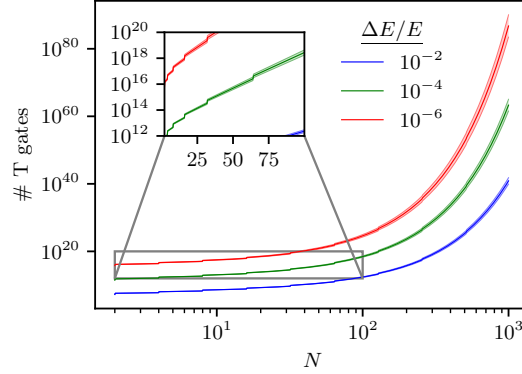


Figure 3.7: Total number of T-gates required for the whole PQSE procedure for the single flavour Schwinger model as a function of lattice size N for different fractional energy errors $\Delta E/E$.

we predict does not hold for larger systems. Further numerical tests, using tensor network methods to study larger system sizes and using Monte-Carlo methods to calculate the true ground state energy (which does not suffer from the sign problem for the single flavour system we consider) may shed more light on this. Whether the resource scaling would be more or less promising is unknown.

3.6 Discussion and conclusions

The aim of this chapter was to perform the first study of the application of the QSE algorithm to a LGT. In particular, we aimed to estimate the resources required to compute the ground state of the single flavour lattice Schwinger model for different system sizes and assess how feasible such a calculation would be on NISQ or early error-corrected quantum processors. Here we assess how successful we've been in achieving these aims.

By developing a novel LCU procedure for block encoding the Schwinger model Hamiltonian and analysing the gate cost required, in combination with numerical simulation, we successfully give a upper bound estimate for the full gate cost of the algorithm. This is the first such work performing a full resource analysis with exact gate costs looking at solving a LGT with a quantum algorithm.

For systems of a few hundred lattice sites, we see that the gate depth for a single circuit is beyond what we expect to be feasible on non-error correct devices in the near term. However, single circuit T-gate costs on the order of 10^4 to 10^6 , along with hundreds of logical qubits required make such circuits appealing targets for early fault tolerant devices.

While such gate depths for individual circuits can be seen as a positive step towards useful application of QSE to LGTs, the exponential measurement cost revealed by our numerical simulations indicates that further algorithmic improvements will be required in order to use such a workflow to compute accurate ground states of LGTs - even for reasonably large error corrected devices. While the exponential measurement cost may occur for other systems and QSE methods, by performing the first rigorous analysis with respect to system size scaling, we have been able to quantify and draw conclusions about the resource cost for different algorithmic approaches for the first time.

Since ground state problems for Fermionic systems tend to be QMA-complete [260–263], such an exponential scaling is not particularly surprising ⁷. Nor should it deter research into using QSE algorithms for LGTs but rather motivate further work aimed at reducing the algorithmic cost of such methods. Importantly, our work presents the first systematic study of the scaling of two QSE algorithms, namely thresholded and partitioned, with system size. We show that PQSE can significantly reduce the exponential scaling of the algorithm compared to the current state of the art (TQSE), serving as an important benchmark for future improvements and demonstrating the necessity to further develop noise-robust subspace methods. One particularly promising method for improving the noise resilience of QSE has been the proposal of the Gaussian power basis [124]. Such a basis could be implemented on a quantum processor using QSVT techniques and the block encoding we present

⁷Although the single flavour Schwinger model can be solved with Monte-Carlo methods that don't suffer from the sign problem. We anticipate quantum algorithms, like the one we present here, being applied to multi-flavour models which are known to suffer from the sign problem and therefore seem unlikely to have efficient computational solutions.

in this work. An investigation of this method, combined with more sophisticated simulation methods, may reduce the resource requirements we predict.

Finally, as part of our comprehensive resource analysis for ground state computation using QSE, we developed novel improvements to the LCU block encoding procedure with particular application to lattice gauge theory Hamiltonians. By exploiting the translational symmetry of the system, something not present in molecular Hamiltonians for example, we can reduce the cost of the LCU procedure when applied to the Schwinger model from $\tilde{\mathcal{O}}(N^2)$ to $\tilde{\mathcal{O}}(N)$. Our procedure could be used to implement other quantum algorithms including phase estimation and Hamiltonian simulation, both of which have seen significant improvements using qubitisation for quantum chemistry in an error corrected setting [38, 230], to study ground states or dynamics of LGTs.

4

Conclusions and outlook

Despite the rapid development of quantum computing hardware and theoretical understanding of quantum algorithms, some recent results have cast doubt on whether various quantum algorithms, as well as various application areas, will lead to useful quantum computation that outperform classical methods. These include large error-correction overheads which outweigh expected quadratic speedups within combinatorial optimisation [9, 264] and so called ‘dequantising’ results which disprove previously believed exponential speedups in machine learning [265–267]. While questions remain about the existence of general exponential speedups in the field [37], studying quantum systems remains a promising candidate for achieving quantum advantage. The aims of this thesis were to develop further applications and algorithms for ground state computation of quantum systems and to assess whether these would be applicable for use on near term devices. In doing so, we focused on systems containing Bosonic degrees of freedom which, in contrast to purely Fermionic systems (typically of molecular electrons), have seen relatively little research. Here we summarise the major contributions of this thesis and describe how these can be developed into further research.

Firstly, we presented the first proposal for how to model intermolecular interactions

on a quantum computer. This expanded the palette of problems in quantum chemistry that can be studied with quantum computers. Through our experiments we demonstrated that, when combined with known and novel error mitigation techniques, the quality of quantum hardware is good enough that such systems can already be studied using real quantum processors. These experiments serve as the first successful demonstration of using quantum computers to model dispersive interactions between two molecules. Since this work was performed, the number of qubits and error rates available on quantum hardware have significantly improved such that larger and more interesting systems can now be studied. By using a method based on VQE, our algorithm will likely also suffer from scalability issues related to barren plateaus. However for the Hamiltonian we consider, the effect of barren plateaus may be less severe than for ab-initio VQE for the following two reasons.

- Since we do not need to map a Fermionic Hilbert space onto to spins such as Bravyi-Kitaev or Jordan-Wigner transforms in which Fermionic ladder operators act on non-local sets qubits, our cost function is local. We may therefore avoid the increased susceptibility that non-local cost functions have to barren plateaus [77].
- Our theoretical analysis demonstrated that, since it contains only two body interaction terms; rather than four body terms present in ab-initio methods for molecular Hamiltonians, the number of shots required to measure expectation values of our Hamiltonian scales far better than for ab-initio methods. This reduction means that the measurement cost, highlighted in Table 1.1, may be far lower for our method, allowing us to resolve the cost function landscape with a far smaller number of shots.

Furthermore, methods aimed at reducing the effect of barren plateaus in variational algorithms for other Hamiltonians, such as adaptive methods [91], should also be applicable to our coarse grained model.

The second contribution of this thesis was the first proposal of using QSE methods to study a LGT extending our understanding of how quantum computers may be used for such systems. Furthermore, to the authors knowledge, this is the first work performing qubitisation for a LGT¹. We performed the most comprehensive analysis of the resources required to compute the ground state of a LGT on quantum computers. This analysis serves as a benchmark for understanding how feasible using such an algorithm is for near-term as well as error corrected devices. While we were able to reduce the resources required for the block encoding procedure significantly, our resource analysis revealed the significant dependence of the algorithm on shot noise. This highlights the need to further develop our understanding of and methods to avoid the issues of numerical stability within the QSE algorithm. While research has begun to understand these issues [120, 122, 124], such works have been very theoretical and have derived scalings that depend upon model problem parameters that are often not well understood and do not systematically study the effect of increasing system size. Our work plays an important role in quantifying how different methods to control errors and numerical instability perform for different system sizes.

Finally, as part of each piece of work we developed an understanding of how to represent truncated Bosonic Hamiltonians on quantum computers using a binary encoding. We developed the theory of how to implement measurements of Bosonic operators within a VQE setting, providing for the first time rigorous upper bounds for the number of individual circuits and measurements that need to be performed. Within an error corrected setting, we performed the first theoretical analysis of how you can perform qubitisation and apply QSVT methods to Boson-Fermion systems for a LGT. Compared to a basic LCU strategy, we were able to considerably reduce the gate cost to do so. As well as helping to understand and reduce the resource costs for the algorithms and Hamiltonians studied in this thesis. Our analysis can

¹A recent work [268] has emerged performing ‘qubitisation’ of LGTs for quantum computers. The word ‘qubitisation’ is used in a different context here to describe the truncation of a Bosonic Hilbert space to finite dimension so that they can be mapped onto a finite number of qubits, rather than the method based on block-encoding the Hamiltonian within a larger effective qubit space [255].

be readily applied to other variational algorithms or other fault tolerant algorithms such as phase estimation and Hamiltonian simulation based on QSVT methods. Furthermore, it can be applied to a range of other interesting and important systems containing Bosonic degrees of freedom. This includes the Hubbard Holstein model for studying photosynthesis [269], related models used to study electron-phonon coupling in candidates for high-temperature superconductivity [270], models for molecular vibrations [60], other gauge theories [271], and even algorithms for tensor principle component analysis within machine learning in which the problem is mapped onto a Bosonic subspace [23].

Appendices

A

In depth analysis for digital quantum simulation of Bosonic systems

A.1 London dispersion in linear molecules and one-dimensional quantum Drude oscillators

This appendix shows how the interaction between two axially symmetric linear molecules can be captured by one-dimensional QDOs. The results and derivations given here can be found in the literature but are combined and included here for the reader.

In the framework developed by London [272], the interaction between two anisotropic linear molecules that are axially symmetric about their respective principle intramolecular bond can be described by ellipsoids of charge. For a pair of identical linear molecules that share components of the dipole polarisability $\alpha_{\parallel}$ ($\alpha_{\perp}$) and characteristic angular frequencies $\omega_{\parallel}$ ($\omega_{\perp}$) parallel (perpendicular) to their intramolecular axis, the London Dispersion energy at large separation R is

$$\Delta E = -\frac{1}{R^6} \left[(C_{\parallel,\parallel} + C_{\perp,\perp} - C_{\parallel,\perp}) \left\{ \sin^2(\theta_a) \sin^2(\theta_b) \cos(\phi_a - \phi_b) - 2 \cos(\theta_a) \cos(\theta_b) \right\}^2 \right. \\ \left. + 3(C_{\parallel,\perp} - C_{\perp,\perp}) \left\{ \cos^2(\theta_a) + \cos^2(\theta_b) \right\} + 2(C_{\parallel,\perp} + 4C_{\perp,\perp}) \right], \quad (\text{A.1})$$

where $\theta_a, \phi_a, \theta_b, \phi_b$ describe the orientation of molecules a and b relative to the intramolecular axis. $C_{\cdot,\cdot}$ coefficients are given by

$$C_{\parallel,\parallel} = \frac{1}{8} \alpha_{\parallel}^2 \hbar \omega_{\parallel}, \quad C_{\perp,\perp} = \frac{1}{8} \alpha_{\perp}^2 \hbar \omega_{\perp}, \quad C_{\parallel,\perp} = \frac{1}{4} \alpha_{\parallel} \alpha_{\perp} \hbar \frac{\omega_{\parallel} \omega_{\perp}}{\omega_{\parallel} + \omega_{\perp}}. \quad (\text{A.2})$$

For further details about the geometry and derivation see Ref. [273]. When the two molecules are aligned parallel to each other and the intermolecular axis $\theta_a = \theta_b = 0$ and $\phi_a = \phi_b$ (which can be set to 0 without loss of generality) in the completely anisotropic limit $\alpha_{\parallel} \gg \alpha_{\perp}$ the London dispersion interaction is given by

$$\Delta E = -\frac{\alpha_{\parallel}^2 \hbar \omega_{\parallel}}{2R^6} \quad (\text{A.3})$$

We see that this matches the leading order dispersive correction for the 1D QDO model in Eqn. (2.5).

A.2 Truncated Fock space

For our one-dimensional QDO system, we want to be sure that using a finite dimensional Fock space for each QDO still gives a reasonable approximation to the ground state energy of the true infinite-dimensional system. Fig. A.1 shows how the ground state energy of the system deviates for different values of the local QDO dimension d . We can see that even for $d = 4$, the ground state energy is approximated well until very large couplings γ . For reference, $\gamma = 2$ corresponds to a Hamiltonian that is no longer positive-semi definite, at this point the charged Drude particles dissociate from the nuclei and the system breaks down. For physically relevant systems, we expect to require value of γ much lower than 2 and we can therefore limit ourselves to small d . For the real device experiments in the main text, the smallest separation R considered corresponds to the largest coupling value of $\gamma = 1.55$.

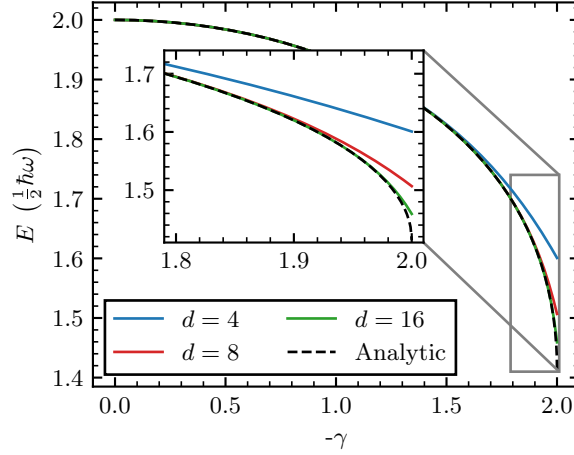


Figure A.1: Ground state energy of one-dimensional QDO pair as a function of coupling constant γ for various values of d ; the dimensionality of the truncated Fock space for each QDO. Analytic refers to exact analytic calculation of the ground state.

A.3 Decomposition and measurement of Hamiltonian

A.3.1 Scaling in the total number of Pauli operators

Theorem A.3.1. *For a d -level harmonic QDO, there are $\left(\frac{1}{2}d \log_2 d\right)^2$ non-zero Pauli terms in the binary encoding of the $\hat{x} \otimes \hat{x}$ interaction.*

Proof. The following single qubit operators can be represented in the Pauli basis

$$\begin{aligned} |0\rangle\langle 0| &= \frac{1}{2}(\mathbb{1} + Z), & |1\rangle\langle 1| &= \frac{1}{2}(\mathbb{1} - Z), \\ |1\rangle\langle 0| &= \frac{1}{2}(X - iY), & |0\rangle\langle 1| &= \frac{1}{2}(X + iY). \end{aligned} \quad (\text{A.4})$$

We first consider a $\hat{x} = \frac{1}{\sqrt{2}}(\hat{a}^\dagger + \hat{a})$ for a single QDO, where the ladder operators $\hat{a}^\dagger$ and $\hat{a}$ for the d -level Fock space are given by

$$\begin{aligned} \hat{a}^\dagger &= \sum_{n=0}^{d-1} \sqrt{n+1} |\underline{n+1}\rangle \langle \underline{n}|, \\ \hat{a} &= \sum_{n=0}^{d-1} \sqrt{n+1} |\underline{n}\rangle \langle \underline{n+1}|, \end{aligned} \quad (\text{A.5})$$

which follow the usual bosonic commutation relations. We begin by considering a single $|\underline{n+1}\rangle \langle \underline{n}|$ which given in terms of k single qubit operators is

$$|\underline{n+1}\rangle \langle \underline{n}| = \left(\bigotimes_{i=k+1}^{m-1} |\text{bin}(n)_i\rangle \langle \text{bin}(n)_i| \right) \otimes |1\rangle \langle 0| \otimes (|0\rangle \langle 1|)^{\otimes k}, \quad (\text{A.6})$$

where $\text{bin}(n)_i \in \{0, 1\}$ is the binary digit at position i of the binary representation of n and $k(n) = \max\{i \mid \text{bin}(n)_i = 0\}$ is the position of the right-most 0. Using the definitions in (A.4) this is equal to

$$|\underline{n+1}\rangle \langle \underline{n}| = \left(\bigotimes_{i=k+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n)_i} Z) \right) \otimes (X - iY) \otimes (X + iY)^{\otimes k}. \quad (\text{A.7})$$

Expanding (A.7) will give 2^m unique Pauli terms of the form $\hat{P} = c\sigma_i^{\otimes m}$ where $\sigma_i^{\otimes m} \in \{\mathbb{1}, X, Y, Z\}^m$ and $c \in \{\pm 1, \pm i\}$. Each $\hat{P}$ will either be hermitian or anti-hermitian, depending on whether $c = \pm 1$ or $c = \pm i$.

For each term in $\hat{x}$ like $(|\underline{n+1}\rangle \langle \underline{n}| + |\underline{n}\rangle \langle \underline{n+1}|)$, we note that $|\underline{n+1}\rangle \langle \underline{n}| = (|\underline{n}\rangle \langle \underline{n+1}|)^\dagger$ and their sum is Hermitian. Therefore the Pauli terms within this summed term are simply two times the Hermitian terms in the expansion of (A.7) (anti-Hermitian terms cancel). Of the 2^m different $\hat{P}$ within (A.7), half will be Hermitian ($c = \pm 1$) and half will be anti-Hermitian ($c = \pm i$). Therefore, there are a total of $\frac{1}{2}2^m$ unique Pauli operators in the compact representation of $(|\underline{n+1}\rangle \langle \underline{n}| + |\underline{n}\rangle \langle \underline{n+1}|)$.

Finally, in the summation

$$\hat{x} = \sum_{n=0}^{d-1} \sqrt{n+1} (|\underline{n+1}\rangle \langle \underline{n}| + |\underline{n}\rangle \langle \underline{n+1}|), \quad (\text{A.8})$$

there are $\log_2 d = m$ unique values of $k(n)$, each of which will contribute $\frac{2^m}{2}$ unique Pauli operators to the binary encoding of $\hat{x}$. Thus there are a total of $\left(\frac{1}{2}d \log_2 d\right)^2$ Pauli operators within $\hat{x} \otimes \hat{x}$, where we have used $m = \log_2 d$. $\square$

A.3.2 Scaling of number of circuits required

Lemma A.3.2. *The $\left(\frac{1}{2}d \log_2 d\right)^2$ Pauli operators for the $\hat{x} \otimes \hat{x}$ interaction in the binary encoding can be measured using $(d-1)^2$ circuits.*

Proof. We define the cover of an M qubit Pauli operator $\hat{P} \in \{\mathbb{1}, X, Y, Z\}^M$ as

$$\text{Cover}(\hat{P}) = \{\hat{B} \in \{\mathbb{1}, X, Y, Z\}^{\otimes M} \mid \hat{B}^{(i)} = \hat{P}^{(i)} \text{ when } \hat{P}^{(i)} \neq \mathbb{1}\}, \quad (\text{A.9})$$

where $\hat{P}^{(i)}$ is Pauli operator acting on the i^{th} qubit. For a collection of Pauli operators $\{\hat{P}_1, \hat{P}_2, \hat{P}_3, \dots\}$ that qubit-wise commute, there exists at least one basis $\hat{B}$ which covers all of the operators:

$$\hat{B} \in \text{Cover}(\hat{P}_1) \cup \text{Cover}(\hat{P}_2) \cup \text{Cover}(\hat{P}_3) \cup \dots \quad (\text{A.10})$$

The expectation values of every operator in $\{\hat{P}_1, \hat{P}_2, \hat{P}_3, \dots\}$ can be measured using a single circuit in which qubit i is measured in the $\hat{B}^{(i)}$ basis, where the ± 1 measurement value is used if $\hat{P}^{(i)} \neq \mathbb{1}$, and discarded otherwise. Here we called bases independent if they do not qubit wise commute.

We show that there are $(d-1)$ independent covers for terms in $\hat{x}$ and therefore $(d-1)^2$ covers needed to measure $\hat{x} \otimes \hat{x}$.

We begin by considering the expansion of $|\underline{n+1}\rangle \langle \underline{n}|$ given in (A.7). Expanding out the $\bigotimes_{i=k+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n)_i} Z)$ term on the first $(m-k-1)$ qubits will give tensor products of only $\mathbb{1}$ and Z . The covers for this $2^{(m-k-1)}$ operator subspace will always contain the basis $\bigotimes_{i=k+1}^{m-1} Z^{(i)}$, which can be used to measure them all. Then, expanding the $(X + iY) \otimes (X - iY)^{\otimes k}$ term on the remaining $(m-k+1)$ qubits, gives $2^{(k+1)}$ terms that do not qubit-wise commute and so each correspond to a unique basis.

Thus, there are $2^{(k+1)}$ independent bases for the terms in $|\underline{n+1}\rangle \langle \underline{n}|$. Using the argument that $|\underline{n+1}\rangle \langle \underline{n}| = |\underline{n}\rangle \langle \underline{n+1}|^\dagger$ as in Theorem A.3.6, only the Hermitian terms remain when summing the two. This halves the number of unique covers and so 2^k bases are required for $|\underline{n+1}\rangle \langle \underline{n}| + |\underline{n}\rangle \langle \underline{n+1}|$.

Again, for the sum of these terms given in (A.8), there are m unique values for $k(n) \in \{0, 1, \dots, m-1\}$ when written in the form of (A.4). For each unique value of $k(n)$, the value of $\text{bin}(n)$ does not affect the coverings required [since the first $(m-k-1)$ qubits are always covered with $Z^{\otimes(m-k-1)}$], so each unique $k(n)$

contributes a unique set of 2^k coverings to the summation. Therefore, the total number of coverings needed for the Pauli terms in $\hat{x}$ are $\sum_{k=0}^{m-1} 2^k = 2^m - 1$, where we have used a standard result for the summation.

Finally, the number of unique coverings for the $\hat{x} \otimes \hat{x}$ interaction is then $(2^m - 1)^2 = (d - 1)^2$. $\square$

Definition A.3.3 (Hypergraphs [274]). *A hypergraph is a generalisation of a graph in which edges can connect arbitrary number (i.e., not just a pair) of vertices. Formally, a hypergraph $G = (V, E)$ consists of V a finite set of vertices and E a set of non-empty subsets of V called hyperedges.*

Definition A.3.4 (Complete Hypergraphs [274]). *For $0 \leq r \leq N$, a complete hypergraph $K_N^r = (V, E)$ is a hypergraph with vertices V and set of hyperedges E such that $|V| = N$ and E contains all r -subsets of V .*

Lemma A.3.5 (Baranyai's Theorem [275]). *For a complete hypergraph K_N^r where $2 \leq r < N$ and r divides N , the N vertices can be partitioned into subsets (called 1-factorisation) of r vertices in $\binom{N-1}{r-1}$ different ways, such that each r -element subset (hyper edge) appears in exactly one partition.*

Proof. See [275, Chapter 38]. $\square$

For example, the complete hypergraph K_4^2 contains four vertices, each connected by two-element hyperedges (i.e, traditional edges). Baranyai's theorem states that the vertices can be partitioned into $\binom{4-1}{2-1} = 3$ partitions. K_4^2 and its corresponding partitions are shown in Fig. A.2.

Theorem A.3.6. *The number of circuits required for measurement of the Hamiltonian of the form*

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (\text{A.11})$$

for $N \geq 2$ d -level QDOs in the binary encoding is upper bounded by

$$\begin{aligned} (d-1)^2(N-1) + 1; & \quad N \text{ even} \\ (d-1)^2N + 1; & \quad N \text{ odd,} \end{aligned} \quad (\text{A.12})$$

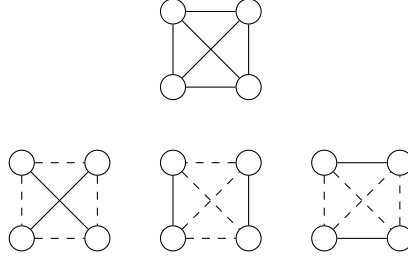


Figure A.2: Fully connected graph K_4^2 along with the three corresponding partitions of edges.

which is $\mathcal{O}(d^2 N)$.

Proof. All $(\hat{x}^2 + p^2) = \sum_{n=0}^{d-1} n |\underline{n}\rangle \langle \underline{n}|$ terms within the Hamiltonian can be decomposed into Pauli operators given in (2.10) of the main text. All of these terms commute within each QDO register as well between QDO registers and can thus be measured using the single covering $Z^{\otimes Nm}$.

For the pair-wise coupling terms, from Theorem A.3.2, a single $\hat{x} \otimes \hat{x}$ term between two QDOs can be measured with $(d-1)^2$ circuits. The bases required to measure pairwise coupling terms between disjoint pairs of QDOs are guaranteed to commute since they act on different qubit registers. Therefore, by measuring each of the bases concurrently for collections of mutually disjoint interactions, the number of measurement bases can be reduced in a predictable way.

We represent interaction terms as a graph in which each vertex corresponds to one of the N QDOs and edges between vertices correspond represent a non-zero pairwise coupling. For all-to-all interactions this will be a fully connected graph. For even N the number of partitions needed to separate interaction edges into disjoint sets is equal by Baranyai's theorem (Lemma A.3.5) for the graph K_N^2 . For odd N the edges cannot be partitioned exactly - there will always be one unpaired vertex when collecting disjoint edges. Considering a dummy vertex where an edge between that and an QDO vertex gives no contribution to the Hamiltonian, the number of partitions is given by Baranyai's theorem for K_{N+1}^2 .

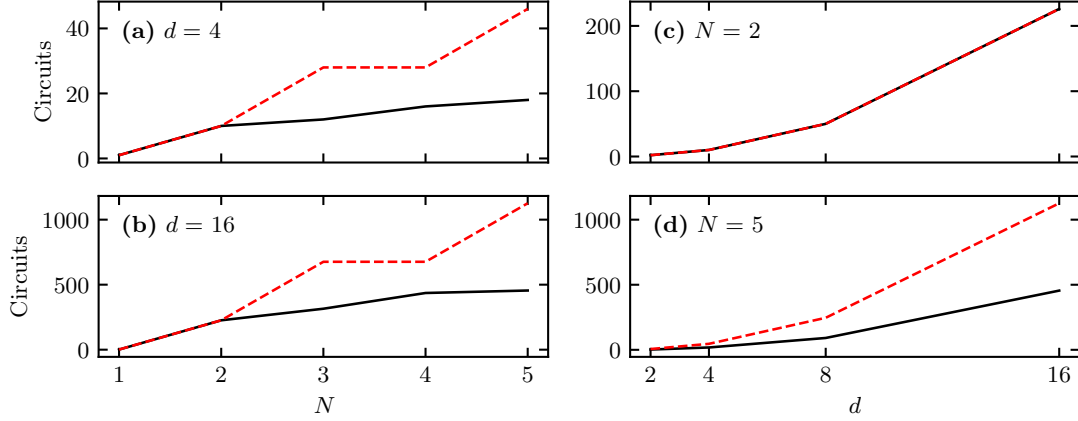


Figure A.3: Number of circuits required to measure Hamiltonian as **(a)** and **(b)**: a function of N for two fixed values of d as well as **(c)** and **(d)**: a function of d for two fixed values of N . Red dashed line: upper bound scaling given by (A.12), black solid line: using largest-degree first colouring for grouping Paulis. In **(c)**, the two lines coincide. Trivially, the number of circuits for $N = 1$ is always 1.

Using these collections, the number of bases that will need to be measured of the pairwise interaction terms is

$$\begin{aligned}
 (d-1)^2 \binom{N-1}{1} &= (d-1)^2 (N-1); & N \text{ even,} \\
 (d-1)^2 \binom{N}{1} &= (d-1)^2 N; & N \text{ odd.}
 \end{aligned}
 \tag{A.13}$$

Combining this with the single basis needed to measure the non-interaction term gives an upper bound for the number of bases that must be measured and thus the result. $\square$

This upper bound is plotted as a function of N and d in Fig. A.3 as well as the number of circuits required when using a heuristic sorting algorithm, namely largest-degree first colouring [179]. We see that for $N = 1$ both curves agree as there are no interaction terms and that the heuristic sorted saturates the upper bound for $N = 2$ QDOs. For $N > 2$ we see that the sorting significantly reduces the number of circuits below the upper bound due to additional qubit wise commutation between non-disjoint interaction terms.

Corollary A.3.7. *The number of circuits required for measurement of the Hamiltonian for $N \geq r$ d -level QDOs with r -order coupling in the binary encoding is $\mathcal{O}(d^2 N^{r-1})$.*

Proof. The proof is the same as that for Theorem A.3.6 except using $r > 2$ for the $\hat{x}^{\otimes r}$ interaction in Lemma A.3.2 and the corresponding hypergraph K_N^r in Lemma A.3.5. This gives an upper bound on the total number of groups of

$$(d-1)^r \binom{N-1}{r-1} + 1, \quad (\text{A.14})$$

assuming r divides N , which is $\mathcal{O}(d^r N^{(r-1)})$. If r does not divide N , the $\mathcal{O}(d^r N^{(r-1)})$ scaling will remain as dummy vertices can be used to give the number of vertices that is divisible for r .

□

A.3.3 Number of circuits for three-dimensional QDOs

Theorem A.3.8. *The number of circuits required for measurement of the Hamiltonian of the form*

$$H = \sum_{i=1}^N \left(\frac{\hat{\mathbf{p}}_i^2}{2\mu} + 2\mu\omega\hat{\mathbf{x}}_i^2 \right) + q^2 \sum_{j>i} \frac{3(\hat{\mathbf{x}}_i \cdot \mathbf{R}_{i,j})(\hat{\mathbf{x}}_j \cdot \mathbf{R}_{i,j}) - \hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_j R_{i,j}^2}{R_{i,j}^5}, \quad (\text{A.15})$$

for N three-dimensional QDOs with d levels per spatial dimension in the binary encoding is upper bounded by

$$\begin{aligned} (d-1)^2(N-1) + 1; & \quad N \text{ even,} \\ (d-1)^2N + 1; & \quad N \text{ odd.} \end{aligned} \quad (\text{A.16})$$

Proof. The uncoupled terms within (A.15) will correspond to operators $\sum_{n=0}^{2^m-1} n |\underline{n}\rangle \langle \underline{n}|$ for each of the three spatial dimensions for each QDO. As before, these require just Z measurements on each of the qubits and can thus be measured using a single circuit.

Written in terms of operators for each one-dimensional component, the interaction part of (A.15) will contain all nine terms of the form $\{\hat{x}_i, \hat{y}_i, \hat{z}_i\} \otimes \{\hat{x}_j, \hat{y}_j, \hat{z}_j\}$, where across each $3m$ qubit register for a single QDO, the operators making up $\hat{\mathbf{x}}_i$ are given by $\hat{x}_i = \hat{x} \otimes \mathbb{1}^{\otimes m} \otimes \mathbb{1}^{\otimes m}$, $\hat{y}_i = \mathbb{1}^{\otimes m} \otimes \hat{x} \otimes \mathbb{1}^{\otimes m}$ and $\hat{z}_i = \mathbb{1}^{\otimes m} \otimes \mathbb{1}^{\otimes m} \otimes \hat{x}$, following a similar pattern for $\hat{\mathbf{x}}_j$, $\hat{\mathbf{p}}_i$ and $\hat{\mathbf{p}}_j$. Since every single term of the form $\hat{x}\hat{x}$, $\hat{x}\hat{y}$, $\hat{y}\hat{z}$ etc. in the coupling interactions will consist of the same decomposition of Pauli operators either acting on the same or different QDO subregisters. These all qubit wise commute and each term each of the $(\frac{d}{2} \log_2 d)^2$ Pauli terms for all oscillator-oscillator interaction terms can be measured together. Thus for each pair of coupled three-dimensional QDOs, only $(\frac{d}{2} \log_2 d)^2$ circuits are needed; the same number required for a pairwise interaction between one-dimensional QDOs.

Finally, using Lemma A.3.5 and following the same logic as Theorem A.3.6, an upper bound for the number of circuits needed to measure in (A.15) is $(N-1) \left(\frac{d}{2} \log_2 d\right)^2 + 1$ or $N \left(\frac{d}{2} \log_2 d\right)^2 + 1$, for even and odd N respectively. $\square$

A.4 Shot requirements

In this section, we derive an upper bound for an estimate of the number of shots required to achieve a given statistical error on the measurement of $\langle \hat{H} \rangle$ for a system of 1D coupled QDOs using the grouping of measurements described in Theorem A.3.6.

We write a Hamiltonian in the form

$$\hat{H} = \sum_{g \in \mathcal{G}} \sum_{i \in g} a_i \hat{P}_i, \quad (\text{A.17})$$

where $\mathcal{G} = \{g | [\hat{P}_i, \hat{P}_j] = 0, \forall i, j \in g\}$ is the set of all collections for the Pauli operators $\hat{P}_i$ in which all operators in a collection qubit wise commute. A standard method [99] using Lagrange multipliers to optimally distribute measurements between circuits gives a total number experimental shots of

$$S = \frac{1}{\epsilon^2} \left[\sum_{g \in \mathcal{G}} \sqrt{\sum_{i \in g} |a_i|^2 \text{Var } \hat{P}_i} \right]^2 \quad (\text{A.18})$$

to achieve an expected measurement error of ϵ , assuming that $\text{Cov}[\hat{P}_i, \hat{P}_j] = 0$ if $i \neq j$. The variance of a Pauli operator is given by $\text{Var } \hat{P}_i = 1 - \langle \hat{P}_i \rangle$. In the absence of knowledge about what state the expectation value is taken with respect to, we take the expectation of this variance with respect to a spherical measure (See [183, Chapter 7] and following a similar method to Ref. [72]):

$$\mathbb{E}[\text{Var } \hat{P}_i] = \begin{cases} 0; & \hat{P}_i = \mathbb{1} \\ 1 - \frac{1}{2^{n_i+1}}; & \text{otherwise} \end{cases} \quad (\text{A.19})$$

where n_i is the number of non-identity Pauli operators in $\hat{P}_i$. Additionally, in the uniform spherical measure, $\mathbb{E}[\text{Cov}(\hat{P}_i, \hat{P}_j)] = 0$ if $i \neq j$. Thus the expected number of shots S to achieve error ϵ is upper bounded by

$$\mathbb{E}[S] = \frac{1}{\epsilon^2} \left[\sum_{g \in \mathcal{G}} \sqrt{\sum_{i \in g} |a_i|^2 \left(1 - \frac{1}{2^{n_i+1}}\right)} \right]^2. \quad (\text{A.20})$$

Theorem A.4.1 (Expected shots over uniform measure when grouping measurements). *Given a Hamiltonian*

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (\text{A.21})$$

for $N \geq 2$ d -level QDOs in the binary encoding and measuring commuting operators according to the procedure described in Theorem A.3.6, the expected number of shots S required to achieve a constant measurement error ϵ is upper bounded by

$$S \leq \frac{1}{\epsilon^2} \left[\frac{\sqrt{2}}{3} \sqrt{N} \sqrt{d^2 - 1} + \left(\sum_{F \in \text{Part}(G)} \sqrt{\sum_{\langle i,j \rangle \in F} \gamma_{i,j}^2} \right) \times \left(2d - [4 + 2 \log_2 d] d^2 + \left[2 + 2 \log_2 d + \frac{1}{2} (\log_2 d)^2 \right] d^3 \right) \right]^2, \quad (\text{A.22})$$

when taken over the spherically symmetric distribution of states, where $\text{Part}(G)$ is the set of coverings for a given 1-factorisation of the interaction graph G . Assuming G is fully connected and that values of $\gamma_{i,j}$ are independent of both N and d and randomly distributed between different $F \in \text{Part}(G)$, then $S \in \mathcal{O}\left(\frac{\gamma^2 N^3 (\log_2 d)^4 d^6}{\epsilon^2}\right)$, where γ is a typical scale for $\gamma_{i,j}$.

Proof. We consider the groupings described in Theorem A.3.6. Since interacting and non-interacting parts of the Hamiltonian have different forms and groupings, we consider each one's contribution to the total shots separately:

$$\epsilon^2 S = \left(\epsilon \tilde{S}_{\text{non-int}} + \epsilon \tilde{S}_{\text{int}} \right)^2. \quad (\text{A.23})$$

Starting with the uncoupled part of the Hamiltonian:

$$\sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) = \sum_{i=1}^N 2^m \mathbb{1}^{\otimes mN} - \sum_{j=0}^{m-1} 2^j Z^{(im+j)}, \quad (\text{A.24})$$

where $Z^{(im+j)}$ is the Z operator acting on the j^{th} qubit of QDO i , all terms form a single collection which gives

$$\epsilon \tilde{S}_{\text{non-int}} = \sqrt{N \sum_{j=0}^{m-1} 2^{2j} \text{Var } Z} = \sqrt{\frac{2}{9} N (2^{2m} - 1)}, \quad (\text{A.25})$$

where we have used that $\mathbb{E}[\text{Var } Z] = \frac{2}{3}$ from (A.19).

For the interacting terms, the collection to which a single Pauli operators given by (A.7) (i.e., which $g \in \mathcal{G}$ it is sorted into) is determined by three things:

1. Which of the $(N-1)$ (even N) or N (odd N) 1-factorisations of the interaction graph the corresponding interaction term belongs to.
2. The value of $k(n)$, the position of the rightmost zero in the binary representation of integer n labelling the Fock state, for *each of the two* the QDOs involved in an interaction term.
3. For a given pair of $k(n)$, which of the possible 2^{k+1} of X and Y operators is applied to the right most $(k+1)$ qubits.

As before, the anti-Hermitian terms in (A.7) are not needed in the sum $|\underline{n+1}\rangle \langle \underline{n}| + |\underline{n}\rangle \langle \underline{n+1}|$.

Separating the terms within $\hat{H}$ according to these criteria (their contributions to the equation labelled with curly braces), the expected number of shots required is

$$\epsilon_{\tilde{S}_{\text{int}}} = \frac{1}{2} \underbrace{\sum_{k_1, k_2=0}^{m-1}}_{1.} \underbrace{\sum_{F \in \text{Part}(G)}}_{2.} \underbrace{\sum_{\substack{\hat{P}_1^{(:k_1+1)} \in \{X,Y\}^{\otimes(k_1+1)} \\ \hat{P}_2^{(:k_2+1)} \in \{X,Y\}^{\otimes(k_2+1)}}}}_{3.} \Gamma(F, k_1, k_2), \quad (\text{A.26})$$

where $\text{Part}(G)$ is the set of the interaction graph G . From Lemma A.3.5, if G is fully connected then $|G| = (N-1)$ or N (for N even or odd). We have defined $\mathcal{N}_{k_1} = \{n | k(n) = k_1, n \in \mathbb{Z}, 0 \leq n < 2^m\}$ as the set of integers satisfying a given $k(n)$ (similarly for $\mathcal{N}_{k_2}$) and $\hat{P}_1^{(:k_1+1)}$ is a (k_1+1) qubit Pauli operator (similarly for $\hat{P}_2^{(:k_2+1)}$). The $1/2$ prefactor results from the cancellation of anti-hermitian terms which make up half of the operators. The value of Γ is given by

$$\Gamma(F, k_1, k_2) = \left[\sum_{\langle i,j \rangle \in F} \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} \sum_{\substack{\hat{P}_1^{(k_1+1:m)} \in \{\mathbf{1}, Z\}^{\otimes(m-k_1-1)} \\ \hat{P}_2^{(k_2+1:m)} \in \{\mathbf{1}, Z\}^{\otimes(m-k_2-1)}}} \right. \\ \left. \left(2\gamma_{i,j} \sqrt{n_1+1} \sqrt{n_2+1} \right)^2 \mathbb{E}[\text{Var } \hat{P}_1 \otimes \hat{P}_2] \right]^{\frac{1}{2}}, \quad (\text{A.27})$$

where, $\hat{P}_1^{(k_1+1:m)}$ is an $(m-(k_1+1))$ qubit Pauli operator and $\hat{P}_1 = \hat{P}_1^{(:k_1+1)} \otimes \hat{P}_1^{(k_1+1:m)}$ (similarly for $\hat{P}_2^{(k_2+1:m)}$). This is the equivalent of the square root term in (A.20); the contribution to the variance from a single collection of commuting operators. In the above, $\langle i, j \rangle$ correspond to edges within 1-factorisation F , n_1 and n_2 correspond to Fock terms satisfying the requirement of $k(n)$; the position of the right-most X or Y operator. Using the expectation on the spherical measure from (A.19) gives

$$\mathbb{E}[\text{Var } \hat{P}_1 \otimes \hat{P}_2] = \left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta} + 1} \right) \quad (\text{A.28})$$

where integers $0 < \alpha_1 < m - k_1$ and $0 < \alpha_2 < m - k_2$ count the number of Z operators in $\hat{P}_1^{(k_1+1:m)}$ and $\hat{P}_2^{(k_2+1:m)}$ respectively.

By recognising that when using the spherical measure, many of the expected variances are equivalent and only depend on the total number of X,Y or Z operators, the expected number of shots simplifies to

$$\epsilon_{\tilde{S}_{\text{int}}} = \frac{1}{2} \sum_{k_1, k_2=0}^{m-1} \sum_{F \in \text{Part}(K_N^2)} 2^{k_1+1} 2^{k_2+1} \Gamma(F, k_1, k_2) \quad (\text{A.29})$$

with

$$\begin{aligned} \Gamma(F, k_1, k_2) = & \left[\sum_{\langle i,j \rangle \in F} \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} \sum_{\alpha=0}^{m-k_1-1} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_1-1}{\alpha} \binom{m-k_2-1}{\beta} \right. \\ & \left. \times \left(2\gamma_{i,j} \sqrt{n_1+1} \sqrt{n_2+1} \right)^2 \left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta}+1} \right) \right]^{\frac{1}{2}}. \end{aligned} \quad (\text{A.30})$$

Now we aim to find a simple upper bound for S . We start by grouping terms

$$\begin{aligned} \epsilon_{\tilde{S}_{\text{int}}} = & \left(\sum_{F \in \text{Part}(G)} \sqrt{\sum_{\langle i,j \rangle \in F} \gamma_{i,j}^2} \right) \sum_{k_1, k_2=1}^m \left[2^{k_1+1} 2^{k_2+1} \sqrt{\sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} (n_1+1)(n_2+1)} \right. \\ & \left. \sqrt{\sum_{\alpha=0}^{m-k_1-1} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_1-1}{\alpha} \binom{m-k_2-1}{\beta} \left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta}+1} \right)} \right]. \end{aligned} \quad (\text{A.31})$$

We note that

$$\begin{aligned} \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} (n_1+1)(n_2+1) &= \left(2^{k_1} \sum_{\gamma=0}^{2^{m-k_1}-1} (2\gamma+1) \right) \left(2^{k_2} \sum_{\delta=0}^{2^{m-k_2}-1} (2\delta+1) \right) \\ &= 2^{-k_1-2} (2^{k_1+1} + 2^m)^2 \cdot 2^{-k_2-2} (2^{k_2+1} + 2^m)^2 \end{aligned} \quad (\text{A.32})$$

and

$$\begin{aligned}
& \sum_{\alpha=0}^{m-k_1-1} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_1-1}{\alpha} \binom{m-k_2-1}{\beta} \left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta}+1}\right) \\
& \leq \sum_{\alpha=0}^{m-k_1-1} \binom{m-k_1-1}{\alpha} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_1-1}{\beta} \\
& = 2^{m-k_1-1} \cdot 2^{m-k_2-1},
\end{aligned} \tag{A.33}$$

where we have used a standard result for the summation of binomial coefficients.

Substituting (A.32) and (A.33) into (A.31) gives

$$\begin{aligned}
\epsilon S_{\text{int}} & \leq \left(\sum_{F \in \text{Part}(G)} \sqrt{\sum_{\langle i,j \rangle \in F} \gamma_{i,j}^2} \right) \sum_{k_1, k_2=0}^{m-1} 2^{m-1} (2^{k_1+k_2+2} + 2^{m+k_1+1} + 2^{m+k_2+1} + 2^{2m}) \\
& = \left(\sum_{F \in (G)} \sqrt{\sum_{\langle i,j \rangle \in F} \gamma_{i,j}^2} \right) \left(2d - [4 + 2 \log_2 d] d^2 + \left[2 + 2 \log d + \frac{1}{2} (\log_2 d)^2 \right] d^3 \right),
\end{aligned} \tag{A.34}$$

using standard results for the summations and substitution $d = 2^m$. Combining this with (A.25) gives the result. $\square$

Theorem A.4.2 (Expected shots over uniform measure when not grouping measurements.). *Given a Hamiltonian*

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \tag{A.35}$$

for $N \geq 2$ d -level QDOs in the binary encoding and measuring each corresponding Pauli operator separately, the expected number of shots required to achieve a constant measurement error ϵ is given by the upper bound

$$\begin{aligned}
S & \leq \frac{1}{\epsilon^2} \left[N \frac{\sqrt{6}}{3} (d-1) + \left(\sum_{\langle i,j \rangle \in G} |\gamma_{i,j}| \right) \right. \\
& \quad \left. \left([3 + 2\sqrt{2}] d^2 - [2 + \sqrt{2}] d^{\frac{5}{2}} - \left[\frac{7}{2} + 3\sqrt{2} \right] d^3 + [1 + \sqrt{2}] d^{\frac{7}{2}} + \left[\frac{3}{2} + \sqrt{2} \right] d^4 \right) \right]^2,
\end{aligned} \tag{A.36}$$

when taken over the spherically symmetric distribution of states, where G is the interaction graph. Assuming G is fully connected and, that values of $\gamma_{i,j}$ are independent of both N and d , then $S \in \mathcal{O}\left(\frac{\gamma^2 N^4 d^8}{\epsilon^2}\right)$ where γ is a typical scale for $\gamma_{i,j}$.

Proof. As in Theorem A.4.1, we consider contributions from non-interaction and interaction terms separately. Starting with the uncoupled part of the Hamiltonian

$$\epsilon \tilde{S}_{\text{non-int}} = N \sum_{j=0}^{m-1} \sqrt{2^{2j} \text{Var } Z} = N \frac{\sqrt{6}}{3} (2^m - 1). \quad (\text{A.37})$$

For the interaction part of the Hamiltonian

$$\begin{aligned} \epsilon \tilde{S}_{\text{int}} &= \frac{1}{2} \sum_{\langle i,j \rangle \in G} \sum_{k_1, k_2=0}^{m-1} \sum_{\substack{\hat{P}_1 \in \{X,Y\}^{\otimes k_1+1} \\ \otimes \{1,Z\}^{\otimes m-k_1-1} \\ \hat{P}_2 \in \{X,Y\}^{\otimes k_2+1} \\ \otimes \{1,Z\}^{\otimes m-k_2-1}}} \left[\sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} 4\gamma_{i,j}^2 (n_1 + 1)(n_2 + 1) \mathbb{E}(\text{Var } \hat{P}_1 \otimes \hat{P}_2) \right]^{\frac{1}{2}} \\ &= \sum_{\langle i,j \rangle \in G} |\gamma_{i,j}| \sum_{k_1, k_2=0}^{m-1} \left[2^{k_1+1} 2^{k_2+1} \sqrt{\sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} (n_1 + 1)(n_2 + 1)} \right. \\ &\quad \left. \sum_{\alpha=0}^{m-k_1-1} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_1-1}{\alpha} \binom{m-k_2-1}{\beta} \sqrt{\left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta} + 1}\right)} \right]. \end{aligned} \quad (\text{A.38})$$

Using (A.33) and (A.32) again and writing $d = 2^m$, we get the inequality

$$\begin{aligned} \epsilon \tilde{S}_{\text{int}} &\leq \left(\sum_{\langle i,j \rangle \in G} |\gamma_{i,j}| \right) \sum_{k_1, k_2=0}^{m-1} \left(2^{k_1+1} \cdot 2^{k_2+1} \right) \left(2^{m-k_1-1} \cdot 2^{m-k_2-1} \right) \\ &\quad \times \left(2^{-\frac{k_1}{2}-1} (2^{k_1+1} + 2^m) \cdot 2^{-\frac{k_2}{2}-1} (2^{k_2+1} + 2^m) \right) \\ &= \left(\sum_{\langle i,j \rangle \in G} |\gamma_{i,j}| \right) \left(\left[3 + 2\sqrt{2} \right] d^2 - \left[2 + \sqrt{2} \right] d^{\frac{5}{2}} - \left[\frac{7}{2} + 3\sqrt{2} \right] d^3 \right. \\ &\quad \left. + \left[1 + \sqrt{2} \right] d^{\frac{7}{2}} + \left[\frac{3}{2} + \sqrt{2} \right] d^4 \right). \end{aligned} \quad (\text{A.39})$$

Combining this with (A.37) gives the result. $\square$

The preceding results assume no knowledge about the states that are being measured and use an expectation over all of Hilbert space with respect to the spherically symmetric measure. Here we make a second estimate for the number of shots required using the pure state $|\psi_0\rangle = |0\rangle^{\otimes mN}$. $|\psi_0\rangle$ corresponds to the product state of all QDOs in their lowest Fock state and is assumed to be reasonably close to the states that we seek using the VQE procedure.

When not using the spherical measure, contributions to the measurement error from the covariance of terms within the same collection may be non-zero. In this case, the expected number of shots is given by

$$S = \frac{1}{\epsilon^2} \left[\sum_{g \in \mathcal{G}} \sqrt{\sum_{i,j \in g} a_i a_j \text{Cov}[\hat{P}_i, \hat{P}_j]} \right]^2. \quad (\text{A.40})$$

where $\text{Cov}[\hat{P}_i, \hat{P}_j] = \langle \hat{P}_i \hat{P}_j \rangle_{|\psi_0\rangle} - \langle \hat{P}_i \rangle_{|\psi_0\rangle} \langle \hat{P}_j \rangle_{|\psi_0\rangle}$ with subscripts indicating the state that expectation values are taken with respect to.

Theorem A.4.3 (Expected shots for uncoupled state when grouping measurements.). *Given a Hamiltonian*

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (\text{A.41})$$

for $N \geq 2$ d -level QDOs in the binary encoding and measuring commuting operators according to the procedure described in Theorem A.3.6, the expected number of shots S required to achieve a constant measurement error ϵ for state $|0\rangle^{\otimes mN}$ is upper bounded by

$$S = \frac{1}{16\epsilon^2} \left(\sum_{F \in \text{Part}(G)} \sqrt{\sum_{\langle i,j \rangle \in F} \gamma_{i,j}^2} \right)^2 d^4, \quad (\text{A.42})$$

where $\text{Part}(G)$ is the set of coverings for a given 1-factorisation of the interaction graph G . Assuming G is fully connected and that values of $\gamma_{i,j}$ are independent of both N and d and randomly distributed between different $F \in \text{Part}(G)$, then $S \in \mathcal{O}\left(\frac{\gamma^2 N^3 d^4}{\epsilon^2}\right)$, where γ is a typical scale for $\gamma_{i,j}$.

Proof. Again we consider non-interacting and interacting contributions separately. For the non interacting contribution, we note that for $|\psi\rangle = |0\rangle^{\otimes mN}$, $\langle Z^{(i)} \rangle = \langle Z^{(j)} \rangle = \langle Z^{(i)} Z^{(j)} \rangle = 1 \forall i, j$. Therefore,

$$\epsilon \tilde{S}_{\text{non-int}} = \sqrt{N \sum_{i,j=0}^{m-1} 2^i 2^j \text{Cov}[Z^{(i)}, Z^{(j)}]} = 0. \quad (\text{A.43})$$

The contribution from the coupling terms is

$$\epsilon \tilde{S}_{\text{int}} = \frac{1}{2} \sum_{k_1, k_2=0}^{m-1} \sum_{F \in \text{Part}(K_N^2)} \sum_{\substack{\hat{P}_1^{(:k_1)} \in \{X, Y\}^{\otimes (k_1+1)} \\ \hat{P}_2^{(:k_2)} \in \{X, Y\}^{\otimes (k_2+1)}}} \Gamma(F, k_1, k_2), \quad (\text{A.44})$$

where $\text{Part}(K_N^2)$ is the set of the $(N-1)$ or N (for N even or odd) partitions of the interaction graph K_N^2 , $\mathcal{N}_{k_1} = \{n | k(n) = k_1, n \in \mathbb{Z}, 0 \leq n < 2^m\}$ is the set of integers satisfying a given $k(n)$ (similarly for $\mathcal{N}_{k_2}$) and $\hat{P}_1^{(:k_1+1)}$ is a (k_1+1) qubit Pauli operator (similarly for $\hat{P}_2^{(:k_2+1)}$). Γ is given by

$$\begin{aligned} \Gamma(G, k_1, k_2) = & \left[\sum_{\langle i, j \rangle \in F} \sum_{\langle i', j' \rangle \in F} \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} \sum_{\substack{n'_1 \in \mathcal{N}_{k_1} \\ n'_2 \in \mathcal{N}_{k_2}}} \right. \\ & 2^2 \gamma_{i,j} \gamma_{i',j'} \sqrt{n_1+1} \sqrt{n_2+1} \sqrt{n'_1+1} \sqrt{n'_2+1} \\ & \left. \times \text{Cov} \left[\hat{A}_{\langle i, j \rangle}(k_1, k_2, n_1, n_2), \hat{A}_{\langle i', j' \rangle}(k_1, k_2, n'_1, n'_2) \right] \right]^{\frac{1}{2}} \end{aligned} \quad (\text{A.45})$$

where

$$\begin{aligned} \hat{A}_{\langle i, j \rangle}(k_1, k_2, n_1, n_2) = & \left(\hat{P}_1^{(:k_1)} \otimes \bigotimes_{\alpha=k_1+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n_1)_\alpha} Z) \right)_{\text{QDO } i} \\ & \otimes \left(\hat{P}_2^{(:k_2)} \otimes \bigotimes_{\beta=k_2+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n_2)_\beta} Z) \right)_{\text{QDO } j} \end{aligned} \quad (\text{A.46})$$

acts on the sub registers corresponding to QDOs i and j .

We note that $\hat{P}_1^{(:k_1)} \in \{\pm 1\} \times \{X, Y\}^{\otimes (k_1+1)}$ (recalling only Hermitian terms remain in $\hat{H}$) and so $\langle \hat{P}_1^{(:k_1)} \rangle_{|\psi_0\rangle} = 0 \forall k_1$ and similarly for $\hat{P}_2^{(:k_2)}$, from which it follows

$$\langle i, j \rangle \neq \langle i', j' \rangle \implies \text{Cov} \left[\hat{A}_{\langle i, j \rangle}(k_1, k_2, n_1, n_2), \hat{A}_{\langle i', j' \rangle}(k_1, k_2, n'_1, n'_2) \right] = 0 \quad (\text{A.47})$$

and

$$\begin{aligned}
& \text{Cov} \left[\hat{A}_{\langle i,j \rangle}(k_1, k_2, n_1, n_2), \hat{A}_{\langle i,j \rangle}(k_1, k_2, n'_1, n'_2) \right] \\
&= \left\langle \hat{A}_{\langle i,j \rangle}(k_1, k_2, n_1, n_2) \hat{A}_{\langle i,j \rangle}(k_1, k_2, n'_1, n'_2) \right\rangle_{|\psi_0\rangle} \\
&= \left\langle \hat{P}_1^{(:,k_1)} \hat{P}_1^{(:,k_1)} \right\rangle_{|\psi_0\rangle} \left\langle \bigotimes_{\alpha=k_1+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n_1)_\alpha} Z) (\mathbb{1} + (-1)^{\text{bin}(n'_1)_\alpha} Z) \right\rangle_{|\psi_0\rangle} \\
&\quad \left\langle \hat{P}_2^{(:,k_2)} \hat{P}_2^{(:,k_2)} \right\rangle_{|\psi_0\rangle} \left\langle \bigotimes_{\alpha=k_1+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n_2)_\alpha} Z) (\mathbb{1} + (-1)^{\text{bin}(n'_2)_\alpha} Z) \right\rangle_{|\psi_0\rangle}, \tag{A.48}
\end{aligned}$$

where all expectation values are with respect to $|\psi_0\rangle$. Trivially $\langle \hat{P}_1^{(:,k_1)} \hat{P}_1^{(:,k_1)} \rangle_{|\psi_0\rangle} = \langle \hat{P}_2^{(:,k_2)} \hat{P}_2^{(:,k_2)} \rangle_{|\psi_0\rangle} = 1$. It can be seen that

$$\left\langle \bigotimes_{\alpha=k_1+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n_1)_\alpha} Z) (\mathbb{1} + (-1)^{\text{bin}(n'_1)_\alpha} Z) \right\rangle_{|\psi_0\rangle} = \begin{cases} 4^{m-k_1-1}, & n_1 = n'_1 = 0, \\ 0, & \text{otherwise,} \end{cases} \tag{A.49}$$

and similarly for n_2 and n'_2 . Therefore

$$\begin{aligned}
& \text{Cov} \left[\hat{A}_{\langle i,j \rangle}(k_1, k_2, n_1, n_2), \hat{A}_{\langle i',j' \rangle}(k_1, k_2, n'_1, n'_2) \right] \\
&= 2^{4m-2k_1-2k_2-4} \delta_{i,i'} \delta_{j,j'} \delta_{n_1,0} \delta_{n_2,0} \delta_{n'_1,0} \delta_{n'_2,0}, \tag{A.50}
\end{aligned}$$

where δ is the Kronecker delta. We note that $\delta_{n_1,0} = \delta_{n_1,0} \delta_{k_1,0} \forall k_1 : n_1 \in \mathcal{N}_{k_1}$ and similarly for $\delta_{n_2,0}$. Then, substituting this into (A.45) gives

$$\Gamma(G, k_1, k_2) = \sum_{\langle i,j \rangle} 2^2 \gamma_{i,j}^2 2^{4m-k_1-k_2-4} \delta_{k_1,0} \delta_{k_2,0}, \tag{A.51}$$

which substituting into (A.44) gives

$$\epsilon \tilde{S}_{\text{int}} = \frac{1}{4} \sum_{F \in \text{Part } G} \sqrt{\sum_{\langle i,j \rangle \in F} \gamma_{i,j}^2} d^4, \tag{A.52}$$

from which the result follows. $\square$

Theorem A.4.4 (Expected shots for uncoupled state when not grouping measurements.). *Given a Hamiltonian*

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \tag{A.53}$$

for $N \geq 2$ d -level QDOs in the binary encoding and measuring each corresponding Pauli operator separately, the expected number of shots S required to achieve a constant measurement error ϵ for state $|0\rangle^{\otimes mN}$ is upper bounded by

$$S = \frac{1}{16\epsilon^2} \left(\sum_{\langle i,j \rangle \in G} |\gamma_{i,j}| \right)^2 d^4, \quad (\text{A.54})$$

where G is the interaction graph. Assuming G is fully connected and that values of $\gamma_{i,j}$ are independent of both N and d and randomly distributed between different $F \in \text{Part}(G)$, then $S \in \mathcal{O}\left(\frac{\gamma^2 N^4 d^4}{\epsilon^2}\right)$, where γ is a typical scale for $\gamma_{i,j}$.

Proof. Again we consider non-interacting and interacting contributions separately. As in theorem A.4.3, the uncoupled contribution is zero.

The contribution from the coupling terms is

$$\epsilon \tilde{S}_{\text{int}} = \frac{1}{2} \sum_{k_1, k_2=0}^{m-1} \sum_{\langle i,j \rangle \in G} \sum_{\substack{\hat{P}_1^{(:k_1)} \in \{X,Y\}^{\otimes(k_1+1)} \\ \hat{P}_2^{(:k_2)} \in \{X,Y\}^{\otimes(k_2+1)}}} \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} \sqrt{2^2 \gamma_{i,j}^2 \text{Var}[\hat{A}_{\langle i,j \rangle}(k_1, k_2, n_1, n_2)]} \quad (\text{A.55})$$

where

$$\begin{aligned} \hat{A}_{\langle i,j \rangle}(k_1, k_2, n_1, n_2) = & \left(\hat{P}_1^{(:k_1)} \otimes \bigotimes_{\alpha=k_1+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n_1)_\alpha} Z) \right)_{\text{QDO } i} \\ & \otimes \left(\hat{P}_2^{(:k_2)} \otimes \bigotimes_{\beta=k_2+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n_2)_\beta} Z) \right)_{\text{QDO } j} \end{aligned} \quad (\text{A.56})$$

acts on the sub registers corresponding to QDOs i and j . Using (A.50), this simplifies to

$$\epsilon \tilde{S}_{\text{int}} = \frac{1}{4} \left(\sum_{\langle i,j \rangle \in G} |\gamma_{i,j}| \right) d^2, \quad (\text{A.57})$$

from which the result follows. $\square$

B

Experimental details on variational quantum algorithms for coarse grained systems

B.1 Noise mitigation for anharmonic experiments

Unlike for the harmonic systems, for anharmonic systems it is not possible to perform the subtraction method given in (2.17). This is because we do not have a good way to estimate λ as the ground state of the zero field system is not simply given by the zero Fock state. Therefore, we use more costly error mitigation techniques in order to reduce the effect of device noise in the final energy measurements. We choose to use a combination of two error mitigation schemes, namely zero-noise extrapolation [197, 198] and an algorithmic error mitigation scheme based on Lanczos extrapolation [199].

To reduce the effect of noise on the measurement of $\langle \hat{H} \rangle$ using zero noise extrapolation, the expectation value is measured multiple times with the noise in the system artificially increased to various levels. The amount that the noise is amplified in the system is parameterised by a variable which we call Λ such that we have values $\{\langle H \rangle_\Lambda\}$ evaluated at each chosen value of Λ . A curve is then fitted to the values of

Λ and $\langle \hat{H} \rangle_\Lambda$ which can then be extrapolated back to $\Lambda = 0$ to hopefully remove the effect of noise on the system. Zero noise extrapolation requires no more qubits than the original system and can be performed fairly easily. The difficulty lies in finding a good method to amplify the noise to enough values of Λ such that a good fit can be found, as well as choosing a good fitting function for your device noise model.

In our experiments we boost the device noise by replacing each two qubit CNOT gate by an odd number of CNOT gates. Since in the noiseless setting an odd number of CNOT gates has the same effect as a single CNOT gate, this has the effect of artificially boosting the two-qubit gate noise - which is expected to be the main source of noise in the superconducting system. We use three different values of Λ which correspond to using one, three and five CNOTs in place of each CNOT in the circuit. A linear fit was found to match values of $\{\langle H \rangle_\Lambda\}$ well and used to extrapolate to the $\Lambda = 0$ to achieve a zero noise estimate.

For the further mitigation based on the Lanczos algorithm, we follow the method described in Ref. [199] to perform an $m = 2$ Krylov subspace expansion of the Hamiltonian. In the $m = 2$ case, this requires measuring $\langle H \rangle$, $\langle \hat{H}^2 \rangle$ and $\langle \hat{H}^3 \rangle$ on the noisy device to create the Krylov basis. Minimisation in this basis (performed by exact matrix diagonalisation) is then used to find a minimum value for $\langle \hat{H} \rangle$ which satisfies the variational principle.

The zero-noise extrapolation and Lanczos algorithm were combined as in Ref. [199] to give an increased level of noise reduction. First zero noise extrapolation was used on expectation values of $\langle \hat{H} \rangle$, $\langle \hat{H}^2 \rangle$ and $\langle \hat{H}^3 \rangle$ which were then used in the Lanczos mitigation method. As with measurements of the cost function $\langle H \rangle$ in the main text, the Pauli strings for each operator were grouped into collections that qubit-wise commuted and each group measured with approximately 16 million shots. Operators $\hat{H}$, $\hat{H}^2$ and $\hat{H}^3$ consist of 6, 7 and 8 Pauli strings which were grouped into 2, 3 and 3 commuting groups respectively.

B.2 IBM quantum hardware

For practical reasons related to device availability and queue-times we used three different quantum processors to evaluate our algorithm. All three processors were provided by IBM via the Qiskit [276] open source framework. We used *ibmq_montreal* (27 qubits, Processor type Falcon r4), *ibmq_lagos* (7 qubits, Processor type Falcon r5.11H) and *ibmq_santiago* (5 qubits, Processor type Falcon r4L). The reported quantum volume of the aforementioned devices is 128, 32 and 32 respectively.

The experiments in Sec. 2.3 use 4 qubits on every processor, namely qubits 13,14, 16 and 19 in *ibmq_montreal*, qubits 1,3,4 and 5 in *ibmq_lagos* and qubits 1,2,3 and 4 in *ibmq_santiago* following the coupling map show in Fig. B.1. The experiments were performed in a timeframe of 2 months and the noise parameters (and qubit properties) were varied along the period of the experiments. The experiments in Sec. 2.5.1 used qubits 14 and 16 on *ibmq_montreal* and were performed over a period of two days.

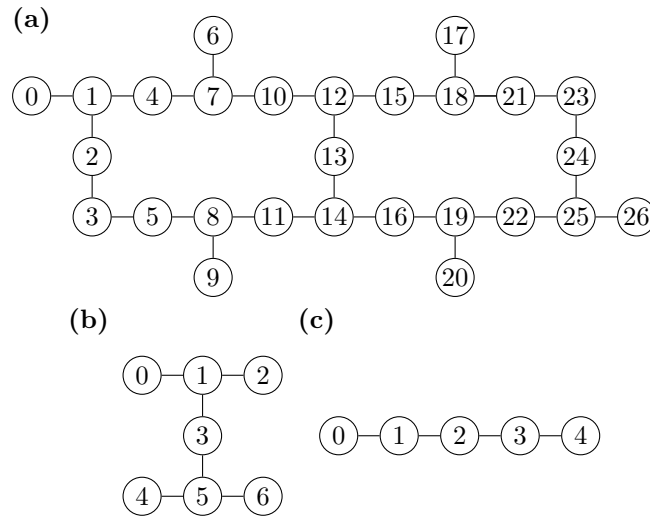


Figure B.1: Coupling map of the IBM Quantum devices used in this work. The devices used are (a) *ibmq_montreal*, (b) *ibmq_lagos* and (c) *ibmq_santiago* with 27, 7 and 5 qubits respectively.

B.3 Symplectic eigenvalues of harmonic oscillators with linear coupling

In the case of purely harmonic oscillators with linear coupling given by the Hamiltonian in Eqn. (2.6), the exact ground state energy (without any truncation of the Fock space) can be found by calculating the eigenspectrum of a $2N \times 2N$ matrix.

The Hamiltonian in (2.6) can be rewritten as

$$\hat{H} = \hat{\mathbf{q}}^T \mathbf{H} \hat{\mathbf{q}}, \quad (\text{B.1})$$

where we have defined the vector operator

$$\hat{\mathbf{q}}^T = (\hat{q}_1, \hat{q}_2, \dots, \hat{q}_{2N-1}, \hat{q}_{2N}) := (\hat{x}_1, \hat{p}_1, \dots, \hat{x}_N, \hat{p}_N) \quad (\text{B.2})$$

and correlation matrix $\mathbf{H} \in \mathbb{R}^{2N \times 2N}$ which has elements $H_{i,j}$ equal to the coefficients of the $\hat{q}_i \hat{q}_j$ operator in $\hat{H}$. In this representation, the bosonic commutation relation can be written as elements of a matrix $[\hat{q}_i, \hat{q}_j] = \Omega_{i,j}$ where

$$\boldsymbol{\Omega} = \bigoplus_{k=1}^N \boldsymbol{\omega} \in \mathbb{R}^{2N \times 2N}, \quad \boldsymbol{\omega} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (\text{B.3})$$

The eigenspectrum of the matrix $|i\boldsymbol{\Omega}\mathbf{H}|$ (where the absolute value is defined in the operatorial sense [277]) is $\{\epsilon_1, \epsilon_2, \dots, \epsilon_{2N}\}$, where $\{\epsilon_i\}$ are known as the symplectic eigenvalues. The ground state energy of the system is given by $E_0 = \frac{1}{2} \sum_{i=1}^{2N} \epsilon_i$. We consider systems of increasing coupling strength up to the point at which $\min_i \epsilon_i = 0$ and the system becomes critical. For further details see [277, 278].

We note that this method for calculating the ground state energy only works for Hamiltonians that can be written in the quadratic form of (B.1) for which ground states and thermal states are Gaussian.

C

Resource analysis for block encoding Schwinger model

C.1 Preparing $|G\rangle$ and U using linear combination of unitaries

In this section, we will recap the procedure that Kirby et al. described within Appendix C. of Ref. [120] to prepare U and $|G\rangle$ for a local Hamiltonian. Later we will modify this procedure to reduce the resource cost for our Hamiltonian.

C.1.1 Implementing U

We need to implement U given by (3.18), for some choice of $\{|i\rangle\}$. Let

$$|i\rangle_a = |\vec{x}, \vec{z}\rangle, \tag{C.1}$$

where $(\vec{x}, \vec{z})$ is the binary representation of n qubit Pauli operator $\hat{P}_i$, defined as follows. For now we will assume that all $\{\alpha_i\}$ are positive in the definition of the Hamiltonian (3.17) and that no $\hat{P}_i$ have a negative sign. Kirby et al. consider applying the required signs to the Pauli operator for their Hamiltonian separately, we will show how to do this for an arbitrary model later.

Definition C.1.1 (Binary representation of Pauli operator). *Given a Pauli operator $\hat{P} \in \{\mathbb{1}, X, Y, Z\}^{\otimes n}$, the binary representation of $\hat{P}$ is a pair of n length binary vectors $(\vec{x}, \vec{z})$ such that a value 1 in position j of $\vec{x}$ means that $\hat{P}$ contains X acting on qubit j , and similarly for $\vec{z}$ and Z . This means $\hat{P}$ can be written as*

$$\hat{P}(\vec{x}, \vec{z}) = \bigotimes_{j=0}^{n-1} i^{x_j z_j} X^{x_j} Z^{z_j}. \quad (\text{C.2})$$

Here, i is the imaginary unit which occurs when $\vec{x}$ and $\vec{z}$ contain 1 in the same position. Both X and Z acting on a qubit, along with the phase, gives Y operator since $ZX = iY$.

Using this representation, U can be implemented as

$$U = \sum_{\vec{x}, \vec{z}} |\vec{x}, \vec{z}\rangle \langle \vec{x}, \vec{z}| \otimes \hat{P}(\vec{x}, \vec{z}) = \prod_{j=0}^{n-1} i^{x_j z_j} C_{x_j} X_j \cdot C_{z_j} Z_j, \quad (\text{C.3})$$

which requires three layers of two qubit gates. Firstly, a layer of controlled gates that applies Z to each qubit j , controlled on the value of auxiliary (qubit z_j), a second layer doing the same for X , and then a layer that applies the two qubit controlled S gate

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & i \end{pmatrix} \quad (\text{C.4})$$

between each pair of qubits (x_j, z_j) .

C.1.2 Preparing $|G\rangle_a$

We must now prepare the block encoding state

$$|G\rangle_a = \sum_{i=0}^{T-1} \sqrt{|\alpha_i|} |(\vec{x}, \vec{z})_i\rangle_a, \quad (\text{C.5})$$

which requires us to prepare a potentially arbitrary real superposition of basis states. This can be achieved using multi-controlled partial CNOT ($C^n \text{NOT}(\vartheta)$) gates as multi-controlled versions of the two qubit partial swap ($\text{PSWAP}(\vartheta)$) gates. The action of the singly controlled partial NOT gate is

$$\text{CNOT}(\vartheta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \vartheta & \sin \vartheta \\ 0 & 0 & \sin \vartheta & -\cos \vartheta \end{pmatrix}, \quad (\text{C.6})$$

and the uncontrolled action of the PSWAP gate on two qubits is

$$\text{PSWAP}(\vartheta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \vartheta & -\sin \vartheta & 0 \\ 0 & \sin \vartheta & \cos \vartheta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (\text{C.7})$$

for some rotation angle ϑ . We note two things, if we have a state containing basis states of Hamming weight less than or equal to n , then a $C^n\text{NOT}(\vartheta)$ gate can rotate an n weight basis state (with 1 corresponding to every control qubit) into an $n + 1$ weight basis state (with the additional Hamming weight coming from the target qubit) and will have no effect on lower weight states. Similarly, a $C^n\text{PSWAP}(\vartheta)$ will rotate between different states of Hamming weight $n + 1$ (those with 1s in the controls and a 1 and a 0 in the two target qubits) and have no effect on lower weight states.

The following procedure can then be used to generate an arbitrary superposition of basis states.

1. Starting with zero state, apply X gate to flip a single qubit.
2. Use $\text{PSWAP}(\vartheta)$ gates, to distribute this single 1 between other qubits to generate all Hamming weight 1 states required.
3. Use a single $\text{CNOT}(\vartheta)$ gate to rotate one of the Hamming weight 1 states into a Hamming weight 2 state.
4. Use $\text{CPSWAP}(\vartheta)$ gates to distribute this Hamming weight 2 state into all Hamming weight 2 states required (Hamming weight 1 states will be unaffected).
5. Repeat steps 3. and 4., increasing the number of controls and the Hamming weight by 1 each time, until all states up to the maximum Hamming weight have been produced.

The choice of ϑ for each gate must be chosen according to the values of α_i required, noting that larger amplitudes than $\sqrt{\alpha_i}$ will be required from some rotations for

basis states that will have CⁿNOT gates applied to access the higher weight basis states.

C.1.3 Applying signs to $\hat{P}_i$'s and preparing $|\tilde{G}_a\rangle$

To prepare the state

$$|\tilde{G}\rangle_a = \sum_{i=0}^{T-1} \text{sgn}(\alpha_i) \sqrt{|\alpha_i|} |(\vec{x}, \vec{z})_i\rangle_a, \quad (\text{C.8})$$

we can use the same procedure and rotation gates as for $|G\rangle$. The only change required is to flip the signs of some of the rotation angles for some of the partial SWAP and partial CNOT gates. By flipping the sign of the rotation angle $\vartheta \rightarrow -\vartheta$ for these gates, we induce a -1 phase shift in the resulting basis state compared to the one used in its generation. This approach ensures the incorporation of the requisite negative signs. It is crucial to monitor the sign evolution of successive bitstrings throughout this non-Markovian process. The sign of a basis state being rotated out may itself be positive or negative, contingent upon the preceding rotations and potential bitflips utilised in its generation.

C.2 Gate costs for block encoding Schwinger model

C.2.1 Some useful mathematics

Before we evaluate the quantum resource required for the block encoding, we prove a few useful lemmas.

Lemma C.2.1. *Given positive integer b ,*

$$\sum_{w_y=0}^{\lfloor \frac{b}{2} \rfloor} \binom{b-w_y}{w_y} = F(b+1), \quad (\text{C.9})$$

where $F(n)$ is the n th Fibonacci number.

Proof. See Ref. [279]. □

Lemma C.2.2. *Defining $F(n)$ to be the n th Fibonacci number, then for any positive integer m ,*

$$\sum_{b=2}^{m+2} F(b+1) = \frac{1}{\sqrt{5}} \left(\phi^{m+5} - (-\phi)^{-m-5} \right) - 3. \quad (\text{C.10})$$

where ϕ is the golden ratio.

Proof. From the definition of the Fibonacci numbers

$$\begin{aligned} F(m+3) &= F(m+5) - F(m+4) \\ F(m+2) &= F(m+4) - F(m+3) \\ &\vdots \\ F(4) &= F(6) - F(5) \\ F(3) &= F(5) - F(4). \end{aligned} \quad (\text{C.11})$$

Summing the L.H.S. and R.H.S. of the above yields

$$\sum_{b=2}^{m+2} F(b+1) = F(m+5) - F(4). \quad (\text{C.12})$$

Using Binet's formula and $F(4) = 3$ yields the result. $\square$

Lemma C.2.3. *For any integer positive integer m ,*

$$\sum_{b=2}^{m+2} bF(b+1) = \frac{m+2}{\sqrt{5}} \left(\phi^{m+5} - (-\phi)^{-m-5} \right) - \frac{1}{\sqrt{5}} \left(\phi^{m+6} - (-\phi)^{-m-6} \right) + 2. \quad (\text{C.13})$$

Proof. From the definition of the Fibonacci numbers

$$\begin{aligned} (m+2)F(m+3) &= (m+2)F(m+5) - (m+2)F(m+4) \\ (m+1)F(m+2) &= (m+1)F(m+4) - (m+1)F(m+3) \\ &\vdots \\ 3F(4) &= 3F(6) - 3F(5) \\ 2F(3) &= 2F(5) - 2F(4). \end{aligned} \quad (\text{C.14})$$

Summing the L.H.S. and R.H.S. of the above equations yields

$$\begin{aligned} \sum_{b=2}^{m+2} bF(b+1) &= (m+2)F(m+5) - \sum_{b=4}^{m+3} F(b+1) - 2F(4) \\ &= (m+2)F(m+5) - F(m+6) + F(3), \end{aligned} \quad (\text{C.15})$$

where we have rewritten the sum in the second line using Lemma C.2.2. Using Binet's formula and $F(3) = 2$ yields the result. $\square$

Corollary C.2.4.

$$\begin{aligned} \sum_{b=2}^{m+2} F(b+1)(\alpha b + \beta) &= \frac{\alpha}{\sqrt{5}} \left(\phi^{m+5} - (-\phi)^{-m-5} \right) m - \frac{\alpha}{\sqrt{5}} \left(\phi^{m+6} - (-\phi)^{-m-6} \right) \\ &\quad + \frac{(\alpha + \beta)}{\sqrt{5}} \left(\phi^{m+5} - (-\phi)^{-m-5} \right) + (2\alpha - 3\beta). \end{aligned} \quad (\text{C.16})$$

Proof. This follows from Lemmas C.2.2 and C.2.3. $\square$

Proposition C.2.5.

$$\sum_{b=0}^m \binom{m}{b} (\alpha b + \beta) = \alpha m 2^{m-1} + \beta 2^m. \quad (\text{C.17})$$

C.2.2 Pauli decomposition of the Hamiltonian

In order to calculate the resources needed for the qubitisation procedure, we begin by calculating the Pauli operators needed to construct the Hamiltonian.

Definition C.2.6 (Weight of a Pauli operator). *Given an operator $\hat{P}$ which can be written as single tensor product of the form $\{1, X, Y, Z\}^{\otimes m}$, we define the weight $W(\hat{P})$ of $\hat{P}$ as the total number of non-identity operators within $\hat{P}$.*

Definition C.2.7 (X, Y and Z weight of a Pauli operator). *For an operator $\hat{P}$ as above, we define the X weight of an operator as the number of X operators in its Pauli decomposition. Similarly for the Y weight and Z weight. If $W(\hat{P})$ is the total Pauli weight of an operator, then $W_x(\hat{P}) + W_y(\hat{P}) + W_z(\hat{P}) = W(\hat{P})$, where $W_{x,y,z}(\hat{P})$ are the X, Y and Z weights respectively.*

Remark C.2.8. *Given an operator $\hat{P}$ with X, Y, Z Pauli weights $W_x(\hat{P}), W_y(\hat{P}), W_z(\hat{P})$. The Hamming weight of the binary representation vector $(\vec{x}, \vec{y})$ of $\hat{P}$ as defined in Def. C.1.1 is equal to $b(\hat{P}) = W_x(\hat{P}) + 2W_y(\hat{P}) + W_z(\hat{P})$.*

To analyse the resource requirements of the qubitisation procedure, we will need to decompose each term in the Hamiltonian into a basis of Pauli operators and calculate the number of constituent operators of a given weight.

C.2.2.1 Gauge field terms

Lemma C.2.9. *Consider the truncated gauge-field operator $L^2 = \sum_{l=0}^{\Lambda-1} (l - \Lambda/2)^2 |l\rangle\langle l|$ where Λ is a power of 2. If we write the operator in the Pauli basis $\{\mathbb{1}, X, Y, Z\}^{\otimes m}$, where $m = \log_2(\Lambda)$, then for every integer $0 \leq w < m$, the number of multi-qubit Pauli operators of weight w is less than or equal to the binomial coefficient $\binom{m}{w}$.*

number of multi-qubit Pauli operators of weight

Proof. Using (A.4), we can rewrite this operator as

$$\sum_{l=0}^{\Lambda-1} (l - \Lambda/2)^2 |l\rangle\langle l| = \sum_{l=0}^{\Lambda-1} (l - \Lambda/2)^2 \bigotimes_{j=0}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(l)_j} Z), \quad (\text{C.18})$$

where $\text{bin}(l)_j \in \{0, 1\}$ denotes the j th bit in the binary representation of integer l . For a given l , expand the number the m qubit tensor product yields all m qubit operators in $\{\mathbb{1}, Z\}^{\otimes m}$ with coefficients ± 1 , the number of unique terms containing a total of W Pauli Z operators is given by the binomial coefficient $\binom{m}{w}$. Every value of l yields operators in $\{\mathbb{1}, Z\}^{\otimes m}$, just with a different set of prefactors. Thus, the total number of unique Pauli operators of W will be at most $\binom{m}{w}$. It may be the case that, depending on the values of Λ and therefore $(l - \Lambda/2)$, some of these term cancel, however the binomial coefficient provides an upper limit on the number of terms. $\square$

Remark C.2.10. *We note that the only Pauli operators within the above are Z operators, there are no X or Y operators.*

C.2.2.2 Interaction terms

Lemma C.2.11. *Consider the truncated raising operator $e^{-i\hat{\theta}} = \sum_{l=0}^{\Lambda} |l+1\rangle\langle l|$ where Λ is a power of 2 and cyclic numbering is used such that $|\underline{\Lambda}\rangle = |0\rangle$. If we write the raising operator in the Pauli basis $\{\mathbb{1}, X, Y, Z\}^{\otimes m}$, where $m = \log_2(\Lambda)$,*

then for every integer $0 \leq w \leq m$, the number of multi-qubit Pauli operators of weight w present is equal to 2^w .

Proof. As noted in Theorem A.3.1, the Pauli decomposition can be written as

$$|l+1\rangle\langle l| = \bigotimes_{j=k(l)+1}^m \left(\mathbb{1} + (-1)^{\text{bin}(l)_j} Z \right) \otimes (X + iY) \otimes (X - iY)^{\otimes k(l)}, \quad (\text{C.19})$$

where $\text{bin}(l)_j \in \{0, 1\}$ denotes the j th bit in the binary representation of integer l and $k(l) = \max\{j \mid \text{bin}(l)_j = 0\}$ is the position of the right-most zero within that representation. We say $k(l) = m$, if l contains no zeros (i.e., $l = \Lambda - 1$). Therefore, we can write the full raising operator as

$$\sum_{l=0}^{\Lambda-1} |l+1\rangle\langle l| = \sum_{l=0}^{\Lambda-1} \bigotimes_{j=k(l)+1}^m \left(\mathbb{1} + (-1)^{\text{bin}(l)_j} Z \right) \otimes (X + iY) \otimes (X - iY)^{\otimes k(l)}, \quad (\text{C.20})$$

which, by rewriting the summations, can be re-expressed as

$$\begin{aligned} \sum_{\{l \mid k(l)=k'\}} |l+1\rangle\langle l| &= \sum_{k'=0}^m \left[\sum_{\{l \mid k(n)=k'\}} \bigotimes_{j=k'+1}^m \left(\mathbb{1} + (-1)^{\text{bin}(l)_j} Z \right) \right] \\ &\quad \otimes (X + iY) \otimes (X - iY)^{\otimes k'}. \end{aligned} \quad (\text{C.21})$$

If we consider first the term within the square braces, we note that

$$\sum_{\{l \mid k(l)=k'\}} \bigotimes_{j=k'+1}^m \left(\mathbb{1} + (-1)^{\text{bin}(l)_j} Z \right) = (2\mathbb{1})^{\otimes (m-k'-1)}. \quad (\text{C.22})$$

We can see that this holds by induction. Firstly, for $m = k' + 1$, we note that there are only four terms (acting on a single qubit) within the summation, these correspond to $n = \Lambda$ and $n = \Lambda + 1 = 0$. The summation is thus equal to $(\mathbb{1} + Z) + (\mathbb{1} - Z) = 2\mathbb{1}$. When $m = k' + 2$, there are two terms within the summation, each acting on two qubits. These are equal to

$$\begin{aligned} &(\mathbb{1} + Z) \otimes (\mathbb{1} + Z) + (\mathbb{1} + Z) \otimes (\mathbb{1} - Z) + (\mathbb{1} - Z) \otimes (\mathbb{1} + Z) + (\mathbb{1} - Z) \otimes (\mathbb{1} - Z) \\ &= [(\mathbb{1} + Z) + (\mathbb{1} - Z)] \otimes [(\mathbb{1} + Z) + (\mathbb{1} - Z)] \\ &= 2\mathbb{1} \otimes 2\mathbb{1}. \end{aligned} \quad (\text{C.23})$$

From this, we can see how (C.22) holds by induction. Therefore,

$$\sum_{\{l|k(l)=k'\}} |\underline{l+1}\rangle\langle l| = \sum_{k'=0}^m (2\mathbb{1})^{\otimes(m-k'-1)} \otimes (X + iY) \otimes (X - iY)^{k'}. \quad (\text{C.24})$$

Looking at the $(X + iY) \otimes (X - iY)^{\otimes k'}$ term, expanding these brackets will yield $2^{k'+1}$ terms in $\{X, Y\}^{\otimes k'+1}$, which all have weight $k' + 1$. There will be $(m + 1)$ unique values of $k' = 0, 1, \dots, m$, each of which contribute $2^{k'+1}$ terms of weight $k' + 1$, hence the result. $\square$

Remark C.2.12. *From the above proof, we can also see that if we consider all multi-qubit Pauli operator of a given w within the above expansion, the number of operators that contain a total of w_y Y operators (or equivalently X operators) is $\binom{w}{w_y}$. There are no Z operators remaining in the expansion.*

Corollary C.2.13. *Consider a single interaction term (for a certain n) given by (3.12) using the truncated raising and lowering operators $e^{-i\hat{\theta}} = \sum_{l=0}^{\Lambda} |\underline{l+1}\rangle\langle l|$ and $e^{i\hat{\theta}} = \sum_{l=0}^{\Lambda} |l\rangle\langle l+1|$. Decomposing the interaction term into Pauli operators onto a qubit register of size $(m + 2)$, such that $m = \log_2(\Lambda)$ qubits are used for the gauge field space, and the remaining two qubits for the two spin degrees of freedom, then for every integer w satisfying $2 \leq w \leq m + 2$, there will be 2^{w-1} Pauli operators of weight w . These operators will always contain an even number w_y of Pauli Y operators, for a given w the number of operators with w_y Pauli Y operators will be $\binom{w}{w_y}$.*

Proof. Decomposing the spin raising and lower operators within a single interaction term, we get

$$\begin{aligned} \sigma^+ e^{i\theta} \sigma^- + \text{h.c.} &= \frac{1}{4} [(X + iY) e^{i\theta} (X - iY) + \text{h.c.}] \\ &= \frac{1}{4} [X e^{i\theta} X + iX e^{i\theta} Y - iY e^{i\theta} X + Y e^{i\theta} Y + \text{h.c.}]. \end{aligned} \quad (\text{C.25})$$

We can see from (C.24), that half of the terms within the Pauli expansion of $e^{i\theta}$ are Hermitian and half are anti-Hermitian. Therefore, exactly half the terms within the expansion $X e^{i\theta} X + iX e^{i\theta} Y - iY e^{i\theta} X - Y e^{i\theta} Y$ will be Hermitian and anti-Hermitian.

Thus, in adding the Hermitian conjugate to get the full interaction term in (C.25), half of the terms (the anti-hermitian ones) will cancel, and the other half (the Hermitian ones) will be repeated. We note that anti-hermitian terms, and therefore those that cancel, always contain an odd number of Pauli Y operators, since in the definition of the raising operators these contain the imaginary prefactors.

Counting up the number of terms in (C.25), multiplying by the weights of the operator in the Pauli decomposition of $e^{i\theta}$ and taking into account the cancellation of half the terms gives the required result. $\square$

Remark C.2.14. *The second term, corresponding to the energy of the gauge field, in the gauged transformed Schwinger model of (3.14) will contain all Pauli operators of the form $\{1, Z\}^{\otimes N}$. There are 2^N of these operators and thus the expected circuit cost for producing the block encoding state $|G\rangle$ with the method described would grow exponentially in system size. Although using a gauge transformed Hamiltonian would reduce the qubit numbers required, this comes at the expense of an exponential increase in circuit depth.*

C.2.2.3 Spin terms

Remark C.2.15. *It is trivial to see that each the Pauli decomposition for an onsite spin term $\mu(-1)^n [1 + \sigma^3(n)] / 2$ consists of a single Z operator only.*

C.2.3 Generating the block encoding state $|G\rangle$

The circuit cost of generating $|G\rangle$ using the method described in Appendix C.1 is entirely determined by the number of operators and the weight of each operator within the Pauli decomposition of the Hamiltonian. Here we will estimate the total number of gates needed to generate G for the 1+1D Schwinger Hamiltonian. We will start by calculating the number of vectors of each Hamming weight required to encode the Pauli decomposition of the Hamiltonian. We will consider gauge field terms, spin terms and interaction terms separately.

We will separate $|G\rangle$ into parts corresponding to each term in the Hamiltonian:

$$|G\rangle = \sum_{n=1}^{N-1} |G^{\text{field}}(n)\rangle + \sum_{n=1}^N |G^{\text{spin}}(n)\rangle + \sum_{n=1}^{N-1} |G^{\text{int}}(n)\rangle, \quad (\text{C.26})$$

where $|G^{\text{field}}(n)\rangle$ is an unnormalised state containing the basis states corresponding to the binary representation of the Pauli decomposition, with appropriate amplitudes, for the Hamiltonian term for the energy of gauge field n . The definitions of $|G^{\text{int}}(n)\rangle$ and $|G^{\text{spin}}(n)\rangle$ follow similarly.

C.2.3.1 Binary representation of gauge field terms $|G^{\text{field}}(n)\rangle$

Consider a single term corresponding to the gauge field energy for a given site n in (3.10), for a given integer $0 \leq w_z \leq m$, there will be $\binom{m}{w_z}$ Pauli operators of Pauli weight $W(\hat{P}) = W_z(\hat{P}) = w_z$. This follows from Lemma C.2.9 and Remark C.2.10. Thus, for all $0 < w_z \leq m$, the gauge field term for a given lattice site n in the Hamiltonian will contribute a total of $\binom{m}{w_z}$ Pauli operators of Pauli weight w_z , which from Remark C.2.8, will require binary representation vector of Hamming weight $b = w_z$. Therefore, we write the number of terms corresponding to Hamming weight b due to the fields as

$$N_b^{\text{field}} = \begin{cases} \binom{m}{b} & \text{for } 0 \leq b \leq m \\ 0 & \text{otherwise.} \end{cases} \quad (\text{C.27})$$

C.2.3.2 Binary representation of spin terms $|G^{\text{spin}}(n)\rangle$

Next, consider a single term corresponding to the onsite energy of a single spin n . The Pauli decomposition of this operator will involve a single qubit Z operator with Pauli weight $W(\hat{P}) = W_z(\hat{P}) = 1$. Thus the N onsite spin energy terms will lead to N operators with binary representation vector of hamming weight one. Therefore, we write the number of terms corresponding to Hamming weight b due to the spins as

$$N_b^{\text{spin}} = \begin{cases} 1 & \text{for } b = 1 \\ 0 & \text{otherwise.} \end{cases} \quad (\text{C.28})$$

C.2.3.3 Binary representation of interaction terms $|G^{\text{spin}}(n)\rangle$

Finally, consider a single term corresponding to an interaction term for a given n . We define $m' := m + 2$. From Corollary C.2.13, for a given $2 \leq w \leq m'$, there will be $\binom{w}{w_y}$ bitstrings of Pauli weight $w_y \leq w$ if w_y is even and zero otherwise. The Hamming weight of the associated binary representations will be $b = w + w_y$, which can therefore take any value between 2 and $2m'$. The number of terms with Hamming weight b will be given by the following summation

$$N_b^{\text{int}} = \sum_{\substack{2 \leq w \leq m' \\ 0 \leq w_y \leq m' \\ \text{s.t. } b=w+w_y \\ w_y \text{ even}}} \binom{w}{w_y} = \sum_{\substack{w_y = \max(0, b-m') \\ w_y \text{ even}}}^{\lfloor b/2 \rfloor} \binom{b-w_y}{w_y}, \quad (\text{C.29})$$

where we arrive at the final line by substituting $w = b - w_y$ and simplifying the ranges on the summations appropriately. We can find an upper bound for this as follows:

$$N_b^{\text{int}} = \sum_{\substack{w_y = \max(0, b-m') \\ w_y \text{ even}}}^{\lfloor b/2 \rfloor} \binom{b-w_y}{w_y} \leq \sum_{\substack{w_y=0 \\ w_y \text{ even}}}^{\lfloor b/2 \rfloor} \binom{b-w_y}{w_y} \leq F(b+1), \quad (\text{C.30})$$

where we have used the result for the summation from Lemma C.2.1. Therefore,

$$N_b^{\text{int}} \leq \begin{cases} F(b+1) & 2 < b \leq 2m' \\ 0 & \text{otherwise.} \end{cases} \quad (\text{C.31})$$

C.2.3.4 Gate cost of associated C^n PSWAPs and C^n NOTs

Now that we have the number of bitstrings for each Hamming weight, we can begin to calculate the gate cost of generating $|G\rangle = \sum_i \sqrt{\alpha_i} |i\rangle$. To do so, we first break down the cost of the multiqubit $C^n\text{NOT}(\vartheta)$ and $C^n\text{PSWAP}(\theta)$ in the following two remarks, before multiplying these by the total number that must be applied according to the number of binary representations required with each Hamming weight.

Remark C.2.16 (n qubit Toffoli gate with one ancilla [280]). *A single $C^{n-1}\text{NOT}$, where $n \geq 3$, can be implemented using $32n - 96$ T -gates, $24n - 72$ CNOT gates and one recyclable ancilla qubit. A $C^{n-1}Z$ gate can be implemented with the same cost plus two additional Hadamard gates.*

Remark C.2.17 (n qubit Toffoli gate with multiple ancillae [280]). A single $C^{n-1}\text{NOT}$, where $n \geq 3$, can be implemented using $4n - 8$ T -gates, $4n - 7$ CNOT gates and $n - 1$ ancilla qubits. A $C^{n-1}Z$ gate can be implemented with the same cost plus two additional Hadamard gates.

Remark C.2.18 (Creating multi-controlled PSWAP and partial CNOT gates). Following the logic of Appendix C3. of Ref. [120], $C^{n-1}\text{PSWAP}(\vartheta)$ and $C^n\text{NOT}(\vartheta)$ gates can be implemented using the following gate sequences:

$$C_{\{j\}}^{n-1}\text{PSWAP}_{k,l}(\vartheta) = e^{\frac{i\vartheta\hat{\mu}_{k,l}}{2}} \cdot C_{\{j\}}^{n-1}Z_k \cdot e^{\frac{-i\vartheta\hat{\mu}_{k,l}}{2}} \cdot C_{\{j\}}^{n-1}Z_k, \quad (\text{C.32})$$

where $C_{\{j\}}^{n-1}$ denotes controlled operator on a set of $n - 1$ qubits $\{j\}$ and the operator $\hat{\mu}_{k,l}$, which generates the two qubit $\text{PSWAP}_{k,l}(\vartheta)$ gate, is defined by

$$\hat{\mu}_{k,l} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (\text{C.33})$$

and

$$C_{\{j\}}^{n-1}\text{CNOT}_{k,l}(\vartheta) = e^{\frac{i\vartheta\hat{\nu}_{k,l}}{2}} \cdot C_{\{j\}}^{n-1}Z_k \cdot e^{\frac{-i\vartheta\hat{\nu}_{k,l}}{2}} \cdot C_{\{j\}}^{n-1}Z_k \cdot C_{\{j\}}^{n-1}S_k^\dagger, \quad (\text{C.34})$$

where $\hat{\nu}_{k,l}$, which generates a two qubit $\text{CNOT}(\vartheta)_{k,l}$ gate, is defined by

$$\hat{\nu}_{k,l} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (\text{C.35})$$

As discussed by Kirby et al. [120], the application of the multi-controlled $S^\dagger$ gate can be included in the decomposition of the multi-controlled Z gate so that it contributes none of its two-qubit or non-Clifford gates.

Remark C.2.19. The two qubit gates $\text{CNOT}(\vartheta)$ and $\text{PSWAP}(\vartheta)$ can be constructed with two CNOTs, six single qubit rotations (i.e. R_z gates) and single qubit gates.

This follows from Vatan and Williams' ([188], Fig. 2) construction of a general $\text{SO}(4)$ gate and counting the number of non-clifford and CNOT gates required.

Lemma C.2.20 (Partial swap and partial CNOT gates.). *A C^{n-1} PSWAP(ϑ) gate and a C^n NOT(ϑ) can be implemented using $8n - 16$ T -gates, $8n - 14$ CNOT gates, $12 R_Z$ gates and single qubit Clifford gates, provided we have access to $n - 1$ ancilla qubits.*

Proof. This follows from Remarks C.2.17, C.2.18 and C.2.19 □

C.2.3.5 Counting the number of C^n PSWAPs, C^n NOTs and C^n Zs

Following the procedure as laid out in Sec. C.1.2, we must apply a C^n NOT(ϑ) once for each value of possible Hamming weight (each time we encounter Step 3) up to the maximum possible values $b_{\max} = 2m + 4$. We note that a single C^n PSWAP(ϑ) gate can only rotate between states that have same Hamming weight and differ in the position of only a single one bit. Thus, if we just apply a single C^n NOT(ϑ) each time we apply Step 3, say to access a Hamming weight $n + 1$ state from a Hamming weight n state, then it is likely that more than one of the 1s within this new state do not match up with the position of one of the 1s in one of the weight $n + 1$ states we need to rotate into with the C^n PSWAP(ϑ) gate. This would require us to apply more than one C^n PSWAP(ϑ) per unique value of $(\vec{x}, \vec{z})$ gate of Hamming weight $n + 1$. However, if we examine the form of the Pauli decompositions for the spin (which is trivial), field (C.18) and interaction terms (C.24) for single lattice site n , we can exploit the structure of the binary representation strings to ensure we only need to apply at most one C^n PSWAP(ϑ) for each unique value of $(\vec{x}, \vec{z})$, at the expense of a small number of C^n NOT(ϑ) gates.

Additionally, we will reduce the overall cost of the algorithm by only using the procedure of [120] (described in Section C.1.2 using C^n NOT(ϑ) and C^n PSWAP(ϑ)) gates to generate the terms of $|G\rangle$ corresponding to gauge field and interaction energy of a single gauge link (and all Fermion only terms). We will then generate the terms related to the other $N - 2$ gauge sites using a procedure described in Section C.2.3.6.

For the gauge field energy of a single site, the binary representation vectors (considering only the m bits corresponding to the gauge field qubits of each auxiliary register) take all possible values of the form $\vec{x} = 0^m, \vec{z} \in \{0, 1\}^m$, where 0^m corresponds to a zero bit repeated m times. Thus, for each unique Hamming weight $0 < b \leq m$, there is at least one binary representation vector present that differs from another of weight $b - 1$ by just a single bit. Thus, for the gauge field terms, one must apply $C^{b-1}\text{NOT}(\vartheta)$ gate for each value of $1 \leq b \leq m$, where we have multiplied by the total number of gauge field registers, to generate the appropriate Hamming weights. This is then combined with $\left[\binom{m}{b} - 1\right] C^{b-2}\text{PSWAP}(\vartheta)$ gates for each $1 \leq b \leq m$ to access all bitstrings of that Hamming weight. Here, the binomial coefficient arises from the number of weight b strings given by (C.27).

For the interaction terms of a single gauge link site, we can see from corollary C.2.13 that the binary representation of the interaction terms have the following form

$$\vec{x} = 1^w \cdot 0^{m'-w}, \quad \vec{z} \in \pi(1^{w_y} \cdot 0^{w-w_y}) \cdot 0^{m'-w}, \quad (\text{C.36})$$

for $2 \leq w_y \leq w \leq m$, with w_y even, where for notational simplicity the first two bits correspond to the two spin sites and the remaining m bits correspond to the gauge field. Here $\pi(1^{w_y} \cdot 0^{w-w_y})$ denotes the set of all permutations of a bitstring containing w_y ones and $w - w_y$ zeros. From this representation it is clear that for every weight $b > 1$ bitstring there is at least one a) bitstring of weight $b - 1$ that is the same except it contains one fewer one in the $\vec{x}$ register or b) bitstring of weight $b - 2$ that is the same except it contains two fewer ones in the $\vec{z}$ register. Thus, every value of b can be accessed by applying a $C^{b-1}\text{NOT}(\vartheta)$ to one of the $b - 1$ states, or a $C^{b-2}\text{NOT}(\vartheta)$ to a weight $b - 2$ state followed by a $C^{b-1}\text{NOT}(\vartheta)$. Therefore, all interaction bitstrings for each Hamming weight $2 \leq b \leq 2m'$ can be generated by applying $N - 1$ total $C^{b-1}\text{NOT}(\vartheta)$ gates and less than or equal to $F(b + 1) - 1$ total $C^{b-2}\text{PSWAP}(\vartheta)$ gates. The latter upper bound comes from the remaining number of terms of weight b from (C.31).

Hamiltonian term	Hamming weights	# $C^{b-1}\text{NOT}(\vartheta)$	# $C^{b-2}\text{PSWAP}(\vartheta)$
Gauge fields	$0 < b \leq m$	1	$\binom{m}{b} - 1$
Interactions	$2 \leq b \leq m + 2$	1	$F(b + 1) - 1$
Spins	$b = 2$	0	$(N - 1)$

Table C.1: Number of multi-controlled gates for each Hamming weight b needed to generate state $(N - 1) |G^{\text{field}}(1)\rangle + (N - 1) |G^{\text{int}}(1)\rangle + \sum_{i=1}^N |G^{\text{spin}}(n)\rangle$ for each term in the Hamiltonian. See Section C.2.3.5 for details.

Hamiltonian term	# $C\text{NOT}(\vartheta)$	# X	# $C\text{SWAP}$	# $C^m\text{NOT}$
Gauge fields + Interactions	$N - 2$	$m(N - 2)$	$(m + 4)(N - 2)$	$N - 2$

Table C.2: Gate counts to partially swap gauge field and interaction terms for a gauge link site into the remaining $N - 2$ sites implementing the transformation $(N - 1) |G^{\text{field}}(1)\rangle + (N - 1) |G^{\text{int}}(1)\rangle + \sum_{i=1}^N |G^{\text{spin}}(n)\rangle \rightarrow |G\rangle$. See Section C.2.3.6 for details.

Finally, all the onsite spin terms are of weight one and can thus be generated from an initial weight one state by $N - 1$ total $\text{PSWAP}(\vartheta)$ gates. These weight one states can be used as the source for the $b = 1$ states for the interaction and gauge field terms.

The total number of $C^{b-1}\text{NOT}(\vartheta)$, $C^{b-2}\text{PSWAP}(\vartheta)$ for the three parts of the Hamiltonian are summarised in Table C.1.

C.2.3.6 Reducing gate cost using translational symmetry

Naively, we could generate $|G\rangle$ by applying the multi-controlled gates within Section C.2.3.5 a total of $(N - 1)$ times for the gauge field and interaction terms (once for each gauge link in the model) and N times for the spin terms. This would lead to a very large number of gates controlled on lots of qubits. However, by noticing that the gauge field and interaction terms have the same form for all $n = 1, 2, \dots, N$, we can significantly reduce the cost in preparing $|G\rangle$ by partially swapping these terms for the $n = 1$ site into all other lattice sites.

Consider following the procedure in the previous section to generate terms corresponding to just the gauge field and interaction terms for the first site. At this

point the state will be proportional to $|G^{\text{field}}(1)\rangle + |G^{\text{int}}(1)\rangle$. Written in terms of the computation basis, each term will be of the form

$$\underbrace{|\alpha, \beta\rangle}_{x_0, z_0 \text{ site } n=1} \otimes \underbrace{|\psi\rangle}_{\text{field } n=1} \otimes \underbrace{|\alpha', \beta'\rangle}_{x_0, z_0 \text{ site } n=2} \otimes \underbrace{|0^m\rangle}_{\text{field } n=2} \otimes \underbrace{|0, 0\rangle}_{x_0, z_0 \text{ site } n=3} \otimes \dots, \quad (\text{C.37})$$

where $\alpha, \beta, \alpha', \beta' \in \{0, 1\}$ and $|\psi\rangle$ represents one of the m binary states corresponding to the gauge field register. If we include an ancilla qubit to which we apply a rotation gate we have

$$\underbrace{(\cos(\vartheta)|0\rangle + \sin(\vartheta)|1\rangle)}_{\text{ancilla}} \otimes |\alpha, \beta\rangle \otimes |\psi\rangle \otimes |\alpha', \beta'\rangle \otimes |0^m\rangle \otimes |0, 0\rangle \otimes \dots \quad (\text{C.38})$$

Next, we apply m CSWAP (Fredkin) gates, controlled on the ancilla qubit, to swap each qubit within the m qubit $n = 1$ field register, with the corresponding qubit in the $n = 2$ field qubit. This yields

$$\begin{aligned} & \cos(\vartheta) |0\rangle \otimes |\alpha, \beta\rangle \otimes |\psi\rangle \otimes |\alpha', \beta'\rangle \otimes |0^m\rangle \otimes |0, 0\rangle \otimes \dots \\ & + \sin(\vartheta) |1\rangle \otimes |\alpha, \beta\rangle \otimes |0^m\rangle \otimes |\alpha', \beta'\rangle \otimes |\psi\rangle \otimes |0, 0\rangle \otimes \dots \end{aligned} \quad (\text{C.39})$$

We have now swapped the first two field registers, conditioned on the ancilla qubit. We remove the ancilla qubit, by applying an X gate to each qubit in the $n = 2$ field register, followed by an $C^m\text{NOT}$ gate conditioned on the $n = 2$ field register and targeting the ancilla qubit, followed by an X gate to each qubit in the $n = 2$ field register again, and finally an X gate to the ancilla qubit. This yields

$$\begin{aligned} & |0\rangle \otimes (\cos(\vartheta) |\alpha, \beta\rangle \otimes |\psi\rangle \otimes |\alpha', \beta'\rangle \otimes |0^m\rangle \otimes |0, 0\rangle \otimes \dots \\ & + \sin(\vartheta) |\alpha, \beta\rangle \otimes |0^m\rangle \otimes |\alpha', \beta'\rangle \otimes |\psi\rangle \otimes |0, 0\rangle \otimes \dots) \end{aligned} \quad (\text{C.40})$$

where we now have an arbitrary superposition (determined by ϑ) over states with the gauge field terms swapped, with the auxiliary qubit unentangled and returned to its initial state. For initial states corresponding to the terms corresponding to the $n = 1$ gauge field energy (C.18), $\alpha = \beta = \alpha' = \beta' = 0$ (as there are no spin terms) and so we have effectively partially swapped the required binary encoding of the gauge field terms from $n = 1$ to $n = 2$. We can ensure the interaction terms (in which some of $\alpha, \beta, \alpha', \beta'$ are non zero) are correctly swapped as well by doing

the following. Apply two Fredkin gates, controlled on the least significant x qubit in the $n = 2$ field register, swapping the registers corresponding to the $n = 2$ spin register with the $n = 3$ spin register. From (C.21), we see that interaction terms will always contain a one within this control qubit, thus for $|\psi\rangle$ corresponding to an interaction term, we get

$$\begin{aligned} & \cos(\vartheta) |\alpha, \beta\rangle \otimes |\psi\rangle \otimes |\alpha', \beta'\rangle \otimes |0^m\rangle \otimes |0, 0\rangle \otimes \dots \\ & + \sin(\vartheta) |\alpha, \beta\rangle \otimes |0^m\rangle \otimes |0, 0\rangle \otimes |\psi\rangle \otimes |\alpha', \beta'\rangle \otimes \dots, \end{aligned} \quad (\text{C.41})$$

where we omit the ancilla qubit which is no longer used. Finally, applying another two Fredkin gates, controlled in the same way to swap the $n = 1$ and $n = 2$ spin registers, we get:

$$\begin{aligned} & \cos(\vartheta) |\alpha, \beta\rangle \otimes |\psi\rangle \otimes |\alpha', \beta'\rangle \otimes |0^m\rangle \otimes |0, 0\rangle \otimes \dots \\ & + \sin(\vartheta) |0, 0\rangle \otimes |0^m\rangle \otimes |\alpha, \beta\rangle \otimes |\psi\rangle \otimes |\alpha', \beta'\rangle. \end{aligned} \quad (\text{C.42})$$

This is the form of the output for both interaction and gauge energy terms (for the latter, since all the spin bits are zero, these Fredkin gates have no effect). Thus, we have taken the bitstrings corresponding to the interaction and gauge field terms for $n = 1$ and used them to generate the $n = 2$ gauge field and interaction terms, avoiding the need to reapply the costly multicontrolled procedure described in Section C.2.3.5. i.e., following the notation in (C.26), we have performed

$$\begin{aligned} \frac{1}{\mathcal{N}} \left(|G^{\text{field}}(1)\rangle + |G^{\text{int}}(1)\rangle \right) & \longrightarrow \frac{1}{\mathcal{N}} \left(\cos \vartheta |G^{\text{field}}(1)\rangle + |G^{\text{int}}(1)\rangle \right) \\ & + \frac{1}{\mathcal{N}} \left(\sin(\vartheta) |G^{\text{field}}(2)\rangle + |G^{\text{int}}(2)\rangle \right), \end{aligned} \quad (\text{C.43})$$

where $\mathcal{N}$ is the norm of unnormalised state $|G^{\text{field}}(1)\rangle + |G^{\text{int}}(1)\rangle$. We can keep repeating this, noticing that the spin only terms are unaffected, and choosing ϑ each time such that we end up with the transformation

$$\begin{aligned} & (N - 1) \left(|G^{\text{field}}(1)\rangle + |G^{\text{int}}(1)\rangle \right) + \sum_{n=1}^N |G^{\text{spin}}(n)\rangle \\ & \longrightarrow \sum_{n=1}^{N-1} |G^{\text{field}}(n)\rangle + \sum_{n=1}^{N-1} |G^{\text{int}}(n)\rangle + \sum_{n=1}^N |G^{\text{spin}}(n)\rangle = |G\rangle. \end{aligned} \quad (\text{C.44})$$

The total gate cost for this is $(N-2) R_x(\vartheta)$ gates, $2m(N-2) X$ gates, $(m+4)(N-2)$ CSWAP gates and $(N-2) C^m\text{NOT}$ gates along with an ancilla qubit that can be reused.

C.2.3.7 Total gate cost for $|G\rangle$

Theorem C.2.21 (Improved gate cost for G). *For the single flavour 1+1D Schwinger model with N lattice sites and gauge fields represented by $m \geq 2$ qubits, the block encoding state $|G\rangle$ can be prepared using the following number of gates.*

$$\begin{aligned} \#T \text{ gates} &\leq 56 - 16m - 8 \frac{\phi^{m+7} - (-\phi)^{-m-7}}{\sqrt{5}} + 8m \frac{\phi^{m+5} - (-\phi)^{-m-5}}{\sqrt{5}} \\ &\quad - 16 \cdot 2^m + 4N + 4m2^m + 8Nm \end{aligned} \quad (\text{C.45})$$

$$\begin{aligned} \#CNOT \text{ gates} &\leq 54 - 22m - 6 \frac{\phi^{m+5} - (-\phi)^{-m-5}}{\sqrt{5}} - 8 \frac{\phi^{m+6} - (-\phi)^{-m-6}}{\sqrt{5}} \\ &\quad + 8m \frac{\phi^{m+5} - (-\phi)^{-m-5}}{\sqrt{5}} - 14 \cdot 2^m + 25N + 4m2^m + 11Nm \end{aligned} \quad (\text{C.46})$$

$$\#R_z \text{ gates} \leq -48 + 12 \frac{\phi^{m+5} - (-\phi)^{-m-5}}{\sqrt{5}} + 12 \cdot 2^m + 6N \quad (\text{C.47})$$

where ϕ is the golden ratio. In addition, we require an unspecified number of single qubit Clifford gates as well as $2N + 2m(N-1)$ qubits to store $|G\rangle$ and m ancilla qubits for the computation.

Proof. The full procedure to generate $|G\rangle$ consists of two parts. First, generate the state

$$(N-1) \left(|G^{\text{field}}(1)\rangle + |G^{\text{field}}(1)\rangle \right) + \sum_{n=1}^N |G^{\text{spin}}(n)\rangle, \quad (\text{C.48})$$

by applying the gates listed in Table C.1 as described in Section C.2.3.5 once for the gauge field and interaction terms, and $N-1$ times for the spin terms.

The total number of T-gate, CNOT gates and R_z gates for gauge field and interaction terms for this procedure can be written as

$$\sum_{b=1}^m \binom{m}{b} (\alpha b + \beta) + \sum_{b=2}^{m+2} F(b+1) (\alpha b + \beta), \quad (\text{C.49})$$

where for $\alpha = 8, \beta = -16$ for T-gates, $\alpha = 8, \beta = -14$ for CNOT gates and $\alpha = 0, \beta = 12$ for R_z gates according to the decomposition of $C^{b-1}\text{NOT}(\vartheta)$ and

$C^{b-2}\text{PSWAP}(\vartheta)$ gates of Proposition C.2.20. In addition, $2(N - 1)$ CNOTs and $6(N - 1)$ R_z gates are required for the spin terms.

Secondly, we must apply the gates listed in Table C.2 as described in Section C.2.3.6 to implement the transformation into $|G\rangle$.

The cost of a single CNOT(ϑ) gate will be six R_z gates and two CNOTs. A single Fredkin gates (CSWAP) can be implemented with two CNOTs along with a Toffoli (CCNOT) gate [281] for a total cost of four T gates and five CNOTs (using remark C.2.17 for the Toffoli decomposition). Each of the $C^m\text{NOT}$ gates requires $4m - 4$ T gates and $4m - 3$ CNOTs from Remark C.2.17. Thus the total cost for this step is

$$8Nm + 4N - 16m - 8 \quad \text{T gates}, \quad (\text{C.50})$$

$$11Nm + 25N - 22m - 4 \quad \text{CNOT gates} \quad (\text{C.51})$$

$$6N - 12 \quad R_z \text{ gates}. \quad (\text{C.52})$$

The required expressions result from using Lemma C.2.4 and Proposition C.2.5 to upper bound (C.48) with the appropriate values of α, β and adding the result to the relevant equation of (C.50), (C.51) and (C.52). $\square$

Remark C.2.22. *If we choose $m = \lceil \log(N) + 1 \rceil$ to correspond to the maximum gauge field occupancy of Remark 3.2.1, then the gate cost of $|G\rangle$ is $\mathcal{O}(N \log N)$ and $(2N - 1) \log N - \log(N)$ qubits are required.*

C.2.4 T gate cost of rotation gates

For error corrected quantum computing, we typically do not have access to an arbitrary gate set but rather a restricted subset from which arbitrary operations must be performed approximately. according to the Solovay Kitaev theorem however, the gate cost of doing so scales favourably with approximation error. In our case,

we need to convert the arbitrary R_z rotations into Clifford + T-gates. To do so, we use the following Lemma.

Proposition C.2.23 (Clifford+T approximations of z-rotations [259]). *There exists an efficient algorithm, which is known, for approximately decomposing an R_z gate up to rotation error ϵ which requires $3 \log \epsilon^{-1} + \mathcal{O}(\log \log \epsilon^{-1})$ T gates.*

We use this to make the following approximation for the number of T gates required to implement a single R_z gate in our algorithm.

Theorem C.2.24. *Assume the following*

1. *The average fractional error required on values of α_i is given by a constant value $\bar{\epsilon}_\alpha$.*
2. *The average value of $\sqrt{\alpha_i}$ is approximately equal to the reciprocal of the number of Pauli terms within the Hamiltonian.*
3. *Recalling that the values of $\sqrt{\alpha_i}$ are generated by successive rotations using partial swap gates, for an average approximation error ϵ per R_z gate, the average error on resulting value of given $\sqrt{\alpha_i}$ is less than approximately $d\epsilon$, where d is the maximum number of R_z gates affecting the value of $\sqrt{\alpha_i}$.*

Then, the average number of T gates required per R_z is approximately less than or equal to

$$3 \log \left((N-1)2^m + N \cdot 2^{m+1} + N \right) + 3 \log(12m + 48) + 3 \log(\bar{\epsilon}_\alpha^{-1}), \quad (\text{C.53})$$

up to $\mathcal{O}(\log \log)$ terms.

Proof. Given that the average value of $\sqrt{\alpha_i}$ equal to the reciprocal of the number of Pauli terms within the Hamiltonian. The absolute error required is approximated by

$$\epsilon_{\text{tot}} = \bar{\epsilon}_\alpha \left[(N-1)2^m + (2^{m-1})(N-1) + N \right]^{-1}, \quad (\text{C.54})$$

where we have used the total number of Pauli strings resulting from Corollary C.2.13 and Remarks C.2.14, C.2.15.

From Lemma C.2.20, each multi-qubit rotation gate for the generation of $|G\rangle$ requires $12R_z$ gates. The maximum number of rotations affecting a single value of α_i is equal to the Hamming weight of the corresponding basis state $|(\vec{x}, \vec{z})_i\rangle$. The maximum Hamming weight present in the Hamiltonian is equal to $2m + 4$ from the interaction terms of (C.25). Thus, to achieve average absolute error on $\sqrt{\alpha_i}$ that is less than or equal to ϵ_{tot} , we require $\epsilon_{R_z} \approx (12m - 48)^{-1} \epsilon_{\text{tot}}$.

From the the approximate values of ϵ_{tot} and ϵ_{R_z} above and using Proposition C.2.23 for the optimal decomposition of R_z gates we get the result. $\square$

Corollary C.2.25. *Assuming $\bar{\epsilon}_\alpha$ is constant, if $m = \lceil \log(N) \rceil$ then this is $\mathcal{O}(\log(N))$ and the total T -gate cost for generating $|G\rangle$, including rotations, is $\mathcal{O}(N \log(N))$. This does not affect the asymptotic scaling of this part of the algorithm.*

For gate cost calculations presented in the main text, we assume that $\bar{\epsilon}_\alpha$ terms are negligible.

C.2.5 Applying the block encoding unitary U

C.2.5.1 Total gate cost for U

Theorem C.2.26. *For the single flavour 1+1D Schwinger model with N lattice sites and gauge fields represented by $m \geq 2$ qubits, the block encoding state $|G\rangle$ can be prepared using the following number of gates.*

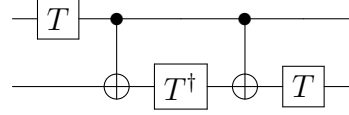
$$\#T \text{ gates} = 6N + 6mN - 6m \quad (\text{C.55})$$

$$\#CNOT \text{ gates} = 8N + 8mN - 8m \quad (\text{C.56})$$

where ϕ is the golden ratio. We require an unspecified number of single qubit Clifford gates and m ancilla qubits for the computation.

Proof. As described in Ref. [120], to implement the Pauli operators in U , we require a total of N_G , the number of qubits in $|G\rangle$, two qubit CNOT, CZ and Controlled S

(see (C.4)) gates. The decompositions of the first two are trivial, for the latter we note that a controlled S gate performing (C.4) can be implemented as:


(C.57)

Therefore, given that $N_G = 2N + 2m(N - 1)$, the number of gates required to apply the controlled Pauli operations of U is

$$3 \cdot (2N + 2m(N - 1)) \quad \text{T gates,} \quad (\text{C.58})$$

$$4 \cdot (2N + 2m(N - 1)) \quad \text{CNOT gates,} \quad (\text{C.59})$$

and zero R_z gates. □

Remark C.2.27. If $m = \lceil \log(N) + 1 \rceil$ as in Remark 3.2.1, then the gate cost of U is $\mathcal{O}(N \log N)$.

C.2.5.2 Applying phases within U

Rather than introducing the signs of α_i within $|\tilde{G}\rangle$, it is possible to follow the method presented in Appendix C of Ref. [120] to apply phases of α_i . To do so, a symmetric block encoding with $|G\rangle = \sum_i \sqrt{|\alpha_i|} |i\rangle$ is used and the required signs are applied within U to give the operation

$$U = \sum_i \text{sgn}(\alpha_i) |i\rangle\langle i| \otimes \hat{P}_i. \quad (\text{C.60})$$

To understand how this is done, we note that we can apply a -1 to a state auxiliary state of Hamming weight n by applying a $C^{n-1}Z$ gate to all auxiliary qubits containing 1 in the basis state of interest. This will also flip the sign of higher Hamming weight states that contain 1s in the same positions as the Hamming weight n state. Thus, the desired phases can be achieved by applying Z gates to qubits in Hamming weight one states that require a -1 phase, then applying CZ gates to Hamming weight two states that need their phases changed, flipping back the phase of any that may have been unnecessarily changed by the application of

the previous single qubit phase gates, then doing the same for weight three states with CCZ gates and so on until all states have the required phase applied to them. The gate overall gate cost is $\mathcal{O}(\log(N)N^2)$ as given in the following theorem.

Theorem C.2.28 (Applying phases within U). *If, we add the required phases for the Block encoding of the single flavour 1+1D Schwinger model using a method analagous to that described in Appendix C of Ref. [120], the gate cost of U is*

$$\begin{aligned} \#T \text{ gates} \leq (N-1) & \left[32 + 6m + 4m \frac{\phi^{m+5} - (-\phi)^{-m-5}}{\sqrt{5}} - 4 \frac{\phi^{m+7} - (-\phi)^{-m-7}}{\sqrt{5}} \right. \\ & \left. - 8 \cdot 2^m + 2m2^m \right] + 6N \end{aligned} \quad (\text{C.61})$$

$$\begin{aligned} \#CNOT \text{ gates} \leq (N-1) & \left[29 + 8m - 4 \frac{\phi^{m+6} - (-\phi)^{-m-6}}{\sqrt{5}} - 3 \frac{\phi^{m+5} - (-\phi)^{-m-5}}{\sqrt{5}} \right. \\ & \left. + 4m \frac{\phi^{m+5} - (-\phi)^{-m-5}}{\sqrt{5}} - 7 \cdot 2^m + 2m2^m \right] + 8N \end{aligned} \quad (\text{C.62})$$

Proof. In addition to the controlled Pauli and phase gates of Theorem C.2.26, we must generate the required minus signs by applying multicontrolled Z gates. In the worst case, every single bitstring of weight b requires a different phase from the total phase of the $b-1$ weight strings differing by just a single bit. Thus, each one must have a $C^b Z$ gate applied to achieve the correct phase. Thus, we require at most $(N-1)\binom{m}{b}$ total $C^{b-1}Z$ gates for each $1 \leq b \leq m$ for gauge field terms and $(N-1)F(b+1)$ total $C^{b-1}Z$ for each $2 \leq b \leq m+2$ gates for interaction terms. These values correspond to the N_b^{field} and N_b^{int} in (C.27), (C.31).

As in Theorem C.2.21, the gate cost of these terms is given by (C.49), this time multiplying by $(N-1)$, where for $C^{b-1}Z$ gates $\alpha = 4, \beta = -8$ for T gates and $\alpha = 4, \beta = -7$ for CNOT gates. No rotation gates are needed. The total cost follows from Lemma C.2.4 and Proposition C.2.5 (C.48). $\square$

C.2.6 Total gate cost for Π_φ and $\tilde{\Pi}_\varphi$

Theorem C.2.29. *For the single flavour 1+1D Schwinger model with N lattice sites and gauge fields represented by $m \geq 2$ qubits, the rotation operator Π_φ for the QSVT*

procedure of Theorem 3.3.2 can be implemented with two times the cost of G from Theorem C.2.21 plus $128mN + 128N - 128m - 192$ T gates, $96mN + 96N - 96m - 144$ CNOT gates and an R_z gate.

Proof. As described in Proposition 3.3.4, the rotation operator can be implemented with two G operators, two $C^{N_G}\text{NOT}$ gates and an R_z gate. From Remark C.2.16 the cost of a $C^{N_G}\text{NOT}$ gate is $32N_G - 96$ T gates and $24N_G - 72$ CNOT gates along with one ancilla qubit which can be reused from the G computation. Using $N_G = 2m(N - 1) + N$ and combining the costs of G s, the $C^{N_G}\text{NOT}$ s and rotation yields the result. $\square$

Remark C.2.30. If $m = \lceil \log(N) + 1 \rceil$ as in Remark 3.2.1, then the gate cost of Π_{φ_j} is $\mathcal{O}(N \log N)$. It may be possible to reduce the multicontrolled gates by using additional ancilla [282] or by reducing the rotation to the subspace of the maximum achievable Hamming weight as in [120]. Here we sketch the steps taken to compute the quantum resource cost.

Corollary C.2.31. Since $\tilde{G}$ can be generated from the same gates as G with some rotation angles differing in the phase, the cost of $\tilde{\Pi}_\varphi$ is the same as the cost of Π given in Theorem C.2.29.

C.2.7 Total gate cost of the full algorithm

Theorem C.2.32. A single measurement needed for the estimation of matrix element $\langle H \rangle$ can be calculated by performed using two applications of the gates listed in Theorem C.2.21 and k applications of the gates in Theorems C.2.26 and C.2.29. Then assuming $m = \log N$, the gate cost of this is $\mathcal{O}(kN \log(N))$.

Proof. This follows from Lemma 3.3.3 and the preceding gates costs for U , $|G\rangle$ and $|\tilde{G}\rangle$. $\square$

D

Additional simulations of QSE with measurement noise

Here we present the results of using thresholded QSE (TQSE) algorithm as given by Ref. [122] for the same systems as in Sec. 3.4.3. These are shown in Fig. D.1 as a direct comparison with Fig. 3.3 in which partitioned QSE (PQSE) [125] was used. Noise was applied in the same way as in the main text and the threshold parameter was chosen to be equal to $\|\Delta_{\mathbf{S}}\|$ which has been found, in practice to be a good choice of thresholding parameter [e]. For input states used the expectation of the Hamiltonian was approximately equal to system size N (as is the true ground state energy), therefore we rescaled expectation values used in $\mathbf{S}$ and $\mathbf{H}$ as $\langle\psi_0|H^k|\psi_0\rangle \longrightarrow \langle\psi_0|H^k|\psi_0\rangle/N^k$, which ensured elements of $\Delta_{\mathbf{S}}$ and $\Delta_{\mathbf{H}}$ were of similar absolute values (as well as the same fractions of corresponding elements in $\mathbf{S}$ and $\mathbf{H}$).

We can see that the TQSE performed approximately an order of magnitude worse than PQSE for the same number of calls to U . Furthermore, fitting and extrapolation to predict the number of calls to U to achieve a given measurement error shows more severe exponential scaling. Comparing the fit of Fig. D.2 to that of Fig. 3.5

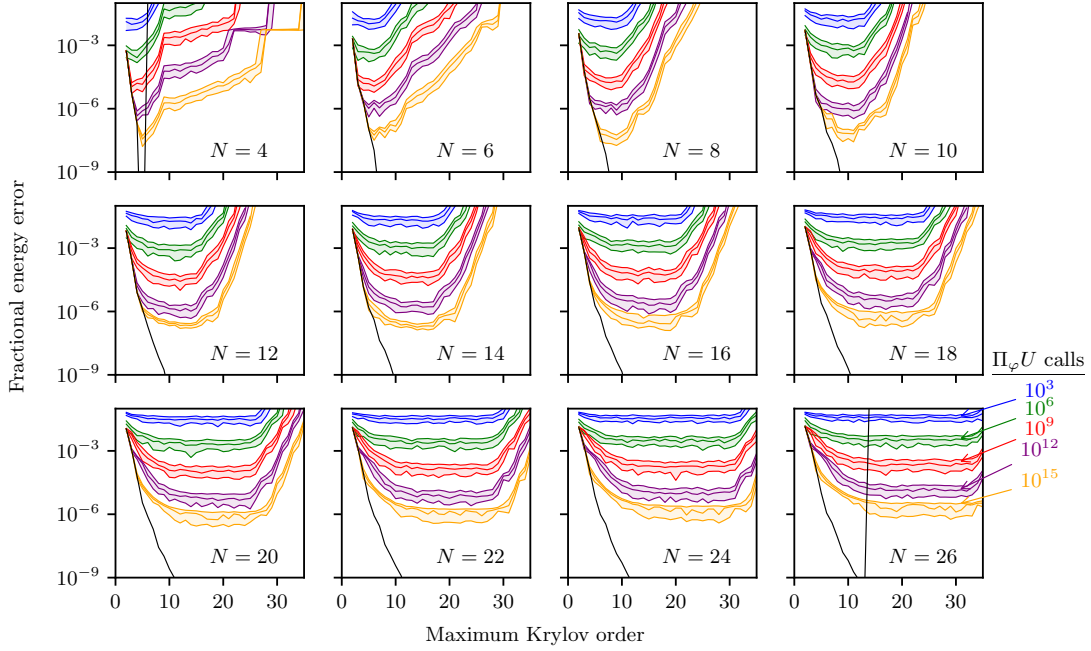


Figure D.1: Fractional energy error using TQSE procedure with finite number of measurements for range of different lattice sizes N . Values are shown as a function of the Krylov basis size for the TQSE procedure and as a function of number of calls to the block encoding operator U (which affect the number of measurements as defined in the main text). Lines and shaded area corresponds to the median as well as upper and lower quartiles of energy error. The black line (labelled as infinite calls to $\Pi_\varphi U$) corresponds to no measurement noise.

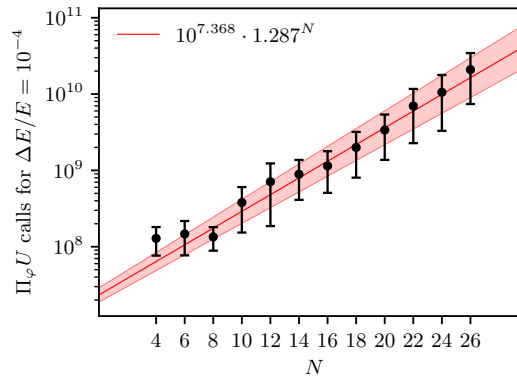


Figure D.2: Estimated number of call to qubitisation procedure with TQSE to achieve fractional energy error of 10^{-4} calculated using data from Fig. D.1. Red line corresponds to exponential fit (straight line on log-log axes) with filled area indicating uncertainty in fitting and extrapolation.

shows that the number of calls to U required to achieve fractional measurement error of 10^{-4} is $\sim 1.287^N$ for TQSE, compared to $\sim 1.119^N$ for PQSE.

E

Hardware information for current generation quantum processors

Below we give details for current generation quantum processors used for analysis in Chapter 3. Gate times and coherence times for four processors, where they could be found or calculated, are given in Table E.1.

For data on the Quantinuum H2-1 processor, two-qubit gate time was estimated in the following way. We take benchmark data for shot times and number of operations for a range of circuits from Table I of Ref. [283]. We take the shot time for the circuit, subtract the quoted initial state preparation time ($32\mu\text{s}$) and divide this by the number of two-qubit gates in the circuit (assuming two-qubit gates dominate the execution time over single qubit and SPAM gates). We take the mean of these values to estimate an average two qubit gate time, excluding the ‘Transport 1Q RB, $l = 64$ ’ circuit which contains no two-qubit gates. Coherence time is calculated by taking the quoted quadratic dephasing rate ($0.043 \cdot 2\pi\text{rad/s}$) from Ref. [284], which is quoted as being applied during qubit transport and idling. We inverted this and multiplied by the average fraction of time spent in transport from Table I of [283] to give an estimate of the typical dephasing time.

For data on the IonQ Forte processor, two-qubit gate times were estimated in the following way. Using run time benchmark given in the supplemental data for Ref. [285], for each type of algorithm listed (each with varying problem sizes), the total single circuit time was divided by the number of native two-qubit gates (assuming that two-qubit gates dominated execution time). The mean two-qubit gate time was then calculated over the algorithm types. Estimated coherence times are taken from Ref. [286].

Using the total CNOT counts given in Fig. 3.6, we get a rough estimate of the time taken to apply a single block encoding step $\Pi_{\varphi_j} U$ by multiplying the CNOT count by the two-qubit gate time for each processor. The values calculated are shown in Table E.2 both as an absolute time, as well as a multiple of the for T1 and T2 times. This makes three assumptions, (1) when compiled to the particular hardware gate set, the number of native two-qubit gates is roughly equal to the number of CNOTs we have counted, (2) total run time is dominated by two-qubit gate operations, and (3) CNOT gates are implemented in serial. In reality, many of these CNOTs could be run in parallel or even removed by recompilation; so this is likely to be an overestimate. We can instead assume that all gates are maximally parallelised such that two-qubit gates are applied in layers of $N_{\text{qubits}}/2$ acting at the same time to give a lower bound on the run time. These values are shown in Table E.3.

Processor	Type	# Qubits	T1	T2	2Q gate time
IBM Eagle r3 [287]	SC	127	275 μ s	117 μ s	636 ns
Google Sycamore [288]	SC	70	22.9 μ s	15.5 μ s	20 ns
Quantinuum H2-1 [283]	Ion trap	32	?	1410 s	6.46 ms
IonQ Forte [285]	Ion trap	32	10–100 s	$\sim$ 1s	931 μ s

Table E.1: Hardware information for range of superconducting (SC) and ion trap quantum processors. T1 and T2 correspond to qubit relaxation and coherence times and 2Q gate time refers to native two-qubit gate execution times. The Quantinuum H2-1 and IonQ Forte processors, T1, T2 and 2Q gate times were calculated as described in main text of this Appendix. T1 time for the Quantinuum HH2-1 could not be found. Values for other processors are median values taken from the corresponding references.

Processor	Block encoding run time (s)			run time as fraction of T1			run time as fraction of T2		
	$N=100$	$N=1,000$	$N=10,000$	$N=100$	$N=1,000$	$N=10,000$	$N=100$	$N=1,000$	$N=10,000$
Eagle r3	5.14×10^{-4}	4.47×10^{-2}	5.92×10^0	1.87×10^0	1.63×10^2	2.15×10^4	4.40×10^0	3.82×10^2	5.06×10^4
Sycamore	1.62×10^{-2}	1.41×10^0	1.86×10^2	7.06×10^2	6.14×10^4	8.13×10^6	1.04×10^3	9.08×10^4	1.20×10^7
H2-1	5.22×10^3	4.54×10^5	6.01×10^7	-	-	-	3.71×10^0	3.22×10^2	4.26×10^4
Forte	7.53×10^2	6.55×10^4	8.66×10^6	7.53×10^1	6.55×10^3	8.66×10^5	7.53×10^2	6.55×10^4	8.66×10^6

Table E.2: Estimated block encoding run time for different quantum processors assuming that two-qubit gates are executed in series for different system sizes. run times are given in seconds as well as proportions of the T1 (where known) and T2 times for each processor using data from Table E.1.

Processor	run time as fraction of T1			run time as fraction of T2		
	$N=100$	$N=1,000$	$N=10,000$	$N=100$	$N=1,000$	$N=10,000$
Eagle r3	1.88×10^{-3}	1.09×10^{-2}	1.08×10^{-1}	4.41×10^{-3}	2.56×10^{-2}	2.54×10^{-1}
Sycamore	7.09×10^{-1}	4.11×10^0	4.08×10^1	1.05×10^0	6.07×10^0	6.02×10^1
H2-1	-	-	-	3.72×10^{-3}	2.16×10^{-2}	2.14×10^{-1}
Forte	7.55×10^{-2}	4.38×10^{-1}	4.35×10^0	7.55×10^{-1}	4.38×10^0	4.35×10^1

Table E.3: Estimated block encoding run time for different quantum processors assuming that two-qubit gates are maximally parallelised. Run times are given as proportions of the T1 (where known) and T2 times for each processor using data from Table E.1.

References

- [1] Richard P Feynman et al. “Simulating physics with computers”. In: *Int. j. Theor. phys* 21.6/7 (1982).
- [2] Peter W Shor. “Scheme for reducing decoherence in quantum computer memory”. In: *Physical review A* 52.4 (1995), R2493.
- [3] Daniel A Lidar and Todd A Brun. *Quantum error correction*. Cambridge university press, 2013.
- [4] Joschka Roffe. “Quantum error correction: an introductory guide”. In: *Contemporary Physics* 60.3 (2019), pp. 226–245.
- [5] Zhenyu Cai et al. “Quantum error mitigation”. In: *arXiv preprint arXiv:2210.00921* (2022).
- [6] Peter W Shor. “Algorithms for quantum computation: discrete logarithms and factoring”. In: *Proceedings 35th annual symposium on foundations of computer science*. Ieee. 1994, pp. 124–134.
- [7] Lov K Grover. “A fast quantum mechanical algorithm for database search”. In: *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*. 1996, pp. 212–219.
- [8] Craig Gidney and Martin Ekerå. “How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits”. In: *Quantum* 5 (2021), p. 433.
- [9] Ryan Babbush et al. “Focus beyond quadratic speedups for error-corrected quantum advantage”. In: *PRX Quantum* 2.1 (2021), p. 010103.
- [10] Frank Arute et al. “Quantum supremacy using a programmable superconducting processor”. In: *Nature* 574.7779 (2019), pp. 505–510.
- [11] Youngseok Kim et al. “Evidence for the utility of quantum computing before fault tolerance”. In: *Nature* 618.7965 (2023), pp. 500–505.
- [12] Feng Pan, Keyang Chen, and Pan Zhang. “Solving the sampling problem of the sycamore quantum circuits”. In: *Physical Review Letters* 129.9 (2022), p. 090502.
- [13] Joseph Tindall et al. “Efficient tensor network simulation of IBM’s kicked Ising experiment”. In: *arXiv preprint arXiv:2306.14887* (2023).
- [14] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. “A quantum approximate optimization algorithm”. In: *arXiv preprint arXiv:1411.4028* (2014).
- [15] Dominic W Berry. “High-order quantum algorithm for solving linear differential equations”. In: *Journal of Physics A: Mathematical and Theoretical* 47.10 (2014), p. 105301.

- [16] Dominic W Berry et al. “Quantum algorithm for linear differential equations with exponentially improved dependence on precision”. In: *Communications in Mathematical Physics* 356 (2017), pp. 1057–1081.
- [17] Andrew M Childs and Jin-Peng Liu. “Quantum spectral methods for differential equations”. In: *Communications in Mathematical Physics* 375.2 (2020), pp. 1427–1457.
- [18] Michael Lubasch et al. “Variational quantum algorithms for nonlinear problems”. In: *Physical Review A* 101.1 (2020), p. 010301.
- [19] Nathan Wiebe, Ashish Kapoor, and Krysta Svore. “Quantum algorithms for nearest-neighbor methods for supervised and unsupervised learning”. In: *arXiv preprint arXiv:1401.2142* (2014).
- [20] Zhikuan Zhao et al. “Bayesian deep learning on a quantum computer”. In: *Quantum Machine Intelligence* 1 (2019), pp. 41–51.
- [21] Peter Wittek and Christian Gogolin. “Quantum enhanced inference in Markov logic networks”. In: *Scientific reports* 7.1 (2017), p. 45672.
- [22] Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme. “A rigorous and robust quantum speed-up in supervised machine learning”. In: *Nature Physics* 17.9 (2021), pp. 1013–1017.
- [23] Matthew B Hastings. “Classical and quantum algorithms for tensor principal component analysis”. In: *Quantum* 4 (2020), p. 237.
- [24] Dan Boneh and Richard J Lipton. “Quantum cryptanalysis of hidden linear functions”. In: *Advances in Cryptology—CRYPTO’95: 15th Annual International Cryptology Conference Santa Barbara, California, USA, August 27–31, 1995 Proceedings 15*. Springer. 1995, pp. 424–437.
- [25] Daniel R Simon. “On the power of quantum computation”. In: *SIAM journal on computing* 26.5 (1997), pp. 1474–1483.
- [26] Chris Lomont. “The hidden subgroup problem-review and open problems”. In: *arXiv preprint quant-ph/0411037* (2004).
- [27] Christoph Dürr et al. “Quantum query complexity of some graph problems”. In: *SIAM Journal on Computing* 35.6 (2006), pp. 1310–1328.
- [28] Frédéric Magniez, Miklos Santha, and Mario Szegedy. “Quantum algorithms for the triangle problem”. In: *SIAM Journal on Computing* 37.2 (2007), pp. 413–424.
- [29] Pawel Wocjan and Jon Yard. “The Jones polynomial: quantum algorithms and applications in quantum complexity theory”. In: *arXiv preprint quant-ph/0603069* (2006).
- [30] Yudong Cao et al. “Quantum chemistry in the age of quantum computing”. In: *Chemical reviews* 119.19 (2019), pp. 10856–10915.
- [31] Sam McArdle et al. “Quantum computational chemistry”. In: *Reviews of Modern Physics* 92.1 (2020), p. 015003.
- [32] Bela Bauer et al. “Quantum algorithms for quantum chemistry and quantum materials science”. In: *Chemical Reviews* 120.22 (2020), pp. 12685–12717.
- [33] Seth Lloyd. “Universal quantum simulators”. In: *Science* 273.5278 (1996), pp. 1073–1078.

- [34] Steven A Cuccaro et al. “A new quantum ripple-carry addition circuit”. In: *arXiv preprint quant-ph/0410184* (2004).
- [35] A Yu Kitaev. “Quantum measurements and the Abelian stabilizer problem”. In: *arXiv preprint quant-ph/9511026* (1995).
- [36] Michael A Nielsen and Isaac Chuang. *Quantum computation and quantum information*. 2002.
- [37] Seunghoon Lee et al. “Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry”. In: *Nature Communications* 14.1 (2023), p. 1952.
- [38] Dominic W Berry et al. “Qubitization of arbitrary basis quantum chemistry leveraging sparsity and low rank factorization”. In: *Quantum* 3 (2019), p. 208.
- [39] Joonho Lee et al. “Even more efficient quantum computations of chemistry through tensor hypercontraction”. In: *PRX Quantum* 2.3 (2021), p. 030305.
- [40] Nick S Blunt et al. “Perspective on the current state-of-the-art of quantum computing for drug discovery applications”. In: *Journal of Chemical Theory and Computation* 18.12 (2022), pp. 7001–7023.
- [41] Aleksei V Ivanov et al. “Quantum computation for periodic solids in second quantization”. In: *Physical Review Research* 5.1 (2023), p. 013200.
- [42] Isaac H Kim et al. “Fault-tolerant resource estimate for quantum chemical simulations: Case study on Li-ion battery electrolyte molecules”. In: *Physical Review Research* 4.2 (2022), p. 023019.
- [43] Modjtaba Shokrian Zini et al. “Quantum simulation of battery materials using ionic pseudopotentials”. In: *Quantum* 7 (2023), p. 1049.
- [44] Earl T Campbell, Barbara M Terhal, and Christophe Vuillot. “Roads towards fault-tolerant universal quantum computation”. In: *Nature* 549.7671 (2017), pp. 172–179.
- [45] Nicholas C Rubin et al. “Fault-tolerant quantum simulation of materials using Bloch orbitals”. In: *PRX Quantum* 4.4 (2023), p. 040303.
- [46] Joshua J Goings et al. “Reliably assessing the electronic structure of cytochrome p450 on today’s classical computers and tomorrow’s quantum computers”. In: *Proceedings of the National Academy of Sciences* 119.38 (2022), e2203533119.
- [47] Alberto Peruzzo et al. “A variational eigenvalue solver on a photonic quantum processor”. In: *Nature communications* 5.1 (2014), p. 4213.
- [48] Andrew G Taube and Rodney J Bartlett. “New perspectives on unitary coupled-cluster theory”. In: *International journal of quantum chemistry* 106.15 (2006), pp. 3393–3401.
- [49] Ilias Magoulas and Francesco A Evangelista. “Unitary Coupled Cluster: Seizing the Quantum Moment”. In: *arXiv preprint arXiv:2305.01097* (2023).
- [50] Edward Farhi and Aram W Harrow. “Quantum supremacy through the quantum approximate optimization algorithm”. In: *arXiv preprint arXiv:1602.07674* (2016).
- [51] Bela Bauer et al. “Hybrid quantum-classical approach to correlated materials”. In: *Physical Review X* 6.3 (2016), p. 031045.

- [52] Juha M Kreula et al. “Few-qubit quantum-classical simulation of strongly correlated lattice fermions”. In: *EPJ Quantum Technology* 3.11 (2016).
- [53] JM Kreula, Stephen R Clark, and D Jaksch. “Non-linear quantum-classical scheme to simulate non-equilibrium strongly correlated fermionic many-body dynamics”. In: *Scientific reports* 6 (2016), p. 32940.
- [54] Stasja Stanisic et al. “Observing ground-state properties of the Fermi-Hubbard model using a scalable algorithm on a quantum computer”. In: *Nature Communications* 13.1 (2022), p. 5743.
- [55] Laura Clinton et al. “Towards near-term quantum simulation of materials”. In: *arXiv preprint arXiv:2205.15256* (2022).
- [56] Seth Lloyd et al. “Quantum algorithm for nonlinear differential equations”. In: *arXiv preprint arXiv:2011.06571* (2020).
- [57] Ben Jaderberg et al. “Let Quantum Neural Networks Choose Their Own Frequencies”. In: *arXiv preprint arXiv:2309.03279* (2023).
- [58] Edward Farhi and Hartmut Neven. “Classification with quantum neural networks on near term processors”. In: *arXiv preprint arXiv:1802.06002* (2018).
- [59] Ben Jaderberg et al. “Quantum self-supervised learning”. In: *Quantum Science and Technology* 7.3 (2022), p. 035005.
- [60] Sam McArdle et al. “Digital quantum simulation of molecular vibrations”. In: *Chemical science* 10.22 (2019), pp. 5725–5735.
- [61] Chae-Yeun Park and Nathan Killoran. “Hamiltonian variational ansatz without barren plateaus”. In: *arXiv preprint arXiv:2302.08529* (2023).
- [62] Abhinav Kandala et al. “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets”. In: *nature* 549.7671 (2017), pp. 242–246.
- [63] Sukin Sim, Peter D Johnson, and Alán Aspuru-Guzik. “Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms”. In: *Advanced Quantum Technologies* 2.12 (2019), p. 1900070.
- [64] Harper R Grimsley et al. “An adaptive variational algorithm for exact molecular simulations on a quantum computer”. In: *Nature communications* 10.1 (2019), p. 3007.
- [65] Ho Lun Tang et al. “qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor”. In: *PRX Quantum* 2.2 (2021), p. 020310.
- [66] Hugh GA Burton et al. “Exact electronic states with shallow quantum circuits from global optimisation”. In: *npj Quantum Information* 9.1 (2023), p. 75.
- [67] Gavin E Crooks. “Gradients of parameterized quantum gates using the parameter-shift rule and gate decomposition”. In: *arXiv preprint arXiv:1905.13311* (2019).
- [68] David Wierichs et al. “General parameter-shift rules for quantum gradients”. In: *Quantum* 6 (2022), p. 677.

- [69] Ikko Hamamura and Takashi Imamichi. “Efficient evaluation of quantum observables using entangled measurements”. In: *npj Quantum Information* 6.1 (2020), p. 56.
- [70] Richard Karp. “Reducibility Among Combinatorial Problems”. In: *Proceedings of a Symposium on the Complexity of Computer Computations*. Ed. by R. E. Miller and J. W. Thatcher. Plenum Press, 1972, pp. 85–103.
- [71] Pranav Gokhale et al. “Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families”. In: *arXiv preprint arXiv:1907.13623* (2019).
- [72] Ophelia Crawford et al. “Efficient quantum measurement of Pauli operators in the presence of finite sampling error”. In: *Quantum* 5 (Jan. 2021), p. 385.
- [73] Jarrod R McClean et al. “Exploiting locality in quantum computation for quantum chemistry”. In: *The journal of physical chemistry letters* 5.24 (2014), pp. 4368–4380.
- [74] Giacomo Torlai et al. “Precise measurement of quantum observables with neural-network estimators”. In: *Physical Review Research* 2.2 (2020), p. 022060.
- [75] Hsin-Yuan Huang, Richard Kueng, and John Preskill. “Predicting many properties of a quantum system from very few measurements”. In: *Nature Physics* 16.10 (2020), pp. 1050–1057.
- [76] Jarrod R McClean et al. “Barren plateaus in quantum neural network training landscapes”. In: *Nature communications* 9 (2018), p. 4812.
- [77] Marco Cerezo et al. “Cost function dependent barren plateaus in shallow parametrized quantum circuits”. In: *Nature communications* 12 (2021), p. 1791.
- [78] AV Uvarov and Jacob D Biamonte. “On barren plateaus and cost function locality in variational quantum algorithms”. In: *Journal of Physics A: Mathematical and Theoretical* 54.24 (2021), p. 245301.
- [79] Samson Wang et al. “Noise-induced barren plateaus in variational quantum algorithms”. In: *Nature Communications* 12 (2021), p. 6961.
- [80] Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe. “Entanglement-Induced Barren Plateaus”. In: *PRX Quantum* 2.4 (2021), p. 040316.
- [81] Zoë Holmes et al. “Connecting ansatz expressibility to gradient magnitudes and barren plateaus”. In: *PRX Quantum* 3.1 (2022), p. 010313.
- [82] Enrico Fontana et al. “The Adjoint Is All You Need: Characterizing Barren Plateaus in Quantum Ansätze”. In: *arXiv preprint arXiv:2309.07902* (2023).
- [83] Michael Ragone et al. “A Unified Theory of Barren Plateaus for Deep Parametrized Quantum Circuits”. In: *arXiv preprint arXiv:2309.09342* (2023).
- [84] Arthur Pesah et al. “Absence of barren plateaus in quantum convolutional neural networks”. In: *Physical Review X* 11.4 (2021), p. 041011.
- [85] Kishor Bharti and Tobias Haug. “Iterative quantum-assisted eigensolver”. In: *Physical Review A* 104.5 (2021), p. L050401.
- [86] Roeland Wiersema et al. “Exploring entanglement and optimization within the hamiltonian variational ansatz”. In: *PRX Quantum* 1.2 (2020), p. 020319.

- [87] Hao-Kai Zhang, Shuo Liu, and Shi-Xin Zhang. “Absence of barren plateaus in finite local-depth circuits with long-range entanglement”. In: *arXiv preprint arXiv:2311.01393* (2023).
- [88] Edward Grant et al. “An initialization strategy for addressing barren plateaus in parametrized quantum circuits”. In: *Quantum* 3 (2019), p. 214.
- [89] Guillaume Verdon et al. “Learning to learn with quantum neural networks via classical neural networks”. In: *arXiv preprint arXiv:1907.05415* (2019).
- [90] Huan-Yu Liu et al. “A Parameter Initialization Method for Variational Quantum Algorithms to Mitigate Barren Plateaus Based on Transfer Learning”. In: *arXiv preprint arXiv:2112.10952* (2021).
- [91] Harper R Grimsley et al. “Adaptive, problem-tailored variational quantum eigensolver mitigates rough parameter landscapes and barren plateaus”. In: *npj Quantum Information* 9.1 (2023), p. 19.
- [92] Abhinav Anand, Matthias Degroote, and Alán Aspuru-Guzik. “Natural evolutionary strategies for variational quantum computation”. In: *Machine Learning: Science and Technology* (2021).
- [93] Taylor L Patti et al. “Entanglement devised barren plateau mitigation”. In: *Physical Review Research* 3.3 (2021), p. 033090.
- [94] Stefan H. Sack et al. “Avoiding barren plateaus using classical shadows”. In: *arXiv preprint arXiv:2201.08194* (2022).
- [95] Antonio Sannia et al. “Engineered dissipation to mitigate barren plateaus”. In: *arXiv preprint arXiv:2310.15037* (2023).
- [96] Bálint Koczor and Simon C Benjamin. “Quantum natural gradient generalised to non-unitary circuits”. In: *arXiv preprint arXiv:1912.08660* 10 (2019).
- [97] James Stokes et al. “Quantum natural gradient”. In: *Quantum* 4 (2020), p. 269.
- [98] Gregory Boyd and Bálint Koczor. “Training variational quantum circuits with CoVaR: covariance root finding with classical shadows”. In: *arXiv preprint arXiv:2204.08494* (2022).
- [99] Dave Wecker, Matthew B Hastings, and Matthias Troyer. “Progress towards practical quantum variational algorithms”. In: *Physical Review A* 92.4 (2015), p. 042303.
- [100] Michael Kuühn et al. “Accuracy and resource estimations for quantum chemistry on a near-term quantum computer”. In: *Journal of chemical theory and computation* 15.9 (2019), pp. 4764–4780.
- [101] Jérôme F Gonthier et al. “Identifying challenges towards practical quantum advantage through resource estimation: the measurement roadblock in the variational quantum eigensolver”. In: *arXiv preprint arXiv:2012.04001* (2020), p. 8.
- [102] Jules Tilly et al. “The variational quantum eigensolver: a review of methods and best practices”. In: *Physics Reports* 986 (2022), pp. 1–128.
- [103] Jarrod R McClean et al. “Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states”. In: *Physical Review A* 95.4 (2017), p. 042308.

- [104] Robert M Parrish and Peter L McMahon. “Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation”. In: *arXiv preprint arXiv:1909.08925* (2019).
- [105] Nicholas H Stair, Renke Huang, and Francesco A Evangelista. “A multireference quantum Krylov algorithm for strongly correlated electrons”. In: *Journal of chemical theory and computation* 16.4 (2020), pp. 2236–2245.
- [106] Cristian L Cortes and Stephen K Gray. “Quantum Krylov subspace algorithms for ground-and excited-state energy estimation”. In: *Physical Review A* 105.2 (2022), p. 022417.
- [107] Nicholas H Stair et al. “Stochastic quantum Krylov protocol with double-factorized Hamiltonians”. In: *Physical Review A* 107.3 (2023), p. 032414.
- [108] Mario Motta et al. “Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution”. In: *Nature Physics* 16.2 (2020), pp. 205–210.
- [109] Kübra Yeter-Aydeniz, Raphael C Pooser, and George Siopsis. “Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and Lanczos algorithms”. In: *npj Quantum Information* 6.1 (2020), p. 63.
- [110] Tyler Takeshita et al. “Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources”. In: *Physical Review X* 10.1 (2020), p. 011004.
- [111] James I Colless et al. “Computation of molecular spectra on a quantum processor with an error-resilient algorithm”. In: *Physical Review X* 8.1 (2018), p. 011021.
- [112] William J Huggins et al. “A non-orthogonal variational quantum eigensolver”. In: *New Journal of Physics* 22.7 (2020), p. 073009.
- [113] Adriano Barenco et al. “Stabilization of quantum computations by symmetrization”. In: *SIAM Journal on Computing* 26.5 (1997), pp. 1541–1557.
- [114] Harry Buhrman et al. “Quantum fingerprinting”. In: *Physical Review Letters* 87.16 (2001), p. 167902.
- [115] Cornelius Lanczos. “An iteration method for the solution of the eigenvalue problem of linear differential and integral operators”. In: (1950).
- [116] Jane K Cullum and Ralph A Willoughby. *Lanczos algorithms for large symmetric eigenvalue computations: Vol. I: Theory*. SIAM, 2002.
- [117] Christopher Conway Paige. “The computation of eigenvalues and eigenvectors of very large sparse matrices.” PhD thesis. University of London, 1971.
- [118] Yousef Saad. “On the rates of convergence of the Lanczos and the block-Lanczos methods”. In: *SIAM Journal on Numerical Analysis* 17.5 (1980), pp. 687–706.
- [119] Kazuhiro Seki and Seiji Yunoki. “Quantum power method by a superposition of time-evolved states”. In: *PRX Quantum* 2.1 (2021), p. 010333.
- [120] William Kirby, Mario Motta, and Antonio Mezzacapo. “Exact and efficient Lanczos method on a quantum computer”. In: *Quantum* 7 (2023), p. 1018.

- [121] Bernhard Beckermann and Alex Townsend. “On the singular values of matrices with displacement structure”. In: *SIAM Journal on Matrix Analysis and Applications* 38.4 (2017), pp. 1227–1248.
- [122] Ethan N Epperly, Lin Lin, and Yuji Nakatsukasa. “A theory of quantum subspace diagonalization”. In: *SIAM Journal on Matrix Analysis and Applications* 43.3 (2022), pp. 1263–1290.
- [123] Gwonhak Lee, Dongkeun Lee, and Joonsuk Huh. “Sampling Error Analysis in Quantum Krylov Subspace Diagonalization”. In: *arXiv preprint arXiv:2307.16279* (2023).
- [124] Zongkang Zhang et al. “Measurement-efficient quantum Krylov subspace diagonalisation”. In: *arXiv preprint arXiv:2301.13353* (2023).
- [125] Tom O’Leary et al. “Partitioned quantum subspace expansion”. In: *In preparation* (2023).
- [126] Miroslav Urbanek et al. “Chemistry on quantum computers with virtual quantum subspace expansion”. In: *Journal of chemical theory and computation* 16.9 (2020), pp. 5425–5431.
- [127] Bencheng Huang et al. “Quantum simulations of Fermionic Hamiltonians with efficient encoding and ansatz schemes”. In: *Journal of Chemical Theory and Computation* 19.5 (2023), pp. 1487–1498.
- [128] Nobuyuki Yoshioka et al. “Hunting for quantum-classical crossover in condensed matter problems”. In: *arXiv preprint arXiv:2210.14109* (2022).
- [129] Andrew Jones. “Quantum Drude oscillators for accurate many-body intermolecular forces”. PhD thesis. University of Edinburgh, 2010.
- [130] Vlad P Sokhan et al. “Signature properties of water: Their molecular electronic origins”. In: *Proceedings of the National Academy of Sciences* 112.20 (2015), pp. 6341–6346.
- [131] Julian Schwinger. “Gauge invariance and mass. II”. In: *Physical Review* 128.5 (1962), p. 2425.
- [132] Kazuya Saito. *Chemical Physics of Molecular Condensed Matter*. Springer, 2020.
- [133] Fritz London. “Zur theorie und systematik der molekularkräfte”. In: *Zeitschrift für Physik* 63.3 (1930), pp. 245–279.
- [134] Anthony M Reilly and Alexandre Tkatchenko. “van der Waals dispersion interactions in molecular materials: beyond pairwise additivity”. In: *Chemical Science* 6.6 (2015), pp. 3289–3301.
- [135] Yaroslav V Shtogun and Lilia M Woods. “Many-body van der Waals interactions between graphitic nanostructures”. In: *The Journal of Physical Chemistry Letters* 1.9 (2010), pp. 1356–1362.
- [136] Ewa Pastorczyk et al. “Intricacies of van der Waals interactions in systems with elongated bonds revealed by electron-groups embedding and high-level coupled-cluster approaches”. In: *Journal of chemical theory and computation* 13.11 (2017), pp. 5404–5419.

- [137] Jan Hermann, Robert A DiStasio Jr, and Alexandre Tkatchenko. “First-principles models for van der Waals interactions in molecules and materials: Concepts, theory, and applications”. In: *Chemical Reviews* 117.6 (2017), pp. 4714–4758.
- [138] Hans Hellmann. “Einführung in die Quantenchemie”. In: *Franz Deuticke, Leipzig* 285 (1937).
- [139] Richard Phillips Feynman. “Forces in molecules”. In: *Physical review* 56.4 (1939), p. 340.
- [140] Max Born and Robert Oppenheimer. “Zur quantentheorie der molekeln”. In: *Annalen der physik* 389.20 (1927), pp. 457–484.
- [141] Roger H French et al. “Long range interactions in nanoscale science”. In: *Reviews of Modern Physics* 82.2 (2010), p. 1887.
- [142] O Anatole Von Lilienfeld et al. “Optimization of effective atom centered potentials for London dispersion forces in density functional theory”. In: *Physical review letters* 93.15 (2004), p. 153004.
- [143] Alexandre Tkatchenko and Matthias Scheffler. “Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data”. In: *Physical review letters* 102.7 (2009), p. 073005.
- [144] Stefan Grimme. “Density functional theory with London dispersion corrections”. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 1.2 (2011), pp. 211–228.
- [145] John C Slater. “The theory of complex spectra”. In: *Physical Review* 34.10 (1929), p. 1293.
- [146] S Francis Boys. “Electronic wave functions-I. A general method of calculation for the stationary states of any molecular system”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 200.1063 (1950), pp. 542–554.
- [147] Peter G Szalay et al. “Multiconfiguration self-consistent field and multireference configuration interaction methods and applications”. In: *Chemical reviews* 112.1 (2012), pp. 108–181.
- [148] Jiří Čížek. “On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods”. In: *The Journal of Chemical Physics* 45.11 (1966), pp. 4256–4266.
- [149] Rodney J Bartlett and Monika Musiał. “Coupled-cluster theory in quantum chemistry”. In: *Reviews of Modern Physics* 79.1 (2007), p. 291.
- [150] Piotr A Pieniazek et al. “Benchmark full configuration interaction and equation-of-motion coupled-cluster model with single and double substitutions for ionized systems results for prototypical charge transfer systems: Noncovalent ionized dimers”. In: *The Journal of chemical physics* 127.16 (2007), p. 164110.
- [151] Vytor Oliveira and Elfi Kraka. “Systematic coupled cluster study of noncovalent interactions involving halogens, chalcogens, and pnictogens”. In: *The Journal of Physical Chemistry A* 121.49 (2017), pp. 9544–9556.

- [152] Giovanni Bistoni. “Finding chemical concepts in the Hilbert space: Coupled cluster analyses of noncovalent interactions”. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 10.3 (2020), e1442.
- [153] Alain Delgado et al. “How to simulate key properties of lithium-ion batteries with a fault-tolerant quantum computer”. In: *arXiv preprint arXiv:2204.11890* (2022).
- [154] Jarrod R McClean et al. “The theory of variational hybrid quantum-classical algorithms”. In: *New Journal of Physics* 18.2 (2016), p. 023023.
- [155] John Preskill. “Quantum Computing in the NISQ era and beyond”. In: *Quantum* 2 (2018), p. 79.
- [156] Marco Cerezo et al. “Variational quantum algorithms”. In: *Nature Reviews Physics* 3.9 (2021), pp. 625–644.
- [157] Peter JJ O’Malley et al. “Scalable quantum simulation of molecular energies”. In: *Physical Review X* 6.3 (2016), p. 031007.
- [158] Yunseong Nam et al. “Ground-state energy estimation of the water molecule on a trapped-ion quantum computer”. In: *npj Quantum Information* 6 (2020), p. 33.
- [159] Pauline J Ollitrault et al. “Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor”. In: *Physical Review Research* 2.4 (2020), p. 043140.
- [160] Frank Arute et al. “Hartree-Fock on a superconducting qubit quantum computer”. In: *Science* 369.6507 (2020), pp. 1084–1089.
- [161] Anton Nykänen et al. “Toward Accurate Post-Born-Oppenheimer Molecular Simulations on Quantum Computers: An Adaptive Variational Eigensolver with Nuclear-Electronic Frozen Natural Orbitals”. In: *arXiv preprint arXiv:2310.01302* (2023).
- [162] Pauline J Ollitrault et al. “Hardware efficient quantum algorithms for vibrational structure calculations”. In: *Chemical science* 11.26 (2020), pp. 6842–6855.
- [163] Ivan Rungger et al. “Dynamical mean field theory algorithm and experiment on quantum computers”. In: *arXiv preprint arXiv:1910.04735* (2019).
- [164] Ben Jaderberg et al. “Minimum hardware requirements for hybrid quantum–classical DMFT”. In: *Quantum Science and Technology* 5.3 (2020), p. 034015.
- [165] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2014.
- [166] C David Sherrill. “Computational scaling of the configuration interaction method with system size”. In: *Quantum Chemistry Lecture Notes, School of Chem. and Biochem., Georgia Inst. of Technol.* (1996).
- [167] C David Sherrill. “Wavefunction theory approaches to noncovalent interactions”. In: *Non-Covalent Interactions in Quantum Chemistry and Physics*. Ed. by Alberto Otero de la Roza and Gino A. DiLabio. Elsevier, 2017, pp. 137–168.
- [168] FS Cipcigan et al. “Electronic coarse graining: Predictive atomistic modeling of condensed matter”. In: *Reviews of Modern Physics* 91.2 (2019), p. 025003.
- [169] Flaviu Cipcigan et al. “Structure and hydrogen bonding at the limits of liquid water stability”. In: *Scientific reports* 8 (2018), p. 1718.

- [170] VP Sokhan et al. “Molecular-scale remnants of the liquid-gas transition in supercritical polar fluids”. In: *Physical review letters* 115.11 (2015), p. 117801.
- [171] Flaviu S Cipcigan et al. “Electronic coarse graining enhances the predictive power of molecular simulation allowing challenges in water physics to be addressed”. In: *Journal of Computational Physics* 326 (2016), pp. 222–233.
- [172] Pranav Gokhale et al. “ $O(N^3)$ Measurement Cost for Variational Quantum Eigensolver on Molecular Hamiltonians”. In: *IEEE Transactions on Quantum Engineering* 1 (2020).
- [173] Marise JE Westbroek et al. “User’s guide to Monte Carlo methods for evaluating path integrals”. In: *American Journal of Physics* 86.4 (2018), pp. 293–304.
- [174] A Jones et al. “Electronically coarse-grained model for water”. In: *Physical review letters* 110.22 (2013), p. 227801.
- [175] Andrew P Jones et al. “Quantum Drude oscillator model of atoms and molecules: Many-body polarization and dispersion interactions for atomistic simulation”. In: *Physical Review B* 87.14 (2013), p. 144103.
- [176] Nicolas PD Sawaya et al. “Resource-efficient digital quantum simulation of d-level systems for photonic, vibrational, and spin-s Hamiltonians”. In: *npj Quantum Information* 6.1 (2020), pp. 1–13.
- [177] Eugene F Dumitrescu et al. “Cloud quantum computing of an atomic nucleus”. In: *Physical review letters* 120.21 (2018), p. 210501.
- [178] Andrew Jena, Scott Genin, and Michele Mosca. “Pauli partitioning with respect to gate sets”. In: *arXiv preprint arXiv:1907.07859* (2019).
- [179] Vladyslav Verteletskyi, Tzu-Ching Yen, and Artur F Izmaylov. “Measurement optimization in the variational quantum eigensolver using a minimum clique cover”. In: *The Journal of chemical physics* 152.12 (2020), p. 124114.
- [180] Artur F Izmaylov et al. “Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method”. In: *Journal of chemical theory and computation* 16.1 (2020), pp. 190–195.
- [181] Andrew Zhao et al. “Measurement reduction in variational quantum algorithms”. In: *Physical Review A* 101.6 (2020), p. 062322.
- [182] Tzu-Ching Yen, Vladyslav Verteletskyi, and Artur F Izmaylov. “Measuring all compatible operators in one series of single-qubit measurements using unitary transformations”. In: *Journal of chemical theory and computation* 16.4 (2020), pp. 2400–2409.
- [183] John Watrous. *The theory of quantum information*. Cambridge university press, 2018.
- [184] George Maroulis et al. “Electrooptical properties and molecular polarization of iodine, I₂”. In: *The Journal of Physical Chemistry A* 101.5 (1997), pp. 953–956.
- [185] Sekika Yamamoto et al. “Pressure dependence of local structure in liquid carbon disulfide”. In: *The journal of chemical physics* 124.14 (2006), p. 144511.
- [186] Bernhard Schropp and Paul Tavan. “The polarizability of point-polarizable water models: density functional theory/molecular mechanics results”. In: *The Journal of Physical Chemistry B* 112.19 (2008), pp. 6233–6240.

- [187] Yongli Huang et al. “Size, separation, structural order and mass density of molecules packing in water and ice”. In: *Scientific reports* 3.1 (2013).
- [188] Farrokh Vatan and Colin Williams. “Optimal quantum circuits for general two-qubit gates”. In: *Physical Review A* 69.3 (2004), p. 032315.
- [189] Jonathan Romero et al. “Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz”. In: *Quantum Science and Technology* 4.1 (2018), p. 014008.
- [190] Marco Majland and Nikolaj Thomas Zinner. “Resource-efficient encoding algorithm for variational bosonic quantum simulations”. In: *arXiv preprint arXiv:2102.11886* (2021).
- [191] Ove Christiansen. “A second quantization formulation of multimode dynamics”. In: *The Journal of chemical physics* 120.5 (2004), pp. 2140–2148.
- [192] Ove Christiansen. “Vibrational coupled cluster theory”. In: *The Journal of chemical physics* 120.5 (2004), pp. 2149–2159.
- [193] Diederik P Kingma and Jimmy Ba. “Adam: A method for stochastic optimization”. In: *arXiv preprint arXiv:1412.6980* (2014).
- [194] Kosuke Mitarai et al. “Quantum circuit learning”. In: *Physical Review A* 98.3 (2018), p. 032309.
- [195] Robert W Boyd. *Nonlinear optics*. Academic press, 2020.
- [196] Nicolas PD Sawaya and Joonsuk Huh. “Quantum algorithm for calculating molecular vibronic spectra”. In: *The journal of physical chemistry letters* 10.13 (2019), pp. 3586–3591.
- [197] Kristan Temme, Sergey Bravyi, and Jay M Gambetta. “Error mitigation for short-depth quantum circuits”. In: *Physical review letters* 119.18 (2017), p. 180509.
- [198] Ying Li and Simon C Benjamin. “Efficient variational quantum simulator incorporating active error minimization”. In: *Physical Review X* 7.2 (2017), p. 021050.
- [199] Philippe Suchsland et al. “Algorithmic error mitigation scheme for current quantum processors”. In: *Quantum* 5 (2021), p. 492.
- [200] BM Axilrod and Ei Teller. “Interaction of the van der Waals type between three atoms”. In: *The Journal of Chemical Physics* 11.6 (1943), pp. 299–300.
- [201] Lennart Bittel and Martin Kliesch. “Training variational quantum algorithms is NP-hard”. In: *Physical Review Letters* 127.12 (2021), p. 120502.
- [202] Avrim Blum and Ronald L Rivest. “Training a 3-node neural network is NP-complete”. In: *Proceedings of the 1st International Conference on Neural Information Processing Systems*. 1988, pp. 494–501.
- [203] Mark Fingerhuth, Tomáš Babej, et al. “A quantum alternating operator ansatz with hard and soft constraints for lattice protein folding”. In: *arXiv preprint arXiv:1810.13411* (2018).
- [204] Anton Robert et al. “Resource-efficient quantum algorithm for protein folding”. In: *npj Quantum Information* 7 (2021), p. 49.

- [205] István Montvay and Gernot Münster. *Quantum fields on a lattice*. Cambridge University Press, 1994.
- [206] Jan Smit. *Introduction to quantum fields on a lattice*. Cambridge University Press, 2002.
- [207] Stephen P Jordan, Keith SM Lee, and John Preskill. “Quantum algorithms for quantum field theories”. In: *Science* 336.6085 (2012), pp. 1130–1133.
- [208] Andy CY Li et al. “Simulating scalar field theories on quantum computers with limited resources”. In: *Physical Review A* 107.3 (2023), p. 032603.
- [209] Matthew D Schwartz. *Quantum field theory and the standard model*. Cambridge university press, 2014.
- [210] David Bailin and Alexander Love. *Cosmology in gauge field theory and string theory*. Taylor & Francis, 2004.
- [211] Eduardo Fradkin. *Field theories of condensed matter physics*. Cambridge University Press, 2013.
- [212] Kenneth G Wilson. “Confinement of quarks”. In: *Physical review D* 10.8 (1974), p. 2445.
- [213] Michael Creutz, Laurence Jacobs, and Claudio Rebbi. “Monte Carlo computations in lattice gauge theories”. In: *Physics Reports* 95.4 (1983), pp. 201–282.
- [214] Matthias Troyer and Uwe-Jens Wiese. “Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations”. In: *Physical review letters* 94.17 (2005), p. 170201.
- [215] Anders W Sandvik and Juhani Kurkijärvi. “Quantum Monte Carlo simulation method for spin systems”. In: *Physical Review B* 43.7 (1991), p. 5950.
- [216] Qi-Zhou Chen and Xiang-Qian Luo. “Chiral-symmetry breaking in the Schwinger model with Wilson fermions”. In: *Physical Review D* 42.4 (1990), p. 1293.
- [217] Kirill Melnikov and Marvin Weinstein. “Lattice Schwinger model: Confinement, anomalies, chiral fermions, and all that”. In: *Physical Review D* 62.9 (2000), p. 094504.
- [218] Mari Carmen Bañuls et al. “The mass spectrum of the Schwinger model with matrix product states”. In: *Journal of High Energy Physics* 2013.11 (2013), pp. 1–21.
- [219] Boye Buyens et al. “Tensor networks for gauge field theories”. In: *arXiv preprint arXiv:1511.04288* (2015).
- [220] Mari Carmen Bañuls et al. “Tensor networks and their use for lattice gauge theories”. In: *arXiv preprint arXiv:1810.12838* (2018).
- [221] Lena Funcke, Karl Jansen, and Stefan Kühn. “Exploring the CP-Violating Dashen Phase in the Schwinger Model with Tensor Networks”. In: *arXiv preprint arXiv:2303.03799* (2023).
- [222] Mari Carmen Bañuls et al. “Chiral condensate in the Schwinger model with matrix product operators”. In: *Physical Review D* 93.9 (2016), p. 094512.
- [223] Thomas Pichler et al. “Real-time dynamics in U(1) lattice gauge theories with tensor networks”. In: *Physical Review X* 6.1 (2016), p. 011023.

- [224] Boye Buyens et al. “Real-time simulation of the Schwinger effect with matrix product states”. In: *Physical Review D* 96.11 (2017), p. 114501.
- [225] Mari Carmen Bañuls et al. “Efficient basis formulation for (1+1)-dimensional SU(2) lattice gauge theory: Spectral calculations with matrix product states”. In: *Physical Review X* 7.4 (2017), p. 041046.
- [226] Kai Zapp and Román Orús. “Tensor network simulation of QED on infinite lattices: Learning from (1+1)d, and prospects for (2+1)d”. In: *Physical Review D* 95.11 (2017), p. 114508.
- [227] Christian W Bauer et al. “Quantum simulation for high-energy physics”. In: *PRX Quantum* 4.2 (2023), p. 027001.
- [228] Mari Carmen Banuls et al. “Simulating lattice gauge theories within quantum technologies”. In: *The European physical journal D* 74 (2020), pp. 1–42.
- [229] Alexander F Shaw et al. “Quantum algorithms for simulating the lattice schwinger model”. In: *Quantum* 4 (2020), p. 306.
- [230] Guang Hao Low and Isaac L Chuang. “Hamiltonian simulation by qubitization”. In: *Quantum* 3 (2019), p. 163.
- [231] Earl Campbell. “Random compiler for fast Hamiltonian simulation”. In: *Physical review letters* 123.7 (2019), p. 070503.
- [232] Nhung H Nguyen et al. “Digital quantum simulation of the schwinger model and symmetry protection with trapped ions”. In: *PRX Quantum* 3.2 (2022), p. 020324.
- [233] Wibe A de Jong et al. “Quantum simulation of nonequilibrium dynamics and thermalization in the Schwinger model”. In: *Physical Review D* 106.5 (2022), p. 054508.
- [234] Natalie Klco et al. “Quantum-classical computation of Schwinger model dynamics using quantum computers”. In: *Physical Review A* 98.3 (2018), p. 032331.
- [235] Christian Kokail et al. “Self-verifying variational quantum simulation of lattice models”. In: *Nature* 569.7756 (2019), pp. 355–360.
- [236] A Avkhadiiev, PE Shanahan, and RD Young. “Accelerating lattice quantum field theory calculations via interpolator optimization using noisy intermediate-scale quantum computing”. In: *Physical review letters* 124.8 (2020), p. 080501.
- [237] Ryan R Ferguson et al. “Measurement-based variational quantum eigensolver”. In: *Physical review letters* 126.22 (2021), p. 220501.
- [238] Shane Thompson and George Siopsis. “Quantum computation of phase transition in the massive Schwinger model”. In: *Quantum Science and Technology* 7.3 (2022), p. 035001.
- [239] Lena Funcke et al. “Exploring the phase structure of the multi-flavor Schwinger model with quantum computing”. In: *arXiv preprint arXiv:2211.13020* (2022).
- [240] Roland C Farrell et al. “Scalable Circuits for Preparing Ground States on Digital Quantum Computers: The Schwinger Model Vacuum on 100 Qubits”. In: *arXiv preprint arXiv:2308.04481* (2023).
- [241] Takis Angelides et al. “First-Order Phase Transition of the Schwinger Model with a Quantum Computer”. In: *arXiv preprint arXiv: 2312.12831* (2023).

- [242] Stefan Kühn, J Ignacio Cirac, and Mari-Carmen Bañuls. “Quantum simulation of the Schwinger model: A study of feasibility”. In: *Physical Review A* 90.4 (2014), p. 042305.
- [243] Philipp Hauke et al. “Quantum simulation of a lattice Schwinger model in a chain of trapped ions”. In: *Physical Review X* 3.4 (2013), p. 041018.
- [244] Michael Meth et al. “Simulating 2D lattice gauge theories on a qudit quantum computer”. In: *arXiv preprint arXiv:2310.12110* (2023).
- [245] Giulia Mazzola et al. “Gauge-invariant quantum circuits for U (1) and Yang-Mills lattice gauge theories”. In: *Physical Review Research* 3.4 (2021), p. 043209.
- [246] Danny Paulson et al. “Simulating 2d effects in lattice gauge theories on a quantum computer”. In: *PRX Quantum* 2.3 (2021), p. 030334.
- [247] Luca Lumia et al. “Two-Dimensional Z₂ Lattice Gauge Theory on a Near-Term Quantum Simulator: Variational Quantum Optimization, Confinement, and Topological Order”. In: *PRX Quantum* 3.2 (2022), p. 020320.
- [248] Jarrod R McClean et al. “Decoding quantum errors with subspace expansions”. In: *Nature communications* 11.1 (2020), p. 636.
- [249] Tom Banks, Leonard Susskind, and John Kogut. “Strong-coupling calculations of lattice gauge theories: (1+1)-dimensional exercises”. In: *Physical Review D* 13.4 (1976), p. 1043.
- [250] CJ Hamer, Zheng Weihong, and J Oitmaa. “Series expansions for the massive Schwinger model in Hamiltonian lattice theory”. In: *Physical Review D* 56.1 (1997), p. 55.
- [251] Yu Tong et al. “Provably accurate simulation of gauge theories and bosonic systems”. In: *Quantum* 6 (2022), p. 816.
- [252] Th Holstein. “Studies of polaron motion: Part I. The molecular-crystal model”. In: *Annals of physics* 8.3 (1959), pp. 325–342.
- [253] Herbert Fröhlich. “Electrons in lattice fields”. In: *Advances in Physics* 3.11 (1954), pp. 325–361.
- [254] Istvan Montvay and Gernot Münster. “Scalar fields”. In: *Quantum Fields on a Lattice*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1994, pp. 54–90.
- [255] John M Martyn et al. “Grand unification of quantum algorithms”. In: *PRX Quantum* 2.4 (2021), p. 040203.
- [256] Yuta Kikuchi et al. “Realization of quantum signal processing on a noisy quantum computer”. In: *npj Quantum Information* 9.1 (2023), p. 93.
- [257] András Gilyén et al. “Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics”. In: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. 2019, pp. 193–204.
- [258] <https://doi.org/10.5281/zenodo.22558>.
- [259] Neil J Ross and Peter Selinger. “Optimal ancilla-free Clifford+T approximation of z-rotations”. In: *arXiv preprint arXiv:1403.2975* (2014).

- [260] Roberto Oliveira and Barbara M Terhal. “The complexity of quantum spin systems on a two-dimensional square lattice”. In: *arXiv preprint quant-ph/0504050* (2005).
- [261] Yi-Kai Liu, Matthias Christandl, and Frank Verstraete. “Quantum computational complexity of the N-representability problem: QMA complete”. In: *Physical review letters* 98.11 (2007), p. 110503.
- [262] Norbert Schuch and Frank Verstraete. “Computational complexity of interacting electrons and fundamental limitations of density functional theory”. In: *Nature physics* 5.10 (2009), pp. 732–735.
- [263] Adam D Bookatz. “QMA-complete problems”. In: *arXiv preprint arXiv:1212.6312* (2012).
- [264] Yuval R Sanders et al. “Compilation of fault-tolerant quantum heuristics for combinatorial optimization”. In: *PRX Quantum* 1.2 (2020), p. 020312.
- [265] Ewin Tang. “A quantum-inspired classical algorithm for recommendation systems”. In: *Proceedings of the 51st annual ACM SIGACT symposium on theory of computing*. 2019, pp. 217–228.
- [266] Ewin Tang. “Quantum principal component analysis only achieves an exponential speedup because of its state preparation assumptions”. In: *Physical Review Letters* 127.6 (2021), p. 060503.
- [267] Nai-Hui Chia et al. “Sampling-based sublinear low-rank matrix arithmetic framework for dequantizing quantum machine learning”. In: *Journal of the ACM* 69.5 (2022), pp. 1–72.
- [268] Andrei Alexandru et al. “Fuzzy Gauge Theory for Quantum Computers”. In: *arXiv preprint arXiv:2308.05253* (2023).
- [269] Benjamin Jaderberg et al. “Recompilation-enhanced simulation of electron–phonon dynamics on IBM quantum computers”. In: *New Journal of Physics* 24.9 (2022), p. 093017.
- [270] Yao Wang et al. “Phonon-mediated long-range attractive interaction in one-dimensional cuprates”. In: *Physical Review Letters* 127.19 (2021), p. 197003.
- [271] Lena Funcke et al. “Review on quantum computing for lattice field theory”. In: *arXiv preprint arXiv:2302.00467* (2023).
- [272] F London. “On Centers of van der Waals Attraction.” In: *The Journal of Physical Chemistry* 46.2 (1942), pp. 305–316.
- [273] Joseph O Hirschfelder et al. *Molecular theory of gases and liquids*. Vol. 165. Wiley New York, 1964.
- [274] Vitaly I Voloshin. *Introduction to graph and hypergraph theory*. Nova Science Publ., 2009.
- [275] Jacobus Hendricus Van Lint and Richard Michael Wilson. *A course in combinatorics*. Cambridge university press, 2001.
- [276] Qiskit contributors. *Qiskit: An Open-source Framework for Quantum Computing*. 2023.
- [277] Stefano Pirandola, Alessio Serafini, and Seth Lloyd. “Correlation matrices of two-mode bosonic systems”. In: *Physical Review A* 79.5 (2009), p. 052327.

- [278] Norbert Schuch, J Ignacio Cirac, and Michael M Wolf. “Quantum states on harmonic lattices”. In: *Communications in mathematical physics* 267.1 (2006), pp. 65–92.
- [279] Sam E. Ganis. “Notes on the Fibonacci Sequence”. In: *The American Mathematical Monthly* 66.2 (1959), pp. 129–130.
- [280] Yong He et al. “Decompositions of n-qubit Toffoli gates with linear circuit complexity”. In: *International Journal of Theoretical Physics* 56 (2017), pp. 2350–2361.
- [281] John A Smolin and David P DiVincenzo. “Five two-bit quantum gates are sufficient to implement the quantum Fredkin gate”. In: *Physical Review A* 53.4 (1996), p. 2855.
- [282] Jonathan M Baker et al. “Decomposing quantum generalized toffoli with an arbitrary number of ancilla”. In: *arXiv preprint arXiv:1904.01671* (2019).
- [283] SA Moses et al. “A race track trapped-ion quantum processor”. In: *arXiv preprint arXiv:2305.03828* (2023).
- [284] *Quantinuum System Model H2: Emulator data sheet, version 1.1*. www.quantinuum.com/hardware/h2. 2023.
- [285] Jwo-Sy Chen et al. “Benchmarking a trapped-ion quantum computer with 29 algorithmic qubits”. In: *arXiv preprint arXiv:2308.05071* (2023).
- [286] *IonQ Forte*. www.ionq.com/quantum-systems/forte. 2022.
- [287] IBM Kyiv processor, accessed 16/09/2023 via IBM Quantum experience.
- [288] A Morvan et al. “Phase transition in random circuit sampling”. In: *arXiv preprint arXiv:2304.11119* (2023).